

Review

# Advances in the Structure of GGGGCC Repeat RNA Sequence and Its Interaction with Small Molecules and Protein Partners

Xiaole Liu <sup>1</sup>, Xinyue Zhao <sup>1</sup>, Jinhan He <sup>1</sup>, Sishi Wang <sup>1</sup>, Xinfei Shen <sup>1</sup>, Qingfeng Liu <sup>1</sup>  
and Shenlin Wang <sup>1,2,\*</sup>

<sup>1</sup> State Key Laboratory of Bioreactor Engineering, East China University of Science and Technology, Shanghai 200237, China; y85210065@mail.ecust.edu.cn (X.L.); 20000624@mail.ecust.edu.cn (X.Z.); y85220142@mail.ecust.edu.cn (J.H.); 20000612@mail.ecust.edu.cn (S.W.); 20000603@mail.ecust.edu.cn (X.S.); 20000608@mail.ecust.edu.cn (Q.L.)

<sup>2</sup> Beijing NMR Center, Peking University, Beijing 100087, China

\* Correspondence: wangshenlin@ecust.edu.cn

**Abstract:** The aberrant expansion of GGGGCC hexanucleotide repeats within the first intron of the *C9orf72* gene represent the predominant genetic etiology underlying amyotrophic lateral sclerosis (ALS) and frontal temporal dementia (FTD). The transcribed r(GGGGCC)<sub>n</sub> RNA repeats form RNA foci, which recruit RNA binding proteins and impede their normal cellular functions, ultimately resulting in fatal neurodegenerative disorders. Furthermore, the non-canonical translation of the r(GGGGCC)<sub>n</sub> sequence can generate dipeptide repeats, which have been postulated as pathological causes. Comprehensive structural analyses of r(GGGGCC)<sub>n</sub> have unveiled its polymorphic nature, exhibiting the propensity to adopt dimeric, hairpin, or G-quadruplex conformations, all of which possess the capacity to interact with RNA binding proteins. Small molecules capable of binding to r(GGGGCC)<sub>n</sub> have been discovered and proposed as potential lead compounds for the treatment of ALS and FTD. Some of these molecules function in preventing RNA–protein interactions or impeding the phase transition of r(GGGGCC)<sub>n</sub>. In this review, we present a comprehensive summary of the recent advancements in the structural characterization of r(GGGGCC)<sub>n</sub>, its propensity to form RNA foci, and its interactions with small molecules and proteins. Specifically, we emphasize the structural diversity of r(GGGGCC)<sub>n</sub> and its influence on partner binding. Given the crucial role of r(GGGGCC)<sub>n</sub> in the pathogenesis of ALS and FTD, the primary objective of this review is to facilitate the development of therapeutic interventions targeting r(GGGGCC)<sub>n</sub> RNA.

**Keywords:** amyotrophic lateral sclerosis; frontotemporal dementia; *C9orf72*; GGGGCC; G4



**Citation:** Liu, X.; Zhao, X.; He, J.; Wang, S.; Shen, X.; Liu, Q.; Wang, S. Advances in the Structure of GGGGCC Repeat RNA Sequence and Its Interaction with Small Molecules and Protein Partners. *Molecules* **2023**, *28*, 5801. <https://doi.org/10.3390/molecules28155801>

Academic Editors: Ramon Eritja, Daniela Montesarchio and Montserrat Terrazas

Received: 26 June 2023

Revised: 21 July 2023

Accepted: 21 July 2023

Published: 1 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Amyotrophic lateral sclerosis (ALS) and frontal temporal dementia (FTD) are two neurodegenerative disorders characterized by progressive degeneration and dysfunction of neuronal architecture [1–5]. Both diseases have a fatality rate typically occurring within three to five years after the onset of symptoms [6,7]. ALS, affecting approximately two individuals per 100,000, is characterized by the degeneration of motor neurons, leading to muscle weakness and atrophy [8,9]. FTD, the second most prevalent form of dementia in individuals under the age of 65, is typically characterized by atrophy of the frontal and/or temporal lobes, manifesting as heterogeneous symptoms encompassing behavioral changes (behavioral variant FTD, bvFTD), language impairment (primary progressive aphasia, PPA), or deterioration in motor skills [10]. Despite considerable efforts, the development of efficacious therapeutic strategies for the treatment of ALS and FTD remains a challenge [11].

The etiologies of ALS and FTD are various. Sporadic ALS (sALS) accounts for 90% of the ALS patients. The remaining 10% of ALS patients are familial ALS (fALS) caused by mutations. The fALS can be caused by the dysfunction of mutated proteins, such as SOD1 mutations, FUS/TLS, and TDP-43 provoked by *TARDBP* mutations [12,13], which

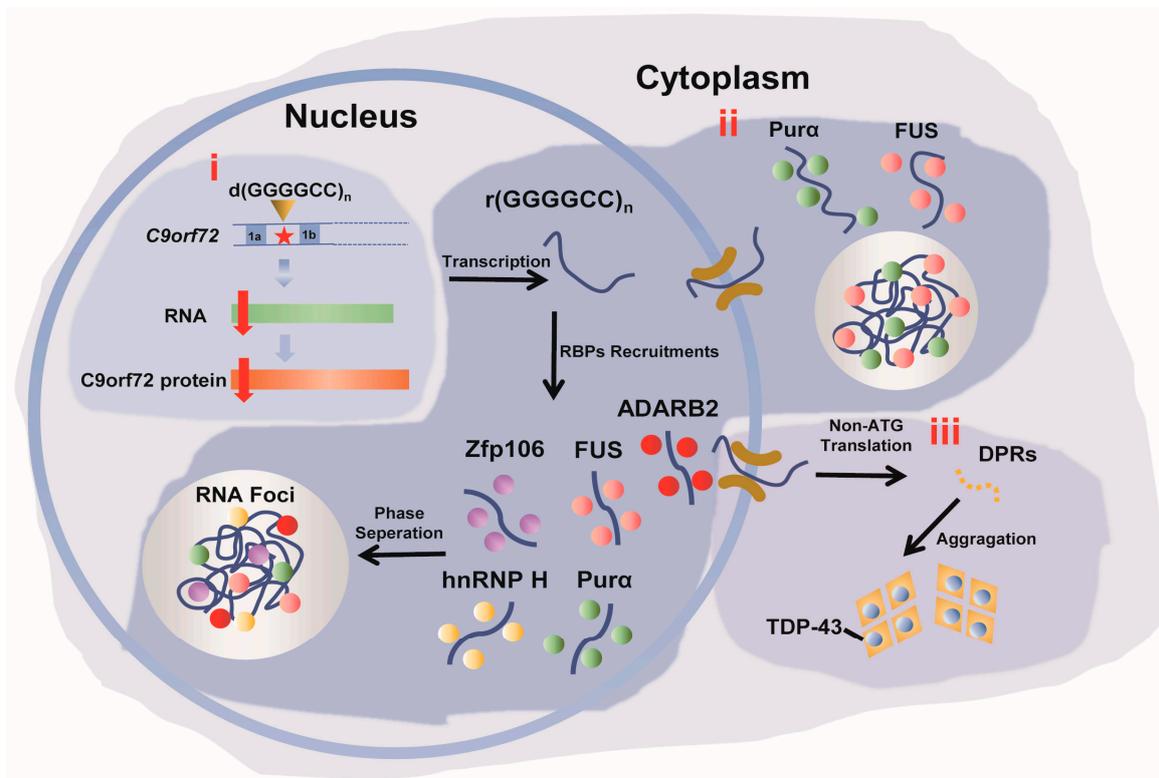
lead to neurotoxicity. The aberrant elongation of the hexanucleotide repeat GGGGCC in the non-coding region of *C9orf72* has also been demonstrated to be causally associated with a ALS and FTD [8,14–22]. Aberrant expansion of the GGGGCC repeats is observed in 8% of sALS patients, as well as in more than 40% of fALS cases [14]. Individuals affected by ALS exhibit an average repeat count ranging from 700 to 1600, whereas healthy individuals possess fewer than 25 repeats [23–26]. As for FTD, approximately one-third of FTDs are familial, with autosomal dominant mutations in three genes accounting for the majority of inheritance, including progranulin (GRN), *C9orf72*, and microtubule-associated protein tau (MAPT) [27]. The co-occurrence of these two disorders among families provides support for their genetic linkage [28].

Moreover, aberrant expansion of short nucleotide repeats has been observed in many neurodegenerative diseases. CTG triplet amplification in 3'-UTR may occur in dystrophin myotonia protein kinase (DMPK) gene, and alternative splicing of junctophilin (JPH) gene exon 2a and ataxin8 (ATXN8) gene, which can, respectively, result in Muscular dystrophy type 1 (DM1), Huntington disease-like 2 (HDL2), and Spinocerebellar Ataxia 8 (SCA8). The amplification of CGG triplet in 5'-UTR of the fragile X mental retardation 1 (FMR1) gene may lead to Fragile X disorders (FXTAS), while (CAG)<sub>n</sub> in the exon of ataxin3 (ATXN3) may cause Spinocerebellar Ataxia 3 (SCA3). In addition, (ATTCT)<sub>n</sub> within an intron of the ataxin 10 (ATXN10) gene and (CCTG)<sub>n</sub> in the first intron of the zinc finger protein 9 (ZNF9) gene may lead to Spinocerebellar Ataxia 10 (SCA10) and Muscular dystrophy type 2 (DM2), respectively [29]. Characterization of pathological mechanisms by which these short nucleotide repeats cause fatal diseases has been the research focus aiming in finding the treatments.

Three mechanisms have been proposed to elucidate the pathological underpinnings of aberrant GGGGCC expansion (Figure 1) [30]. Firstly, these abnormal expansions can lead to a gain or loss of function in the associated gene [7,31]. Secondly, the transcribed r(GGGGCC)<sub>n</sub> RNA forms RNA foci that recruit RNA binding proteins (RBPs), consequently impairing protein function and ultimately triggering intracellular cytotoxicity [7,32–36]. Lastly, non-ATG translation of the r(GGGGCC)<sub>n</sub> sequence produces dipeptide repeats (DPRs) that exert neurotoxic effects within the central nervous system [7,37–42]. Among the three proposed mechanisms, the formation of RNA foci and recruitment of RBPs have garnered the most attention. This process involves the spontaneous liquid–liquid separation of r(GGGGCC)<sub>n</sub>, followed by a sol–gel phase transition by increased interactions [43]. The RNA foci can recruit various RBPs, including hnRNP H [44], Zfp106 [45], ADARB2 [46], Purα [47,48], and FUS [49,50], ultimately leading to disruptions in the intracellular environment [43,51]. The aberrant phase separation and spread of hnRNP H within r(GGGGCC)<sub>n</sub> in ALS patients are key features in the pathogenesis of the disease [52].

Consequently, the investigation of the r(GGGGCC)<sub>n</sub> RNA repeats structures and the interactions between r(GGGGCC)<sub>n</sub> and small molecules is currently a highly prominent area of research. The primary objective is to identify lead compounds for the treatment of FTD and ALS [53]. The r(GGGGCC)<sub>n</sub> sequence exhibits characteristic structural polymeric and has the ability to adopt either a hairpin or G-quadruplex (G4) structure [54]. Several small molecules have been discovered to possess strong binding affinity for r(GGGGCC)<sub>n</sub>. A majority of these compounds incorporate polyaromatic ring conjugated systems that effectively stabilize the G4 structures of r(GGGGCC)<sub>n</sub>, thereby inhibiting phase separation, disrupting protein–RNA interactions [55] and/or preventing non-ATG translation of DPRs [53]. Additionally, two drugs, riluzole and edaravone, have received approval for the treatment of ALS [56]. While these drugs can delay disease progression, they do not specifically target the r(GGGGCC)<sub>n</sub> RNA [56]. Moreover, recent research by Meijboom and colleagues used the adeno-associated virus vector system to deliver CRISPR/Cas9 gene editing system into neuron cells, and successfully removed the hexanuclear repeat expansion from the *C9orf72* gene in the mouse model (500–600 repeats), as well as the patient-derived Induced Pluripotent Stem Cell (iPSC) motor neuron and brain organoid (450 repeats). This led to a reduction in RNA foci, DPRs and haploinsufficiency, major

hallmarks of C9-ALS/FTD, making this a promising therapeutic approach to ALS/FTD diseases [57].



**Figure 1.** Three pathogenic mechanisms associated with *C9orf72*-related FTD and ALS. (i) Aberrant expansion of the  $d(GGGGCC)_n$  of *C9orf72* within contains two non-coding exons (1a and 1b), and under pathological conditions, repress transcription, resulting in the reduction in *C9orf72* protein. (ii) The transcribed  $r(GGGGCC)_n$  aggregates in the nucleus to form RNA foci that recruit RBPs, affecting the intra cellular functions of RBPs, i.e., splicing. (iii) The  $r(GGGGCC)_n$  RNA is transported into the cytoplasm and undergoes repeat-associated non-ATG translation, resulting in the synthesis of DPRs. The DPRs forms aggregation and associate TDP-43, which induce cytotoxic effects in cells. The labels (i–iii) correspond to the three pathological mechanisms. The red star highlights the hexanucleotide repeat GGGGCC in the non-coding region of *C9orf72*.

This review provides a comprehensive overview of the recent progress made in understanding the structures of  $r(GGGGCC)_n$ , and the interactions between  $r(GGGGCC)_n$  and small molecules and between  $r(GGGGCC)_n$  and protein partners. We focus on elucidating the structural diversity of  $r(GGGGCC)_n$  and its implications for partner binding. Given the crucial role of  $r(GGGGCC)_n$  in the pathogenesis of ALS and FTD, the primary objective of this review is to support the development of drugs targeting  $r(GGGGCC)_n$  RNA.

