Abnormal Pathological Mechanisms of TDP-43 Protein and its Associated Pathway in Association with ALS

Yawen Gao

College of Life Sciences, Shandong Normal University, Jinan, Shandong 250000, China
Corresponding author: 202213210113@stu.sdnu.edu.cn

Abstract:

TAR DNA binding protein 43 (TDP-43) is a nuclear protein with multiple RNA/DNA binding and editing functions. It is a key protein that causes Amyotrophic lateral sclerosis (ALS), frontotemporal lobe degeneration (FTLD), and other neurodegenerative diseases (NDDs). It is present in the patient's tissue in the form of abnormal amyloid deposits and has been identified as the cause of the disease. The mutation site of TDP-43 is known to be associated with possible abnormal folding, and it has been determined that the cause of its formation of precipitation is the disruption of the phase separation state. The TDP-43 protein itself is well studied, but its interaction with other genes to cause ALS is relatively scarce. In this paper, the interaction of TDP-43 protein with HSPB1 and UNC13A will be introduced, which indirectly leads to the increased risk of ALS. And whether the association triggered by CE of UNC13A can specifically select and detect TDP-43 lesions at the early stage of the disease. By summarizing part of the pathogenic mechanism of TDP-43, this paper provides a reference for future further studies.

Keywords: TDP-41; pathological mechanism; ALS.

1. Introduction

NDDs are diseases caused by the loss of function of neurons or myelin sheath, and the pathological changes accompanying their onset are often irreversible. Once the disease appears, its course is usually in the middle and late stages, and treatment can only delay the development of the disease. Still, it can not improve the already diseased tissue. Among them, ALS is a chronic, progressive neurological degenerative disease with a high fatality rate.

According to the survey, in all ALS cases, familial ALS (fALS) patients only account for 5%-10%, and there are few cases related to TARDBP gene mutations in fALS (5%-10%), which are mostly caused by other gene mutations, such as C9orf72, SOD1, etc. [1]. However, sporadic ALS (sALS) accounts for the vast majority of the total ALS cases (90-95%), and the cases involving abnormal deposition of TAR DNA binding protein 43 (TDP-43) account for the vast majority of sALS patients (97%) [1].

TDP-43 protein, encoded by the TARDBP gene on chromosome 1, is a multifunctional binding protein, one of which is involved in RNA metabolism.TDP-43 was identified as an HIV-1 transcription-related suppressor protein and a CFTR exon splice-related protein in 1995 and 2001, respectively [1]. TDP-43 belongs to the hnRNP family, which is highly conserved, universally expressed, and has

great sequence specificity. Reduced expression of TDP-43 protein in the brain and spinal cord is considered to be a signature pathological feature of ALS and FTLD [2]. The study found that mutated TDP-43 reduces its self-regulatory ability, resulting in its gradually increasing levels of the protein; this affects how TDP-43 binds and regulates different RNAs [3]. Later in the progression of ALS, the activity of TDP-43 decreases, resulting in abnormal localization or aggregation within the nucleus, which affects its interaction with all RNAs. In addition, [3]. TDP-43 was found to regulate multiple mRNAs associated with ALS or TDP-43 regulation; these include G3BP1, a core protein associated with stress particle formation, Gemin6, a Gem protein complex component associated with spinal muscular atrophy, and Ser/Thr kinase CK-1, which phosphorylates TDP-43 in response to stress [3]. These findings provide new insights into the pathological mechanisms of TDP-43, including abnormal phosphorylation, abnormal cytoplasmic localization, and TDP-43 aggregation, which are characteristic of ALS [3]. ALS and FTLD-TDP are both late-onset neurodegenerative diseases with several common clinical, neuropathological, and genetic features, but they affect different regions of the nervous system.

The researchers found that abnormal translation of the TDP-43 protein would greatly increase the possibility of being recruited into the cytoplasm by the cytoplasmic

protein TDP-35 and forming abnormal globular polymers. Without TDP-43 to assist in mRNA cutting, the hidden exon of UNC13A would participate in translation, resulting in reduced expression of UNC13A protein. Resulting in abnormal synaptic and neuronal function [4]. Therefore, TDP-43 can remove the hidden exon part of UNC13A mRNA of synapse-related genes, avoid the meaningless mediated decay and loss of UNC13A protein, and maintain the function of mRNA and the expression of UNC13A protein so as to maintain the normal function of synapses and neurons. These all indicate that studying the pathological mechanism of TDP-43 protein abnormality may bring new ideas and directions for ALS treatment.

