Review

G-Quadruplex Structures Formed by Human Telomere and *C9orf72* GGGGCC Repeats

Bing Yan [1], Monica Ching Suen [1], Naining Xu [1,2], Chao Lu [3], Changdong Liu [1,2,*] and Guang Zhu [1,2,*]

1   State Key Laboratory of Molecular Neuroscience, Institute for Advanced Study, Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong SAR, China; byanab@connect.ust.hk (B.Y.); cmsuen@connect.ust.hk (M.C.S.); xunaining@ust.hk (N.X.)
2   HKUST Shenzhen Research Institute, Hi-Tech Park, Nanshan, Shenzhen 518057, China
3   Department of Chemistry, Faculty of Science, The University of Hong Kong, Hong Kong SAR, China; luchaohk@connect.hku.hk
*   Correspondence: lcd@ust.hk (C.L.); gzhu@ust.hk (G.Z.)

Abstract: G-quadruplexes (G4s) are unique nucleic acid structures composed of guanine-rich (G-rich) sequences that can form diverse topologies based on the arrangement of their four strands. G4s have attracted attention for their potential roles in various biological processes and human diseases. In this review, we focus on the G4 structures formed by human telomeric sequences, $(GGGTTA)_n$, and the hexanucleotide repeat expansion, $(GGGGCC)_n$, in the first intron region of the chromosome 9 open reading frame 72 (*C9orf72*) gene, highlighting their structural diversity and biological significance. Human telomeric G4s play crucial roles in telomere retention and gene regulation. In particular, we provide an in-depth summary of known telomeric G4s and focus on our recently discovered chair-type conformation, which exhibits distinct folding patterns. The chair-type G4s represent a novel folding pattern with unique characteristics, expanding our knowledge of telomeric G4 structural diversity and potential biological functions. Specifically, we emphasize the G4s formed by the $(GGGGCC)_n$ sequence of the *C9orf72* gene, which represents the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). The thorough structural analysis in this review advances our comprehension of the disease mechanism and provides valuable insights into developing targeted therapeutic strategies in ALS/FTD.

Keywords: G-quadruplex; telomere; neurodegenerative disease; *C9orf72*; ALS/FTD

1. Introduction

G-quadruplexes (G4s) represent remarkable four-stranded helical structures that emerge from G-rich nucleotide sequences, characterized by distinctively stacked two or more G-tetrad planes that comprise four guanines interconnected through Hoogsteen hydrogen bonds [1,2]. The structural integrity of G4s is preserved mostly by coordination through monovalent cations, typically, K^+ and Na^+, which occupy central ionic channels and confer necessary stabilization [3]. G4s can adopt several different conformations that can be classified as parallel, antiparallel, or hybrid types by the relative orientation of the strands [4]. Factors that influence this structural polymorphism include sequence composition, loop length, ionic conditions, and molecular crowding effects [5–8].

Currently, over 700,000 regions capable of forming G4 structures throughout the human genome have been extensively documented by computational and deep-sequencing

approaches, which are notably enriched in functionally significant genomic regions, including promoters, telomeres, and regulatory elements [9–11]. The physiological relevance of G4 structures has been definitively established through multiple experimental approaches, including immunological detection, single-molecule fluorescence visualization, and small-molecule probing methodologies [12,13]. Accumulated studies have progressively unveiled the pivotal roles of G4s in biological processes including telomere retention, replication dynamics, transcriptional regulation, and genome stability [14–16].

Importantly, dysfunction in G4-mediated processes has been implicated in numerous pathological conditions, particularly various cancers and neurodegenerative diseases [17,18]. For example, the *C9orf72* hexanucleotide repeat expansion forms stable G4 structures linked to the neurodegenerative diseases ALS and FTD [17]. However, the G4 motifs in *FMR1* contribute to Fragile X syndrome [19]. Inherited disorders such as Werner and Bloom syndromes demonstrate compromised G4 resolution capabilities, highlighting their role in genomic stability [20]. Interestingly, the G4 elements discovered in *c-MYC* and *KRAS* genes regulate the biological function of these key oncogenes influencing cancer progression [21]. In particular, the telomeric G4 structures, which are crucial factors in cancer development, impact cellular senescence and immortalization [22]. These studies on G4s expand our understanding of their pathological significance and potential therapeutic applications. Notably, G4 structures have also been identified within viral genomes [23], such as SARS-CoV-2 [24,25], where their vital regulatory functions in viral processes and viral latency have been demonstrated [26]. These findings have paved new avenues for antiviral therapeutic strategies. Telomeres and the *C9orf72* gene are pivotal models for studying G4 structures due to their unique structural properties, biological significance, and direct implications in human diseases.

Moreover, extensive efforts have been made to target G4s as a therapeutic approach given their profound implication in diseases such as cancer and neurodegenerative diseases [27–29]. Early examples of G4-targeting ligands, such as quarfloxin and TMPyP4, demonstrated initial promise but faced significant limitations [21,30]. Recent developments have introduced more selective and viable G4 ligands, such as naphthalene diimides (NDIs), standing out as potent G4-binding ligands due to their ability to simultaneously target multiple G4s, exhibiting strong and selective anticancer activity [31]. Additionally, the tetra-substituted naphthalene diimide QN-302 is currently in clinical trials for pancreatic ductal adenocarcinoma (PDAC), highlighting its potential as a first-in-class G4-targeting therapeutic [32]. However, G4 conformations characterized by high precision are the key step of the therapeutic strategy. Therefore, the structure of G4s at high resolution is not only crucial for elucidating the pathogenic mechanisms of G4-related diseases, but also significant for the design of drugs targeting G4s in the treatment of such diseases. Specially, structural biology techniques, including high-resolution NMR spectroscopy and X-ray crystallography, combined with computational approaches have highlighted the structural diversity and dynamics of G4s, which has become crucial in developing novel therapeutic strategies for various G4-related pathologies.

