Sequence effects of single base loops in intramolecular quadruplex DNA

Phillip A. Rachwal^a, Tom Brown^b, Keith R. Fox^{a,*}

^a School of Biological Sciences, University of Southampton, Bassett Crescent East, Southampton SO16 7PX, UK ^b School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK

Received 7 February 2007; revised 8 March 2007; accepted 16 March 2007

Available online 28 March 2007

Edited by Hans Eklund

Abstract We have examined the properties of intramolecular G-quadruplexes in which the G3 tracts are separated by single base loops. The most stable complex contained 1',2'-dideoxyribose in all three loops, while loops containing T and C were slightly less stable (by about 2 °C). Quadruplexes containing loops with single A residues were less stable by 8 °C for each T to A substitution. These folded sequences display similar CD spectra, which are consistent with the formation of parallel stranded complexes with double-chain reversal loops. These results demonstrate that loop sequence, and not just length, affects quadruplex stability.

© 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved.

Keywords: DNA quadruplex; DNA structure

1. Introduction

DNA sequences that contain four or more G-tracts can fold to form intramolecular structures that consist of stacks of Gquartets [1-5]. The four G-tracts are separated by different length loops, which can be as short as a single nucleotide [6-8]. The complexes can adopt a range of structures in which the bases are either anti or syn and in which the strands run parallel or antiparallel [4,5]. The loops can be arranged in several different ways; double chain reversal (propeller) loops link two adjacent parallel strands by a connection between the top and bottom G-tetrads [9], while edgewise or diagonal loops link two antiparallel strands [10]. Some structures contain both edgewise and propeller loops [11-13]. G-rich sequences with the potential to adopt these structures are found in several gene promoters; most notably the c-myc [8,14,15], bcl2 [16,17] and c-kit [18,19] oncogenes and several of these contain single base loops. Although it has been shown that loop length affects quadruplex stability and structure [6,20-23] there have been few studies on the effect of loop sequence, though single base changes can have a significant effect (e.g. changing TTA in the human telomeric repeat to TTG [10] in the Tetrahymena repeat [24,25]).

One very stable intramolecular quadruplex contains four G_3 tracts that are linked by single T residues [7,26]. We have investigated the effect of loop sequence on its stability using sequences of the type d(TGGGNGGGNGGGNGGGT), where

2. Materials and methods

2.1. Oligonucleotides

All oligonucleotides were synthesized on an Applied Biosystems ABI 394 automated DNA/RNA synthesizer on the 0.2 or 1 μ mole scale using the standard cycles of acid-catalysed detritylation, coupling, capping and iodine oxidation procedures. Phosphoramidite monomers and other reagents were purchased from Applied Biosystems or Link Technologies. The sequences of the oligonucleotides used in this work are shown in Table 1. All oligonucleotides were prepared with 5'-fluorescein and 3'-dabcyl (fluorescein C6 phosphoramidite and dabcyl cpg purchased from Link Technologies Ltd.) for use in the fluorescence melting experiments and the same sequences were used for the circular dichroism studies. The bases adjacent to the fluorophore and quencher were the same (T) for all the oligonucleotides to ensure that the terminal base did not affect quadruplex formation and stability. Inclusion of this base also hinders any fluorescence quenching between G and fluorescein.

2.2. Fluorescence melting

Fluorescence melting curves were determined in a Roche Light-Cycler as previously described [22,26,27] in a total reaction volume of 20 µL. Oligonucleotides (final concentration 0.25 µM) were prepared in 10 mM lithium phosphate pH 7.4, which was supplemented with various concentrations of potassium or sodium chloride. The LightCycler has one excitation source (488 nm) and the changes in fluorescence were measured at 520 nm. In order to avoid hysteresis between heating and cooling curves melting experiments were performed at a slow rate of temperature change (0.2 °C min⁻¹). This was achieved by changing the temperature in 1 °C steps and leaving the samples to equilibrate for 5 min at each temperature before recording the fluorescence. In a typical experiment the oligonucleotides were first denatured by heating to 95 °C for 5 min. They were then annealed by cooling to 30 °C at 0.2 °C min⁻¹ and melted by heating to 95 °C at the same rate. The fluorescence was recorded during both the annealing and melting steps.