Therefore, to review and integrate the current understanding of the pathology of TDP-43 in ALS, this paper will start with the structural basis of TDP-43 agglutination. In this paper, the mutation sites of its related genes and the probability of occurrence in cases were summarized, as the discovery of the mechanism by which HSPB1 regulates the cytoplasm TDP-43 phase separation and aggregation and the relationship between TDP-43 and splicing inhibition revealed by fluid biomarkers were introduced, and the basic molecular mechanism of the effect of TDP-43 protein on ALS through UNC13A was described. Finally, we discuss novel strategies for TDP-43 breakdown and ALS treatment [1].

2. Structural Basis of TDP-43 Agglutination

TARDBP consists of 414 amino acids, including an n-terminal domain (NTD) with nuclear localization signals, two RNA recognition motifs RRM1 and RRM2 (RRMs), and a C-terminal domain (CTD) [1]. The abnormal deposition of TDP-43 protein is closely related to its structure, so it is very important to introduce its structure [1].

2.1 NTDs

Many experimental data have proved that TDP-43 exists in the form of a dimer in the natural state, and its monomer-dimer equilibrium transformation is closely related to its aggregation [1]. The dimer form of TDP-43 is the basis for many of its functions; the most important and ALS-related one is to participate in RNA splicing [1]. By observing the structure of TARDBP, it can be seen that the formation of TDP-43 dimer is directly related to the interaction of n-terminal residues. The N-terminal region presents a ubiquitin-like fold consisting of an alpha helix and six beta chains arranged in the order of $\beta1$-$\beta2$-$\alpha1$-$\beta3$-$\beta4$-$\beta5$-$\beta6$ [1]. The dimer is formed by the head-to-head interaction of two NTDs of TDP-43, while the RRM2 domain will extend outward [1]. Among them, the first ten residues of NTDs play an important role in the forma-

tion of functional homo-dimers and also participate in the aggregation of full-length TDP-43 [1]. In the regulation of RNA synthesis splicing, the N-terminal region can promote self-oligomerization in a concentration-dependent manner, thus regulating the binding and separation of nucleic acids [1].

2.2 RRMs

These RRM domains consist of two alpha helices and five beta chains arranged in the sequence $\beta1$ -- $\beta2$ -- $\beta3$ -- $\alpha2$ -- $\beta4$ -- $\beta5$, which bind to homologous RNA/DNA molecules and exhibit greater specificity for short and high-ug /TG sequences [1]. At the same time, two ALS-associated mutations were detected in the relevant regions, suggesting that the RRM2 domain may also be involved in the formation of the TDP-43 protein dimer [1]. Since the function of RRMs is to recognize RNA, once it mutates, it will affect the RNA binding ability of its protein, resulting in changes in mRNA processing, RNA output, and RNA stability [1]. Therefore, changes in the function of rbp, such as TDP-43 caused by mutations in RRM, may prove that RNA metabolism disorder is a pathogenic factor of NDDs [1].

2.3 CTDs

The C-terminal region of the TDP-43 protein is highly disordered, containing a region rich in glycine and a fragment rich in uncharged polar amino acids, glutamate, and aspartate [1]. (Q/N) This similar structural region is commonly seen in the prion-like domain of yeast proteins, which can change from disordered constructs to self-templated, crossed beta-rich amyloid conformations, sometimes as an adaptive physiological response [1]. About a third to a quarter of all human proteins with prion-like domains are RNA/ DNA-binding proteins containing RRM motifs, including many proteins known to be associated with NDDs pathogenesis, such as TDP-43 and FUS [1]. This greatly enhances the relevance of the C-terminal region of TDP-43 to its pathogenic molecular mechanism, which contains fragments that lead to protein aggregation, mutated gene sites, and potentially cytotoxic substances [1]. First, its domain is the same as the prion-like domain, and its nature is disordered and prone to aggregation. Second, most of the TARDBP mutations and phosphorylation sites associated with ALS are located within the C-terminal region [1]. Third, some 25-35 KDa C-terminal fragments produced by TDP-43 with abnormal cystease activity are highly cytotoxic and have been found as major presence substances in the brain inclusion of ALS patients [1]. At the same time, the C-terminal region of TDP-43 also includes a short, highly dynamic, and unstable helio-transhelical region, from which peptides can form amyloid fi-

