Fast MAS NMR Spectroscopy Can Identify G-Quartets and Double-Stranded Structures in Aggregates Formed by GGGGCC RNA Repeats

Sara Zager[1], Nataša Medved[1,2], Mirko Cevec[1], Urša Čerček[4,5], Boris Rogelj[4,6], Janez Plavec[1,2,3], Jaka Kragelj[1]

1 Slovenian NMR Centre, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia

2 University of Ljubljana, Faculty of Chemistry and Chemical Technology, Ljubljana

3 EN-FIST Centre of Excellence, Ljubljana, Slovenia

4 Department of Biotechnology, Jožef Stefan Institute, Ljubljana, 1000, Slovenia

5 Graduate School of Biomedicine, Faculty of Medicine, University of Ljubljana, Ljubljana, 1000, Slovenia

6 Faculty of Chemistry and Chemical Technology, University of Ljubljana, Ljubljana, 1000, Slovenia

The expansion of GGGGCC repeats within the *C9orf72* gene has been linked to amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD). In neurons from patients with expanded repeats, *C9orf72* GGGGCC repeat RNA predominantly forms nuclear foci, and *in vitro*, repeat-containing RNA can self-aggregate. Two structural motifs have been proposed to provide the interstrand interactions that drive aggregation: G-quartet (G4) structures and double-strand interactions with GG mismatches. Using *in vitro* transcribed RNA with physiologically relevant number of repeats, we were able to form gel-like aggregates suitable for investingation using fast MAS NMR spectroscopy. This approach enabled us to characterize the dominant interstrand interactions within the RNA gels. Both Watson–Crick and Hoogsteen base pairs were identified in RNA gels formed by RNA with 48 GGGGCC repeats. Their relative abundance shifted upon reconstitution in the presence of different divalent cations or nuclear extracts, underscoring the dynamic equilibrium between G-quadruplex and duplex interactions in GGGGCC RNA aggregation.

Introduction

Expansion of GGGGCC repeat sequences within the *C9orf72* gene is associated with neurodegenerative diseases, such as amyotrophic lateral sclerosis (ALS) and frontotemporal dementia (FTD) [1–4]. Repeat expansions with hundreds of repeats have been observed in some patients with FTD/ALS, whereas unaffected individuals have fewer than 24 repeats and most commonly only two repeats[1]. Disease mechanism of the mutation is quite complex and involves haploinsufficiency, RNA toxicity and/or toxicity of translated dipeptide repeats[5–9]. In the motor neurons of FTD/ALS patients and in model cell lines, GGGGCC repeats have been shown to form nuclear foci, which have paraspeckle like properties[1,10,11]. Interestingly, the foci are larger and more numerous when the RNA repeats are longer, and they do not form in neuronal cell lines when shorter RNA is used (e.g. 8 repeats)[9]. In addition, the long repeats form aggregates resembling gels *in vitro*, whereas short repeats do not[9,12].

The *in vitro* aggregation of RNA repeats raises the question of which structural motifs are responsible for the intermolecular association of RNA. The structure of GGGGCC repeats, both in DNA and RNA form, has been intensively studied in recent years. A wide variety of structures of GGGGCC repeats have been

determined. G-quartets (G4) have been observed to fold from a single strand or assemble as a tetramer of GGGGCC oligonucleotides[6,13–20]. Hairpins and double-stranded structures with internal loops were also observed[14,21,22]. The double-stranded structures can form in two ways: They contain internal loops with either single GG mismatches or double GG mismatches[23–25]. All these structural motifs can occur both intramolecularly and intermolecularly. For long RNAs with a large number of interacting repeats, such interactions lead to phase separation and sol-gel transition[26]. This process can be even more complex *in vivo* due to additional interactions between RNA and protein binders that are multivalent [9,14,15,27–29].

The number of repeats is an important factor when studying the large-scale organization of repeat RNA[1,9]. Biophysical techniques such as CD, NMR spectroscopy, and X-ray crystallography have been successfully used to study RNA oligonucleotides with up to 8 GGGGCC repeats[6,15,16,22,30]. Obtaining structural information on longer RNAs, ideally with more than 24 repeats in length, would be beneficial as RNAs with 8 GGGGCC repeats do not form foci and do not aggregate *in vitro*[9,12].

We show that a technique called NMR spectroscopy under fast Magic Angle Spinning (fast MAS NMR) can be used to obtain structural information on GGGGCC repeat RNAs of physiological lenghts. Fast MAS NMR can be used to obtain structural information on biomolecules regardless of their size. This means that information can be obtained for samples in "solid" phases such as fibrils, microcrystals or sedimented samples[31–35]. Very fast MAS NMR has previously been used to study RNA molecules[35–39].

We used GGGGCC repeat RNA with 48 repeats (48xG4C2) to form aggregates that resemble gels and are suitable for very fast MAS NMR spectroscopy. Both Hoogsteen and Watson–Crick base pairs were present in these gels. The balance between the two shifted depending on the divalent cation added to trigger aggregation or upon addition of nuclear extracts, underscoring the influence of the cellular environment on foci formation.

