Grover, J.; Trujillo, C.; Saad, M.; Emandi, G.; Stipaničev, N.; Bernhard, S. S. R., Guédin, A.; Mergny, J.-L.; Senge, M. O.; Rozas, I. (2020):

Dual-binding conjugates of diaromatic guanidines and porphyrins for recognition of G-quadruplexes. *Organic & Biomolecular Chemistry* 18, 5617–5624. doi: 10.1039/D0OB01264E

PAPER

Engineering conjugates of diaromatic guanidinium systems and porphyrins for recognition of G-quadruplexes

Received 00th January 20xx,
Accepted 00th January 20xx

DOI: 10.1039/x0xx00000x

Jagdeep Grover,^a Mona Saad,^b Ganapathi Emandi,^a Cristina Trujillo,^a Nikolina Stipaničev,^a Stefan S. R. Bernhard,^a Jean-Louis Mergny,^b Mathias O. Senge,^a Isabel Rozas^{a,*}

A firstly conceptualised class of dual-binding guanine quadruplexes binders has been designed, synthesised and biophysically studied. These compounds, which combine diaromatic guanidinium systems and a classical binding moiety for guanine quadruplexes (porphyrins), discriminate between guanine quadruplexes and double-stranded DNA. An extensive screening of a variety of guanine quadruplex structures and double stranded DNA *via* UV-vis, FRET and CD experiments revealed the conjugates preference towards guanine quadruplexes.

Introduction

The human genome accommodates hundreds of thousands of potential guanine quadruplex (G4) forming sequences.¹ Notably, these G-rich sequences are prevalent at chromosomal ends, telomeres,² and in promoter regions of some oncogenes such as *BCL2*,³ *KRAS*,⁴ *c-MYC*⁵ and *c-KIT*.⁶ These proto-oncogenes code for different protein kinases, which, upon activation, become engaged in an array of physiological processes such as proliferation, migration, maturation and survival,⁷ provoking tumour cells to develop independently from growth and survival signals.⁸

Formation of G4s in gene promoters often results in the suppression of transcription; therefore, ligands that stabilized G4s in those gene promoters should inhibit transcriptional activation. However, considering that G4s are far less abundant than double stranded DNA (dsDNA), selectivity for G4s over dsDNA represents one of the key challenges in the development of new ligands. To be effective, selectivity must be implemented at an early stage in the ligands' design, for example, by incorporating selective G4 binding motifs.⁹

Porphyrins are known to superbly stabilize several types of G4s. Their unique structural features, including the planar macrocycle, endow them with potential to bind to G4 structures by stacking on the G-tetrads.¹⁰ The most explored porphyrin is the tetra cationic TMPyP4 (5,10,15,20-tetrakis(*N*-methyl-4-pyridyl)porphine), which is able to interact with some G4s by strong end-stacking of the charged porphyrin core as well as by weak external loop binding.¹¹ Unfortunately, the therapeutic potential of TMPyP4 is diminished by its low

specificity, as it also binds to single-strands, duplexes and triplexes of DNA.^{12,13} In addition to the mentioned need for selectivity over dsDNA, one of the utmost demands in the field is to achieve also selectivity among different G4 structural subclasses. In this context, it is worth mentioning the pentacationic Mn(III) porphyrin developed by Dixon *et al.* that showed high selectivity for G4s over dsDNA,¹⁴ or anionic porphyrins like NMM which bind preferentially to parallel G4s.¹⁵

During the last 15 years we have developed diaromatic bis-guanidinium-like systems with good affinity towards the DNA minor groove.¹⁶ We have shown that, due to the almost co-planar orientation of the diaromatic core and the binding ability of the guanidinium system, these systems are ideal to fall into a groove.¹⁷ Thus, we envision that the combination of π -stacking classical porphyrins with groove binding motifs such as diaromatic guanidinium systems will yield novel and efficient *dual-mode* binding systems with good affinity and selectivity for G4s. Herein, we present firstly conceptualised diaromatic guanidinium-porphyrin conjugates (Fig. 1) that specifically recognize G4 sequences.

Fig. 1 General structure of the diaromatic guanidinium-porphyrin conjugates prepared in this work. The corresponding Zn(II) porphyrin complexes of **3a** and **3e** were also prepared (**4a** and **4e**, see ESI).

^a School of Chemistry, Trinity Biomedical Science Institute, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland. E-mail: rozasi@tcd.ie.

^b Université de Bordeaux, ARNA Laboratory, F-33000 Bordeaux, France Inserm U1212, CNRS UMR 5320, IECB, F-33600 Pessac, France

Electronic Supplementary Information (ESI) available: Experimental procedures and characterization data for all new compounds. See DOI: 10.1039/x0xx00000x

Scheme 1. Synthesis of diaromatic guanidinium-porphyrin conjugates and Zn(II) complexes.

Results and discussion

Synthesis

Considering this hypothesis, we coupled several diaromatic guanidinium systems (based on compounds previously developed by our group¹⁶) with suitably substituted porphyrins via a *semi-rigid* linker that increase the overall steric crowding, thus precluding interactions with dsDNA and improving G4 selectivity. The present design is distinct from any other reported G4 ligand centred on heteroaromatic cores accompanying positive charges¹⁸ and represents an original approach in the design of G4 ligands.

Our strategy involved reacting bis-Boc-guanidines with aromatic amines to generate amidinouras through the isocyanate mechanism (see Scheme 1).¹⁹ Accordingly, a variety of diverse bis-Boc-protected guanidines (**1a-1f**) were prepared and reacted with 5,10,15,20-tetrakis-(4-aminophenyl)porphyrin (*p*-NH₂-TPP) at high temperature to afford Boc-protected amidinouras (**2a-2f**) in moderate to good yields (47-65%). Subsequent Boc-deprotection of **2a-2f** with a solution of HCl in dry 1,4-dioxane (4M) yielded the hydrochloride salt of the aimed conjugates (**3a-3f**) in good yields (Scheme 1).

