*Published in : Chemical Communications (2009)*
*Status : Postprint (Author's version)*

# Selectivity in small molecule binding to human telomeric RNA and DNA quadruplexes†

Gavin Collie,[a] Anthony P. Reszka,[a] Shozeb M. Haider,[a] Valérie Gabelica,[b] Gary N. Parkinson [a] and Stephen Neidle[a]

[a] *Cancer Research UK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London, UK WC1N 1AX*
[b] *Mass Spectrometry Laboratory, GIGA-Research, University of Liège, Belgium*

Quadruplex RNAs are less well understood than their DNA counterparts, yet of potentially high biological relevance. The interactions of several quadruplex-binding ligands with telomeric RNA quadruplexes are reported and compared with their binding to the analogous DNA quadruplexes.

DNA quadruplexes (G4s), both within gene promoter sequences and at telomeres, have been extensively studied as potential small-molecule therapeutic targets.[1] A variety of topologies have been found for both inter- and intramolecular DNA G4s, and detailed structural studies by X-ray crystallography, molecular simulation and NMR methods have highlighted the role played by the loops in small-molecule recognition.[2] By contrast, there have been few studies until recently on RNA quadruplexes. Putative RNA G4 sequences have been found in 5'-untranslated sequences of a number of genes,[3] and may be plausible therapeutic targets for appropriate small molecules. Most recently surprising observations have been made that telomeric DNA can be transcribed into telomeric RNA sequences,[4] which appear to play an important role in regulating telomerase function. It is plausible that in cellular conditions they can spontaneously fold into G4 arrangements.[5] A bimolecular G4 formed from two separate strands, each having two telomeric RNA repeats, has a parallel topology in both $Na^+$ and $K^+$ solutions.[6] An NMR analysis[7] has shown that the structure is analogous to the parallel telomeric DNA G4 crystal structures.[8] Both 23- and 24-mer four-repeat telomeric RNA sequences also form a stable parallel topology in either $Na^+$ or $K^+$ solution.[7,9] We report here on biophysical and molecular modeling studies of small-molecule interactions with the four-repeat intramolecular human telomeric RNA sequence r[AGGG(UUAGGG)$_3$]. We have used the experimental telomeric G4 drug BRACO-19 (Fig. 1) and three tetrasubstituted naphthalene diimides[10] (compounds 1-3). In each case the side-chains comprise two or more $CH_2$ groups terminating in an amine group. The binding of BRACO-19 to G4 DNAs has been previously studied *in vitro* and *in vivo* by a variety of biophysical and biological methods.[11] A crystal structure[12] of a complex between a bimolecular human telomeric G4 of sequence d(TAGGGTTAGGGT) and BRACO-19 has shown that this DNA G4 has a parallel topology in the crystalline state, in common with the native 12-mer and 22-mer telomeric DNA sequences,[8] and that the substituents interact extensively with the TTA loop regions in the complex, promoting their conformational remodelling. The naphthalene diimide ligands have high affinity for telomeric DNA G4 as well as some promoter G4s.[10b,c] If telomeric quadruplexes are to be useful as druggable targets then binding to (as well as selectivity between) G4 DNAs and RNAs is required. Telomeric RNAs inhibit telomerase activity, possibly by directly binding to the telomerase RNA template, so their stabilization as quadruplex-ligand complexes would enhance this inhibition and thus would be a viable therapeutic strategy. Circular dichroism (CD) spectra show that the 22-mer RNA telomeric G4 has a characteristic parallel topology[13] in $K^+$ solution (Fig. 2). The spectrum of the RNA 22-mer shows a positive band at 265 nm and a minimum at 245 nm, in accord with that reported for the bimolecular telomeric G4 RNA 12-mer[6,7] and the four-repeat 23- and 24-mer[6,7,9] telomeric G4 RNAs. This conservation of topology is in striking contrast with the highly polymorphic behaviour of four-repeat telomeric DNA G4s, which depends on factors such as the precise nature of the flanking sequences, molecular crowding and concentration.[14] We conclude that the parallel topology of RNA G4s, now observed in 22-, 23- and 24-mers, is an intrinsic feature that does not depend on flanking sequence.

