OBC
www.rsc.org/obc

Structure and K^+ ion-dependent stability of a parallel-stranded DNA quadruplex containing a core A-tetrad

Mark S. Searle,*[a] Huw E. L. Williams,[a,b] Cathal T. Gallagher,[a] Richard J. Grant[a] and Malcolm F. G. Stevens[b]

[a] *School of Chemistry, Centre for Biomolecular Sciences, University Park, Nottingham, UK NG7 2RD*

[b] *School of Pharmaceutical Sciences, Centre for Biomolecular Sciences, University Park, Nottingham, UK NG7 2RD*

Received 12th November 2003, Accepted 30th January 2004
First published as an Advance Article on the web 18th February 2004

The c-MYC oncogene, containing the DNA 5′-GGAGG repeat sequence, is emerging as an attractive target for anti-cancer therapeutics with evidence to suggest that DNA quadruplex formation is involved in regulating gene transcription. We describe NMR studies and molecular dynamics (MD) simulations on intermolecular quadruplex structures containing the purine-rich GGAGG motif which we show readily assembles to form stable parallel-stranded quadruplexes containing A- and G-tetrads with evidence for weak sequence-specific K^+ ion binding at the 5′-GA step.

DNA quadruplex structures have been implicated in the regulation of a variety of processes within the cell cycle including replication, transcription and recombination.[1a] The observation that the telomeric repeats at the ends of chromosomes (5′-TTAGGG in humans) are able to assemble in the presence of monovalent cations into folded structures containing stacked G-tetrads has bought them into focus as a drug target, notably to interfere with telomere maintenance in immortalised human tumour-derived cell lines.[1] It is now apparent that the four stranded quadruplex structure is more versatile than the early G-rich sequences suggested, with examples in the literature of stable assemblies containing combinations of stacked G–G–G–G, G–C–G–C and A–T–A–T tetrads with the latter two formed through the major groove alignment of Watson-Crick G–C and A–T hydrogen bonded base pairs.[2] Others have also demonstrated that a T–T–T–T tetrad can be accommodated when sandwiched between adjacent G-tetrads, although when placed at the ends of G-stacks the T-tetrads are largely disordered.[3] The stability of A-tetrads also appears to be context-dependent.[4] The GGAGG repeat sequence is found within the 21-nucleotide nuclease hypersensitive element upstream of the c-MYC oncogene promoter region with evidence to suggest that quadruplex formation functions as a negative regulator of gene transcription.[5] c-MYC is emerging as an attractive target for anti-cancer therapeutics;[5] thus, the range of possible DNA conformations accessible in solution is of some interest in the design of sequence selective ligands. We describe NMR studies and molecular dynamics (MD) simulations on intermolecular quadruplex structures containing the purine-rich GGAGG motif. These sequences readily assemble to form stable parallel-stranded quadruplexes containing A- and G-tetrads, with evidence that K^+ ions can bind weakly in a sequence-specific manner to the A-tetrad.

1D and 2D NMR spectra of $d(TGGAGGC)_4$ (pH 7.0, 100 mM KCl, 10 mM phosphate) show that the structure has a four-fold symmetry with four well-resolved NH resonances consistent with the presence of four stable G-tetrads. We observe an unusually downfield shifted non-exchangeable resonance at 9.37 ppm that we assign to A^4H2, which gives strong NOEs to the imino proton of G^3 and G^5. The sequential NOE connectivity pathway along the DNA backbone (base H8 ↔ sugar H1′; see Fig. 1A) gives patterns and intensities of NOEs consistent with the adenine base stacking within the quadruplex with all purine glycosidic torsion angles in the *anti* range. Two possible geometries are envisaged in which the A-tetrad is stabilised through putative interstrand hydrogen bonding between the 6-NH_2 group and N7, or 6-NH_2 group and N1, with the latter shown in Fig. 1B. The latter arrangement brings the adenine H2 of one strand into close proximity to the N7 of an adjacent adenine, potentially accounting for the >1 ppm downfield shift observed for the A^4H2. Chemical shift calculations[6] based on various A-tetrad geometries, are in qualitative agreement with the 6-NH_2 to N1 hydrogen bonding

