J|A|C|S COMMUNICATIONS

Published on Web 11/18/2010

Hydration Is a Major Determinant of the G-Quadruplex Stability and Conformation of the Human Telomere 3′ Sequence of d($AG_3(TTAG_3)_3$)

M. Clarke Miller, Robert Buscaglia, Jonathan B. Chaires, Andrew N. Lane, and John O. Trent*

Department of Medicine, Clinical Translational Research Building, University of Louisville, 505 Hancock Street, Louisville, Kentucky 40202, United States

Received June 16, 2010; E-mail: john.trent@louisville.edu

Abstract: The factors that determine the conformation and stability of G-quadruplex forming sequences remain poorly understood. Here we demonstrate the influence of cosolvents on the conformation and stability of the human telomeric sequence d(A(GGGTTA)$_3$GGG)) in both K^+ and Na^+ containing solutions using a combination of circular dichroism, NMR, and thermodynamics. Molecular crowding arguments have previously been used to suggest that the parallel quadruplex form may be biologically relevant. However, the small cosolvents previously used, PEG 200 and 400, are actually dehydrating agents. We have used acetonitrile as a non-hydrogen-bonding dehydrating agent; similar conformational transitions were observed in K^+ solution. Moreover, NMR analysis shows that the resulting structure contains non-anti guanine glycosyl torsion angles suggesting that the conformation present in acetonitrile is not identical to the all-parallel crystal structure, despite the supposed parallel type CD spectrum.

The factors that determine the conformation and stability of G-quadruplex forming sequences remain poorly understood.[1−4] Here we demonstrate the influence of cosolvents on the conformation and stability of the human telomeric sequence d($AG_3(TTAG_3)_3$) in both K^+ and Na^+ containing solutions using a combination of circular dichroism (CD), nuclear magnetic resonance (NMR), and thermodynamics.

In the presence of certain monovalent cations, guanine-rich DNA of the general type $(G_{2+X}N_n)_{4+Y}$ can form three-dimensional structures known as G-quadruplexes. These are made up of stacks of two or more square planar arrays of four Hoogsteen hydrogen-bonded guanines called G-quartets. G-quadruplexes are known to be stabilized by coordinating a cation such as Na^+ or K^+ to the O6 of the guanines.[4] Although most often associated with the ends of telomeres, potential quadruplex-forming sequences have been found throughout the genome.[5]

Quadruplex formation is intrinsically complex. When factors such as strand orientation, loop type and arrangement, and glycosyl torsion angles are considered, even the simple human telomeric sequence d$(GGGTTA)_n$ can theoretically fold into more than 200 unimolecular conformations.[2,6] Formation of tetramolecular and bimolecular structures of the G-quartets can lead to an even greater degree of polymorphism. Factors such as ion type, DNA concentration, the presence of other molecules such as organic solvents and various biological molecules, and even annealing protocols can determine what quadruplex conformations are formed.[2,3,7,8]

Here we report the influence of solvation on the stability and conformation of the human telomeric quadruplex d($AG_3(TTAG_3)_3$), which has strong implications for the state of this sequence in the conditions existing in the nucleus. Several structures based on this sequence have been reported, and the behavior, folding kinetics, and biophysical properties of this sequence have been extensively studied.[4,8−12] Recent studies concerning the role of small molecules in quadruplex stability have raised new questions about the properties of this sequence in solution particularly with regard to the stabilization/destabilization of the K^+ solution form and the species revealed by the crystal structure, which is not significantly populated in dilute aqueous solution.[3,4]

