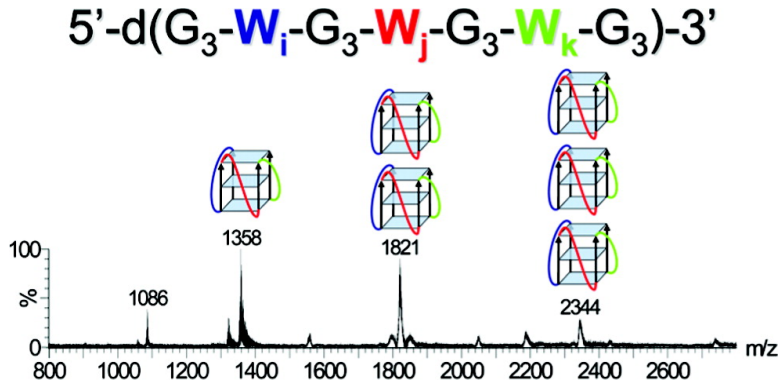
Article

G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation

Nicolas Smargiasso, Fre#de#ric Rosu, Wei Hsia, Pierre Colson, Erin Shammel Baker, Michael T. Bowers, Edwin De Pauw, and Vale#rie Gabelica

J. Am. Chem. Soc., **2008**, 130 (31), 10208-10216 • DOI: 10.1021/ja801535e • Publication Date (Web): 16 July 2008



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
ARTICLES

Published on Web 07/16/2008

G-Quadruplex DNA Assemblies: Loop Length, Cation Identity, and Multimer Formation†

Nicolas Smargiasso,‡ Frédéric Rosu,‡ Wei Hsia,§ Pierre Colson,§ Erin Shammel Baker,|| Michael T. Bowers,§ Edwin De Pauw,‡ and Valérie Gabelica*,‡

Mass Spectrometry Laboratory, GIGA-Research, University of Liège, Belgium, Biospectroscopy Laboratory, University of Liège, Belgium, and Department of Chemistry and Biochemistry, University of California at Santa Barbara, California

Received February 29, 2008; E-mail: v.gabelica@ulg.ac.be

Abstract: G-rich DNA sequences are able to fold into structures called G-quadruplexes. To obtain general trends in the influence of loop length on the structure and stability of G-quadruplex structures, we studied oligodeoxynucleotides with random bases in the loops. Sequences studied are $dGGGW_iGGGW_jGGGW_kGGG$, with W = thymine or adenine with equal probability, and *i*, *j*, and *k* comprised between 1 and 4. All were studied by circular dichroism, native gel electrophoresis, UV-monitored thermal denaturation, and electrospray mass spectrometry, in the presence of 150 mM potassium, sodium, or ammonium cations. Parallel conformations are favored by sequences with short loops, but we also found that sequences with short loops form very stable multimeric quadruplexes, even at low strand concentration. Mass spectrometry reveals the formation of dimers and trimers. When the loop length increases, preferred quadruplex conformations tend to be more intramolecular and antiparallel. The nature of the cation also has an influence on the adopted structures, with K^+ inducing more parallel multimers than NH_4^+ and Na^+. Structural possibilities are discussed for the new quadruplex higher-order assemblies.

Introduction

Guanine-rich DNA sequences are able to fold into G-quadruplex structures, composed of stacked guanine tetrads which are stabilized by Hoogsteen-type hydrogen bonds between the guanines and by interactions with cations located between the tetrads (Figure 1A).[1–4] G-quadruplexes are involved in different biological phenomena such as gene regulation and telomere maintenance.[5–18] Calculations of the number of sequences containing four tracks of three or more guanines separated by loops containing at least one base have shown that over 376000 G-quadruplexes could be potentially formed simultaneously in the human genome.[19,20] A total of 40% of genes contain putative G-quadruplex forming sequences in their promoter,[21]and these sequences seem correlated with gene functions: they are less present in tumor suppressor genes and more present in proto-oncogenes.[22] This finding supports the hypothesis that these structures play a role in gene regulation.

Several such putative G-quadruplex forming sequences spotted in the genome were isolated and subsequently studied in vitro, and indeed found to form stable intramolecular G-quadruplex structures, such as in c-Myc,[9,18,23–25] c-kit,[10,15] and BCL-2 oncogenes.[17,26] However, because of the huge number

† A preliminary version of our results was presented at the first International Meeting on Quadruplex DNA in April 2007 under the title "Influence of loop length on the structure and stability of intramolecular G-quadruplexes".

‡ Mass Spectrometry Laboratory, GIGA-Research, University of Liège.

§ Biospectroscopy Laboratory, University of Liège.

|| University of California at Santa Barbara.

(1) Williamson, J. R.; Raghuraman, M. K.; Cech, T. R. *Cell* **1989**, *59*, 871–880.
(2) Nagesh, N.; Chatterji, D. *J. Biochem. Biophys. Methods* **1995**, *30*, 1–8.
(3) Chen, F. M. *Biochemistry* **1992**, *31*, 3769–3776.
(4) Pedroso, I. M.; Duarte, L. F.; Yanez, G.; Baker, A. M.; Fletcher, T. M. *Biochem. Biophys. Res. Commun.* **2007**, *358*, 298–303.
(5) Sen, D.; Gilbert, W. *Nature* **1988**, *334*, 364–366.
(6) Henderson, E.; Hardin, C. C.; Walk, S. K.; Tinoco, I., Jr.; Blackburn, E. H. *Cell* **1987**, *51*, 899–908.
(7) Murchie, A. I.; Lilley, D. M. *Nucleic Acids Res.* **1992**, *20*, 49–53.
(8) Hammond-Kosack, M. C.; Kilpatrick, M. W.; Docherty, K. *J. Mol. Endocrinol.* **1992**, *9*, 221–225.
(9) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593–11598.
(10) Rankin, S.; Reszka, A. P.; Huppert, J.; Zloh, M.; Parkinson, G. N.; Todd, A. K.; Ladame, S.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2005**, *127*, 10584–10589.
(11) Guo, K.; Pourpak, A.; Beetz-Rogers, K.; Gokhale, V.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2007**, *129*, 10220–10228.
(12) Cogoi, S.; Xodo, L. E. *Nucleic Acids Res.* **2006**, *34*, 2536–2549.
(13) Maizels, N. *Nat. Struct. Mol. Biol.* **2006**, *13*, 1055–1059.
(14) De Armond, R.; Wood, S.; Sun, D.; Hurley, L. H.; Ebbinghaus, S. W. *Biochemistry* **2005**, *44*, 16341–16350.
(15) Phan, A. T.; Kuryavyi, V.; Burge, S.; Neidle, S.; Patel, D. J. *J. Am. Chem. Soc.* **2007**, *129*, 4386–4392.
(16) Zhao, Y.; Du, Z.; Li, N. *FEBS Lett.* **2007**, *581*, 1951–1956.
(17) Dai, J.; Dexheimer, T. S.; Chen, D.; Carver, M.; Ambrus, A.; Jones, R. A.; Yang, D. *J. Am. Chem. Soc.* **2006**, *128*, 1096–1098.
(18) Simonsson, T.; Pecinka, P.; Kubista, M. *Nucleic Acids Res.* **1998**, *26*, 1167–1172.
(19) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2005**, *33*, 2908–2916.
(20) Todd, A. K.; Johnston, M.; Neidle, S. *Nucleic Acids Res.* **2005**, *33*, 2901–2907.
(21) Huppert, J. L.; Balasubramanian, S. *Nucleic Acids Res.* **2007**, *35*, 406–413.
(22) Eddy, J.; Maizels, N. *Nucleic Acids Res.* **2006**, *34*, 3887–3896.

