Click chemistry assembly of G-quadruplex ligands incorporating a diarylurea scaffold and triazole linkers†

William C. Drewe and Stephen Neidle*

Received (in Cambridge, UK) 21st August 2008, Accepted 22nd September 2008
First published as an Advance Article on the web 30th September 2008
DOI: 10.1039/b814576h

A series of diarylurea ligands were designed to interact selectively with G-quadruplexes and were synthesised using copper(I) catalysed 'click' chemistry to incorporate the 1,4-substituted 1,2,3-triazole ring into the core of the ligands; the optimal ligands demonstrate a high degree of selective telomeric G-quadruplex stabilisation and are not cytotoxic in several cancer cell lines.

Guanine (G) rich DNA sequences are able to fold into higher-order DNA structures, G-quadruplexes, which arise from the self-association of guanine DNA bases stabilised by Hoogsteen hydrogen bonds and monovalent cations.[1,2] G-rich DNA sequences and putative G-quadruplexes (G4s) have been located throughout the human genome;[2–5] the most studied sequences are in human telomeres[6–8] and the promoter regions of proto-oncogenes such as *c-kit*[9–11] and *c-myc*.[12,13] Targeting such G4 DNAs with small molecules is a potential route to novel cancer-selective chemotherapeutic intervention, by inhibiting telomerase activity[14] and selectively interfering with telomere maintenance in cancer cells,[15,16] or regulating the transcriptional activity of proto-oncogenes.[13,17]

We report here on the rational design, synthesis and initial evaluation of a series of novel diarylurea-based ligands, which have high telomeric G4 selectivity combined with low affinity for duplex DNA so that short-term acute toxicity to both cancer and normal cells is minimised. Our goal has been to devise ligands that do not have the planar aromatic polycyclic features characteristic of most G4-binding ligands, but instead incorporate more drug-like features. Initial qualitative modelling suggested that the 1,3-diphenyl urea scaffold (as found in a wide range of clinically-approved drugs) would be a suitable pharmacophore. It also indicated that the introduction of the heteroaromatic 1,4-substituted 1,2,3-triazole ring (using click chemistry[18,19]) could enhance G4 affinity/selectivity, as previously found in other series of ligands.[20,21]

Molecular modelling using a docking protocol indicates that these ligands have several low-energy binding conformations when bound onto the exposed G-quartet at the end of a G4 structure, with the urea oxygen atom located over the ion channel and the 1,3-diphenyl urea moiety well positioned for effective π-stacking interactions. These models also indicate that the 1,2,3-triazole heteroaromatic rings could be positioned in an optimal orientation for efficient π-stacking interactions on the G-quartet surface, with triazole–phenyl bond rotation allowing deep penetration of the side chains into the G4 grooves (Fig. 1). The modelling further suggests that the *ortho*- (**1–2**) and *meta*-substituted analogues (**3–9**) can adopt a 'square-planar' conformation that can optimize G4 DNA overlap/affinity, while the *para*-substituted ligands (**10–13**) have a conformation which is sterically too large for optimal interaction between the triazole rings and the G-quartet. The *ortho*-substituted analogues are conformationally constrained by the formation of an intramolecular hydrogen bond between the urea-NH and the triazole-N3 atoms, enforcing the desired 'square-planar' conformation.

A simple, convergent synthetic route to the target ligands has been devised (Scheme 1). The symmetrical 1,3-bis(ethynylphenyl)urea building blocks were synthesised from commercially available ethynylbenzenamines by reaction with 1,1′-carbonyldiimidazole (CDI) in refluxing THF, to give yields of 93–100%. The azide side chains were synthesised from the previously reported amine side chains *via* a 'one-pot' diazotization reaction with subsequent azide substitution,[20,21] to give yields of 26–93%. The target ligands were synthesised from these alkyne and azide building blocks by copper(I) catalysed 'click chemistry', yielding the desired 1,4-substituted 1,2,3-triazole linked products with yields of 56–95%. Each ligand was purified by semi-preparative HPLC to afford samples of suitable purity for biological evaluation (ESI†).