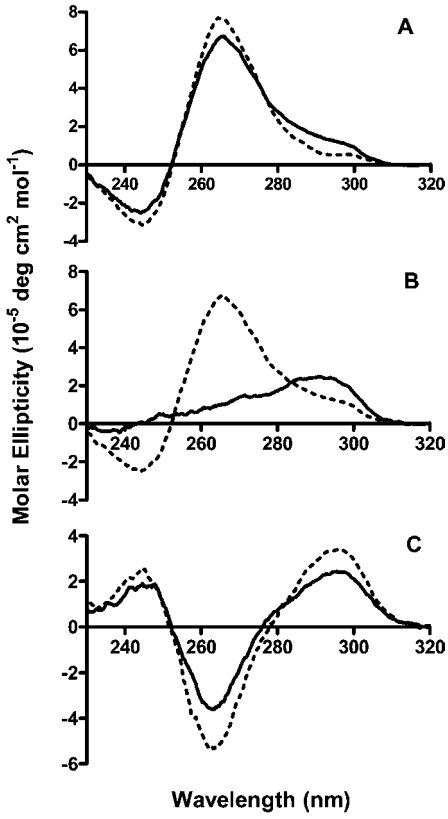

Figure 1. CD spectra of d($AG_3(TTAG_3)_3$). Spectra were recorded on a JASCO J-810 spectropolarimeter at 293 K in buffers containing 25 mM KCl or NaCl and 10 mM phosphate buffer, pH 7. (A) K^+ with 50% (v/v) acetonitrile (solid line), K^+ with PEG 400 (dashed line). (B) K^+ (solid line), K^+ with 50% (v/v) acetonitrile (dashed line). (C) Na^+ (solid line), Na^+ with 50% (v/v) acetonitrile (dashed line).

The crystallization procedure for the K^+ form used PEG 400 as a precipitating agent.[9] Since PEG 400 is inconvenient for NMR and has a low dielectric constant, we chose an alternative solvent, acetonitrile, because it is completely miscible with water, is an excellent general solvent, has a high dielectric constant and low viscosity, does not participate in hydrogen bonding, and is readily available in deuteriated form.

The CD of the oligonucleotide in the K^+ form changes dramatically in the presence of PEG or acetonitrile (Figure 1A and 1B), consistent with a change to a new conformation, that is very similar

 © 2010 American Chemical Society

for the two cosolvents. In contrast, the sodium form, which has a radically different CD and structure from those in K^+ solution, showed only small changes on adding CH_3CN (Figure 1C). This behavior was mirrored in the proton NMR spectra as shown in Figure 2. The NMR spectrum of the aqueous K^+ solution is the most complex (Figure 2A), which by integration comprises 23−24 iminoproton resonances. As the sequence contains 12 G residues, this is consistent with the presence of two major conformations. From the integration of resolved resonances, these are approximately equally populated. This finding agrees with previously reported NMR data for this sequence under these conditions.[13] In contrast, in aqueous Na^+ solutions the NMR spectrum is much simpler (Figure 2C) comprising 11−12 narrow GN1H resonances, consistent with a single quadruplex structure that differs from the mix of species present in K^+ solution.

Addition of CH_3CN to 50% v/v caused a large change in the NMR spectrum in K^+ solution, to a single species containing approximately 11 GN1H resonances (Figure 2B). In contrast, the spectral changes from 50% v/v CH_3CN in Na^+ solution were much smaller (Figure 2D). These small differences may be attributable to the change in bulk susceptibility/dielectric constant.

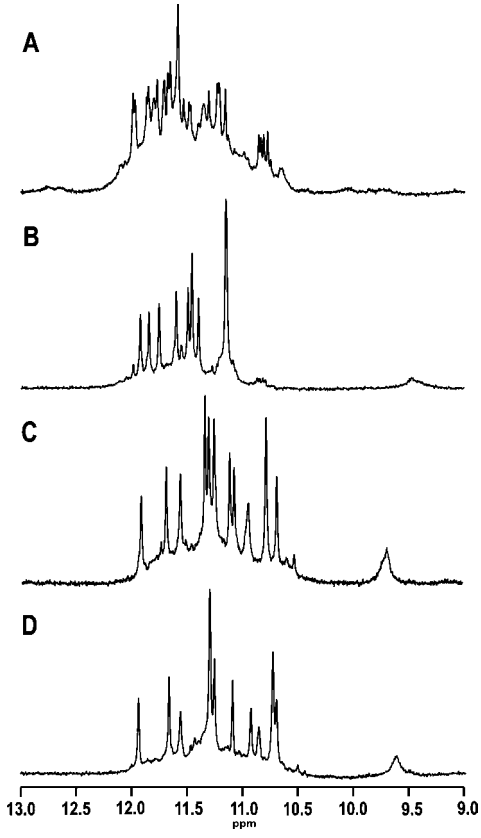

Figure 2. ^{1}H NMR of the imino protons of d($AG_3(TTAG_3)_3$) under different solution conditions. NMR spectra were recorded at 293 K and 800 MHz on a Varian Inova spectrometer. (A) K^+; (B) K^+ with 50% CD_3CN; (C) Na^+; (D) Na^+ with 50% CD_3CN.