10.1021/ja801535e CCC: $40.75

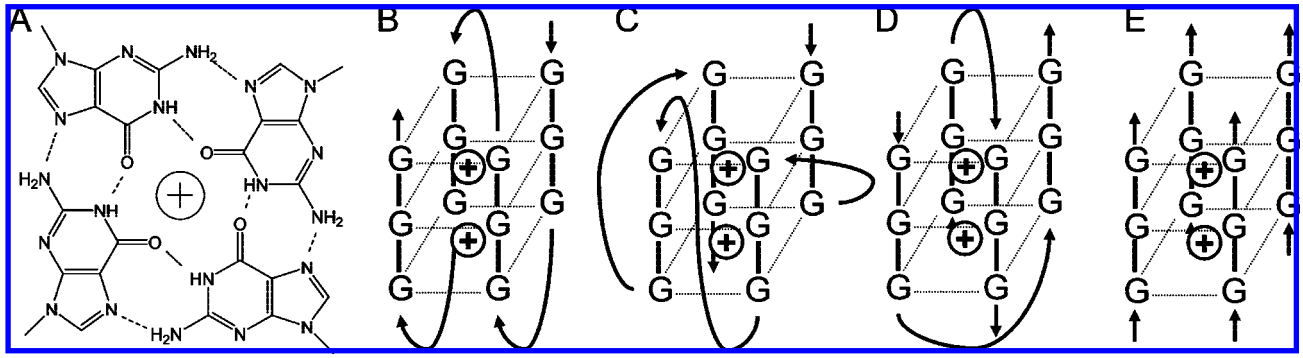


Figure 1. (A) Chemical structure of a G-quartet. (B−E) Schematic representations of an intramolecular antiparallel G-quadruplex (B), an intramolecular parallel G-quadruplex (C), a bimolecular antiparallel G-quadruplex (D), and a tetramolecular parallel G-quadruplex (E).

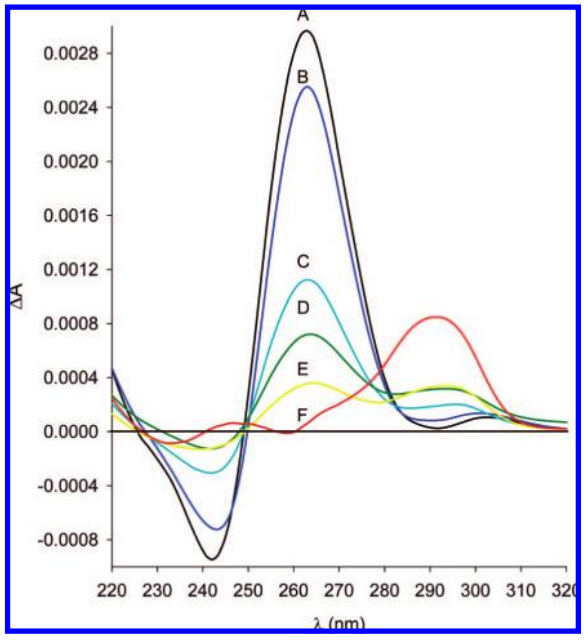


Figure 2. Circular dichroism spectra of six mixtures of oligonucleotides (total strand concentration of 5 μM) in 150 mM KCl: 1-1-2 (A, black), 1-3-1 (B, dark blue), 2-2-1 (C, cyan), 3-2-1 (D, green), 3-3-1 (E, yellow), 3-3-3 (F, red).

of sequences able to fold into G-quadruplexes, it would be interesting to predict, a priori, the structure and the stability of these G-quadruplexes, without having to study each particular sequence. In this context, in vitro studies aimed at establishing predictive rules taking into account the influence of loop length and the nature of the cation on G-quadruplex structure and stability could be valuable for predicting the formation of G-quadruplexes structures in vivo.

Different G-quadruplex topologies have been characterized in vitro. When the guanine tracts are oriented in opposite directions with loops above and below the terminal G-quartets, the G-quadruplex is called "antiparallel" (Figure 1B). When all guanines tracts are oriented in the same direction with loops located on the sides of the G-quartets, the G-quadruplex is called "parallel" (Figure 1C).[27,28] Hybrid structures combining both kinds of loops and strand orientations also exist. Moreover, multimeric G-quadruplexes can be formed by the association of two (Figure 1D) or four (Figure 1E) strands.[29]

Previous studies have shown that, in intramolecular G-quadruplexes, the loops that connect each tract of guanines play a key role in determining the conformation and stability.[30–32] For example, in the presence of potassium, sequences of the form $dTG_3T_nG_3T_nG_3T_nG_3T$ adopt a parallel conformation when $n = 1$ and a mixture of parallel and antiparallel conformations when $n = 2$, while sequences with longer loops are antiparallel when $n \geq 3$.[33] Influence of loop sequence on stability of G-quadruplexes was also studied by base substitution in sequences containing four tracts of three guanines with one base loop, indicating dramatic stability reduction for each substitution of T by A.[34] Finally, experiments were reported with sequences containing non-nucleosidic loops in order to determine the role of bases in the loops. In the case of intramolecular structures, it was demonstrated that non-nucleosidic loops favor parallel and more stable structures than their nucleosidic analogues.[35] However, for dimeric structures, the presence of non-nucleosidic loops destabilizes the G-quadruplex.[36]

Using experimental results on individual sequences to predict G-quadruplex DNA conformations and stability in the genome remains a challenge, given the huge number of different sequences that potentially form G-quadruplexes. Here, we propose an experimental approach to improve such predictions by studying oligodeoxynucleotides separated by loops containing random bases. Sequences studied in this manuscript have the general formula $dGGGW_iGGGW_jGGGW_kGGG$, with W = thymine or adenine with equal probability. In each sample, the number of distinct sequences is equal to $2^{(i+j+k)}$. In this study, loop length were systematically varied between 1 and 3, giving 27 possible W-containing sequences. These 27 mixtures of sequences contain a total of 2744 individual sequences. We

(23) Gabelica, V.; Baker, E. S.; Teulade-Fichou, M. P.; De Pauw, E.; Bowers, M. T. *J. Am. Chem. Soc.* **2007**, *129*, 895–904.
(24) Grand, C. L.; Han, H.; Munoz, R. M.; Weitman, S.; Von Hoff, D. D.; Hurley, L. H.; Bearss, D. J. *Mol. Cancer. Ther.* **2002**, *1*, 565–573.
(25) Seenisamy, J.; Rezler, E. M.; Powell, T. J.; Tye, D.; Gokhale, V.; Joshi, C. S.; Siddiqui-Jain, A.; Hurley, L. H. *J. Am. Chem. Soc.* **2004**, *126*, 8702–8709.
(26) Dexheimer, T. S.; Sun, D.; Hurley, L. H. *J. Am. Chem. Soc.* **2006**, *128*, 5404–5415.
(27) Simonsson, T. *Biol. Chem.* **2001**, *382*, 621–628.
(28) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
(29) Baker, E. S.; Bernstein, S. L.; Gabelica, V.; De Pauw, E.; Bowers, M. T. *Int. J. Mass Spectrom.* **2006**, *253*, 225–237.
(30) Wang, Y.; Patel, D. J. *J. Mol. Biol.* **1993**, *234*, 1171–1183.
(31) Wang, Y.; Patel, D. J. *Structure* **1993**, *1*, 263–282.
(32) Bugaut, A.; Balasubramanian, S. *Biochemistry* **2008**, *47*, 689–697.
(33) Hazel, P.; Huppert, J.; Balasubramanian, S.; Neidle, S. *J. Am. Chem. Soc.* **2004**, *126*, 16405–16415.
(34) Rachwal, P. A.; Brown, T.; Fox, K. R. *FEBS Lett.* **2007**, *581*, 1657–1660.
(35) Risitano, A.; Fox, K. R. *Nucleic Acids Res.* **2004**, *32*, 2598–2606.
(36) Cevec, M.; Plavec, J. *Biochemistry* **2005**, *44*, 15238–15246.