† Electronic supplementary information (ESI) available: Details of molecular dynamics simulation methods, backbone interactions in the RNA quadruplex and plots of structural stability, groove dimensions and nitrogen-oxygen interactions against time. See DOI: 10.1039/b901889a

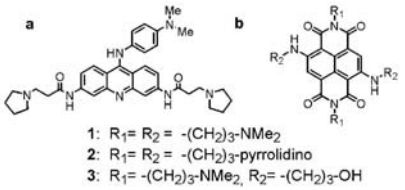


***Fig. 1*** *Structures of (a) BRACO-19 and (b) naphthalene diimides.*

The melting behaviour of the 22-mer sequence was examined by UV and CD methods, in $K^+$ solution, and shows a hypochromic shift characteristic of quadruplex formation[15] and a $T_m$ of 74.7 °C. The $T_m$ for the DNA 22-mer is 66.5 °C in identical conditions (50 mM $K^+$ buffer). Interactions of the 22-mer RNA quadruplex with BRACO-19 were examined using CD, and no significant spectral changes were observed (Fig. 2). The melting behaviour of the 22-mer in 50 mM $K^+$ solution with BRACO-19 at various molar ratios was monitored by a FRET method,[16] and compared with the behaviour under identical conditions with the 22-mer DNA quadruplex. The $\Delta T_m$ values at a 1 μM ligand concentration for the DNA and RNA G4s are 27.8° and 11.3°, respectively; this preferential stabilization of the DNA G4 is maintained over a range of ligand concentrations.

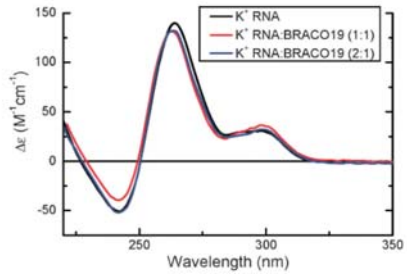


***Fig. 2*** *CD spectra of G4 RNA with BRACO-19 in 50 mM $K^+$ buffer.*

We have used electrospray mass spectrometry methods[17] to quantitate the interactions of all four ligands with DNA and RNA G4s. Spectra corresponding to stable G4 monomers with two $NH_4^+$ ions were observed with both intramolecular G4s (Fig. 3). Ligands were added to G4 solutions prior to injection and incubated at room temperature for 1 min at ligand: G4 ratios of 2.5:5, 5:5 and 10:5 (μM). Equilibrium dissociation constants ($K_d$) were determined from the peak areas[18] (Table 1).

The ranking order for the $K_d$ values obtained by ESI-MS for the two BRACO-19 G4 complexes is in qualitative accord with the FRET data (and with the modelling studies—see below) as BRACO-19 was found to bind to the telomeric DNA G4 with *ca.* 2-fold higher affinity than to the RNA one. All three naphthalene diimide ligands bind to the DNA G4 with high affinity, whereas the RNA G4 forms a strong complex only with ligand 3.

***Table 1*** *Equilibrium dissociation constants (in μM), with esds calculated from repeat experiments*

| | | $K_d1$ | $K_d2$ |
|---|---|---|---|
| r[AGGG(UUAGGG)$_3$] | 1 | 63.1 ± 8.5 | — |
| | 2 | 25.1 ± 2.0 | 12.6 ± 2.9 |
| | 3 | 4.0 ± 0.5 | 5.0 ± 1.4 |
| | BRACO-19 | 15.8 ± 1.0 | 4.0 ± 0.5 |
| d[AGGG(TTAGGG)$_3$] | 1 | 12.6 ± 3.0 | — |
| | 2 | 5.0 ± 1.0 | 12.6 ± 2.0 |
| | 3 | 4.0 ± 1.0 | 12.0 ± 2.2 |
| | BRACO-19 | 7.9 ± 1.4 | 5.3 ± 1.0 |