The melting of the oligonucleotide measured by CD (Figure 3) shows that the mixture in the presence of potassium is more stable than the sodium form and that acetonitrile greatly increased the stability of both forms. At 298 K, the ΔG of folding of the sodium form increased from −3.74 to −7.1 kcal/mol. The stabilization of the K^+ form was substantially greater; the ΔG increased from −4.9 to −10.7 kcal/mol in 50% CH_3CN (Supporting Information, Table S1). This is associated with a change in conformation (see above).

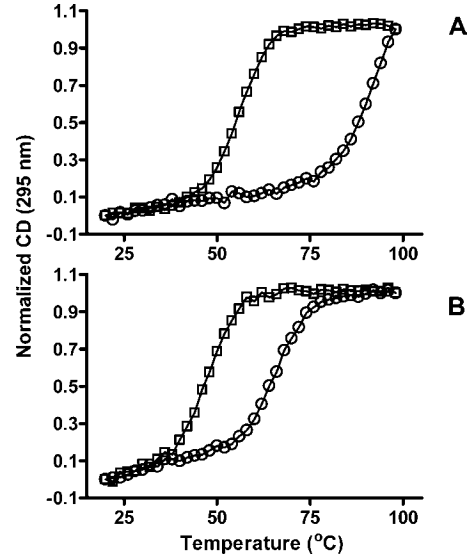

Figure 3. CD melting of d($AG_3(TTAG_3)_3$) under different conditions. Squares: buffer; circles: 50% acetonitrile. (A) K^+ in the absence and presence of 50% CH_3CN. Data were fitted to a two-state model including sloping baselines.[14] T_m = 372 K, ΔH = 53.7 kcal/mol (CH_3CN); T_m = 329 K, ΔH = 51.9 kcal/mol (buffer). (B) Na^+ in the absence and presence of 50% CH_3CN. T_m = 339 K, ΔH = 58.1 kcal/mol (CH_3CN); T_m = 321 K, ΔH = 51.8 kcal/mol (buffer). Solid lines are nonlinear regression fits.

The NMR line widths and cross-relaxation rate constants determined from NOESY and truncated 1D NOE build-up curves showed that the rotational correlation time of this species (<2 ns at 293 K) implies that the quadruplex is monomeric (and see Supporting Information). NOESY data showed that the K^+-form in 50% CH_3CN contains G-quartets and that some nucleotides have glycosyl torsion angles different from those of the high anti range but are also not pure syn (Supporting Information Figures S1, S2, S3, S4 and Table S2). The latter suggests that the conformation is not identical to the all-parallel form reported for the crystal structure.[9] Similarly, assignment of the parallel structures in cosolvents such as PEG-200 based solely based on CD spectra[8] is not definitive for structural classification[1,15] because CD reports on base stacking interactions, not directly on strand polarity.[16] Conventional interpretation of quadruplex CD spectra holds that parallel quadruplexes generally display a negative peak near 240 nm and a positive peak near 265 nm, and that antiparallel quadruplexes are characterized by a negative peak near 260 nm and a positive peak near 295 nm.[17−20] However, there are many exceptions to this rule[15,21,22] so no assignment to parallel or antiparallel strand orientation is really possible without supporting structural data.

We recently reported that CD can hide the underlying complexity and polymorphism of quadruplex-forming sequences.[1] There were at least eight resolvable species present in solution, which collectively showed a CD spectrum that would be classified as being "parallel" according to this convention, yet individual components have a wide variety of CD spectra[1] This illustrated the need for a high-resolution technique, such as NMR, to be used when investigating quadruplex topology.