Table 1. Clustering of Studied Sequences in Six Groups and Average Melting Temperatures Obtained in 150 mM KCl, NaCl, or NH_4OAc

group	sequences	average melting temperature (°C)		
		KCl	NaCl	NH_4OAc
I	1−1−1, 1−2−1, 2−1−1, 1−1−2 (two loops of one base and one loop of <3 bases)	>80	48.6 ± 2.7	60 ± 6
II	1−3−1, 3−1−1, 1−1−3, 1−4−1 (two loops of one base and one loop of ≥3 bases)	>80	46 ± 3	50.0 ± 0.8
III	2−2−1, 1−2−2, 2−1−2 (one loop of one base and two of two bases)	>80	47 ± 5	48 ± 3
IV	1−2−3, 1−3−2, 2−1−3, 2−3−1, 3−1−2, 3−2−1 (one loop of one, of two and of three bases)	69.5 ± 0.8	44 ± 3	45.3 ± 1.7
V	3−3−1, 1−3−3, 3−1−3 (one loop of one base and two of three bases)	64.3 ± 0.8	43.0 ± 2.7	43.1 ± 0.6
VI	2−2−2, 3−2−2, 2−3−2, 2−2−3, 3−3−2, 3−2−3, 3−3−2, 3−3−3 (No loop of one base)	65 ± 4	51.3 ± 2.8	48.2 ± 1.2

present here the global view of the influence of loop length and cation on G-quadruplex topology and stability that emerges from the study of such sequence mixtures. The formation of dimeric and trimeric G-quadruplex assemblies with parallel conformation in sequences containing short loops is reported for the first time. This constitutes a novel type of G-quadruplex-based nanostructure. Finally, we demonstrate the predictive character of rules established with the sequences containing random bases by comparison with results obtained on the human telomeric sequence and the Pu22myc G-quadruplex.

Materials and Methods

Oligonucleotides Sequences. All oligonucleotides were ordered from Eurogentec (Seraing, Belgium) with Oligold quality. Their general sequence structure is 5′-$dG_3W_iG_3W_jG_3W_kG_3$-3′, with W = thymine or adenine with equal probability. Sequences were named *i-j-k*, according to the number of bases in loops connecting the four tracts of guanines (for example, 2-1-2 corresponds to $dG_3W_2G_3WG_3W_2G_3$). Three additional sequences 5′-$dW_2G_3W_iG_3W_jG_3W_kG_3$WT-3′ were used to study the influence of the presence of flanking sequences. Finally, we studied the human telomeric sequence dGGGTTAGGGTTAGGGTTAGGG, and the Pu22myc sequence dGAGGGTGGGGAGGGTGGGGAAG.

Sequences were received lyophilized and stock solutions were prepared in bidistilled water with 300 μM total strand concentration. For all experiments, the stock solution was heated at 80 °C during 5 min, diluted using a cold aqueous solution containing either KCl, NaCl, or NH_4OAc to reach the desired DNA concentration in 150 mM cation, and then cooled rapidly on ice. Lithium cacodylate (10 mM, pH = 7.4) was added in thermal denaturation and circular dichroism experiments. Lithium cacodylate was obtained by neutralizing lithium hydroxide by cacodylic acid.

Circular Dichroism. Experiments were performed on a Jobin Yvon CD6 dichrograph using 1-cm path length quartz cells (Hellma, type No. 120-QS, France). The final concentration of oligonucleotides was 5 μM in a buffer containing 150 mM salt and 10 mM lithium cacodylate, pH 7.4. For each sample, three spectra were recorded from 220 to 350 nm with a scan rate of 0.25 nm/s.

Thermal Denaturation. Thermal denaturation experiments were carried out on a Uvikon XS spectrophotometer (Secomam), using 1-cm path length quartz cells (Hellma, type No. 115B-QS, France). Oligonucleotides final concentration was 5 μM in 150 mM salt and 10 mM lithium cacodylate, pH = 7.4. Absorbance was monitored as a function of the temperature at 295, 240, 260 nm for the determination of the melting temperature (T_m)[37] and at 405 nm as control wavelength. Gradient was 0.3 °C/min between 10 and 90 °C. When hysteresis was observed, experiments were repeated at 0.2 °C/min to minimize this effect. Melting temperatures were determined using the method described by Marky and Breslauer.[38]

Gel Electrophoresis. Gel electrophoresis experiments were performed on 24% TBE acrylamide−bisacrylamide gels. To maintain G-quadruplex structures intact during the run (native gel electrophoresis), both the gel and the running buffer were supplemented by salt (20 mM). For denaturing gel electrophoresis experiments, gel was supplemented by 5 M urea. Oligonucleotide samples were prepared at 5 μM final concentration in a buffer containing the appropriate salt at 150 mM. Sucrose was added (10% final) to facilitate sample loading in the wells. Gels were run during maximum 3 h, stained with SYBR Green I (Roche), and scanned on a Phosphor Imager (Bio-Rad). Molecular weight markers $d(T)_{10}$, $d(T)_{15}$, $d(T)_{20}$, and $d(T)_{30}$ were used to facilitate comparisons between different gels.

Electrospray Mass Spectrometry (ESI-MS). All measurements were carried out on a Q-TOF Ultima Global mass spectrometer (Micromass, now Waters, Manchester, U.K.), using the electrospray ionization (ESI) source in negative mode, as described previously.[39] Source conditions were optimized to avoid in-source fragmentation: capillary voltage = −2.2 kV, cone voltage = 100 V, RF lens = 150 V, source block temperature = 80 °C, and desolvation gas temperature = 100 °C. Source pressure readback was set to 3.9 mbar and collision cell pressure readback was 3.0×10^{-5} mbar. Oligonucleotide samples were first prepared at 50 μM final concentration in NH_4OAc 150 mM. Just before injection in the mass spectrometer, they were further diluted to 10 μM in 150 mM NH_4OAc and 20% methanol. The role of methanol is to increase ion signals. Comparisons with spectra acquired without methanol confirmed that methanol influences only the absolute intensities and not the relative intensities of the different species. To calculate the relative abundances of the different peaks, raw spectra of each sequence were background subtracted, smoothed, and integrated using the MassLynk 4.0 software.

Ion Mobility Spectrometry Determination of Collision Cross Section. The instrument used for the ion mobility measurements[40] and the particular experimental conditions used for the Pu22myc sequence dGAGGGTGGGGAGGGTGGGGAAG have been described previously.[23] Briefly, the ion mobility mass spectrometer consists of a nanoelectrospray source to produce ions from the sample solution, an ion funnel to focus and form the ion packet, a drift cell which the ion packet travels through under the influence of a constant electric field, a quadrupole that mass-selects the ions, and a detector which records the ion arrival time distribution (ATD). From the ion's ATD, one can deduce the velocity of the ions in the drift cell. The drift cell is filled with helium at a pressure of 5 Torr. Collisions with helium produce a frictional force proportional to their velocity, such that the ions reach a constant drift velocity v_d that is proportional to the electric field E ($v_d = K \times E$). The proportionality constant K is termed the mobility of the ions and since K is dependent on the number density of the buffer gas, it is

(37) Mergny, J. L.; Phan, A. T.; Lacroix, L. *FEBS Lett.* **1998**, *435*, 74–78.
(38) Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601–1620.
(39) Rosu, F.; Gabelica, V.; Houssier, C.; Colson, P.; De Pauw, E. *Rapid Commun. Mass Spectrom.* **2002**, *16*, 1729–1736.
(40) Wyttenbach, T.; Kemper, P. R.; Bowers, M. T. *Int. J. Mass Spectrom.* **2001**, *212*, 13–23.

usually standardized (as shown in eq 1) with respect to molecular number density. Thus, the reduced mobility, K_o, is usually reported with T being the drift cell temperature in Kelvin and p the buffer gas pressure in Torr.

$$K_o = K\frac{p}{760}\frac{273.16}{T} \quad (1)$$

A value for K_o can be easily obtained since it is linearly related to the arrival time, t_A, of the ions through the instrument as shown in eq 2 where l is the length of the drift cell, V is the voltage drop across the cell, and t_o is the time the ion spends between the exit of the drift cell and the detector where the shape of the ion does not matter.[23]

$$t_A = \frac{l^2}{K_o}\frac{273.16}{760T}\frac{p}{V} + t_o \quad (2)$$

In the experiments detailed in this manuscript, K_o was acquired by collecting ATDs at 5 different drift voltages and t_A was extracted from the center of the ATD peaks. t_A was then plotted against the five different p/V values and the slope from this plot was utilized to calculate K_o. t_A versus p/V plots display a high degree of linearity with correlation values of at least 0.9999 indicating that the drift time (e.g., mobility) measurements are independent of the electric field.