 $T_{\rm m}$ values were obtained from the maxima of the first derivatives of the melting profiles using the LightCycler software or, together with ΔH , from van't Hoff analysis of the melting profiles [23,26–28]. The fraction folded was calculated as previously described [28] from the difference between the measured fluorescence and the upper and lower baselines. All reactions were performed at least twice and the $T_{\rm m}$ values differed by <0.5 °C with a 5% variation in ΔH . Since ΔG = 0 at the $T_{\rm m}$, ΔS was estimated as $\Delta H/T_{\rm m}$. It should therefore be noted that ΔS is not determined independently of ΔH and $T_{\rm m}$. This analysis assumes a simple two-state equilibrium between the folded and unfolded forms. The presence of polymorphic quadruplex structures will lead to shallower melting profiles and therefore smaller apparent values for ΔH .

2.3. Circular dichroism

CD spectra were measured on a Jasco J-720 spectropolarimeter as previously described [22]. Oligonucleotides solutions (5 μ M) were prepared in 10 mM lithium phosphate pH 7.4, containing either 100 mM potassium chloride or 100 mM sodium chloride. The samples were

0014-5793/\$32.00 © 2007 Federation of European Biochemical Societies. Published by Elsevier B.V. All rights reserved. doi:10.1016/j.febslet.2007.03.040

N is each base in turn (except G) together with 1',2'-dideoxyribose, Φ .

E-mail address: k.r.fox@soton.ac.uk (K.R. Fox).

Table 1 Sequences of the oligonucleotides used in this work

Name	Oligonucleotide sequence
TTT	d(Fam-TGGGTGGGTGGGTGGGT-dabcyl)
CCC	d(Fam-TGGGCGGGCGGGCGGGT-dabcyl)
AAA	d(Fam-TGGGAGGGAGGGAGGGT-dabcyl)
$\Phi \Phi \Phi$	d(Fam-TGGG Φ GGG Φ GGG Φ GGG T -dabcyl)
TAT	d(Fam-TGGGTGGGAGGGTGGGT-dabcyl)
ATA	d(Fam-TGGGAGGGTGGGAGGGT-dabcyl)

Fam = fluorescein; $\Phi = 1', 2'$ -dideoxyribose.

heated to 95 °C and annealed by slowly cooling to 15 °C over a period of 12 hours. Spectra were recorded between 220 and 320 nm in 5 mm path length cuvettes. Spectra were averaged over 10 scans, which were recorded at 100 nm min⁻¹ with a response time of 1 s and 1 nm bandwidth. A buffer baseline was subtracted from each spectrum and the spectra were normalized to have zero ellipticity at 320 nm.

3. Results

We have investigated the effect of sequence on the stability of quadruplexes that contain single nucleotide loops using oligonucleotides of the type (TGGGNGGGNGGGNGGGT), where N is each base in turn. The sequences of these oligonucleotides are shown in Table 1; they were each labelled at the 5'-end with fluorescein and with dabcyl at the 3'-end. Fluorescence melting curves were used to assess the thermal stability of these complexes. When the quadruplex is folded the fluorophore and quencher are in close proximity and the fluorescence is quenched [22,26–28]. When the structure unfolds the fluorophore and quencher are separated and there is an increase in fluorescence. In the presence of 100 mM KCl all these sequences formed structures that melted above 90 °C. However, by using only 1 mM KCl the melting temperatures decreased to a measurable range. Although this concentration is lower than that used in most quadruplex studies, higher potassium ion concentrations produced melting curves that were too stable to measure. The fluorescence melting profiles (in which the fluorophore and quencher are close together) and the similar patterns of stability in both 1mM KCl and 100 mM NaCl, strongly suggest that these sequences adopt an intramolecular quadruplex. Melting profiles for the four oligonucleotides that contain the same base in each loop (TTT, CCC, AAA or $\Phi\Phi\Phi$)



Fig. 1. Fraction of the different oligonucleotides folded at different temperatures (A) in 10 mM lithium phosphate pH 7.4 containing 1 mM KCl. (B) in 10 mM lithium phosphate pH 7.4 containing 100 mM NaCl. Solid line, TTT; dotted line, $\Phi\Phi\Phi$; dot and dash line, CCC; dashed line, AAA. These plots were derived from the melting curves by assuming that the complexes are folded at low temperature and single stranded at elevated temperatures [27].

