G-quadruplex recognition by bis-indole carboxamides†

Jyotirmayee Dash, Pravin S. Shirude and Shankar Balasubramanian*

Received (in Cambridge, UK) 9th April 2008, Accepted 29th April 2008
First published as an Advance Article on the web 5th June 2008
DOI: 10.1039/b806042h

Herein we report the *de novo* design and synthesis of a geometrically flexible bis-indole carboxamide and a constrained derivative, as a novel class of small molecule scaffold that exhibits high stabilization potential for DNA G-quadruplex sequences associated with the promoters of c-kit2 and c-myc.

Targeting G-quadruplex nucleic acids with small molecules is emerging as a potential strategy for anti-cancer drug design.[1,2] G-quadruplex sequence motifs are widespread in genomic regions such as telomeres,[3] and in gene promoters.[4,5] Small molecules that selectively bind and stabilize the telomeric quadruplex can inhibit telomerase, an enzyme up-regulated in cancer cells.[2,3] There is evidence that the binding of small molecules to the G-quadruplex associated with proto-oncogenes including c-myc,[5a] KRAS,[5b] PDGF-A[5c] and c-kit,[6,7] can modulate transcription. Recent data suggest that RNA G-quadruplex in the 5′ untranslated regions may also offer another target for small molecule intervention.[8]

Significant progress has been made during the past few years in the design of small molecules that target G-quadruplex DNA, and a number of classes of ligands have been reported.[9] Most of them comprise a planar, aromatic core presumed to stack on the terminal tetrads of G-quadruplex. The scope and utility of G-quadruplex selective molecules were further broadened by a very recent report on a nucleic acid catalyst based on small molecule–G-quadruplex recognition.[10] Herein we report the design, synthesis and biophysical evaluation of a conformationally flexible bis-indole carboxamide and a constrained analogue as a new class of G-quadruplex binding ligands (Fig. 1).

Indole derivatives are privileged scaffolds for medicinal chemistry,[11] thus we envisioned that designing G-quadruplex ligands based on such a structure might lead to novel non-toxic compounds suitable for cell biology. Moreover, compounds **1** and **2** exhibit a planar central core,[12] having comparable molecular size to the G-quartet (Supporting Information, S.I.†).[13] Another relevant feature in our design is the potential for multiple H-bonding sites in the flat bis-indole core of **1** and **2**. Ligand **2** may contain internally organized H-bonds between the NH of the indole ring and the pyridine lone pair electrons, whereas ligand **1** can adopt different conformations due to the geometrical freedom (Fig. 1, S.I.†).

Department of Chemistry, University of Cambridge, Lensfield Road, Cambridge, UK CB2 1EW. E-mail: sb10031@cam.ac.uk; Fax: +44 1223 336913; Tel: +44 1223 336347

† Electronic supplementary information (ESI) available: Experimental procedure for FRET, CD and NMR data of indole amides. See DOI: 10.1039/b806042h

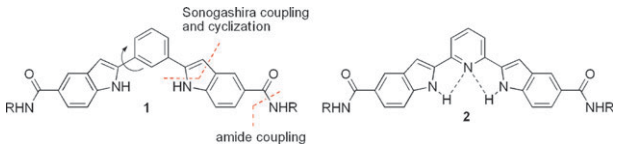
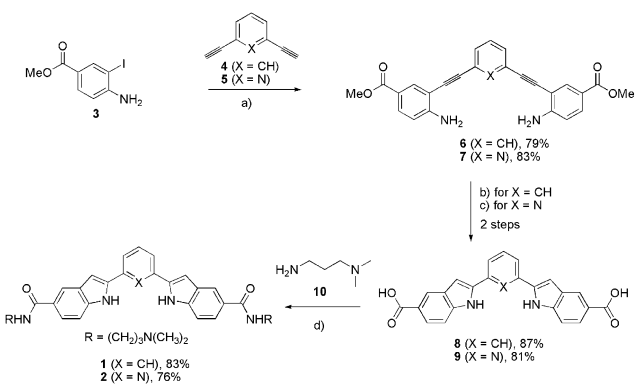


Fig. 1 Design of cationic bis-indole derivatives **1** and **2**.