Since the reduced mobility of an ion is dependent on the number of collisions it encounters with the buffer gas, information about the ion's shape and size, or in other words collision cross section, can be determined. The relationship between the mobility of an ion and its collision cross section has been derived in detail using kinetic theory[23] and is given by

$$K_o = \frac{3q}{16N_o}\left(\frac{2\pi}{\mu k_b T}\right)^{1/2}\frac{1}{\Omega} \quad (3)$$

where q is the ion charge, N_o is the buffer gas density at standard temperature and pressure (STP), μ is the reduced mass of the collision partners, k_b is Boltzmann's constant and Ω is the momentum transfer collision integral also termed the collision cross section.[23] To minimize error in the experimental evaluation of Ω, multiple cross section measurements were made on each system studied. Less than 1% variation was observed between measurements with the small variation most likely due to small pressure fluctuations.

Calculation of Collision Cross Section for Model Structures. To provide a structural interpretation of the systems studied, the experimental cross sections determined from the ATDs must be compared to the calculated cross section of theoretical models. To generate theoretical structures, molecular dynamics simulations using the AMBER 7 set of programs[41] were performed on structural models created from the PDB structure 1XAV of the Pu22myc sequence.[42] Each structure was then energy minimized and its cross section calculated using hard-sphere scattering and trajectory models developed by the Jarrold group.[43,44] In the calculations, the starting structures eventually converge to give one steady-state structure where the cross section remains relatively constant. The calculated cross sections reported here are the average cross sections of the final 50−100 structures (<2% standard deviation). An example is given as Figure S7 in Supporting Information.

Results

Circular Dichroism: Parallel versus Antiparallel Conformations. Circular dichroism (CD) experiments give information about the strand orientation (parallel, antiparallel, or hybrid) of G-quadruplexes, because the CD signal changes with the syn/anti orientation about glycosilic bonds. In parallel G-quadruplexes, all guanines have anti conformation about glycosilic bonds, and their CD spectrum exhibits a positive peak around 260 nm and a negative peak around 240 nm. By contrast, antiparallel G-quadruplexes have both syn and anti orientation about glycosilic bonds, and their CD spectrum displays a negative peak around 260 nm and a positive peak at 295 nm.[45–47] If the strand orientation depends on the loop length, the CD spectra should also vary with loop length in a concerted manner.

The CD spectra of six different sequences containing random bases in KCl are shown in Figure 2. Sequences 1-1-2 (A in Figure 2) and 1-3-1 (B in Figure 2) form parallel structures with positive peaks at 260 nm and negative peaks near 240 nm. Sequences with longer loops become more antiparallel as their loop length increases (D−F in Figure 2) as indicated by the growth of the peak near 295 nm and the reduction in the features at 260 and 240 nm. The complete set of spectra (28 sequences containing random bases in 150 mM KCl, NH_4OAc, or NaCl) is provided as Supporting Information Figures S1−S3. The general trends can be summarized as follows. First, sequences with short loops prefer to form parallel structures, whereas sequences with long loops form antiparallel structures. The reason is that loops of one base are too short to bridge two adjacent or opposite guanines from the same G-quartet, but are able to form propeller-like loops bridging guanines of different G-quartets. Second, the nature of the cation also influences the strand orientation, with parallel structures favored in K^+, antiparallel structures favored in Na^+, and an intermediate situation in NH_4^+. These CD results should be considered as an overview of the preferred structures and their evolution when loop length increases, rather than the exact conformation of all the sequences of the mixture. For example, it has been shown recently that the CD spectrum could change depending on the nature of W (adenine or thymine) for the sequence dGGG-WGGGTTTGGGWGGG in sodium.[48]

From the CD spectra, the sequences were clustered in six groups defined by the (i,j,k) values (Table 1). In agreement with previously published data,[33] sequences with short loops (groups I and II) adopt parallel structures. The only exception is sequence 1-4-1 that can also adopt a partially antiparallel structure in Na^+ solution. This highlights the role of the length of the central loop in the topology. CD spectra of sequences with intermediate loops (groups III, IV, and V) were characterized for all three cations by an evolution from mostly parallel to more antiparallel structures. Finally, antiparallel arrangements are favored when loops are long (group VI). In NaCl and NH_4OAc, G-quadruplexes are always antiparallel when loops are long. Antiparallel structures are formed with 2-2-2 and longer loop sequences in NH_4^+, and even with sequences containing one loop of one base (e.g., 2-3-1 or 3-1-1) in Na^+.

Thermal Denaturation: Relative Stability of G-Quadruplex Folds. Thermal denaturation experiments allow determining the melting temperature (T_m) of oligonucleotides, that is, the

(41) Case, D. et al. *Amber 7*; University of California: San Francisco, 2002.
(42) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. *Biochemistry* **2005**, *44*, 2048–2058.
(43) Shvartsburg, A. A.; Jarrold, M. F. *Chem. Phys. Lett.* **1996**, *261*, 86–91.
(44) Mesleh, M. F.; Hunter, J. M.; Shvartsburg, A. A.; Schatz, G. C.; Jarrold, M. F. *J. Phys. Chem.* **1996**, *100*, 16082–16086.
(45) Balagurumoorthy, P.; Brahmachari, S. K.; Mohanty, D.; Bansal, M.; Sasisekharan, V. *Nucleic Acids Res.* **1992**, *20*, 4061–4067.
(46) Balagurumoorthy, P.; Brahmachari, S. K. *J. Biol. Chem.* **1994**, *269*, 21858–21869.
(47) Lu, M.; Guo, Q.; Kallenbach, N. R. *Biochemistry* **1992**, *31*, 2455–2459.
(48) Guedin, A.; De Cian, A.; Gros, J.; Lacroix, L.; Mergny, J. L. *Biochimie* **2008**, *90*, 686–696.