## 2. The Structure of $r(GGGGCC)_n$ RNA Repeats and the RNA within the RNA Foci of $r(GGGGCC)_n$

### 2.1. The Solution Structures of $r(GGGGCC)_n$ RNA

The  $r(GGGGCC)_n$  RNA is a guanine-rich sequence, which promotes the formation of G4 structures. While the tertiary structure of  $r(GGGGCC)_n$  remains to be fully elucidated, the secondary structures have been extensively studied. Circular Dichroism (CD) spectra are commonly used to demonstrate the G4 structures and the topology of G4, which for instance, the spectral patterns would provide the evidence of parallel, antiparallel, or other types of topologies (Figure 2). Nuclear Magnetic Resonance (NMR) spectra are able to provide more structural details, some of which yield full structural determination or provide evidence of co-existence of different conformations (Figure 2a). Depending on the sequence

length and solution conditions,  $r(\text{GGGGCC})_n$  can adopt different secondary structures, including G4 [58,59] and hairpin conformations [60]. In 2012, Adrian M. Isaacs and colleagues demonstrated that  $r(\text{GGGGCC})_3\text{GGGGC}$  can fold into G4 or double-stranded structures, with the topology being influenced by the presence of cation ions in the solution [61]. In a  $\text{K}^+$  buffer, it forms a stable parallel intramolecular G4 structure, while it becomes less stable in  $\text{Na}^+$  and  $\text{Li}^+$  solutions.

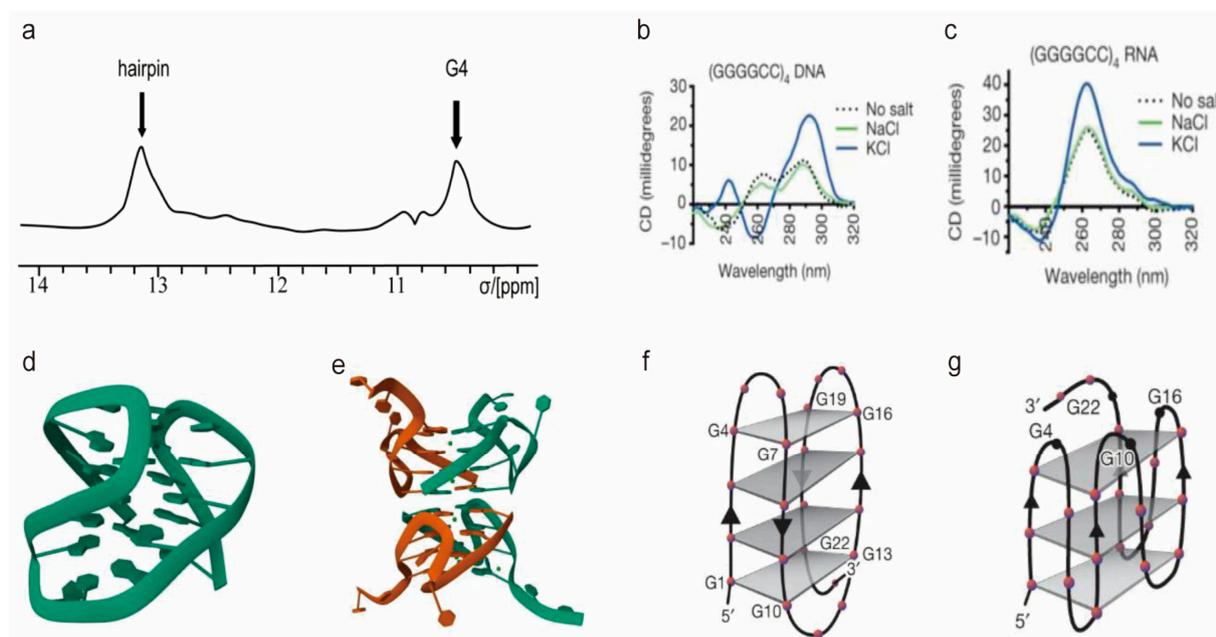
Further investigations by Pearson and colleagues employed circular dichroism (CD) spectroscopy (Figure 2a) and gel-shift assays, revealing that  $r(\text{GGGGCC})_n$  ( $n = 2, -5, -6,$  and  $-8$ ) predominantly adopt highly stable uni- and multi-molecular parallel G4 structures [62]. The abundance of G4 structures is influenced by the repeat number and RNA concentrations, with the proportion of multi-molecular G4 structures increasing as the number of repetitions rises.

The equilibrium between G4 and hairpin structures has also been observed in  $r(\text{GGGGCC})_n$ . In the absence of  $\text{K}^+$  ions,  $r(\text{GGGGCC})_4$  RNA forms a hairpin conformation [62], featuring single-stranded bulges within the RNA chain. However, in a  $\text{K}^+$  buffer (Figure 2c), it adopts a parallel G4 structures (Figure 2g) [59]. This equilibrium between hairpin and G4 structures is suggested to be linked to the presence of an abortive transcript containing hexanucleotide repeats [55]. The G4 structure may hinder the transcription of full-length RNA and recruit RBPs in cells, contributing to disease pathogenesis. The equilibrium is biased towards the hairpin conformation with a higher repeat number of  $r(\text{GGGGCC})_n$ . Specifically,  $r(\text{GGGGCC})_4$  predominantly adopts a G4 topology, while  $r(\text{GGGGCC})_8$  RNA exhibits both G4 and hairpin structures, even in a  $\text{K}^+$  buffer, as confirmed by various biophysical methods. However, in a  $\text{Na}^+$  buffer,  $r(\text{GGGGCC})_8$  RNA solely adopts a hairpin structure [55]. Furthermore,  $r(\text{GGGGCC})_4$  undergoes a monomer-dimer equilibrium in a pH-dependent manner. At pH 6.0 and 25 °C, it exists as both a homodimer and a hairpin structure. Decreasing the temperature increases the population of dimeric RNA, which exhibits distinct structural differences compared to G4 structures in the presence of  $\text{K}^+$  [63]. Conversely, at neutral pH,  $r(\text{GGGGCC})_4$  primarily adopts a hairpin conformation.

## 2.2. Structure of $d(\text{GGGGCC})_n$ DNA

High-resolution structures of  $d(\text{GGGGCC})_n$  have been successfully determined [64,65]. Janez Plavec and colleagues utilized NMR spectroscopy to elucidate the structure of  $d[(\text{GGGGCC})_3\text{GG}^{\text{Br}}\text{GG}]$  (represented by PDB codes 2N2D) [66]. The incorporation of a bromine-substituted guanine residue ( $\text{G}^{\text{Br}}$ ) contributed to the stabilization of the conformation, leading to a more rigid structure amenable to structural analysis. The  $d[(\text{GGGGCC})_3\text{GG}^{\text{Br}}\text{GG}]$  sequence adopted an antiparallel G4 topology (Figure 2b) [67].

In 2015, Guang Zhu and colleagues employed CD, NMR, and native polyacrylamide gel electrophoresis (PAGE) to investigate the structures of  $d(\text{GGGGCC})_n$  repeats. Their studies revealed distinct G4 folding patterns in the presence of  $\text{K}^+$  ions. Notably,  $d(\text{GGGGCC})\text{GGGG}$ ,  $d(\text{GGGGCC})_2$ , and  $d(\text{GGGGCC})_3$  did not exhibit stable G4 structures. Instead,  $d(\text{GGGGCC})_2$  and  $d(\text{GGGGCC})_3$  displayed mixed forms of parallel and antiparallel G4 folding. On the other hand,  $d(\text{GGGGCC})_4$  and  $d(\text{GGGGCC})_5$  formed stable G4 structures. Specifically,  $d(\text{GGGGCC})_5$  exhibited a combination of parallel and antiparallel G4 folds, while  $d(\text{GGGGCC})_4$  adopted a homogeneous monomeric form characterized by a chair-type G4 structure [10]. In 2021, this group determined the crystal structure of  $d(\text{GGGGCC})_2$  in both  $\text{Ba}^{2+}$  and  $\text{K}^+$  solutions, revealing an eight-layer parallel G4 structure for  $d(\text{GGGGCC})_2$  (represented by PDB codes 7ECF and 7ECG) (Figure 2e) [68]. Jiou Wang and colleagues (Figure 2d) also confirmed that  $d(\text{GGGGCC})_4$  adopts an antiparallel G4 (Figure 2f) [59].



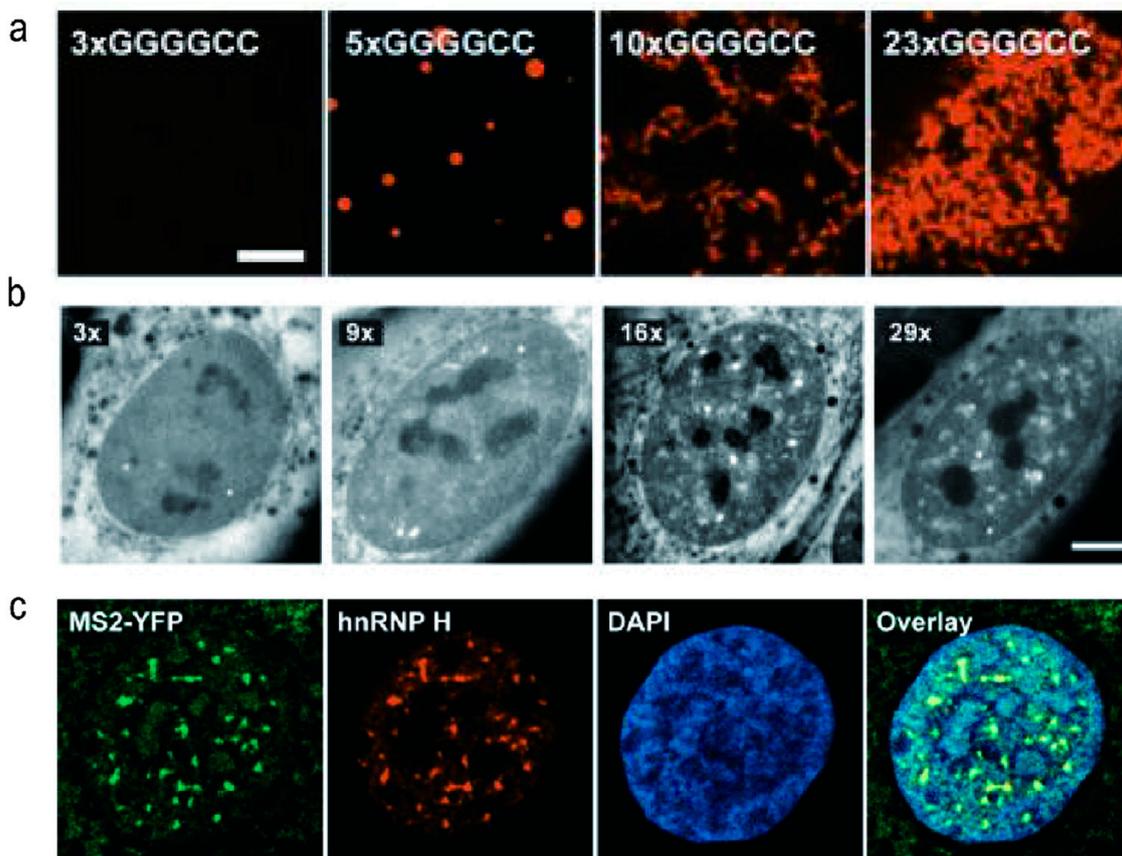
**Figure 2.** Biophysical methods in studying the  $r(\text{GGGGCC})_n$  and the  $d(\text{GGGGCC})_n$  structures. (a)  $^1\text{H}$  NMR spectra of  $r(\text{GGGGCC})$ , with highlighting the cross-peaks of the hairpin and G4 structures. (b,c) The CD spectra with the characteristic antiparallel G4 topology of  $d(\text{GGGGCC})_4$  (b) and the parallel G4 for the  $r(\text{GGGGCC})_4$  in the presence of 100 mM KCl. The ions dependent of G4 structures were shown. (d,e) The high-resolution crystal structures of  $d[(\text{GGGGCC})_3\text{GG}^{\text{Br}}\text{GG}]$  (d) and  $d(\text{GGGGCC})_2$  (e). (f,g) The proposed topology for the antiparallel DNA G4 formed by  $(\text{GGGGCC})_4$  (f) and the parallel G4 topology formed by the  $r(\text{GGGGCC})_4$  RNA (g). The (b,c,f,g) were reprinted from the reference [59]. (d,e) were reprinted from the reference [67,68], respectively.

### 2.3. Biological Phase Separation and Transition of $r(\text{GGGGCC})_n$

Biological liquid–liquid phase separation is a widely observed phenomenon in cells and plays a critical role in the formation of membraneless organelles, signal transduction, and DNA packaging [69–72]. As the strength of interactions in phase separation systems increases, a transition from a liquid to a solid state often occurs, resulting in the formation of insoluble gel-like states, many of which are associated with diseases [73]. Jain and colleagues demonstrated that  $r(\text{GGGGCC})_n$  can undergo phase separation both in vivo and in vitro [43]. They found that phase separation of  $r(\text{GGGGCC})_n$  occurs once a specific threshold of repeat value is reached, leading to a solution-gel phase transition as the strength of multi-base interactions increases (Figure 3). The formation of RNA foci is dependent on solution conditions and is reinforced by  $\text{Mg}^{2+}$  but impaired by monovalent cations such as  $\text{K}^+$  or  $\text{Na}^+$ . The authors proposed that inter-chain hydrogen bonds stabilize intermolecular G4s, which serve as the building blocks of RNA foci. However, direct evidence of the secondary structure of  $r(\text{GGGGCC})_n$  within RNA foci is still lacking.