Fable 2	2
---------	---

 $T_{\rm m}$ values (°C) for the various oligonucleotides, determined in 10 mM lithium phosphate pH 7.4 containing different concentrations of KCl or NaCl

		TTT	AAA	CCC	$\Phi\Phi\Phi$	ATA	TAT
KCl	0	46.0	<30	44.4	49.5	36.9	37.9
	1	73.9	50.0	71.8	77.0	59.6	67.3
	5	85.6	64.4	83.2	88.1	72.1	79.3
	10	>90	70.0	>90	>90	78.0	84.2
	50	>90	82.2	>90	>90	>90	>90
NaCl	1	48.3	<30	47.3	53.2	<30	40.0
	10	55.4	<30	53.6	60.5	39.6	46.3
	50	63.7	<30	62.3	68.0	47.0	55.6
	100	68.3	43.7	67.1	72.5	52.2	60.5
	200	73.3	50.2	72.1	77.0	57.8	65.6

The oligonucleotide concentration was 0.25 μM and the samples were heated and cooled at 0.2 °C min^{-1}. All T_m values are ±0.5 °C.

are shown in Fig. 1 in the presence of 1 mM KCl or 100 mM NaCl and the melting temperatures determined at a range of ionic strengths are summarized in Table 2. At this rate of heating ($0.2 \,^{\circ}$ C min⁻¹) there was no hysteresis between the melting and annealing profiles and the $T_{\rm m}$ values were independent of oligonucleotide concentration between 0.1 and 10 μ M.

It can be seen that loops with T, C and 1',2'-dideoxyribose produce the most stable complexes ($\Phi > T > C$), while the $T_{\rm m}$ of the complex with As in the loops is about 25 °C lower. The complexes are less stable in sodium containing buffers, but the rank order of stabilities is the same. The complex with A-containing loops has a similar stability to a complex containing propanediol linkers [22]. Thermodynamic parameters for the folding of each complex in the presence of potassium were estimated from van't Hoff analysis of the melting profiles and these are presented in Table 3. The values are consistent with those previously determined for intramolecular quadruplexes, which are typically between 65 and 100 kJ mol⁻¹ per quartet [23,26]. These show that the complexes with the lower stability are characterized by a smaller enthalpy and higher entropy.

To further study the effect of loop sequence on stability we examined the properties of oligonucleotides with two Ts in the loops and one A (TAT), or one T in the loop and two As (ATA). The melting profiles of these sequences are shown in Fig. 2, alongside those for AAA and TTT and the thermodynamic parameters are presented in Table 3. Each T to A modification reduces the melting temperature by about 8 °C in both sodium and potassium-containing buffers, with a change in enthalpy of 46 kJ mol^{-1} per A substitution in 1 mM KCl and 33 kJ mol⁻¹ in 100 mM NaCl.

CD spectral signatures are often used as indicators of the folding pattern of intramolecular quadruplexes; parallel structures (with *anti* glycosidic bonds) typically have positive maxima around 260 nm, while antiparallel structures (containing both *anti* and *syn* bonds) show maxima around 295 nm [29–32]. CD spectra of these sequences in the presence of 100 mM KCl are shown in Fig. 3 and show that the complexes produce very similar CD spectra with clear single maxima at 265 nm. These CD spectra were identical in sodium and potassium-containing buffers and strongly suggest that all these complexes adopt the same topology, which is likely to be the

	Table 3						
	Thermodynamic	parameters	for folding	of the quadruples	forming of	oligonucleotides	5
1							100