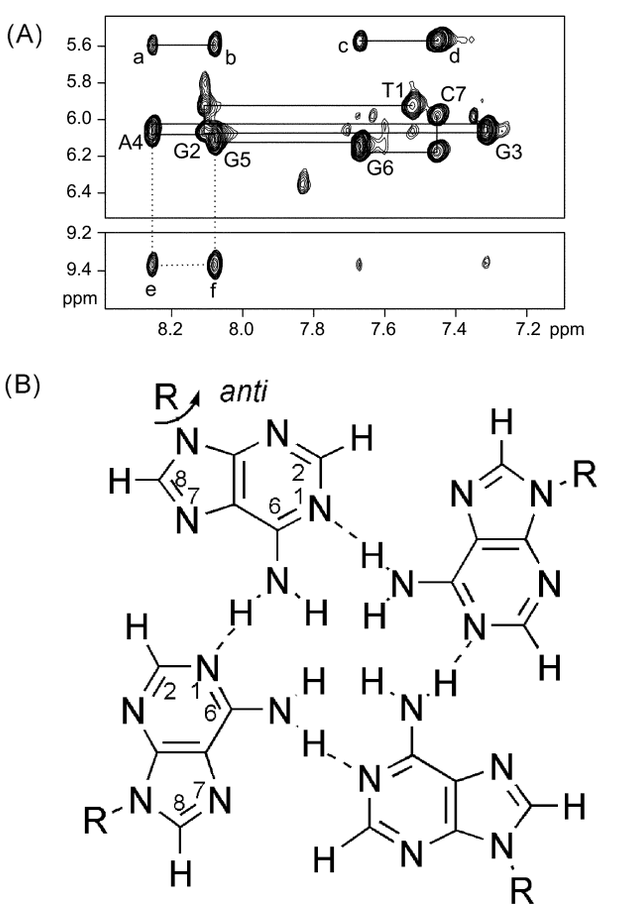

Fig. 1 (A) Portion of the 300 ms NOESY spectrum (600 MHz) of $d(T^1G^2G^3A^4G^5G^6C^7)_4$ at 298 K, pH 7.0, 100 mM KCl showing sequential connectivities *via* the base H6/H8 ↔ deoxyribose H1′ pathway, other NOEs are labelled a, G^5NH_2–A^4H8, b, G^5NH_2–G^5H8, c, C^7H6–G^6H8, d, C^7H5–H6, e, G^5NH_2–A^4H8, f, G^5NH_2–G^5H8; (B) Schematic showing 6-NH_2 to N1 hydrogen bonding in the A-tetrad.

DOI: 10.1039/b314559j

arrangement, which predicts a downfield shift of the A^4H2 of >0.5 ppm. Although two broadened resonances are resolved for each of the guanine 2-NH_2 groups within each G-tetrad, with a chemical shift difference between them of >3 ppm reflecting hydrogen bonding and non-hydrogen bonding environments, the 6-NH_2 of adenine can not be resolved unambiguously, suggesting that strong hydrogen bonding may not be a feature of A-tetrad stabilisation.

The imino proton region of the 1D spectrum of $d(TGGAGGC)_4$ shows a large chemical shift dispersion between 10.5 to 11.5 ppm with the resonances of G^5 and G^6 substantially upfield of those of G^2 and G^3. Particularly strong NOEs are observed between A^4H8-G^5NH_2 and A^4H8-G^5NH suggesting strong stacking interactions across the 5′-AG step (Fig. 1A). These interactions are reflected in the relative stabilities of the G-tetrads as evident from NH ⟶ ND exchange experiments were we observe the disappearance of signals in the 1D spectrum after redissolving the lyophilised sample in D_2O. Analogous experiments have reported the high kinetic stability of G-tetrads in quadruplex structures, and have identified bound cations between the G-tetrad planes.[7] Our studies show that exchange is extremely slow at 298 K but can be accelerated at elevated temperatures (>330 K). Under these conditions we see marked sequence-specific effects that show that the G^5N1-H resonance persists considerably longer than other guanine NH signals. The reverse experiment, in which the quadruplex has first been melted at high temperature and re-annealed in D_2O to trap N–D groups and then redissolved in H_2O solution, also shows that the G^5N1–H signal is very slow to reappear, indicating specific protection and stabilisation against ND/NH exchange. These observations are again consistent with particularly good base stacking interactions at the 5′-AG step.