Initially the click reaction was conducted under microwave irradiation at 130 °C for 30 min; however this led to the

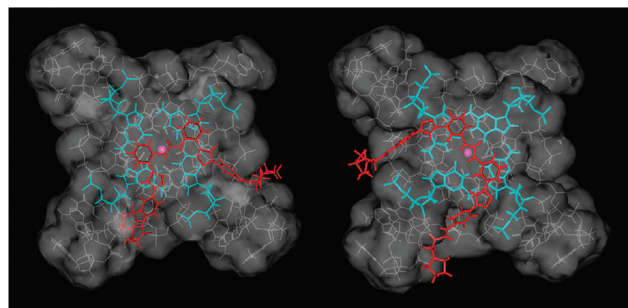
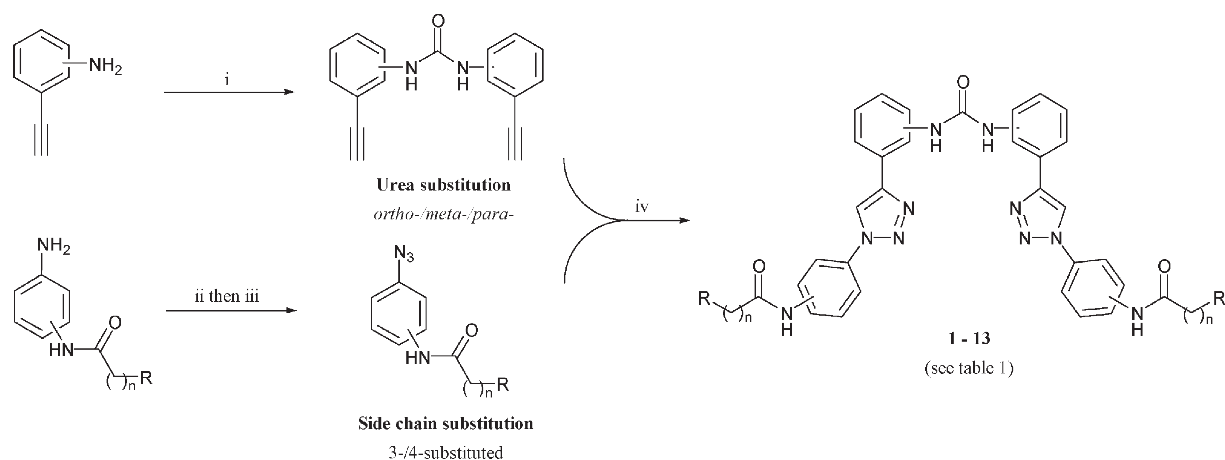


Fig. 1 The low-energy binding conformations of ligands **4** and **7** as suggested by molecular modelling and docking calculations, using the human intramolecular telomeric G4 crystal structure (PDB 1KF1). Ligands are shown in red, the 3′G-quartet in cyan and a K^+ ion in mauve.

CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, 29-39 Brunswick Square, London, UK WC1N 1AX. E-mail: stephen.neidle@pharmacy.ac.uk; Fax: +44 (0) 207 7535970; Tel: +44 (0) 207 7535971

† Electronic supplementary information (ESI) available: Experimental procedures for synthesis, FRET and cell culture; HPLC purity; FRET melting curves. See DOI: 10.1039/b814576h

This journal is © The Royal Society of Chemistry 2008



Scheme 1 The synthetic route to the novel urea-based ligands **1–13** (ESI†); (i) CDI, THF, reflux (93–100%); (ii) ^{t}BuONO, HCl, THF, 0 °C; (iii) NaN_3, H_2O, 0 °C to rt (26–93%; 2 steps); (iv) $CuSO_4$, sodium ascorbate, H_2O–^{t}BuOH, μW or rt (56–95%).

formation of undesired mono-eliminated side products for the $n = 2$ side chain analogues, of which $\mathbf{3_{Elim}}$ and $\mathbf{5_{Elim}}$ were isolated for assessment. This side-chain elimination was however overcome by conducting the reactions at ambient temperature for increased time periods; however under these conditions the *ortho*-substituted 1,3-bis(2-ethynylphenyl)urea proved un-reactive. Hence the $n = 2$ side chain analogues of compounds **1** and **2** were not able to be synthesised by this route.