Telomeric DNA sequences have a high propensity to form stable G4 structures under physiological conditions, and these G4 structures are well characterized, providing a robust model for studying G4 formation and stabilization [33]. Furthermore, the (GGGGCC)n repeat in *C9orf72* can form both unimolecular and multimeric G4s, which are highly stable and can drive the formation of biomolecular condensates [34–36]. The ability of these repeats to form G4s has been well documented, making them an excellent model for studying the structural and functional roles of G4s. In particular, small molecules that stabilize G4s in telomere and *C9orf72* have shown promise in inhibiting telomerase activity for cancer therapy and reducing the toxic effects of the repeat expansion in ALS/FTD models [27,28]. These studies demonstrate the potential of targeting G4s for developing

drugs that can treat G4-related diseases. Therefore, G4s in telomere and *C9orf72* are particularly relevant models for studying G4s due to their unique structural properties, biological significance, and therapeutic potential. These models provide valuable insights into the role of G4 in both normal cellular functions and disease pathogenesis.

2. Human Telomeric G4 Structures

Telomeres, the repetitive DNA sequences located at the termini of eukaryotic chromosomes, are crucial to preserve chromosomal stability and genome integrity [37,38]. Extensive research has highlighted the significance of human telomeric DNA in cancer development and cellular aging mechanisms [39,40]. Telomeres in humans comprise tandem GGGTTA repeats spanning up to 10,000 base pairs, making them the richest sources of sequences that can form G4s [16]. In recent decades, the structural diversity and biological significance of human telomeric G4s have been extensively investigated. It is well known that the human telomeric d[GGGTTA]n sequence can fold into parallel, hybrid, and antiparallel structures including basket- and chair-type and (2 + 2) G4s [8,41–49]. We briefly summarized the reported telomeric G4 structures listed in Table 1. These diverse structural arrangements highlight the adaptability of telomeric sequences in response to different cellular environments and regulatory requirements.
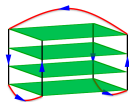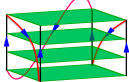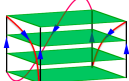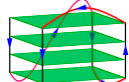
Table 1. Telomeric G4 structures.
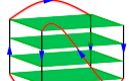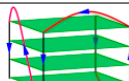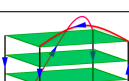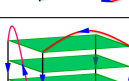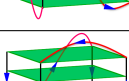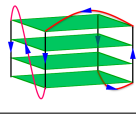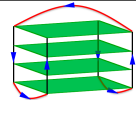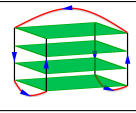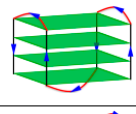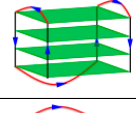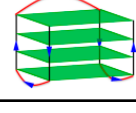
PDB ID	Sequence	Structure Type	Solution Conditions	Method	Structure
143D	AGGG(TTAGGG)$_3$	antiparallel basket-type	Na$^+$	NMR	
1KF1	d[(AGGGTT)$_3$AGGG]	parallel	K$^+$	XRD	
1K8P	d[^{Br}UAGGG^{Br}UTAGGGT]	parallel	K$^+$/Na$^+$	XRD	
2GKU	d[TTGGG(TTAGGG)$_3$A]	hybrid	K$^+$	NMR	
2HY9	d[AAAGGG(TTAGGG)$_3$AA]	hybrid-1	K$^+$	NMR	
2JPZ	d[(TTAGGG)$_4$TT]	hybrid-2	K$^+$	NMR	
2JSM	d[TAGGG(TTAGGG)$_3$]	hybrid-1	K$^+$	NMR	
2JSL	d[TAGGG(TTAGGG)$_3$TT]	hybrid-2	K$^+$	NMR	
2JSK	d[TAGGGTTAGGGTTAG(^{Br}G)GTTAGGG]	hybrid-1	K$^+$	NMR	

Table 1. *Cont.*

PDB ID	Sequence	Structure Type	Solution Conditions	Method	Structure
2JSQ	d[TAGGGTTAGGGTTA(^{Br}G)GGTTAGGGTT]	hybrid-2	K^+	NMR	
2KF8	d[(GGGTTA)$_3$GGGT]	basket-type	K^+	NMR	
2KF7	d[GGGTTA(^{Br}G)GGTTAGGGTTAGGGT]	basket-type	K^+	NMR	
2KKA	d[(AGGGTT)$_2$AIGGTTAGGGT]	basket-type	K^+	NMR	
5YEY	d[(GGGTTA)$_2$GGGTTTGGG]	chair-type	K^+	NMR	
6JKN	d[GGGTTAG(^{Br}G)GTTAGGGTTAG(^{Br}G)G]	chair-type	K^+	XRD	

NMR: Nuclear magnetic resonance, XRD: X-ray diffraction.

Recently, we reported a novel chair-type G4 structure adopted by *htel21*_Br-8,20, d[GGGTTAG(^{Br}G$_8$)GTTAGGGTTAG(^{Br}G$_{20}$)G], two 8Br-dG substitutions at 8 and 20 positions of *htel21* (Figure 1A,C) [50]. This three-layer chair-type G4 features trinucleotide edgewise loops, with T5/T17 forming a water-mediated sandwich structure through K^+ coordination and A6·A18 pairing, while a Hoogsteen A12·T10 pair caps the G-tetrad core, resulting in antiparallel strand arrangements. Interestingly, we also discovered that the same chair-type structure can be adopted by *htel21*T$_{18}$, d[(GGGTTA)$_2$GGGTTTGGG], a T substitution at A18 of *htel21* [51] (Figure 1B,D). This variant DNA adopts three-layer chair-type G4 featuring edgewise loops connected via reverse Watson–Crick A6·T18 pairing, with characteristic *syn·anti·syn·anti* glycosidic conformations in each G-terad. Strikingly, *htel21*_Br-8,20 and *htel21*T$_{18}$, while sharing identical overall chair-type topology, exhibit reversed donor–acceptor directionalities in their individual G-tetrad layers (Figure 1C,D), demonstrating the subtle complexity of G4 structural arrangements.