The role of K^+ ion binding in A-tetrad stabilisation is at present unclear and was investigated by molecular dynamics simulations using the AMBER 6.0 package employing an explicit solvent model.[8] We first derived an NMR structural model of $d(TGGAGGC)_4$ on the basis of a large number of NOE-derived distance restraints in which K^+ ions were inserted only between G-tetrads (G·G–A–G·G) (Fig. 2). This structure gives a stable MD trajectory over at least 1 ns of simulation. We subsequently modeled the structure with K^+ ions inserted between each pair of tetrads (G·G·A·G·G). After careful equilibration of the solvent, the DNA co-ordinates were slowly released to undergo free dynamics at 300 K (see Experimental methods). Very rapidly (<2 ps) we observe that one K^+ ion is ejected from the structure with repeat simulations showing this to occur at either the 5′-AG or 5′-GA step without apparent discrimination. The resulting structures with one vacant site (G·G–A·G·G and G·G·A–G·G) are stable for at least 1 ns of unrestrained dynamics and are of comparable stability to the initial structure with both 5′-GA and 5′-AG sites vacant. Dissection of the electrostatic energy components shows that repulsions between K^+ ions bound between adjacent tetrads is large. However, when the tetrads consist of guanines (G·G·G), favourable interactions between K^+ ions and the G bases more than offset the repulsion between charges. This is not the case when K^+ ions are bound on either face of an A-tetrad (G·A·G). Relatively weak electrostatic interactions with the nitrogen lone pair on the adenine 6-NH_2 group do not appear to compensate for the repulsion between K^+ ions occupying adjacent sites. Analogous experiments and simulations with the homologous sequence $d(TGGGAGGC)_4$ result in identical observations and K^+ binding properties.

Surprisingly, altering the sequence of the unstructured 3′- and 5′-termini to d(TAGGAGGT) results in the partial stabilisation of a minor conformer (~20%) in which the guanine imino protons resonate over a much narrower chemical shift range (11.0–11.4 ppm; Fig. 3A and 3B). 2D NOESY spectra at 298 K (pH 7.0, 100 mM KCl) reveal strong chemical exchange cross-peaks between the two species which we estimate to be interconverting on a time scale of a few seconds (Fig. 3A). The major conformer shows the same pattern of protection towards NH/ND exchange seen for $d(TGGAGGC)_4$ (Fig. 3B). The observed time scale of interconversion between conformers demonstrated in Fig. 3A is fast enough to preclude a conformational equilibrium that involves disruption of A- or G-tetrads, since strand dissociation and re-annealing has been shown to occur over a period of hours at 298 K. On this basis we can exclude the possibility that the minor conformer arises from an anti-parallel quadruplex structure in equilibrium with the parallel-stranded form.

The MD simulations appear to preclude the possibility of K^+ ions binding simultaneously on both faces of the central A-tetrad since we are unable to generate a stable complex with this metal ion configuration. However, the data can be rational-