Oligo	1 mM KCl			100 mM NaCl			
	$T_{\rm m}$ (°C)	$\Delta H (\mathrm{kJ} \mathrm{mol}^{-1})$	$\Delta S (\text{kJ mol}^{-1} \text{K}^{-1})$	$T_{\rm m}$ (°C)	$\Delta H (\mathrm{kJ}\mathrm{mol}^{-1})$	$\Delta S (\text{kJ mol}^{-1} \text{ K}^{-1})$	
TTT	73.9	-278 ± 8	-0.80	68.3	-206 ± 3	-0.60	
CCC	71.8	-257 ± 3	-0.75	67.1	-181 ± 10	-0.53	
AAA	50.0	-142 ± 8	-0.44	43.7	-104 ± 9	-0.33	
$\Phi\Phi\Phi$	77.0	-296 ± 10	-0.85	72.5	-229 ± 8	-0.66	
TAT	67.3	-247 ± 4	-0.73	60.5	-158 ± 6	-0.47	
ATA	59.6	-198 ± 8	-0.60	52.2	-131 ± 11	-0.40	

All $T_{\rm m}$ values are ±0.5 °C; Experiments were performed in 10 mM lithium phosphate pH 7.4 containing 1 mM KCl or 100 mM NaCl. ΔH was calculated from van't Hoff analysis of the melting profiles, assuming a two-state equilibrium and ΔS was estimated as $\Delta H/T_{\rm m}$.



Fig. 2. Fraction of the different oligonucleotides folded at different temperatures (A) in 10 mM lithium phosphate pH 7.4 containing 1 mM KCl. (B) in 10 mM lithium phosphate pH 7.4 containing 100 mM NaCl. Solid line, TTT: dot and dash line, TAT; dotted line, ATA; dashed line, AAA. These plots were derived from the melting curves by assuming that the complexes are folded at low temperature and single stranded at elevated temperatures [27].



Fig. 3. CD spectra of the different oligonucleotides in 10 mM lithium phosphate, pH 7.4 containing 100 mM KCl. The oligonucleotide concentration was $5 \,\mu$ M. solid line, TTT; dotted line, AAA; dashed line, CCC; dot and dash line, $\Phi \Phi \Phi$.

parallel form containing double-chain reversal, propeller loops.

4. Discussion

These results demonstrate that the loop sequence has a large effect on quadruplex stability. Structures with loops containing T, C or 1',2'-dideoxyribose (Φ) are the most stable ($\Phi > T > C$) and these are about 20 °C more stable than those with A in the loops. Each A-containing loop decreases the T_m by about 8 °C.

The size of the base may be significant as purine bases (AAA) produce less stable complexes than pyrimidines (TTT and CCC). Surprisingly the most stable complex is produced with 1',2'-dideoxyribose linkers, suggesting that specific interactions with the loop bases are not necessary for forming these intramolecular quadruplexes. However inclusion of the non-nucleosidic propanediol linker (which is the same length as a single nucleotide) has the same effect as A, though in this case the increased flexibility of the loop may limit the stability. An alternative explanation is that the structure of the unfolded state may itself be sequence dependent. It is known that adenine stacks particularly well at the ends of duplexes [33] and any base specific stacking will need to be reversed in order to form a parallel quadruplex. Studies with longer loops have also observed the stacking of adenine against the terminal quartets [8,14], which would not be possible for a parallel quadruplex containing single nucleotide loop. The CD spectra of these sequences are also very similar, suggesting that the differences in stability do not arise from gross differences in the folding topology, for which the most abundant form in each case is probably a parallel orientation of the strands.

Structural studies of various quadruplex forming sequences have shown several examples of single nucleotide loops that contain single T, C or A residues [8,14,17,34]. Searches for potential quadruplex-forming sequences in the human genome reveal that the most common linker between the G-tracts is single base A or T [35,36]. It is therefore clear that the lower stability of quadruplexes with A-containing loops does not preclude intramolecular quadruplex formation, but the inclusion of A instead of T may moderate the stability, thereby permitting the interconversion between different structural forms.



Fig. 4. Relationship between ΔH and ΔS determined for the complexes in the presence of 1 mM KCl (filled circles) and 100 mM NaCl (open circles).

Comparison of the thermodynamic parameters for this closely related series of quadruplexes (Table 3) reveals a correlation between the stability, enthalpy and entropy. The most stable complexes (higher T_m) have higher enthalpies which are compensated by less favourable entropic contributions (more negative ΔS). The relationship between enthalpy and entropy is illustrated in Fig. 4, which reveals that the same effect is seen in the presence of both potassium and sodium. It appears that the larger enthalpic contribution, which may arise from better stacking between the quartets and/or the loops, is accompanied by less flexibility, maybe by forming more ordered loop structures [23,37].

