



HAL
open science

Self-Assembly of DNA and RNA Building Blocks Explored by Nitrogen-14 NMR Crystallography: Structure and Dynamics

Diego Carnevale, Marcel Hollenstein, Geoffrey Bodenhausen

► **To cite this version:**

Diego Carnevale, Marcel Hollenstein, Geoffrey Bodenhausen. Self-Assembly of DNA and RNA Building Blocks Explored by Nitrogen-14 NMR Crystallography: Structure and Dynamics. ChemPhysChem, 2020, 21 (10), pp.1044-1051. 10.1002/cphc.201901214 . pasteur-02625685

HAL Id: pasteur-02625685

<https://hal-pasteur.archives-ouvertes.fr/pasteur-02625685>

Submitted on 23 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

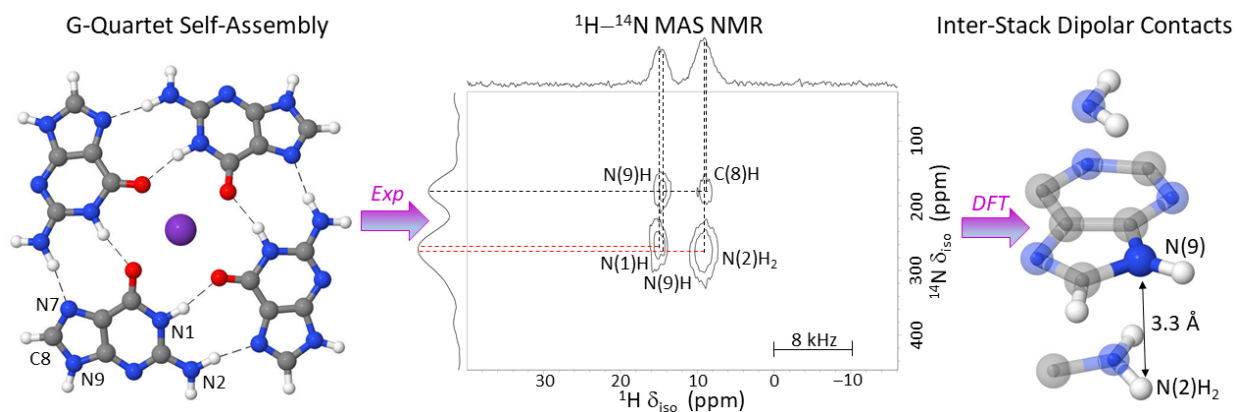
L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Self-Assembly of DNA and RNA Building Blocks Explored by Nitrogen-14 NMR Crystallography: Structure and Dynamics

Diego Carnevale,^{1,*} Marcel Hollenstein² and Geoffrey Bodenhausen¹

¹ *Laboratoire des biomolécules, LBM, Département de chimie, École normale supérieure, PSL University, Sorbonne Université, CNRS, 75005 Paris, France*

² *Institut Pasteur, Department of Structural Biology and Chemistry, Laboratory for Bioorganic Chemistry of Nucleic Acids, CNRS UMR 3523, 28, rue du Docteur Roux, 75724 Paris Cedex 15, France*



To be cited as: ChemPhysChem 10.1002/cphc.201901214

Link to VoR: <http://dx.doi.org/10.1002/cphc.201901214>

Corresponding author: Diego Carnevale (diego.carnevale@ens.fr)

Abstract

The isotopic enrichment of nucleic acids with nitrogen-15 is often carried out by solid-phase synthesis of oligonucleotides using phosphoramidites precursors that are synthetically demanding and expensive. These synthetic challenges, combined with the overlap of chemical shifts, explain the lag of nitrogen-15 NMR studies of nucleic acids behind those of proteins. For the structural characterization of DNA and RNA-related systems, new NMR methods that exploit the naturally occurring 99.9% abundant nitrogen-14 isotope are therefore highly desirable. In this study, we have investigated nitrogen-14 spectra of self-assembled quartets based on the nucleobase guanine in the solid state by means of magic-angle spinning NMR spectroscopy. The network of dipolar proton-nitrogen couplings between neighboring stacked purine units is probed by 2D spectra based on $^1\text{H} \rightarrow ^{14}\text{N} \rightarrow ^1\text{H}$ double cross-polarization. Interplane dipolar contacts are identified between the stacked G quartets. The assignment is supported by density functional theory (DFT) calculations of the anisotropic chemical shifts and quadrupolar parameters. The experimental spectra are fully consistent with internuclear distances obtained *in silico*. Averaging of chemical shifts due to internal motions can be interpreted by semiempirical calculations. This method can easily be extended to synthetic G quartets based on nucleobase or nucleoside analogs and potentially to oligonucleotides.

Key words

Nucleic Acids, Nitrogen-14 NMR, Solid-State NMR, Guanine Quartets, Self-Assembly, Stacking Interactions, DFT Calculations of NMR Parameters, Quadrupolar Couplings, NH_2 Dynamics, Counterion Mobility.

Introduction

Playing a fundamental role in the structural organization of macromolecular assemblies such as the DNA double helix, double-stranded RNA, proteins, protein complexes, and virus capsids, self-assembly processes lay at the very heart of biochemical systems. The possibility to utilize a small number of building blocks that undergo self-organization to form superstructures with complex functionalities is very interesting from a synthetic perspective.^{1,2} The understanding of how intermolecular interactions such as H-bonds, π - π stacking or coordination to metal centers guide self-assembly processes is crucial for a rational design of novel supramolecular machines.³ The quest to rationalize and purposely exploit these interactions and recognition motifs is particularly evident in crystal engineering, where one seeks to predict the packing of single molecules to achieve a specific self-organized and functional solid-state structure.^{4,5}

In this context, nucleic acids are prime examples of biopolymers that can self-assemble into a broad range of three-dimensional structures, including hairpins, bulges, cruciform and four-strand Holliday junctions, and of course double- or triple-stranded helices.⁶ Of the different conformations that can be adopted by nucleic acids, G-quadruplexes that comprise four guanine bases play a major role in a number of critical biological processes such as DNA replication, gene recombination, translation, transcription, and genome and telomere maintenance.⁷⁻⁹ Due to their omnipresence and biological relevance, the structural investigation of RNA^{10,11} and DNA^{12,13} G-quadruplexes is of particular interest.^{14,15} Besides their biological relevance, nucleoside analogs capable of forming short synthetic G quartets have attracted increased interest for the development of novel materials such as organic frameworks or hydrogels.¹⁶⁻¹⁹

The structural diversity adopted by nucleic acids calls for the development of potent and reliable methods for their structural characterization at atomic resolution. In this context, CD spectroscopy, X-ray crystallography, Small Angle X-Ray Scattering (SAXS), ¹H and ¹³C NMR, and most recently cryo-electron microscopy (cryo-EM) are the main biophysical methods used to study conformational changes and structural features of nucleic acids,²⁰ in order to better understand their biological relevance and roles. However, the chemical nature of nucleic acids

imposes a certain number of limitations and challenges to each of these methods. For instance, nucleic acids may be difficult to crystallize due to their rather flexible nature²¹ while the repetitive structural motives reduce the efficiency of crystal packing and often complicate molecular replacements protocols, so that the outcome of X-ray crystallography is fraught with uncertainties.²² One of the advantages of NMR compared to other biophysical methods is that it permits not only the elucidation of the structure of the ground state²³ but also the structures of conformers and short-lived intermediates,²⁴ thus giving valuable insight into the dynamics of nucleic acids in solution.²⁵ However, the structural elucidation of longer DNA and RNA sequences by ¹H NMR, or even of ¹H/¹³C or ¹H/¹⁵N correlation spectra, is impaired by severe signal overlap in proton spectra due to signal degeneracy, the narrow range of the chemical shifts of non-exchangeable protons, low overall proton density and, when exploiting Overhauser effects, the limited number of long-range ¹H contacts.^{26–28} Hence, structural studies of longer oligonucleotides heavily rely on the labeling of these biopolymers with stable ¹³C and ¹⁵N isotopes.^{29–31} However, the isotopic ¹⁵N/¹³C enrichment of nucleosides incorporated into oligonucleotides by solid-phase synthesis is complex, labor-intensive and expensive, particularly in longer sequences. Indeed, the introduction of these isotopes into nucleosides requires either modified phosphoramidite building blocks for solid phase synthesis or nucleoside triphosphate analogs for enzymatic DNA or RNA production.^{32–36} Since the preparation of these isotope-enriched building blocks requires multistep synthetic protocols, the possibility of characterizing nitrogen environments by exploiting the naturally abundant 99.9% ¹⁴N isotope without recourse to isotopic enrichment is particularly appealing.