Christopher E. Shaw and colleagues discovered that  $r(\text{GGGGCC})_n$  RNA foci were detected in neuronal cell lines and zebrafish embryos expressing 38 or 72 repeats but not in those expressing 8 repeats [6]. This finding indicates that longer  $r(\text{GGGGCC})_n$  sequences lead to nuclear retention of transcripts and the formation of RNA foci, which are resistant to the enzyme ribonuclease (RNase) [6,52]. Extended  $r(\text{GGGGCC})_n$  sequences exhibit significant neurotoxicity and bind to hnRNP H and other RBPs. RNA toxicity and sequestration of RBPs may impair RNA processing and contribute to neurodegenerative diseases. In a study conducted by Simon Alberti and colleagues, it was demonstrated that RNA plays a crucial role in regulating the phase behavior of prion-like RBPs [74]. Lower RNA to protein ratios promote the separation of RBPs into liquid droplets, whereas higher ratios prevent droplet formation in vitro. When nuclear RNA levels are reduced or RNA binding is genetically ablated, excessive phase separation occurs, leading to the

formation of cytotoxic solid-like assemblies in cells. The researchers proposed that the nucleus functions as a buffered system, with high RNA concentrations maintaining RBPs in a soluble state. Disruptions in RNA levels or the RNA binding abilities of RBPs result in abnormal phase transitions [75].



**Figure 3.** Representative results of biological phase separation of  $r(\text{GGGGCC})_n$ . (a) The in vitro fluorescence imaging of  $r(\text{GGGGCC})_n$  RNA clusters at indicated number of  $r(\text{GGGGCC})_n$ . (b) Representative fluorescence micrographs and corresponding quantification of the total volume of foci per cell in U-2OS cells transduced with  $r(\text{GGGGCC})_n$  RNA with the indicated number of  $r(\text{GGGGCC})_n$ . (c) Representative immunofluorescence images illustrating that the  $r(\text{GGGGCC})_{29}$  recruited endogenous hnRNP H. Figure 3 was reprinted from the reference [43].

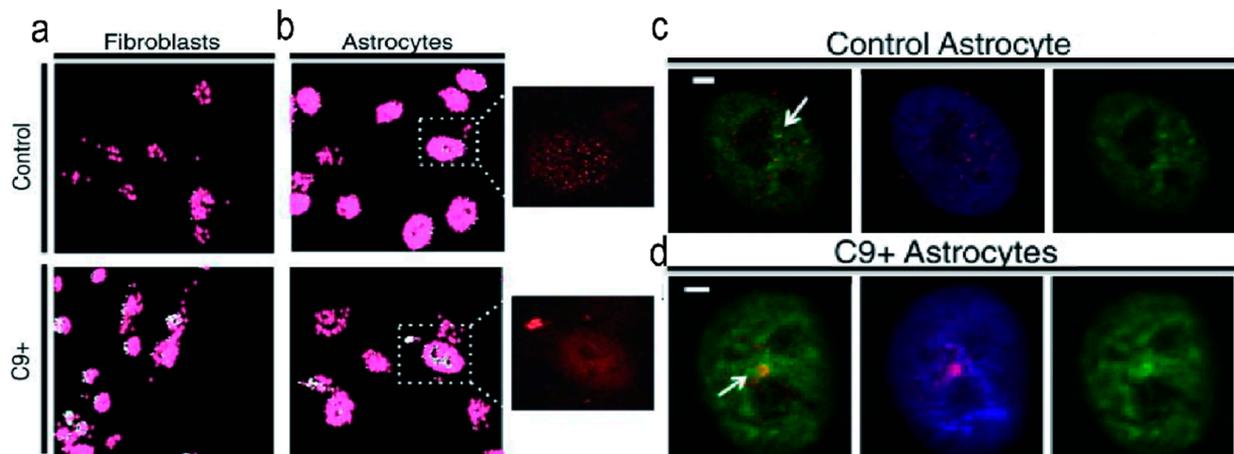
### 3. Disease Related RBPs That Bind to $r(\text{GGGGCC})_n$

#### 3.1. hnRNP H and TDP-43

Heterogeneous nuclear ribonucleoprotein H (hnRNP H) is a member of the hnRNP family and functions as a multifunctional RBP involved in mRNA maturation at various stages [76]. It contains a modular domain consisting of tandem quasi-RNA recognition motifs (HqRRM1,2) at the N-terminus and a third qRRM3 at the C-terminus, situated between two glycine-rich segments [44,77,78]. The hnRNP H has the ability to bind G-rich RNA sequences containing at least three consecutive guanines [44]. In the brain cells of ALS patients, hnRNP H has been found associated with insoluble aggregation of  $r(\text{GGGGCC})_n$ , leading to aberrant alternative splicing [52]. This phenotype has been utilized as a biomarker for disease diagnosis. Furthermore, ALS/FTD patients exhibit splicing alterations in several key targets and insoluble hnRNP H, indicating that modifications along this axis are critical aspects of disease etiology [52].

James L. Manley and colleagues demonstrated that hnRNP H binds to  $r(\text{GGGGCC})_n$  in vitro, and this interaction is dependent on the formation of G4s. The hnRNP H colocalizes with G4 aggregates in C9 patient-derived fibroblasts and astrocytes, but not in

control cells, as proven by imaging on BG4, a G4 structure-specific antibody (Figure 4) [79]. Another study by Donald C. Rio and colleagues revealed that in sporadic ALS/FTD patients, insolubility of hnRNP H was associated with altered splicing of a wide range of targets [52]. Numerous ALS/FTD brains show high levels of insoluble hnRNP H sequestered in r(GGGGCC)<sub>4</sub> RNA foci, resulting from RNA splicing defects involving intron retention [52]. These findings highlight previously unreported splicing abnormalities in extremely insoluble hnRNP H-related ALS brains, suggesting a potential feedback relationship between effective RBP concentrations and protein quality control in all ALS/FTD cases.



**Figure 4.** Quantification of stained BG4-foci and area was performed in fibroblasts and astrocytes derived from patients with ALS/FTD and healthy controls. Representative images of non-ALS and ALS fibroblasts (a) and astrocytes (b) are shown, with the ‘BG4 Count’ projection representing all stained areas above the determined threshold (showed in red), and areas of particularly dense staining shown in white. The inset displays the source image, highlighting only the red channel corresponding to BG4-FLAG staining. (c) Control astrocytes exhibit single, small nuclear hnRNP H/BG4 foci. (d) Patient astrocytes demonstrate nuclear hnRNP H/BG4 foci. Figure 4 was reprinted from the reference [79].

TAR DNA binding protein 43 (TDP-43), another member of the hnRNP family, possesses two RNA recognition motifs (RRMs), a nuclear localization signal (NLS), and a prion-like domain at the C-terminus [80]. Numerous mutations in TDP-43 have been associated with ALS and FTD [81,82]. The accumulation of TDP-43 is a major pathological feature of ALS and FTD [83–85], and inclusion bodies are observed in patients with abnormal expansions of r(GGGGCC)<sub>n</sub>, serving as a histopathological marker in 97% of ALS cases and 45% of FTD cases.

In contrast to hnRNP H, which directly associates with r(GGGGCC)<sub>n</sub>, the pathogenic mechanism of TDP-43 in ALS/FTD is believed to involve its interaction with DPRs, which are non-ATG translation products of r(GGGGCC)<sub>n</sub> [15,86]. Edward B. Lee and colleagues discovered that DPRs induce TDP-43 protein lesions in an ALS/FTD model and trigger the onset and progression of FTD [81]. The amount and characteristics of produced DPRs, rather than the length of r(GGGGCC)<sub>n</sub> repeats, determine the duration and severity of TDP-43 dysfunction.

### 3.2. FUS

Sarcoma fusion protein (FUS) is a 526-amino acid residue protein. [87] It is predominantly expressed in neurons and is involved in DNA and RNA metabolism through its interactions with motor proteins kinesin [88] and myosin-Va [89]. Missense mutations in the FUS gene have been associated with ALS [90,91], although the prevalence of FUS gene variants in the familial ALS population is low.

Sua Myong and colleagues conducted investigations on the binding of wild-type FUS to single-stranded RNAs, including r(GGGGCC)<sub>4</sub>, in a length-dependent manner. They observed the formation of a highly dynamic protein–RNA complex. The FUS–RNA interaction involves two mechanisms: (i) stable binding of FUS monomers to single-stranded RNA (ssRNA), and (ii) weak interaction of two FUS units with RNA, resulting in a highly dynamic interaction.

Higuro and workers observed the formation and phase transition of FUS condensates *in vitro* using purified full-length wild-type and mutant FUS proteins and r(GGGGCC)<sub>4</sub>. They found that FUS specifically forms complexes with r(GGGGCC)<sub>4</sub> in a G4 structure-dependent manner, leading to a transition from liquid–liquid separation to liquid–solid transitions. Importantly, amino acid mutations associated with ALS significantly impact G4-dependent FUS condensation. These findings provide insights into the relationship between protein aggregation and dysfunction of FUS in ALS [49].

### 3.3. Zfp106

Zfp106 is a C2H2 zinc finger protein characterized by the presence of seven WD40 domains and four putative zinc fingers [92]. It plays a crucial role in maintaining neuromuscular signaling. Knockout mice exhibit gene expression patterns indicative of neuromuscular degeneration in their muscles and spinal cords. Interestingly, this phenotype can be reversed through motor neuron-specific repair of the Zfp106 transgene, highlighting its essential role in biological processes [93]. The functional acquisition model of C9orf72 neurodegeneration has been investigated in a *Drosophila* model [94], where Zfp106 effectively mitigates the neurotoxicity associated with the expression of GGGGCC repeat in C9orf72 ALS *Drosophila*. This suggests that Zfp106 acts as a repressor of neurodegeneration in C9orf72 ALS models and demonstrates a functional interaction between Zfp106 and the r(GGGGCC)<sub>n</sub> sequence. Furthermore, Brian L. Black and colleagues conducted pull-down assays and mobility shift assays, providing evidence that Zfp106 specifically binds to r(GGGGCC)<sub>8</sub> but not to the sequence of r(AAAACC)<sub>8</sub>. The ability of Zfp106 to regulate normal cellular functions and inhibit ALS by binding to r(GGGGCC)<sub>n</sub> makes it a potential drug target for treating ALS [45]. However, the mechanisms through which Zfp106 regulates normal cellular processes via RNA binding and how it inhibits ALS progression by interacting with r(GGGGCC)<sub>n</sub> are still being investigated to guide drug design efforts [45].

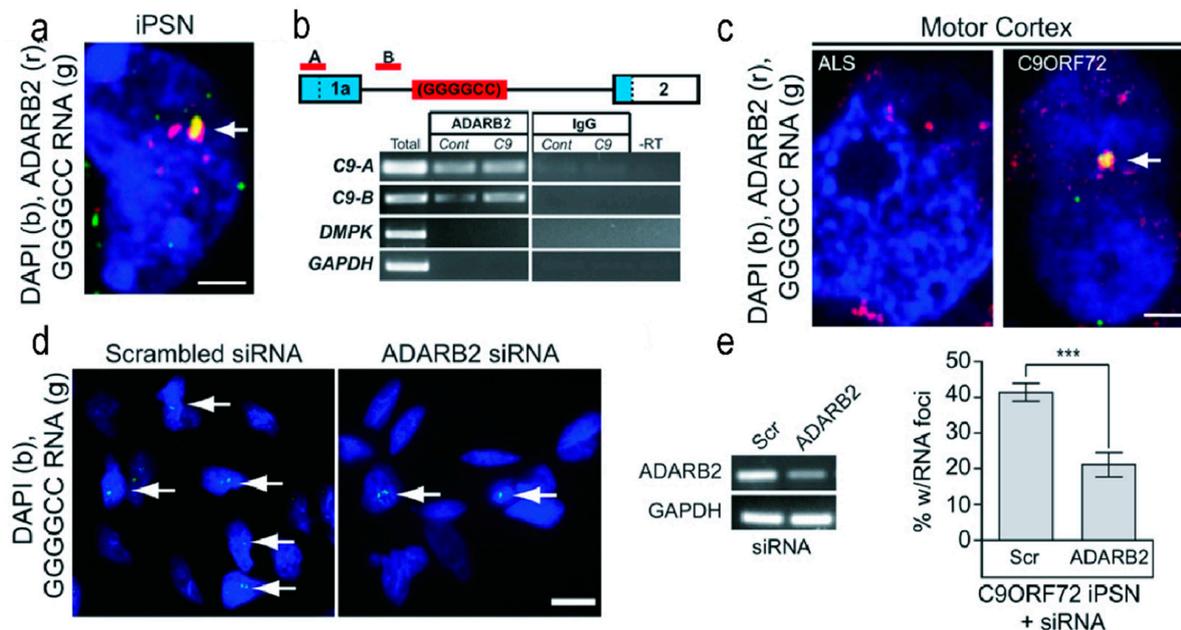
### 3.4. ADARB2

ADARB2 is a member of the CNS-rich adenosine deaminase family, known for its role in mediating A-to-I (adenosine to inosine) editing of RNA [95]. It consists of two double-stranded-specific adenosine deaminase repeats, three double-stranded RNA-binding domains, and one editase domain spanning from the N- to C-terminus. The A-to-I editing activity primarily occurs within the 16–130 nucleotide interval. This enzyme selectively deaminates adenosine (A) residues in the double-stranded region of mRNA, converting them to inosine (I), which is recognized as guanine by the cellular translation machinery, resulting in codon alterations within the synthesized protein [46] (Figure 5).