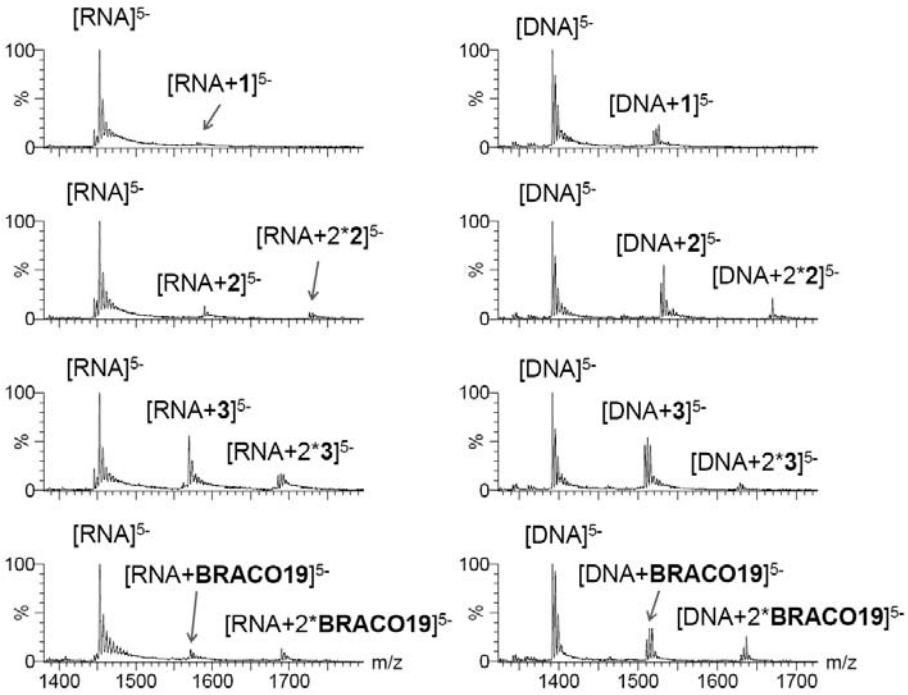


***Fig. 3*** *Spectra of G4-ligand complexes. RNA, r[AGGG(UUAGGG)]$_3$; DNA, d[AGGG(TTAGGG)]$_3$. Both at concentrations of 5 μM; the ligand concentration was 5 μM. Buffer was 100 mM $NH_4^+$.*

A molecular dynamics approach[19] has been used to simulate the structures of all eight DNA and RNA intramolecular G4 ligand complexes. The crystal structure of the native 22-mer DNA G4[8] (PDB id 1KF1) and the bimolecular G4 complex with BRACO-19[12] (PDB id 3CE5) were used as starting-points for the study. The RNA 22-mer G4 was generated from the DNA structure by direct addition of 2'-OH groups. Simulations for the native DNA and RNA quadruplexes and the eight ligand complexes all employed identical protocols.

The slightly lower RMSD value (1.90 Å) of the native 22-mer RNA G4 following its simulation suggest that it is more stable than that of the DNA G4 «RSMD) of 2.01 Å), in accord with the $T_m$ data presented here, and with thermodynamic studies on DNA and RNA tetramolecular G4s.[20] In addition to the eight H-bonds that form the co-planar array in each G-quartet, the presence of the 2'-OH groups in the G4 RNA was shown by the MD simulations to impart rigidity to the backbone by invoking interactions with the 05' groups and water molecules in the grooves. The overall structures of the 22-mer G4 DNA and RNA are closely similar, in accord with the recent NMR analysis[7] and more detailed X-ray studies of a bimolecular RNA G4 (Collie *et al.*, to be published).

