




Article

Structural Diversity of Sense and Antisense RNA Hexanucleotide Repeats Associated with ALS and FTLD

Tim Božič¹, Matja Zalar^{1,†} , Boris Rogelj^{2,3}, Janez Plavec^{1,4,5,*}  and Primož Šket^{1,*} 

¹ Slovenian NMR Centre, National Institute of Chemistry, Hajdrihova 19, SI-1000 Ljubljana, Slovenia; timbozicbiochemistry@gmail.com (T.B.); m.zalar@beatson.gla.ac.uk (M.Z.)

² Department of Biotechnology, Jožef Stefan Institute, Jamova 39, SI-1000 Ljubljana, Slovenia; boris.rogelj@ijs.si

³ Biomedical Research Institute BRIS, Puhova 10, SI-1000 Ljubljana, Slovenia

⁴ EN-FIST Center of Excellence, Trg OF 13, SI-1000 Ljubljana, Slovenia

⁵ Faculty of Chemistry and Chemical Technology, University of Ljubljana, Večna pot 113, SI-1000 Ljubljana, Slovenia

* Correspondence: primoz.sket@ki.si (P.Š.); janez.plavec@ki.si (J.P.); Tel.: +386-1-4760223 (P.Š.); +386-1-4760353 (J.P.)

† Current address: CRUK Beatson Institute, Switchback Rd, Bearsden, Glasgow G61 1BD, UK.

Academic Editor: Ta-Chau Chang

Received: 12 December 2019; Accepted: 24 January 2020; Published: 25 January 2020



Abstract: The hexanucleotide expansion GGGGCC located in *C9orf72* gene represents the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar dementia (FTLD). Since the discovery one of the non-exclusive mechanisms of expanded hexanucleotide G₄C₂ repeats involved in ALS and FTLD is RNA toxicity, which involves accumulation of pathological sense and antisense RNA transcripts. Formed RNA foci sequester RNA-binding proteins, causing their mislocalization and, thus, diminishing their biological function. Therefore, structures adopted by pathological RNA transcripts could have a key role in pathogenesis of ALS and FTLD. Utilizing NMR spectroscopy and complementary methods, we examined structures adopted by both guanine-rich sense and cytosine-rich antisense RNA oligonucleotides with four hexanucleotide repeats. While both oligonucleotides tend to form dimers and hairpins, the equilibrium of these structures differs with antisense oligonucleotide being more sensitive to changes in pH and sense oligonucleotide to temperature. In the presence of K⁺ ions, guanine-rich sense RNA oligonucleotide also adopts secondary structures called G-quadruplexes. Here, we also observed, for the first time, that antisense RNA oligonucleotide forms i-motifs under specific conditions. Moreover, simultaneous presence of sense and antisense RNA oligonucleotides promotes formation of heterodimer. Studied structural diversity of sense and antisense RNA transcripts not only further depicts the complex nature of neurodegenerative diseases but also reveals potential targets for drug design in treatment of ALS and FTLD.

Keywords: NMR; RNA; ALS; FTLD; *C9orf72*

Academic Editor: Ta-Chau Chang

1. Introduction

The pathological hexanucleotide expansion GGGGCC located in the first intron or promoter region of the *C9orf72* gene on chromosome 9p21 represents the most common genetic cause of amyotrophic lateral sclerosis (ALS) and frontotemporal lobar dementia (FTLD)—two fatal neurodegenerative diseases with progressive loss of motor neurons in the brain and spinal cord [1–7]. Healthy individuals possess up to 19 repeats, while patients with ALS or FTLD may carry up to 5000 repeats [3–6]. To date, three main non-exclusive mechanisms of expanded hexanucleotide G₄C₂ repeats involved in ALS and FTLD have been proposed: protein *C9orf72* haploinsufficiency, accumulation of dipeptide repeat proteins

(DPRs) and RNA toxicity [7]. While haploinsufficiency and DPRs are both protein driven mechanisms, RNA toxicity involves accumulating sense and antisense RNA transcripts. Accumulation leads to formation of RNA foci, which sequesters RNA-binding proteins [8]. Bound RNA-binding proteins are responsible for nuclear transport, splicing, and translation, all of which are affected by dysregulated RNA metabolism [9,10]. Myotonic dystrophy and fragile X-associated tremor/ataxia syndrome are also repeat expansion disorders, where RNA toxicity plays a crucial role in pathogenesis [11,12]. This indicates that structures and structural motifs adopted by different repeat expansions direct the final outcome as they dictate the binding affinity and specificity of protein domains. Latter hypothesis led to several structural studies of DNA and RNA repeat expansion variations of *C9orf72* gene [1,2,10,13–18]. While G-rich sense RNA repeat expansion variations seem to form mainly parallel G-quadruplexes, C-rich antisense RNA repeat expansion is presumably limited to hairpins and A-form-like dimer with regularly spaced tandem C:C mismatches [16–18]. Although i-motifs have been proposed for antisense DNA repeat expansion variations [1], they have not been reported yet for antisense RNA repeat expansion, to the best of our knowledge.

By utilizing NMR spectroscopy and complementary methods we demonstrated that sense $r(G_4C_2)_4$ and antisense $r(G_2C_4)_4$ RNA oligonucleotides composed of four hexanucleotide repeats can form dimers and hairpins under conditions approaching physiological relevance. While structures of antisense RNA oligonucleotide show pH dependency, structures of sense RNA oligonucleotide are temperature sensitive. Furthermore, sense RNA oligonucleotide forms structurally different G-quadruplexes in the presence of KCl and antisense RNA oligonucleotide indicates i-motif formation at low pH and temperature. Simultaneous presence of both RNA oligonucleotides revealed formation of heterodimer although structures adopted by individual RNA oligonucleotides seem to persist but are a minority.

2. Results

2.1. Homodimer–Hairpin Equilibrium of $r(G_2C_4)_4$ is pH-Dependent

Analysis of NMR spectra of $r(G_2C_4)_4$ allowed us to propose secondary structure of homodimer consisting of G-C base pairs and C-C mismatches adopted at pH 7.0 (Figure 1a). Dimeric nature of structure was identified based on translational diffusion coefficient (D_t) value of $0.91 \pm 0.05 \times 10^{-10} \text{ m}^2\text{s}^{-1}$. 1D ^1H -NMR spectrum reveals six well resolved signals between δ 12.6 and 13.6 ppm corresponding to guanine imino protons involved in Watson-Crick base pairs (Figure 1b and Figure S1a). Relative integral ratio of signals suggests that homodimer consists of 16 G-C base pairs. Predominant formation of a dimeric symmetric structure is in agreement with observation of 16 cross-peaks corresponding to cytosine H5/H6 connectivities in 2D TOCSY NMR spectrum of the studied oligonucleotide with 16 cytosine residues (Figure S2). Examination of imino regions of ^1H -NMR and CD spectra, as well as UV melting experiment on $r(G_2C_4)_4$ at pH 7.0 showed that homodimer is thermally stable in temperature range from 5 to 70 °C (Figure 1c, Figures S3 and S4). In oligonucleotide concentration range from 2.6 to 9.3 μM per strand, temperatures of half transition obtained from melting curves during unfolding were between 73 and 76 °C. At 20 μM oligonucleotide concentration, homodimer was not completely unfolded even at 90 °C.