The 2,6-bis(2′-indolyl)pyridine core of ligand **2** has been used as a tridentate ligand[12,14] and as synthetic receptors.[14] While several methodologies exist for the construction of indole ring systems,[11a] only a few approaches[12,14,15] have been reported for the synthesis of the 2,6-bis(2′-indolyl)pyridine core of ligand **2**. Our synthesis was based on the application of Sonogashira coupling,[16] and 5-*endo-dig* cyclization[17] (Fig. 1, Scheme 1), and would be amenable to the preparation of diverse functionalized indole derivatives. The Sonogashira coupling reaction of iodoaniline derivative **3** with 1,3-diethynyl-benzene (**4**) and 2,6-diethynyl-pyridine (**5**) proceeds with excellent yield to afford the corresponding diester derivatives **6** and **7**. The diacids **8** and **9** were acquired in high yields over two steps involving a 5-*endo-dig* type cyclization[18] and base mediated hydrolysis of the corresponding indole diester intermediates. Treatment of diacids **8** and **9** with amine **10** gave the corresponding bis-amides **1** and **2**. The syntheses of **1** and **2** were efficiently achieved using 4-step protocols in 57% and 51% overall yield respectively starting from easily accessible starting materials **3–5** (Scheme 1).

We evaluated the ability of ligands **1** and **2** to interact with four different G-quadruplex forming oligonucleotides based on human genomic sequences using fluorescence resonance energy transfer (FRET) melting.[19] FRET melting analysis determines the ligand-induced stabilization of a folded quadruplex by measurement of the ligand induced increase in the melting temperature (ΔT_m). The G-quadruplex forming



a) $PdCl_2(PPh_3)_2$, CuI, Et_3N, DMF, rt, 9 h; b) 1) $NaAuCl_4.2H_2O$, EtOH, 70 °C, 10 h, 2) NaOH (2N), THF:MeOH (1:1), rt, 18 h; c) 1) CuI, NMP, 190 °C, microwave, 1 h; 2) NaOH (2N), THF:MeOH (1:1), rt, 18 h; d) **10**, EDC, HOBT, NMM, DMF, rt, 12 h.

Scheme 1 Synthesis of water soluble bis-indole derivatives **1** and **2**.

Table 1 DNA G-quadruplexes and duplex stabilization by **1** and **2** determined by FRET melting experiments

Ligand	ΔT_m at 1 μM ligand concentration (°C)				
	h-telo (59 ± 1)[a]	c-kit1 (57 ± 1)[a]	c-kit2[b] (71 ± 1)[a]	c-myc[b] (77 ± 1)[a]	dup (58 ± 1)[a]
1	21.5	27.7	21.7[b]	16.5[b]	4.2
2	9.8	22.1	16.0	10.5	0.5

[a] T_m in 60 mM potassium cacodylate buffer, pH 7.4 without ligand. [b] Maximum measurable ΔT_m that can be measured for c-kit2 is 24 °C (*i.e.* T_m = 95 °C) and for c-myc is 18 °C (*i.e.* T_m = 95 °C).

sequences comprised the human telomeric G-quadruplex DNA (h-telo, *FAM*-d($G_3[T_2AG_3]_3$)-*TAMRA*), two G-quadruplexes found in the promoter of c-kit (c-kit1,[6b] *FAM*-d($G_3AG_3CGCTG_3AG_2AG_3$)-*TAMRA*; and c-kit2,[6a] *FAM*-d($G_3CG_3CGCGAG_3AG_4$)-*TAMRA*) and the G-quadruplex found in the promoter of c-myc[5a] (c-myc, *FAM*-d($TGAG_3TG_3TAG_3TG_3TA_2$)-*TAMRA*). We have also evaluated the melting of a duplex DNA (dup, *FAM*-d(TATAGC-TATA-HEG-TATAGCTATA)-*TAMRA*) as a non-quadruplex control. The quadruplex melting data are summarized in Table 1 and Fig. 2 (S.I.†).