On the other hand, the significantly lower RMSD values of the 22-mer DNA G4 complexes with ligands 1, 2 and BRACO-19 (1.77 *vs.* 2.18 Å for BRACO-19) suggest that they are more stable than their RNA G4 complexes, and are in qualitative accord with the binding data presented here. The primary interactions of all the ligands with both G4s are π-π stacking interactions of the aromatic chromophore with a terminal G-quartet and electrostatic interactions between the positively-charged ligand side-chain and the negatively-charged G4 backbone in the loops. These strong electrostatic interactions are responsible for stabilizing the otherwise highly mobile loops. The lower affinity shown by ligands 1, 2 and BRACO-19 for RNA G4 is explained by the presence of the additional 2'-OH groups which constrict the space available to the ligand side-chains to interact with the loops, by reducing the depth and the width of the UUA loops.

The structural stability of the BRACO-19-DNA G4 complex is enhanced by the interactions of the terminal 3' thymine 04 atom with a side-chain amide group in BRACO-19, as observed in the co-crystal structure.[12] This maintains BRACO-19 asymmetrically on one half of a G-quartet face. Analogous interactions have been observed in antiparallel crystal structures of an *Oxytricha nova* telomeric DNA G4 in complexes with

disubstituted acridine ligands.[21] In the RNA G4-BRACO-19 simulated complex, the side-chain amide group is sandwiched between the terminal uracil 04 atom and the 2'-OH group of a guanine (G17) on which the drug is stacked, so that an amide nitrogen atom of BRACO-19 interacts with the carbonyl oxygen atom (04) from U24. However the amide carbonyl oxygen atom is also directly opposite 2'-OH from G17. The close proximity of the two oxygen atoms causes unfavourable electrostatic repulsion. This results in the side chain of the drug oscillating during the simulation between the amide nitrogen-04 and 2'-OH stable interactions, and repulsion from oxygen atoms at the opposite end. This behaviour further destabilises the complex and as a consequence the side-chains are unable to interact with the charged loops (Fig. 4).

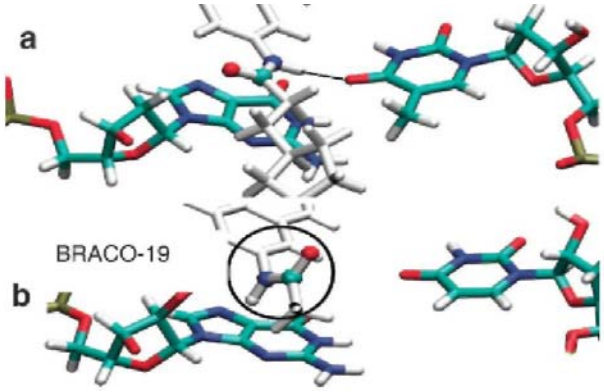


***Fig. 4*** *Interactions of the amide group in BRACO-19 (white) in the complexes with (a) DNA and (b) RNA G4s. The amide group (circled) is unable to form stable interactions in the RNA complex due to the 2'-OH group on G17. The figure shows a time-averaged structure.*

The naphthalene diimide ligands were originally designed as high-affinity G4 DNA binding molecules. A surprising observation is that only one of them (3) forms a strong complex with a telomeric RNA G4 monomer (Table 1). The substitution of -$NMe_2$ for an -OH group increases RNA G4-ligand binding 15-fold, but gives only a 3-fold increase in DNA G4-ligand binding. The effect of -OH groups on RNA *vs.* DNA G4-ligand interactions is an important structural feature to consider when designing selective G4 interacting compounds. This could be exploited, if targeting RNA G4s (*e.g.* the 5' UTR of oncogene promoters) or avoided, if targeting DNA G4s (*e.g.* telomeric DNA).

In summary, this study has shown that small-molecule ligands can discriminate between DNA and RNA G4s. It also confirms that four-repeat 22-mer telomeric RNA sequences fold into a parallel G4 in solution that is more stable than their DNA counterparts, and that flanking sequence changes do not affect their ability to form parallel G4s. Ligand binding selectivity is a consequence of the 2'-OH groups in the RNA and their effects on groove and loop widths so that ligand side-chains interact less effectively than with the DNA G4. One can anticipate that enhanced selectivity between DNA and RNA G4s may be achieved by appropriate ligands; their detailed design will require experimental structural data on RNA G4 complexes.