Jeffrey D. Rothstein and colleagues conducted RNA fluorescence *in situ* hybridization (RNA FISH) and immunofluorescence labeling of RBP simultaneously in the induced pluripotent stem neuron (IPSN) cell line derived from C9orf72-related cases. Their study revealed the co-localization of ADARB2 protein with nuclear r(GGGGCC)<sub>n</sub> RNA foci, while mRNA levels remained unchanged. Co-precipitation of ADARB2 with r(GGGGCC)<sub>n</sub> repeats was also observed *in vivo*.

*In vitro* investigations utilizing recombinant ADARB2 through gel shift assays clearly demonstrated its binding to r(GGGGCC)<sub>n</sub>, implying the possible formation of ADARB2–RNA complexes. These collective findings indicate a strong binding between ADARB2 and r(GGGGCC)<sub>n</sub>. Furthermore, this team verified *in vivo* that the formation of r(GGGGCC)<sub>n</sub> RNA foci requires the involvement of ADARB2 protein. Treatment of the IPSN line with

specific siRNA targeting ADARB2 significantly reduced the number of RNA foci. However, further experimental evidence is still needed to fully elucidate ADARB2's *in vivo* function [96]. Another unresolved aspect of ADARB2 function is the speculation that ADARB2 may lose its editing activity upon interaction with r(GGGGCC)<sub>n</sub>, although experimental validation of its downstream editing effects is currently lacking.



**Figure 5.** ADARB2 Protein Binds to the r(GGGGCC)<sub>n</sub>. (a) Colocalization of r(GGGGCC)<sub>n</sub> RNA foci with ADARB2 signal in IPSN cells. (b) Co-immunoprecipitation (co-IP) of ADARB2-bound RNA isolated from control and *C9orf72*-induced cell lines. RT-PCR of the co-IP RNA using two primer sets (A and B, red), located upstream of the r(GGGGCC)<sub>n</sub> repeat, demonstrated ADARB2 binding to *C9orf72* RNA in both control and *C9orf72* cell lines. (c) Colocalization of r(GGGGCC)<sub>n</sub> RNA foci and ADARB2 was observed in postmortem motor cortex tissue from *C9orf72* patients. (d,e) Knockdown of ADARB2 using siRNA significantly reduced the percentage of nuclear RNA foci (indicated by arrows). siRNA knockdown of ADARB2 results in a significant reduction in the percent of iPSNs with nuclear RNA foci (arrows). Data in (E) indicate mean  $\pm$  SEM (\*\*\*)  $p < 0.001$ ). Figure 5 was reprinted from the reference [46].

### 3.5. Pur $\alpha$

Pur-alpha (Pur $\alpha$ ) is a highly conserved DNA and RNA binding protein in eukaryotic cells [97]. It performs diverse physiological functions, including transcription activation or inhibition, cell growth, and translation [98,99]. While predominantly localized in the nucleus, Pur $\alpha$  is also widely distributed in the cytoplasm of neurons, particularly in synaptic branches [88]. In the nucleus, Pur $\alpha$  stimulates gene transcription by binding to mRNA transcripts and accompanying them to the cytoplasm. It remains associated with the mRNA during transport over considerable distances and functions at specific sites of mRNA translation [100]. The absence of Pur $\alpha$  can lead to various neurological disorders [101,102].

The r(GGGGCC)<sub>n</sub> repeat can sequester Pur $\alpha$ , thereby impairing its normal functions such as gene transcription and mRNA translation, ultimately resulting in cell death [103]. In an ALS/FTD zebrafish model, Swinnen and colleagues demonstrated that the Pur2 domain of Pur $\alpha$  binds to r(GGGGCC)<sub>90</sub> repeat RNA [37]. Peng Jin and colleagues conducted studies on the pathogenesis of ALS/FTD, revealing that r(GGGGCC)<sub>10</sub> can sequester Pur $\alpha$ , a major component of RBPs, from the whole-cell lysate of mouse spinal cord [47]. Rossi and colleagues found that Pur $\alpha$  can aggregate into cytosolic and nuclear granules in HeLa cells transiently transfected with a plasmid expressing r(GGGGCC)<sub>31</sub>. Nonetheless, due

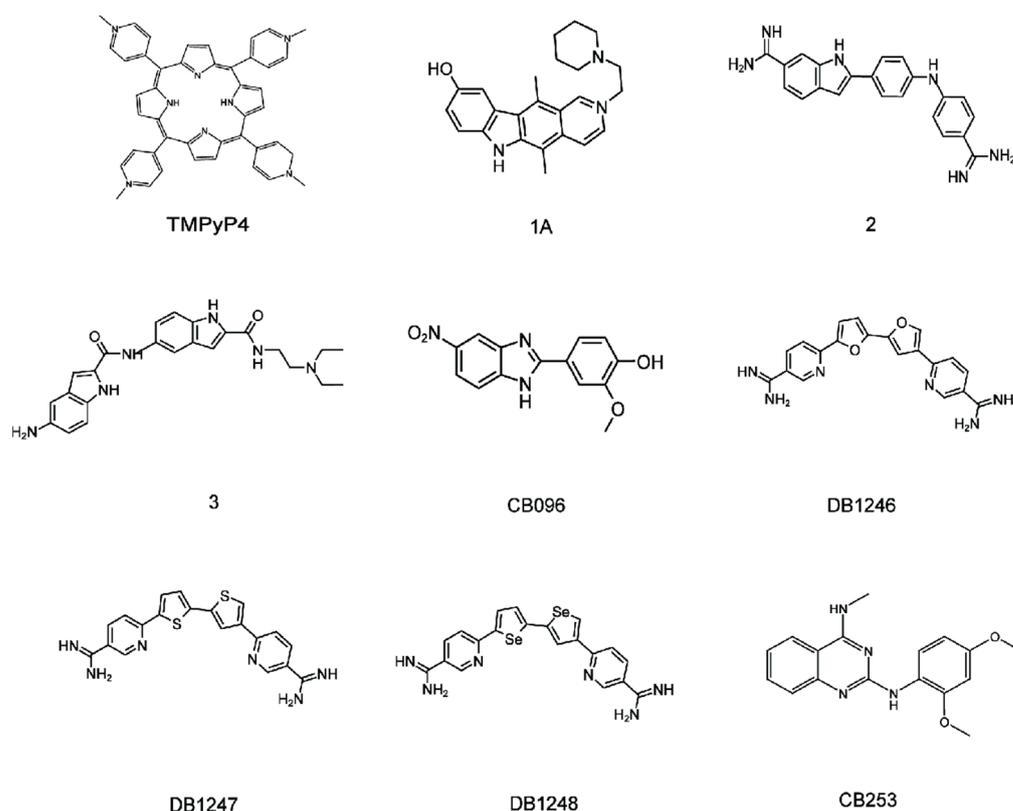
to the specific interaction between Pur $\alpha$  and r(GGGGCC)<sub>n</sub>, it is conceivable that Pur $\alpha$  may influence the outcome of RAN translation. Consequently, in ALS, reduced protein levels amplify certain cellular characteristics. Over-expression of Pur $\alpha$  in mammalian and *Drosophila* model systems can rescue r(GGGGCC)<sub>n</sub> repeat-induced neurodegeneration [47].

Furthermore, Pur $\alpha$  also interacts with the C-terminal region of FUS, another protein recruited by r(GGGGCC)<sub>n</sub> [104]. In vivo expression of Pur $\alpha$  in various *Drosophila* tissues significantly exacerbates neurodegeneration caused by mutated FUS. Conversely, reducing Pur $\alpha$  expression in neurons expressing mutated FUS significantly improves the climbing ability of *Drosophila* flies. This suggests that downregulation of Pur $\alpha$  ameliorates locomotion defects, a classical symptom of ALS resulting from mutant FUS expression. These findings indicate that Pur $\alpha$  may contribute to the pathogenesis of ALS mediated by FUS. However, it remains unclear which functional domains or subdomains of Pur $\alpha$  are involved in mediating its interaction with FUS [105].

Binding of Pur $\alpha$  to other cellular proteins can directly impact the expression of the *PURA* gene. Pur $\alpha$  itself can bind to GC/GA-rich sequences in its own promoter and inhibit gene expression [106]. Similarly, binding of Pur $\alpha$  to expanded polynucleotide repeat RNA may also affect the expression of the *PURA* gene. In both scenarios, the mechanism of action may involve the combination of Pur $\alpha$  with cellular components, resulting in a reduction in effective intracellular Pur $\alpha$  levels. The reduction in Pur $\alpha$  could trigger a feedback mechanism of the *PURA* gene, although it is unknown whether this compensates for Pur $\alpha$  sequestration [100].

#### 4. Lead Small Molecules Binds to r(GGGGCC)<sub>n</sub>

Given the pharmacological advantages of r(GGGGCC)<sub>n</sub> formation of RNA foci and their recruitment of RBPs, small molecules present an attractive option for targeting r(GGGGCC)<sub>n</sub>. Therefore, it is interesting to investigate the binding of r(GGGGCC)<sub>n</sub> to small molecules (Figure 6). Currently, a number of the small molecules contain aromatic rings have been found to bind to r(GGGGCC)<sub>n</sub>.



**Figure 6.** r(GGGGCC)<sub>n</sub> small molecular structure bound to small molecules.

#### 4.1. Binding of $r(\text{GGGGCC})_8$ with the TMPyP4

The G4 structure has been shown to bind to 5,10,15,20-tetra(N-methyl-4-pyridyl) porphyrin (TMPyP4), as demonstrated before [107,108]. TMPyP4 binds a variety of G4 structures of DNA or RNA [109,110]. In 2014, Christopher E. Pearson and colleagues found that TMPyP4 could bind and distort the G4 formed by  $r(\text{GGGGCC})_8$ , inhibiting the interaction of some proteins with the repeat [23]. Several studies have shown that TMPyP4 disrupts the binding of hnRNPA1 to the  $r(\text{GGGGCC})_8$  repeat, that are supposed to link to ALS/FTD pathogenesis [23]. Therefore, it may be possible to develop therapeutic treatments using TMPyP4 to disrupt the interaction of RBPs. However, TMPyP4 may either stabilize or destabilize RNA G4. Kelly and colleagues used molecule dynamics simulations to analyze RNA G4 structure and speculated that TMPyP4 might interact with RNA G4 in three different ways: top-stacking, bottom-stacking, and side-binding, maintaining stability under certain conditions [111]. However, the specific structure and binding mode of the complex have not been reported. Therefore, further study on the interaction between TMPyP4 and  $r(\text{GGGGCC})_n$  RNA, as well as the destruction of RBPs binding which may cause toxicity, will be one of the directions for the development of related small molecule drugs.

#### 4.2. Binding of $r(\text{GGGGCC})_8$ with Other Ligands

Matthew D. Disney and colleagues has discovered three lead compounds, **1a**, **2**, and **3**, that bind with  $r(\text{GGGGCC})_8$  in vitro, with Kds of 9.7, 10, and 16  $\mu\text{M}$ , respectively [55]. These three small molecules were obtained by Hoechst or *bis*-benzimidazole query, and were derived from the small molecule library established by chemical similarity search. This library is enriched in compounds that have the potential to recognize RNA  $1 \times 1$  nucleotide internal loops, among which **1a** has been proven to bind  $1 \times 1$  GG internal loops present in  $r(\text{CGG})^{\text{exp}}$ , and improve fragile X-associated tremor/ataxia syndrome (FXTAS)-associated defects [112].

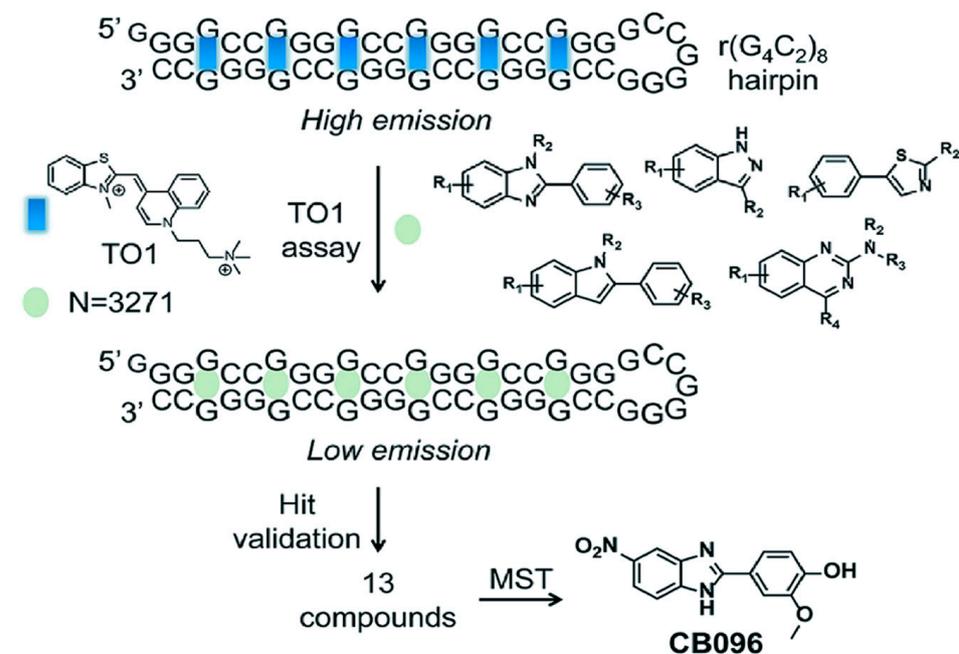
As  $r(\text{GGGGCC})_8$  RNA experiences dynamical equilibrium between hairpin and parallel G4 structure in solution, the binding constants of these lead compounds with RNA were evaluated in either  $\text{K}^+$  containing buffer (favorable for G4 structure) or  $\text{Na}^+$  buffer (favorable for hairpin). The 3–10 times higher Kds of **1a** and **3** were obtained in the presence of  $\text{K}^+$  than the  $\text{Na}^+$  buffer, demonstrating their favor binding to G4 structures of  $r(\text{GGGGCC})_8$ . In contrast, a  $\text{Na}^+$ -dependent affinity of **2** was not affected by  $r(\text{GGGGCC})_8$ , but it significantly decreased with  $\text{K}^+$ , showing the specific binding with hairpin structures. The optical melting data further demonstrated that compound **3** has no influence on the stability of  $r(\text{GGGGCC})_8$ , while compounds **1a** and **2** improve it.