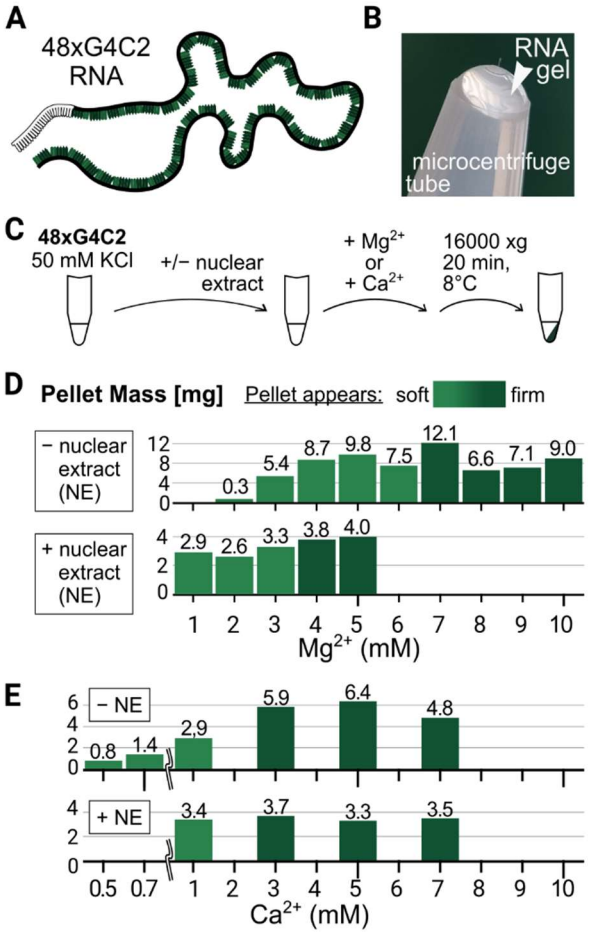

Figure 1: RNA containing 48 GGGGCC repeats forms gels upon addition of divalent ions. A) Schematic of the transcribed 48xG4C2 construct: GGGGCC repeats are colored green. B) A representative photo of an inverted Eppendorf tube containing the gel formed by 48xG4C2. C) Gels were prepared from purified RNA by addition of divalent ions (Mg^{2+} or Ca^{2+}) and by pelleting the sample in a tabletop centrifuge. D, E) Pellet properties. Higher concentrations of divalent ions result in pellets of bigger size that appear firmer and less fluid.

Results

Divalent ions trigger 48xG4C2 aggregation

Template DNA vectors for repeat RNA transcription were prepared as previously described[9]. The 48xG4C2 construct contained 48 repeats of GGGGCC (Figure 1A), preceded by a 25-base-pair flanking sequence containing all four nucleotides and ending with "GGG" at the 3' end. The transcribed and purified 48xG4C2

formed gels upon addition of the divalent ions (Figure 1). Shorter RNAs with 8 repeats or 24 GGGGCC repeats did not form gels under the same conditions (Supplementary Fig. 1). The addition of at least 5 mM Mg^{2+} was required for significant 48xG4C2 aggregation. Higher concentrations of $MgCl_2$ yielded gels that were larger and firmer. All gel formation experiments were performed in the presence of 50 mM KCl, a concentration low enough to permit gel formation under the conditions tested. The role of potassium chloride was to approach the intracellular ionic conditions and to provide conditions under which G-quadruplexes could potentially form. Compared to magnesium, calcium ions were more efficient in triggering aggregation, causing 48xG4C2 aggregation at concentrations as low as 0.5 mM $CaCl_2$ (Figure 1E). The formed gels, which were redissolved caused smeared migration of 48xG4C2 RNA on TBE-urea PAGE and accumulation of RNA in the gel pockets (Supplementary Fig. 2). This demonstrated that the interactions were strong enough to resist complete denaturation in the presence of 7M Urea.

The addition of nuclear extract (HeLaScribe® Nuclear Extract, Promega, E3092) to 48xG4C2 prior to the addition of $MgCl_2$ lowered the concentration threshold required for aggregation of 48xG4C2 (Figure 1D). This effect was not observed when $CaCl_2$ was added to 48xG4C2 pre-mixed with nuclear extracts (Figure 1E).

NMR spectroscopy of model RNA repeats

The GGGGCC repeats have the ability to form various structures, such as double-stranded RNA through Watson–Crick base-pairing, hairpins, and also G4s through Hoogsteen base pairing. To obtain representative NMR spectra that could later be compared with those of 48xG4C2, we designed two model oligonucleotides: a G4 and a double-stranded RNA with an internal loop (Figure 2A, 2B). The G4 forming sequence was $(GGGGCC)_3GGGG$, allowing for a monomolecular G4 formation. The double-strand sequence CCGGGGCCGG was designed to form a symmetric homodimer of with a double GG mismatch in the middle. After synthesis, the oligonucleotides were refolded under conditions that resulted in the most homogenous specturm. To obtain the G4 fold, we