Additionally, and in order to explore the differential recognition of G4s by free base porphyrins and their matching metal-derivatives, the synthesis was extended to the corresponding Zn(II) derivatives of some of the conjugates (**4a** and **4e**) using zinc acetate in DCM/MeOH, obtaining the aimed compounds in reasonable yields (Scheme 1).

Biophysical studies

The ligand-induced thermal stabilization of several G4 forming oligonucleotides (Table S1, ESI) was first evaluated with the prepared conjugates **3a-3f**, and the Zn(II) derivatives **4a** and **4e** using UV-melting experiments (Table 1 and Fig. S1, ESI). As mentioned before, G4s can be found at the DNA of human telomers, thus, two human telomeric sequences (24TTA and 22AG) were chosen for these studies. Additionally, since G4s are present in promoter regions of oncogenes and regulatory genes, the sequences for *c-MYC* and *c-KIT*(1 and 2) promoters were also selected for these experiments.

The thermal unfolding of the G4s and dsDNA studied was recorded at 295 and 260 nm, respectively, in the absence and presence of each ligand (at 10 μM ligand in 2:1 ligand:oligo ratio). We observed that with the human telomeric sequence 24TTA, none of the ligands exhibited strong stabilization except for **3e**, which showed a $\Delta T_m = 7.7$ °C.

Table 1 Screening of G4 selectivity of ligands by UV and FRET thermal melting experiments^{a,b}

Ligands	UV ΔT_m (°C)						FRET ΔT_m (°C) ^f						
	24TTA ^c	22AG ^c	c-MYC ^d	c-KIT2 ^d	c-KIT1 ^e	st-DNA	F21T	F21RT	Fkit2T	F21CTAT	FHiv32T	FHiv321T	FdxT
3a	0.0	-3.1	1.5	8.5	2.3	-	5.1	-	-	-	-	-	-
3b	0.0	-2.7	3.3	>20.0	2.2	0.0	3.1	5.2	7.2	3.9	2.7	2.7	0.4
3c	0.9	-4.9	1.3	>20.0	1.0	0.8	2.1	4.4	5.6	2.2	2.3	2.9	0.2
3d	1.2	0.5	0.4	-1.2	0.0	-	1.3	-	-	-	-	-	-
3e	7.7	0.0	0.3	8.4	2.4	-	2.9	-	-	-	-	-	-
3f	0.9	-3.0	2.0	11.7	2.7	0.3	3.1	-	-	-	-	-	-
4a	4.0	-1.4	3.8	13.0	-1.8	-	4.0	5.7	6.1	3.5	2.6	3.9	0.5
4e	2.3	3.6	4.3	4.5	2.4	-	2.6	-	-	-	-	-	-
p-NH₂TPP	1.0	3.1	1.5	0.0	4.5	-1.6	-	-	-	-	-	-	-
TMPyP4	>15	5.3	-7.5	9.0	3.2	5.5	-	-	-	-	-	-	-

^aSee ESI for detailed conditions. ^b ΔT_m represents the difference in melting temperature [$\Delta T_m = T_m(\text{DNA} + \text{ligand}) - T_m(\text{DNA})$]. Ligands were tested at 10 μM and 1 μM for UV-melting and FRET melting assays, respectively. The T_m values of DNAs alone are 24TTA = 51.0 °C, 22AG = 56.4 °C, c-MYC = 76.2 °C, c-KIT2 = 60.5 °C, c-KIT1 = 61.0 °C, st-DNA = 68.5 °C. All experiments were duplicated, and the values reported are average of two measurements. ^c100 mM NaCl buffer. ^d20 mM KCl buffer. ^e100 mM KCl. ^f10 mM KCl and 90 mM LiCl.

Remarkably, most of the compounds studied (5 out of the 8 derivatives analysed) destabilize the telomeric 22AG structure ($\Delta T_m < 0$ °C). Although the mechanism of disruption of 22AG is not clear, destabilization of an RNA G4 in the presence of TMPyP4 has already been reported by Pearson,²⁰ who proposed that this tetra-cationic porphyrin must induce a multimeric G4 instead of a monomeric one.

Next, we screened these ligands against the promoter region of c-MYC and c-KIT(1 and 2) oncogenes using the same conditions of ligand concentration and ligand:oligo ratio (Table 1). To our delight, all compounds demonstrated a good stabilization of the c-KIT2 sequence, except for **3d**. In particular, compounds **3b** and **3c**, carrying an unrestricted diaromatic guanidinium arm, stabilized c-KIT2 to a large extent ($\Delta T_m > 20$ °C), and the metalloporphyrin **4a** showed also a significant effect ($\Delta T_m = 13.0$ °C). Upon switching from a conformationally unrestricted diaromatic core to the restricted fluorene (**3f**), we observed a decline in stabilization ($\Delta T_m = 11.7$ °C), presumably because rotational freedom in the peripheral flexible diaromatic moiety allows optimal interaction with the negative backbone of the loops or with the G4 grooves. In the case of c-MYC and c-KIT1 G4s, most of the ligands exhibited poor stabilising properties ($\Delta T_m < 5$ °C, Table 1), at the concentration used in these experiments.