brils extracellular and exhibit prion-like infective seeding ability to cells expressing soluble TDP-43, causing TDP-43 to agglutinate in other normal cells [1]. The C-terminal region can also undergo liquid-liquid phase separation (LLPS) to form TDP-43 protein droplets, and C-terminal residues exhibit slight transient interactions in these dynamic protein droplets [1]. When stress conditions change, such as mutations or for too long, these droplets may undergo liquid-solid separation (LSPS), resulting in irreversible aggregates that eventually lead to the formation of amyloid protein precipitates [1].

2.4 Stress granules (SG)

SG is a mechanism that eukaryotic cells form after being stimulated by stress, such as oxidative stress, viral infection, or chemical stimulation, and can protect cells from cell damage. It is essentially a non-membrane structure formed by liquid-liquid phase separation and aggregation, and the formation process is dynamically reversible [5]. Generally speaking, SGs should dissolve and dissipate after the end of stress. Otherwise, the wrong aggregation will cause SGs to form irreversible amyloid deposition. Because nerve cells are very susceptible to stress, abnor-

mal stress responses are highly likely to convert SGs into amyloid deposits, and such protein deposits have been found in the brains of patients with NDDs such as ALS and FTD. TDP-43 has a protective effect on cell damage, which is reflected in its ability to assemble and maintain SG and regulate the expression of SG nucleation key proteins. At the same time, ALS-related mutations can affect SG dynamics [1].

3. Missense Mutations Associated with ALS and Their Corresponding Functions

As shown in Figure 1, many of the genes for TARDBP have been identified as being included in disease-causing genes for NDDs such as ALS [1]. Most ALS-associated mutations are located on exon 6 of TARDBP, the most common being A382T and M337V [1]. The genes that exist in both sALS and fALS are G294V, G348C, A328T, and so on [1]. The effects of all these mutations on TDP-43 protein traits are reflected in increased aggregation factors, increased cytoplasmic mislocalization, and changes in protease resistance [1].
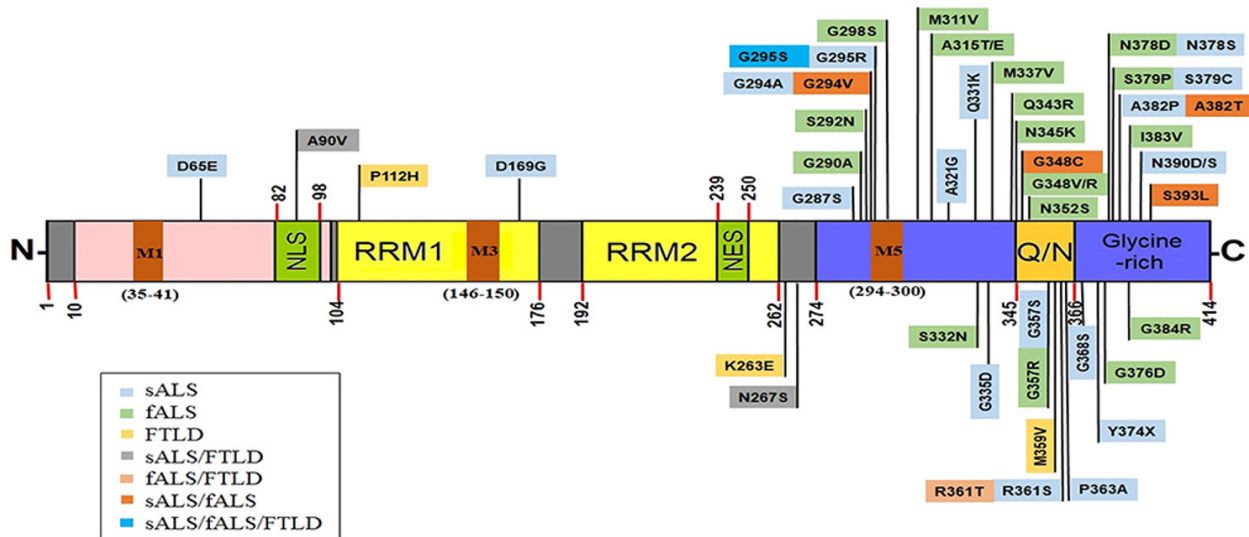

Fig. 1 Structural composition of TDP-43, mutation sites related to ALS and FTLD on TARDBP. sALS is sporadic ALS, fALS is familial ALS [1].