Furthermore, we discovered that the human cell division cycle 6 (Cdc6) protein, a crucial DNA initiation factor, specifically recognizes the *htel21*T$_{18}$ G4 structure [52]. The Cdc6 protein is composed of an N-terminal intrinsically disordered region (IDR), an ATPase-active AAA+ domain, and a C-terminal winged-helix domain (WHD). Through comprehensive biophysical analyses, we demonstrated that the N-terminal IDR (residues 7–20) of Cdc6 specifically binds to *htel21*T$_{18}$. Our NMR structure of Cdc6 7–20 in complex with *htel21*T$_{18}$ shows a hook-type conformation of the Cdc6 peptide, which is located at the bottom of *htel21*T$_{18}$ (Figure 1E). The high-resolution NMR structure reveals the key residues including F14, P15, K16, R17, K18, L19, and S20, which are vital for Cdc6 in the recognition of G4. The biological significance of this interaction was confirmed through in vitro experiments demonstrating Cdc6 co-localization with G4 structures in cells [52]. These findings establish a previously uncharacterized mechanism by which Cdc6 recognizes specific G4 structures, potentially linking DNA replication regulation with G4-mediated genomic processes.
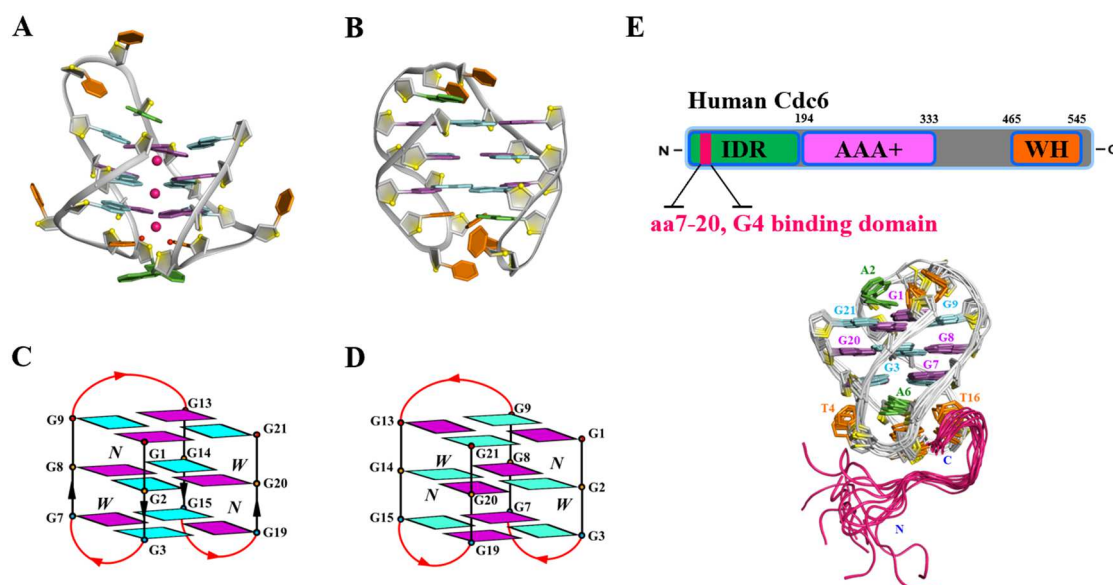
Figure 1. G4s formed by human telomeric DNA sequences and its complex with protein Cdc6. (**A**) The chair-type G4 structure formed by *htel21*_Br-8,20, d[GGGTTAG(^{Br}G$_8$)GTTAGGGTTAG(^{Br}G$_{20}$)G], in K$^+$ solution (PDB: 6JKN). (**B**) The chair-type G4 structure formed by *htel21*T$_{18}$, d[(GGGTTA)$_2$GGGTTTGGG], in K$^+$ solution (PDB: 5YEY). (**C**) The topology of *htel21*_Br-8,20. (**D**) The topology of *htel21*T$_{18}$. (**E**) Schematic representation delineating the domain organization of human Cdc6 and the NMR structure of Cdc6 7–20 in complex with *htel21*T$_{18}$.

The study of telomeric G4 structures and their protein interactions have significant implications for understanding telomere biology and regulation. The discovery of the chair-type topology in telomeric G4 represents a significant advancement in studying G4 structural diversity. The high-resolution structure of the Cdc6-*htel21*T$_{18}$ complex has revealed detailed molecular mechanisms of protein–G4 recognition, including specific base contacts and conformational adaptations, providing novel perspectives on the biological roles of telomeric G4 structures in DNA replication and cell cycle control. From a therapeutic perspective, these structural insights provide a foundation for designing small molecules or peptide-like compounds that could either stabilize specific G4s or modulate protein–G4 interactions, offering valuable insights for developing targeted drugs to treat diseases associated with telomeric G4s.

3. G4s Formed by *C9orf72* Hexanucleotide Repeats (G4C2)$_n$

The abnormal expansion of a hexanucleotide, (GGGGCC)n (n from 30 up to thousands), in the first non-coding region of the *C9orf72* gene, is the predominant genetic factor in ALS and FTD [53]. These fatal neurodegenerative diseases, characterized by progressive neuronal dysfunction causing ultimate death, currently have no effective treatments available [54,55]. Accumulated molecular characterization indicates that *C9orf72* G4C2 repeat sequences, at both DNA and RNA levels, show remarkable structural polymorphism [56]. These sequences can adopt diverse secondary structures such as hairpin and G4 [57,58]. For example, the oligonucleotide d[(G4C2)$_3$G4], serving as a minimal model for d(G4C2) repeats capable of unimolecular G4 formation, adopts two predominant G4 conformations alongside multiple minor species, all of which coexist in potassium solutions [59,60]. Most importantly, the G4 formation is notably associated with the toxicity of *C9orf72* G4C2 repeats in ALS/FTD, which is a gain-of-function mechanism through RNA foci formation sequestering RNA binding proteins (RBPs) [61]. Notably, RNA foci formed by *C9orf72* G4C2 repeats, characteristic pathological features of these disorders, are

primarily composed of RNA G4s [17,62]. These structures appear to serve as nucleation sites for aberrant protein sequestration and aggregation, potentially initiating cascading cellular dysfunction [63]. Furthermore, *C9orf72* G4C2 repeat sequences form RNA condensates through multimolecular G4 structures. These RNA condensates play a critical role in disease pathology, potentially mediating neurotoxicity via liquid–liquid phase separation (LLPS) [36]. Moreover, *C9orf72* G4C2 creates an unstable folate-sensitive fragile site, FRA9A, in the genome. This fragile site is associated with abnormalities in DNA replication and repair, potentially leading to genomic instability and contributing to the pathogenesis of neurodegenerative diseases such as ALS/FTD [64].