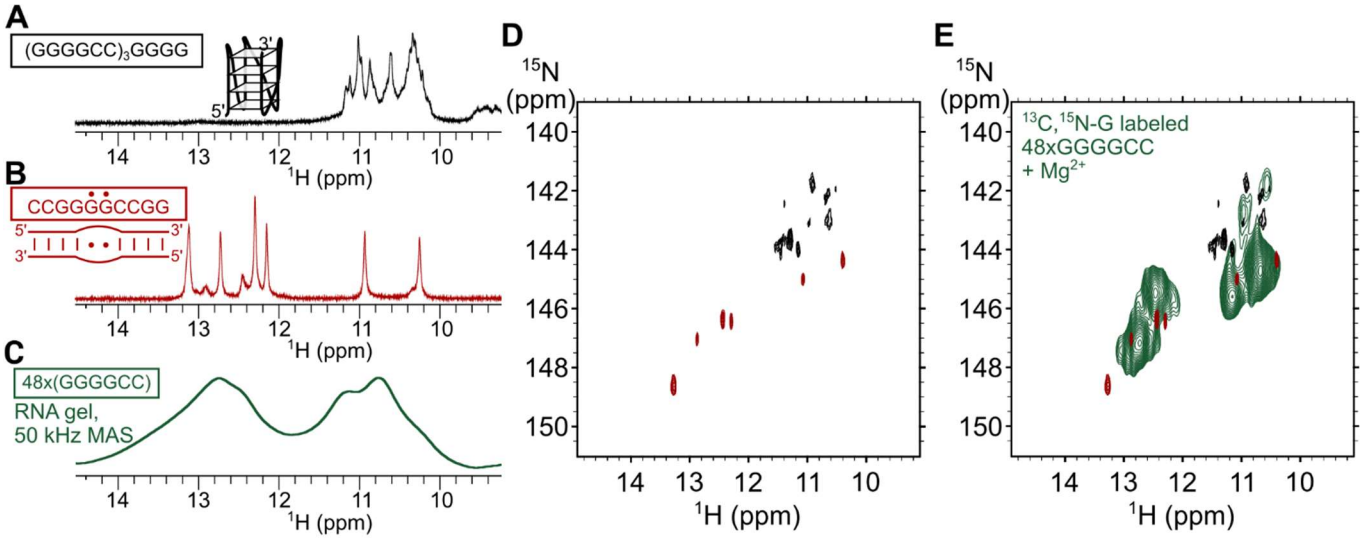

Figure 2: Solution-state NMR spectra of model RNA oligonucleotides and fast MAS NMR spectra of 48xG4C2 RNA gels. A) 1D ^1H spectrum of an oligonucleotide that folds into a G-quadruplex fold. B) 1D ^1H spectrum of an oligonucleotide folded into a double-stranded RNA with an GG internal loop. C) 1D ^1H MAS NMR spectrum of 48xG4C2 RNA gel D) SOFAST HMQC ^1H–^{15}N correlation spectrum of the double-stranded RNA oligonucleotide (red) and G-quadruplex oligonucleotide (black). E) MAS NMR hNH spectrum of 48xG4C2 RNA gel formed in the presence of 10 mM $MgCl_2$.

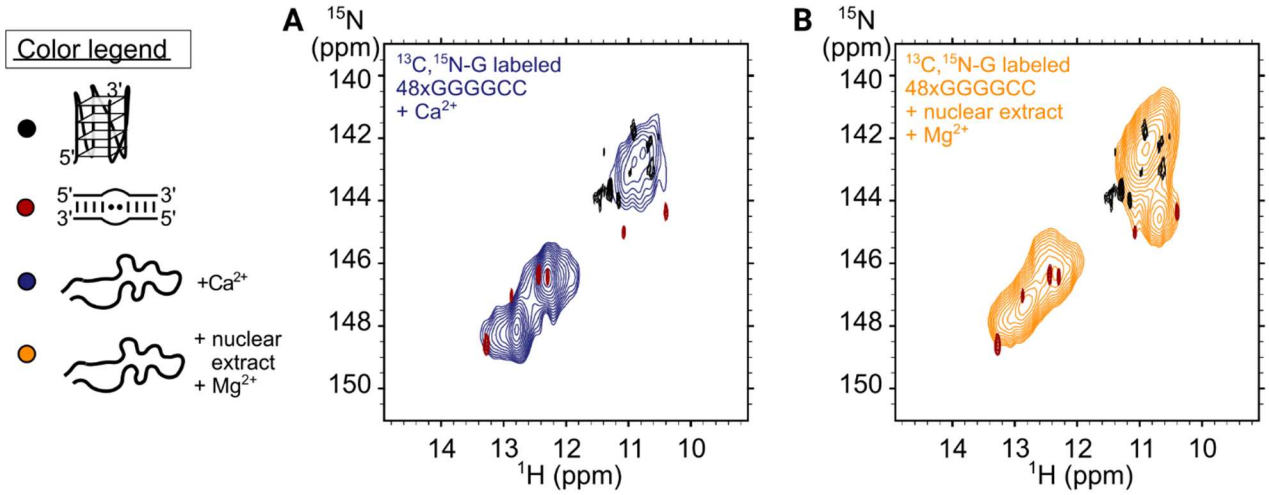

Figure 3: 2D ^{1}H–^{15}N correlation spectra of aggregated GGGGCC RNA display differences, specifically in region marking the formation G4 quartet planes through Hoogsteen base pairs. A) MAS NMR hNH spectrum of 48xG4C2 RNA gel formed in the presence of 10 mM $CaCl_2$. (B) MAS NMR hNH spectrum of 48xG4C2 RNA gel formed in the presence of nuclear extracts and 10 mM $MgCl_2$.

diluted the oligonucleotide to 10 µM, heated it to 95 °C, and then slowly annealed it to room temperature. The double-stranded homodimer was annealed at high concentration in the presence of 50 mM NaCl.

We recorded solution-state ^{1}H NMR spectra of these two model RNA oligonucleotides (Figure 2A, 2B). The G4 fold gave rise to signals in the region between 10.4 and 11.5 ppm corresponding to imino protons involved in Hoogsteen base pairs (Figure 2A). The double-stranded dimer gave rise to a series of signals in the region between 12.0 and 13.4 ppm, which is typical of Watson–Crick base pairs (Figure 2B). Two additional signals appeared at around 10.4 and 11.9 ppm, indicating GG pairing inside an internal loop. Despite some resonance overlap in 1D ^{1}H NMR spectra, the two model oligonucleotides exhibited distinct ^{15}N chemical shifts (Figure 2D). All G4 related cross-peaks appeared at chemical shifts below δN 144.1 ppm.

In summary, the spectra of the model oligonucleotides showed that double-stranded and G4 structures each gave rise to well resolved signals with distinct chemical shifts. By leveraging the additional ^{15}N dimension in 2D ^{1}H–^{15}N correlation spectra, the GG-mismatch and the GQs cross-peaks could be distinguished.

Obtaining fast MAS spectra of RNA gels

For solid-state NMR spectroscopy measurements, we prepared 48xG4C2 RNA by transcription in the presence of ^{13}C,^{15}N-GTP, resulting in 48xG4C2 containing ^{13}C,^{15}N-isotopically labeled guanines at all positions. This RNA was aggregated into gels in the presence of 10 mM $MgCl_2$. We recorded a selective 1D-^{1}H MAS NMR spectrum of 48xG4C2 RNA gels. As expected, the highly heterogeneous nature of the 48xG4C2 gels combined with spinning rate of 50 kHz gave rise to broad cross-peaks (Figure 2C). We could not determine with certainty whether G4 formation occurred in the 48xG4C2 gels. To obtain better resolution, we recorded 2D ^{1}H–^{15}N correlation MAS NMR spectra of 48xG4C2 gels (Figure 2E). The additional ^{15}N dimension allowed us to more clearly map the fingerprints of model oligonucleotides to the spectrum of 48xG4C2 which overlapped better with the peaks of the double-stranded model oligonucleotide. A weak cross-peak density was observed in the region corresponding to the G4 fold (a chemical shift below δ_N 144 ppm). However, the cross-peaks corresponding to the G4 fold were much weaker than others, suggesting that most of the base contacts

in the 48xG4C2 gel occurred through Watson–Crick base-paired guanines (Figure 2E).