3.1 HSPB1 is a Regulator of Cytoplasmic TDP-43 Phase Separation and Aggregation

Normally, most TDP-43 plays a role in the nucleus, and a small part is involved in life activities in the cytoplasm. However, the researchers found abnormal cytoplasmic inclusion bodies in the brains of almost all of the patients studied, and these diseased tissues demonstrate that TDP-43 enters the cytoplasm and forms irreversible amyloid precipitates as one of the causative factors of ALS. Don W.

Cleveland et al. used isotope-labeled mass spectrometry to identify the binding of phase-isolated cytoplasmic TDP-43 to small heat shock protein HSPB1.
The researchers found that cytoplasmic TDP-43, when subjected to stress such as oxidation or reduced proteasome activity, leads to phase separation and protein state changes. [6]. Further experiments showed that the increase of cytoplasmic TDP-43 led to the formation of cytoplasmic TDP-43 droplets, which could block or even consume

nuclear TDP-43 [6]. Small heat shock protein HSPB1 was then identified in combination with APeX proximity labeling, quantitative mass spectrometry using isotopically labeled TMT (based on MS3), co-expression, and immunofluorescence [6]. The results showed that HSPB1 was recruited into cytoplasmic TDP-43 droplets, bound to cytoplasmic TDP-43 under arsenite stress, and decomposed into droplets/gels with it [6]. Finally, by studying the motor neurons of ALS patients, the authors found that the content of nuclear TDP-43 was significantly decreased, and the level of HSPB1 was also significantly decreased [6]. This study demonstrated that chaperone activity of HSPB1 and HSP70 is closely related to the liquid state of the TDP-43 droplets separated in the cytoplasm. This discovery brings new therapeutic ideas: by increasing the level of heat shock chaperones and other components of the protein quality control mechanism to reduce the formation of cytoplasmic irreversible amyloid precipitation, thereby slowing the progression of ALS patients and even curing patients.

3.2 The Loss of TDP-43 Splicing Inhibition Occurred before the Onset of Related Symptoms

Since synapses are the key organizations connecting the information pathways between neurons, maintaining the normal transmission of synaptic signals is an important prerequisite to ensure the normal function of neurons. UNC13A is a key gene related to synaptic function, and its functional mutation will lead to abnormal synaptic function, which may have a negative impact on neuronal function, thereby increasing the risk of NDDs such as ALS. Anna-Leigh Brown et al. took note of this gene and found that a decrease in TDP-43 led to a large amount of implicit exon (CEs) expression in UNC13A, which ultimately resulted in a decrease in the nonsense-mediated UNC13A protein [4].

In vitro, they used human induced pluripotent stem cell (iPS) -derived cortical like iNeuron to reduce the expression of TDP-43 with CRUSPR inhibition (CRISPRi). They performed RNA sequencing (RNA-Seq) on these cells [4]. By comparison, the researchers identified UNC13A as the only ALS risk gene, and in situ hybridization determined that CEs of UNC13A in iNeuron were mainly present in neurons with TDP-43 knocked out in the nucleus [4]. To complete the experiment, the researchers constructed several independent cell lines, and some cells in each cell line were knocked down with different degrees of TDP-43 [4]. The results showed that only the treated cells expressed CEs, and the expression level was related to the level of TDP-43 knocked down [4]. At the same time, in order to prove the direct regulation of TDP-43

on UNC13A, CE of UNC13A was detected, and multiple TDP-43 binding peaks and intron retention were found within and downstream of its domain [4]. RNA expression of UNC13A was reduced in cells where TDP-32 was knocked down [4].