To assess the potential selectivity for G4s vs. dsDNA, we next examined the binding affinities of these systems to the prototypical salmon testes DNA (st-DNA). None of compounds selected (**3b**, **3c** and **3f**) displayed any binding to st-DNA ($\Delta T_m < 1$ °C, Table 1), highlighting their preference for G4. Further, we screened the control TMPyP4 and the single porphyrin p-NH₂-TPP to test the effect of the linker and diaromatic guanidinium appendage on G4 stabilization and selectivity (Table 1). Indeed, TMPyP4, which was reported to recognize all forms of G4 structures via π - π interactions, albeit with poor selectivity,^{12,13} showed However, p-NH₂-TPP did not display strong binding to any G4 ($\Delta T_m < 5.0$ °C). This result highlights the importance of incorporating a suitable linker to segregate the central aromatic core (stacking element) and the side

diaromatic cationic group (groove binding element) to prevent interactions with dsDNA, thereby enhancing G4 selectivity.

Then, we screened all these conjugates against F21T (human telomeric DNA) using FRET melting assays (using a ligand concentration of 1 μM ; Table 1, Fig. 2) and a sample of them (**3b**, **3c** and **4a**) were also tested against other G4s such as F21RT (human telomeric RNA), Fkit2T (c-KIT2 human oncogene promoter), F21CTAT (mutant human telomeric DNA), FHiv32T (HIV PRO2 sequence), FHiv321T (HIV PRO1 sequence) and FdxT (intramolecular duplex) (Tables 1 and S2, Fig. S2, ESI). Gratifyingly, these ligands exhibited the highest degree of stabilization of Fkit2T and little or no stabilization of the duplex FdxT sequence ($\Delta T_m < 1$ °C) (Fig. 2), confirming the G4 selectivity of these conjugates over dsDNA.

Fig. 2 Results of FRET-melting experiments for the binding of compounds **3b**, **3c** and **4a** to different G4s and a double helix control (FdxT).

To further address this issue, we performed a FRET competition assay in the presence of 3 μM (15 equiv.) and 10 μM (50 equiv.) of dsDNA (Fig. S3, ESI). These assays allow to evaluate the specificity (*S*) of the ligands for G4 over dsDNA (Table 2); a *S* value >0.6 indicates G4 selectivity.²¹ The F21T stabilization was not affected even when 50

equiv. of dsDNA were present confirming the exquisite selectivity for G4 structures over dsDNA.

Table 2 FRET-melting competition assay and specificity (S^a) of ligands **3b**, **3c** and **4a** for *F21T* with and without the presence of ds26 at different concentrations

	ΔT_m			S^a	
	0 μ M ds26	3 μ M ds26	10 μ M ds26	3 μ M	10 μ M
3b	2.9	3.5	4.2	1.2	1.4
3c	2.2	3.5	4.4	1.6	2.0
4a	4.1	2.7	3.9	0.7	0.9

^a S is the ratio of ligand-induced ΔT_m of G4 in the presence and absence of a competitor (ds26).

Hence, UV and FRET thermal melting studies revealed the preferential recognition of G4s over dsDNA models by these ligands and indicated *c-KIT2* as a good target for further analysis of these conjugates binding.

Circular dichroism (CD) spectroscopy can be used to study not only the ligands' binding to G4s, but also the molecular architecture of these oligomers reflecting structural features of a conformation rather than those of the complete sequence. The G4s' characteristic signals in the 210–340 nm wavelength range of the CD spectrum result from the π - π stacking between the G-tetrads and the glycosidic angle torsion (*syn* or *anti*).²² Thus, the CD signature of parallel motifs displays a positive peak near 260 nm and a negative one around 240 nm due to the *syn:syn* or *anti-anti* glycosidic bond conformation, whereas antiparallel motifs display a positive peak near 290 nm and a negative one at 260 nm due to the *syn:anti* or *anti:syn* conformations.²³

Therefore, we investigated the molecular recognition of the best ligands found so far (**3b** and **3c**) to the *c-KIT2* oligo using CD (Fig. 3). The CD signature of *c-KIT2* displayed a positive peak at 264 nm and a negative one at 240 nm, characteristic of a parallel conformation. CD titrations of ligands **3b** (Fig. 3A) and **3c** (Fig. 3B) with *c-KIT2* showed a negative induced CD (iCD) signal in the Soret band of porphyrins (400–450 nm) that enhanced upon increasing the number of equivalents of the ligand from 1 to 5.²⁴

Fig. 3 CD (top) and UV-Vis (bottom) titrations of (A) and (C) **3b**: *c-KIT2* and (B) and (D) **3c**: *c-KIT2*. The oligonucleotide concentration for (A) and (B) is 5 μ M; the ligand concentration for (C) and (D) is 3 μ M.

This, indeed, suggested a change in the chemical environment of the ligands induced by strong interactions of the conjugates with the G-tetrads. Moreover, the CD ellipticity of *c-KIT2* was preserved even after addition of 5 equivs. of **3b**, whereas **3c** diminished the CD ellipticity. This is an interesting aspect that distinguishes our conjugates from the cationic TMPyP4, which is known to disrupt the conformation of human telomeric sequences at high drug:DNA ratio.^{10e}

In order to quantify the G4 binding of conjugates **3b** and **3c**, UV-vis titrations were performed with all the G4s studied and ds12 as a double helix model (Figs. 3 and S5–S16, ESI). The binding constants were calculated by sequential addition of the oligo into the ligand's solution, with the absorption spectra recorded after each addition (Table 3). Conjugates **3b** and **3c** showed strong absorption in the Soret band distinct from the G4 absorption (240–265 nm). Hence, the binding constant of **3b** and **3c** to the oligos were calculated by using the peaks in the Soret region. Both conjugates displayed binding constants to G4s in the 10^5 M⁻¹ range (Table 3) and, low affinity towards dsDNA (ds12 binding in the 10^4 M⁻¹ range, Table 3). The titrations of **3b** (Fig. 3C) and **3c** (Fig. 3D) with *c-KIT2* displayed a 2 nm (from 429 to 431 nm) and 4 nm (from 428 to 432 nm) red shift associated with 5% and 25% hypochromism, respectively. These spectral features indicate end-stacking as a plausible mode of binding of these conjugates to *c-KIT2*. UV-vis titrations performed to **3c** with ds12, displayed a red shift of 2 nm associated with 5% hypochromism. These studies show ligand **3c** to have an almost 32-fold selectivity for *c-KIT2* over ds12.