Magic-angle spinning (MAS) NMR spectroscopy has become an invaluable tool for structural investigations in either ordered or amorphous solid materials. A wealth of information about chemical environments on the atomic scale is accessible by monitoring anisotropic interactions such as those due to the chemical shift anisotropy (CSA), dipole-dipole (DD) or quadrupolar couplings.³⁷ This is particularly useful when dealing with amorphous or heterogeneous samples where diffraction techniques are not straightforward.³⁸ In combination with density functional theory (DFT) calculations of CSA and quadrupolar tensors, MAS-NMR has

emerged as a new field of research that has been tentatively termed ‘NMR crystallography’.³⁹ In this context, DFT methods have proven to be extremely helpful in NMR studies aimed at the elucidation of the structure and dynamics in solids.^{40–43} We have successfully utilized this combined approach in a variety of studies of self-assembly ranging from boroxoaromatics for crystal engineering,⁴⁴ benzene-1,3,5-tricarboxamide for columnar discotic systems,⁴⁵ to arene-ruthenium rectangles stabilized by H-bonding and coordination interactions.⁴⁶

Nitrogen-14 is a highly abundant nucleus that occurs regularly in all biological macromolecules, i.e., in DNA, RNA and proteins. Unfortunately, besides being a low- γ nucleus ($\gamma(^{14}\text{N})/\gamma(^1\text{H}) = 0.07$, which may be compared with $\gamma(^{15}\text{N})/\gamma(^1\text{H}) = 0.10$), it has a quadrupolar spin with $I = 1$ and a quadrupolar coupling $\omega_Q/2\pi$ that can be as large as 5 MHz. In solution state, ^{14}N spectra are therefore so severely broadened that they cannot be observed, except in a few highly symmetrical molecules with a nearly tetrahedral structure where the quadrupole coupling is small. In solids, nuclei with spin $I = 1$ are particularly challenging for NMR spectroscopy. In contrast to nuclei with half-integer spins $I = 3/2, 5/2$, etc., that have a central transition $m_1 = \pm 1/2 \leftrightarrow m_1 = \mp 1/2$ which is not affected by the quadrupolar interaction to first order, nuclei with integer spins $I = 1$ only have two allowed single-quantum transitions $m_1 = \pm 1 \leftrightarrow m_1 = 0$ that are displaced by $\pm\omega_Q$ with respect to the center of the spectrum, which is itself affected by a second-order quadrupolar term. Nonetheless, NMR methodologies for the acquisition of ^{14}N spectra in the solid state have been developed over the past 15 years for both static and spinning samples.^{47–63} Techniques involving overtone irradiation have been explored.^{64–72} The use of signal enhancement by DNP has also been evaluated.^{62,71,73} **Structural investigations of guanosine self-assemblies utilizing HMQC techniques have also been reported.**^{74,75} However, NMR studies exploiting the ^{14}N isotope are far from routine applications. We have recently investigated the performance and efficiency of $^1\text{H} \rightarrow ^{14}\text{N} \rightarrow ^1\text{H}$ double cross-polarization (DCP) schemes for the indirect detection of ^{14}N spectra, and found these methods to be remarkably straightforward compared to earlier techniques.⁷⁶ We have illustrated the power of this DCP technique by investigating networks of ($^1\text{H}, ^{14}\text{N}$) dipole-dipole couplings in powders of the cyclic undecapeptide cyclosporin, an important immunosuppressant drug.⁷⁷

In this work, we focus on ($^1\text{H},^{14}\text{N}$) dipole-dipole couplings that offer a means to probe internuclear spatial proximities in powdered samples of G **quartet self-assemblies formed by the nucleobase guanine** by means of DCP experiments. Both intra- and inter-plane dipolar contacts can readily be identified when comparing short- and long-range 2D correlation experiments. Inter-plane dipolar contacts provide direct evidence for stacking interactions which drive the self-assembly of guanine monomers. DFT calculations support spectral assignment and result in good agreement with the experimental evidence. The dynamics of amino groups and chlorine counterions lead to averaging of the NMR observables that can be rationalized by means of single-point energy (SPE) calculations performed at the semiempirical level.

Results and discussion

The structure of a single layer of a guanine quartet, shown in Fig. 1(a) and (b), was optimized with DFT methods by means of the B3LYP-D functional^{78–80} so as to take long-range dispersion forces into account. Subsequently, the three-layer stack of Fig. 1(c) and (d) was optimized at the same level of theory and utilized for the calculation of NMR parameters with the GIAO method.^{81,82} As these computations were carried out in vacuum, where enantiomeric structures yield equivalent NMR parameters, the handedness of the self-assembly was not considered in this study. The first and third layers have the sole purpose of mimicking the π – π stacking environment of the central layer, and the quadrupolar parameters of the outer layers have not been taken into account in our analysis. The contribution of the stacking interactions to the total energy of the assembly is expressed by the quantity $\Delta E_{\text{stacking}} = (E_{3\text{G}4}/3) - E_{\text{G}4} = 8.0$ kcal/mol obtained by our geometry optimizations at this level of theory, where $E_{\text{G}4}$ and $E_{3\text{G}4}$ are the energies of a single layer and of the entire three-layer stack, respectively. It is interesting to notice that the isolated guanine quartet seen in Fig. 1(b) shows a marked convexity with respect to the position of the K^+ cation. When the three-layer stack is formed (Fig. 1(c, d)), this convexity is greatly attenuated, resulting in a quasi-planar configuration of each layer, where the K^+ ions are sandwiched between the planes. Such a structure is in good agreement with those reported for G quartets⁸³ of nucleosides and nucleotides and for RNA⁸⁴ and DNA⁸⁵ sequences that contain G **quartets**.⁸⁶ The NMR parameters calculated for the central layer of Fig. 1(c, d) allow one to assign

the experimental (^1H , ^{14}N) correlation spectra. Chemical shift and quadrupolar parameters calculated by DFT methods for different NH environments in the central layer in Fig. 1(c, d) are given in Table 1. The NMR parameters of the two non-equivalent N(2)H₂ proton environments were averaged. This assumption is consistent with the experimental data presented in this study and is justified by fast rotation of the amino group about the C(2)–N(2) bond **in our experimental conditions. It is worth noting that considerable heating may occur in samples spinning at $\nu_R > 60$ kHz due to friction.⁸⁷ Therefore our data is also consistent with Adrian *et al.*,⁸⁸ who reported a coalescence temperature of ca. 315 K for proton peaks of guanine amino groups of G quadruplexes in solution.** As discussed below, the effect of counterions on the NMR parameters was also considered for the N(9)H environment. This was accomplished by including a single Cl[−] anion positioned along one of the N(9)–H vectors of the central layer. The position of the Cl[−] anion was optimized while keeping all other nuclear coordinates frozen. In this structure, only proton and ^{14}N parameters related to the N(9)H---Cl[−] environment were considered, i.e., the purpose of this structure is not to represent a physical system, but to estimate of the variations of NMR parameters due to counter ions.