The effect of $CaCl_2$ on the spectra of RNA gels

We formed 48xG4C2 RNA gels in the presence of $CaCl_2$ and measured 2D ^{1}H–^{15}N correlation MAS NMR spectra (Figure 3A). The spectrum was similar to that of 48xG4C2 gels formed in the presence of $MgCl_2$, but with more intense and prominent cross-peaks corresponding to the G4 fold (chemical shift below δ_N 144 ppm). The peak density centered at around δ_H 11 ppm overlapped better with the peaks from the G4 model oligonucleotide (black) than with the GG pair of a double-stranded GGGGCC (red). Based on this observation, it is possible that the gels formed by presence of $CaCl_2$ have a greater proportion of G4 stacking interactions than those formed in the presence of $MgCl_2$. However, the double-strand related G-C interactions remained the dominant structural feature based on dominating peak density in the downfield region (Figure 3A, between δ_H 12 and 13 ppm and δ_N 145.5 and 149.0 ppm).

The effect of nuclear extracts on RNA gels

Next, we wanted to see if components from the cell nuclei can change the conformation of the 48xG4C2 RNA. We added nuclear extracts at a volume ratio of 1:1 to ^{13}C,^{15}N-guanine labeled 48xG4C2 (Figure 1C). After a short incubation period, we induced gel formation by adding 10 mM $MgCl_2$. We then performed two washing steps (resuspension and pelleting) to remove the non-bound nuclear components. We packed the resulting 48xG4C2 gel, formed in the presence of nuclear extract, into a 1.3 mm MAS NMR rotor and acquired ^{13}C-^{15}N correlation spectra (Figure 3B). The cross-peaks mapped well to both the G4 and double-strand model oligonucleotides, without a clear preference for either of them (Figure 3B). Western blotting and SDS-PAGE analysis confirmed that proteins remained bound to 48xG4C2 gel after washes (Supplementary Fig. 3).

Conclusion

The aim of this study was to assess the suitability of fast MAS NMR for investigating RNA with a physiologically relevant number of GGGGCC repeats. We successfully identified conditions for preparing gel-like aggregates from RNA that would be suitable for analysis using solid-state NMR spectroscopy. We established an analytical approach, though still qualitative in nature, that can give us information on structure of repeat RNA in gel-like aggregates. Our findings demonstrate that fast MAS NMR spectroscopy enables atomic-resolution monitoring of structural changes in GGGGCC repeat RNA of physiological lengths.

We prepared ^{13}C,^{15}N-labeled 48xG4C2 RNA, which formed gel-like aggregates upon the addition of Mg^{2+} or Ca^{2+} ions. Fast MAS NMR spectroscopy revealed that, in the presence of Mg^{2+}, the RNA predominantly adopted a double-stranded conformation, whereas Ca^{2+} favored G4 structures. Gels formed in the presence of nuclear extracts contained both conformations. These results underscore how strongly ions and environmental factors influence the structural equilibrium of 48xG4C2 RNA. Since the conformation adopted by full-length GGGGCC repeat RNA in ALS-affected neurons may be central to disease progression, understanding this equilibrium could yield crucial insights into ALS pathogenesis.

Materials and Methods

RNA transcription and purification

The pcDNA3 plasmids containing G4C2 repeats in multiples of 8, 24, and 48 were previously constructed and described elswhere[9,10]. Transcription templates were generated by linearizing the plasmids using the BamHI

restriction enzyme, rather than by PCR amplification. Linearized products were purified and concentrated using a PCR cleanup kit.

RNA transcribed from 48xG4C2 plasmid contained a starting sequence of GAGACCCAAGCTTGTGTGGTTTA, followed by 48 G4C2 repeats, and terminated with GGG at the 3' end.

Following optimization to ensure that variations in buffer and ion composition did not significantly affect RNA yield, we selected a potassium-containing transcription buffer (20 mM HEPES pH 7.9, 100 mM KCl, 3 mM $MgCl_2$, 1 mM spermidine, 10 mM DTT, 0.05 mM EDTA, YIPP at 1:150000 dilution, 500 µM of each rNTP). For solid-state NMR samples, rGTP was replaced with ^{13}C,^{15}N-rGTP.

Each 1 ml transcription reaction contained 50 µg of linearized template and 1,650 units of T7 RNA polymerase (NEB, cat. M0251L). For example, a 10 ml reaction used 333 µl of enzyme.

Following transcription, excess nucleotides were removed by dialisys in dialysis buffer (20 mM HEPES pH 7.9, 100 mM KCl, 3 mM MgCl2, 2 mM DTT). The RNA was subsequently purified via phenol–chloroform extraction and reddisolved in water at high concentration (> 9 ug/ul).

RNA gel formation

Test precipitations were performed in microcentrifuge tubes at a final volume of 30 µl containing 20 mM HEPES, 100 mM KCl, and varying concentrations of $MgCl_2$. The final concentration of RNA was maintained at 6 µg/µl. After a 10-minute incubation at 4 °C, the samples were centrifuged at 16,000 ×g to pellet the formed RNA gel. After removing the supernatant, the resulting pellets were weighed to determine their mass. Samples of pure gelified RNA used for MAS NMR spectroscopy were prepared under the same conditions.

For RNA gels formed in the presence of nuclear extract, the HeLa Scribe Nuclear Extract (Promega, cat. E3110, now discontinued) was first clarified by centrifugation at 1,000 ×g for 10 minutes at 15 °C. Prior to gel formation, 30 µl of RNA (containing 100 mM KCl and 1 mM DTT) was pre-incubated with 24 µl of nuclear extract. The mixture of RNA and nuclear extract was then added to a new tube containing 2 µl of $MgCl_2$ and vigorously mixed to obtain a final reaction volume of 60 µl with a final $MgCl_2$ concentration of 10 mM. The resulting gel was then washed twice with wash solution (20 mM HEPES, 100 mM KCl, 10 mM $MgCl_2$, 1 mM DTT, 10% ^{13}C-depleted glycerol).