The effects of three ligands on non-ATG translation of  $r(\text{GGGGCC})_n$  were tested in HEK293 cells expressing  $r(\text{GGGGCC})_{66}$  [55]. It was found that poly(GP) and poly(GA) proteins, but not poly(GR) proteins, were produced in the system. Compound **3** (100  $\mu\text{M}$ , 24 h) was shown to moderately limit poly(GP) synthesis while having no effect on poly(GA). Compounds **1a** and **2**, on the other hand, drastically reduced the amounts of GP and GA proteins, which dramatically lowered the percentage of positive cells in the lesions. This suggests that ligand binding to  $r(\text{GGGGCC})_n$  could be a potentially effective cure for FTD/ALS.

#### 4.3. Binding of $r(\text{GGGGCC})_8$ with CB096

Disney and colleagues discovered a benzimidazole derivative CB096 that binds to  $r(\text{GGGGCC})_n$ . NMR, structure–activity relationship (SAR) studies, and molecular dynamics (MD) simulations with  $r(\text{GGGGCC})_n$  hairpin structure have been used to determine the molecular interaction between CB096 and  $r(\text{GGGGCC})_n$  (Figure 7) [113]. When  $r(\text{GGGGCC})_n$  is folded, CB096 can specifically bind to the repeating  $1 \times 1$  GG inner ring structure of  $5'\text{CGG}\backslash 3'\text{GGC}$ . The TO-PRO-1 (TO-1) fluorescent dye replacement assay and microscale thermoelectrophoresis (MST) were used to screen the ligands bound to the  $r(\text{GGGGCC})_8$  hairpin. CB096 binds to  $5'\text{CGG}/3'\text{GGC}$  of the  $r(\text{GGGGCC})_n$  hairpin and

breaks the base pair as shown by NMR. To bind to the  $r(\text{GGGGCC})_n$  hairpin structure, the chemical 5'-NO<sub>2</sub> group and 2-methoxyphenyl are crucial. In ALS/HEK293T FTD's cells, CB096 slowed RAN translation and reduced poly(GP) DPR formation, but did not affect  $r(\text{GGGGCC})_{66}$  mRNA levels. In conclusion, the researchers showed that CB096 binds particularly to the  $1 \times 1$  GG inner ring 5'CGG\3'GGC generated during the expansion of  $r(\text{GGGGCC})_n$ .



**Figure 7.** CB096 specifically bound to the repeated  $1 \times 1$  GG inner loop structure 5'CGG/3'GGC in the  $r(\text{GGGGCC})_8$  hairpin structures. Figure 7 was reprinted from the reference [113].

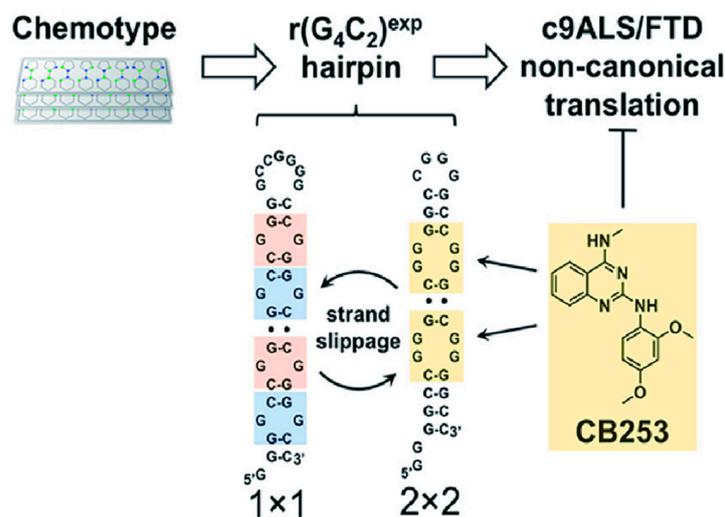
#### 4.4. Binding of $r(\text{GGGGCC})_n$ with DB1246, DB1247, and DB1273

Isaacs and colleagues screened a chemical library of small molecules to find the  $r(\text{GGGGCC})_4$  binding ligands [53]. They identified 44 hits out of 138 small molecules by a FRET-based G4 melting assay. Among those hitting compounds, three molecules are structurally similar (DB1246, DB1247, and DB1273) and have the ability to bind and stabilize G4s structure, as shown by temperature dependent CD spectroscopy [53]. Treatment with these compounds led to a significant reduction in both RNA foci formation and dipeptide repeat protein levels in *Drosophila* carrying  $r(\text{GGGGCC})_{36}$  and improved survival in vivo [53]. These findings suggest that targeting the  $r(\text{GGGGCC})_n$  G4 using small molecules may be a promising therapeutic approach to alleviate two key pathologies associated with FTD/ALS.

#### 4.5. Binding of $r(\text{GGGGCC})_n$ with CB253

Andrei and colleagues incorporated <sup>19</sup>F modified nucleotides to replace the C6 residue in  $r(\text{GGGGCC})_2$  duplex model (5'CCGGGG/3'GGGGCC) to investigate the binding mechanism of CB253 to  $r(\text{GGGGCC})_n$  (Figure 8) [114]. The replacement of <sup>19</sup>F nucleotide enables the use of <sup>19</sup>F NMR spectroscopy to investigate the structure and interactions. Two types of inner ring,  $1 \times 1$  GG and  $2 \times 2$  GG, were detected and verified in the  $r(\text{GGGGCC})_2$  hairpin structure. Among them, the  $1 \times 1$  GG was the main conformation, and the two conformations could slowly transform into each other to achieve an equilibrium. Addition of CB253 stabilizes the  $2 \times 2$  GG inner ring structure of  $r(\text{GGGGCC})_2$  duplex, which becomes a stable dominant conformation. CB253 can form key interactions with N1-H of G3 and combine with  $r(\text{GGGGCC})_2$  at a 2:1 ratio. The precise 2,4-diamino substitution pattern within CB253's quinazoline scaffold is crucial for binding the  $r(\text{GGGGCC})_n$  hairpin RNA. In HEK293T and lymphoblastoid cells from C9orf72 patients, CB253 reduced the

formation of stress granules induced by  $r(\text{GGGGCC})_{66}$  and inhibited RAN translation in a dose-dependent manner, leading to a significant reduction in poly(GP) DPR levels. These findings indicate that CB253 is a promising chemical probe that can specifically bind to and stabilize the  $2 \times 2$  GG inner ring of  $r(\text{GGGGCC})_n$  hairpin structure, and inhibit various *C9orf72*-specific pathological mechanisms by directly engaging  $r(\text{GGGGCC})_n$ .



**Figure 8.** CB253 that selectively binds the hairpin form of  $r(\text{GGGGCC})_n$ . Figure 8 was reprinted from the reference [114].

## 5. Summary and Perspective

In this review, we provide a comprehensive overview of the advancements in understanding the structure of  $r(\text{GGGGCC})_n$  and  $d(\text{GGGGCC})_n$ , the phase separation and transition of  $r(\text{GGGGCC})_n$ , the interactions of  $r(\text{GGGGCC})_n$  with RBPs, and the discovered ligands capable of inhibiting the non-ATG translations of  $r(\text{GGGGCC})_n$  and/or the interactions between  $r(\text{GGGGCC})_n$  and RBPs.

The relationship between the fatal neurodegenerative diseases ALS/FTD, the structure of  $r(\text{GGGGCC})_n$  RNA, and their interactions have garnered significant research attention. When the repeat number exceeds the threshold,  $r(\text{GGGGCC})_n$  RNA undergoes phase separation and transition, leading to the formation of nuclear RNA foci. These RNA foci recruit RBPs, disrupting the physiological functions of RNA splicing and maturation. Another pathogenic mechanism by which  $r(\text{GGGGCC})_n$  contributes to ALS or FTD is the cytotoxicity of repetitive dipeptide proteins generated through non-ATG translation. Aggregates of these repetitive dipeptide proteins, can recruit numerous 26S proteasome complexes and stabilize a transient substrate-processing conformation of the 26S proteasome, suggesting impaired degradation processes [115].

Characterizing the repeat structure of  $r(\text{GGGGCC})_n$  RNA and elucidating the structure-function relationship are key areas of research in understanding the pathogenic causes.  $r(\text{GGGGCC})_n$  can adopt diverse structures, including hairpin and parallel G4 topologies, with equilibrium between them depending on solution conditions. However, the three-dimensional structures of  $r(\text{GGGGCC})_n$  RNA are still unknown. Achieving a dominant conformation for structural studies may require sequence and solution condition optimization. Another challenging aspect is determining the secondary structures of  $r(\text{GGGGCC})_n$  within RNA foci or gel-like states. Due to the non-crystalline solid state and heterogeneous nature of RNA foci, commonly used high-resolution structure determination methods such as X-ray crystallography or solution NMR are not applicable [116,117]. To date, the RNA structures within RNA foci remain unidentified. Advancements in RNA structure determination methodologies, such as solid-state NMR [118–120], are needed to overcome this limitation.

Several small molecules that bind to r(GGGGCC)<sub>n</sub> have been discovered to block RBP interactions, inhibit phase separation, and/or hinder non-ATG translation, as evidenced both in vivo and in vitro. Understanding the structural details of the interactions between r(GGGGCC)<sub>n</sub> RNA and ligands is crucial for facilitating the design of lead compounds to treat ALS/FTD. Similar to the challenges faced in studying r(GGGGCC)<sub>n</sub> RNA, the complex structure determination of r(GGGGCC)<sub>n</sub> RNA repeats and small molecules is lacking, necessitating further developments to gain insights into drug design.

Another known treatment approach for ALS/FTD involves the use of antisense RNA. Single-dose injections of antisense oligonucleotides (ASOs) targeting repeat-containing RNAs, while preserving mRNA levels encoding *C9orf72*, have resulted in sustained reductions in RNA foci and dipeptide-repeat proteins, leading to the amelioration of behavioral deficits. These efforts have identified the gain of toxicity as a central disease mechanism caused by repeat-expanded *C9orf72* and established the feasibility of ASO-mediated therapy [16]. ALS brains treated with ASO therapeutics targeting the *C9orf72* transcript or repeat expansion showed mitigation despite the presence of repeat-associated non-ATG translation products [46]. Moreover, the introduction of mRNA that encodes r(GGGGCC)<sub>n</sub> binding proteins into ALS/FTD cells has the potential to restore RBP functions by augmenting the intracellular pool of RBPs recruited by RNA foci. This approach represents an alternative strategy for treating ALS by targeting r(GGGGCC)<sub>n</sub> RNA. Lastly, gene editing system by CRISPR/Cas9 has successfully removed the GGGGCC repeat expansion in *C9orf72*, leading to reduction in RNA foci and DPR formations, proving a promising approach in ALS treatments.

**Author Contributions:** Conceptualization, S.W. (Shenlin Wang); writing-original draft preparation: X.L., X.Z., J.H., X.S., Q.L., S.W. (Sishi Wang) and S.W. (Shenlin Wang); visualization, X.L., J.H. and X.Z.; writing-review and editing, X.L., J.H. and S.W. (Shenlin Wang); supervision, S.W. (Shenlin Wang); funding acquisition, S.W. (Shenlin Wang). All authors have read and agreed to the published version of the manuscript.

**Funding:** The work was supported by the National Natural Science Foundation of China (22274050) and the Fundamental Research Funds for the Central Universities.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Sample Availability:** Not applicable.