There are three possible physical explanations for the effects of CD_3CN on stability and conformation: differential binding; effects on hydration; and steric crowding. Crowding effects require a sufficiently large difference in volume between the conformational states. Because the difference in excluded volumes among various folded isomers of the 22-mer are very small,[3] molecular crowding per se is unlikely to be important, and any steric crowding effect

will favor the most compact state.[23] As the all-parallel form present in the crystal state[9] is less compact than other known structures[3] it seems unlikely that this structure can be produced by crowding. Simulations of the effect of noninteracting macrosolutes on the CH_3CN-induced conformational transition observed in K^+ solutions show that indeed the effect on the equilibrium is very small and in the direction that destabilizes the all-parallel form (Supporting Information, Figure S5). A more plausible interpretation is that the species differ in their hydration properties, as has been proposed for the thrombin aptamer.[8,24] In duplex DNA, the B-form is stabilized by solvation;[25] decreasing the water activity causes a transition to the A-form and, ultimately, to the strand state. In contrast G-quadruplex structures appear to be destabilized by hydration; the thermodynamic stability increases as the water activity is reduced,[24,26] and the surface area accessible to water is minimized. The transition between two compact molecular forms implies that there is a difference in solvent accessible surface area and/or new sites of different affinity are exposed or hidden during the transition. The solvent accessible surface areas, as calculated by Naccess,[27] for the known structures of the human telomere, i.e., the basket form (143D.pdb[28] 3749 Å^2), the parallel form (1KF1.pdb[9] 4182 Å^2), hybrid1 (2JSM.pdb[4] 3814 Å^2, 2HY9.pdb[29] 3552 Å^2, 2GKU.pdb[30] 4066 Å^2), and hybrid2 (2JPZ.pdb[13] 3581 Å^2, 2JSL.pdb[4] 3843 Å^2) when truncated to the d(GGGTTA)$_3$GGG quadruplex, are different and do not support the parallel form being the most compact. In an additional measure of compactness, the radius of gyration, as calculated by Hydropro,[31] for the basket form (143D.pdb 1.17×10^{-7} cm), the parallel form (1KF1.pdb 1.36×10^{-7} cm), hybrid1 (2JSM.pdb 1.19×10^{-7} cm, 2HY9.pdb 1.16×10^{-7} cm, 2GKU.pdb 1.21×10^{-7} cm) and hybrid2 (2JPZ.pdb 1.15×10^{-7} cm, 2JSL.pdb 1.23×10^{-7} cm) does not support the parallel form being the most compact.

To assess further experimentally the role of steric crowding on the conformational transition in K^+ solution, we have recorded the NMR spectrum of the oligonucleotide in K^+ buffer in the presence of 300 mg/mL bovine serum albumin (BSA) ($M_w = 66\,382$). At neutral pH BSA is negatively charged[32] and thus is unlikely to bind significantly to the negatively charged quadruplex. The protein contains no tryptophan residues and has no NMR resonances in the 10−12 ppm region that could overlap those of the DNA imino protons. Although the solution viscosity increases approximately 4-fold, thereby broadening the NMR spectrum, the NMR spectrum is more similar to that of the K^+ form in water than that observed in 50% CD_3CN (Figure 2B and Supporting Information, Figure S6). Even at this high cosolute concentration, the molar concentration is only about 4.5 mM, and its effect on water activity is thus very small.[32,33] Therefore in this system hydration effects are not significant, and therefore high concentrations of BSA should influence equilibria only by steric crowding. As the spectral changes in the presence of BSA do not resemble the formation of the species present in 50% CD_3CN, that species is most likely stabilized by dehydration, and not by steric crowding. The steric crowding effect is also small, as expected for an isomerization between two or more similarly compact states (Supporting Information, Figure S6).

In conclusion acetonitrile is an excellent solvent for nucleic acids and is suited for detailed biophysical characterization of the solute. Compact G-quadruplexes appear to be stabilized by dehydration, which has relevance to intracellular conditions where the absolute water activity must be much lower than 55.5 M.[34]

Acknowledgment. This work was supported by the National Institutes of Health (CA113735-01 to J.O.T. and GM077422 to J.B.C.), National Institutes of Health Grant Number P20RR018733 from the National Center for Research Resources, and the Kentucky Challenge for Excellence.