Detailed structural analyses employing NMR and circular dichroism (CD) spectroscopy have revealed remarkable conformational diversity among *C9orf72* G4C2 sequences (Figure 2). *C9orf72* G4C2 DNAs demonstrate length-dependent topological preferences: d(G4C2)G4 adopts exclusively parallel conformations, while d(G4C2)$_2$, d(G4C2)$_3$, and d(G4C2)$_5$ exhibit mixed topological arrangements [34] (Figure 2A,B). Notably, d(G4C2)$_4$ uniquely assumes an antiparallel configuration, highlighting the sequence-specific nature of G4 folding patterns [34] (Figure 2A,B). In contrast to their DNA counterparts, *C9orf72* G4C2 RNAs adopt parallel G4s with different lengths of G4C2 repeats [56,57,65] (Figure 2C,D). This structural consistency in RNA G4s suggests distinct folding principles governing RNA versus DNA conformational preferences, potentially reflecting their divergent biological roles. The elucidation of high-resolution structure for these *C9orf72* G4C2 G4s is significant for multiple aspects of ALS/FTD pathobiology. These structural insights are crucial for understanding the molecular mechanisms underlying RNA-binding protein sequestration in ALS/FTD. Moreover, structural characterization provides essential foundations for de novo drug design, particularly enabling the development of structure-specific ligands or agents targeting distinct G4s.
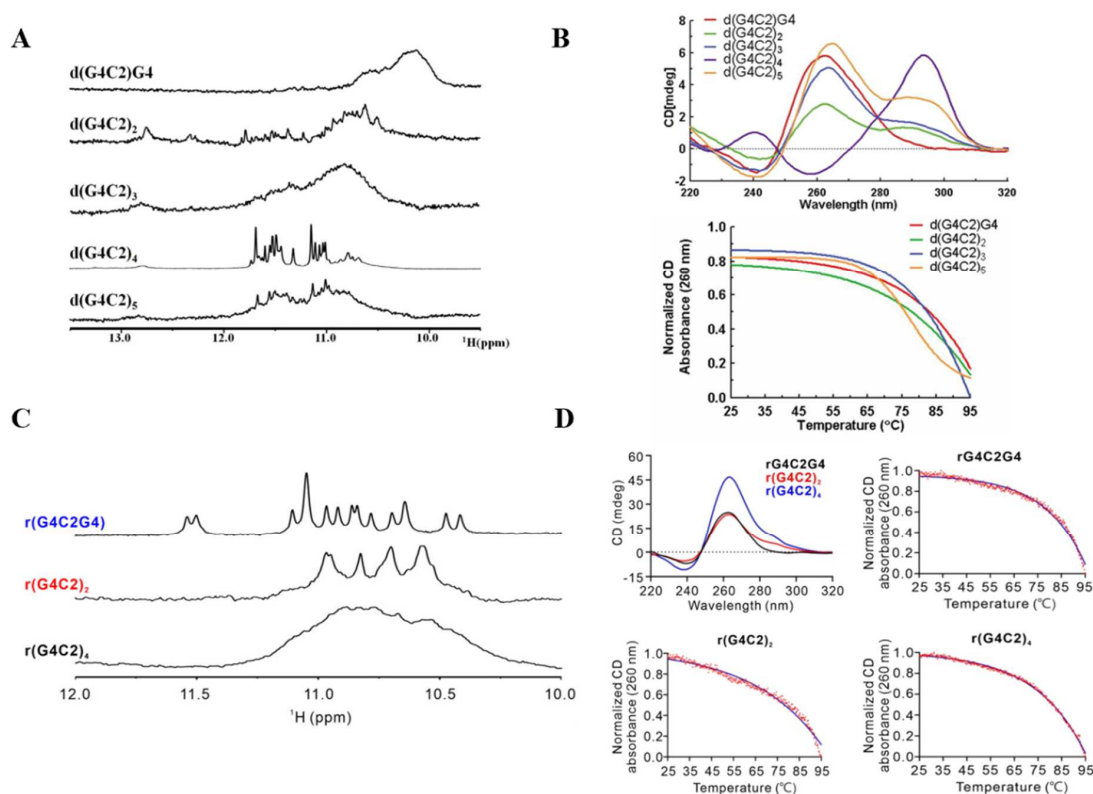

Figure 2. The characterization of different (G4C2)$_n$ sequences in the *C9orf72* gene. (**A**) 1D ^{1}H-NMR spectra of (G4C2)$_n$ DNA sequences. (**B**) CD spectra of (G4C2)$_n$ DNA. (**C**) 1D ^{1}H-NMR spectra of (G4C2)$_n$ RNA. (**D**) CD spectra of (G4C2)$_n$ RNA.