The (^1H , ^{14}N) correlation spectrum of a polycrystalline sample of the G **quartets** obtained by DCP is shown in Fig. 2(a). A rapid transfer with a short cross-polarization contact time $\tau_c = 6 \tau_r = 96 \mu\text{s}$ ensures that only short-range one-bond (^1H , ^{14}N) correlations can be detected. The three expected protonated nitrogen sites are clearly identified, highlighted by dashed lines. Figure 2(b) shows an analogous spectrum where long-range (^1H , ^{14}N) correlations were detected by means of a longer contact time $\tau_c = 10 \tau_r = 160 \mu\text{s}$. An additional correlation signal between the C(8)H proton and the N(9)H nitrogen is clearly detected. The assignment of this correlation is confirmed by higher resolution F_1 (^{14}N) projections given in Supporting Information. Furthermore, the one-bond correlation signal due to the N(1)H site broadens towards higher frequency in the vertical F_1 (^{14}N) dimension, indicating a long-range contact with the N(2)H₂ site. Moreover, the N(1)H signal broadens towards lower frequency in the horizontal F_2 (^1H) dimension, indicating dipolar contacts with the N(9)H and N(2)H₂ sites. This may be appreciated by the two dashed red lines in 2(b) which are closer than those of 2(a). Figure 2(c) shows a simulated (^1H , ^{14}N) one-bond correlation

spectrum (to be compared with Fig. 2(a)), constructed by assuming the shift and quadrupolar parameters calculated by DFT methods for the central layer of Fig. 1(c, d). Analogously, Fig. 2(d) shows a simulated long-range correlation map for comparison with the experiment of 2(b), constructed by taking into account long-range proton-nitrogen distances between various proton-carrying sites within a cut-off distance of 4 Å, as shown in Fig. 3. Importantly, there is clear evidence for an *inter-plane* correlation (blue lines). The agreement between the experiments of Fig. 2(a,b) and (c,d) is generally good, with the exception of an inversion of the order of the proton shifts of C(8)H and N(2)H₂. The quadrupolar coupling constant of the N(1)H nitrogen environment is underestimated by ca. 10%. The correlation between experimental and calculated proton and ¹⁴N parameters is shown in Figs. 4(a) and (b). The white empty arrows indicate the improvements obtained by including the Cl⁻ anion in the structure considered in this study for data points associated with the N(9)H proton and ¹⁴N environments. In Fig. 4(b), the centers of gravity (first moments M_1) of the experimental ¹⁴N lineshapes were correlated with the isotropic shifts ν_{iso}^Q resulting from the second-order quadrupolar interaction calculated by DFT methods.^{77,89} The calculated frequencies for the N(9)H proton and ¹⁴N result from averaging over the two N(9)H and N(9)H---Cl⁻ environments given in Table 1. This averaging has been assumed as it significantly improves the correlation between experiments and calculations. Possible dynamic processes that lead to averaging of NMR observables – like fast chemical exchange of Cl⁻ anions – require further investigations. Solid-state NMR parameters calculated by DFT as in Fig. 4 must be considered with care when they do not take into consideration packing effects via periodic conditions. Nonetheless, effects of boundary environments on the NMR parameters can be mimicked by the inclusion of additional ionic or molecular species such as counterions or explicit solvent molecules, as testified by the improvement of the correlations of Fig. 4 when the Cl⁻ anion was considered.

The experimental evidence gathered in this study and the NMR parameters calculated by DFT methods suggest that extensive dynamics occur in G **quartets** at room temperature. More specifically, *i*) fast hopping of N(2)H₂ groups about their two-fold axis and *ii*) some mobility of the chorine counterions. In order to get insight into these processes, SPE calculations were carried

out at the semiempirical level. Figure 5(a) shows the trimeric G **quartet** assembly of Fig. 1 where an N(2)H₂ group in the central layer was progressively rotated about its two-fold axis in steps of 18°. Figure 5(b) shows the corresponding energy profile in red, with an activation energy of 61.2 kJ/mol. It is worth noting that the equilibrium conformation of the C(2)–N(2)H₂ moiety is not planar. This results in the observed asymmetry of the energy profile. If one proton of the amino group is allowed to relax by geometry optimization (while the other proton maintains the proper dihedral angle), one obtains a more symmetric profile indicated by the blue data points, with an activation energy reduced to 48.6 kJ/mol (blue arrow). These results are consistent with a fast rotation of the N(2)H₂ group at room temperature and with the averaging of the chemical shifts of the two protons observed experimentally.

Figure 5(c) shows a G quartet **analogous** to that of (a), where the chlorine ion was progressively displaced along a straight path in steps of 0.859 Å from its initial equilibrium position along the N(9)H vector ($d(\text{Cl}^{\text{---}}\text{HN}(9)) = 1.79 \text{ \AA}$) of the central quartet towards the closest analogous environment in the upper quartet. The profile resulting from single-point energy calculations is shown by a red line, with maximum activation energy of 51.4 kJ/mol (red arrow). An analogous profile where the position of the Cl[−] anions in the four highest-energy geometries was slightly displaced from the straight path is shown by a blue line. These structural adjustments (displacements always smaller than 0.04 Å) were required to *i*) re-establish the distance $d(\text{Cl}^{\text{---}}\text{HN}(9)) = 1.79 \text{ \AA}$ found in the equilibrium geometry and *ii*) decrease electrostatic repulsion between the chlorine anion and the lone pair of the N(3) site. Cartesian coordinates of the chlorine ion in these structures, an electrostatic potential isosurface and further details are given in Supporting Information. When these adjustments to the path of the chlorine ion are considered, the rate-determining highest activation energy of the considered motion is 20.4 kJ/mol (blue arrow). The energy profiles of Fig. 5(d) indicate that the chlorine atoms are highly mobile in G quartets at room temperature, thus supporting the averaging of DFT parameters between two Cl[−]–HN(9) and HN(9) configurations assumed in this study. **Further computational investigations of the influence of packing effects on quadrupolar parameters are given in Supporting Information.**

Conclusions

We have investigated the ^{14}N environments of guanine **quartet** self-assemblies in the solid state by means of ($^1\text{H},^{14}\text{N}$) correlation spectra obtained with the DCP method. Stacking between guanine tetramers is highlighted by the experimental observation of inter-plane dipolar contacts between a N(2)H₂ nitrogen in one plane and a N(9)H proton in a neighbouring plane. DFT calculations of shielding and quadrupolar parameters agree very well with the experimental spectra. Furthermore, the observed long-range ($^1\text{H},^{14}\text{N}$) correlations can be easily interpreted by considering internuclear ($^1\text{H},^{14}\text{N}$) distances up to ca. 4 Å, as predicted from the structure of the triple layer optimized by DFT. Since the ^{15}N isotopic enrichment of nucleic acids is cumbersome and expensive, our ($^1\text{H},^{14}\text{N}$) correlation **method could be applied as a structural tool for nucleic acids or synthetic G quartets. In addition, the microcrystalline sample preparation is straightforward especially when compared to the lengthy protocol required for the growth of crystals for X-ray crystallographic studies.** The averaging of NMR parameters is supported by a series of SPE calculations at the semiempirical level to model the rotation of the N(2)H₂ groups about their two-fold axis and the mobility of the chlorine counterions. **Synthetic G quartets have been suggested as novel biomaterials and this method could be used for their structural characterization. We are currently evaluating whether ($^1\text{H},^{14}\text{N}$) correlation spectra obtained with the DCP method could be used in biologically more relevant structures such as G quadruplexes formed by short oligonucleotides.**