Synthesis and refolding of model oligonucleotides

Oligonucleotides were synthesized on a DNA/RNA synthesizer H8 (K&A Laborgeraete GbR, Germany) using standard solid-phase phosphoramidite chemistry in DMT-on mode. Oligonucleotides were cleaved from the solid support and fully deprotected using an aqueous ammonia/methylamine (AMA) solution (1:1 v/v) at 65 °C for 15 minutes. The RNA was lyophilized, and the TBS protecting groups were removed by dissolution in dimethyl sulfoxide (DMSO), followed by incubation with triethylamine (TEA) and triethylamine trihydrofluoride (TEA·3HF) at 65 °C for 15 minutes. The RNA oligonucleotides were then immediately purified using GlenPak RNA cartridges following the standard GlenPak procedure. Finally, the samples were lyophilized and resuspended in H_2O (HPLC grade). The concentration of oligos was determined spectroscopically (OD_{260}).

The oligonucleotides folding into a quadruplex structure with a sequence $(GGGGCC)_3GGGG$ was diluted to 10 µM in 10 mM KCl in H_2O. After heating to 95 °C for 5 minutes in a water bath, the sample was left in the switched-off water

bath overnight to cool to room temperature. The annealed oligonucleotide was concentrated using Amicon ultrafilters.

The oligonucleotide folding into the duplex with a sequence CCGGGGCCGG was refolded without dilution at concentrations greater than 100 µM in the presence of 50 mM NaCl and 0.5 mM $MgCl_2$. After heating to 95 °C for 5 minutes, the sample was cooled to room temperature by removing the microcentrifuge tube from the water bath.

NMR spectroscopy

Solution-state NMR titration experiments were performed using Bruker AVANCE NEO 600 MHz spectrometers equipped with cryogenic probes. 1D ^{1}H NMR experiments were performed with a spectral width of 19 ppm, 512 scans and a relaxation delay of 1 sec. Water suppression was achieved using excitation sculpting. Samples for NMR were prepared at concentrations ranging between 0.1 mM and 0.8 mM RNA in 95% H_2O/5% $^{2}H_2O$ supplemented with 10 mM potassium phosphate buffer (pH 7.0) and 10 mM KCl. The RNA samples were annealed at 90 °C for 5 minutes and slowly cooled overnight.

RNA gels were packed into 1.3 mm rotors by centrifugation (67,000 ×g, 16 hours, 10 °C). NMR experiments were conducted on a 23.5 T Bruker spectrometer (ω_0,H/2π = 950 MHz) equipped with a 1.3 mm HCN probe and an NEO console. During measurement, the rotor was spun at 50 kHz. Frictional heating was compensated with a gas flow of 1350 lph, at a stator temperature of 250 K, resulting in an estimated internal sample temperature at 35 °C. The 2D ^{1}H–^{15}N and ^{1}H–^{13}C correlation spectra were acquired using standard ^{1}H-detected Cross-Polarisation (CP, linear ramp) based HSQC experiment (available in NMRlib 2.0 as av_hNH_cp_cp_miss.IBS).

The 48xG4C2 gel formed in the presence of $CaCl_2$ was measured at the CRMN Lyon facility under slightly modified conditions (e.g., sample packing and spinning rate). RNA gels were packed into 1.3 mm rotors by centrifugation (150,000 ×g, 1.5 hours, 10 °C) with a specifically designed insertion tool (GiottoBiotech). The NMR experiments were conducted on a 23.5 T Bruker spectrometer (ω_0,H/2π = 1,000 MHz) equipped with a 1.3 mm HCN probe. During measurement, the rotor was spun at 60 kHz. Frictional heating was compensated with a gas flow of 1000 lph, at a stator temperature of 240 K, resulting in an estimated internal sample temperature at 25 °C. The 2D ^{1}H–^{15}N and ^{1}H–^{13}C correlation spectra were acquired using standard ^{1}H-detected Cross-Polarisation (CP)-based HSQC experiment.

Immunoblotting

48xG4C2 samples were diluted with precipitation buffer lacking $MgCl_2$ (20 mM HEPES pH 7.9, 100 mM KCl, 2 mM DTT, 5% glycerol). Dissolution of 48xG4C2 gels improved with the addition of 0.2 mM EDTA. Loading amounts were estimated based on the final dilution factor of the nuclear extract. Samples and the All Blue Prestained Protein Standard (Bio-Rad, 1610373) were loaded onto 12% stain-free gels (Bio-Rad, Hercules, California, USA) and transferred to nitrocellulose membrane using the semi-dry transfer method.

Blocking was performed for 1 hour at room temperature in blocking solution (5% skim milk in TBS with 0.05% Tween-20). Membrane was incubated overnight at 4°C with the primary anti-hnRNPH antibody (Proteintech, 67375-1-Ig) diluted 1:1000 in blocking solution. The next day, the membrane was washed three times for 10 min with TBST, then incubated with Peroxidase AffiniPure Goat Anti-Rabbit IgG (H+L) secondary antibody (111-035-144, Jackson ImmunoResearch) diluted 1:5000 in blocking solution. Clarity Max Western ECL Substrate (Bio-Rad, Hercules,

California, USA) was used for signal detection and images were captured using GelDoc System and ImageLab software (Bio-Rad, Hercules, CA, USA).

Additional author information

ORCID:

Boris Rogelj: 0000-0003-3898-1943

Janez Plavec: 0000-0003-1570-8602

Jaka Kragelj: 0000-0001-6472-8429

Author contributions

Sara Zager prepared G4C2 RNA, synthesized RNA oligos, and conducted gelation assays. Nataša Medved synthesized RNA oligos. Mirko Cevec carried out solution-state NMR measurements. Urša Čerček prepared plasmids, performed western blots, and contributed to manuscript revisions. Boris Rogelj and Janez Plavec provided project guidance and feedback, as well as critical manuscript revisions. Jaka Kragelj designed the project goals, established the protocols for sample preparations, prepared G4C2 RNA and samples for MAS NMR, analyzed data, and authored the manuscript.