## References

1. Ling, S.-C.; Polymenidou, M.; Cleveland, D.W. Converging Mechanisms in ALS and FTD: Disrupted RNA and Protein Homeostasis. *Neuron* **2013**, *79*, 416–438. [[CrossRef](#)]
2. Sagui, C. Structure and Dynamics of DNA and RNA Double Helices Obtained from the GGGGCC and CCCC GG Hexanucleotide Repeats That Are the Hallmark of C9FTD/ALS Diseases. *ACS Chem. Neurosci.* **2017**, *8*, 578–591.
3. Ash, P.E.A.; Bieniek, K.F.; Gendron, T.F.; Caulfield, T.; Lin, W.-L.; DeJesus-Hernandez, M.; van Blitterswijk, M.M.; Jansen-West, K.; Paul, J.W., III; Rademakers, R.; et al. Unconventional Translation of *C9ORF72* GGGGCC Expansion Generates Insoluble Polypeptides Specific to c9FTD/ALS. *Neuron* **2013**, *77*, 639–646. [[CrossRef](#)] [[PubMed](#)]
4. van der Ende, E.L.; Jackson, J.L.; White, A.; Seelaar, H.; van Blitterswijk, M.; Van Swieten, J.C. Unravelling the clinical spectrum and the role of repeat length in *C9ORF72* repeat expansions. *J. Neurol. Neurosurg. Psychiatry* **2021**, *92*, 502–509. [[CrossRef](#)]
5. Wang, C.; Chen, Z.; Yang, F.; Jiao, B.; Peng, H.; Shi, Y.; Wang, Y.; Huang, F.; Wang, J.; Shen, L.; et al. Analysis of the GGGGCC Repeat Expansions of the *C9orf72* Gene in SCA3/MJD Patients from China. *PLoS ONE* **2015**, *10*, e0130336. [[CrossRef](#)]
6. Lee, Y.-B.; Chen, H.-J.; Peres, J.N.; Gomez-Deza, J.; Attig, J.; Štálekár, M.; Troakes, C.; Nishimura, A.L.; Scotter, E.L.; Vance, C.; et al. Hexanucleotide repeats in ALS/FTD form length-dependent RNA foci, sequester RNA binding proteins, and are neurotoxic. *Cell Rep.* **2013**, *5*, 1178–1186. [[CrossRef](#)]
7. Balendra, R.; Isaacs, A.M. *C9orf72*-mediated ALS and FTD: Multiple pathways to disease. *Nat. Rev. Neurol.* **2018**, *14*, 544–558. [[CrossRef](#)] [[PubMed](#)]

8. DeJesus-Hernandez, M.; Mackenzie, I.R.; Boeve, B.F.; Boxer, A.L.; Baker, M.; Rutherford, N.J.; Nicholson, A.M.; Finch, N.A.; Flynn, H.; Adamson, J.; et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of *C9ORF72* causes chromosome 9p-linked FTD and ALS. *Neuron* **2011**, *72*, 245–256. [[CrossRef](#)] [[PubMed](#)]
9. Dodd, K.C.; Power, R.; Ealing, J.; Hamdalla, H. FUS-ALS presenting with myoclonic jerks in a 17-year-old man. *Amyotroph. Lateral Scler. Front. Degener.* **2019**, *20*, 278–280. [[CrossRef](#)] [[PubMed](#)]
10. Zhou, B.; Liu, C.; Geng, Y.; Zhu, G. Topology of a G-quadruplex DNA formed by *C9orf72* hexanucleotide repeats associated with ALS and FTD. *Sci. Rep.* **2015**, *5*, 16673. [[CrossRef](#)]
11. Calcoen, D.; Elias, L.; Yu, X. What does it take to produce a breakthrough drug? *Nat. Rev. Drug Discov.* **2015**, *14*, 161–162.
12. Tamaki, Y.; Urushitani, M. Molecular Dissection of TDP-43 as a Leading Cause of ALS/FTLD. *Int. J. Mol. Sci.* **2022**, *23*, 12508. [[CrossRef](#)]
13. van Zundert, B.; Brown, R.H., Jr. Silencing strategies for therapy of SOD1-mediated ALS. *Neurosci. Lett.* **2017**, *636*, 32–39. [[CrossRef](#)]
14. Renton, A.E.; Majounie, E.; Waite, A.; Simon-Saánchez, J.; Rollinson, S.; Gibbs, J.R.; Schymick, J.C.; Laaksovirta, H.; van Swieten, J.C.; Myllykangas, L.; et al. A Hexanucleotide Repeat Expansion in *C9ORF72* Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72*, 257–268. [[CrossRef](#)] [[PubMed](#)]
15. O’rourke, J.G.; Bogdanik, L.; Muhammad, A.; Gendron, T.F.; Kim, K.J.; Austin, A.; Cady, J.; Liu, E.Y.; Zarrow, J.; Grant, S.; et al. *C9orf72* BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* **2015**, *88*, 892–901.
16. Jiang, J.; Zhu, Q.; Gendron, T.F.; Saberi, S.; McAlonis-Downes, M.; Seelman, A.; Stauffer, J.E.; Jafar-Nejad, P.; Drenner, K.; Schulte, D.; et al. Gain of Toxicity from ALS/FTD-Linked Repeat Expansions in *C9ORF72* Is Alleviated by Antisense Oligonucleotides Targeting GGGGCC-Containing RNAs. *Neuron* **2016**, *90*, 535–550. [[CrossRef](#)] [[PubMed](#)]
17. Zhu, Q.; Jiang, J.; Gendron, T.F.; McAlonis-Downes, M.; Jiang, L.; Taylor, A.; Garcia, S.D.; Dastidar, S.G.; Rodriguez, M.J.; King, P.; et al. Reduced *C9ORF72* function exacerbates gain of toxicity from ALS/FTD-causing repeat expansion in *C9orf72*. *Nat. Neurosci.* **2020**, *23*, 615–624. [[CrossRef](#)]
18. Rodriguez, C.; Todd, P. New pathologic mechanisms in nucleotide repeat expansion disorders. *Neurobiol. Dis.* **2019**, *130*, 104515. [[CrossRef](#)] [[PubMed](#)]
19. Zhang, K.; Daigle, J.G.; Cunningham, K.M.; Coyne, A.N.; Ruan, K.; Grima, J.C.; Bowen, K.E.; Wadhwa, H.; Yang, P.; Rigo, F.; et al. Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell* **2018**, *173*, 958–971. [[PubMed](#)]
20. Zaepfel, B.L.; Zhang, Z.; Maulding, K.; Coyne, A.N.; Cheng, W.; Hayes, L.R.; Lloyd, T.E.; Sun, S.; Rothstein, J.D. UPF1 reduces *C9orf72* HRE-induced neurotoxicity in the absence of nonsense-mediated decay dysfunction. *Cell Rep.* **2021**, *34*, 108925. [[CrossRef](#)]
21. Wen, X.; Westergard, T.; Pasinelli, P.; Trotti, D. Pathogenic determinants and mechanisms of ALS/FTD linked to hexanucleotide repeat expansions in the *C9orf72* gene. *Neurosci. Lett.* **2017**, *636*, 16–26. [[CrossRef](#)]
22. Mizielinska, S.; Grönke, S.; Niccoli, T.; Ridler, C.E.; Clayton, E.L.; Devoy, A.; Moens, T.; Norona, F.E.; Woollacott, I.O.C.; Pietrzyk, J.; et al. *C9orf72* repeat expansions cause neurodegeneration in *Drosophila* through arginine-rich proteins. *Science* **2014**, *345*, 1192–1194. [[CrossRef](#)]
23. Zamiri, B.; Reddy, K.; Macgregor, R.B., Jr.; Pearson, C.E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)<sub>n</sub> repeat of the *C9orf72* gene and blocks interaction of RNA-binding proteins. *J. Biol. Chem.* **2014**, *289*, 4653–4659. [[CrossRef](#)]
24. West, R.J.; Sharpe, J.L.; Voelzmann, A.; Munro, A.L.; Hahn, I.; Baines, R.A.; Pickering-Brown, S. Co-expression of *C9orf72* related dipeptide-repeats over 1000 repeat units reveals age- and combination-specific phenotypic profiles in *Drosophila*. *Acta Neuropathol. Commun.* **2020**, *8*, 158. [[CrossRef](#)]
25. Nordin, A.; Akimoto, C.; Wuolikainen, A.; Alstermark, H.; Jonsson, P.; Birve, A.; Marklund, S.L.; Graffmo, K.S.; Forsberg, K.; Brännström, T.; et al. Extensive size variability of the GGGGCC expansion in *C9orf72* in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Hum. Mol. Genet.* **2015**, *24*, 3133–3142. [[CrossRef](#)]
26. Goodman, L.D.; Bonini, N.M. New Roles for Canonical Transcription Factors in Repeat Expansion Diseases. *Trends Genet.* **2020**, *36*, 81–92. [[CrossRef](#)]
27. Wang, J.; Wang, B.; Zhou, T. The Advance on Frontotemporal Dementia (FTD)’s Neuropathology and Molecular Genetics. *Mediat. Inflamm.* **2022**, *2022*, 5003902. [[CrossRef](#)]
28. Abramzon, Y.A.; Fratta, P.; Traynor, B.J.; Chia, R. The Overlapping Genetics of Amyotrophic Lateral Sclerosis and Frontotemporal Dementia. *Front. Neurosci.* **2020**, *14*, 42. [[CrossRef](#)]
29. Echeverria, G.V.; Cooper, T.A. RNA-binding proteins in microsatellite expansion disorders: Mediators of RNA toxicity. *Brain Res.* **2012**, *1462*, 100–111. [[CrossRef](#)]
30. Gitler, A.D.; Tsuiji, H. There has been an awakening: Emerging mechanisms of *C9orf72* mutations in FTD/ALS. *Brain Res.* **2016**, *1647*, 19–29. [[CrossRef](#)]
31. Gijssels, I.; Van Langenhove, T.; van der Zee, J.; Slegers, K.; Philtjens, S.; Kleinberger, G.; Janssens, J.; Bettens, K.; Van Cauwenberghe, C.; Pereson, S.; et al. A *C9orf72* promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: A gene identification study. *Lancet Neurol.* **2012**, *11*, 54–65. [[CrossRef](#)] [[PubMed](#)]
32. Fay, M.M.; Anderson, P.J.; Ivanov, P. ALS/FTD-Associated *C9ORF72* Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Rep.* **2017**, *21*, 3573–3584. [[CrossRef](#)]