Experimental details

Aqueous solutions of guanine (200 µL at 100 µM in H₂O) and KCl (200 µL at 5 mM in H₂O) were combined in an Eppendorf tube. The solvent was slowly evaporated over 48 hours by an air flow in a fume hood. All DCP MAS spectra were recorded on a wide-bore Bruker 800 spectrometer (18.8 T) with an Avance-III console, using 1.3 mm ZrO₂ rotors spinning at $\nu_R = 62.5$ kHz in a double-resonance MAS probe modified to apply ^{14}N pulses at 57.8 MHz. The proton *rf*-field strength was $\nu_1 = 179$ kHz (90° pulse width $\tau_p = 1.4$ µs.) The constant *rf*-field strengths during both $^1\text{H} \rightarrow ^{14}\text{N} \rightarrow ^1\text{H}$ cross polarization intervals were $\nu_1(^1\text{H}) = 53$ and $\nu_1(^{14}\text{N}) = 83$ kHz. The lengths of the contact

pulses were $\tau_{cp} = 160$ and $640 \mu\text{s}$ (10 and 40 rotor periods), for the spectra of Fig. 2(a) and 2(b), respectively. A recovery delay of 1 s was used in all cases. Four transients were averaged for each of the 64 t_1 increments, for an overall experimental time of 4 min 16 s. Phase cycling for coherence pathway selection was applied to the ^{14}N channel only, with phases $\phi_1 = \{0\ 2\}$, $\phi_2 = \{0\ 0\ 2\ 2\}$ for the first and second contact pulses, respectively, with a receiver phase $\phi_{\text{rec}} = \{0\ 2\ 2\ 0\}$. **A diagram of the pulse sequence is shown in Supporting Information.** The chemical shifts were referenced with respect to adamantane for ^1H and NH_4Cl for ^{14}N ; the magic angle was adjusted using KBr.

The G quartet structures of Fig. 1 were optimized *in vacuo* with the B3LYP-D functional^{78–80} as implemented in Gaussian09 Rev. A01.⁹⁰ The Pople basis sets⁹¹ utilized in this study were 6-31+G(d) for C and O, 6-31++G(2d,p) for H and N, whereas the pseudopotential LANL2DZ^{92–94} was chosen for K and Cl atoms. When chlorine anions were considered, a further optimization was performed where all atoms were fixed except for chlorine. The optimized geometries were subsequently utilized for the calculation of magnetic shielding and electric field gradient tensors with the GIAO method^{81,82} at the same level of theory. The SPE calculations of Fig. 5 were performed with the PM7⁹⁵ semiempirical method as implemented in MOPAC2016.⁹⁶ Calculations were performed on a Ubuntu Linux 12.04 workstation equipped with 16 CPUs running at 2.27 GHz and 24 GB of RAM.

The simulations of the NMR spectra based on DFT parameters were carried out with Simpson.⁹⁷ Powder averaging involved 320×8 crystal orientations sampled with the Repulsion scheme.⁹⁸ The time-domain signals were simulated by starting with an initial state $\rho(0) = I_x$ evolving under chemical shift and quadrupolar interactions. The offset of the carrier frequency ν_{rf} in these simulations matched the experimental conditions. The simulated 2D spectra of Fig. 2(c,d) represent pure adsorptive correlations between ideal ^{14}N and ^1H lineshapes, and results from the sum of individual 2D spectra associated with different ($^1\text{H}, ^{14}\text{N}$) distances obtained from the DFT-optimized structures.⁷⁷ Each of these individual 2D spectra was constructed by forming the product $\text{Re}[S(\omega_1)] \cdot \text{Re}[S(\omega_2)]$, where the column vector $S(\omega_1)$ is the spectrum for an ideal ^{14}N second-order quadrupolar lineshape simulated with Simpson assuming NMR parameters

calculated with DFT methods, and the row vector $S(\omega_2)$ is a simple Lorentzian lineshape simulated with Simpson, associated with a ^1H spin with an isotropic shift given by DFT methods. The anisotropy of the shift interaction was considered only for the ^{14}N spectra.

Acknowledgements

The authors thank Nicolas Birlirakis and Mathieu Baudin for assistance with experiments. This work was supported by the CNRS, the European Research Council (ERC, contract 'dilute para-water', grant agreement number 339754), and the French "Equipements d'Excellence" Paris-en-Resonance, contract ANR-10-EQPX-09.

References

- 1 P. A. Gale and J. W. Steed, in *Supramolecular Chemistry: From Molecules to Nanomaterials*, John Wiley & Sons, Ltd, 2012, p. 4014.
- 2 D. Philp and J. F. Stoddart, *Angew. Chemie Int. Ed. English*, 1996, **35**, 1154–1196.
- 3 J. M. Lehn, *Top. Curr. Chem.*, 2012, **322**, 1–32.
- 4 L. M. Greig, B. M. Kariuki, S. Habershon, N. Spencer, R. L. Johnston, K. D. M. Harris and D. Philp, *New J. Chem.*, 2002, **26**, 701–710.
- 5 G. R. Desiraju, *Angew. Chemie Int. Ed. English*, 1995, **34**, 2311–2327.
- 6 P. Belmont, J. F. Constant and M. Demeunynck, *Chem. Soc. Rev.*, 2001, **30**, 70–81.
- 7 H. Liu, R. Wang, X. Yu, F. Shen, W. Lan, P. Haruehanroengra, Q. Yao, J. Zhang, Y. Chen, S. Li, B. Wu, L. Zheng, J. Ma, J. Lin, C. Cao, J. Li, J. Sheng and J. Gan, *Nucleic Acids Res.*, 2018, **46**, 11627–11638.
- 8 X. Xie, O. Reznichenko, L. Chaput, P. Martin, M.-P. Teulade-Fichou and A. Granzhan, *Chem. - A Eur. J.*, 2018, **24**, 12638–12651.
- 9 J.-L. Mergny and D. Sen, *Chem. Rev.*, 2019, acs.chemrev.8b00629.
- 10 J. Zhou, S. Amrane, F. Rosu, G. F. Salgado, Y. Bian, H. Tateishi-Karimata, E. Largy, D. N. Korkut, A. Bourdoncle, D. Miyoshi, J. Zhang, H. Ju, W. Wang, N. Sugimoto, V. Gabelica and J.-L. Mergny, *J. Am. Chem. Soc.*, 2017, **139**, 7768–7779.
- 11 K. L. Chan, B. Peng, M. I. Umar, C. Y. Chan, A. B. Sahakyan, M. T. N. Le and C. K. Kwok, *Chem. Commun.*, 2018, **54**, 10878–10881.
- 12 A. Kerkour, J. Marquevielle, S. Ivashchenko, L. A. Yatsunyk, J. L. Mergny and G. F. Salgado, *J. Biol. Chem.*, 2017, **292**, 8082–8091.
- 13 G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876–880.
- 14 J. Brcic and J. Plavec, *Nucleic Acids Res.*, 2018, **46**, 11605–11617.
- 15 A. Maity, F. R. Winnerdy, W. D. Chang, G. Chen and A. T. Phan, *Nucleic Acids Res.*, 2020, 1–13.
- 16 M. Kasper, M. Gerlach, A. F. L. Schneider, C. Groneberg, P. Ochtrup, S. Boldt, D. Schumacher, J. Helma, H.