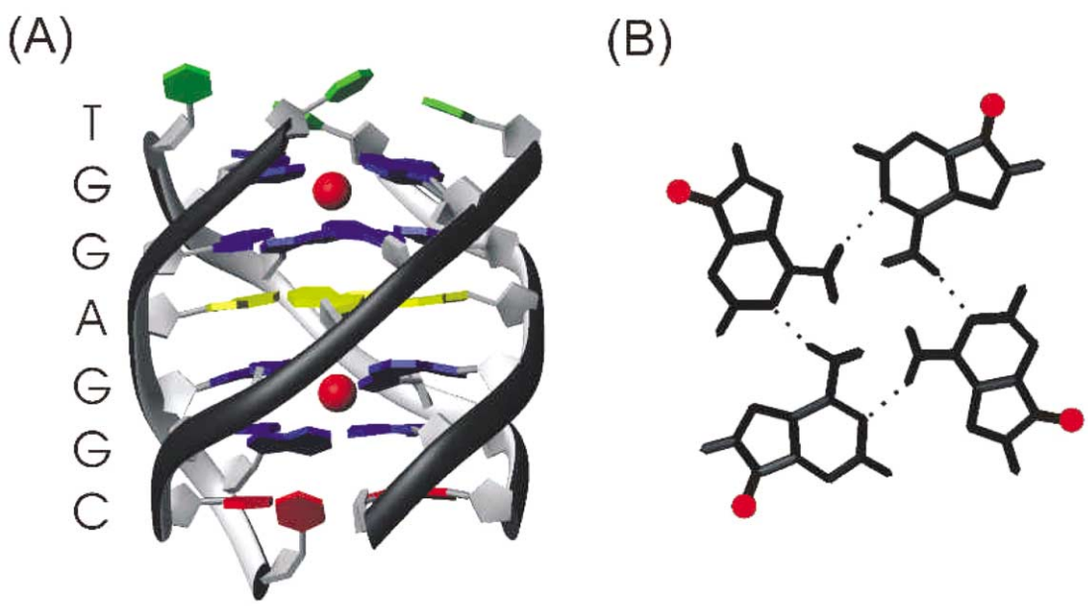

Fig. 2 (A) Mean structure of $d(TGGAGGC)_4$ after 1 ns of MD simulation with K^+ ions bound only at the 5′-GG steps; (B) conformation of the core A-tetrad in the mean structure with the position of the sugar residues indicated by circles. Both structures are displayed with MOLMOL (see ref. 9).

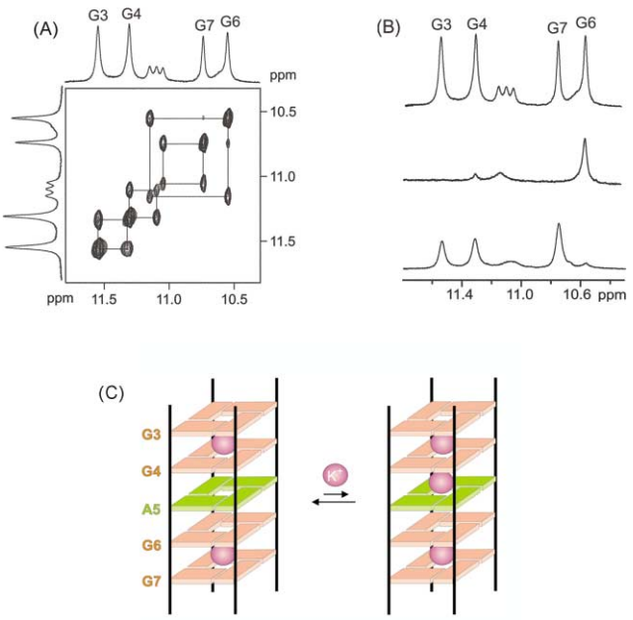

Fig. 3 (A) 600 MHz 2D NOESY spectrum (mixing time 300 ms) of $d(T^1A^2G^3G^4A^5G^6G^7T^8)_4$ in 90% H_2O (298 K, pH 7.0) showing chemical exchange cross-peaks between imino proton resonances 10–11.5 ppm in two conformations (major conformer labelled); (B) 1D ^1H NMR spectra of $d(TAGGAGGT)_4$ (298 K, pH 7.0) showing imino proton resonances (top), after NH ⟶ ND exchange (centre), and after ND ⟶ NH exchange (bottom); (C) schematic representation of the major (left) and minor (right) conformers with K^+ binding at the G4–A5 step.

