Controlled-folding of a small molecule modulates DNA G-quadruplex recognition†

Sebastian Müller, G. Dan Pantoş, Raphaël Rodriguez and Shankar Balasubramanian*

Received (in Cambridge, UK) 26th September 2008, Accepted 24th October 2008
First published as an Advance Article on the web 12th November 2008
DOI: 10.1039/b816861j

Differential recognition of diverse G-quadruplex structures can be achieved by controlling the folding of a small molecule.

The human genome contains guanine-rich sequences which have been shown to fold into G-quadruplexes *in vitro*. These structures have been identified in biologically relevant regions of the genome, such as the promoters of various proto-oncogenes that include, among others c-kit,[1] c-myc,[2] K-ras,[3] and at telomeres.[4] These particular structures have been associated with function and have thus emerged as motifs to be targeted with small molecules.[2,5] G-quadruplexes vary in loop length and sequence, providing the opportunity for differential molecular recognition. There is a need to design small molecules that exhibit specificity with respect to duplex DNA and also between different quadruplexes. Numerous G-quadruplex binders have been reported that generally comprise an aromatic core with side chains to improve molecular recognition.[6] The specificity of such molecules has been enhanced by varying the nature,[6] position[7] and stereochemistry[8] of the appendages, but the challenge to improve specificity between different quadruplexes still remains. Herein, we describe a new G-quadruplex ligand family for which we demonstrated that controlled-folding of the main scaffold can control specificity between the aforementioned diverse nucleic acid structures.

We designed a new class of small molecules (Fig. 1) based on known structural features of classical quadruplex binding ligands.[9] We included two urea functionalities which link two quinolines—moieties that have been shown to enhance the binding of other quadruplex ligands—to a central benzene core.[5d,10] The urea bonds confer a degree of conformational freedom to these molecules.[11] Furthermore, we introduced methyl substituents to the central benzene ring of some derivatives to favour particular conformations as a consequence of steric clash with the oxygens of the urea bonds.[12]

The molecules were synthesised using a short and high yielding synthetic procedure (see ESI†). 2-Amino-quinolinone was reacted with *N*-boc-ethanolamine or *N*-(2-hydroxyethyl)-pyrrolidine under Mitsunobu reaction conditions[11b] to introduce the amine side chains. The products were then reacted with commercially available 1,3-phenylene diisocyanate, 2,4-toluene diisocyanate or 2,6-toluene diisocyanate giving

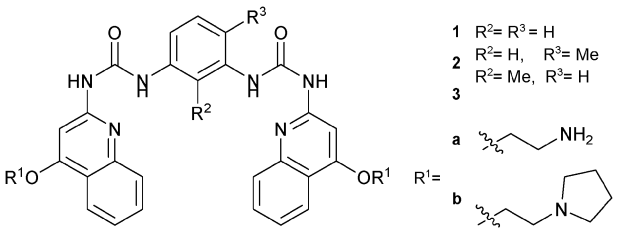

Fig. 1 Molecular structure of the ligands.

rise to the expected bis-ureas. Two analogous series were synthesised, one with a primary amine appendage **a**, and the other with a pyrrolidine appendage **b**. This would allow us to also assess whether the nature of the side-chain could influence the nucleic acid recognition properties of differently folded ligands.

To determine how the methyl substitution would affect the folded state of the molecule, we performed molecular modelling on molecules **1a**, **2a** and **3a** using the PM3 semi-empirical method including a solvent model, followed by 1D and 2D NOESY NMR experiments (see ESI).† According to molecular modeling, **1a** adopts the coiled conformation as depicted in Fig. 2.

In contrast, **2b** shows a twisted conformation, with one of the quinolines pointing upwards; a conformation that was supported by the presence of nOe cross peaks between the hydrogens of the methyl substituent and one of the hydrogens on one of the quinolines as shown in Fig. 3.

Molecule **3a** was predicted to have a W-shaped conformation that drastically differs from the conformation of **1a** and **2a**. This was also supported by nOe cross peaks between the hydrogens of the methyl substituent and one of the hydrogens on both quinolines as depicted in Fig. 4.

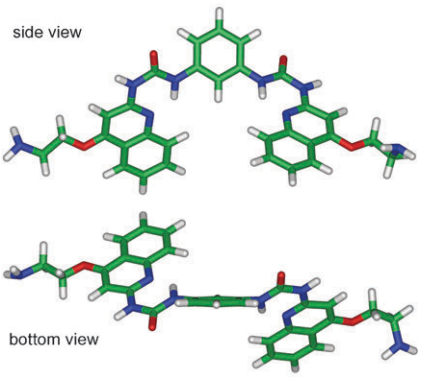

Fig. 2 Coiled conformation adopted by **1a**.