In vivo, the researchers examined RNA-seq data from brain and spinal cord tissue from a large number of patients and controls at the New York Genome Center [4]. Data showed that UNC13A CE was detected only in patient tissues and not in the control group [4]. These experiments proved that TDP-43 plays an important role in maintaining the correct splicing of UNC13A and maintaining normal synaptic function [4]. This discovery opens up new possibilities for the detection and prevention of ALS and FTLD [4]. This discovery provides a new possibility and starting point for the treatment of ALS caused by TDP-43 deficiency.

4. CE Test before Symptoms

Abnormal inclusion bodies of TDP-43 protein have been detected in postmortem tissues of patients, but this does not provide a new way to treat ALS. Whether this lesion can be detected when the early symptoms are not obvious is the key to determining whether it can be detected in advance and intervene in the development process of ALS. Since abnormal precipitation of TDP-43 is known to lead to splicing inhibition and CE expression, the researchers attempted to create a fluid biological test marker that can specifically detect the protein encoded by CE and thus reflect the pathological changes of TDP-43 [7].

Using RNA sequencing, Katherine E. Irwin et al. identified a series of CE's associated with TDP-43 and screened some of the commonly expressed CE's to produce monoclonal antibodies. [7]. The results showed that CE of hepatocellular carcinoma-derived growth factor-like protein (HDGFL2) was the most likely target protein site and could be detected in C9orf72 spinal fluid and blood before symptoms. [7]. Further experiments proved that HDGFL2 CE was elevated in both sALS and fALS, [7] which brought new possibilities for the detection and prevention of ALS and FTLD. This indicates that the content of HDGFL2 CE may be used as an indicator to reflect whether TDP-43 protein is abnormal. In the early stage, without obvious symptoms, splicing inhibition and abnormal deposition of TDP-43 can be determined by detecting the expression of HDGFL2 CEs in patients' spinal fluid or blood.

5. Conclusion

This paper summarizes some recent articles on the TDP-43 protein causing ALS, aiming to expand the study of

TDP-43 from itself to its interaction with other genes and proteins by summarizing the latest research progress so as to apply some references for future studies.

TDP-43 protein also has many other pathogenic pathways, such as the destruction of nerve cells by causing mitochondrial damage; these approaches currently have relatively mature research and applications. It is known that many NDDs seem to have the presence of TDP-43 lesions, whether this provides a possible way to study other NDDs. As one of the universally expressed proteins, TDP-43 protein is also a pathogenic factor in many other diseases, and it may be possible to focus on the common features of TDP-43 lesions in these diseases as a common approach to treating multiple diseases. At the same time, it is suggested that later studies can shift the focus to the pathological study of the interaction of TDP-43 protein with other genes to cause the incidence of ALS and continue to enrich the understanding of the pathogenesis of TDP-43.

References

[1]Prasad A, Bharathi V, Sivalingam V, Girdhar A and Patel BK Molecular Mechanisms of TDP-43 Misfolding and Pathology in Amyotrophic Lateral Sclerosis. Front. Mol. Neurosci. 2019, 12:25.

[2]Ma, X.R., Prudencio, M., Koike, Y. et al. TDP-43 represses cryptic exon inclusion in the FTD–ALS gene UNC13A. Nature 2022, 603, 124–130.

[3]Hallegger M, Chakrabarti A M, Lee F C Y, et al. TDP-43 condensation properties specify its RNA-binding and regulatory repertoire. Cell, 2021, 184(18): 4680-4696. e22.

[4]Brown, AL., Wilkins, O.G., Keuss, M.J. et al. TDP-43 loss and ALS-risk SNPs drive mis-splicing and depletion of UNC13A. Nature 603, 2022, 131–137.

[5]Advani V M, Ivanov P. Stress granule subtypes: an emerging link to neurodegeneration. Cellular and Molecular Life Sciences, 2020, 77(23): 4827-4845.

[6]Lu, S., Hu, J., Arogundade, O.A. et al. Heat-shock chaperone HSPB1 regulates cytoplasmic TDP-43 phase separation and liquid-to-gel transition. Nat Cell Biol 24, 2022, 1378–1393.

[7]Irwin, K.E., Jasin, P., Braunstein, K.E. et al. A fluid biomarker reveals loss of TDP-43 splicing repression in presymptomatic ALS–FTD. Nat Med 2024, 30, 382–393.