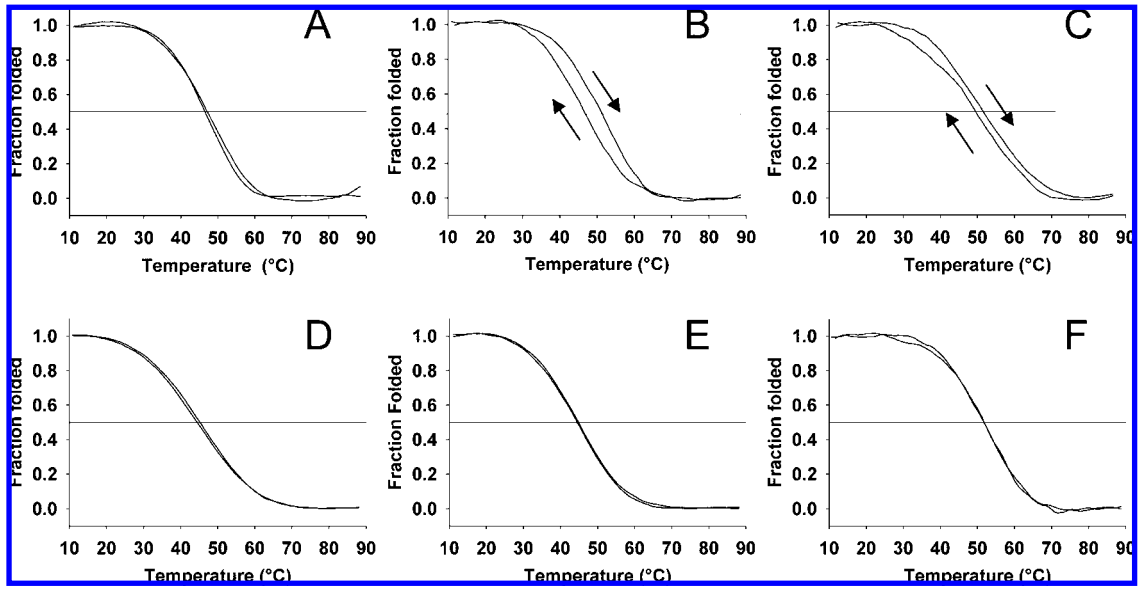


Figure 3. Thermal denaturation curves recorded by monitoring absorbance at 295 nm in NaCl 150 mM, for one sequence of each group (total strand concentration is 5 μM). 1-1-2 (A), 1-3-1 (B), 2-2-1 (C), 3-2-1 (D), 3-3-1 (E) and 3-3-3 (F). In samples B and C, arrows differentiate heating and cooling curves.

temperature at which the oligonucleotides are half-denaturated. The T_m indicates the relative stability of structures adopted by the oligonucleotides. It is well documented that the nature of the bases in the loops influence the thermal stability of intramolecular quadruplexes, with thymines being more stabilizing than adenines in loops.[34,48] Therefore, if the loop length has no particular influence on the stability of G-quadruplex folds, melting curves of our mixtures of sequences would be spread over a large temperature range. However, if the length of the loop is a key factor influencing the stability of intramolecular G-quadruplexes, the T_m values would cluster into similar groups as defined above.

Examples of thermal denaturation curves recorded in NaCl 150 mM are shown in Figure 3. Heating and cooling curves overlap for the majority of samples, but hysteresis was observed in a few cases (see for example Figure 3B,C). This indicates slow denaturation/renaturation processes, most probably due to the presence of multimeric structures.[49] Moreover, for several oligonucleotides (mainly of group IV in KCl, see the denaturation/renaturation curves in Supporting Information Figure S4), two transitions are observed upon heating but only one was observed on cooling. This suggests a mixture of a rapidly folding structure (probably an intramolecular structure) and a slowly folding less stable structure (probably a multimeric structure).[50] The occurrence of multimers will be addressed later. In this section, we discuss T_m values determined from the cooling curves to compare the relative stability of the structures and study the influence of loop length.

For each group defined by circular dichroism (Table 1), the mean T_m was determined for each cation. The standard deviation was calculated from the average of all samples in the group. The complete list of T_m values is provided in Supporting Information Table S1. These data reveal more subtle trends within the groups. For example, 2-2-2 often systematically has a significantly higher T_m than the other members of group VI. Despite these particularities, the low standard deviations indicate that the clustering established based on the CD experiments remains valid for the relative stability of the G-quadruplexes.

The general trends can be summarized as follows. (i) In KCl, the stability is inversely dependent on loop length, a result consistent with a prior report.[32] The decrease in stability therefore mirrors the evolution from parallel structures to hybrid structures. (ii) In NaCl, however, the antiparallel quadruplexes formed by sequences of group VI, where all loops have at least two bases, are the most stable. For sequences of other groups, the T_m only slightly decreases when the loop length increases. (iii) In NH_4OAc, the situation is intermediate between sodium and potassium: the most stable structures are the parallel G-quadruplexes of group I, followed by the antiparallel G-quadruplexes of group VI, and the least stable structures are those formed by sequences with intermediate loop length. (iv) For all sequences, the T_m values are higher for potassium ions than for sodium or ammonium ions. This trend is in line with previous reports indicating that potassium better stabilizes G-quadruplexes than other monovalent cations.[51–53] (v) Whether sodium or ammonium ion best stabilizes the quadruplex depends on the loop length. Sequences with short loops are more stable in NH_4^+ than in Na^+, whereas it is the opposite for sequences with long loops. For sequences with intermediate loop lengths, ΔT_m between the two cations depends subtly on the arrangement of loops in the sequence.

In conclusion, thermal denaturation experiments show that loop length plays a central role in the stability of G-quadruplexes. However, because of the presence of hysteresis and multistep transitions in these experiments, formation of multimeric structures may be involved. Further experiments using gel electrophoresis and mass spectrometry were carried out to investigate this possibility.

Native Gel Electrophoresis: Detection of Multimers. Native gel electrophoresis experiments have been performed to check

(49) Mergny, J. L.; De Cian, A.; Ghelab, A.; Sacca, B.; Lacroix, L. *Nucleic Acids Res.* **2005**, *33*, 81–94.
(50) Yu, H. Q.; Miyoshi, D.; Sugimoto, N. *J. Am. Chem. Soc.* **2006**, *128*, 15461–15468.
(51) Wlodarczyk, A.; Grzybowski, P.; Patkowski, A.; Dobek, A. *J. Phys. Chem. B* **2005**, *109*, 3594–3605.
(52) Risitano, A.; Fox, K. R. *Biochemistry* **2003**, *42*, 6507–6513.
(53) Ross, W. S.; Hardin, C. C. *J. Am. Chem. Soc.* **1994**, *116*, 6070–6080.

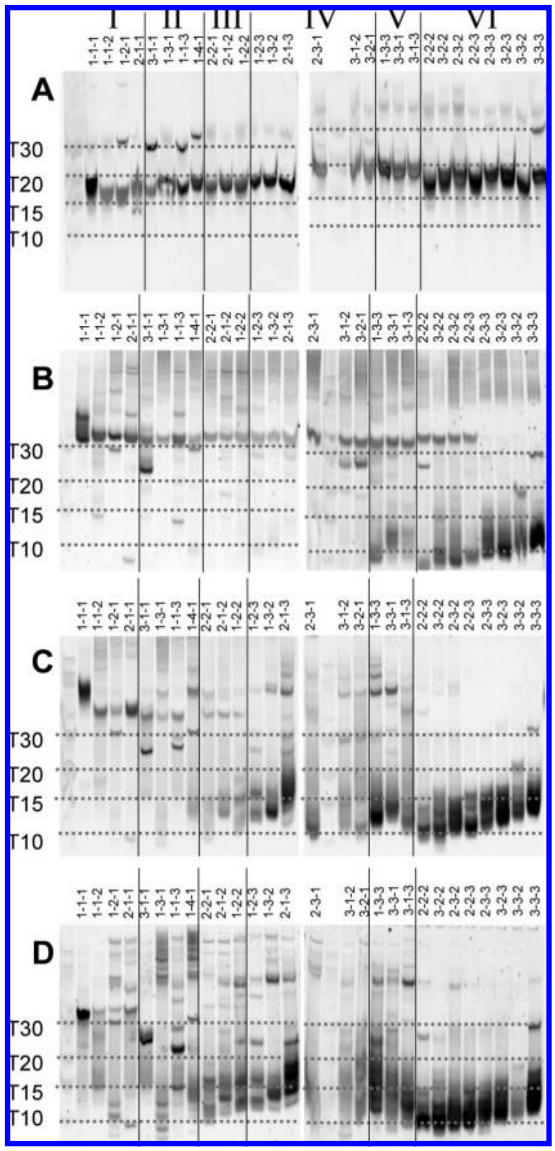


Figure 4. Denaturating gel electrophoresis in urea 5 M (A) and native gel electrophoresis in KCl 150 mM (B), NaCl 150 mM (C), and NH_4OAc 150 mM (D). The sequences are indicated on top of each gel. The horizontal dashed lines indicate the position of the molecular weight markers.

for possible formation of G-quadruplex aggregates. An experiment in denaturing conditions was performed to determine the positions of the unfolded oligonucleotides on the gels (Figure 4A). All unfolded monomer oligonucleotides migrate between the markers for dT_{15} and dT_{20}. In native conditions, it is then probable that oligonucleotides that migrate faster than dT_{15} are intramolecular G-quadruplexes. It is also reasonable to assume that oligonucleotides that migrate slower than dT_{20} are aggregates of the quadruplex structures.