The University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge, CB2 1EW, UK.
E-mail: sb10031@cam.ac.uk; Fax: +44 (0)1223 336913
† Electronic supplementary information (ESI) available: Further experimental details, NOESY spectra, FRET-melting experiments and HyperChem® modeling. See DOI: 10.1039/b816861j

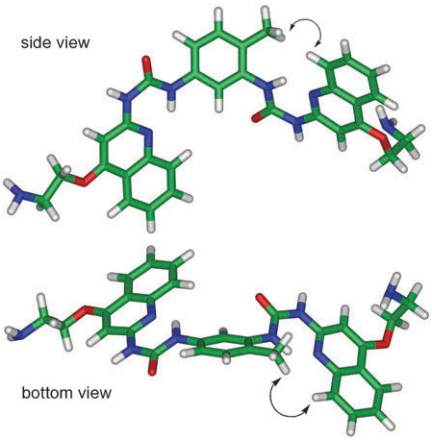

Fig. 3 Twisted conformation adopted by **2a**. The nOe cross peaks are indicated with arrows.

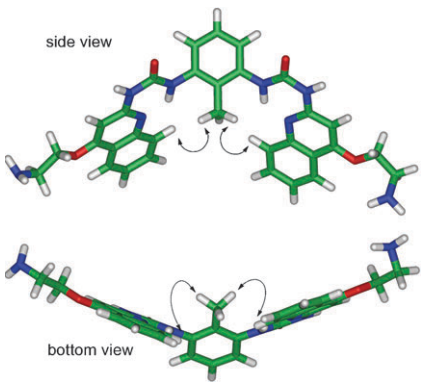

Fig. 4 W-Shaped conformation of **3a**. The nOe cross peaks are indicated with arrows.

These structural data support our hypothesis that conformation can be controlled by the methyl substitution on the benzene ring.

We next explored how distinct folded conformations of these molecules influenced their ability to interact with DNA. Small molecule–DNA interactions were evaluated, using fluorescence resonance energy transfer (FRET) melting experiments, that provide a measure of the ability of a molecule to stabilise the folded form of a DNA structure.[13] We evaluated the aptitude of the small molecules to stabilise an array of different biologically relevant quadruplex forming DNA sequences, occurring in the promoter regions of c-kit (which includes two quadruplex forming sequences: c-kit1 and c-kit2), c-myc, K-ras and the human telomeric quadruplex h-Telo. Fig. 5–7 show thermal shift profiles in which the increase in the transition temperature (ΔT_m) is plotted as a function of ligand concentration. Each DNA target exhibits a different basal T_m value,[14] thus one cannot directly compare absolute ΔT_m values for these targets. We found that different folded molecules of the family induced distinct profiles, suggesting that the folding of these ligands influences their molecular mode of recognition. The ligands do not adopt a planar conformation, and are thus unlikely to interact with the quadruplexes *via* stacking, but may rather interact with the loops of different G-quadruplexes. The selectivity observed

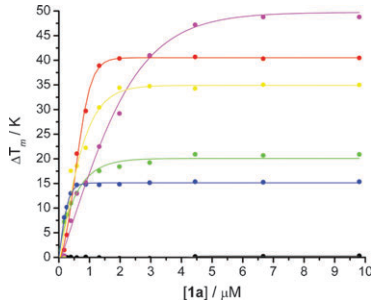

Fig. 5 ΔT_m at different concentrations of ligand **1a**: (pink) K-ras, (red) c-kit1, (yellow) h-Telo, (green) c-kit2, (blue) c-myc, (black) ds-DNA.

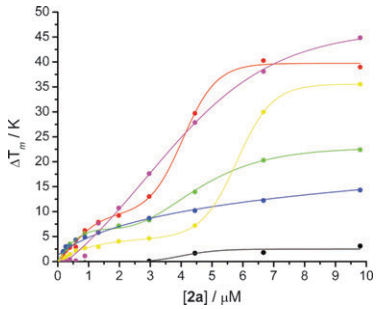

Fig. 6 ΔT_m at different concentrations of ligand **2a**: (pink) K-ras, (red) c-kit1, (yellow) h-Telo, (green) c-kit2, (blue) c-myc, (black) ds-DNA.