Ligand **1** exhibits a high stabilization potential for the c-kit2 quadruplex (ΔT_m = 21.7 °C at 1 μM, *i.e.* a T_m of 93 °C) comparable in magnitude to data obtained for the natural product telomestatin (ΔT_m = 20.4 °C at 1 μM) and for a quinoline macrocycle (ΔT_m = 21.4 °C at 1 μM), which are two of the best reported quadruplex ligands.[20] The stabilization potential of **1** for c-myc also saturates at 1 μM concentration (ΔT_m = 16.5 °C, *i.e.* T_m = 93.5 °C). Only a few ligands are reported to interact and modulate c-myc G-quadruplex.[9] The ΔT_m of ligand **1** for the c-kit1 quadruplex (27.7 °C) was found to be comparable with that of one of the best ligands (27.1 °C) from our recently reported isoalloxazine series, that also shows biological activity against c-kit.[7] It is interesting to note the comparative ΔT_m values for ligands **1** and **2** (phenylene *vs.* pyridyl). While ligand **1** shows detectable duplex stabilization (ΔT_m = 4.2 °C), diminished duplex stabilization (ΔT_m = 0.5 °C) was observed for the more geometrically constrained ligand **2**. Similarly there is a large variation of ΔT_m of ligand **1** and **2** for h-telo quadruplex. The observed ΔT_m of ligand **1** for h-telo was 21.5 °C. Ligand **2** shows stronger stabilization for c-kit1 (ΔT_m = 22.1 °C) and c-kit2 (ΔT_m = 16.0 °C) as compared to h-telo (ΔT_m = 9.8 °C) G-quadruplex DNA at 1 μM concentration.

Circular dichroism (CD) spectroscopy on DNA sequences known to fold into G-quadruplexes can be used to monitor the formation of G-quadruplex structure.[21] The real time CD titrations of **1** and **2** were conducted for h-telo, c-kit1, c-kit2 and c-myc quadruplex forming DNA sequences at room temperature in the absence of added stabilizing salts[22] (S.I.†) to find out any selectivity for a particular G-quadruplex conformation. The c-kit2 quadruplex exists as a predominantly parallel structure with a small proportion of an antiparallel conformation.[6a] The CD spectrum of c-kit2 quadruplex in the absence of any added salt shows a major positive peak at 260 nm and a minor peak at 293 nm, suggestive of a mixture of parallel and antiparallel structures.[6a,20] Upon titration with ligand **2** in real time‡ (Fig. 3B), it was noteworthy in the CD spectroscopic analysis that the ligand dependent increase in ellipticity at 260 nm is not accompanied by an associated increase in the peak at 293 nm, supportive of selective induction of a parallel G-quadruplex structure.§ The effect was more prominent at high ligand concentration with the disappearance of the peak at 293 nm. When the titration experiments were carried out with ligand **1**, the peaks at both 260 nm and 293 nm were increased, suggesting that ligand **1** does not show any preference to induce folding of a particular quadruplex structure (Fig. 3A). In contrast, for the cases of h-telo, c-kit1 and c-myc an incremental ligand dependent induction in G-quadruplex was not significant and an apparent quadruplex unfolding was observed at high ligand concentration (S.I.†).

We have developed an expedient synthetic route to water soluble bis-indole amides, that are a novel class of *de novo* G-quadruplex binding small-molecule ligands. These ligands have shown a level of G-quadruplex stabilization that compares favourably with the best reported G-quadruplex ligands. CD spectroscopic analysis further reveals strong interaction of the ligands with c-kit2 G-quadruplex DNA. The properties of these ligands make them attractive probes to explore hypotheses for the biological function of G-quadruplexes and such investigations are currently underway.

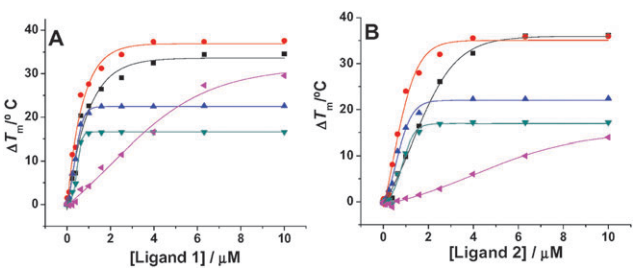


Fig. 2 FRET stabilization curves for ligands (A) **1** and (B) **2** upon binding to h-telo (■), c-kit1 (●), c-kit2 (▲), c-myc (▲) and double stranded (▲) DNA.