G4 DNA affinity and selectivity was assessed by *in vitro* fluorescence resonance energy transfer (FRET)-based assays.[22,23] Each ligand was examined at a range of concentrations (ESI†) against the G4-forming oligonucleotide originating from four repeats of human telomeric DNA [$^{5'}$FAM-G_3($TTAG_3$)$_3$-TAMRA$^{3'}$] where FAM is 6-carboxyfluorescein and TAMRA is 6-carboxy-*N*,*N*,*N′*,*N′*-tetramethylrhodamine, in order to measure the change in thermal stability (ΔT_m) produced by ligand binding. The results of these experiments (Table 1; ESI†) showed that the series has a wide range of telomeric G4 DNA affinities, with the best compounds (**5–7**) comparable to the trisubstituted acridine compound BRACO-19.

In terms of urea bond substitution pattern, the *ortho*- (**1–2**) and *meta*-analogues (**3–6**) have significantly enhanced G4 DNA stabilising affinity over the more extended *para*-analogues (**10–13**), suggesting that the 'square-planar' conformation of the ligand core is critical for efficient G4 interaction, consistent with molecular modelling predictions. The *ortho*-substituted analogues (**1–2**) have reduced G4 stabilisation relative to the *meta*-substituted analogues (**4** and **6**).

Alterations of the basic terminus of the side chains show that the pyrrolidino analogue (**5**) had weakly enhanced affinity for G4 DNA over the dimethylamino (**7**) and piperidino (**8**) analogues, with the less basic morpholino analogue (**9**) and mono-eliminated side products ($\mathbf{3_{Elim}}$ and $\mathbf{5_{Elim}}$) proving detrimental to G4 affinity. These results demonstrate the importance of side chain basicity and the requirement of two basic groups for strong G4 interactions.

Table 1 The *in vitro* biological effects of ligands **1–13**[a]

	Substitution pattern				$\Delta Tm_{1\mu M}$		IC_{50}/μM		
Compound	Urea	Side chain	*n*	R	G4 DNA	ds DNA	MCF7	A549	WI38
BRACO-19	N/A	N/A	2	Pyr	25.9	11.2	2	2	11
1	*Ortho-*	3-	3	Pyr	12.0	0.0	23	>25	>25
2	*Ortho-*	4-	3	Pyr	15.9	0.0	>25	>25	>25
3	*Meta-*	3-	2	Pyr	16.6	0.6	>25	>25	>25
$\mathbf{3_{Elim}}$	*Meta-*	3-	2	Pyr/C=C	1.4	0.1	Nd	Nd	Nd
4	*Meta-*	3-	3	Pyr	17.1	0.5	>25	>25	>25
5	*Meta-*	4-	2	Pyr	18.4	1.1	6	>25	>25
$\mathbf{5_{Elim}}$	*Meta-*	4-	2	Pyr/C=C	0.7	0.0	Nd	Nd	Nd
6	*Meta-*	4-	3	Pyr	19.4	1.8	16	>25	>25
7	*Meta-*	4-	2	Dimeth	17.5	0.5	>25	>25	>25
8	*Meta-*	4-	2	Pip	16.4	1.8	4	>25	>25
9	*Meta-*	4-	2	Morpho	4.3	0.4	>25	>25	>25
10	*Para-*	3-	2	Pyr	3.3	0.0	13	>25	>25
11	*Para-*	3-	3	Pyr	4.3	1.0	>25	>25	>25
12	*Para-*	4-	2	Pyr	0.4	0.3	>25	>25	>25
13	*Para-*	4-	3	Pyr	1.7	2.1	>25	>25	>25