The intra- or intermolecular character of all 28 sequence mixtures was determined in the presence of K^+ (Figure 4B), Na^+ (Figure 4C), and NH_4^+ (Figure 4D). The large number of bands is partly due to the presence of many different sequences in each well because A-rich and T-rich oligonucleotides migrate differently.[54] Nevertheless, clear trends emerge. For all three cations, sequences with long loops tend to form intramolecular G-quadruplexes, while sequences with short loops tend to form multimers. Some differences occur however according to the cation used. Potassium favors intermolecular arrangements, and

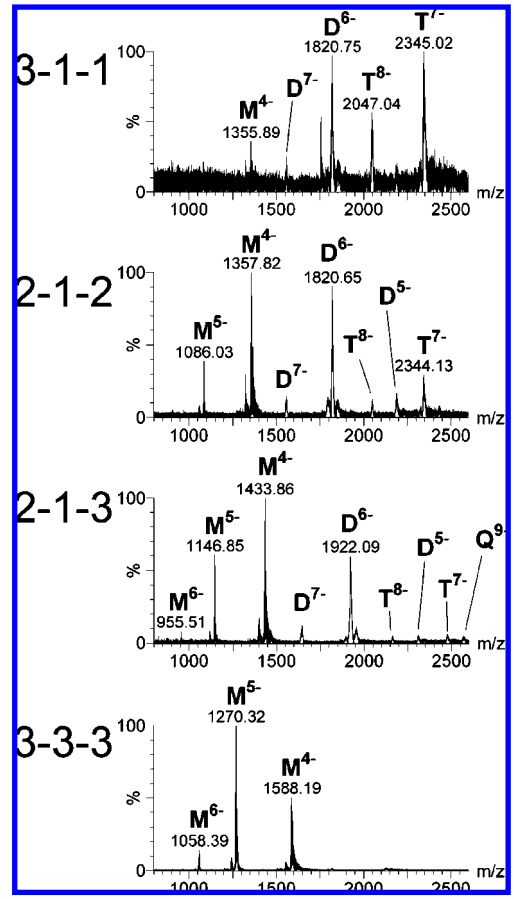


Figure 5. Electrospray mass spectra of samples 3-1-1, 2-1-2, 2-1-3, and 3-3-3. Legend: M = monomer, D = dimer, T = trimer, Q = tetramer. *Y*-axis is in arbitrary units.

only longer sequences of group VI adopt intramolecular conformations exclusively (Figure 4B). In sodium (Figure 4C) sequences with short loops (groups I and II) exhibit both inter- and intramolecular conformations, and sequences of other groups are mainly intramolecular. Ammonium (Figure 4D) has intermediate behavior.

Electrospray Mass Spectrometry: Determination of Strand Stoichiometry. Determining the strand stoichiometry from the gel electrophoresis experiments is difficult, but the presence of several bands well above the dT_{20} marker suggests the formation of species even larger than dimers. To refine this interpretation, the strand stoichiometry was determined by electrospray mass spectrometry, in NH_4OAc 150 mM. The gentle electrospray source conditions used here allow large hydrogen-bonded complexes to be transferred intact from the solution to the mass analyzer.[29,39,55–58] Typical spectra obtained using 150 mM ammonium acetate are shown in Figure 5. Even though the samples consist of a mixture of sequences (the mass difference between T and A is 9 Da) and bound ammonium cations are detected (each adduct gives a mass difference of 17 Da),

(54) Kejnovska, I.; Kypr, J.; Vorlickova, M. *Biochem. Biophys. Res. Commun.* **2007**, *353*, 776–779.
(55) Goodlett, D. R.; Camp, D. G., II; Hardin, C. C.; Corregan, M.; Smith, R. D. *Biol. Mass. Spectrom.* **1993**, *22*, 181–183.
(56) Daniel, J. M.; Friess, S. D.; Rajagopalan, S.; Wendt, S.; Zenobi, R. *Int. J. Mass Spectrom.* **2002**, *216*, 1–27.
(57) Rueda, M.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **2006**, *128*, 3608–3619.
(58) Vairamani, M.; Gross, M. L. *J. Am. Chem. Soc.* **2003**, *125*, 42–43.

the mass spectra remain fairly simple. Knowing the average molecular mass $<M>$ of the oligonucleotide mixture, the number of strands (n) is deduced from the average mass-to-charge ratio ($<m/z>$) using the following equation:

$$n = \{<m/z \times z\}/<M>$$

The charge state z can be deduced either from the isotopic distribution when the peaks are well resolved (usually the case for monomers), from the spacing between consecutive peaks corresponding to different T/A content (separated by $9/z$ Da; these peaks are usually well resolved in the case of monomers and dimers), or alternatively from the smallest integer values of n and z that can account for a given $<m/z>$.

The ESI-MS spectra (Figure 5) reveal unambiguously the presence of monomers, dimers, trimers, and tetramers. In line with the gel electrophoresis results, the tendency to form monomers, dimers, and trimers varies according to the loop length. Supporting Information Table S2 gives the relative abundances of monomer, dimer, and trimer peaks for all sequences. Although the observed relative abundances are not exactly equal to the relative concentrations of each species in solution (due to differential ionization effects), the data are useful for detecting trends as a function of the loop size and position.

The tendency to form larger oligomers increases as the loop length decreases, in agreement with the native gel electrophoresis results. Sequences with short loops predominantly form dimers and trimers. Since the CD spectra of the three sequences are characteristic of parallel structures (Figure S3), it follows that the dimer and trimer structures are parallel. Furthermore, a closer examination of Table S2 reveals that the loop position is also important: (i) The comparison of relative abundance of monomer of sequences 1-1-1, 1-2-1, 1-3-1, and 1-4-1 indicates that a longer central loop favors monomer structure. (ii) The length of the third loop (3′ end) has a larger influence than the length of the first (5′ end) loop: all sequences having a single base in their third loop systematically form higher-order oligomers. (iii) Sequences with a short second loop and a long third loop prefer dimeric conformations.

Discussion

Influence of Loop Length and Cation on G-Quadruplex Topology. For each cation, the evolution of structure and stability according to the loop length can be summarized as follows. Potassium favors parallel structures: sequences with short loops are very stable intermolecular parallel G-quadruplexes. As the loop length increases, intramolecular and less stable hybrid or mixed G-quadruplexes are observed. Sodium, however, strongly favors antiparallel conformations: the most stable G-quadruplex structures are the intramolecular antiparallel G-quadruplexes formed by sequences with long loops. Ammonium has intermediate behavior: G-quadruplex structures are intermolecular and parallel when loops are short and intramolecular and antiparallel when they are long.

The induction of preferentially parallel structures in potassium and antiparallel structures in sodium is well-known.[45,46] This is generally attributed to the different coordination modes of these two cations. Because of their ionic radii, potassium cations can only be accommodated between two quartets in an 8-fold coordination with guanines, whereas smaller sodium cations can also be located in the middle of a G-quartet in a square planar coordination. Sodium cations located in the center of terminal quartets can therefore interact with loop bases,[59] and this is why antiparallel conformations are more stable than their parallel counterpart in sodium.

The second factor influencing the parallel ↔ antiparallel equilibrium is loop length.[33–35,60,61] The longer the loop length, the greater the tendency to form antiparallel structures. When there are at least two loops with one base, the structure is always parallel whatever the cation. Single-base loops are too short to bridge guanines from the same quartet, but can bridge guanines of different quartets in an edgewise conformation. Furthermore, we have shown here that sequences with short loops can also form multimeric assemblies.