33. Mehta, A.R.; Selvaraj, B.T.; Barton, S.K.; McDade, K.; Abrahams, S.; Chandran, S.; Smith, C.; Gregory, J.M. Improved detection of RNA foci in *C9orf72* amyotrophic lateral sclerosis post-mortem tissue using BaseScope™ shows a lack of association with cognitive dysfunction. *Brain Commun.* **2020**, *2*, fcaa009. [[CrossRef](#)]
34. Malnar, M.; Rogelj, B. SFPQ regulates the accumulation of RNA foci and dipeptide repeat proteins from the expanded repeat mutation in *C9orf72*. *J. Cell Sci.* **2021**, *134*, jcs256602. [[CrossRef](#)]
35. Nedelsky, N.B.; Taylor, J.P. Bridging biophysics and neurology: Aberrant phase transitions in neurodegenerative disease. *Nat. Rev. Neurol.* **2019**, *15*, 272–286. [[CrossRef](#)]
36. Wang, X.; Goodrich, K.J.; Conlon, E.G.; Gao, J.; Erbse, A.; Manley, J.L.; Cech, T.R. *C9orf72* and triplet repeat disorder RNAs: G-quadruplex formation, binding to PRC2 and implications for disease mechanisms. *RNA* **2019**, *25*, 935–947. [[CrossRef](#)]
37. Swinnen, B.; Bento-Abreu, A.; Gendron, T.F.; Boeynaems, S.; Bogaert, E.; Nuyts, R.; Timmers, M.; Scheveneels, W.; Hersmus, N.; Wang, J.; et al. A zebrafish model for *C9orf72* ALS reveals RNA toxicity as a pathogenic mechanism. *Acta Neuropathol.* **2018**, *135*, 427–443. [[CrossRef](#)]
38. Tao, Z.; Wang, H.; Xia, Q.; Li, K.; Li, K.; Jiang, X.; Xu, G.; Wang, G.; Ying, Z. Nucleolar stress and impaired stress granule formation contribute to *C9orf72* RAN translation-induced cytotoxicity. *Hum. Mol. Genet.* **2015**, *24*, 2426–2441. [[CrossRef](#)]
39. Mori, K.; Weng, S.-M.; Arzberger, T.; May, S.; Rentzsch, K.; Kremmer, E.; Schmid, B.; Kretzschmar, H.A.; Cruts, M.; Van Broeckhoven, C.; et al. The *C9orf72* GGGGCC Repeat Is Translated into Aggregating Dipeptide-Repeat Proteins in FTL/ALS. *Science* **2013**, *339*, 1335–1338. [[CrossRef](#)]
40. Liu, H.; Lu, Y.-N.; Paul, T.; Periz, G.; Banco, M.T.; Ferré-D'amaré, A.R.; Rothstein, J.D.; Hayes, L.R.; Myong, S.; Wang, J. A Helicase Unwinds Hexanucleotide Repeat RNA G-Quadruplexes and Facilitates Repeat-Associated Non-AUG Translation. *J. Am. Chem. Soc.* **2021**, *143*, 7368–7379. [[CrossRef](#)]
41. Flores, B.N.; Dulchavsky, M.E.; Krans, A.; Sawaya, M.R.; Paulson, H.L.; Todd, P.K.; Barmada, S.J.; Ivanova, M.I. Distinct *C9orf72*-Associated Dipeptide Repeat Structures Correlate with Neuronal Toxicity. *PLoS ONE* **2016**, *11*, e0165084. [[CrossRef](#)]
42. Cheng, W.; Wang, S.; Mestre, A.A.; Fu, C.; Makarem, A.; Xian, F.; Hayes, L.R.; Lopez-Gonzalez, R.; Drenner, K.; Jiang, J.; et al. *C9ORF72* GGGGCC repeat-associated non-AUG translation is upregulated by stress through eIF2 $\alpha$  phosphorylation. *Nat. Commun.* **2018**, *9*, 51. [[CrossRef](#)]
43. Jain, A.; Vale, R.D. RNA phase transitions in repeat expansion disorders. *Nature* **2017**, *546*, 243–247. [[CrossRef](#)]
44. Penumutthu, S.R.; Chiu, L.-Y.; Meagher, J.L.; Hansen, A.L.; Stuckey, J.A.; Tolbert, B.S. Differential Conformational Dynamics Encoded by the Linker between Quasi RNA Recognition Motifs of Heterogeneous Nuclear Ribonucleoprotein H. *J. Am. Chem. Soc.* **2018**, *140*, 11661–11673. [[CrossRef](#)]
45. Celona, B.; Dollen, J.V.; Vatsavayai, S.C.; Kashima, R.; Johnson, J.R.; Tang, A.A.; Hata, A.; Miller, B.L.; Huang, E.J.; Krogan, N.J.; et al. Suppression of *C9orf72* RNA repeat-induced neurotoxicity by the ALS-associated RNA-binding protein Zfp106. *eLife* **2017**, *6*, e19032. [[CrossRef](#)]
46. Donnelly, C.J.; Zhang, P.W.; Pham, J.T.; Haeusler, A.R.; Mistry, N.A.; Vidensky, S.; Daley, E.L.; Poth, E.M.; Hoover, B.; Fines, D.M.; et al. RNA Toxicity from the ALS/FTD *C9ORF72* Expansion Is Mitigated by Antisense Intervention. *Neuron* **2013**, *80*, 415–428. [[CrossRef](#)]
47. Xu, Z.; Poidevin, M.; Li, X.; Li, Y.; Shu, L.; Nelson, D.L.; Li, H.; Hales, C.M.; Gearing, M.; Wingo, T.S.; et al. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 7778–7783. [[CrossRef](#)]
48. Shen, J.; Zhang, Y.; Zhao, S.; Mao, H.; Wang, Z.; Li, H.; Xu, Z. Pur $\alpha$  Repaired Expanded Hexanucleotide GGGGCC Repeat Noncoding RNA-Caused Neuronal Toxicity in Neuro-2a Cells. *Neurotox. Res.* **2018**, *33*, 693–701. [[CrossRef](#)]
49. Ishiguro, A.; Lu, J.; Ozawa, D.; Nagai, Y.; Ishihama, A. ALS-linked FUS mutations dysregulate G-quadruplex-dependent liquid-liquid phase separation and liquid-to-solid transition. *J. Biol. Chem.* **2021**, *297*, 101284. [[CrossRef](#)]
50. Murakami, T.; Qamar, S.; Lin, J.Q.; Schierle, G.S.K.; Rees, E.; Miyashita, A.; Costa, A.R.; Dodd, R.B.; Chan, F.T.S.; Michel, C.H.; et al. ALS/FTD Mutation-Induced Phase Transition of FUS Liquid Droplets and Reversible Hydrogels into Irreversible Hydrogels Impairs RNP Granule Function. *Neuron* **2015**, *88*, 678–690. [[CrossRef](#)]
51. Cooper-Knock, J.; Walsh, M.J.; Higginbottom, A.; Robin Highley, J.; Dickman, M.J.; Edbauer, D.; Ince, P.G.; Wharton, S.B.; Wilson, S.A.; Kirby, J.; et al. Sequestration of multiple RNA recognition motif-containing proteins by *C9orf72* repeat expansions. *Brain* **2014**, *137*, 2040–2051. [[CrossRef](#)]
52. Wang, Q.; Conlon, E.G.; Manley, J.L.; Rio, D.C. Widespread intron retention impairs protein homeostasis in *C9orf72* ALS brains. *Genome Res.* **2020**, *30*, 1705–1715. [[CrossRef](#)] [[PubMed](#)]
53. Simone, R.; Balendra, R.; Moens, T.G.; Preza, E.; Wilson, K.M.; Heslegrave, A.; Woodling, N.S.; Niccoli, T.; Gilbert-Jaramillo, J.; Abdelkarim, S.; et al. G-quadruplex-binding small molecules ameliorate *C9orf72* FTD/ALS pathology in vitro and in vivo. *EMBO Mol. Med.* **2018**, *10*, 22–31. [[CrossRef](#)] [[PubMed](#)]
54. Millevoi, S.; Moine, H.; Vagner, S. G-quadruplexes in RNA biology. *Wiley Interdiscip. Rev. RNA* **2012**, *3*, 495–507. [[CrossRef](#)]
55. Su, Z.; Zhang, Y.; Gendron, T.F.; Bauer, P.O.; Chew, J.; Yang, W.-Y.; Fostvedt, E.; Jansen-West, K.; Belzil, V.V.; Desaro, P.; et al. Discovery of a biomarker and lead small molecules to target r(GGGGCC)-associated defects in c9FTD/ALS. *Neuron* **2014**, *83*, 1043–1050. [[CrossRef](#)]
56. Jaiswal, M.K. Riluzole and edaravone: A tale of two amyotrophic lateral sclerosis drugs. *Med. Res. Rev.* **2019**, *39*, 733–748. [[CrossRef](#)]

57. Meijboom, K.E.; Abdallah, A.; Fordham, N.P.; Nagase, H.; Rodriguez, T.; Kraus, C.; Gendron, T.F.; Krishnan, G.; Esanov, R.; Andrade, N.S.; et al. CRISPR/Cas9-mediated excision of ALS/FTD-causing hexanucleotide repeat expansion in *C9ORF72* rescues major disease mechanisms in vivo and in vitro. *Nat. Commun.* **2022**, *13*, 6286. [[CrossRef](#)]
58. Maity, A.; Winnerdy, F.R.; Chen, G.; Phan, A.T. Duplexes Formed by  $G_4C_2$  Repeats Contain Alternate Slow- and Fast-Flipping G-G Base Pairs. *Biochemistry* **2021**, *60*, 1097–1107. [[CrossRef](#)]
59. Haeusler, A.R.; Donnelly, C.J.; Periz, G.; Simko, E.A.; Shaw, P.G.; Kim, M.-S.; Maragakis, N.J.; Troncoso, J.C.; Pandey, A.; Sattler, R.; et al. *C9orf72* nucleotide repeat structures initiate molecular cascades of disease. *Nature* **2014**, *507*, 195–200. [[CrossRef](#)]
60. Brčić, J.; Plavec, J. G-quadruplex formation of oligonucleotides containing ALS and FTD related GGGGCC repeat. *Front. Chem. Sci. Eng.* **2016**, *10*, 222–237. [[CrossRef](#)]
61. Fratta, P.; Mizielinska, S.; Nicoll, A.J.; Zloh, M.; Fisher, E.M.C.; Parkinson, G.; Isaacs, A.M. *C9orf72* hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2012**, *2*, srep01016. [[CrossRef](#)]
62. Reddy, K.; Zamiri, B.; Stanley, S.Y.; Macgregor, R.B.; Pearson, C.E. The disease-associated r(GGGGCC)<sub>n</sub> repeat from the *C9orf72* gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *J. Biol. Chem.* **2013**, *288*, 9860–9866. [[CrossRef](#)]
63. Božič, T.; Zalar, M.; Rogelj, B.; Plavec, J.; Šket, P. Structural Diversity of Sense and Antisense RNA Hexanucleotide Repeats Associated with ALS and FTL. *Molecules* **2020**, *25*, 525. [[CrossRef](#)]
64. Brčić, J.; Plavec, J. ALS and FTD linked GGGGCC-repeat containing DNA oligonucleotide folds into two distinct G-quadruplexes. *Biochim. Biophys. Acta (BBA)-Gen. Subj.* **2017**, *1861*, 1237–1245. [[CrossRef](#)]
65. Thys, R.G.; Wang, Y.-H. DNA Replication Dynamics of the GGGGCC Repeat of the *C9orf72* Gene. *J. Biol. Chem.* **2015**, *290*, 28953–28962. [[CrossRef](#)]
66. Brčić, J.; Plavec, J. Solution structure of a DNA quadruplex containing ALS and FTD related GGGGCC repeat stabilized by 8-bromodeoxyguanosine substitution. *Nucleic Acids Res.* **2015**, *43*, 8590–8600. [[CrossRef](#)]
67. Brčić, J.; Plavec, J. NMR structure of a G-quadruplex formed by four d( $G_4C_2$ ) repeats: Insights into structural polymorphism. *Nucleic Acids Res.* **2018**, *46*, 11605–11617. [[CrossRef](#)]
68. Geng, Y.; Liu, C.; Cai, Q.; Luo, Z.; Miao, H.; Shi, X.; Xu, N.; Fung, C.P.; Choy, T.T.; Yan, B.; et al. Crystal structure of parallel G-quadruplex formed by the two-repeat ALS- and FTD-related GGGGCC sequence. *Nucleic Acids Res.* **2021**, *49*, 5881–5890. [[CrossRef](#)]
69. Gao, Z.; Zhang, W.; Chang, R.; Zhang, S.; Yang, G.; Zhao, G. Liquid-Liquid Phase Separation: Unraveling the Enigma of Biomolecular Condensates in Microbial Cells. *Front. Microbiol.* **2021**, *12*, 751880. [[CrossRef](#)]
70. Boeynaems, S.; Alberti, S.; Fawzi, N.L.; Mittag, T.; Polymenidou, M.; Rousseau, F.; Schymkowitz, J.; Shorter, J.; Wolozin, B.; van den Bosch, L.; et al. Protein Phase Separation: A New Phase in Cell Biology. *Trends Cell Biol.* **2018**, *28*, 420–435. [[CrossRef](#)]
71. Bertrand, E.; Demongin, C.; Dobra, I.; Rengifo-Gonzalez, J.C.; Singatulina, A.S.; Sukhanova, M.V.; Lavrik, O.I.; Pastré, D.; Hamon, L. FUS fibrillation occurs through a nucleation-based process below the critical concentration required for liquid–liquid phase separation. *Sci. Rep.* **2023**, *13*, 7772. [[CrossRef](#)] [[PubMed](#)]
72. Watanabe, S.; Inami, H.; Oiwa, K.; Murata, Y.; Sakai, S.; Komine, O.; Sobue, A.; Iguchi, Y.; Katsuno, M.; Yamanaka, K. Aggresome formation and liquid-liquid phase separation independently induce cytoplasmic aggregation of TAR DNA-binding protein 43. *Cell Death Dis.* **2020**, *11*, 909. [[CrossRef](#)] [[PubMed](#)]
73. Shorter, J. Phase separation of RNA-binding proteins in physiology and disease: An introduction to the JBC Reviews thematic series. *J. Biol. Chem.* **2019**, *294*, 7113–7114. [[CrossRef](#)]
74. Portz, B.; Lee, B.L.; Shorter, J. FUS and TDP-43 Phases in Health and Disease. *Trends Biochem. Sci.* **2021**, *46*, 550–563. [[CrossRef](#)]
75. Maharana, S.; Wang, J.; Papadopoulos, D.K.; Richter, D.; Pozniakovskiy, A.; Poser, I.; Bickle, M.; Rizk, S.; Guillén-Boixet, J.; Franzmann, T.M.; et al. RNA buffers the phase separation behavior of prion-like RNA binding proteins. *Science* **2018**, *360*, 918–921. [[CrossRef](#)] [[PubMed](#)]
76. Markovtsov, V.; Nikolic, J.M.; Goldman, J.A.; Turck, C.W.; Chou, M.Y.; Black, D.L. Cooperative Assembly of an hnRNP Complex Induced by a Tissue-Specific Homolog of Polypyrimidine Tract Binding Protein. *Mol. Cell. Biol.* **2000**, *20*, 7463–7479. [[CrossRef](#)]
77. Wang, E.; Cambi, F. Heterogeneous Nuclear Ribonucleoproteins H and F Regulate the Proteolipid Protein/DM20 Ratio by Recruiting U1 Small Nuclear Ribonucleoprotein through a Complex Array of G Runs. *J. Biol. Chem.* **2009**, *284*, 11194–11204. [[CrossRef](#)]
78. Van Dusen, C.M.; Yee, L.; McNally, L.M.; McNally, M.T. A glycine-rich domain of hnRNP H/F promotes nucleocytoplasmic shuttling and nuclear import through an interaction with transportin 1. *Mol. Cell. Biol.* **2010**, *30*, 2552–2562. [[CrossRef](#)]
79. Conlon, E.G.; Lu, L.; Sharma, A.; Yamazaki, T.; Tang, T.; Shneider, N.A.; Manley, J.L. The *C9ORF72* GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *eLife* **2016**, *5*, e17820. [[CrossRef](#)]
80. Liao, Y.-Z.; Ma, J.; Dou, J.-Z. The Role of TDP-43 in Neurodegenerative Disease. *Mol. Neurobiol.* **2022**, *59*, 4223–4241. [[CrossRef](#)]
81. Lee, E.B.; Lee, V.M.-Y.; Trojanowski, J.Q. Gains or losses: Molecular mechanisms of TDP43-mediated neurodegeneration. *Nat. Rev. Neurosci.* **2011**, *13*, 38–50. [[CrossRef](#)] [[PubMed](#)]
82. Conicella, A.E.; Zerze, G.H.; Mittal, J.; Fawzi, N.L. ALS Mutations Disrupt Phase Separation Mediated by  $\alpha$ -Helical Structure in the TDP-43 Low-Complexity C-Terminal Domain. *Structure* **2016**, *24*, 1537–1549. [[CrossRef](#)] [[PubMed](#)]