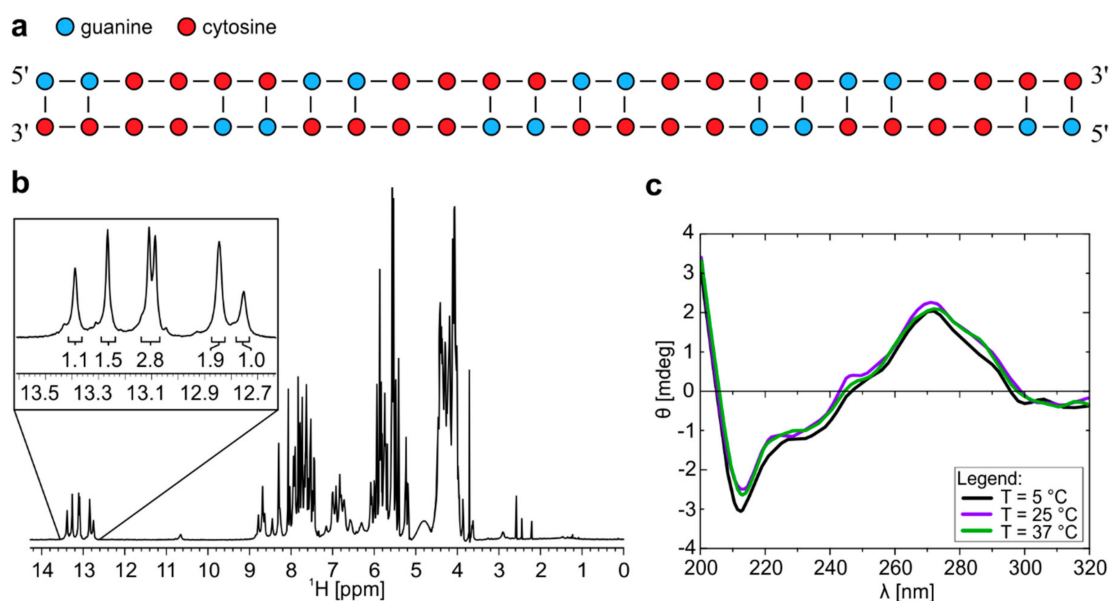


Figure 1. $^1\text{H-NMR}$ and CD spectra of $r(\text{G}_2\text{C}_4)_4$ at pH 7.0 forms dimeric structure. (a) Schematic presentation of homodimer with G-C base pairs and C-C mismatches. (b) $^1\text{H-NMR}$ spectrum of $r(\text{G}_2\text{C}_4)_4$ in 10% $^2\text{H}_2\text{O}$ at pH 7.0, 25 °C and 0.9 mM oligonucleotide concentration per strand. Numbers below the signals represent integral values. c) CD spectra of homodimer with concentration of 100 μM per strand at 5, 25 and 37 °C.

CD spectra of $r(\text{G}_2\text{C}_4)_4$ at pH 7.0 reveal minimum at wavelength of 212 nm, maxima at 245 and 265 nm as well as shoulder at 285 nm (Figure 1c), which is in good agreement with CD spectra of A-form dsRNA and $r((\text{C}_4\text{G}_2)_n\text{CCCC})$ repeats [18,19]. A shoulder in CD spectra at 285 nm, however, suggests the presence of hairpin structure in lesser extent, which is not observable by NMR [20].

$^1\text{H-NMR}$ spectra of $r(\text{G}_2\text{C}_4)_4$ acquired at 25 °C as a function of pH ranging between 7.0 and 4.5 show appearance of signals, which indicate shift in equilibrium towards formation of a new structure upon lowering pH (Figure S5). At pH 4.5 signals between δ 12.5 and 14.0 ppm correspond to guanine imino protons involved in Watson-Crick G-C base pairs, while signal with chemical shift of δ 16.45 ppm corresponds to imino protons of hemi-protonated C^+-C base pairs (Figure S1b). At pH 4.5 the measured value of D_t of $1.40 \pm 0.05 \times 10^{-10} \text{ m}^2\text{s}^{-1}$ is in accordance with formation of a monomeric structure (Table S1). Using coordinates from 3D models of homodimer and hairpin of $r(\text{G}_2\text{C}_4)_4$ predicted by MC-Fold/MC-Sym enabled to calculate D_t with HydroPro program [21]. Comparison of experimental and calculated D_t values unambiguously confirmed formation of hairpin and homodimer at pH 4.5 and 7.0, respectively. Homodimer–hairpin equilibrium was also observed in the presence of 10% *w/v* PEG, however, imino signal corresponding to hemi-protonated C^+-C base pairs at δ 16.45 ppm was not observed even at pH 4.5 (Figure S6).

Structural changes caused by variation of pH were also indicated in CD spectra acquired at 25 °C where lowering of pH to 4.5 resulted in less distinct negative band at 212 nm, shift of positive band from 270 nm to 267 nm, while shoulder at 285 nm became more prominent (Figure 2a). Overall, signals in CD spectrum at pH 4.5 were attributed to a hairpin like structure. Both $^1\text{H-NMR}$ and CD spectra after increasing of pH back to 7.0 show that the process is reversible (Figure S7).

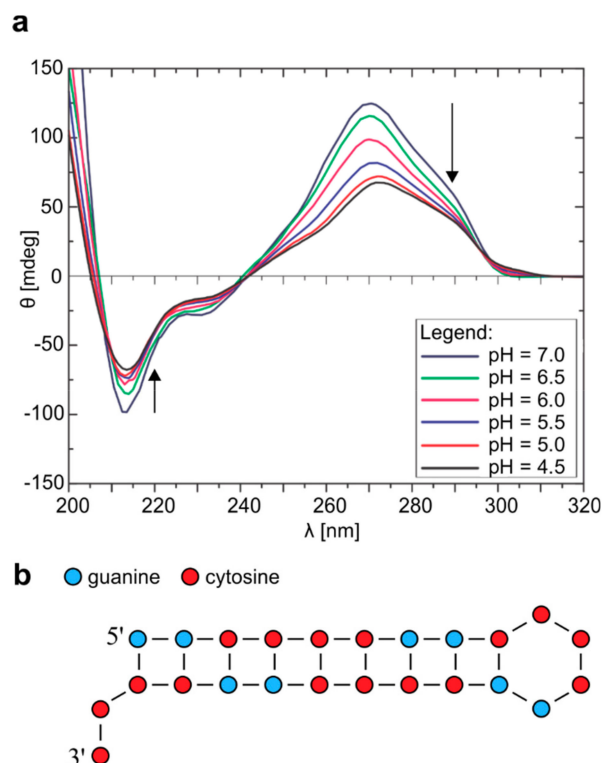


Figure 2. CD spectra of $r(G_2C_4)_4$ and proposed hairpin adopted at pH 4.5. (a) CD spectra of $r(G_2C_4)_4$ with pH ranging from 7.0 to 4.5. All CD spectra were acquired at 25 °C and concentration of 100 μ M per strand. Arrows represent the direction of change in pH. (b) Schematic presentation of hairpin formed by $r(G_2C_4)_4$ at pH 4.5 and 25 °C.

Homodimer–hairpin equilibrium was confirmed by 20% native PAGE at pH 6.0 and 5 °C, where two bands were observed (Figure S8). One band migrating as approximately 30 bp construct corresponds to homodimer, while the other band observed just below 15 bp corresponds to a hairpin structure.

NOE contact between signal at δ 16.45 ppm and signal corresponding to imino protons of guanines involved in Watson-Crick G-C base pairs at δ 13.18 ppm in 2D NOESY spectrum acquired at pH 4.5, 5 °C and mixing time of 100 ms is in agreement with formation of hairpin structure consisting of hemi-protonated C⁺-C base pairs (Figure S9). Imino-imino NOE contacts between three signals resonating from δ 13.0 to 13.3 ppm and signal at δ 13.87 ppm indicate that the latter signal corresponds to imino protons of guanine residues involved in G-C base pairs adjacent to G-C base pairs next to C⁺-C base pairs (Figure 2b).

¹H-NMR spectra at pH 4.5 as well as UV melting experiment on $r(G_2C_4)_4$ at pH 5.0 show that hairpin is thermally stable in temperature range between 5 and 70 °C (Figures S10 and S11). In oligonucleotide concentration range from 2.6 to 20 μ M the temperature of half transition of hairpin adopted by $r(G_2C_4)_4$ was estimated to be approximately 90 °C as completely unfolded state was not reached.