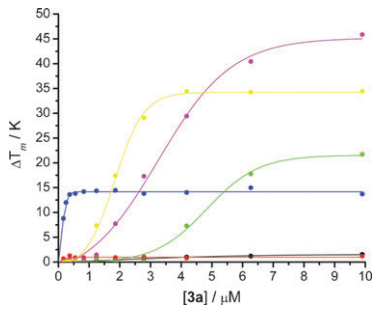

Fig. 7 ΔT_m at different concentrations of ligand **3a**: (pink) K-ras, (red) c-kit1, (yellow) h-Telo, (green) c-kit2, (blue) c-myc, (black) ds-DNA.

may be the result of the diversity observed between these loop sequences and structures.[15] None of the ligands showed any detectable stabilisation of ds-DNA at any of the concentrations measured, consistent with a high level of discrimination between quadruplex and double-stranded DNA.

Compound **1a**, that has a coiled conformation, showed very high changes in melting temperature for all the quadruplexes studied, as displayed in Fig. 5 and Table 1.

This suggests very high stabilisation potentials but little discrimination between different G-quadruplexes. The stabilisation of the human telomeric quadruplex with a ΔT_m of 27.3 K is in the range of the most potent ligands reported in the literature; such as telomestatin and a quinoline macrocycle, which show ΔT_m values of 30.3 K and 33.8 K, respectively.[10] It is noteworthy, that this ligand showed the highest ΔT_m that can be measured by this method for all quadruplexes tested,[14] which is visualised by the flattening of the curves in Fig. 5. The

Table 1 Concentrations of ligands **1a**, **1b**, **2a**, **2b**, **3a** and **3b** in μM at half maximal ΔT_m (with ΔT_m at 1 μM ligand in K shown in parentheses)[17]

	1a	2a	3a	1b	2b	3b
K-ras	1.49 (16.7)	4.00 (4.0)	3.68 (2.7)	2.00 (16.0)	3.56 (9.4)	1.47 (20.8)
c-kit1	0.62 (34.0)	3.74 (6.6)	>10.0 (0.0)	1.07 (19.8)	0.88 (20.1)	>10.0 (0.7)
h-Telo	0.53 (27.3)	5.59 (2.8)	1.89 (3.9)	1.00 (16.7)	0.87 (20.0)	0.48 (25.9)
c-kit2	0.38 (16.2)	3.41 (4.2)	4.75 (1.0)	1.16 (9.3)	1.86 (6.1)	0.46 (14.9)
c-myc	0.17 (17.7)	2.25 (6.2)	0.16 (14.1)	1.00 (7.5)	1.03 (7.1)	0.25 (13.8)
ds-DNA	>10.0 (0.0)	>10.0 (0.0)	>10.0 (0.0)	>10.0 (0.0)	>10.0 (0.5)	>10.0 (0.4)

concentrations to reach this saturation were in the low micromolar range: 0.8 μM for c-myc, 2.0 μM for c-kit 1, 2.5 μM for c-kit2, 2.8 μM for h-Telo and 5 μM for K-ras.

Ligand **2a**, with its twisted conformation, did not show a strong stabilisation potential for any of the quadruplexes investigated at low ligand concentrations, but showed very good stabilisation potential at higher ligand concentrations as seen in Fig. 6.

Much higher ligand concentrations were required to reach saturation of the curves compared to **1a**, making this ligand less potent.[16]

In contrast, **3a** with its W-shaped structure, showed selective stabilisation for a particular G-quadruplex. It stabilised the c-myc quadruplex with a ΔT_m of 14.1 K at a concentration of 0.3 μM as depicted in Fig. 7. Ligand **3a** stabilised the other sequences with a very low ΔT_m at low ligand concentrations (*i.e.* <1 μM). Similar ligand concentrations for **3a** compared to **2a** were required to reach the maximum measurable stabilisation for K-ras and c-kit2. However, ligand **3a** was more potent for h-Telo than **2a** as it reached the maximal ΔT_m at ligand concentrations of 3 μM compared to 8 μM. It is noteworthy that **3a** does not stabilise the c-kit1 quadruplex at any of the ligand concentrations used. To the best of our knowledge, this is the first example of a small molecule that exhibits significant stabilisation potential for one or more quadruplexes, while showing no detectable stabilisation for another. These results show that it is possible to achieve excellent quadruplex discrimination by controlling the folding of a particular molecular scaffold.