[a] Abbreviations: pyrrolidino (pyr), dimethylamino (dimeth), piperidino (pip), morpholino (morpho), mono-eliminated side chain (pyr/C=C), nd is not determined. Notes: $\Delta Tm_{1\mu M}$ is the change in melting temperature at a ligand concentration of 1 μM ± 1 °C. IC_{50} is the ligand concentration required to inhibit cell population growth by 50% ± 1 μM. Maximum solubility of ligands in cell media was 25 μM. The 3,6,9-trisubstituted acridine compound BRACO-19 was used as a control (see ref. 16 and references therein). BRACO-19 IC_{50} data was taken from ref. 16.

 This journal is © The Royal Society of Chemistry 2008

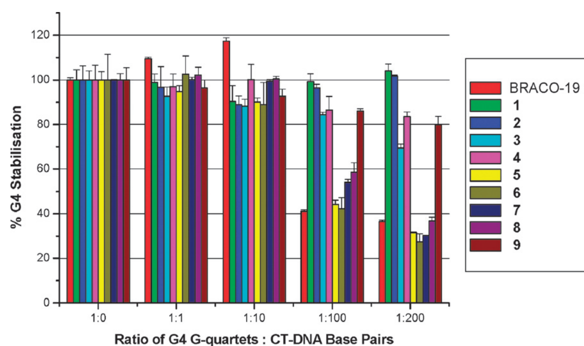


Fig. 2 The FRET-based competition assay where G4 stabilisation is shown as the percentage of $\Delta Tm_{1\mu M}$ retained upon addition of CT-DNA.

G4 *vs.* duplex DNA selectivity was also assessed by FRET-based methods, initially screening each of the ligands against a duplex-forming oligonucleotide containing a range of intercalative binding sites [$^{5'}$FAM-$(TA)_2GC(TA)_2T_6(TA)_2GC$-$(TA)_2$-TAMRA$^{3'}$] (ds DNA). This showed that these ligands have significantly enhanced G4 over duplex stabilisation relative to BRACO-19 (Table 1; ESI†). The G4 selectivity of some of the more potent ligands was further assessed by a FRET-based competition assay[20,24] where the ability of 1 μM of ligand to retain G4 stabilising affinity was challenged by excess ratios of calf thymus DNA (CT-DNA). This shows (Fig. 2) that the *ortho*-substituted ligands (**1**–**2**) retain almost complete G4 selectivity compared to BRACO-19 (∼100% *vs.* ∼40% selective) up to a 200-fold excess of base pairs. In addition, the *meta*-substituted ligands with 3-substituted side chains (**3**–**4**) were significantly more G4 selective (∼70–80%) relative to their 4-substituted analogues (**5**–**6**; ∼30% selective). Altering the basic group of a side chain had little influence on G4 selectivity other than for the morpholino analogue (**9**; ∼80% selective). The combined results of these assays demonstrate the importance of ligand core conformation and side chain substitution pattern for G4 selectivity over duplex DNA, with the *ortho*-urea substitution pattern being optimal.

The effects of compounds in this series on short-term cell proliferation have been examined in two cancer cell lines and one "normal" fibroblast cell line (Table 1). The majority of compounds show no toxic cellular effects up to the limits of aqueous solubility (>25 μM), across all three cell lines. Some activity is observed in the MCF7 breast carcinoma cell line, generally for those compounds that have some effects on duplex DNA. The behaviour of the series is in contrast to that of the trisubstituted acridine compound BRACO-19 and a number of other G-quadruplex ligands. A desirable property of such compounds is that their selective effects on cancer cell growth arise from inhibition of telomerase function and not from more generalised effects on duplex (*i.e.* genomic) DNA.[25] It is also desirable from the perspective of potential therapeutic use that generalised toxicity is minimised. This study has shown that suitable substituents on the diarylurea scaffold can give a combination of high telomeric quadruplex selectivity and low acute toxicity so that for example compounds **2**, **3** or **4** in this present series are suitable for more detailed studies of their effects on cell growth and telomere maintenance. These will be reported elsewhere.