At a temperature of 5 °C and pH 4.5 imino region of ¹H-NMR spectrum revealed also a signal of low intensity at δ 15.75 ppm corresponding to additional hemi-protonated C⁺-C base pairs (Figure S10). According to the chemical shift value, these hemi-protonated C⁺-C base pairs are part of i-motif structure.

2.2. Temperature Changes Affect Population of Structures Formed by $r(G_4C_2)_4$

1D ¹H-NMR spectrum of $r(G_4C_2)_4$ recorded at pH 6.0 and 25 °C reveals four well resolved signals between δ 12 and 14 ppm corresponding to guanine imino protons of Watson-Crick base pairs (Figure 3a). 2D DOSY experiment resulted in two sets of aromatic and imino signals among which three imino signals between δ 13.0 and δ 13.4 ppm with D_t value of $0.94 \pm 0.05 \times 10^{-10} \text{ m}^2\text{s}^{-1}$

correspond to G-C base pairs of homodimer, while signal resonating at δ 12.41 ppm with D_t value of $1.40 \pm 0.05 \times 10^{-10} \text{ m}^2\text{s}^{-1}$ corresponds to G-C base pairs of hairpin. Noteworthy, experimentally observed D_t values are in agreement with values calculated for the proposed models of homodimer and hairpin using HydroPro program (*vide supra*).

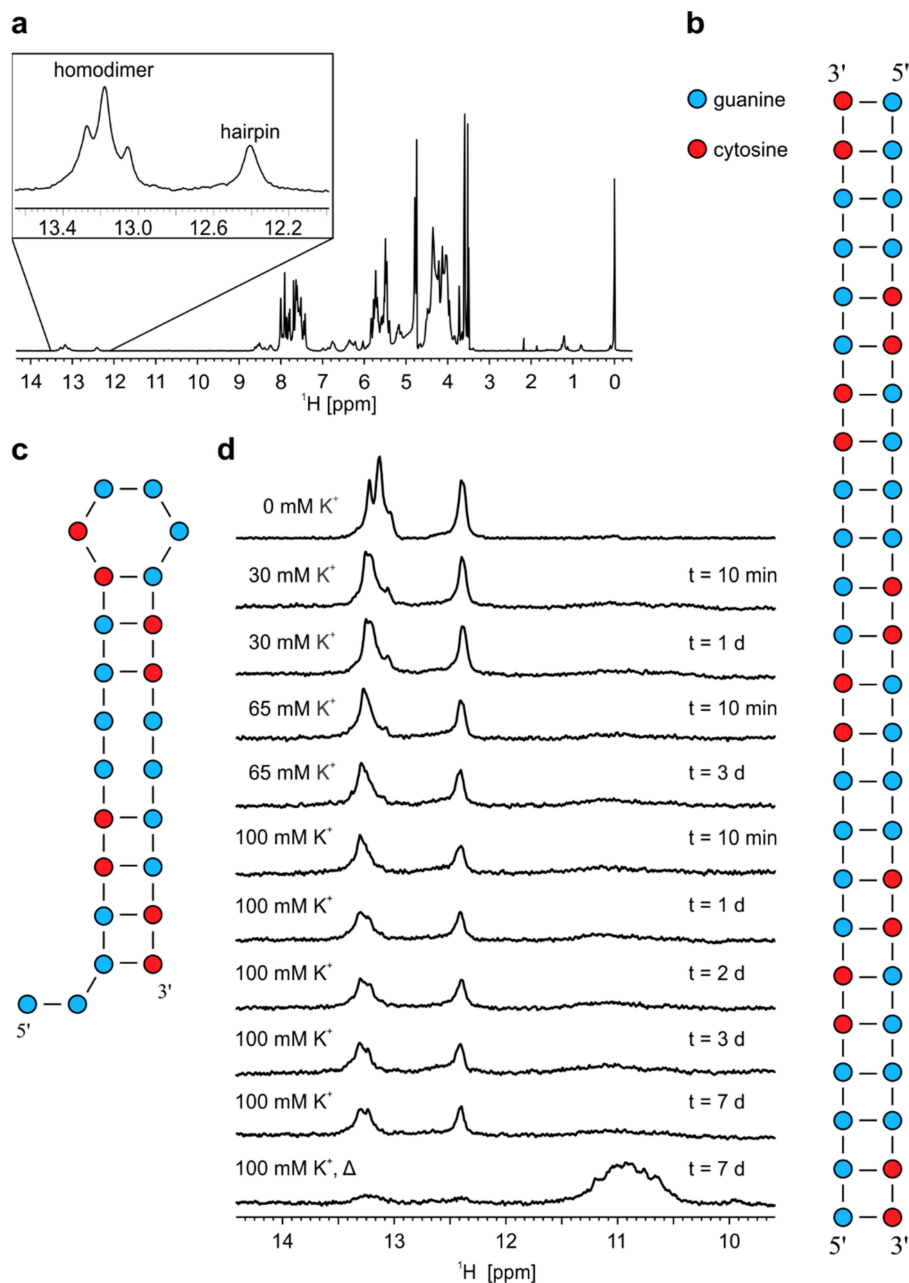


Figure 3. ^1H -NMR spectra and proposed models of structures adopted by $r(\text{G}_4\text{C}_2)_4$. (a) ^1H -NMR spectrum of $r(\text{G}_4\text{C}_2)_4$ in 10% $^2\text{H}_2\text{O}$ at pH 6.0, 25 °C and 0.9 mM oligonucleotide concentration per strand. Schematic presentation of (b) homodimer and (c) hairpin adopted by $r(\text{G}_4\text{C}_2)_4$. (d) Imino regions of ^1H -NMR spectra of $r(\text{G}_4\text{C}_2)_4$ in 10% $^2\text{H}_2\text{O}$ at different concentrations of K^+ ions. KCl solution was gradually titrated into the sample up to the stated concentration of K^+ ions, displayed on the left of the NMR spectra. Δ represents NMR spectrum after annealing while t on the right represents time after each addition of K^+ ions until acquisition of NMR spectrum. All spectra were acquired at 25 °C, pH 6.0 and 0.2 mM oligonucleotide concentration per strand.

Decrease of temperature to 0 °C results in disappearance of imino signal at δ 12.41 ppm corresponding to G-C base pairs of hairpin, while three signals between δ 13.0 and 13.4 ppm corresponding to G-C base pairs of homodimer are still present (Figure S12). Additionally, a new set of imino signals between δ 10.5 and 11.0 ppm starts to appear, which correspond to N1-carbonyl G-G base pairs stabilized at lower temperatures within homodimer (Figure S1c). Integration of imino signals in $^1\text{H-NMR}$ spectrum of $r(\text{G}_4\text{C}_2)_4$ at pH 6.0 and 0 °C shows that the ratio between G-C and G-G base pairs of homodimer is 2:1, which suggests that homodimer of $r(\text{G}_4\text{C}_2)_4$ consists of 16 G-C base pairs and eight G-G base pairs (Figure S13 and Figure 3b). On the other hand, most favourable secondary structure of hairpin adopted by $r(\text{G}_4\text{C}_2)_4$ with the highest amount of G-C base pairs is proposed in Figure 3c.

$^1\text{H-NMR}$ spectra at different pH values show that homodimer and hairpin of $r(\text{G}_4\text{C}_2)_4$ are stable within narrow pH range from 5.0 to 6.0 (Figure S14). At pH 7.0 only two imino signals belonging to homodimer persist, while third signal at δ 13.10 ppm is significantly reduced. Moreover, signal corresponding to hairpin at δ 12.41 ppm is absent. Lowering of pH to 4.0 results in almost complete disappearance of signals indicating guanine imino protons of homodimer and hairpin. Signals belonging to homodimer and hairpin reappear after increasing pH back to 6.0, showing reversibility.

Homodimer-hairpin equilibrium is preserved in concentration range between 0.05 and 1.1 mM of $r(\text{G}_4\text{C}_2)_4$ at pH 6.0 and 25 °C based on integral ratios of imino signals in $^1\text{H-NMR}$ spectra (Figure S15). Homodimer-hairpin equilibrium was confirmed by 20% native PAGE at pH 6.0, 5 °C and 0.3, 0.2 and 0.1 mM oligonucleotide concentrations per strand (Figure S16). Slower migrating band at 35 bp corresponds to homodimer, while constructs with band at 18 bp affirms hairpin formation.