Changing the primary amine of the side chain to pyrrolidine did alter the quadruplex stabilisation potential. It decreased the potency of ligands of type **1** but improved the molecular recognition properties for molecules of type **2** and **3**. The c-kit1 quadruplex was not stabilised by **3b**, a feature displayed by the ligands of the family possessing a W-shaped structure. However, this molecule showed better molecular recognition of the K-ras (ΔT_m of 20.8 K), h-Telo (ΔT_m of 25.9 K) and c-kit2 (ΔT_m of 14.9 K) quadruplexes and still showed good stabilisation of c-myc with a ΔT_m of 13.8 K. These results show that it is possible to combine a controlled-folding approach with side-chain variation to fine-tune the molecular recognition properties of a particular scaffold.

In conclusion, we have described a novel small molecule family that significantly stabilises an array of quadruplexes. Some of the small molecules compare favourably to the most potent ligands reported in the literature. We demonstrated how controlling the ligand conformation has a dramatic effect on the molecular recognition properties of a molecule, and serves as a new approach to achieve differential recognition between G-quadruplexes.

We thank Cancer Research UK for programme funding and for a studentship (S. M.), and Pembroke College for a Stokes Junior Research Fellowship (G. D. P.). We are thankful to Prof. J. K. M. Sanders for helpful discussions.

Notes and references

1 (*a*) S. Rankin, A. P. Reszka, J. Huppert, M. Zloh, G. N. Parkinson, A. K. Todd, S. Ladame, S. Balasubramanian and S. Neidle, *J. Am. Chem. Soc.*, 2005, **127**, 10584; (*b*) H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle and S. Balasubramanian, *Biochemistry*, 2006, **45**, 7854.

2 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593.

3 S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536.

4 A. M. Zahler, J. R. Williamson, T. R. Cech and D. M. Prescott, *Nature*, 1991, **350**, 718.

5 (*a*) D. Gomez, M.-F. O'Donohue, T. Wenner, C. Douarre, J. Macadré, P. Koebel, M.-J. Giraud-Panis, H. Kaplan, A. Kolkes, K. Shin-ya and J.-F. Riou, *Cancer Res.*, 2006, **66**, 6908; (*b*) M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 12926; (*c*) 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; (*d*) R. Rodriguez, S. Müller, J. A. Yeoman, C. Trentesaux, J.-F. Riou and S. Balasubramanian, *J. Am. Chem. Soc.*, 2008, DOI: 10.1021/ja805615w.

6 A. Bugaut, K. Jantos, J.-L. Wietor, R. Rodriguez, J. K. M. Sanders and S. Balasubramanian, *Angew. Chem., Int. Ed.*, 2008, **47**, 2677.

7 G. Zagotto, C. Sissi, S. Moro, D. Dal Ben, G. N. Parkinson, K. R. Fox, S. Neidle and M. Palumbo, *Bioorg. Med. Chem.*, 2008, **16**, 354.

8 K. Jantos, R. Rodriguez, S. Ladame, P. S. Shirude and S. Balasubramanian, *J. Am. Chem. Soc.*, 2006, **128**, 13662.

9 D. Monchaud and M.-P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627.

10 P. S. Shirude, E. R. Gillies, S. Ladame, F. Godde, K. Shin-Ya, I. Huc and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 11890.

11 (*a*) V. Berl, I. Huc, R. G. Khoury and J.-M. Lehn, *Chem.–Eur. J.*, 2001, **7**, 2810; (*b*) H. Jiang, J.-M. Leger and I. Huc, *J. Am. Chem. Soc.*, 2003, **125**, 3448; (*c*) L. Xing, U. Ziener, T. C. Sutherland and L. A. Cuccia, *Chem. Commun.*, 2005, 5751.

12 S. Yagi, I. Yonekura, M. Awakura, M. Ezoe and T. Tagagishi, *Chem. Commun.*, 2001, 557.

13 J.-L. Mergny and J.-C. Maurizot, *ChemBioChem*, 2001, **2**, 124.

14 The T_m of the quadruplexes in 60 mM K^+ in the absence of ligand are: 45.9 °C (K-ras), 54.5 °C (c-kit 1), 72.7 °C (c-kit 2), 79.3 °C (c-myc), 59.4 °C (h-Telo), 62.7 °C ds-DNA; the life cycler can measure temperatures up to 94.0 °C.

15 As a recent example of a small molecule that interacts with G-quadruplex loops, see: G. N. Parkinson, R. Ghosh and S. Neidle, *Biochemistry*, 2007, **46**, 2390–2397.

16 For the h-Telo, c-kit1 and c-kit2 quadruplexes the thermal shifts exhibited a biphasic profile which may be indicative of a different mode of binding at higher ligand concentrations.

17 The values are an average of three independent measurements.