Acknowledgement: PAR is supported by a research studentship from BBSRC.

References

- Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure. Nucleic Acids Res. 34, 5402–5415.
- [2] Davis, J.T. (2004) G-quartets 40 years later: from 5'-GMP to molecular biology and supramolecular chemistry. Angew. Chem., Int. Ed. Engl. 43, 668–698.
- [3] Keniry, M.A. (2000) Quadruplex structures in nucleic acids. Biopolymers 56, 123–146.
- [4] Phan, A.T., Kuryavyi, V. and Patel, D.J. (2006) DNA architecture: from G to Z. Curr. Opin. Struct. Biol. 16, 288–298.
- [5] Simonsson, T. (2001) G-quadruplex DNA structures variations on a theme. Biol. Chem. 382, 621–628.
- [6] Hazel, P., Huppert, J., Balasubramanian, S. and Neidle, S. (2004) Loop-length-dependent folding of G-quadruplexes. J. Am. Chem. Soc. 126, 16405–16415.
- [7] Jing, N.J., Rando, R.F., Pommier, Y. and Hogan, M.E. (1997) Ion selective folding of loop domains in a potent anti-HIV oligonucleotide. Biochemistry 36, 12498–12505.
- [8] Phan, A.T., Modi, Y.S. and Patel, D.J. (2004) Propeller-type parallel-stranded G-quadruplexes in the human c-myc promoter. J. Am. Chem. Soc. 126, 8710–8716.
- [9] Parkinson, G.N., Lee, M.P.H. and Neidle, S. (2002) Crystal structure of parallel quadruplexes from human telomeric DNA. Nature 417, 876–880.
- [10] Wang, Y. and Patel, D.J. (1993) Solution structure of the human telomeric repeat d[AG₃(T₂AG₃)₃] G-tetraplex. Structure 1, 263– 282.
- [11] Ambrus, A., Chen, D., Dai, J.X., Bialis, T., Jones, R.A. and Yang, D.Z. (2006) Human telomeric sequence forms a hybridtype intramolecular G-quadruplex structure with mixed parallel/ antiparallel strands in potassium solution. Nucleic Acids Res. 34, 2723–2735.
- [12] Luu, K.N., Phan, A.T., Kuryavyi, V., Lacroix, L. and Patel, D.J. (2006) Structure of the human telomere in K⁺ solution: an intramolecular (3+1) G-quadruplex scaffold. J. Am. Chem. Soc. 128, 9963–9970.
- [13] Phan, A.T., Luu, K.N. and Patel, D.J. (2006) Different loop arrangements of intramolecular human telomeric (3 + 1) Gquadruplexes in K⁺ solution. Nucleic Acids Res. 34, 5715–5719.
- [14] Ambrus, A., Chen, D., Dai, J.X., Jones, R.A. and Yang, D.Z. (2005) Solution structure of the biologically relevant G-quadruplex element in the human c-myc promoter implications for Gquadruplex stabilization. Biochemistry 44, 2048–2058.
- [15] Siddiqui-Jain, A., Grand, C.L., Bearss, D.J. and Hurley, L.H. (2002) Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-myc transcription. Proc. Natl. Acad. Sci. USA 99, 11593–11598.
- [16] Dai, J.X., Chen, D., Jones, R.A., Hurley, L.H. and Yang, D.Z. (2006) NMR solution structure of the major G-quadruplex structure formed in the human BCL2 promoter region. Nucleic Acids Res. 34, 5133–5144.