2.3. Formation of $r(\text{G}_4\text{C}_2)_4$ G-Quadruplex is Favoured by Annealing in the Presence of K^+ Ions

$^1\text{H-NMR}$ spectra of $r(\text{G}_4\text{C}_2)_4$ acquired at 25 °C with increasing concentration of K^+ ions show that homodimer and hairpin persist even after exposure of sample to 100 mM concentration of K^+ ions for more than 7 days (Figure 3d). Only after annealing of the sample in the presence of 100 mM concentration of K^+ ions signals belonging to imino protons of G-C base pairs within homodimer and hairpin almost disappear. At the same time signals between δ 10.5 and 11.5 ppm corresponding to imino protons of guanines involved in G-quartets appear in $^1\text{H-NMR}$ spectrum. D_t values of aromatic proton signals at 25 °C were in range between 1.1 and $1.5 \times 10^{-10} \text{ m}^2\text{s}^{-1}$, which suggesting formation of G-quadruplexes of different sizes.

2.4. Antisense hybridizes with sense RNA oligonucleotide

Imino region of the $^1\text{H-NMR}$ spectrum of equimolar mixture of $r(\text{G}_2\text{C}_4)_4$ and $r(\text{G}_4\text{C}_2)_4$ recorded at pH 6.0 and 25 °C reveals five well resolved imino signals between δ 12.5 and 13.5 ppm, with slightly different chemical shifts as observed for imino signals of individual sense and antisense RNA corresponding to G-C base pairs of heterodimer (Figure 4a). In the same chemical shift range small overlapping imino signals of homodimers formed individually by $r(\text{G}_2\text{C}_4)_4$ and $r(\text{G}_4\text{C}_2)_4$ can still be observed, while the presence of small imino signal at δ 12.45 ppm corresponds to hairpin of $r(\text{G}_4\text{C}_2)_4$. Decrease in temperature to 7 °C leads to observation of additional signals between δ 10.5 and 11.0 ppm corresponding to N1-carbonyl G-G base pairs, which are stabilized within homodimer of $r(\text{G}_4\text{C}_2)_4$. Signal of low intensity at δ 16.45 ppm corresponds to hemi-protonated $\text{C}^+\text{-C}$ base pairs that are part of hairpin formed by $r(\text{G}_2\text{C}_4)_4$ (Figure 4b). Existence of a low population of hairpins in equimolar mixture was also confirmed with CD spectrum at pH 6.0 and 25 °C, where a shoulder at wavelength of 285 nm can be observed in addition to a negative signal at 212 and positive one at 265 nm (Figure S17) [20]. After annealing, signals corresponding to G-G base pairs and hemi-protonated $\text{C}^+\text{-C}$ base pairs disappear completely in $^1\text{H-NMR}$ spectra acquired at pH 6.0 and 7 °C as well as 25 °C (Figure 4b and Figure S18). In CD spectrum recorded at pH 6.0 and 25 °C after annealing signals at wavelengths of 212 and 265 nm persist, while shoulder at 285 nm is significantly reduced (Figure S17).

2D DOSY acquired on annealed sample at pH 6.0 and 25 °C afforded D_t of $1.05 \pm 0.05 \times 10^{-10} \text{ m}^2\text{s}^{-1}$, which is in accordance with hydrodynamic properties of dimeric structure and calculated D_t value by HydroPro program. Therefore, secondary structure of heterodimer formed by $r(\text{G}_2\text{C}_4)_4/r(\text{G}_4\text{C}_2)_4$ consists of 24 Watson-Crick G-C base pairs with indications that some G-C base pairs are more flexible than others, which helps in interpretation of a broad nature of signals in ^1H -NMR spectrum (Figure 4c and Figure S18).

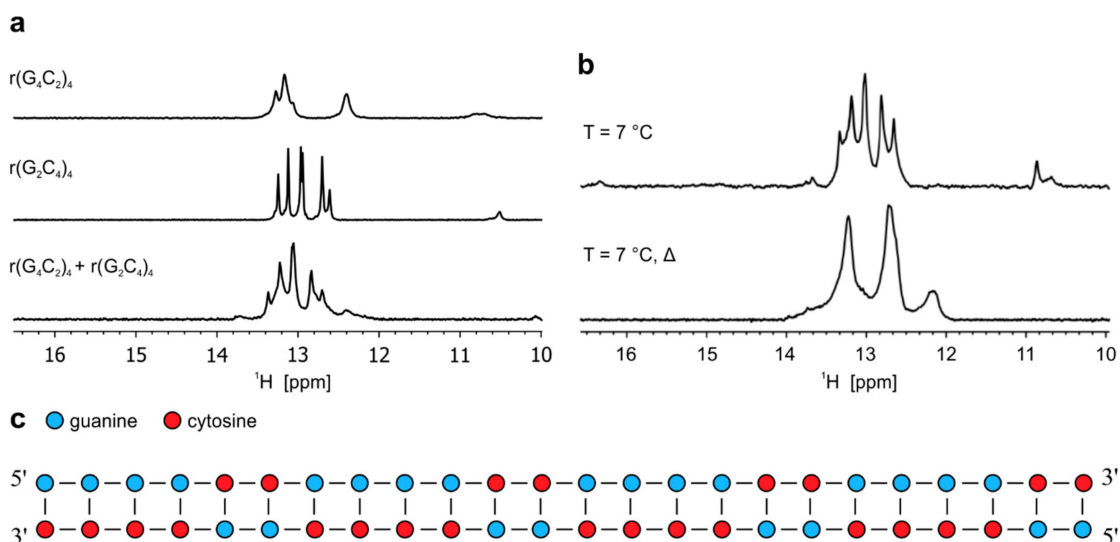


Figure 4. Formation of heterodimer by oligonucleotides $r(\text{G}_4\text{C}_2)_4$ and $r(\text{G}_2\text{C}_4)_4$. (a) ^1H -NMR spectra of $r(\text{G}_4\text{C}_2)_4$, $r(\text{G}_2\text{C}_4)_4$ and their equimolar mixture ($r(\text{G}_4\text{C}_2)_4 + r(\text{G}_2\text{C}_4)_4$) at pH 6.0 and 25 °C. (b) ^1H -NMR spectra ($r(\text{G}_4\text{C}_2)_4 + r(\text{G}_2\text{C}_4)_4$) at 7 °C before and after annealing of the sample. Δ represents NMR spectrum after annealing. All spectra were acquired at 0.2 mM oligonucleotide concentration per sense and antisense strand. (c) Schematic presentation of heterodimer adopted by $r(\text{G}_2\text{C}_4)_4$ and $r(\text{G}_4\text{C}_2)_4$.

Comparison of equimolar mixture of $r(\text{G}_2\text{C}_4)_4$ and $r(\text{G}_4\text{C}_2)_4$ with samples of individual RNA oligonucleotides on 20% native PAGE at pH 6.0 and 5 °C reveals a slow migrating band at approximately 40 bp, which corresponds to heterodimer (Figure S19). Low-intensity bands corresponding to hairpins formed by individual RNA oligonucleotides are also present at approximately 18 and 15 bp.