- Leonhardt, M. Christmann and C. P. R. Hackenberger, *ChemBioChem*, 2020, **21**, 113–119.
- 17 Y. L. Wu, N. E. Horwitz, K. S. Chen, D. A. Gomez-Gualdrón, N. S. Luu, L. Ma, T. C. Wang, M. C. Hersam, J. T. Hupp, O. K. Farha, R. Q. Snurr and M. R. Wasielewski, *Nat. Chem.*, 2017, **9**, 466–472.
- 18 A. Ghossoub and J.-M. Lehn, *Chem. Commun.*, 2005, 5763.
- 19 L. Stefan and D. Monchaud, *Nat. Rev. Chem.*, 2019, **3**, 650–668.
- 20 P. Murat, Y. Singh and E. Defrancq, *Chem. Soc. Rev.*, 2011, **40**, 5293–5307.
- 21 P. Röthlisberger and M. Hollenstein, *Adv. Drug Deliv. Rev.*, 2018, **134**, 3–21.
- 22 L. Lin, J. Sheng and Z. Huang, *Chem. Soc. Rev.*, 2011, **40**, 4591–4602.
- 23 S. Brown, M. M. Lockart, C. S. Thomas, M. K. Bowman, S. A. Woski and J. B. Vincent, *ChemBioChem*, 2019, **20**, 1–5.
- 24 A. Rangadurai, E. S. Szymaski, I. J. Kimsey, H. Shi and H. M. Al-Hashimi, *Prog. Nucl. Magn. Reson. Spectrosc.*, 2019, 112–113, 55–102.
- 25 M. Marušič, J. Schlagnitweit and K. Petzold, *ChemBioChem*, ,
DOI:10.1002/CBIC.201900072@10.1002/(ISSN)1439-7633.CHEMBIOTALENTS.
- 26 K. Lu, Y. Miyazaki and M. F. Summers, *J. Biomol. NMR*, 2010, **46**, 113–125.
- 27 M. P. Latham, D. J. Brown, S. A. McCallum and A. Pardi, *ChemBioChem*, 2005, **6**, 1492–1505.
- 28 F. Wachowius and C. Höbartner, *ChemBioChem*, 2010, **11**, 469–480.
- 29 S. Neuner, T. Santner, C. Kreutz and R. Micura, *Chem. - A Eur. J.*, 2015, **21**, 11634–11643.
- 30 C. Wan, W. Fu, H. Jing and N. Zhang, *Nucleic Acids Res.*, 2019, **47**, 1544–1556.
- 31 A. Marchanka, B. Simon, G. Althoff-Ospelt and T. Carlomagno, *Nat. Commun.*, 2015, **6**, 7024.
- 32 A. Dallmann, A. V. Beribisky, F. Gnerlich, M. Rübhelke, S. Schiesser, T. Carell and M. Sattler, *Chem. - A Eur. J.*, 2016, **22**, 15350–15359.
- 33 J. Kremser, E. Strebiter, R. Plangger, M. A. Juen, F. Nußbaumer, H. Glasner, K. Breuker and C. Kreutz, *Chem. Commun.*, 2017, **53**, 12938–12941.
- 34 F. Nußbaumer, M. A. Juen, C. Gasser, J. Kremser, T. Müller, M. Tollinger and C. Kreutz, *Nucleic Acids Res.*, 2017, **45**, 9178–9192.
- 35 Y. Liu, E. Holmstrom, J. Zhang, P. Yu, J. Wang, M. A. Dyba, D. Chen, J. Ying, S. Lockett, D. J. Nesbitt, A. R. Ferré-D’Amaré, R. Sousa, J. R. Stagno and Y. X. Wang, *Nature*, 2015, **522**, 368–372.
- 36 M. Flamme, L. K. McKenzie, I. Sarac and M. Hollenstein, *Methods*, 2019, **161**, 64–82.
- 37 M. J. Duer, Ed., *Solid-State NMR Spectroscopy Principles and Applications*, Blackwell Science Ltd, Oxford, UK, 2001.
- 38 W. Y. Chow, R. Rajan, K. H. Muller, D. G. Reid, J. N. Skepper, W. C. Wong, R. A. Brooks, M. Green, D. Bihan, R. W. Farndale, D. A. Slatter, C. M. Shanahan and M. J. Duer, *Science (80-.)*, 2014, **344**, 742–746.
- 39 R. K. (Robin K. Harris, R. E. Wasylshen and M. J. Duer, *NMR crystallography*, Wiley, 2009.
- 40 S. E. Ashbrook and D. McKay, *Chem. Commun.*, 2016, **52**, 7186–7204.
- 41 D. Carnevale, S. E. Ashbrook and G. Bodenhausen, *RSC Adv.*, , DOI:10.1039/c4ra09992c.
- 42 M. R. Mitchell, D. Carnevale, R. Orr, K. R. Whittle and S. E. Ashbrook, *J. Phys. Chem. C*, ,
DOI:10.1021/jp2105133.
- 43 D. Carnevale, P. Pelupessy and G. Bodenhausen, *J. Phys. Chem. Lett.*, , DOI:10.1021/acs.jpcclett.9b00914.

- 44 D. Carnevale, V. Del Amo, D. Philp and S. E. Ashbrook, *Tetrahedron*, , DOI:10.1016/j.tet.2010.05.065.
- 45 E. Banach, C. Invernizzi, M. Baudin, R. Neier and D. Carnevale, *Phys. Chem. Chem. Phys.*, , DOI:10.1039/c6cp05598b.
- 46 D. Appavoo, N. Raja, R. Deschenaux, B. Therrien and D. Carnevale, *Dalt. Trans.*, , DOI:10.1039/c5dt04179a.
- 47 Z. Gan, *J. Am. Chem. Soc.*, 2006, **128**, 6040–6041.
- 48 S. Cavadini, A. Lupulescu, S. Antonijevic and G. Bodenhausen, *J. Am. Chem. Soc.*, 2006, **128**, 7706–7707.
- 49 L. A. O’Dell, R. W. Schurko, K. J. Harris, J. Autschbach and C. I. Ratcliffe, *J. Am. Chem. Soc.*, 2011, **133**, 527–546.
- 50 V. Vitzthum, M. A. Caporini, S. Ulzega, J. Trébosc, O. Lafon, J.-P. Amoureux and G. Bodenhausen, *J. Magn. Reson.*, 2012, **223**, 228–236.
- 51 J. A. Jarvis, I. M. Haies, P. T. F. Williamson and M. Carravetta, *Phys. Chem. Chem. Phys.*, 2013, **15**, 7613–7620.
- 52 S. L. Veinberg, Z. W. Friedl, K. J. Harris, L. A. O’Dell and R. W. Schurko, *CrystEngComm*, 2015, **17**, 5225–5236.
- 53 G. N. M. Reddy, M. Malon, A. Marsh, Y. Nishiyama and S. P. Brown, *Anal. Chem.*, 2016, **88**, 11412–11419.
- 54 Y. Hong, T. Asakura and Y. Nishiyama, *ChemPhysChem*, 2018, **19**, 1841–1845.
- 55 J. A. Jarvis, M. Concistre, I. M. Haies, R. W. Bounds, I. Kuprov, M. Carravetta and P. T. F. Williamson, *Phys. Chem. Chem. Phys.*, 2019, **21**, 5941–5949.
- 56 S. Cavadini, S. Antonijevic, A. Lupulescu and G. Bodenhausen, *J. Magn. Reson.*, 2006, **182**, 168–172.
- 57 S. Cavadini, S. Antonijevic, A. Lupulescu and G. Bodenhausen, *ChemPhysChem*, 2007, **8**, 1363–1374.
- 58 Z. Gan, *J. Magn. Reson.*, 2007, **184**, 39–43.
- 59 S. Cavadini, A. Abraham, S. Ulzega and G. Bodenhausen, *J. Am. Chem. Soc.*, 2008, **130**, 10850–10851.
- 60 S. Cavadini, A. Abraham and G. Bodenhausen, *J. Magn. Reson.*, 2008, **190**, 160–164.
- 61 S. Cavadini, V. Vitzthum, S. Ulzega, A. Abraham and G. Bodenhausen, *J. Magn. Reson.*, 2010, **202**, 57–63.
- 62 V. Vitzthum, M. A. Caporini and G. Bodenhausen, *J. Magn. Reson.*, 2010, **205**, 177–179.
- 63 V. Vitzthum, M. A. Caporini, S. Ulzega and G. Bodenhausen, *J. Magn. Reson.*, 2011, **212**, 234–239.
- 64 R. Tycko, P. L. Stewart and S. J. Opella, *J. Am. Chem. Soc.*, 1986, **108**, 5419–5425.
- 65 R. Tycko and S. J. Opella, *J. Am. Chem. Soc.*, 1986, **108**, 3531–3532.
- 66 R. Tycko and S. J. Opella, *J. Chem. Phys.*, 1987, **86**, 1761–1774.
- 67 L. A. O’Dell and C. I. Ratcliffe, *Chem. Phys. Lett.*, 2011, **514**, 168–173.
- 68 L. A. O’Dell, R. He and J. Pandohee, *CrystEngComm*, 2013, **15**, 8657–8667.
- 69 L. A. O’Dell and A. Brinkmann, *J. Chem. Phys.*, , DOI:10.1063/1.4775592.
- 70 Y. Nishiyama, M. Malon, Z. Gan, Y. Endo and T. Nemoto, *J. Magn. Reson.*, 2013, **230**, 160–164.
- 71 A. J. Rossini, L. Emsley and L. A. O’Dell, *Phys. Chem. Chem. Phys.*, 2014, **16**, 12890–12899.
- 72 I. M. Haies, J. A. Jarvis, H. Bentley, I. Heinmaa, I. Kuprov, P. T. F. Williamson and M. Carravetta, *Phys. Chem. Chem. Phys.*, 2015, **17**, 6577–6587.
- 73 J. A. Jarvis, I. Haies, M. Lelli, A. J. Rossini, I. Kuprov, M. Carravetta and P. T. F. Williamson, *Chem. Commun.*, 2017, **53**, 12116–12119.