Binding of these diarylureas to non-telomeric quadruplexes, for example the two found in the *c-kit* promoter,[9–11] is very dependent on the nature of the substituents and the particular quadruplex, suggesting that the compounds may be tuneable for a given target quadruplex (to be published).

This work was supported by CRUK and the EU (FP6 Project on Molecular Cancer Medicine). We are grateful to Dr Mekala Gunaratnam for help with cell biology and Dr JE Moses for much useful advice on click chemistry.

Notes and references

1 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402.
2 D. J. Patel, A. T. Phan and V. Kuryavyi, *Nucleic Acids Res.*, 2007, **35**, 7429.
3 J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2005, **33**, 2908.
4 J. L. Huppert and S. Balasubramanian, *Nucleic Acids Res.*, 2007, **35**, 406.
5 A. K. Todd, M. Johnston and S. Neidle, *Nucleic Acids Res.*, 2005, **33**, 2901.
6 G. N. Parkinson, M. P. Lee and S. Neidle, *Nature*, 2002, **417**, 876.
7 A. T. Phan, K. N. Luu and D. J. Patel, *Nucleic Acids Res.*, 2006, **34**, 5715.
8 A. T. Phan, V. Kuryavyi, K. N. Luu and D. J. Patel, *Nucleic Acids Res.*, 2007, **35**, 6517.
9 H. Fernando, A. P. Reszka, J. Huppert, S. Ladame, S. Rankin, A. R. Venkitaraman, S. Neidle and S. Balasubramanian, *Biochemistry*, 2006, **45**, 7854.
10 A. T. Phan, V. Kuryavyi, S. Burge, S. Neidle and D. J. Patel, *J. Am. Chem. Soc.*, 2007, **129**, 4386.
11 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.
12 A. T. Phan, V. Kuryavyi, H. Y. Gaw and D. J. Patel, *Nat. Chem. Biol.*, 2005, **1**, 167.
13 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593.
14 J. Reed, M. Gunaratnam, M. Beltran, A. P. Reszka, R. Vilar and S. Neidle, *Anal. Biochem.*, 2008, **380**, 99.
15 A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J. F. Riou and J. L. Mergny, *Biochimie*, 2008, **90**, 131.
16 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.
17 M. Bejugam, S. Sewitz, P. S. Shirude, R. Rodriguez, R. Shahid and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 12926.
18 H. C. Kolb, M. G. Finn and K. B. Sharpless, *Angew. Chem., Int. Ed.*, 2001, **40**, 2004.
19 J. E. Moses and A. D. Moorhouse, *Chem. Soc. Rev.*, 2007, **36**, 1249.
20 A. D. Moorhouse, A. M. Santos, M. Gunaratnam, M. Moore, S. Neidle and J. E. Moses, *J. Am. Chem. Soc.*, 2006, **128**, 15972.
21 A. D. Moorhouse, S. Haider, M. Gunaratnam, D. Munnur, S. Neidle and J. E. Moses, *Mol. Biosyst.*, 2008, **4**, 629.
22 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.
23 J. L. Mergny and J. C. Maurizot, *ChemBioChem*, 2001, **2**, 124.
24 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.
25 (*a*) P. Phatak and A. M. Burger, *Br. J. Pharmacol.*, 2007, **152**, 1003–1011; (*b*) 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**, 963–975; (*c*) A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou and J.-L. Mergny, *Biochimie*, 2008, **90**, 131–155.