Table 3 Ligand-G4 binding constants ($K_b \times 10^5$, M⁻¹) obtained by UV-vis titrations^a

	24TTA	22AG	<i>c-MYC</i>	<i>c-KIT2</i>	<i>c-KIT1</i>	ds12
3b	1.15	1.01	1.27	4.34	1.19	0.39
3c	0.18	0.65	0.52	6.09	0.55	0.19

^aBinding constant were measured at $\lambda_{\max}$ between 425–435 nm.

Modelling studies

Finally, for a better understanding of the binding mode of these conjugates to G4s and considering that the 3D structure of *c-KIT2* is known,²⁵ modelling studies were carried out with conjugates **3b** and **3c** and *c-KIT2* as a target. Thus, the structure of the two ligands was optimised at DFT level (B3LYP/6-31G*,²⁶ using SDM-water to model solvation) (Fig. S17, ESI). Also, the monomeric human *c-KIT2* proto-oncogene promoter NMR structure gathered in the PDB (PDB:2KYP) was used as the target where the optimised conjugate structures were docked using Autodock Vina²⁷. The best 'poses' obtained from this docking studies (see the **3c** case in Fig. 4, the **3b** case in Fig. S18 and the scores in Table S3, ESI) show how the porphyrin core of each conjugate rest on top of the upper G-tetrad, while the diaromatic guanidinium *arm* nicely falls within one of the side grooves of this parallel G4, as we had hypothesised.

Fig. 4 Computational model of **3c** (optimised at B3LYP/6-31G* level, SMD-water) docked in the 3D structure of the *c-KIT2* G4 system.

However, it seems that delicate changes within the diaromatic linker (Ph-X-Ph: X= O or S corresponding to **3b** and **3c**, respectively) provokes a big change in the conformation adopted by the conjugates and, therefore, they exhibit different binding mode to the G4 (Figure S19, ESI). It is also important to point out how the *semi-rigid* amidinourea linker bends to facilitate the double interaction with this particular G4.

Structure-function Relationship Discussion

The fact that in CD and UV-vis titration experiments conjugate **3c** seem to form stronger interactions with *c-KIT2* than **3b** contrasts with the results obtained with FRET melting experiments. On the one hand, we could consider that the differences in stabilization observed in the FRET experiments are within the error of the measurements (see Figure S19). On the other hand, similar situations have been previously reported by Waller et al. (J Am Chem Soc. 2009 September 9; 131(35): 12628–12633. doi:10.1021/ja901892u.) ...

Considering all the biophysical experiments performed it should be kept in mind that the concentration of ligand used for each experiment was different (10 μ M for UV thermal melting; 1 μ M for FRET; 1, 2 and 5 equiv. for CD titrations and 3 μ M for UV titrations) and for that reason the magnitude of the responses obtained is not comparable; however, similar trends are observed. Therefore, some general structure-function relationships can be deduced; first, all compounds show a preferential binding to G4s over different models of double stranded DNA (*st-DNA*, *FdxT* or *ds12*). Second, conjugates with diaromatic systems connected by tetrahedral monoatomic linkers such as –O– (**3b**) and –S– (**3c**) provide the best results in affinity towards G4s in general and *c-KIT2* in particular. The conjugate with a rigid diaromatic system such as the fluorene core (**3f**) seems to follow in terms of binding strength, next is that with a tetrahedral –CH₂– (**3a**) followed by **3e** with a more flexible diaromatic system connected by a –CH₂–CH₂– linker.

Some exceptions are observed with different G4s; thus, compound **3e** with a highly flexible linker shows better affinity for *24TTA* than the conformationally restricted **3f** or the tetrahedral linked **3a**, **3b** or **3c**. Additionally, the –CH₂– linked conjugate **3a** shows preferential binding to *F21T*. Finally, the conjugate with a diaromatic system

connected by a pyramidal linker such as –NH– (**3d**) shows a poor binding to all G4s.

Experimental

All commercial chemicals were obtained from either Sigma-Aldrich or Fluka and used without further purification. Deuterated solvents for NMR spectroscopy use were purchased from Apollo. Column chromatography was performed using Sigma-Aldrich silica gel 100–200 mesh. Solvents for synthesis purposes were used at GPR grade. Analytical TLC was performed using either Merck Kieselgel 60 F254 silica gel plates or Polygram Alox N/UV254 aluminium oxide plates. Visualization was by UV light (254 nm).

NMR spectra were recorded on Bruker DPX-400 Avance spectrometers, operating at 400.13 and 600.1 MHz for ¹H NMR; 100.6 and 150.9 MHz for ¹³C NMR. Shifts are referenced to the internal solvent signals.¹ NMR data were processed using Bruker TOPSPIN software. HRMS spectra were measured on a Micromass LCT electrospray TOF instrument with a WATERS 2690 autosampler and methanol/acetonitrile as carrier solvent. Melting points were determined using a Stuart SP10 melting point apparatus and are uncorrected. Infrared spectra were recorded on a PerkinElmer Spectrum One FT-IR spectrometer equipped with a Universal ATR sampling accessory. UV-Vis spectra were recorded on Specord 250 spectrophotometer in the 200–700 nm range at 25 °C with a 1 cm path-length quartz cell.