ised in terms of a K^+ ion binding at either the 5′-GA or 5′-AG site. The G^4H8 corresponding to the guanine at the 5′-GA step is perturbed more significantly than any other base proton. Although the minor conformer is only weakly populated, the G^4H8 is clearly split into two resolved signals consistent with a K^+ ion binding with low affinity at the 5′-GA step (G·G·A–G·G), rather than at the 5′-AG step where the purine–purine stacking interactions appear to be stronger (see above). Based on the ratio of major and minor conformers observed by NMR (Fig. 3), we estimate a binding energy of only ~3 kJ mol^{-1}, although the kinetic barrier to exchange is clearly significant since the two conformers are in slow exchange. It is uncertain why a terminal 5′-TA *versus* 5′-T should produce such a subtle difference in K^+ binding affinity at the 5′-GA step resulting in the observation of two conformers for $d(TAGGAGGT)_4$ but only one for $d(TGGAGGC)_4$.

We have shown that A-tetrads are non-disruptive within G-rich quadruplex-forming sequences and are able to bind K^+ ions weakly in a sequence-specific manner preferring to intercalate at the 5′-GA step rather than the more stable 5′-AG step. This suggests the possibility that drug molecules may also be able to bind to these sequences by displacing the weakly bound cation. Previous studies have described the formation of A-tetrads stacked at the ends of runs of guanines, and have also demonstrated some sequence specific effects in dictating preferred A-tetrad geometries.[3] In these cases, the observation of an averaged signal for the 6-NH_2 group has suggested a dynamic A-tetrad weakly stabilised by hydrogen bonding interactions in one of two possible hydrogen bonding geometries.[3,4b] In neither case have large perturbations to NMR δ-values for AH2 resonances been reported. These observations are consistent with a dynamic A-tetrad that stacks relatively weakly on the ends of the G-rich quadruplex. Our chemical shift calculations suggest that the AH2 δ-value is highly distance and geometry-dependent and that dynamic averaging will largely negate any significant ring current effects. In this work, a downfield ring-current shift for the AH2 resonance (>1 ppm) appears to be diagnostic of a stabilised A-tetrad with a putative 6-NH_2 to N1 hydrogen bonding geometry. The absence of clearly resolved resonances for the 6-NH_2 group, analogous to those of the 2-NH_2 group of guanine, suggests that π-stacking interactions between A and G bases, rather than hydrogen bonding between adenines, may be largely responsible for A-tetrad stabilisation in this context.[9]

Acknowledgements

We thank the EPSRC and Cancer Research Campaign of the UK for financial support and the School of Chemistry for supporting NMR facilities.

References

1 (*a*) S. Neidle and G. Parkinson, *Nat. Rev., Drug Disc.*, 2002, **1**, 383; (*b*) R. T. Wheelhouse, D. Sun, H. Han, F. X. Han and L. H. Hurley, *J. Am. Chem.Soc.*, 1998, **120**, 3261; (*c*) J.-L. Mergny, P. Mailliet, F. Lavelle, J.-F. Riou, A. Laoui and C. Helene, *Anti-Cancer Drug Des.*, 1999, **14**, 327; (*d*) R. A. Heald, C. Modi, J. C. Cookson, I. Hutchinson, C. A. Laughton, S. M. Gowan, L. R. Kelland and M. F. G. Stevens, *J. Med. Chem.*, 2002, **45**, 590; (*e*) E. Gavathiotis, R. A. Heald, M. F. G. Stevens and M. S. Searle, *Angew. Chem.*, 2001, **40**, 4749; (*f*) E. Gavathiotis, R. A. Heald, M. F. G. Stevens and M. S. Searle, *J. Mol. Biol.*, 2003, **334**, 25; (*g*) S. M. Haider, G. N. Parkinson and S. Neidle, *J. Mol. Biol.*, 2003, **326**, 117; (*h*) J. A. Schouten, S. Ladame, S. J. Mason, M. A. Cooper and S. Balasubramanian, *J. Am. Chem. Soc.*, 2003, **125**, 5594.