To test the ideas developed here from idealized sequences, we compared our results with a well studied system: the human telomeric sequence dGGGTTAGGGTTAGGGTTAGGG, which belongs to the 3-3-3 group of sequences. The human telomeric sequence has been extensively studied. The circular dichroism spectrum of 3-3-3 in KCl exhibits a major peak around 295 nm and a minimum around 233 nm, in agreement with the CD spectrum of the telomeric sequence, which has a hybrid structure in KCl.[62–66] The CD results on 3-3-3 are also consistent with the fully antiparallel structure of telomeric sequence in sodium solution.[31] The telomeric sequence with 3.5 repeats is always observed as a monomer in ESI-MS, similar to the 3-3-3 sequences.

Structure of the Multimeric Assemblies. A surprising result of the present study is the detection of abundant multimers in sequences containing short loops. Nanoassembly of guanine-rich sequences into G-wire structures is a well-known phenomenon,[67–71] and the classical representation of these structures is shown in Figure 6A. However, trimer formation by sequences containing four tracts of guanines has never been reported or proposed previously. The unambiguous detection of trimers by mass spectrometry requires a revised view of G-quadruplex higher-order structures, beyond the classical representations shown in Figure 6A. The correlation between the formation of multimers and the formation of parallel G-quadruplexes strongly suggests that the multimeric assemblies are parallel.

A number of structures for the multimers are possible. Since all sequences form monomer quadruplexes, the simplest multimer structure is formed by stacking of parallel monomers as shown in Figure 6B. In that model the trimer is therefore an intermediate en route to the formation of higher-order structures.

(59) Schultze, P.; Hud, N. V.; Smith, F. W.; Feigon, J. *Nucleic Acids Res.* **1999**, *27*, 3018–3028.
(60) Rachwal, P. A.; Brown, T.; Fox, K. R. *Biochemistry* **2007**, *46*, 3036–3044.
(61) Rachwal, P. A.; Findlow, I. S.; Werner, J. M.; Brown, T.; Fox, K. R. *Nucleic Acids Res.* **2007**, *35*, 4214–4222.
(62) Xu, Y.; Noguchi, Y.; Sugiyama, H. *Bioorg. Med. Chem.* **2006**, *14*, 5584–5591.
(63) Phan, A. T.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2006**, *34*, 5715–5719.
(64) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723–2735.
(65) Dai, J.; Punchihewa, C.; Ambrus, A.; Chen, D.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 2440–2450.
(66) Luu, K. N.; Phan, A. T.; Kuryavyi, V.; Lacroix, L.; Patel, D. J. *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970.
(67) Marsh, T. C.; Henderson, E. *Biochemistry* **1994**, *33*, 10718–10724.
(68) Marsh, T. C.; Vesenka, J.; Henderson, E. *Nucleic Acids Res.* **1995**, *23*, 696–700.
(69) Miyoshi, D.; Nakao, A.; Sugimoto, N. *Nucleic Acids Res.* **2003**, *31*, 1156–1163.
(70) Miyoshi, D.; Wang, Z. M.; Karimata, H.; Sugimoto, N. *Nucleic Acids Symp. Ser. (Oxford)* **2005**, *49*, 43–44.
(71) Miyoshi, D.; Karimata, H.; Wang, Z. M.; Koumoto, K.; Sugimoto, N. *J. Am. Chem. Soc.* **2007**, *129*, 5919–5925.

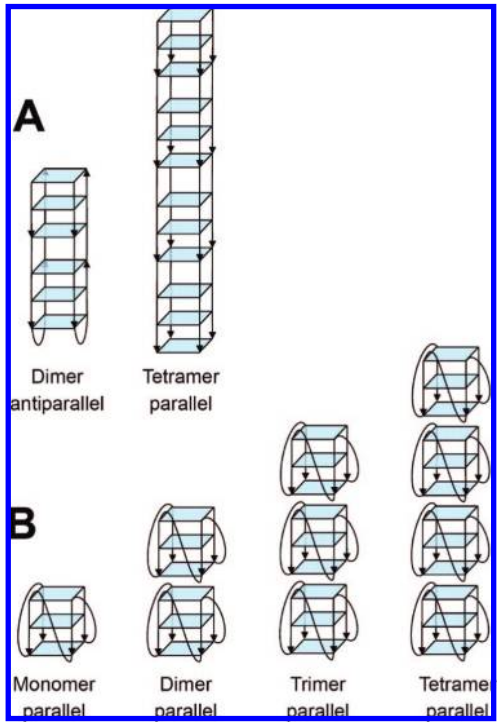


Figure 6. (A) Classical representation of G-quadruplex dimer and tetramer. (B) Possible structure of the parallel dimer, trimer, and tetramer formed with sequences $dGGGW_iGGGW_jGGGW_kGGG$.

This simple model readily accounts for the preferential multimer formation for by parallel G-quadruplexes since there is no loop impediment to stacking. That model also implies that the first step is the formation of parallel intramolecular structures, then the assembly into multimers, so that multimer formation is a consequence of the formation of parallel structures. Moreover, this stacking is highly favored by our sequences starting and ending with guanines involved in G-tetrads.

However, it is expected that additional bases or covalent groups on both sides of our sequences would interfere with simple stacking and that multimer formation would be disfavored. For example, the apparent discrepancy between our results for $dG_3W_iG_3W_jG_3W_kG_3$ sequences, where we found essentially all multimer formation for $i,j,k = 1$, and previously published results on $dTG_3TG_3TG_3TG_3T$, where only monomer formation was found even in potassium,[60,61] can be explained by the fact that the sequences studied by Rachwal and co-workers were fluorescently labeled on both ends.

To evaluate this effect we studied three sequences with additional bases at the extremities: 5′-WW-[1-1-1]-WT-3′, 5′-WW-[3-1-1]-WT-3′ and 5′-WW-[1-2-2]-WT-3′. Circular dichroism spectra are virtually the same with and without flanking sequences (Supporting Information Figure S5), but the strand stoichiometries observed by mass spectrometry in 150 mM NH_4^+ are different (Table 2). Monomers are favored when flanking sequences are added, while trimers are no longer observed. However, dimeric structures can still be formed. For example, the sequence WW-1-2-2-WT still forms nearly as much dimer as 1-2-2. The same trends are observed in K^+ and Na^+ by gel electrophoresis (Supporting Information Figure S6).

The presence of dimers but not trimers when bases are added on both sides suggests that other dimeric structures might be formed in addition to or in the place of the stacked dimers shown in Figure 6B. For example, if 5′-to-5′ or 3′-to-3′ stacking of two monomers is much more favorable than the 3′-to-5′ stacking, oligomerization does not go beyond the dimer. The formation of 5′-5′ dimers has already been reported for the octamer $(GGGT)_8$, consisting of two interlocked $(GGGT)_4$ units,[72] in the X-ray crystal structure of $[(TGGGGT)_4]_2$,[73] and for the 99del dGGGGTGGGAGGAGGGT aptamer, which forms an interlocked dimeric parallel-stranded structure.[74] Compared to intramolecular G-quadruplexes, interlocked structures are stabilized by extra hydrogen bonds between bases of different sequences. In the present study of mixtures of sequences, the possibility exists that higher-order structures may be favored when a sequence with a given loop encounters a different sequence, suggesting the possibility that higher order structures may be favored when interstructure loop complementarity occurs. In addition, while formation of parallel intramolecular structures is a prerequisite for the formation of stacked multimers, the situation is different for interlocked dimers or multimers. Indeed, in this case, the conformation would be determined by the competition between intramolecular folding and intermolecular hydrogen bonding, and nonparallel multimers could also be envisaged. The present experiments do not allow distinguishing between the different possible dimer structures, but the subject of oligomer formation by short G-quadruplex forming sequences clearly warrants further attention, because these short sequences are usually studied in vitro as models of the putative structures formed in vivo.