The HPLC purity analysis of prepared ligands was carried out using Alliance 2695 HPLC system (Waters Corporation, Milford, MA, USA) equipped with PDA detector and automatic injector (5 μ L). The stationary phase consist of Atlantis T3 C18 column (Waters Corporation, Milford, USA, 100 mm $\times$ 2.1 mm; 3 μ m particle size) and elution was carried out using stepwise gradient system, eluting at 0.3 mL per min: acetonitrile (A) and aqueous formate buffer (30 mM, pH 3.0, B) starting with A:B (50:50) to (100: 0) for 30 min. All the compounds analysed were found to be >95% pure.

General procedure for the preparation of Boc-protected diaromatic guanidine-porphyrin conjugates (2a-f)

To a stirred solution of *p*-NH₂-TPP (0.3 mmol) in CHCl₃ (6–7 mL), **1a-f** (1.2 equiv.) was added and the resulting mixture was refluxed for 24 h. Upon completion, the excess solvent was evaporated under vacuum and further purification was performed using column chromatography (eluted in hexane:DCM, 60:40). The intermediate **2d** was not purified owing complex mixture with Boc-protected conjugates, so used as such for Boc-deprotection.

General procedure for preparation of hydrochloride salts (3a-f)

To Boc-protected guanidine precursors (**2a-f**, 1.0 equiv.) were added 4 M HCl/1,4-dioxane (6.0 equiv. per Boc group) and CHCl₃ (2–3 mL). The mixture was stirred at 60 °C until completion, as judged by disappearance of starting material in TLC. Solvent and excess HCl were then removed under vacuum, and the crude salt was washed with minimum amount of water to remove excess HCl and recrystallised from CH₂Cl₂ to furnish pure compounds (**3a-f**).

3a: Yield = 40 mg, 0.04 mmol, 55%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 416 (125,893), 513 (6310), 547 (3981), 592 (1995), 649 nm (2512 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3362, 1719, 1680, 1629, 1482, 1185. ¹H NMR (400 MHz, MeOD): δ. 8.83-8.78 (m, 8H), 8.60 (s, 8H), 8.27 (s, 2H), 8.07 (s, 9H), 7.47 (d, *J* = 7.7 Hz, 2H), 7.42-7.38 (m, 4H), 7.24 (d, *J* = 7.9 Hz, 2H), 4.11 ppm (s, 2H). ¹³C NMR (100 MHz, MeOD): δ. 157.0, 154.6, 151.9, 145.8, 145.7, 142.4, 140.8, 140.5, 140.4, 139.8, 139.78, 139.74, 139.6, 139.5, 138.7, 138.6, 136.2, 133.2, 132.5, 131.4, 130.9, 130.6, 130.5, 128.9, 128.6, 128.2, 126.7, 125.8, 125.7, 122.9, 122.6, 119.7, 40.7 ppm. HRMS (ESI): *m/z* Calcd. for C₆₀H₄₈N₁₁O (M+H)⁺: 938.4043, Found: 938.4041. HPLC: *t*_R 18.75 min., 97.64%.

3b: Yield = 55 mg, 0.056 mmol, 68%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 430 (79,433), 521 (10,000), 558 (6310), 596 (3981), 653 nm (3162 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3340, 1725, 1674, 1629, 1583, 1492. ¹H NMR (400 MHz, MeOD): δ. 8.87 (s, 8H), 8.62 (s, 7H), 8.28 (s, 2H), 8.06 (s, 9H), 7.50 (d, *J* = 8.3 Hz, 2H), 7.34 (d, *J* = 8.6 Hz, 3H), 7.25-7.20 ppm (m, 4H). ¹³C NMR (100 MHz, MeOD): δ. 157.5, 157.0, 156.1, 154.5, 151.5, 145.5, 145.47, 145.44, 139.3, 138.57, 138.55, 138.52, 138.50, 138.4, 130.35, 130.34, 130.29, 128.6, 128.3, 128.1, 127.79, 127.72, 127.68, 122.8, 122.3, 120.3, 120.2, 120.1, 199.9 ppm. HRMS (ESI): *m/z* Calcd. for C₅₉H₄₆N₁₁O₂ (M+H)⁺: 940.3836, Found: 940.3818. HPLC: *t*_R 9.30 min., 98.03%.

3c: Yield = 45 mg, 0.045 mmol, 55%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 429 (100,000), 519 (15,849), 557 (10,000), 591 (7943), 650 nm (6310 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3311, 1731, 1680, 1487. ¹H NMR (400 MHz, MeOD): δ. 8.83 (s, 9H), 8.63 (s, 6H), 8.30 (s, 2H), 8.07 (s, 9H), 7.94-7.87 (m, 1H), 7.54-7.48 (s, 6H), 7.32 ppm (s, 2H). ¹³C NMR (100 MHz, MeOD): δ. 156.6, 154.3, 151.4, 146.4, 146.3, 146.06, 146.04, 145.97, 145.94, 145.90, 139.54, 139.51, 138.4, 138.3, 137.0, 134.6, 133.6, 133.1, 132.8, 132.3, 132.0, 131.7, 130.2, 129.6, 129.5, 129.4, 128.4, 127.2, 126.6, 126.0, 125.9, 125.3, 122.9, 122.5, 119.46, 119.45 ppm. HRMS (ESI): *m/z* Calcd. for C₅₉H₄₆N₁₁OS (M+H)⁺: 956.3608, Found: 956.3597. HPLC: *t*_R 20.29 min., 97.98%.