Heterodimer formed by $r(\text{G}_2\text{C}_4)_4$ and $r(\text{G}_4\text{C}_2)_4$ was exposed to low pH and to the presence of K^+ ions that could induce structural changes leading to formation of i-motifs and G-quadruplexes, respectively. However, lowering of pH did not result in any significant difference of imino signals in ^1H -NMR spectrum (Figure 5a). On the other hand, addition of K^+ ions to 100 mM concentration at pH 7.0 leads to additional stabilization of G-C base pairs since two new signals between δ 12.0 and 12.5 ppm appear. Imino signals that would indicate G-quadruplex formation are not observed even after additional annealing performed at 100 mM KCl (Figure 5a). Decrease in pH to 4.5 at 100 mM K^+ ions leads to slightly less distinct imino signals with no apparent change in their chemical shifts. Addition of PEG at pH 7.0 and 100 mM K^+ ions did not promote G-quadruplex formation as signals belonging to heterodimer between δ 12.0 and 13.5 ppm remain unchanged (Figure 5b). Two imino signals corresponding to G-C base pairs at δ 12.12 and 12.20 ppm disappear after lowering pH to 4.5 in the presence of PEG, while the intensity of signal at δ 13.40 ppm is significantly reduced. At the same time, signals most likely corresponding to N1-carbonyl G-G base pairs appear between δ 10.9 and 11.2 ppm, since signals corresponding to guanine imino protons within such base pairs are stabilized by lower pH [22]. Relative ratio of integrals of imino signals corresponding to G-C base pairs between δ 12.2 and 13.6 ppm and G-G base pairs between δ 10.9 and 11.2 ppm is 16:3. At low pH, it would be expected that signals corresponding to imino protons are more intense due to slower exchange with protons of bulk water in comparison to higher pH. Based on this ratio we concluded that heterodimer

exposed to low pH and PEG most likely experienced structural changes, which resulted in one base pair shift, causing a reduction of initial 24 G-C base pairs to 16 G-C base pairs and increase from zero to three G-G base pairs (Figure 5c).

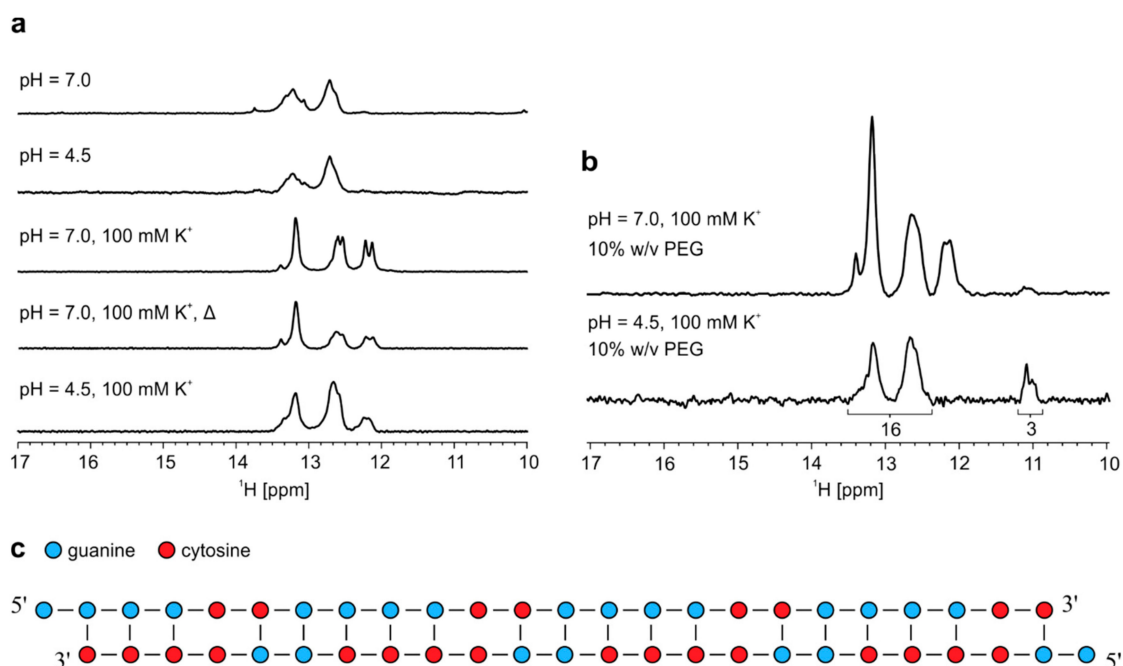


Figure 5. Effect of pH, presence of K⁺ ions and PEG on heterodimer. (a) ¹H-NMR spectra of annealed equimolar mixture of r(G₄C₂)₄ and r(G₂C₄)₄. Δ represents NMR spectrum after second annealing. All spectra were acquired on an 800 MHz NMR spectrometer at 25 °C and 0.2 mM oligonucleotide concentration per strand. (b) ¹H-NMR spectra of annealed equimolar mixture of r(G₄C₂)₄ and r(G₂C₄)₄ in the presence of 100 mM K⁺ and PEG at different pHs acquired at 25 °C and 0.2 mM oligonucleotide concentration per sense and antisense strand. Numbers under signals in the bottom spectrum represent integral values. (c) Proposed model of heterodimer in the presence of 10% w/v PEG at pH 4.5.

3. Discussion

Here we have shown that sense and antisense RNA oligonucleotides containing four hexanucleotide repeats of pathological repeat expansion of *C9orf72* gene can adopt diverse structures depending on the conditions of solution. ¹H-NMR and CD spectra together with diffusion data at pH 7.0 and 25 °C indicate that r(G₂C₄)₄ forms a homodimer with 16 Watson-Crick G-C base pairs. Homodimer was previously observed for closely related oligonucleotide r((C₄G₂)₃C₄) for which the crystal structure revealed 12 Watson-Crick G-C base pairs and 6 hemi-protonated C⁺-C base pairs [18]. Lowering of pH to 4.5 shifts the equilibrium towards formation of a hairpin with G-C base pairs, hemi-protonated C⁺-C base pairs and CCCG loop. This process was also observed in the presence of 10% w/v PEG, which was used to simulate molecular crowding conditions. The formation of hairpin adopted by r(C₄G₂)₄ was previously proposed only through data obtained with CD spectra [17]. Although i-motif adopted by hexanucleotide repeats was previously described for antisense DNA oligonucleotides [1], it has never been observed for antisense RNA oligonucleotides comprising the repeat expansion(s). In the present study ¹H-NMR spectra of r(G₂C₄)₄ at pH 4.5 and 5 °C reveal the presence of imino signal at δ 15.75 ppm. According to its chemical shift value, this signal could correspond to imino protons of hemi-protonated C⁺-C base pairs within i-motif.

Previous studies of RNA dimers showed that G-G base pairs are the most stable compared to other mismatches [23]. However, single G-G base pair is more stable in comparison to two consecutive G-G base pairs likely due to an unfavorable stacking interaction between adjacent G-G pairs [24]. Our data show that r(G₄C₂)₄ forms two structures, homodimer and hairpin, at pH 6.0 and 25 °C. Their equilibrium

populations are temperature-dependent. Lowering the temperature favors the formation of homodimer and stabilization of its inherent G-G base pairs in N1 carbonyl symmetric geometry. Both structures are stable within narrow pH range between 5.0 and 6.0. While further decrease in pH leads to disappearance of all imino signals, increase to pH 7.0 favors homodimer. In contrast to cytosine-rich antisense oligonucleotide $r(G_2C_4)_4$, guanine-rich sense oligonucleotide $r(G_4C_2)_4$ has the potential to form secondary structures called G-quadruplexes. Several studies have shown that G-quadruplexes adopted by pathological RNA molecules influence many essential cell processes such as translation, splicing, nuclear transport and function of RNA-binding proteins [8,10,25–27]. In solution RNA G-quadruplex can be stabilized by various cations, which enable their distinct folding topologies [28]. To approach physiological conditions $r(G_4C_2)_4$ was observed in the presence of K^+ ions, which are the most abundant in cells. Even in the presence of 100 mM K^+ ions for more than one week, $r(G_4C_2)_4$ does not form G-quadruplexes. Only after annealing, during which homodimer and hairpin are almost completely unfolded, a group of new signals corresponding to imino protons of guanines involved in G-quartets appears at 25 °C, thus indicating the presence of G-quadruplexes. Latter observations are in agreement with recent spectroscopic data combined with chemical and enzymatic probing of $r(G_4C_2)$ repeat expansion, which together suggest temperature-dependent equilibrium between stable hairpins formed at 37 °C and G-quadruplexes adopted at higher annealing temperatures [10,29]. Diffusion data also indicate that intra- and inter-molecular G-quadruplexes are present, which is in agreement with electrophoretic mobility shift assay where both G-quadruplexes adopted by $r(G_4C_2)_4$ in the presence of 200 mM K^+ ions at pH 7.4 were observed previously [17]. Additionally, 1H -NMR and CD spectroscopy together with results from UV melting experiments imply that oligonucleotide $r((G_4C_2)_3GGGGC)$ in the presence of 40 mM KCl forms a unimolecular four-stacked parallel-stranded G-quadruplex with propeller loops [16]. Furthermore, CD spectroscopy and RNase protection assay showed that $r(G_4C_2GGGG)$ at 100 mM K^+ ions adopt a dimeric three-stacked parallel-stranded G-quadruplex with propeller loops [10].