This work was supported by Cancer Research UK (Programme Grant No. C129/A4489), by European COST action MP0802, by the University of Liege (FRSD 08/10) and the FNRS (CC 1.5.286.09.F).

**Notes and references**

1. 1 See for example: A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.,* 2002, 99, 11593; T.-M. Ou, Y.-J. Lu, C. Zhang, Z.-S. Huang, X.-D. Wang, J.-H. Tan, Y. Chen, D.-L. Ma, K.-Y. Wong, J. C.-O. Tang, A. S.-C. Chan and L.-Q. Gu, *J. Med. Chem.,* 2007, 50, 1465; M.-K. Cheng, C. Modi, J. C. Cookson, I. Hutchinson, R. A. Heald, A. J. McCarroll, S. Missailidis, F. Tanious, W. D. Wilson, J.-L. Mergny, C. A. Laughton and M. F. G. Stevens, *J. Med. Chem.,* 2008, 51, 963; also see recent reviews: A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou and J.-L. Mergny, *Biochimie,* 2008, 90, 131; D. Monchaud and M.-P. Teulade-Fichou, *Org. Biomol. Chem.,* 2008, 6, 627; T. M. Ou, Y. J. Lu, J. H. Tan, Z. S. Huang, K. Y. Wong and L. Q. Gu, *Chem Med Chem,* 2008, 3, 690.

2. 2 D. J. Patel, A. T. Phan and V. Kuryavyi, *Nucleic Acids Res.,* 2007, 35, 7429; S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.,* 2006, 34, 5402.

3. 3 O. Kikin, Z. Zappala, L. D'Antonio and P. S. Bagga, *Nucleic Acids Res.,* 2008, D141; S. Kumari, A. Bugaut, J. L. Huppert and S. Balasubramanian, *Nat. Chem. Biol.,* 2007, 3, 218; S. Kumari, A. Bugaut and S. Balasubramanian, *Biochemistry,* 2008, 47, 12664; M. Wieland and J. S. Hartig, *Chem. Biol,* 2007, 14, 757; A. Arora, M. Dutkiewicz, V. Scaria, M. Hariharan, S. Maiti and J. Kurreck, *RNA,* 2008, 14, 1290; M. J. Morris and S. Basu, *Biochemistry,* 2009,48, 5313.

4. 4 C. M. Azzalin, P. Reichenbach, L. Khoriauli, E. Giulotto and J. Lingner, *Science,* 2007, 318, 798; S. Schoeftner and M. A. Blasco, *Nat. Cell Biol,* 2008, 10, 228; B. Luke and J. Lingner, *EMBO J.,* 2009, 28, 2503; S. Schoeftner and M. A. Blasco, *EMBO J.,* 2009, 28, 2323.

5. 5 A. Randall and J. D. Griffith, *J. Biol. Chem.,* 2009, 284, 13980.

6. 6 Y. Xu, K. Kaminaga and M. Komiyama, *J. Am. Chem. Soc,* 2008, 130, 11179; Y. Xu, K. Kaminaga and M. Komiyama, *Nucleic Acids Symp. Ser.,* 2008, 52, 175.

7. 7 H. Martadinata and A. T. Phan, *J. Am. Chem. Soc,* 2009, 131, 2570.

8. 8 G. N. Parkinson, M. P. Lee and S. Neidle, *Nature,* 2002, 417, 876.

9. 9 A. Joachimi, A. Benz and J. S. Hartig, *Bioorg. Med. Chem.,* 2009, 17, 6811.