3d: Yield = 22 mg, 0.022 mmol, 45%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 425 (79,433), 520 (7943), 552 (3981), 591 (3162), 648 nm (2512 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3226, 1612, 1490, 1289. ¹H NMR (400 MHz, MeOD): δ. 8.75 (s, 9H), 8.57 (s, 5H), 8.29 (s, 3H), 8.04 (s, 7H), 7.29-7.14 ppm (m, 11H). ¹³C NMR (100 MHz, MeOD): δ. 157.4, 154.9, 151.9, 146.1, 146.0, 145.9, 145.8, 144.7, 143.6, 143.2, 139.7, 138.69, 138.68, 136.1, 135.94, 135.93, 130.6, 130.0, 129.72, 129.70, 129.59, 129.58, 129.46, 129.43, 128.8, 128.2, 128.1, 127.8, 127.54, 127.51, 127.1, 126.6, 124.6, 123.9, 123.2, 122.85, 122.81, 119.7, 118.8, 118.7, 118.2, 118.1, 117.9 ppm. HRMS (ESI): *m/z* Calcd. for C₅₉H₄₇N₁₂O (M+H)⁺: 939.3996, Found: 939.3988. HPLC: *t*_R 15.14 min., 97.70%.

3e: Yield = 25 mg, 0.025 mmol, 61%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 417 (125,893), 523 (15,849), 557 (10,000), 591 (5012), 648 nm (6310 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3328, 1731, 1674, 1485, 1295. ¹H NMR (400 MHz, MeOD): δ. 8.76 (s, 9H), 8.57 (s, 7H), 8.27 (s, 2H), 8.05 (s, 9H), 7.45-7.36 (m, 6H), 7.19 (d, *J* = 6.7 Hz, 2H), 3.03 ppm (s, 4H). ¹³C NMR (100 MHz, MeOD): δ. 156.7, 154.3, 141.5, 145.9, 145.88, 145.85, 145.83, 145.81, 145.7, 145.65, 145.62, 145.59, 145.56, 142.7, 141.3, 139.4, 138.33, 138.31, 138.29, 138.24, 135.6, 132.5, 130.8, 130.2, 129.8, 129.4, 129.3, 129.26, 129.22, 129.17.

129.14, 128.5, 126.2, 125.3, 122.7, 122.3, 119.4 ppm. HRMS (ESI): *m/z* Calcd. for C₆₁H₄₉N₁₁O (M+H)⁺: 951.4122, Found: 951.4161. HPLC: *t*_R 20.17 min., 100.00%.

3f: Yield = 32 mg, 0.032 mmol, 70%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 417 (79,433), 514 (10,000), 549 (5012), 590 (3981), 625 nm (3162 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3306, 1668, 1583, 1487, 1221. ¹H NMR (400 MHz, MeOD): δ. 8.79 (s, 9H), 8.60 (d, *J* = 4.4 Hz, 8H), 8.28 (d, *J* = 7.5 Hz, 2H), 8.06 (s, 10H), 7.73 (s, 1H), 7.57 (s, 1H), 7.52 (d, *J* = 7.8 Hz, 1H), 7.35 (d, *J* = 7.9 Hz, 1H), 4.11 ppm (s, 2H). ¹³C NMR (100 MHz, MeOD): δ. 157.1, 154.8, 151.9, 146.1, 145.9, 141.7, 140.4, 140.2, 139.8, 139.78, 139.71, 139.6, 138.7, 138.67, 138.64, 138.62, 138.5, 135.9, 134.3, 132.1, 130.7, 130.6, 129.7, 129.5, 129.47, 129.41, 128.8, 125.6, 124.6, 123.5, 123.2, 122.8, 122.5, 121.8, 121.7, 121.4, 119.6, 36.9 ppm. HRMS (ESI): *m/z* Calcd. for C₆₀H₄₆N₁₁O (M+H)⁺: 936.3887, Found: 936.3876. HPLC: *t*_R 12.89 min., 97.59%.

General procedure for porphyrin metallation (4a and 4e)

Metallation of porphyrins was achieved by stirring porphyrins (**3a** and **3e**, 0.05 mmol) and Zn(II)acetate dihydrate (0.4 mmol) in 3:1 dichloromethane/methanol. The solution was stirred overnight at room temperature under nitrogen atmosphere. The solvent was evaporated under vacuum and obtained residues were washed with minimum water (2 × 5 mL) to remove excess of Zn(OAc)₂ followed by recrystallisation from DCM to furnish the pure compounds as acetate salts.

4a: Yield = 17 mg, 0.016 mmol, 50%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 420 (100,000), 555 (7943), 598 nm (2512 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3311, 1731, 1668, 1583, 1532, 1306. ¹H NMR (400 MHz, MeOD): δ. 8.90 (s, 1H), 8.89 (s, 1H), 8.82-8.80 (m, 6H), 8.18 (d, *J* = 6.7 Hz, 7H), 8.07 (d, *J* = 8.0 Hz, 2H), 7.86 (d, *J* = 7.9 Hz, 2H), 7.75-7.73 (m, 10H), 7.36 (d, *J* = 8.0 Hz, 2H), 7.21-7.09 (m, 4H), 3.92 ppm (s, 2H). ¹³C NMR (100 MHz, MeOD): δ. 150.2, 150.0, 143.5, 141.3, 134.5, 134.2, 131.9, 131.0, 130.0, 129.3, 126.9, 126.0, 125.3, 123.1, 120.3, 40.2 ppm. HRMS (ESI): *m/z* Calcd. for C₆₀H₄₅N₁₁OZn (M)⁺: 999.3100, Found: 999.3072. HPLC: *t*_R 21.44 min., 97.22%.