In 2021, we reported the first crystal structure of d(G4C2)$_2$ in K$^+$ and Ba^{2+} solution, which adopts a parallel-stranded tetrameric eight-layer G4 composed of two dimeric G4 units (Figure 3A) [66]. The crystal structure of d(G4C2)$_2$-Ba is obtained in the C2221 and F222 space groups, which adopt different arrangements of chains. The d(G4C2)$_2$-Ba in the C2221 reveals that four chains assemble into two dimeric G4 units, which further stack in a 5′-to-5′ orientation to form a tetrameric G4 through π–π interactions. The d(G4C2)$_2$-Ba in the F222 reveals asymmetric units consisting of three chains, which further organize into parallel-stranded tetrameric eight-layer G4s. Based on the structure of d(G4C2)$_2$-Ba (F222), the d(G4C2)$_2$-K structure in space group F222 was solved, which demonstrates isomorphous packing with similar structural features, differentiated primarily by their coordinating ions. In particular, our NMR studies demonstrated that the π–π stacking via 5′-to-5′ arrangement in crystal structure exists in solution [67]. Recently, we solved the first crystal structure of r(G4C2)$_2$, which adopts a parallel eight-layer G4 (Figure 3B). The crystal structure indicated that two chains form parallel dimeric units that undergo 5′-to-5′ coaxial stacking with another symmetric dimeric unit to form an eight-layer parallel tetrameric G4 structure in K$^+$ solution. This stacking is facilitated by π–π interactions (Figure 3B) [35].
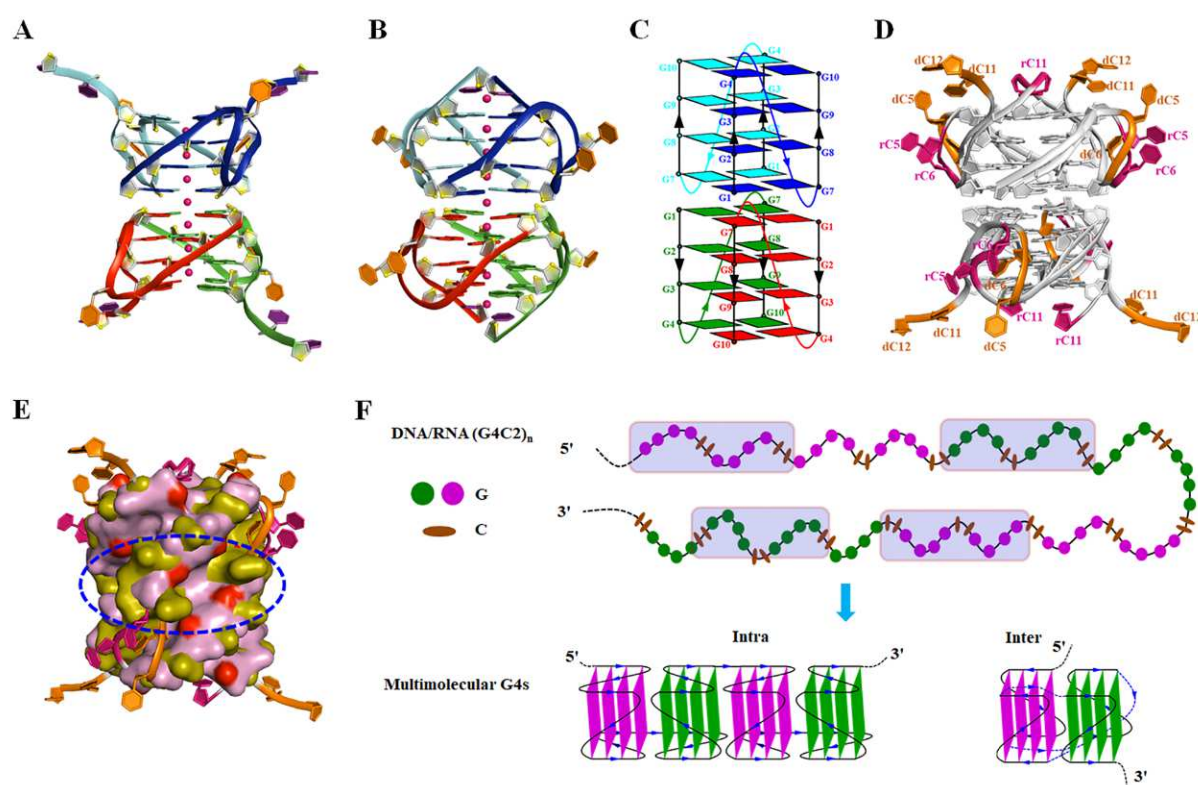

Figure 3. G4s formed by *C9orf72* HRE G4C2 DNA/RNA. Cartoon representation of tetrameric G4s formed by (**A**) d(G4C2)$_2$ and (**B**) r(G4C2)$_2$. (**C**) Topology of d(G4C2)$_2$ and r(G4C2)$_2$. Overlay of r(G4C2)$_2$ (PDB:8X0S) and d(G4C2)$_2$ (PDB:7ECH) G4 in (**D**) cartoon and (**E**) surface mode (the core of DNA in olive and RNA in pink), in which the 2′OH group of r(G4C2)$_2$ is colored red and the CC bases are colored hot pink in RNA and orange in DNA, respectively. (**F**) Schematic diagram showing the formation of multimeric G4s by the DNA/RNA (G4C2)n, in which the sequential (G4C2)$_4$ can form intramolecular G4s and the (G4C2)$_2$ in rectangles colored pink can form intermolecular G4s.

Although the G4 structures of d(G4C2)$_2$ and r(G4C2)$_2$ adopt the same topology as depicted in Figure 3C, the G-core of these two G4s is nearly identical, with a root-mean-square deviation (RMSD) of ~2.42 Å. However, the C5 and C6 bases in these two G4s adopt different conformations and are located at distinct positions (Figure 3D). In the case of d(G4C2)$_2$, the C6 base is situated within the medium groove of the G4 structure. Conversely,

for r(G4C2)$_2$, the C5 and C6 bases protrude outward, away from the G-core. Strikingly, the presence of the 2′OH groups and the 5′-to-5′ stacking mode in RNA leads to a shallower groove compared to DNA, which may represent unique and significant structural features for RBPs and ligands to recognize (Figure 3E).

Based on the G4 structures, we determined that both *C9orf72* DNA and RNA G4C2 sequences in vitro have the ability to form intramolecular four-layer parallel G4s [35,66]. These intramolecular structures sequentially stack together through 5′-to-5′ stacking, resulting in the formation of compact, higher-order G4 structures. Additionally, four neighboring G4C2 repeats along the same chain in different positions can associate to form intermolecular eight-layer parallel G4 structures. In these intermolecular structures, two four-layer parallel G4 units are connected by extended G4C2 repeats (Figure 3F).