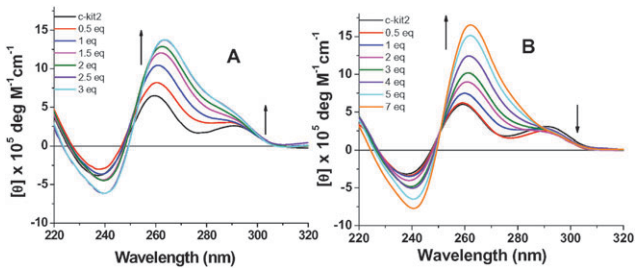


Fig. 3 CD spectra of a 12.5 μM solution of c-kit2 d[$G_3CG_2GCGCGAG_3AG_4$] in Tris buffer (pH 7.4), 0–7 equivalent of ligands (A) **1** and (B) **2**.

We thank the BBSRC for project funding, Cancer Research UK for programme funding and the European Commission for a Marie Curie Incoming Postdoctoral Fellowship to JD.

Notes and references

‡ CD spectra of ligand **2** annealed with c-kit2 quadruplex and CD titration of ligand **2** with c-kit2 in the presence of K^+ are presented in S.I.†

§ The UV/Vis spectrum for ligand **1** shows absorption at 257, 272, 323 nm, and that of ligand **2** shows absorption at 264, 341 nm. Upon titration with G-quadruplex DNA, both the ligands show peaks at 260 nm. We cannot rule out the possibility that ICD bands belonging to **1** and **2** (at 272 and 264 nm) might overlap in this region.