Acknowledgments

We would like to thank Tanguy Le Marchand from CRMN Lyon and Alicia Vallet from IBS Grenoble for performing MAS NMR spectroscopy measurements. We would like to thank the members of the Jernej Ule lab, particularly Tajda Klobučar, for their assistance with analyzing RNA sample quality and performing analytical RNA purifications. Additionally, we thank Marjetka Podobnik, Head of the Department of Molecular Biology and Nanobiotechnology, for providing access to microbiology equipment, and Tanja Peric for her help with the cultures required for plasmid purification.

Funding

This project was funded through ARIS program “Pečat odličnosti” (ARRS-MS-MSCA-SE-2021). Authors B.R. and U.Č. and J.N. were supported by Slovenia Research and Innovation Agency (P4-0127, J3-4503, J3-60057, J7-60125, J1-50026, J2-60047 and J3-3065).

This work benefited from access to RALF-NMR. MAS NMR measurements have been performed at CRMN in Lyon and at IBS in Grenoble and were supported by iNEXT-Discovery (project number 22024, project number 871037), funded by the Horizon 2020 program of the European Commission.

The authors gratefully acknowledge the financial support from the Slovenian Research and Innovation Agency (ARIS, Grants P1-0242 and J1-60019). The authors acknowledge the CERIC-ERIC consortium for the access to experimental facilities and financial support.

References

(1) 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.; Kouri, N.; Wojtas, A.; Sengdy, P.; Hsiung, G.-Y. R.; Karydas, A.; Seeley, W. W.; Josephs, K. A.; Coppola, G.; Geschwind, D. H.; Wszolek, Z. K.; Feldman, H.; Knopman, D. S.; Petersen, R. C.; Miller, B. L.; Dickson, D. W.; Boylan, K. B.; Graff-Radford, N. R.; Rademakers, R. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **2011**, *72* (2), 245–256. https://doi.org/10.1016/j.neuron.2011.09.011.

(2) 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.; Kalimo, H.; Paetau, A.; Abramzon, Y.; Remes, A. M.; Kaganovich, A.; Scholz, S. W.; Duckworth, J.; Ding, J.; Harmer, D. W.; Hernandez, D. G.; Johnson, J. O.; Mok, K.; Ryten, M.; Trabzuni, D.; Guerreiro, R. J.; Orrell, R. W.; Neal, J.; Murray, A.; Pearson, J.; Jansen, I. E.; Sondervan, D.; Seelaar, H.; Blake, D.; Young, K.; Halliwell, N.; Callister, J. B.; Toulson, G.; Richardson, A.; Gerhard, A.; Snowden, J.; Mann, D.; Neary, D.; Nalls, M. A.; Peuralinna, T.; Jansson, L.; Isoviita, V.-M.; Kaivorinne, A.-L.; Hölttä-Vuori, M.; Ikonen, E.; Sulkava, R.; Benatar, M.; Wuu, J.; Chiò, A.; Restagno, G.; Borghero, G.; Sabatelli, M.; ITALSGEN Consortium; Heckerman, D.; Rogaeva, E.; Zinman, L.; Rothstein, J. D.; Sendtner, M.; Drepper, C.; Eichler, E. E.; Alkan, C.; Abdullaev, Z.; Pack, S. D.; Dutra, A.; Pak, E.; Hardy, J.; Singleton, A.; Williams, N. M.; Heutink, P.; Pickering-Brown, S.; Morris, H. R.; Tienari, P. J.; Traynor, B. J. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **2011**, *72* (2), 257–268. https://doi.org/10.1016/j.neuron.2011.09.010.

(3) Gijselinck, I.; Van Langenhove, T.; van der Zee, J.; Sleegers, K.; Philtjens, S.; Kleinberger, G.; Janssens, J.; Bettens, K.; Van Cauwenberghe, C.; Pereson, S.; Engelborghs, S.; Sieben, A.; De Jonghe, P.; Vandenberghe, R.; Santens, P.; De Bleecker, J.; Maes, G.; Bäumer, V.; Dillen, L.; Joris, G.; Cuijt, I.; Corsmit, E.; Elinck, E.; Van Dongen, J.; Vermeulen, S.; Van den Broeck, M.; Vaerenberg, C.; Mattheijssens, M.; Peeters, K.; Robberecht, W.; Cras, P.; Martin, J.-J.; De Deyn, P. P.; Cruts, M.; Van Broeckhoven, C. A C9orf72 Promoter Repeat Expansion in a Flanders-Belgian Cohort with Disorders of the Frontotemporal Lobar Degeneration-Amyotrophic Lateral Sclerosis Spectrum: A Gene Identification Study. *Lancet Neurol* **2012**, *11* (1), 54–65. https://doi.org/10.1016/S1474-4422(11)70261-7.

(4) Waite, A. J.; Bäumer, D.; East, S.; Neal, J.; Morris, H. R.; Ansorge, O.; Blake, D. J. Reduced C9orf72 Protein Levels in Frontal Cortex of Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration Brain with the *C9ORF72* Hexanucleotide Repeat Expansion. *Neurobiology of Aging* **2014**, *35* (7), 1779.e5-1779.e13. https://doi.org/10.1016/j.neurobiolaging.2014.01.016.

(5) Vatovec, S.; Kovanda, A.; Rogelj, B. Unconventional Features of C9ORF72 Expanded Repeat in Amyotrophic Lateral Sclerosis and Frontotemporal Lobar Degeneration. *Neurobiology of Aging* **2014**, *35* (10), 2421.e1-2421.e12. https://doi.org/10.1016/j.neurobiolaging.2014.04.015.

(6) Šket, P.; Pohleven, J.; Kovanda, A.; Štalekar, M.; Župunski, 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* (2), 1091–1096. https://doi.org/10.1016/j.neurobiolaging.2014.09.012.

(7) Kovanda, A.; Zalar, M.; Šket, P.; Plavec, J.; Rogelj, B. Anti-Sense DNA d(GGCCCC)n Expansions in C9ORF72 Form i-Motifs and Protonated Hairpins. *Sci Rep* **2015**, *5* (1), 17944. https://doi.org/10.1038/srep17944.