Altogether, our structural findings are crucial for understanding the pathological mechanisms associated with the abnormal expansion of *C9orf72* DNA/RNA G4C2 repeats. Furthermore, the elucidated structures of *C9orf72* G4s demonstrate particular significance in multiple therapeutic aspects. First, they enable structure-based drug design approaches, facilitating the identification of small-molecule ligands with optimized binding properties for specific G4 structures. Second, these structures guide the development of therapeutic agents targeting distinct structures with specificity, potentially allowing targeted intervention at distinct stages of disease progression. Recent advances in structure-guided drug discovery have already yielded promising lead compounds demonstrating selective recognition of *C9orf72* G4 structures, which is significant for developing targeted therapies for ALS/FTD [27].

4. Conclusions and Future Perspectives

In recent years, studies on the biological role of G4s have validated the potential of designing specific ligands to target G4s in the treatment of disease. The structural diversity in telomeric G4s and *C9orf72* G4s particularly exemplifies the complexity and specificity requirements for targeted therapeutic development. The principal challenge in G4-targeted therapeutics is the selective recognition of specific structural motifs. This requirement of specificity highlights the importance of high-resolution structural characterization for each identified G4. Detailed structural information enables rational approaches to design novel ligands with optimized binding properties and reduced off-target effects. The expanding recognition of G4 involvement in human pathologies, from neurodegenerative disorders to cancer, positions structural studies at the forefront of therapeutic innovation. Future investigations focusing on G4 structures and interaction with cellular factors promise to advance our understanding of disease mechanisms and therapeutic strategies. This evolving field presents exciting opportunities for structure-guided drug discovery and the development of novel therapeutic modalities targeting specific G4 structures in various diseases.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Largy, E.; Mergny, J.L.; Gabelica, V. Role of Alkali Metal Ions in G-Quadruplex Nucleic Acid Structure and Stability. *Met. Ions Life Sci.* **2016**, *16*, 203–258. [PubMed]
2. Neidle, S.; Balasubramanian, S. *Quadruplex Nucleic Acids*; Royal Society of Chemistry: London, UK, 2006; Volume 7.
3. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.* **2020**, *21*, 459–474. [CrossRef] [PubMed]
4. Lightfoot, H.L.; Hagen, T.; Tatum, N.J.; Hall, J. The diverse structural landscape of quadruplexes. *FEBS Lett.* **2019**, *593*, 2083–2102. [CrossRef]
5. Phan, A.T.; Kuryavyi, V.; Darnell, J.C.; Serganov, A.; Majumdar, A.; Ilin, S.; Raslin, T.; Polonskaia, A.; Chen, C.; Clain, D.; et al. Structure-function studies of FMRP RGG peptide recognition of an RNA duplex-quadruplex junction. *Nat. Struct. Mol. Biol.* **2011**, *18*, 796–804. [CrossRef]
6. Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. Loop-length-dependent folding of G-quadruplexes. *J. Am. Chem. Soc.* **2004**, *126*, 16405–16415. [CrossRef]
7. Heddi, B.; Phan, A.T. Structure of human telomeric DNA in crowded solution. *J. Am. Chem. Soc.* **2011**, *133*, 9824–9833. [CrossRef]
8. Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R.A.; Yang, D. Human telomeric sequence forms a hybrid-type intramolecular G-quadruplex structure with mixed parallel/antiparallel strands in potassium solution. *Nucleic Acids Res.* **2006**, *34*, 2723–2735. [CrossRef]
9. Murat, P.; Balasubramanian, S. Existence and consequences of G-quadruplex structures in DNA. *Curr. Opin. Genet. Dev.* **2014**, *25*, 22–29. [CrossRef]
10. Pavlova, A.V.; Kubareva, E.A.; Monakhova, M.V.; Zvereva, M.I.; Dolinnaya, N.G. Impact of G-Quadruplexes on the Regulation of Genome Integrity, DNA Damage and Repair. *Biomolecules* **2021**, *11*, 1284. [CrossRef]
11. Chambers, V.S.; Marsico, G.; Boutell, J.M.; Di Antonio, M.; Smith, G.P.; Balasubramanian, S. High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.* **2015**, *33*, 877–881. [CrossRef]
12. Di Antonio, M.; Ponjavic, A.; Radzevicius, A.; Ranasinghe, R.T.; Catalano, M.; Zhang, X.Y.; Shen, J.Z.; Needham, L.M.; Lee, S.F.; Klenerman, D.; et al. Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.* **2020**, *12*, 832–837. [CrossRef] [PubMed]
13. Hansel-Hertsch, R.; Di Antonio, M.; Balasubramanian, S. DNA G-quadruplexes in the human genome: Detection, functions and therapeutic potential. *Nat. Rev. Mol. Cell. Biol.* **2017**, *18*, 279–284. [CrossRef] [PubMed]
14. Ma, Y.; Iida, K.; Nagasawa, K. Topologies of G-quadruplex: Biological functions and regulation by ligands. *Biochem. Biophys. Res. Commun.* **2020**, *531*, 3–17. [CrossRef] [PubMed]
15. Prioleau, M.N. G-Quadruplexes and DNA Replication Origins. *Adv. Exp. Med. Biol.* **2017**, *1042*, 273–286. [PubMed]
16. Rhodes, D.; Lipps, H.J. G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.* **2015**, *43*, 8627–8637. [CrossRef]
17. Simone, R.; Fratta, P.; Neidle, S.; Parkinson, G.N.; Isaacs, A.M. G-quadruplexes: Emerging roles in neurodegenerative diseases and the non-coding transcriptome. *FEBS Lett.* **2015**, *589*, 1653–1668. [CrossRef]
18. Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-quadruplexes: A promising target for cancer therapy. *Mol. Cancer* **2021**, *20*, 40. [CrossRef]
19. Cammas, A.; Millevoi, S. RNA G-quadruplexes: Emerging mechanisms in disease. *Nucleic Acids Res.* **2017**, *45*, 1584–1595. [CrossRef]
20. Wu, Y.; Brosh, R.M., Jr. G-quadruplex nucleic acids and human disease. *FEBS J.* **2010**, *277*, 3470–3488. [CrossRef]
21. Balasubramanian, S.; Hurley, L.H.; Neidle, S. Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug Discov.* **2011**, *10*, 261–275. [CrossRef]
22. Yang, D.; Okamoto, K. Structural insights into G-quadruplexes: Towards new anticancer drugs. *Future Med. Chem.* **2010**, *2*, 619–646. [CrossRef] [PubMed]
23. Metifiot, M.; Amrane, S.; Litvak, S.; Andreola, M.L. G-quadruplexes in viruses: Function and potential therapeutic applications. *Nucleic Acids Res.* **2014**, *42*, 12352–12366. [CrossRef] [PubMed]
24. Cui, H.; Zhang, L. G-Quadruplexes Are Present in Human Coronaviruses Including SARS-CoV-2. *Front. Microbiol.* **2020**, *11*, 567317. [CrossRef]
25. Zhao, C.; Qin, G.; Niu, J.; Wang, Z.; Wang, C.; Ren, J.; Qu, X. Targeting RNA G-Quadruplex in SARS-CoV-2: A Promising Therapeutic Target for COVID-19? *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 432–438. [CrossRef]
26. Ruggiero, E.; Richter, S.N. Viral G-quadruplexes: New frontiers in virus pathogenesis and antiviral therapy. *Annu. Rep. Med. Chem* **2020**, *54*, 101–131.
27. Cheng, A.; Liu, C.; Ye, W.; Huang, D.; She, W.; Liu, X.; Fung, C.P.; Xu, N.; Suen, M.C.; Ye, W.; et al. Selective C9orf72 G-Quadruplex-Binding Small Molecules Ameliorate Pathological Signatures of ALS/FTD Models. *J. Med. Chem.* **2022**, *65*, 12825–12837. [CrossRef]
28. Neidle, S. Quadruplex nucleic acids as targets for anticancer therapeutics. *Nat. Rev. Chem.* **2017**, *1*, 41. [CrossRef]