- 74 G. N. M. Reddy, A. Marsh, J. T. Davis, S. Masiero and S. P. Brown, *Cryst. Growth Des.*, 2015, **15**, 5945–5954.
- 75 A. L. Webber, S. Masiero, S. Pieraccini, J. C. Burley, A. S. Tatton, D. Iuga, T. N. Pham, G. P. Spada and S. P. Brown, *J. Am. Chem. Soc.*, 2011, **133**, 19777–19795.
- 76 D. Carnevale, X. Ji and G. Bodenhausen, *J. Chem. Phys.*, 2017, **147**, 184201.
- 77 D. Carnevale, B. Grosjean and G. Bodenhausen, *Commun. Chem.*, 2018, **1**, 1–9.
- 78 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- 79 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785–789.
- 80 S. Grimme, *J. Comput. Chem.*, 2006, **27**, 1787–1799.
- 81 D. Zeroka and H. F. Hameka, *J. Chem. Phys.*, 1966, **45**, 300–311.
- 82 R. Ditchfield, *J. Chem. Phys.*, 1972, **56**, 5688–5691.
- 83 S. P. Brown, *Curr. Opin. Colloid Interface Sci.*, 2018, **33**, 86–98.
- 84 N. Vasilyev, A. Polonskaia, J. C. Darnell, R. B. Darnell, D. J. Patel and A. Serganov, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E5391–E5400.
- 85 S.-T. D. Hsu, P. Varnai, A. Bugaut, A. P. Reszka, S. Neidle and S. Balasubramanian, *J. Am. Chem. Soc.*, 2009, **131**, 13399–13409.
- 86 A. T. Phan, *FEBS J.*, 2010, **277**, 1107–1117.
- 87 D. Zhang, B. Itin and A. E. McDermott, *J. Magn. Reson.*, 2019, **308**, 106574.
- 88 M. Adrian, F. R. Winnerdy, B. Heddi and A. T. Phan, *Biophys. J.*, 2017, **113**, 775–784.
- 89 A. Samoson, *Chem. Phys. Lett.*, 1985, **119**, 29–32.
- 90 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. Montgomery, J. A., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian, Inc., Wallingford CT*, 2009.
- 91 R. Ditchfield, W. J. Hehre and J. A. Pople, *J. Chem. Phys.*, 1971, **54**, 724–728.
- 92 P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 270–283.
- 93 W. R. Wadt and P. J. Hay, *J. Chem. Phys.*, 1985, **82**, 284–298.
- 94 P. J. Hay and W. R. Wadt, *J. Chem. Phys.*, 1985, **82**, 299–310.
- 95 J. J. P. Stewart, *J. Mol. Model.*, 2013, **19**, 1–32.
- 96 Stewart Computational Chemistry - MOPAC Home Page, <http://openmopac.net/>, (accessed 10 December 2019).
- 97 M. Bak, J. T. Rasmussen and N. C. Nielsen, *J. Magn. Reson.*, 2000, **147**, 296–330.
- 98 M. Bak and N. C. Nielsen, *J. Magn. Reson.*, 1997, **125**, 132–139.

Table 1 NMR parameters of all ^{14}N and ^1H sites in the central layer of the trimeric G quartet columnar assembly of Fig. 1(c,d) calculated by DFT methods. The isotropic chemical shieldings were referenced to 200 ppm for ^{14}N and 33 ppm for ^1H nuclei, according to $\delta_{iso}^{(i)} = \sigma_{ref} - \sigma_{iso}^{(i)}$ for the i -th nucleus.

Residue	^{14}N				^1H
	δ_{iso} (ppm)	$ C_Q $ (MHz)	η_Q	$\nu_{iso}^{(Q)}$ (ppm)	δ_{iso} (ppm)
N(1)H	91.3	2.96	0.43	300.6	15.3
N(2)H ₂	14.6	3.88	0.43	373.4	8.0
N(3)	120.8	2.88	0.95	363.0	-
N(7)	161.3	3.27	0.24	405.5	-
N(9)H	92.0	3.03	0.07	298.8	9.8
N(9)H---Cl ⁻	137.4	2.13	0.28	242.2	19.8
C(8)H	-	-	-	-	8.7