(8) Solomon, D. A.; Stepto, A.; Au, W. H.; Adachi, Y.; Diaper, D. C.; Hall, R.; Rekhi, A.; Boudi, A.; Tziortzouda, P.; Lee, Y.-B.; Smith, B.; Bridi, J. C.; Spinelli, G.; Dearlove, J.; Humphrey, D. M.; Gallo, J.-M.; Troakes, C.; Fanto, M.; Soller, M.; Rogelj, B.; Parsons, R. B.; Shaw, C. E.; Hortobágyi, T.; Hirth, F. A Feedback Loop between Dipeptide-Repeat Protein, TDP-43 and Karyopherin-α Mediates C9orf72-Related Neurodegeneration. *Brain* **2018**, *141* (10), 2908–2924. https://doi.org/10.1093/brain/awy241.

(9) Lee, Y.-B.; Chen, H.-J.; Peres, J. N.; Gomez-Deza, J.; Attig, J.; Stalekar, M.; Troakes, C.; Nishimura, A. L.; Scotter, E. L.; Vance, C.; Adachi, Y.; Sardone, V.; Miller, J. W.; Smith, B. N.; Gallo, J.-M.; Ule, J.; Hirth, F.; Rogelj, B.; Houart,

C.; Shaw, C. E. Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *Cell Rep* **2013**, *5* (5), 1178–1186. https://doi.org/10.1016/j.celrep.2013.10.049.

(10) Malnar Črnigoj, M.; Čerček, U.; Yin, X.; Ho, M. T.; Repic Lampret, B.; Neumann, M.; Hermann, A.; Rouleau, G.; Suter, B.; Mayr, M.; Rogelj, B. Phenylalanine-tRNA Aminoacylation Is Compromised by ALS/FTD-Associated C9orf72 C4G2 Repeat RNA. *Nat Commun* **2023**, *14* (1), 5764. https://doi.org/10.1038/s41467-023-41511-3.

(11) Bajc Česnik, A.; Darovic, S.; Prpar Mihevc, S.; Štalekar, M.; Malnar, M.; Motaln, H.; Lee, Y.-B.; Mazej, J.; Pohleven, J.; Grosch, M.; Modic, M.; Fonovič, M.; Turk, B.; Drukker, M.; Shaw, C. E.; Rogelj, B. Nuclear RNA Foci from C9ORF72 Expansion Mutation Form Paraspeckle-like Bodies. *Journal of Cell Science* **2019**, *132* (5), jcs224303. https://doi.org/10.1242/jcs.224303.

(12) Jain, A.; Vale, R. D. RNA Phase Transitions in Repeat Expansion Disorders. *Nature* **2017**, *546* (7657), 243–247. https://doi.org/10.1038/nature22386.

(13) Brčić, J.; Plavec, J. NMR Structure of a G-Quadruplex Formed by Four d(G4C2) Repeats: Insights into Structural Polymorphism. *Nucleic Acids Res* **2018**, *46* (21), 11605–11617. https://doi.org/10.1093/nar/gky886.

(14) Haeusler, A. R.; Donnelly, C. J.; Periz, G.; Simko, E. A. J.; Shaw, P. G.; Kim, M.-S.; Maragakis, N. J.; Troncoso, J. C.; Pandey, A.; Sattler, R.; Rothstein, J. D.; Wang, J. C9orf72 Nucleotide Repeat Structures Initiate Molecular Cascades of Disease. *Nature* **2014**, *507* (7491), 195–200. https://doi.org/10.1038/nature13124.

(15) 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* (14), 9860–9866. https://doi.org/10.1074/jbc.C113.452532.

(16) Zhou, B.; Geng, Y.; Liu, C.; Miao, H.; Ren, Y.; Xu, N.; Shi, X.; You, Y.; Lee, T.; Zhu, G. Characterizations of Distinct Parallel and Antiparallel G-Quadruplexes Formed by Two-Repeat ALS and FTD Related GGGGCC Sequence. *Sci Rep* **2018**, *8* (1), 2366. https://doi.org/10.1038/s41598-018-20852-w.

(17) 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* (1), 1016. https://doi.org/10.1038/srep01016.

(18) Fay, M. M.; Anderson, P. J.; Ivanov, P. ALS/FTD-Associated C9ORF72 Repeat RNA Promotes Phase Transitions In Vitro and in Cells. *Cell Reports* **2017**, *21* (12), 3573–3584. https://doi.org/10.1016/j.celrep.2017.11.093.

(19) Zhang, Y.; Yang, M.; Duncan, S.; Yang, X.; Abdelhamid, M. A. S.; Huang, L.; Zhang, H.; Benfey, P. N.; Waller, Z. A. E.; Ding, Y. G-Quadruplex Structures Trigger RNA Phase Separation. *Nucleic Acids Res* **2019**, *47* (22), 11746–11754. https://doi.org/10.1093/nar/gkz978.

(20) Geng, Y.; Liu, C.; Cai, Q.; Luo, Z.; Miao, H.; Shi, X.; Xu, N.; Fung, C. P.; Choy, T. T.; Yan, B.; Li, N.; Qian, P.; Zhou, B.; Zhu, G. Crystal Structure of Parallel G-Quadruplex Formed by the Two-Repeat ALS- and FTD-Related GGGGCC Sequence. *Nucleic Acids Res* **2021**, *49* (10), 5881–5890. https://doi.org/10.1093/nar/gkab302.

(21) 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.; Johnston, A.; Overstreet, K.; Oh, S.-Y.; Todd, P. K.; Berry, J. D.; Cudkowicz, M. E.; Boeve, B. F.; Dickson, D.; Floeter, M. K.; Traynor, B. J.; Morelli, C.; Ratti, A.; Silani, V.; Rademakers, R.; Brown, R. H.; Rothstein, J. D.; Boylan, K. B.; Petrucelli, L.; Disney, M. D. Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **2014**, *83* (5), 1043–1050. https://doi.org/10.1016/j.neuron.2014.07.041.