29. Wang, E.; Thombre, R.; Shah, Y.; Latanich, R.; Wang, J.O. G-Quadruplexes as pathogenic drivers in neurodegenerative disorders. *Nucleic Acids Res.* **2021**, *49*, 4816–4830. [CrossRef]

30. Zamiri, B.; Reddy, K.; Macgregor, R.B., Jr.; Pearson, C.E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNA-binding proteins. *J. Biol. Chem.* **2014**, *289*, 4653–4659. [CrossRef]

31. Platella, C.; Napolitano, E.; Riccardi, C.; Musumeci, D.; Montesarchio, D. Disentangling the Structure-Activity Relationships of Naphthalene Diimides as Anticancer G-Quadruplex-Targeting Drugs. *J. Med. Chem.* **2021**, *64*, 3578–3603. [CrossRef]

32. Figueiredo, J.; Mergny, J.L.; Cruz, C. G-quadruplex ligands in cancer therapy: Progress, challenges, and clinical perspectives. *Life Sci.* **2024**, *340*, 122481. [CrossRef] [PubMed]

33. Xu, Y.; Komiyama, M. G-Quadruplexes in Human Telomere: Structures, Properties, and Applications. *Molecules* **2023**, *29*, 174. [CrossRef] [PubMed]

34. Zhou, B.; Liu, C.D.; Geng, Y.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] [PubMed]

35. Geng, Y.; Liu, C.; Xu, N.; Suen, M.C.; Miao, H.; Xie, Y.; Zhang, B.; Chen, X.; Song, Y.; Wang, Z.; et al. Crystal structure of a tetrameric RNA G-quadruplex formed by hexanucleotide repeat expansions of C9orf72 in ALS/FTD. *Nucleic Acids Res.* **2024**, *52*, 7961–7970. [CrossRef]

36. Raguseo, F.; Wang, Y.; Li, J.; Petric Howe, M.; Balendra, R.; Huyghebaert, A.; Vadukul, D.M.; Tanase, D.A.; Maher, T.E.; Malouf, L.; et al. The ALS/FTD-related C9orf72 hexanucleotide repeat expansion forms RNA condensates through multimolecular G-quadruplexes. *Nat. Commun.* **2023**, *14*, 8272. [CrossRef]

37. Bailey, S.M.; Murnane, J.P. Telomeres, chromosome instability and cancer. *Nucleic Acids Res.* **2006**, *34*, 2408–2417. [CrossRef]

38. Bryan, T.M. G-Quadruplexes at Telomeres: Friend or Foe? *Molecules* **2020**, *25*, 3686. [CrossRef]

39. Jafri, M.A.; Ansari, S.A.; Alqahtani, M.H.; Shay, J.W. Roles of telomeres and telomerase in cancer, and advances in telomerase-targeted therapies. *Genome Med.* **2016**, *8*, 69. [CrossRef]

40. Boccardi, V.; Marano, L. Aging, Cancer, and Inflammation: The Telomerase Connection. *Int. J. Mol. Sci.* **2024**, *25*, 8542. [CrossRef]

41. Dai, J.X.; Carver, M.; Punchihewa, C.; Jones, R.A.; Yang, D.Z. Structure of the Hybrid-2 type intramolecular human telomeric G-quadruplex in K^+ solution: Insights into structure polymorphism of the human telomeric sequence. *Nucleic Acids Res.* **2007**, *35*, 4927–4940. [CrossRef]

42. Dai, J.X.; Punchihewa, C.; Ambrus, A.; Chen, D.; Jones, R.A.; Yang, D.Z. Structure of the intramolecular human telomeric G-quadruplex in potassium solution: A novel adenine triple formation. *Nucleic Acids Res.* **2007**, *35*, 2440–2450. [CrossRef]