Supporting Information Available: Supporting figures, tables, and methods. This information is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Dailey, M. M.; Miller, M. C.; Bates, P. J.; Lane, A. N.; Trent, J. O. *Nucleic Acids Res.* **2010**, *38*, 4877–4888.
(2) Lane, A. N.; Chaires, J. B.; Gray, R. D.; Trent, J. O. *Nucleic Acids Res.* **2008**, *36*, 5482–5515.
(3) Li, J.; Correia, J. J.; Wang, L.; Trent, J. O.; Chaires, J. B. *Nucleic Acids Res.* **2005**, *33*, 4649–4659.
(4) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2007**, *35*, 6517–6525.
(5) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908–2916.
(6) Webba da Silva, M. *Chemistry* **2007**, *13*, 9738–45.
(7) Vorlickova, M.; Chladkova, J.; Kejnovska, I.; Fialova, M.; Kypr, J. *Nucleic Acids Res.* **2005**, *33*, 5851–60.
(8) Xue, Y.; Kan, Z. Y.; Wang, Q.; Yao, Y.; Liu, J.; Hao, Y. H.; Tan, Z. *J. Am. Chem. Soc.* **2007**, *129*, 11185–11191.
(9) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876–880.
(10) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263–282.
(11) Gray, R. D.; Petraccone, L.; Trent, J. O.; Chaires, J. B. *Biochemistry* **2010**, *49*, 179–194.
(12) Gray, R. D.; Li, J.; Chaires, J. B. *J. Phys. Chem. B* **2009**, *113*, 2676–2683.
(13) Dai, J.; Carver, M.; Punchihewa, C.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 4927–40.
(14) Asensio, J. L.; Lane, A. N.; Dhesi, J.; Bergqvist, S.; Brown, T. *J. Mol. Biol.* **1998**, *275*, 811–822.
(15) Dapic, V.; Abdomerovic, V.; Marrington, R.; Peberdy, J.; Rodger, A.; Trent, J. O.; Bates, P. J. *Nucleic Acids Res.* **2003**, *31*, 2097–2107.
(16) Masiero, S.; Trotta, R.; Pieraccini, S.; De Tito, S.; Perone, R.; Randazzo, A.; Spada, G. P. *Org. Biomol. Chem.* **2010**, *8*, 2683–2692.
(17) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405–16415.
(18) Giraldo, R.; Suzuki, M.; Chapman, L.; Rhodes, D. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 7658–7662.
(19) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061–4067.
(20) Kypr, J.; Kejnovska, I.; Renciuk, D.; Vorlickova, M. *Nucleic Acids Res.* **2009**, *37*, 1713–25.
(21) Dapic, V.; Bates, P. J.; Trent, J. O.; Rodger, A.; Thomas, S. D.; Miller, D. M. *Biochemistry* **2002**, *41*, 3676–3685.
(22) Jing, N. J.; Rando, R. F.; Pommier, Y.; Hogan, M. E. *Biochemistry* **1997**, *36*, 12498–12505.
(23) Minton, A. P. *Biophys. J.* **2005**, *88*, 971–985.
(24) Miyoshi, D.; Karimata, H.; Sugimoto, N. *J. Am. Chem. Soc.* **2006**, *128*, 7957–7963.
(25) Spink, C. H.; Chaires, J. B. *Biochemistry* **1999**, *38*, 496–508.
(26) Olsen, C. M.; Gmeiner, W. H.; Marky, L. A. *J. Phys. Chem. B* **2006**, *110*, 6962–6969.
(27) Hubbard, S. J.; Thornton, J. M. *Department of Biochemistry and Molecular Biology*; University College: London, 1993.
(28) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263–82.
(29) Dai, J.; Punchihewa, C.; Ambrus, A.; Chen, D.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 2440–50.
(30) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963–70.
(31) Garcia De La Torre, J.; Huertas, M. L.; Carrasco, B. *Biophys. J.* **2000**, *78*, 719–30.
(32) Courtenay, E. S.; Capp, M. W.; Anderson, C. F.; Record, M. T., Jr. *Biochemistry* **2000**, *39*, 4455–4471.
(33) Kiyosawa, K. *Biophys. Chem.* **2003**, *104*, 171–188.
(34) Hancock, R. *Biol. Cell* **2004**, *96*, 595–601.

JA105259M