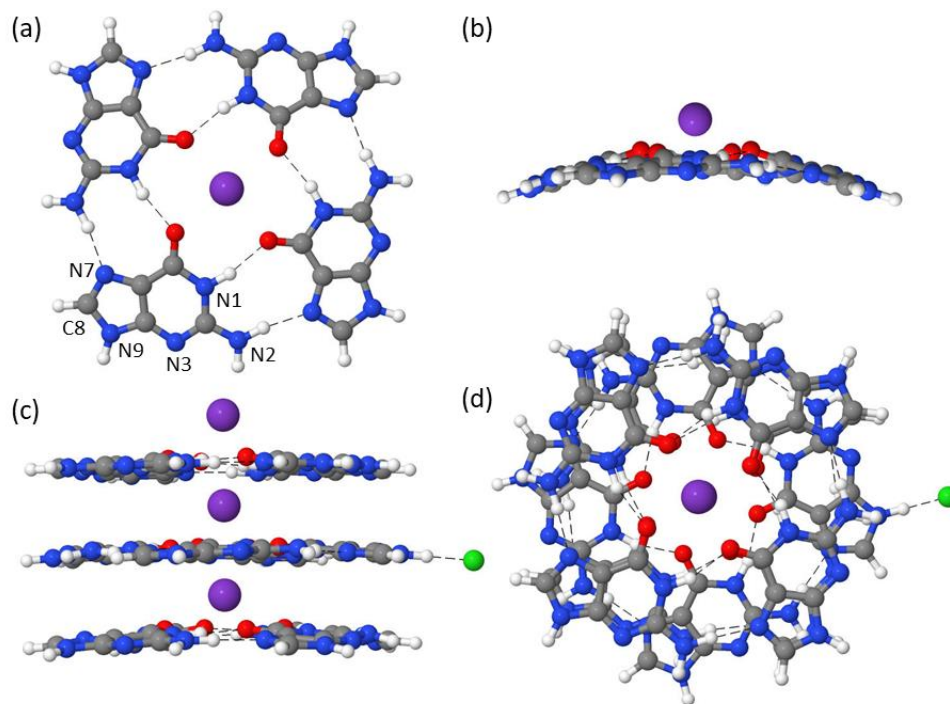


Figure 1 (a) Structure of a G quartet comprising four guanidine bases, optimized by DFT methods *in vacuo* as viewed down the stacking axis, showing numbering of all relevant nitrogen and carbon sites. (b) Side view of the structure of (a). (c) Stack of three layers of the same G quartet, again optimized by DFT. (d) Same structure as (c) viewed down the stacking axis. Hydrogen, carbon, oxygen, nitrogen, potassium and chlorine atoms are shown in white, grey, red, blue, purple and green, respectively. Hydrogen bonds are represented by dashed lines. The position of the chlorine anion was optimized at the semiempirical level.

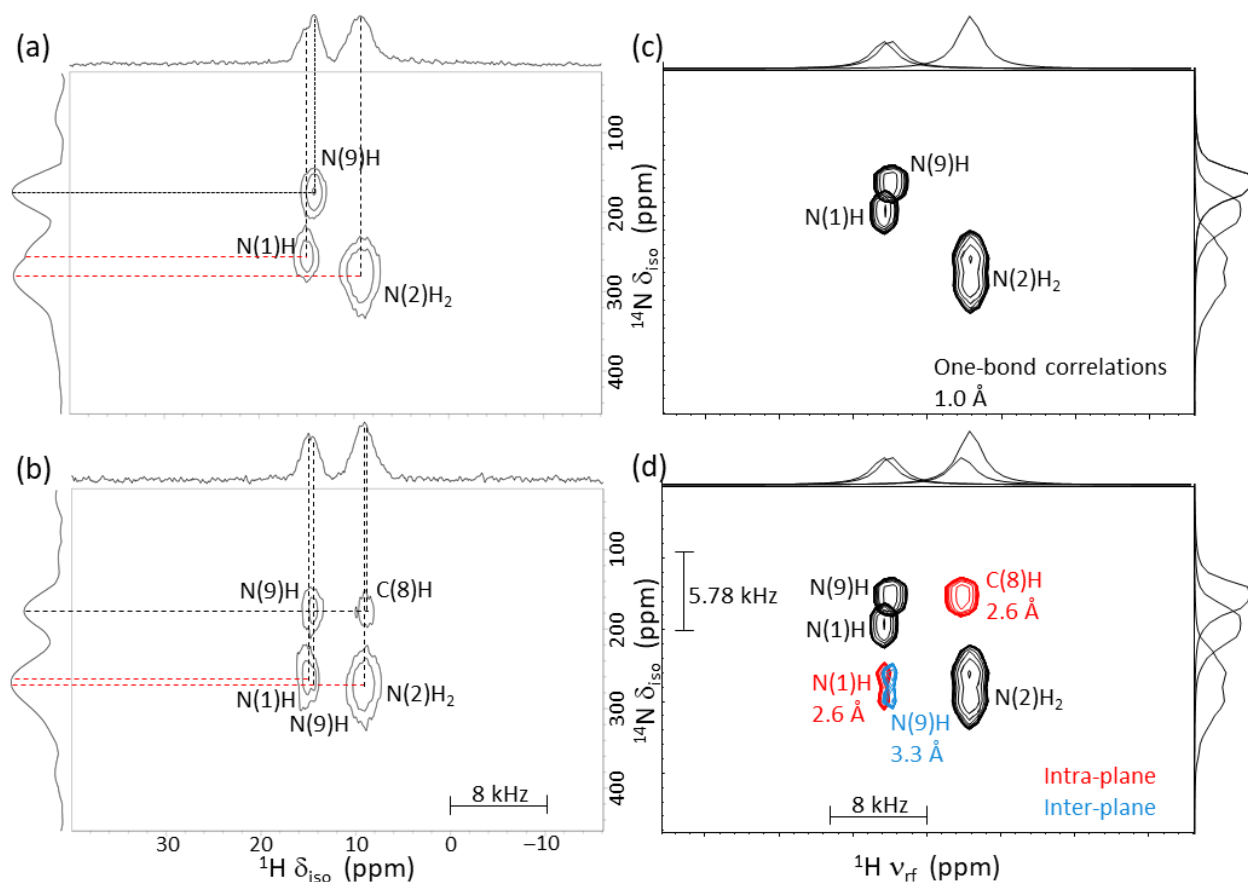


Figure 2 (a) Experimental (^1H , ^{14}N) correlation spectrum of a powdered sample of the G **quartet** crystallized in the presence of an excess KCl, spinning at 62.5 kHz about the magic angle (MAS), obtained with double cross-polarization (DCP) with two short rotor-synchronized cross-polarization contact intervals of 160 μs each. Only one-bond correlations are visible. (b) Analogous spectrum obtained with longer contacts of 640 μs , where long-range correlations can also be detected. (c-d) Spectra analogous to those of (a) and (b) simulated assuming NMR parameters calculated by DFT methods for short and long contacts, respectively. The overall experimental time for each of the two spectra (a) and (b) was 4 min 30 s.

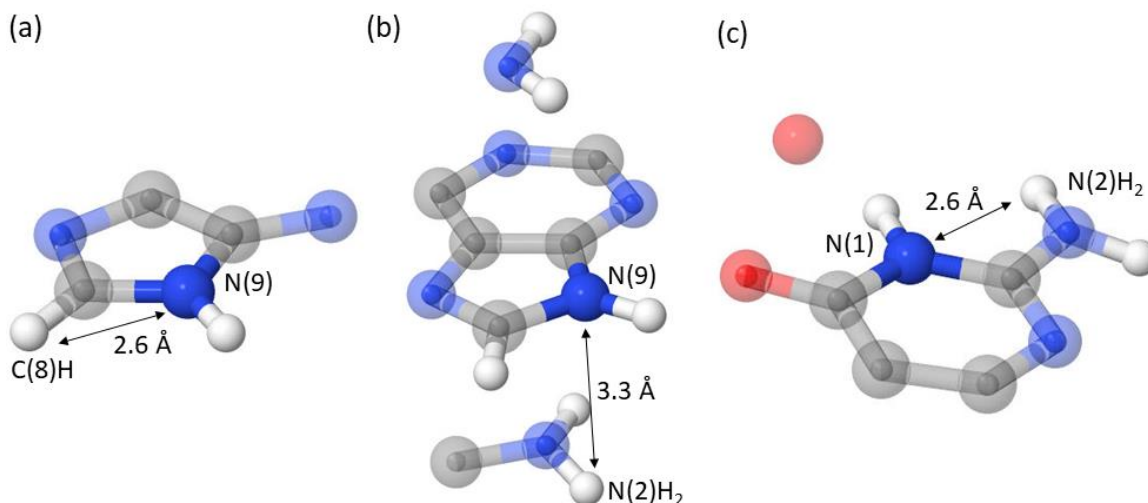


Figure 3 (a) The proton-carrying N(9)H nitrogen site showing the C(8)H at a (^1H , ^{14}N) distance of 2.6 Å. (b) Analogous picture as (a) but showing (^1H , ^{14}N) distances up to 4.0 Å, including an inter-plane distance of 3.3 Å. (c) The hydrogen-bonded proton-carrying N(1)H...O site shows two N(2)H₂ protons at a minimum (^1H , ^{14}N) distance of 2.6 Å. Color coding as in Fig. 1.