Table 2. Influence of Flanking Sequences on the Relative Abundance of Monomer (% M), Dimer (% D), and Trimer (% T), Determined by Electrospray Mass Spectrometry in 150 mM NH_4OAc

sequences	% M	% D	% T
1-1-1	14	34	52
WW-1-1-1-WT	91	9	0
3-1-1	6	47	47
WW-3-1-1-WT	70	30	0
1-2-2	50	39	11
WW-1-2-2-WT	63.5	36.5	0

Parallel Dimer Formation by the Pu22myc Quadruplex. This will be illustrated in the case of the well-known Pu22myc sequence dGAGGGTGGGGAGGGTGGGGAAG. This sequence is close to the WW-1-2-2-WT group of sequences. The structure of Pu22myc G-quadruplex was reported to be predominantly parallel,[25,42,75] which is in line with the results on WW-1-2-2-WT. However, on the basis of the results obtained in the present study, the formation of a minor parallel dimer is predicted. When performing ESI-MS experiments on the Pu22myc sequence in ammonium acetate, the presence of the dimer is indeed unambiguously confirmed by a peak at $m/z = 2030$, corresponding to a $Dimer^{7-}$ (Figure 7A). The dimer-to-monomer ratio depends on the oligonucleotide concentration during annealing.

To determine if the minor dimer species is parallel or antiparallel, ion mobility spectrometry experiments were performed to determine the collision cross section of the $Dimer^{7-}$ (Figure 7B), and the experimental value was compared with theoretical values computed for an antiparallel dimer model

(72) Krishnan-Ghosh, Y.; Liu, D.; Balasubramanian, S. *J. Am. Chem. Soc.* **2004**, *126*, 11009–11016.
(73) Caceres, C.; Wright, G.; Gouyette, C.; Parkinson, G.; Subirana, J. A. *Nucleic Acids Res.* **2004**, *32*, 1097–1102.
(74) Phan, A. T.; Kuryavyi, V.; Ma, J. B.; Faure, A.; Andreola, M. L.; Patel, D. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 634–639.
(75) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710–8716.

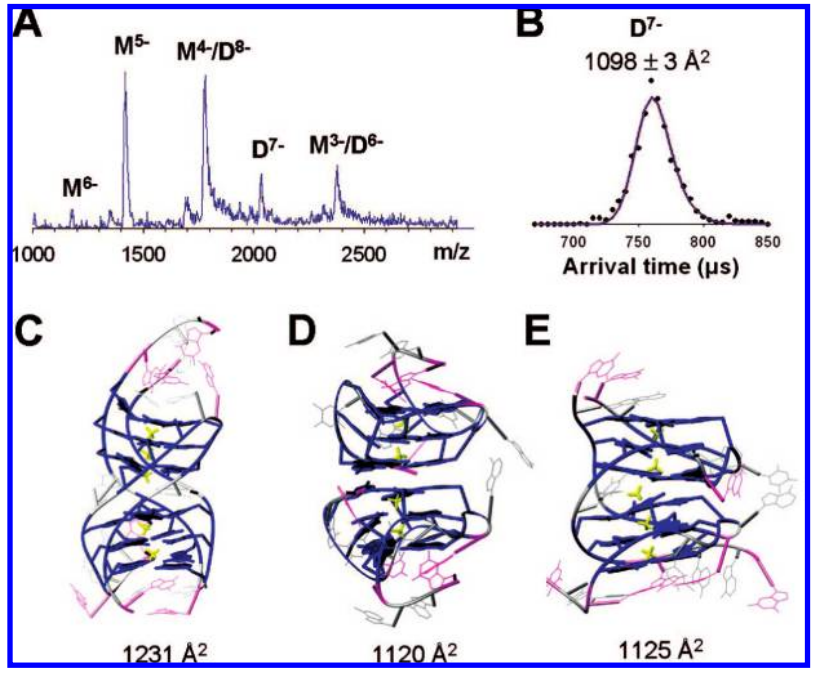


Figure 7. Electrospray ion mobility spectrometry of the Pu22myc dimer. (A) ESI-MS spectrum of Pu22myc in ammonium acetate prepared by annealing at [DNA] = 50 μM, final concentration is 5 μM. (B) Arrival time distribution of the $Dimer^{7-}$ and experimental collision cross section. (C−E) Theoretical models for an antiparallel dimer (C), a stacked parallel dimer (D), and an interlocked parallel dimer (E) and the corresponding calculated collision cross sections.

(Figure 7C), a stacked dimer model (Figure 7D), and an interlocked dimer model (Figure 7E). The value calculated for the stacked parallel dimer (1120 ± 10 Å^2) and the intertwined parallel dimer (1125 ± 10 Å^2) are much closer to the experimental value (1098 ± 3 Å^2) than the value calculated for any antiparallel dimer (1231 ± 15 Å^2). These results therefore confirm the formation of the parallel dimer of the Pu22myc sequence in vitro. Whether the Pu22myc dimer is stacked or interlocked cannot be distinguished with the current methods of investigation. The stacked dimer is slightly favored both in cross section (see above) and thermodynamically (AMBER calculations, Supporting Information Figure S7). It is also kinetically significantly favored if the dimer formation mechanism is association of two preformed quadruplex monomers. However, the fraction of dimer in Pu22myc depends more strongly on denaturing conditions than on the concentration of Pu22myc in the ESI spray solution suggesting at least some dimer may be irreversibly formed with an interlocking structure during annealing. Further work needs to be done to unambiguously assign the dimer structure in this system.

Conclusions

The study of model G-rich DNA sequences containing mixtures of linking bases is a reliable approach for predicting the structure and stability of particular G-quadruplex forming sequences. In particular, this experimental design is a cheap and easy way to study the influence of loop length without unexpected biases due to particular choices in the base composition of the loops. The present study classified the structure and stability of the sequences with the general motif $dGGGW_iGGGW_jGGGW_kGGG$ (i,j,k = 1−3 or 4). The major findings can be summarized as follows:

(1) Parallel G-quadruplex structures are favored by short loops and by $K^+ > NH_4^+ > Na^+$. Longer loops favor antiparallel structures.

(2) Formation of higher-order G-quadruplex assemblies is also favored by short loops and by $K^+ > NH_4^+ > Na^+$. These assemblies are therefore most likely parallel.

(3) For the first time, trimers have been observed for sequences containing four tracks of three guanines. This observation led to the proposal of a novel type of G-quadruplex higher-order structure: stacked parallel structures are formed by sequences beginning and ending with guanines and constitute a novel G-wire motif.

(4) This highlights one of the caveats in using short oligonucleotides as models for in vitro studies of particular genomic sequences. If the model oligonucleotide is chosen too short (e.g., when the G-quadruplex unit is starting and ending with a guanine), unexpected phenomena can occur in vitro (e.g., aggregation).

(5) Stacked monomer structures appear to be the most likely structures for the multimers but some contribution of interlocked dimers cannot be completely ruled out, especially for sequences containing flanking sequences.

(6) Based on the observations made, we predicted and then demonstrated that the well-known Pu22myc sequence, a model for the NHE III_1 region of the c-myc oncogene, can also form some parallel dimer structures in vitro.

Acknowledgment. Acknowledgement to the FNRS and FRIA is also made (V.G. is a FNRS research associate, F.R. is a FNRS postdoctoral researcher, and N.S. is a FRIA doctoral fellow). The authors acknowledge the financial contribution of the Fonds de la Recherche Scientifique-FNRS (FRFC 2.4.623.05 to EDP and PC; CC 1.5.096.08 to VG) and of the National Science Foundation (MTB).

Supporting Information Available: Complete ref 41; Table S1 and S2; circular dichroism spectra; AMBER molecular dynamics data. This material is available free of charge via the Internet at http://pubs.acs.org.

JA801535E