- [17] Dai, J.X., Dexheimer, T.S., Chen, D., Carver, M., Ambrus, A., Jones, R.A. and Yang, D.Z. (2006) An intramolecular Gquadruplex structure with mixed parallel/antiparallel G-strands formed in the human BCL-2 promoter region in solution. J. Am. Chem. Soc. 128, 1096–1098.
- [18] Fernando, H., Reszka, A.P., Huppert, J., Ladame, S., Rankin, S., Venkitaraman, A.R., Neidle, S. and Balasubramanian, S. (2006) A conserved quadruplex motif located in a transcription activation site of the human c-kit oncogene. Biochemistry 45, 7854– 7860.
- [19] Rankin, S., Reszka, A.P., Huppert, J., Zloh, M., Parkinson, G.N., Todd, A.K., Ladame, S., Balasubramanian, S. and Neidle, S. (2005) Putative DNA quadruplex formation within the human ckit oncogene. J. Am. Chem. Soc. 127, 10584–10589.
- [20] Cevec, M. and Plavec, J. (2005) Role of loop residues and cations on the formation and stability of dimeric DNA G-quadruplexes. Biochemistry 44, 15238–15246.
- [21] Hazel, P., Parkinson, G.N. and Neidle, S. (2006) Topology variation and loop structural homology in crystal and simulated structures of a bimolecular DNA quadruplex. J. Am. Chem. Soc. 128, 5480–5487.
- [22] Risitano, A. and Fox, K.R. (2004) Influence of loop size on the stability of intramolecular DNA quadruplexes. Nucleic Acids Res. 32, 2598–2606.
- [23] Smirnov, I. and Shafer, R.H. (2000) Effect of loop sequence and size on DNA aptamer stability. Biochemistry 39, 1462–1468.
- [24] Wang, Y. and Patel, D.J. (1994) Solution structure of the Tetrahymena telomeric repeat $d(T_2G_4)_4$ G-tetraplex. Structure 2, 1141–1156.
- [25] Miyoshi, D., Karimata, H. and Sugimoto, N. (2005) Drastic effect of a single base difference between human and Tetrahymena telomere sequences on their structures under molecular crowding conditions. Angew. Chem., Int. Ed. Engl. 44, 3740–3744.
- [26] Risitano, A. and Fox, K.R. (2003) Stability of intramolecular DNA quadruplexes: comparison with DNA duplexes. Biochemistry 42, 6507–6513.
- [27] Darby, R.A.J., Sollogoub, M., McKeen, C., Brown, L., Risitano, A., Brown, N., Barton, C., Brown, T. and Fox, K.R. (2002) High throughput measurement of duplex, triplex and quadruplex melting curves using molecular beacons and a LightCycler. Nucleic Acids Res. 30, e39.
- [28] Mergny, J.L. and Lacroix, L. (2003) Analysis of thermal melting curves. Oligonucleotides 13, 515–537.
- [29] Balagurumoorthy, P., Brahmachari, S.K., Mohanty, D., Bansal, M. and Sasisekharan, V. (1992) Hairpin and parallel quartet structures for telomeric sequences. Nucleic Acids Res. 20, 4061– 4067.
- [30] Balagurumoorthy, P. and Brahmachari, S.K. (1994) Structure and stability of human telomeric sequence. J. Biol. Chem. 269, 21858– 21869.
- [31] Lu, M., Guo, Q. and Kallenbach, N.R. (1992) Structure and stability of sodium and potassium complexes of dT_4G_4 and dT_4G_4T . Biochemistry 31, 2455–2459.
- [32] Rujan, I.N., Meleney, J.C. and Bolton, P.H. (2005) Vertebrate telomere repeat DNAs favor external loop propeller quadruplex structures in the presence of high concentrations of potassium. Nucleic Acids Res. 33, 2022–2031.
- [33] Bommarito, S., Peyret, N. and SantaLucia, J. (2000) Thermodynamic parameters for DNA sequences with dangling ends. Nucleic Acids Res. 28, 1929–1934.
- [34] Phan, A.T., Kuryavyi, V., Ma, J.B., Faure, A., Andreola, M.L. and Patel, D.J. (2005) An interlocked dimeric parallel-stranded DNA quadruplex: a potent inhibitor of HIV-1 integrase. Proc. Natl. Acad. Sci. USA 102, 634–639.
- [35] Huppert, J.L. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. Nucleic Acids Res. 33, 2908– 2916.
- [36] Todd, A.K., Johnston, M. and Neidle, S. (2005) Highly prevalent putative quadruplex sequence motifs in human DNA. Nucleic Acids Res. 33, 2901–2907.
- [37] Olsen, C.M., Gmeiner, W.H. and Marky, L.A. (2006) Unfolding of G-quadruplexes: energetic, and ion and water contributions of G-quartet stacking. J. Phys. Chem. B 110, 6962–6969.