1. *Quadruplex Nucleic Acids*, ed. S Neidle and S Balasubramanian, Royal Society of Chemistry, Cambridge, UK, 2006.
2. (*a*) A. M. Burger, F. Dai, C. M. Schultes, A. P. Reszka, M. J. B. Moore, J. A. Double and S. Neidle, *Cancer Res.*, 2005, **65**, 1489; (*b*) E. Mandine, J.-F. Riou, J.-L. Mergny, P. Mailliet and F. Boussin, *Oncogene*, 2005, **24**, 2917; (*c*) H. Han and L. H. Hurley, *TiPS*, 2000, **21**, 136.
3. D. Y. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle and L. H. Hurley, *J. Med. Chem.*, 1997, **40**, 2113.
4. (*a*) J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406; (*b*) J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2005, **33**, 2908; (*c*) A. K. Todd, M. Johnston and S. Neidle, *Nucleic Acids Res.*, 2005, **33**, 2901.
5. (*a*) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593; (*b*) S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536; (*c*) Y. Qin, E. M. Rezler, V. Gokhale, D. Sun and L. H. Hurley, *Nucleic Acids Res.*, 2007, **35**, 1.
6. (*a*) H. Fernando, A. P. Reszka, J. L. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle and S. Balasubramanian, *Biochemistry*, 2006, **45**, 7854; (*b*) 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.
7. M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 12926.
8. S. Kumari, A. Bugaut, J. L. Huppert and S. Balasubramanian, *Nat. Chem. Biol.*, 2007, **3**, 218.
9. For reviews, see: (*a*) D. Monchaud and M.-P. Teulade-Fichou, *Org. Biomol. Chem.*, 2008, **6**, 627; (*b*) J. Cuesta, M. A. Read and S. Neidle, *Mini-Rev. Med. Chem.*, 2003, **3**, 11; (*c*) For specific examples see: Z. A. E. Waller, P. S. Shirude, R. Rodriguez and S. Balasubramanian, *Chem. Commun.*, 2008, 1467; (*d*) J. E. Reed, S. Neidle and R. Vilar, *Chem. Commun.*, 2007, 4366; (*e*) D. P. N. Gonçalves, R. Rodriguez, S. Balasubramanian and J. K. M. Sanders, *Chem. Commun.*, 2006, 4685; (*f*) K. Jantos, R. Rodriguez, S. Ladame, P. S. Shirude and S. Balasubramanian, *J. Am. Chem. Soc.*, 2006, **128**, 13662; (*g*) J. Seenisamy, S. Bashyam, V. Gokhale, H. Vankayalapati, D. Sun, A. Siddiqui-Jain, N. Streiner, K. Shin-ya, E. White, W. D. Wilson and L. H. Hurley, *J. Am. Chem. Soc.*, 2005, **127**, 2944.
10. Z. Tang, D. P. N. Gonçalves, M. Wieland, A. Marx and J. S. Hartig, *ChemBioChem*, 2008, **9**, 1061.
11. For reviews, see: (*a*) G. R. Humphrey and J. T. Kuethe, *Chem. Rev.*, 2006, **106**, 2875; (*b*) D. A. Horton, G. T. Bourne and M. L. Smythe, *Chem. Rev.*, 2003, **103**, 893; (*c*) H.-J. Knölker and K. R. Reddy, *Chem. Rev.*, 2002, **102**, 4303.
12. For the X-ray crystal structure of the 2,6-bis(2′-indolyl)pyridine core, see: W.-L. Jia, Q.-D. Liu, R. Wang and S. Wang, *Organometallics*, 2003, **22**, 4070.
13. Molecular dimensions of indole derivative **1** are determined from an energy minimized chem 3D diagram and compared with the X-ray crystal structure of G-tetrad. (*a*) A. D. Cian, E. DeLemos, J.-L. Mergny, M.-P. Teulade-Fichou and D. Monchaud, *J. Am. Chem. Soc.*, 2007, **129**, 1856; (*b*) G. N. Parkinson, M. P. H. Lee and S. Neidle, *Nature*, 2002, **417**, 876.
14. (*a*) Q. Liu, L. Thorne, I. Kozin, D. Song, C. Seward, M. D'Iorio, Y. Tao and S. Wang, *J. Chem. Soc., Dalton Trans.*, 2002, 3234; (*b*) R. P. Thummel and V. Hegde, *J. Org. Chem.*, 1989, **54**, 1720; (*c*) V. Hegde, C.-Y. Hung, P. Madhukar, R. Cunningham, T. Hopfner and R. P. Thummel, *J. Am. Chem. Soc.*, 1993, **115**, 812.
15. (*a*) S. Tollari, S. Cenini, A. Rossi and G. Palmisano, *J. Mol. Catal. A: Chem.*, 1998, **135**, 241; (*b*) G. Palmisano and M. Santagostino, *Helv. Chim. Acta*, 1993, **76**, 2356.
16. (*a*) For reviews, see: K. Sonogashira, in *Metal-Catalyzed Cross-Coupling Reactions*, ed. F Diederich and P. J. Stang, Wiley-VCH, New York, 1998, ch. 5; (*b*) E.-I. Negishi and L. Anastasia, *Chem. Rev.*, 2003, **103**, 1979.
17. For a review, see: T. E. Müller and M. Beller, *Chem. Rev.*, 1998, **98**, 675.
18. (*a*) A. Arcadi, G. Bianchi and F. Marinelli, *Synthesis*, 2004, 610; (*b*) S. E. Pearson and S. Nandan, *Synthesis*, 2005, 2503.
19. (*a*) R. A. J. Darby, M. Sollogoub, C. McKeen, L. Brown, A. Risitano, N. Brown, C. Barton, T. Brown and K. R. Fox, *Nucleic Acids Res.*, 2002, **30**, e39; (*b*) A. De Cian, L. Guittat, M. Kaiser, B. Sacca, S. Amrane, A. Bourdoncle, P. Alberti, M.-P. Teulade-Fichou, L. Lacroix and J.-L. Mergny, *Methods*, 2007, **42**, 183.
20. P. S. Shirude, E. R. Gilles, S. Ladame, F. Godde, K. Shin-ya, I. Huc and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 11890.
21. (*a*) P. Balagurumoorthy and S. K. Brahmachari, *J. Biol. Chem.*, 1994, **269**, 21858; (*b*) R. Jin, B. L. Gaffney, C. Wang, R. A. Jones and K. J. Breslauer, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 8832.
22. (*a*) E. M. Rezler, J. Seenisamy, S. Bashyam, M.-Y. Kim, E. White, W. D. Wilson and L. H. Hurley, *J. Am. Chem. Soc.*, 2005, **127**, 9439; (*b*) R. Rodriguez, G. D. Pantos, D. P. N. Gonçalves, J. K. M. Sanders and S. Balasubramanian, *Angew. Chem., Int. Ed.*, 2007, **46**, 5405; (*c*) B. Fu, J. Huang, L. Ren, X. Weng, Y. Zhou, Y. Du, X. Wu, X. Zhou and G. Yang, *Chem. Commun.*, 2007, 3264.