4e: Yield = 22 mg, 0.020 mmol, 45%. m.p. >300 °C. UV-vis (CH₃OH): λ_{max} (ε) 423 (100,000), 558 (6310), 598 nm (1995 L mol⁻¹ cm⁻¹). IR (ATR), ν_{max} (cm⁻¹): 3345, 1735, 1668, 1515, 1323. ¹H NMR (400 MHz, MeOD): δ. 8.90 (s, 1H), 8.89 (s, 1H), 8.82-8.81 (m, 6H), 8.18 (s, 7H), 8.08 (d, *J* = 8.3 Hz, 2H), 7.87 (d, *J* = 4 Hz, 2H), 7.75-7.73 (m, 10H), 7.33 (d, *J* = 7.3 Hz, 2H), 7.20 (d, *J* = 8.4 Hz, 4H), 2.82 ppm (s, 4H). ¹³C NMR (100 MHz, MeOD): δ. 148.7, 148.5, 148.4, 142.0, 139.8, 133.0, 132.6, 129.6, 129.5, 128.3, 128.2, 128.1, 127.5, 125.4, 124.5, 123.7, 123.6, 123.5, 121.9, 118.8, 118.7, 35.2, 35.1 ppm. HRMS (ESI): *m/z* Calcd. for C₆₁H₄₇N₁₁OZn (M)⁺: 1013.3257, Found: 1013.3263. HPLC: *t*_R 13.98 min., 96.26%.

Conclusions

In summary, we have engineered first conceptualised and diverse conjugates by combining peripheral cationic diaromatic guanidinium arms with a central porphyrin unit connected by a semi-rigid linker as probes for G4 recognition over dsDNA.

The binding of these conjugates to various G4 structures was studied by UV and FRET melting experiments, the ligands **3b** and **3c** emerging as the most promising with maximum $\Delta T_m > 20^\circ\text{C}$ (UV-melting; 10 μM) and $\Delta T_m = 7.2$ and 5.6°C (FRET melting; 1 μM), respectively, towards *c-KIT2*. We can conclude that these conjugates show significant selectivity for G4 structures over dsDNA.

Furthermore, CD studies confirmed the binding to *c-KIT2*. UV-vis titrations allowed measuring the corresponding K_b of **3b** and **3c** to different G4s and dsDNA showing not only excellent selectivity for G4s vs. dsDNA, but also that compound **3c** discriminate among numerous G4 structures.

Finally, docking studies of compounds **3b** and **3c** with the 3D structure of *c-KIT2* suggest that the binding mechanism of these conjugates proceeds by stacking of the porphyrin system and fitting of the diaromatic guanidinium arm within the grooves of the G4, as hypothesised.

The success of these unconventional diaromatic guanidinium-porphyrin architectures as selective G4 binders opens new avenues for decoration of existing G4 ligands centred on a porphyrin unit with diaromatic guanidinium systems, thereby leading to improved selectivity. Studies pertaining to gain more insights about the G4 interaction of these ligands as well as their potential application as fluorescent probes for G4s are in progress.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

We acknowledge the Irish Research Council for a postdoctoral fellowship (GOIPD/2016/380, J.G.) and a postgraduate award (GOIPG/2018/2336, N.S.), Science Foundation Ireland (IvP-13/IA/1894) for support, and ICHEC for providing computational facilities. We thank Prof. D. Grayson for generous gift of TPP, P. Chenne for her help preparing starting materials, A. Guédin for her support on FRET studies and Drs. J. O'Brien and G. Hessman for performing NMR and HRMS experiments.