83. Chew, J.; Cook, C.; Gendron, T.F.; Jansen-West, K.; del Rosso, G.; Daugherty, L.M.; Castanedes-Casey, M.; Kurti, A.; Stankowski, J.N.; Disney, M.D.; et al. Aberrant deposition of stress granule-resident proteins linked to *C9orf72*-associated TDP-43 proteinopathy. *Mol. Neurodegener.* **2019**, *14*, 9. [[PubMed](#)]
84. Solomon, D.A.; Stepto, A.; Au, W.H.; Adachi, Y.; Diaper, D.C.; Hall, R.; Rekhi, A.; Boudi, A.; Tziortzouda, P.; Lee, Y.B.; et al. A feedback loop between dipeptide-repeat protein, TDP-43 and karyopherin- $\alpha$  mediates *C9orf72*-related neurodegeneration. *Brain* **2018**, *141*, 2908–2924. [[CrossRef](#)]
85. Mori, K.; Lammich, S.; Mackenzie, I.R.A.; Forné, I.; Zilow, S.; Kretzschmar, H.; Edbauer, D.; Janssens, J.; Kleinberger, G.; Cruts, M.; et al. hnRNP A3 binds to GGGGCC repeats and is a constituent of p62-positive/TDP43-negative inclusions in the hippocampus of patients with *C9orf72* mutations. *Acta Neuropathol.* **2013**, *125*, 413–423.
86. Peters, O.M.; Cabrera, G.T.; Tran, H.; Gendron, T.F.; McKeon, J.E.; Metterville, J.; Weiss, A.; Wightman, N.; Salameh, J.; Kim, J.; et al. Human *C9ORF72* Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* **2015**, *88*, 902–909.
87. Lagier-Tourenne, C.; Polymenidou, M.; Cleveland, D.W. TDP-43 and FUS/TLS: Emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **2010**, *19*, R46–R64. [[CrossRef](#)]
88. Kanai, Y.; Dohmae, N.; Hirokawa, N. Kinesin transports RNA: Isolation and characterization of an RNA-transporting granule. *Neuron* **2004**, *43*, 513–525. [[CrossRef](#)]
89. Yoshimura, A.; Fujii, R.; Watanabe, Y.; Okabe, S.; Fukui, K.; Takumi, T. Myosin-Va facilitates the accumulation of mRNA/protein complex in dendritic spines. *Curr. Biol.* **2006**, *16*, 2345–2351.
90. Vance, C.; Rogelj, B.; Hortobágyi, T.; De Vos, K.J.; Nishimura, A.L.; Sreedharan, J.; Hu, X.; Smith, B.; Ruddy, D.; Wright, P.; et al. Mutations in FUS, an RNA Processing Protein, Cause Familial Amyotrophic Lateral Sclerosis Type 6. *Science* **2009**, *323*, 1208–1211.
91. Niaki, A.G.; Sarkar, J.; Cai, X.; Rhine, K.; Vidaurre, V.; Guy, B.; Hurst, M.; Lee, J.C.; Koh, H.R.; Guo, L.; et al. Loss of Dynamic RNA Interaction and Aberrant Phase Separation Induced by Two Distinct Types of ALS/FTD-Linked FUS Mutations. *Mol. Cell* **2020**, *77*, 82–94. [[CrossRef](#)] [[PubMed](#)]
92. Grasberger, H.; Bell, G.I. Subcellular recruitment by TSG118 and TSPYL implicates a role for zinc finger protein 106 in a novel developmental pathway. *Int. J. Biochem. Cell Biol.* **2005**, *37*, 1421–1437. [[CrossRef](#)] [[PubMed](#)]
93. Anderson, D.M.; Cannavino, J.; Li, H.; Anderson, K.M.; Nelson, B.R.; McAnally, J.; Bezprozvannaya, S.; Liu, Y.; Lin, W.; Liu, N.; et al. Severe muscle wasting and denervation in mice lacking the RNA-binding protein ZFP106. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E4494–E4503. [[CrossRef](#)] [[PubMed](#)]
94. reibaum, B.D.; Lu, Y.; Lopez-Gonzalez, R.; Kim, N.C.; Almeida, S.; Lee, K.-H.; Badders, N.; Valentine, M.; Miller, B.L.; Wong, P.C.; et al. GGGGCC repeat expansion in *C9orf72* compromises nucleocytoplasmic transport. *Nature* **2015**, *525*, 129–133.
95. Hideyama, T.; Yamashita, T.; Aizawa, H.; Tsuji, S.; Kakita, A.; Takahashi, H.; Kwak, S. Profound downregulation of the RNA editing enzyme ADAR2 in ALS spinal motor neurons. *Neurobiol. Dis.* **2012**, *45*, 1121–1128. [[CrossRef](#)]
96. Chen, C.-X.; Cho, D.-S.C.; Wang, Q.; Lai, F.; Carter, K.C.; Nishikura, K. A third member of the RNA-specific adenosine deaminase gene family, ADAR3, contains both single- and double-stranded RNA binding domains. *RNA* **2000**, *6*, 755–767.
97. White, M.K.; Johnson, E.M.; Khalili, K. Multiple roles for Pur- $\alpha$  in cellular and viral regulation. *Cell Cycle* **2009**, *8*, 414–420. [[CrossRef](#)]
98. Johnson, E.M.; Kinoshita, Y.; Weinreb, D.B.; Wortman, M.J.; Simon, R.; Khalili, K.; Winckler, B.; Gordon, J. Role of Pur $\alpha$  in targeting mRNA to sites of translation in hippocampal neuronal dendrites. *J. Neurosci. Res.* **2006**, *83*, 929–943. [[CrossRef](#)]
99. Daniel, D.C.; Wortman, M.J.; Schiller, R.J.; Liu, H.; Gan, L.; Mellen, J.S.; Chang, C.F.; Gallia, G.L.; Rappaport, J.; Khalili, K.; et al. Coordinate effects of human immunodeficiency virus type 1 protein Tat and cellular protein Pur $\alpha$  on DNA replication initiated at the JC virus origin. *J. Gen. Virol.* **2001**, *82*, 1543–1553. [[CrossRef](#)]
100. Daniel, D.C.; Johnson, E.M. PURA, the gene encoding Pur-alpha, member of an ancient nucleic acid-binding protein family with mammalian neurological functions. *Gene* **2018**, *643*, 133–143. [[CrossRef](#)]
101. Khalili, K.; Del Valle, L.; Muralidharan, V.; Gault, W.J.; Darbinian, N.; Otte, J.; Meier, E.; Johnson, E.M.; Daniel, D.C.; Kinoshita, Y.; et al. Pur $\alpha$  is essential for postnatal brain development and developmentally coupled cellular proliferation as revealed by genetic inactivation in the mouse. *Mol. Cell. Biol.* **2003**, *23*, 6857–6875. [[CrossRef](#)]
102. Barbe, M.F.; Krueger, J.J.; Loomis, R.; Otte, J.; Gordon, J. Memory Deficits, Gait Ataxia and Neuronal Loss in the Hippocampus and Cerebellum in Mice That Are Heterozygous for Pur-Alpha. *Neuroscience* **2016**, *337*, 177–190. [[CrossRef](#)]
103. Rossi, S.; Serrano, A.; Gerbino, V.; Giorgi, A.; Di Francesco, L.; Nencini, M.; Bozzo, F.; Schininà, M.E.; Bagni, C.; Cestra, G.; et al. Nuclear accumulation of mRNAs underlies G<sub>4</sub>C<sub>2</sub> repeat-induced translational repression in a cellular model of *C9orf72* ALS. *J. Cell Sci.* **2015**, *128*, 1787–1799. [[CrossRef](#)]
104. Daigle, J.G.; Krishnamurthy, K.; Ramesh, N.; Casci, I.; Monaghan, J.; McAvoy, K.; Godfrey, E.W.; Daniel, D.C.; Johnson, E.M.; Monahan, Z.; et al. Pur-alpha regulates cytoplasmic stress granule dynamics and ameliorates FUS toxicity. *Acta Neuropathol.* **2016**, *131*, 605–620. [[CrossRef](#)]
105. Di Salvio, M.; Piccinni, V.; Gerbino, V.; Mantoni, F.; Camerini, S.; Lenzi, J.; Rosa, A.; Chellini, L.; Loreni, F.; Carri, M.T.; et al. Pur-alpha functionally interacts with FUS carrying ALS-associated mutations. *Cell Death Dis.* **2015**, *6*, e1943. [[CrossRef](#)]
106. Muralidharan, V.; Sweet, T.; Nadraga, Y.; Amini, S.; Khalili, K. Regulation of Pur $\alpha$  gene transcription: Evidence for autoregulation of Pur $\alpha$  promoter. *J. Cell. Physiol.* **2001**, *186*, 406–413. [[CrossRef](#)]

107. Martino, L.; Pagano, B.; Fotticchia, I.; Neidle, S.; Giancola, C. Shedding Light on the Interaction between TMPyP4 and Human Telomeric Quadruplexes. *J. Phys. Chem. B* **2009**, *113*, 14779–14786. [[CrossRef](#)] [[PubMed](#)]
108. Morris, M.J.; Wingate, K.L.; Silwal, J.; Leeper, T.C.; Basu, S. The porphyrin TmPyP4 unfolds the extremely stable G-quadruplex in MT3-MMP mRNA and alleviates its repressive effect to enhance translation in eukaryotic cells. *Nucleic Acids Res.* **2012**, *40*, 4137–4145. [[CrossRef](#)] [[PubMed](#)]
109. Zahler, A.M.; Williamson, J.R.; Cech, T.R.; Prescott, D.M. Inhibition of telomerase by G-quartet DNA structures. *Nature* **1991**, *350*, 718–720. [[CrossRef](#)] [[PubMed](#)]
110. Izbicka, E.; Wheelhouse, R.; Raymond, E.; Davidson, K.K.; Lawrence, R.A.; Sun, D.; Windle, B.E.; Hurley, L.H.; Von Hoff, D.D. Effects of cationic porphyrins as G-quadruplex interactive agents in human tumor cells. *Cancer Res.* **1999**, *59*, 639–644.
111. Mulholland, K.; Sullivan, H.J.; Garner, J.; Cai, J.; Chen, B.; Wu, C. Three-Dimensional Structure of RNA Monomeric G-Quadruplex Containing ALS and FTD Related G<sub>4</sub>C<sub>2</sub> Repeat and Its Binding with TMPyP4 Probed by Homology Modeling based on Experimental Constraints and Molecular Dynamics Simulations. *ACS Chem. Neurosci.* **2020**, *11*, 57–75. [[CrossRef](#)]
112. Disney, M.D.; Liu, B.; Yang, W.-Y.; Sellier, C.; Tran, T.; Charlet-Berguerand, N.; Childs-Disney, J.L. A small molecule that targets r(CGG)<sup>exp</sup> and improves defects in fragile X-associated tremor ataxia syndrome. *ACS Chem. Biol.* **2012**, *7*, 1711–1718. [[CrossRef](#)]
113. Ursu, A.; Wang, K.W.; Bush, J.A.; Choudhary, S.; Chen, J.L.; Baisden, J.T.; Zhang, Y.J.; Gendron, T.F.; Petrucelli, L.; Yildirim, I.; et al. Structural Features of Small Molecules Targeting the RNA Repeat Expansion That Causes Genetically Defined ALS/FTD. *ACS Chem. Biol.* **2020**, *15*, 3112–3123. [[CrossRef](#)] [[PubMed](#)]
114. Ursu, A.; Baisden, J.T.; Bush, J.A.; Taghavi, A.; Choudhary, S.; Zhang, Y.J.; Gendron, T.F.; Petrucelli, L.; Yildirim, I.; Disney, M.D. A Small Molecule Exploits Hidden Structural Features within the RNA Repeat Expansion That Causes c9ALS/FTD and Rescues Pathological Hallmarks. *ACS Chem. Neurosci.* **2021**, *12*, 4076–4089. [[CrossRef](#)]
115. Guo, Q.; Lehmer, C.; Martínez-Sánchez, A.; Rudack, T.; Beck, F.; Hartmann, H.; Pérez-Berlanga, M.; Frottin, F.; Hipp, M.S.; Hartl, F.U.; et al. In Situ Structure of Neuronal C9orf72 Poly-GA Aggregates Reveals Proteasome Recruitment. *Cell* **2018**, *172*, 696–705. [[CrossRef](#)] [[PubMed](#)]
116. Zhao, S.; Yang, Y.; Zhao, Y.; Li, X.; Xue, Y.; Wang, S. High-resolution solid-state NMR spectroscopy of hydrated non-crystallized RNA. *Chem. Commun.* **2019**, *55*, 13991–13994. [[CrossRef](#)] [[PubMed](#)]
117. Marchanka, A.; Simon, B.; Althoff-Ospelt, G.; Carlomagno, T. RNA structure determination by solid-state NMR spectroscopy. *Nat. Commun.* **2015**, *6*, 7024. [[CrossRef](#)] [[PubMed](#)]
118. Zhao, S.; Li, X.; Wen, Z.; Zou, M.; Yu, G.; Liu, X.; Mao, J.; Zhang, L.; Xue, Y.; Fu, R.; et al. Dynamics of base pairs with low stability in RNA by solid-state nuclear magnetic resonance exchange spectroscopy. *iScience* **2022**, *25*, 105322. [[CrossRef](#)]
119. Yang, Y.; Wang, S. RNA Characterization by Solid-State NMR Spectroscopy. *Chemistry* **2018**, *24*, 8698–8707. [[CrossRef](#)]
120. Yang, Y.; Xiang, S.; Liu, X.; Pei, X.; Wu, P.; Gong, Q.; Li, N.; Baldus, M.; Wang, S. Proton-detected solid-state NMR detects the inter-nucleotide correlations and architecture of dimeric RNA in microcrystals. *Chem. Commun.* **2017**, *53*, 12886–12889. [[CrossRef](#)]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.