2 (*a*) N. Zhang, A. Gorin, A. Majumdar, A. Kettani, N. Chernichenko, E. Skripkin and D. J. Patel, *J. Mol. Biol.*, 2001, **312**, 1073; (*b*) A. Kettani, A. Kumar and D. J. Patel, *J. Mol. Biol.*, 1995, **254**, 638.

3 P. K. Patel and R. V. Hosur, *Nucl. Acids. Res.*, 1999, **27**, 2457.

4 (*a*) P. K. Patel, A. S. R. Koti and R. V. Hosur, *Nucleic Acids Res.*, 1999, **27**, 3836; (*b*) E. Gavathiotis and M. S. Searle, *Org. Biomol. Chem.*, 2003, **1**, 1650; (*c*) A. I. M. Murchie and D. J. M. Lilley, *EMBO J.*, 1994, **13**, 993.

5 (*a*) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 11593; (*b*) A. Rangan, O. Y. Federoff and L. H. Hurley, *J. Biol. Chem.*, 2001, **276**, 4640.

6 D. Sitkoff and D. A. Case, *Prog. Nucl. Magn. Reson. Spect.*, 1998, **32**, 165.

7 (*a*) Y. Wang and D. J. Patel, *Biochemistry*, 1992, **31**, 8112; (*b*) F. W. Smith and J. Feigon, *Nature*, 1992, **356**, 164; (*c*) N. V. Hud, P. Schultz, V. Sklenar and J. Feigon, *J. Mol. Biol.*, 1999, **285**, 233–243.

8 (*a*) D. A. Case, D. A. Pearlman, J. W. Caldwell, Cheatham T. E. III, W. S. Ross, C. L. Simmerling, T. L. Darden, K. M. Marz, R. V. Stanton, A. L. Cheng, J. J. Vincent, M. Crowley, V. Tsui, R. J. Radmer, Y. Duan, J. Pitera, I. Massova, G. L. Seibel, U. C. Dingh, P. K. Weiner, P. A. Kollman, AMBER 6, 1999, University of California, San Francisco; (*b*) T. E. Cheatham, J. L. Miller, T. Fox, T. A. Darden and P. A. Kollman, *J. Am. Chem. Soc.*, 1995, **117**, 4193.

9 Oligonucleotides were synthesised using automated solid-phase methods and purified by reverse-phase HPLC; NMR data were collected on a Bruker Avance600 NMR spectrometer using standard phase-sensitive experiments.[1f,1g,4b] Derivation of NOE distance restraints and structure modeling protocols have been described previously,[1f,4b] using the AMBER 6.0 suite of programs[8] employing the AMBER 98 force field with modifications (T. E. Cheatham, P. Cieplak and P. A. Kollman, *J. Biomol. Struct. Dyn.*, 1999, **16**, 845–862) and Particle Mesh Ewald (PME) method for the treatment of long-range electrostatics (D. M. Yang, W. T. Lee, H. Darden and L. G. Pedersen, *J. Am. Chem. Soc.*, 1995, **117**, 5001–5002). The starting model for $d(TGGAGGC)_4$ was generated using the LEaP module within AMBER and was allowed to equilibrate fully before restrained molecular dynamics simulations were performed. NOE restraints (155 per strand, total 620) were introduced gradually in the form of square well potentials with a force constant of 20 kcal mol^{-1} $Å^{-2}$ over a period of 10 ps as the temperature was raised from 100–300 K. Restrained molecular dynamics continued for 1 ns with the whole set of NOE restraints active followed by energy minimisation. Snapshots from the final 100 ps were extracted every picosecond. In the final averaged energy minimised structure only 6 of the 155 NOEs per strand showed restraint violations >0.5 Å but <0.8 Å, all corresponded to very weak NOEs. The mean pairwise RMSD calculated over the final 100 snapshots was 0.78 (±0.1) Å over all heavy atoms. The ANAL module of AMBER 6.0[8] was used to analyse electrostatic interaction energies involving bound K^+ ions and the structures were displayed using MOLMOL (R. Koradi, M. Billeter and K. Wuthrich, *J. Mol. Graph.*, 1996, **14**, 51–55).