(22) Božič, T.; Zalar, M.; Rogelj, B.; Plavec, J.; Šket, P. Structural Diversity of Sense and Antisense RNA Hexanucleotide Repeats Associated with ALS and FTLD. *Molecules* **2020**, *25* (3), 525. https://doi.org/10.3390/molecules25030525.

(23) Maity, A.; Winnerdy, F. R.; Chen, G.; Phan, A. T. Duplexes Formed by G4C2 Repeats Contain Alternate Slow- and Fast-Flipping G·G Base Pairs. *Biochemistry* **2021**, *60* (14), 1097–1107. https://doi.org/10.1021/acs.biochem.0c00916.

(24) Taghavi, A.; Baisden, J. T.; Childs-Disney, J. L.; Yildirim, I.; Disney, M. D. Conformational Dynamics of RNA G4C2 and G2C4 Repeat Expansions Causing ALS/FTD Using NMR and Molecular Dynamics Studies. *Nucleic Acids Research* **2023**, *51* (11), 5325–5340. https://doi.org/10.1093/nar/gkad403.

(25) Ursu, A.; Baisden, J. T.; Bush, J. A.; Taghavi, A.; Choudhary, S.; Zhang, Y.-J.; Gendron, T. F.; Petrucelli, L.; Yildirim, I.; Disney, M. D. A Small Molecule Exploits Hidden Structural Features within the RNA Repeat Expansion That Causes c9ALS/FTD and Rescues Pathological Hallmarks. *ACS Chem. Neurosci.* **2021**, *12* (21), 4076–4089. https://doi.org/10.1021/acschemneuro.1c00470.
(26) Li, P.; Banjade, S.; Cheng, H.-C.; Kim, S.; Chen, B.; Guo, L.; Llaguno, M.; Hollingsworth, J. V.; King, D. S.; Banani, S. F.; Russo, P. S.; Jiang, Q.-X.; Nixon, B. T.; Rosen, M. K. Phase Transitions in the Assembly of Multi-Valent Signaling Proteins. *Nature* **2012**, *483* (7389), 336–340. https://doi.org/10.1038/nature10879.
(27) 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. https://doi.org/10.7554/eLife.17820.
(28) Zamiri, B.; Reddy, K.; Macgregor, R. B.; 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* (8), 4653–4659. https://doi.org/10.1074/jbc.C113.502336.
(29) Currie, S. L.; Rosen, M. K. Using Quantitative Reconstitution to Investigate Multicomponent Condensates. *RNA* **2022**, *28* (1), 27–35. https://doi.org/10.1261/rna.079008.121.
(30) Beseiso, D.; Chen, E. V.; McCarthy, S. E.; Martin, K. N.; Gallagher, E. P.; Miao, J.; Yatsunyk, L. A. The First Crystal Structures of Hybrid and Parallel Four-Tetrad Intramolecular G-Quadruplexes. *Nucleic Acids Res* **2022**, *50* (5), 2959–2972. https://doi.org/10.1093/nar/gkac091.
(31) Martin, R. W.; Zilm, K. W. Preparation of Protein Nanocrystals and Their Characterization by Solid State NMR. *Journal of Magnetic Resonance* **2003**, *165* (1), 162–174. https://doi.org/10.1016/S1090-7807(03)00253-2.
(32) Marulanda, D.; Tasayco, M. L.; Cataldi, M.; Arriaran, V.; Polenova, T. Resonance Assignments and Secondary Structure Analysis of E. Coli Thioredoxin by Magic Angle Spinning Solid-State NMR Spectroscopy. *J. Phys. Chem. B* **2005**, *109* (38), 18135–18145. https://doi.org/10.1021/jp052774d.
(33) Reif, B.; Ashbrook, S. E.; Emsley, L.; Hong, M. Solid-State NMR Spectroscopy. *Nat Rev Methods Primers* **2021**, *1* (1), 1–23. https://doi.org/10.1038/s43586-020-00002-1.
(34) Tang, M.; Comellas, G.; Rienstra, C. M. Advanced Solid-State NMR Approaches for Structure Determination of Membrane Proteins and Amyloid Fibrils. *Acc. Chem. Res.* **2013**, *46* (9), 2080–2088. https://doi.org/10.1021/ar4000168.
(35) Le Marchand, T.; Schubeis, T.; Bonaccorsi, M.; Paluch, P.; Lalli, D.; Pell, A. J.; Andreas, L. B.; Jaudzems, K.; Stanek, J.; Pintacuda, G. 1H-Detected Biomolecular NMR under Fast Magic-Angle Spinning. *Chem. Rev.* **2022**, *122* (10), 9943–10018. https://doi.org/10.1021/acs.chemrev.1c00918.
(36) Aguion, P. I.; Marchanka, A. Strategies for RNA Resonance Assignment by 13C/15N- and 1H-Detected Solid-State NMR Spectroscopy. *Front. Mol. Biosci.* **2021**, *8*. https://doi.org/10.3389/fmolb.2021.743181.
(37) Aguion, P. I.; Kirkpatrick, J.; Carlomagno, T.; Marchanka, A. Identification of RNA Base Pairs and Complete Assignment of Nucleobase Resonances by Proton-Detected Solid-State NMR Spectroscopy at 100 kHz MAS. *Angewandte Chemie International Edition* **2021**, *60* (44), 23903–23910. https://doi.org/10.1002/anie.202107263.
(38) Zhao, S.; Li, X.; Wen, Z.; Zou, M.; Yu, G.; Liu, X.; Mao, J.; Zhang, L.; Xue, Y.; Fu, R.; Wang, S. Dynamics of Base Pairs with Low Stability in RNA by Solid-State Nuclear Magnetic Resonance Exchange Spectroscopy. *iScience* **2022**, *25* (11), 105322. https://doi.org/10.1016/j.isci.2022.105322.
(39) Yang, Y.; Xiang, S.; Liu, X.; Pei, X.; Wu, P.; Gong, Q.; Li, N.; Baldus, M.; Wang, S. Proton-Detected Solid-State NMR Detects the Inter-Nucleotide Correlations and Architecture of Dimeric RNA in Microcrystals. *Chem. Commun.* **2017**, *53* (96), 12886–12889. https://doi.org/10.1039/C7CC07483B.