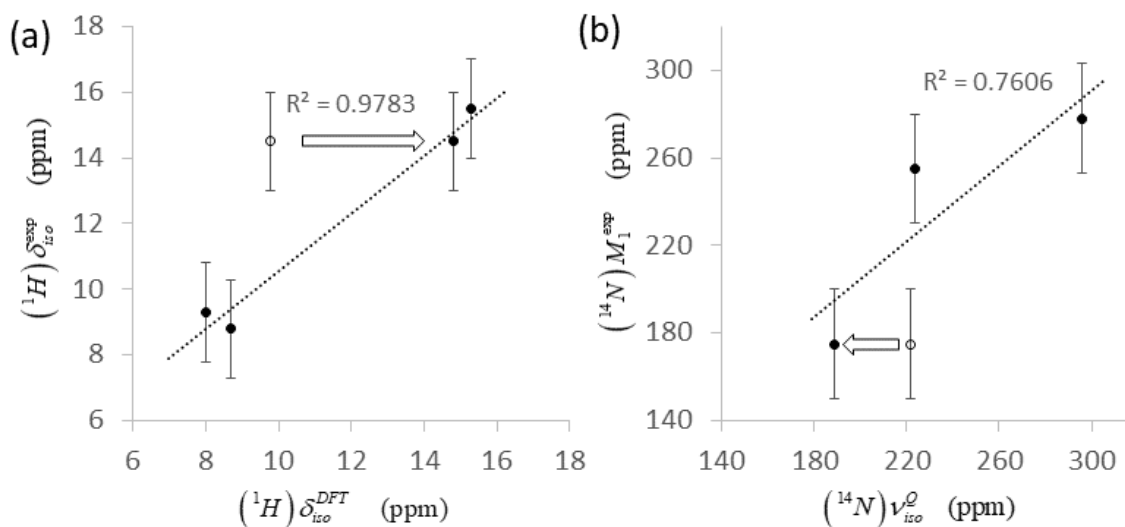


Figure 4 (a) Correlation between experimental and calculated isotropic ^1H shifts. (b) Correlation between experimental ^{14}N frequencies and calculated ^{14}N isotropic shifts ν_{iso}^Q resulting from the second-order quadrupolar interaction. Only the proton-carrying nitrogen environments of Table 1 were considered. The experimental data points refer to the center of gravity (first moment M_1) of the experimental lineshapes. In both cases, the arrows indicate the effect of the inclusion of the Cl^- anion along the N(9)H vector at 1.8 Å from the proton site (green atoms in Fig. 1 (c), (d).) The error bars correspond to the average experimental full widths at half height. The calculated frequencies for the N(9)H proton and ^{14}N result from averaging over the two N(9)H and N(9)H... Cl^- environments given in Table 1. The plotted values of ν_{iso}^Q were obtained by referencing the values in Table 1 so as to match the centers of gravity of the experimental lineshapes.

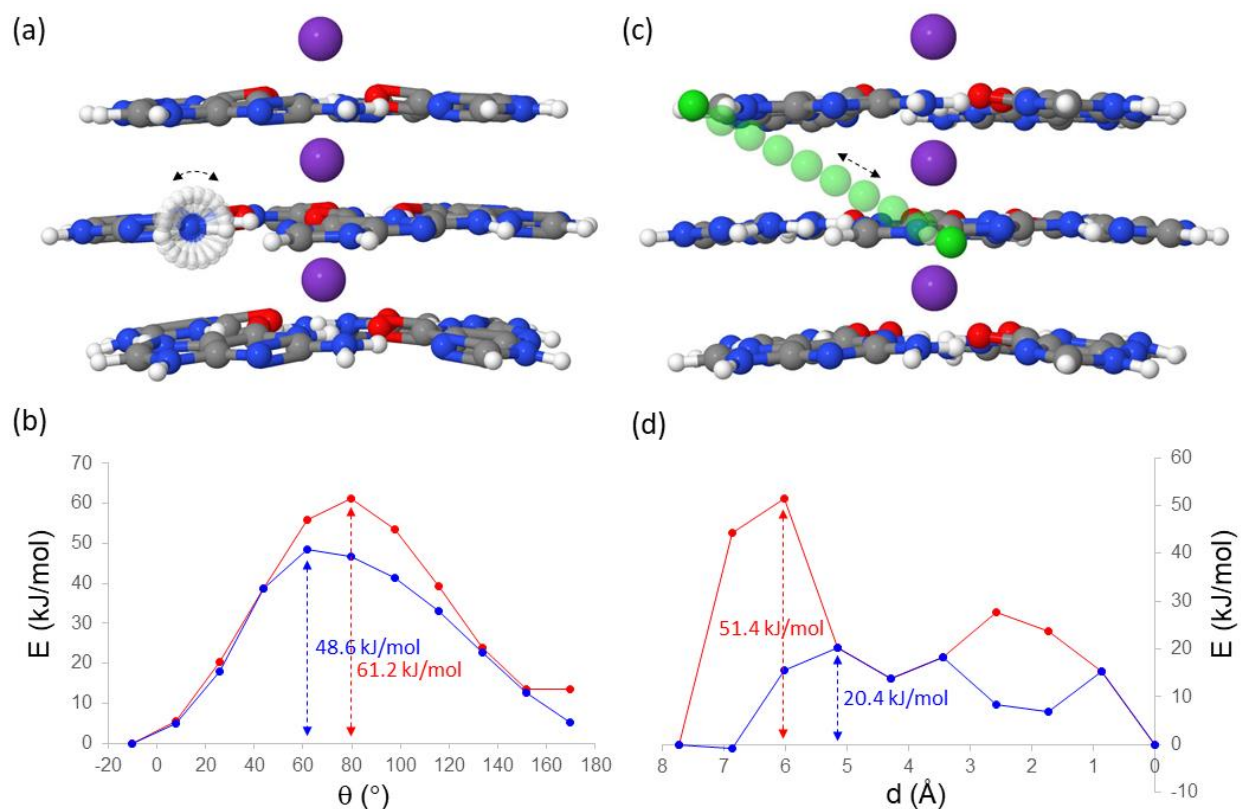


Figure 5 (a) Stack of three layers of G quartets analogous to Fig. 1 (c), where one N(2)H₂ of the central layer was progressively rotated about its two-fold axis in 10 steps of 18°. (b) Red dots: energy profile resulting from single point energy calculations at the semiempirical level of the structures of (a). In blue: profile derived from geometry optimization of one proton of the N(2)H₂ moiety, where the other proton maintains a suitable dihedral angle. (c) Stack of three layers of G quartets analogous to Fig. 1 (c), where the chlorine ion was progressively displaced in 10 steps of 0.859 Å from its initial position in the central quartet along a linear trajectory towards an analogous environment in the upper quartet. (d) Red dots: corresponding energy profile. An analogous profile where the position of the Cl⁻ anions in the four higher-energy geometries was slightly adjusted is shown by blue line. The Cartesian coordinates of the hypothetical path of the chlorine ions are given in Supporting Information.