Transcription of the *C9orf72* gene occurs on both DNA strands leading to production of sense and antisense RNA transcripts, which are simultaneously present in the nucleus. Moreover, RNA foci of sense and antisense RNA transcripts appear to colocalize meaning that the presence of complementary strand could influence the structures formed independently by individual RNA transcripts [30,31]. In the present study 1H -NMR and CD spectra as well as PAGE indicate that sense and antisense RNA oligonucleotides form a heterodimer, while structures adopted by individual RNA oligonucleotide are still present as minor species. Exposure of heterodimer to physiological concentration of K^+ ions as well as its annealing did not lead to G-quadruplex formation. In both cases, additional G-C base pairs within the heterodimer were stabilized, which implies that heterodimer is the favorable structure under these conditions. When equimolar mixture of DNA oligonucleotides $d(G_2C_4)_8$ and $d(G_4C_2)_8$ in solution with 100 mM K^+ and low pH was exposed to PEG, signals of imino protons involved in G-quartets, G-C base pairs and hemi-protonated C^+-C base pairs within i-motifs were observed [1]. On the contrary, RNA heterodimer formed by $r(G_2C_4)_4$ and $r(G_4C_2)_4$ does not undergo structural changes at pH 4.5 and formation of hemi-protonated C^+-C base pairs associated with i-motif. Even in the presence of 100 mM K^+ ions at pH 4.5, the heterodimer maintained its structure. The effect of low pH and K^+ ions on heterodimer was also observed in the presence of PEG. While 100 mM K^+ ion concentration in the presence of 10% *w/v* PEG at pH 7.0 did not promote G-quadruplex formation, the reduction of pH to 4.5 under these conditions led to destabilization of certain G-C base pairs and most likely stabilization of N1-carbonyl G-G base pairs resulting in formation of heterodimer with one base pair shift.

Finally, described RNA structures and dynamic equilibria they are involved in could represent potential drug targets for modulation of DPRs and RNA toxicity-dependent mechanism. Structures of sense and antisense RNA transcripts could play a key role in translation where the production of DPRs could be initiated or repressed through formation of protein-binding hairpins, i-motifs and G-quadruplexes [10]. Recently, small binding molecules have been identified that reduce the overall

RNA foci and DPRs by specifically targeting G_4C_2 repeat RNA G-quadruplexes, which ultimately prolonged the life expectancy of ALS/FTLD fruit fly model [27]. In addition, study on interaction between G-quadruplex adopted by $r(G_4C_2)_8$ and porphyrin TMPyP4 has shown that TMPyP4 distorts the formation of G-quadruplex, which prevents binding of proteins associated with ALS such as hnRNPA1 [32]. Self-dimerization, intra- and inter-molecular association of described structures could, therefore, play an important part in formation of RNA foci and sequestration of RNA-binding proteins. Over the years, several studies targeting guanine-rich sense RNA transcripts with antisense DNA oligonucleotides (ASOs) turned out to be more or less successful [33,34]. ASOs targeting intronic region downstream of sense RNA repeats were developed, which reduced the overall RNA foci in fibroblasts and iPSC-derived neurons [8]. Furthermore, ASOs have been identified that significantly reduce pathogenic sense RNA foci but do not affect the level of C9orf72 encoding RNAs [35,36]. In one case reduction of RNA foci using ASO816 also led to significant correction of abnormally expressed genes after treatment of ALS patient iPSC-derived motor neurons [36]. However, designing an ASO that would specifically target and reduce antisense RNA foci remains a challenge for future studies. On the other hand, RNA duplexes were constructed, which were able to target and degrade both sense and antisense RNA foci by exploiting the RNAi pathway [37]. Although ASOs and RNAi show immense promise in treatment of ALS and FTL, half-life of RNA duplexes and long term effects of ASOs needs to be carefully assessed. Therefore, the complexity of pathological RNA molecules of C9orf72 repeat expansion associated with ALS and FTL demands further comprehensive studies as structures adopted by RNA transcripts act not only as important promoters of RNA toxicity-dependent mechanism but also present promising therapeutic targets in ALS, FTL and other repeat-associated neurodegenerative diseases.

4. Materials and Methods

4.1. Sample Preparation

Reverse-phase HPLC purified RNA oligonucleotides $r(G_2C_4)_4$ and $r(G_4C_2)_4$ were purchased from Eurogentec (Seraing, Belgium). After addition of 5 mL of 2 M LiCl samples were heated to 90 °C for 5 min. Samples were extensively dialyzed against autoclaved H_2O through four 30 min cycles and then concentrated using an Amicon ultrafilter (Merck Millipore, Herfordshire, UK). Following the overnight lyophilization, samples were diluted in 90% of autoclaved H_2O and 10% of 2H_2O . The concentrations of the 300 μ L NMR samples were 0.9, 0.7 and 0.3 mM for $r(G_2C_4)_4$ and 1.1, 0.7, 0.3, 0.2, 0.1 and 0.05 mM for $r(G_4C_2)_4$. In the case of equimolar mixing experiments, equimolar quantities of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ were diluted to final oligonucleotide concentration of 0.2 mM per sense and antisense strand. Samples of $r(G_4C_2)_4$ and mixed RNA oligonucleotides were also heated to 90 °C for 5 min prior to annealing. For molecular crowding experiments, sample of $r(G_2C_4)_4$ and mixed RNA oligonucleotides were diluted in 10% *w/v* PEG (8000 MW) (Sigma Aldrich, Munich, Germany) in autoclaved H_2O to 0.3 and 0.2 mM oligonucleotide concentration per strand, respectively. The pH of samples was adjusted by the addition of 0.1 mM LiOH or HCl.

4.2. Circular Dichroism Spectroscopy

CD spectra were acquired on Applied Photophysics Chirascan CD spectrometer using a 0.1 cm path length quartz cell. The wavelength was varied from 200 to 320 nm with 0.5 nm per step. Time-per-point was set to 0.5 s. A blank containing only autoclaved H_2O was used for baseline correction. For CD spectroscopy the sample of $r(G_2C_4)_4$ was measured immediately after dilution in autoclaved H_2O at 5, 25, and 37 °C, at pH 7.0. Measurements at different pHs were performed in autoclaved H_2O at 25 °C and pH 4.5, 5.0, 5.5, 6.0, 6.5 and 7.0. The pH of samples was adjusted by the addition of 0.1 mM LiOH or HCl. CD spectra of equimolar mixture of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ in autoclaved H_2O before and after annealing were recorded at 25 °C. CD measurements were carried out on a sample of $r(G_2C_4)_4$ and mixed RNA oligonucleotides diluted to concentration of 100 μ M per sense and antisense strand.