10. 10 *(a) G.* N. Parkinson, F. Cuenca and S. Neidle, *J. Mol. Biol.,* 2008, 381, 1145; *(b)* F. Cuenca, O. Greciano, M. Gunaratnam, S. Haider, D. Munnur, R. Nanjunda, W. D. Wilson and S. Neidle, *Bioorg. Med. Chem. Lett.,* 2008, 18, 1668; (c) M. Gunaratnam, S. Swank, S. M. Haider, K. Galesa, A. P. Reszka, M. Beltran, F. Cuenca, J. A. Fletcher and S. Neidle, *J. Med. Chem.,* 2009, 52, 3774.

11. 11 M. Read, R. J. Harrison, B. Romagnoli, F. A. Tanious, S. H. Gowan, A. P. Reszka, W. D. Wilson, L. R. Kelland and S. Neidle, *Proc. Natl. Acad. Sci. U. S. A.,* 2001, 98, 4844; A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. Moore, J. A. Double and S. Neidle, *Cancer Res.,* 2005, 65, 1489; M. Gunaratnam, O. Greciano, C. Martins, A. P. Reszka, C. M. Schultes, H. Morjani, J. F. Riou and S. Neidle, *Biochem. Pharmacol.,* 2007, 74, 679.

12. 12 N. H. Campbell, G. N. Parkinson, A. P. Reszka and S. Neidle, *J. Am. Chem. Soc,* 2008, 130, 6722.

13. 13 S. Paramasivan, I. Rujan and P. H. Bolton, *Methods,* 2007, 43, 324.

14. 14 A. T. Phan, V. Kuryavyi, K. N. Luu and D. J. Patel, *Nucleic Acids Res.,* 2007, 35, 6517; A. Ambrus, D. Chen, J. Dai, T. Bialis, R. A. Jones and D. Yang, *Nucleic Acids Res.,* 2006, 34, 2723; K. W. Lim, S. Amrane, S. Bouaziz, W. Xu, Y. Mu, D. J. Patel, K. N. Luu and A. T. Phan, *J. Am. Chem. Soc,* 2009, 131, 4301; D. Renciuk, I. Kejnovská , P. Skoláková , K. Bednárová , J. Motlová and M. Vorlicková , *Nucleic Acids Res.,* 2009, 37, 6625.

15. 15 J.-L. Mergny, A. T. Phan and L. Lacroix, *FEBS Lett.,* 1998, 435, 74; P. A. Rachwal and K. R. Fox, *Methods,* 2007, 43, 291.

16. 16 B. Guyen, C. M. Schultes, P. Hazel, J. Mann and S. Neidle, *Org. Biomol. Chem.,* 2004, 2, 981.

17. 17 F. Rosu, V. Gabelica, K. Shin-ya and E. De Pauw, *Chem. Commun.,* 2003, 2702; N. Smargiasso, F. Rosu, W. Hsia, P. Colson, E. S. Baker, M. T. Bowers, E. De Pauw and V. Gabelica, *J. Am. Chem. Soc,* 2008, 130, 10208.

18. 18 F. Rosu, E. De Pauw and V. Gabelica, *Biochimie,* 2008, 90, 1074.

19. 19 The AMBER 9.0 suite was used (W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell and P. A. Kollman, *J. Am. Chem. Soc,* 1995, 117, 5179http://ambermd.org). Positions of the $K^+$ ions located within the central channel in the crystal structures were retained. Additional $K^+$ ions were added such that the net —ve charge was zero. The complex was solvated in a waterbox whose boundary was at least 10 A away from the solute. Equilibration involved 220 ps of minimisation and dynamics followed by 20 ns of unrestrained MD simulation using the ParmBSC nucleic acid forcefield.

20. 20 J.-L. Mergny, A. De Cian, A. Ghelab, B. Saccá and L. Lacroix, *Nucleic Acids Res.,* 2005, 33, 81.

21. 21 S. M. Haider, G. N. Parkinson and S. Neidle, *J. Mol. Biol,* 2003, 326, 117; N. H. Campbell, M. Patel, A. Tofa, R. Ghosh, G. N. Parkinson and S. Neidle, *Biochemistry,* 2008, 130, 6722.