43. Lim, K.W.; Amrane, S.; Bouaziz, S.; Xu, W.X.; Mu, Y.G.; Patel, D.J.; Luu, K.N.; Phan, A.T. Structure of the Human Telomere in K+ Solution: A Stable Basket-Type G-Quadruplex with Only Two G-Tetrad Layers. *J. Am. Chem. Soc.* **2009**, *131*, 4301–4309. [CrossRef]

44. Lim, K.W.; Ng, V.C.M.; Martin-Pintado, N.; Heddi, B.; Phan, A.T. Structure of the human telomere in Na+ solution: An antiparallel (2+2) G-quadruplex scaffold reveals additional diversity. *Nucleic Acids Res.* **2013**, *41*, 10556–10562. [CrossRef]

45. Liu, C.; Geng, Y.; Miao, H.; Shi, X.; You, Y.; Xu, N.; Zhou, B.; Zhu, G. G-quadruplex structures formed by human telomeric DNA and C9orf72 hexanucleotide repeats. *Biophys. Rev.* **2019**, *11*, 389–393. [CrossRef]

46. Parkinson, G.N.; Lee, M.P.H.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876–880. [CrossRef]

47. Phan, A.T.; Kuryavyi, V.; Luu, K.N.; Patel, D.J. Structure of two intramolecular G-quadruplexes formed by natural human telomere sequences in K^+ solution. *Nucleic Acids Res.* **2007**, *35*, 6517–6525. [CrossRef]

48. Wang, Y.; Patel, D.J. Solution structure of the human telomeric repeat d[AG3(T2AG3)3] G-tetraplex. *Structure* **1993**, *1*, 263–282. [CrossRef]

49. Zhang, Z.J.; Dai, J.X.; Veliath, E.; Jones, R.A.; Yang, D.Z. Structure of a two-G-tetrad intramolecular G-quadruplex formed by a variant human telomeric sequence in K^+ solution: Insights into the interconversion of human telomeric G-quadruplex structures. *Nucleic Acids Res.* **2010**, *38*, 1009–1021. [CrossRef]

50. Geng, Y.; Liu, C.; Zhou, B.; Cai, Q.; Miao, H.; Shi, X.; Xu, N.; You, Y.; Fung, C.P.; Din, R.U.; et al. The crystal structure of an antiparallel chair-type G-quadruplex formed by Bromo-substituted human telomeric DNA. *Nucleic Acids Res.* **2019**, *47*, 5395–5404. [CrossRef]

51. Liu, C.; Zhou, B.; Geng, Y.; Yan Tam, D.; Feng, R.; Miao, H.; Xu, N.; Shi, X.; You, Y.; Hong, Y.; et al. A chair-type G-quadruplex structure formed by a human telomeric variant DNA in $K^{(+)}$ solution. *Chem. Sci.* **2019**, *10*, 218–226. [CrossRef]

52. Geng, Y.Y.; Liu, C.D.; Xu, N.N.; Shi, X.; Suen, M.C.; Zhou, B.; Yan, B.; Wu, C.M.; Li, H.; Song, Y.J.; et al. The N-terminal region of Cdc6 specifically recognizes human DNA G-quadruplex. *Int. J. Biol. Macromol.* **2024**, *260*, 129487. [CrossRef]

53. 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]

54. Rademakers, R.; Neumann, M.; Mackenzie, I.R. Advances in understanding the molecular basis of frontotemporal dementia. *Nat. Rev. Neurol.* **2012**, *8*, 423–434. [CrossRef]

55. Rowland, L.P.; Shneider, N.A. Amyotrophic lateral sclerosis. *N. Engl. J. Med.* **2001**, *344*, 1688–1700. [CrossRef]

56. 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]

57. Reddy, K.; Zamiri, B.; Stanley, S.Y.R.; Macgregor, R.B.; Pearson, C.E. The Disease-associated r(GGGGCC)(n) Repeat from the C9orf72 Gene Forms Tract Length-dependent Uni- and Multimolecular RNA G-quadruplex Structures. *J. Biol. Chem.* **2013**, *288*, 9860–9866. [CrossRef]

58. Wang, Z.F.; Ursu, A.; Childs-Disney, J.L.; Guertler, R.; Yang, W.Y.; Bernat, V.; Rzuczek, S.G.; Fuerst, R.; Zhang, Y.J.; Gendron, T.F.; et al. The Hairpin Form of r(G(4)C(2))(exp) in c9ALS/FTD Is Repeat-Associated Non-ATG Translated and a Target for Bioactive Small Molecules. *Cell Chem. Biol.* **2019**, *26*, 179–190. [CrossRef]

59. Brcic, J.; Plavec, J. NMR structure of a G-quadruplex formed by four d(G4C2) repeats: Insights into structural polymorphism. *Nucleic Acids Res.* **2018**, *46*, 11605–11617.

60. Brcic, 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]

61. Schmitz, A.; Pinheiro Marques, J.; Oertig, I.; Maharjan, N.; Saxena, S. Emerging Perspectives on Dipeptide Repeat Proteins in C9ORF72 ALS/FTD. *Front. Cell. Neurosci.* **2021**, *15*, 637548. [CrossRef]

62. 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]

63. Swinnen, B.; Robberecht, W.; Van Den Bosch, L. RNA toxicity in non-coding repeat expansion disorders. *EMBO J.* **2020**, *39*, e101112. [CrossRef]

64. Mirceta, M.; Schmidt, M.H.M.; Shum, N.; Prasolava, T.K.; Meikle, B.; Lanni, S.; Mohiuddin, M.; McKeever, P.M.; Zhang, M.; Liang, M.; et al. C9orf72 repeat expansion creates the unstable folate-sensitive fragile site FRA9A. *NAR Mol. Med.* **2024**, *1*, ugae019. [CrossRef]

65. 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*, 1016. [CrossRef]

66. 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]

67. Liu, C.; Zhou, B.; Xu, N.; Fung, C.P.; Yan, B.; Suen, M.C.; Huang, Z.; Zhu, G. The parallel tetrameric DNA G-quadruplex formed by the two-repeat C9orf72 GGGGCC sequence in solution. *Magn. Reson. Lett.* **2022**, *2*, 196–204. [CrossRef]