4.3. UV Spectroscopy

Melting experiments of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ were performed on a Varian Cary 100 Bio UV-VIS spectrometer (Varian Inc.) equipped with a thermoelectric temperature controller. UV melting experiments were performed on samples diluted to 2.6, 3.0, 7.7, 9.3 and 20 μM oligonucleotide concentration per strand in 20 mM lithium cacodylate buffer at pH 5.0 and 7.0. Each sample was placed in 1, 0.5 or 0.1 cm path-length quartz cell. Mineral oil and a tightly fixed cuvette cap were used to seal the quartz cells in order to prevent evaporation and sample loss due to high temperatures. To prevent condensation at lower temperatures a stream of nitrogen was applied throughout the measurements. Folding/unfolding processes were followed between 20 and 90 $^\circ\text{C}$ by measuring absorbance at 260 nm. Data were acquired using scan rate of 0.1 $^\circ\text{C min}^{-1}$. Temperature of half transition ($T_{1/2}$) for homodimer was determined using first derivative method, while temperature of half transition for hairpin was estimated from obtained melting curves.

4.4. Native PAGE

Native PAGE of $r(G_2C_4)_4$, $r(G_4C_2)_4$ and equimolar mixture of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ was performed on 20% polyacrylamide gel (5 $^\circ\text{C}$ at 100 V) in 1 \times TBE (pH 6.0) buffer. Samples for native PAGE were prepared by mixing 5 or 10 μL of stock or diluted RNA oligonucleotide solution, 5 μL of loading buffer 15% Ficoll and diluted to 20 μL with autoclaved H_2O . The approximate band size was determined by using the GeneRuler Ultra Low range DNA Ladder (Thermo Scientific, Waltham, MA, USA). The samples were treated at room temperature overnight before loading, except for the equimolar mixture sample, which was heated to 90 $^\circ\text{C}$ and cooled slowly at room temperature prior to loading. Following the overnight electrophoresis, the gel was stained with Stains-All gel stain solution (Sigma Aldrich).

4.5. NMR Spectroscopy

NMR spectra were obtained with Agilent Technologies VNMRS 800 MHz and DD2 600 MHz NMR spectrometers equipped with triple resonance cold probe. Standard 1D ^1H -NMR spectra were acquired with the use of DPGSE or watergate 3919 solvent suppression. Diffusion coefficient measurements were performed by a spin-echo pulse sequence with PFG gradient strengths between 1.3 and 32.5 G cm^{-1} . NOESY spectra were recorded using mixing times of 100, 80, 40 and 20 ms. TOCSY spectrum was recorded using mixing times of 80 ms. ^1H -NMR spectra were acquired at 0, 5, 7, 15, 20, 25, 37 and 40 $^\circ\text{C}$. Effect of different pHs was observed by recording ^1H -NMR spectra at pH 4.5, 5.0, 5.5, 6.0, 6.5 and 7.0. Effect of molecular crowding conditions was observed by recording ^1H -NMR spectra in the presence of 10% *w/v* PEG (8000 MW) in autoclaved H_2O . In experiments with K^+ ions 3M KCl was diluted into the sample up to the desired concentration of K^+ ions. 1D and 2D NMR spectra were processed and analysed using VNMRJ (Varian Inc.), MestReNova and Sparky (UCSF) software. All NMR spectra were acquired on 600 MHz NMR spectrometer unless stated otherwise. DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) was used as external reference in all NMR spectra.

5. Conclusions

Exploiting NMR and complementary methods allowed determination of structures adopted by guanine rich $r(G_4C_2)_4$ sense and cytosine rich $r(G_2C_4)_4$ antisense RNA oligonucleotides associated with ALS and FTL. Under conditions approaching physiological relevance, both sense and antisense RNA oligonucleotides form dimers and hairpins. Equilibrium of structures adopted by antisense RNA oligonucleotides is pH-dependent, hence lowering the pH favours hairpin formation. On the other hand, equilibrium of structures adopted by sense RNA oligonucleotides is temperature-dependent owing to prevalent formation of dimer at lower temperatures. Furthermore, sense RNA oligonucleotides form structurally different G-quadruplexes in the presence of KCl and antisense RNA oligonucleotides indicate i-motif formation at low pH and temperature. Simultaneous presence of both RNA

oligonucleotides revealed formation of heterodimers although structures adopted by individual RNA oligonucleotides seem to persist in the minority. Adopted structures of pathological RNA oligonucleotides could play a key role in sequestering RNA binding proteins and, therefore, represent an important biological target of drug design in the combat against ALS and FTLD.

Supplementary Materials: The following are available online, Figure S1: Canonical and non-canonical base pairs, Figure S2: Cytosine H5/H6 correlation region of 2D TOCSY NMR spectrum of $r(G_2C_4)_4$, Figure S3: 1H -NMR spectra of $r(G_2C_4)_4$ in 10% 2H_2O recorded at 5, 25 and 37 °C, Figure S4: UV melting experiment of homodimer adopted by $r(G_2C_4)_4$, Figure S5: Imino regions of 1H -NMR spectra of $r(G_2C_4)_4$ in 10% 2H_2O at different pH, Figure S6: Imino regions of 1H -NMR spectra of $r(G_2C_4)_4$ in 10% 2H_2O at different pH and in the presence of 10% w/v PEG, Figure S7: CD spectra of $r(G_2C_4)_4$ with pH ranging from 4.5 to 7.0, Figure S8: 20% native PAGE of $r(G_2C_4)_4$ at pH 6.0 and 5 °C, Figure S9: Imino-imino spectral region of 2D NOESY spectrum of $r(G_2C_4)_4$, Figure S10: Imino regions of 1H -NMR spectra of $r(G_2C_4)_4$ in 10% 2H_2O at 5, 25 and 40 °C, Figure S11: UV melting experiment of hairpin adopted by $r(G_2C_4)_4$, Figure S12: 1H -NMR spectra of $r(G_4C_2)_4$ in 10% 2H_2O at 0, 7, 15, 20, 25 and 37 °C, Figure S13: Imino regions of 1H -NMR spectra of $r(G_4C_2)_4$ in 10% 2H_2O at 0 and 25 °C, Figure S14: Imino and aromatic regions of 1H -NMR spectra of $r(G_4C_2)_4$ in 10% 2H_2O at different pH, Figure S15: Imino regions of 1H -NMR spectra of $r(G_4C_2)_4$ in 10% 2H_2O at concentrations of 0.05, 0.1, 0.2, 0.3, and 1.1 mM per strand, Figure S16: 20% native PAGE of $r(G_4C_2)_4$ at pH 6.0 and 5 °C, Figure S17: CD spectra of equimolar mixture of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ before and after annealing, Figure S18: 1H -NMR spectra of equimolar mixture of $r(G_2C_4)_4$ and $r(G_4C_2)_4$ before and after annealing, Figure S19: 20% native PAGE of $r(G_4C_2)_4$, $r(G_2C_4)_4$ and $r(G_4C_2)_4 + r(G_2C_4)_4$ at pH 6.0 and 5 °C, Table S1: Translational diffusion coefficients (D_t) of $r(G_2C_4)_4$ at different pH values.

Author Contributions: Conceptualization, P.Š., J.P and B.R.; investigation, T.B. and M.Z.; writing—original draft preparation, T.B., P.Š. and J.P.; supervision, P.Š. and J.P.; funding acquisition, P.Š., J.P and B.R. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by Slovenian Research Agency [ARRS, grants P1-0242, P4-0127, J1-7108, J1-1704 and J3-9263].

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

References

1. Kovanda, A.; Zalar, M.; Sket, P.; Plavec, J.; Rogelj, B. Anti-sense DNA $d(GGCCCC)_n$ expansions in C9orf72 form i-motifs and protonated hairpins. *Sci. Rep.* **2015**, *5*, 17944. [[CrossRef](#)] [[PubMed](#)]
2. Sket, P.; Pohleven, J.; Kovanda, A.; Stalekar, M.; Zupunski, V.; Zalar, M.; Plavec, J.; Rogelj, B. Characterization of DNA G-quadruplex species forming from C9orf72 G4C2-expanded repeats associated with amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **2015**, *36*, 1091–1096. [[CrossRef](#)]
3. 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 non-coding region of C9ORF72 causes chromosome 9p-linked frontotemporal dementia and amyotrophic lateral sclerosis. *Neuron* **2011**, *72*, 245–256. [[CrossRef](#)] [[PubMed](#)]
4. Renton, A.E.; Majounie, E.; Waite, A.; Simón-Sánchez, 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](#)]
5. Gómez-Tortosa, E.; Gallego, J.; Guerrero-López, R.; Marcos, A.; Gil-Neciga, E.; Sainz, M.J.; Díaz, A.; Franco-Macías, E.; Trujillo-Tiebas, M.J.; Ayuso, C.; et al. C9ORF72 hexanucleotide expansions of 20–22 repeats are associated with frontotemporal deterioration. *Neurology* **2013**, *80*, 366–370. [[CrossRef](#)] [[PubMed](#)]
6. van Blitterswijk, M.; DeJesus-Hernandez, M.; Niemantsverdriet, E.; Murray, M.E.; Heckman, M.G.; Diehl, N.N.; Brown, P.H.; Baker, M.C.; Finch, N.A.; Bauer, P.O.; et al. Associations of repeat sizes with clinical and pathological characteristics in C9ORF72 expansion carriers (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol.* **2013**, *12*, 978–988. [[CrossRef](#)]
7. Vatovec, S.; Kovanda, A.; Rogelj, B. Unconventional features of C9ORF72 expanded repeat in amyotrophic lateral sclerosis and frontotemporal lobar degeneration. *Neurobiol. Aging* **2014**, *35*, 2421.e1–2421.e12. [[CrossRef](#)]
8. Donnelly, C.J.; Zhang, P.W.; Pham, J.T.; Heusler, 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](#)]

9. Freibaum, 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. [[CrossRef](#)]
10. 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](#)]
11. Pettersson O., J.; Aagaard, L.; Jensen, T.G.; Damgaard, C.K. Molecular mechanisms in DM1 - A focus on foci. *Nucleic Acids Res.* **2015**, *43*, 2433–2441. [[CrossRef](#)]
12. Galloway, J.; Nelson, D. Evidence for RNA-mediated toxicity in the fragile X-associated tremor/ataxia syndrome. *Future Neurol.* **2009**, *4*, 785–798. [[CrossRef](#)] [[PubMed](#)]
13. 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](#)]
14. Zamiri, B.; Mirceta, M.; Bomsztyk, K.; Macgregor, R.; Pearson, C. Quadruplex formation by both G-rich and C-rich DNA strands of the C9orf72 (GGGGCC)₈•(GGCCCC)₈ repeat: Effect of CpG methylation. *Nucleic Acids Res.* **2015**, *43*, 10055–10064.
15. Brcic, J.; Plavec, J. G-quadruplex formation of oligonucleotides containing ALS and FTD related GGGGCC repeat. *Front. Chem. Sci. Eng.* **2016**, *10*, 222–237. [[CrossRef](#)]
16. Fratta, P.; Mizielska, S.; Nicoll, A.; Zloh, M.; Fisher, E.; Parkinson, G.; Isaacs, A. C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2012**, *2*, 1016. [[CrossRef](#)]
17. Reddy, K.; Zamiri, B.; Stanley, S.; Macgregor, R.; Pearson, C. 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](#)]
18. Dodd, D.; Tomchick, D.; Corey, D.; Gagnon, K. Pathogenic C9ORF72 antisense repeat RNA forms a double helix with tandem C:C mismatches. *Biochemistry* **2016**, *55*, 1283–1286. [[CrossRef](#)]
19. Steely, H.; Gray, D.; Lang, D.; Maestre, M. Circular dichroism of double-stranded RNA in the presence of salt and ethanol. *Biopolymers* **1986**, *25*, 91–117. [[CrossRef](#)] [[PubMed](#)]
20. Kypr, J.; Kejnovská, I.; Renciuik, D.; Vorlíčková, M. Circular dichroism and conformational polymorphism of DNA. *Nucleic Acids Res.* **2009**, *37*, 1713–1725. [[CrossRef](#)] [[PubMed](#)]
21. Parisien, M.; Major, F. The MC-Fold and MC-Sym pipeline infers RNA structure from sequence data. *Nature* **2008**, *452*, 51–55. [[CrossRef](#)] [[PubMed](#)]
22. Ceru, S.; Sket, P.; Prislán, I.; Lah, J.; Plavec, J. A new pathway of DNA G-quadruplex formation. *Angew. TChem. Int. Ed.* **2014**, *53*, 4881–4884. [[CrossRef](#)] [[PubMed](#)]
23. Burkard, M.; Turner, D. NMR structures of r(GCAGGCGUGC)₂ and determinants of stability for single guanosine-guanosine base pairs. *Biochemistry* **2000**, *39*, 11748–11762. [[CrossRef](#)] [[PubMed](#)]
24. Burkard, M.; Xia, T.; Turner, D. Thermodynamics of RNA internal loops with a guanosine-guanosine pair adjacent to another noncanonical pair. *Biochemistry* **2001**, *40*, 2478–2483. [[CrossRef](#)]
25. Zhang, K.; Donnelly, C.J.; Haeusler, A.R.; Grima, J.C.; Machamer, J.B.; Steinwald, P.; Daley, E.L.; Miller, S.J.; Cunningham, K.M.; Vidensky, S.; et al. The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **2015**, *525*, 56–61. [[CrossRef](#)] [[PubMed](#)]
26. Conlon, E.; Lu, L.; Sharma, A.; Yamazaki, T.; Tang, T.; Shneider, N.; Manley, J. The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *eLife* **2016**, *5*, 1–28. [[CrossRef](#)]
27. 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](#)]
28. Adrian, M.; Heddi, B.; Phan, A. NMR spectroscopy of G-quadruplexes. *Methods* **2012**, *57*, 11–24. [[CrossRef](#)]
29. 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* **2015**, *83*, 1043–1050. [[CrossRef](#)]
30. Cooper-Knock, J.; Higginbottom, A.; Stopford, M.; Highley, J.; Ince, P.; Wharton, S.; Pickering-Brown, S.; Kirby, J.; Hautbergue, G.; Shaw, P. Antisense RNA foci in the motor neurons of C9ORF72-ALS patients are associated with TDP-43 proteinopathy. *Acta Neuropathol.* **2015**, *130*, 63–75. [[CrossRef](#)]

31. Mizielinska, S.; Lashley, T.; Norona, F.; Clayton, E.; Ridler, C.; Fratta, P.; Isaacs, A. C9orf72 frontotemporal lobar degeneration is characterised by frequent neuronal sense and antisense RNA foci. *Acta Neuropathol.* **2013**, *126*, 845–857. [[CrossRef](#)] [[PubMed](#)]
32. Zamiri, B.; Reddy, K.; Macgregor, R.; Pearson, C. 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](#)]
33. Mathis, S.; Masson, G. RNA-targeted therapies and amyotrophic lateral sclerosis. *Biomedicines* **2018**, *6*. [[CrossRef](#)] [[PubMed](#)]
34. Lagier-Tourenne, C.; Baughn, M.; Rigo, F.; Sun, S.; Liu, P.; Li, H.R.; Jiang, J.; Watt, A.T.; Chun, S.; Katz, M.; et al. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E4530–E4539. [[CrossRef](#)] [[PubMed](#)]
35. 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](#)]
36. Sareen, D.; O'Rourke, J.G.; Meera, P.; Muhammad, A.K.; Grant, S.; Simpkinson, M.; Bell, S.; Carmona, S.; Ornelas, L.; Sahabian, A.; et al. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9orf72 repeat expansion. *Sci. Transl. Med.* **2013**, *5*, 208ra149. [[CrossRef](#)]
37. Hu, J.; Liu, J.; Li, L.; Gagnon, K.; Corey, D. Engineering duplex RNAs for challenging targets recognition of GGGGCC/CCCCGG repeats at the ALS/FTD C9ORF72 locus. *Chem. Biol.* **2015**, *22*, 1505–1511. [[CrossRef](#)]

Sample Availability: Not available.



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).