Notes and references

- (a) A.K. Todd, M. Johnston and S. Neidle, *Nucleic Acids Res.*, 2005, **33**, 2901; (b) G. Biffi, D. Tannahill, J. McCafferty and S. Balasubramanian, *Nat. Chem.*, 2013, **5**, 182; (c) V.S. Chambers, G. Marsico, J.M. Boutell, M. Di-Antonio, G.P. Smith and S. Balasubramanian, *Nat. Biotechnol.*, 2015, **33**, 877; (d) A. Bedrat, L. Lacroix and J.L. Mergny, *Nucleic Acids Res.*, 2016, **44**, 1746; (e) R. Hansel-Hertsch, D. Beraldi, S.V. Lensing, G. Marsico, K. Zyner, A. Parry, M. Di-Antonio, J. Pike, H. Kimura, M. Narita, D. Tannahill and S. Balasubramanian, *Nat. Genet.*, 2016, **48**, 1267.
- (a) A. De-Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J. -F. Riou and J. -L. Mergny, *Biochimie*, 2008, **90**, 131; (b) Y. Xu, *Chem. Soc. Rev.*, 2011, **40**, 2719.
- (a) J. X. Dai, T. S. Dexheimer, D. Chen, M. Carver, A. Ambrus, R.A. Jones and D.Z. Yang, *J. Am. Chem. Soc.*, 2006, **128**, 1096; (b) S. Balasubramanian, L. H. Hurley and S. Neidle, *Nat. Rev. Drug Discov.*, 2011, **10**, 261.
- S. Cogoi and L. E. Xodo, *Nucleic Acids Res.*, 2006, **34**, 2536.
- (a) A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. USA*, 2002, **99**, 11593; (b) I. Wierstra and J. Alves, *Adv. Cancer Res.*, 2008, **99**, 113; (c) E. Hatzakis, K. Okamoto and D. Yang, *Biochemistry*, 2010, **49**, 9152.
- (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; (c) A.T. Phan, V. Kuryavyi, S. Burge, S. Neidle and D. J. Patel, *J. Am. Chem. Soc.*, 2007, **129**, 4386; (d) P. S. Shirude, B. Okumus, L. Ying, T. Ha and S. Balasubramanian, *J. Am. Chem. Soc.*, 2007, **129**, 7484; (e) V. Kuryavyi, A. T. Phan and D. J. Patel, *Nucleic Acids Res.*, 2010, **38**, 6757. (f) D. Wei, G. N. Parkinson, A. P. Reszka and S. Neidle, *Nucleic Acids Res.*, 2012, **40**, 4691.
- E. Gregory-Bryson, E. Bartlett, M. Kiupel, S. Hayes and V. Yuzbasiyan-Gurkan, *BMC Cancer*, 2010, **10**, 559.
- (a) P. Pittoni, S. Piconese, C. Tripodo and M. P. Colombo, *Oncogene*, 2011, **30**, 757; (b) J. Lennartsson and L. Rönstrand, *Physiol. Rev.*, 2012, **92**, 1619.
- D. Drygin, A. Siddiqui-Jain, S. O'Brien, M. Schwaeb, A. Lin, J. Bliesath, C.B. Ho, C. Proffitt, K. Trent, J. P. Whitten, J. K. C. Lim, D. Von Hoff, K. Anderes and W.G. Rice, *Cancer Res.*, 2009, **69**, 7653.
- Free-base porphyrins: (a) H. Y. Han, D. R. Langley, A. Rangan and L. H. Hurley, *J. Am. Chem. Soc.*, 2001, **123**, 8902; (b) V. Gabelica, E. S. Baker, M.-P. Teulade-Fichou, E. D. Pauw and M. T. Bowers, *J. Am. Chem. Soc.*, 2007, **129**, 895. Metallo porphyrins: (c) L. R. Keating and V. A. Szalai, *Biochemistry*, 2004, **43**, 15891; (d) S. N. Georgiades, N. H. Abd Karim, K. Suntharalingam and R. Vilar, *Angew. Chem. Int. Ed.*, 2010, **49**, 4020; (e) E. Boschi, S. Davis, S. Taylor, A. Butterworth, L. A. Chirayath, V. Purohit, L. K. Siegel, J. Buenaventura, A. H. Sheriff, R. Jin, R. Sheardy, L. A. Yatsunyk and M. Azam, *J. Phys. Chem. B*, 2016, **120**, 12807.
- (a) A. Arora and S. Maiti, *J. Phys. Chem. B*, 2008, **112**, 8151; (b) S. Manaye, R. Eritja, A. Avino, J. Jaumot and R. Gargallo, *Biochim. Biophys. Acta*, 2012, **1820**, 1987.
- (a) S. Zhang, Y. Wu and W. Zhang, *ChemMedChem*, 2014, **9**, 899; (b) S.A. Ohnmacht and S. Neidle, *Bioorg. Med. Chem. Lett.*, 2014, **24**, 2602; (c) S. Neidle, *Nat. Rev. Chem.*, 2017, **1**, 41.
- J. Ren and J. B. Chaires, *Biochemistry*, 1999, **38**, 16067.
- I.M. Dixon, F. Lopez, A.M. Tejera, J.-P. Estève, M.A. Blasco, G. Pratviel and B. Meunier, *J. Am. Chem. Soc.*, 2007, **129**, 1502.
- J.M. Nicoludis, S.P. Barrett, J.-L. Mergny and L.A. Yatsunyk, *Nucleic Acids Res.*, 2012, **40**, 5432.
- (a) P. S. Nagle, F. Rodriguez, A. Kahvedžić, S. J. Quinn and I. Rozas, *J. Med. Chem.*, 2009, **52**, 7113; (b) P. S. Nagle, F. Rodriguez, B. Nguyen, W. D. Wilson and I. Rozas, *J. Med. Chem.*, 2012, **55**, 4397; (c) C. McKeever, M. Kaiser and I. Rozas, *J. Med. Chem.*, 2013, **56**, 700; (d) P. S. Nagle, C. McKeever, F. Rodriguez, B. Nguyen, W. D. Wilson and I. Rozas, *J. Med. Chem.*, 2014, **57**, 7663; and references therein.
- P. S. Nagle, F. Rodriguez, S. J. Quinn, D. H. O'Donovan, J. M. Kelly, B. Nguyen, W. D. Wilson and I. Rozas, *Org. Biomol. Chem.*, 2010, **8**, 5558-5567.
- (a) S. Neidle and G. Parkinson, *Nat. Rev. Genet.*, 2002, **1**, 383; (b) J.-L. Mergny, J.-F. Riou, P. Mailliet, M.-P. Teulade-Fichou and E. Gilson, *Nucleic Acids Res.*, 2002, **30**, 839; (c) 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; (d) E. S. Baker, J. T. Lee, J. L. Sessler and M. T. Bowers, *J. Am. Chem. Soc.*, 2006, **128**, 2641; and refs. therein.

- 19 (a) H. Miel and S. Rault, *Tetrahedron Lett.*, 1998, **39**, 1565; (b) P. Prabhakaran and G. J. Sanjayan, *Tetrahedron Lett.*, 2007, **48**, 1725.
- 20 B. Zamiri, K. Reddy, R.B. Macgregor Jr. and C.E. Pearson, *J. Biol. Chem.*, 2014, **289**, 4653.
- 21 A. Ou, A. Guedin, B.W. Skelton, S. Amrane, C.W. Evans, M. Norret, K.S. Iyer, J.L. Mergny and N.M. Smith, *Chem. Comm.*, 2018, **54**, 9647.
- 22 (a) N. J. Greenfield, *Nat. Protoc.*, 2007, **1**, 2876–2890; (b) A. I. S. Holm, L. M. Nielsen, S. V. Hoffmann and S. B. Nielsen, *Phys. Chem. Chem. Phys.*, 2010, **12**, 9581.
- 23 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.
- 24 No CD signal was observed for ligands **3b** and **3c** alone (Figure S4, ESI).
- 25 PDB:2KYP (Human c-kit2 promoter), V. Kuryavyi, A.T. Phan and D.J. Patel, *Nucleic Acids Res.*, 2010, **38**, 6757.
- 26 (a) A.D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648; (b) C. Lee, W. Yang and R.G. Parr, *Phys. Rev. B*, 1988, **37**, 785; (c) M.J. Frisch, J. A. Pople and J. S. Binkley, *J. Chem. Phys.*, 1984, **80**, 3265.
- 27 O. Trott and A. J. Olson, *J. Comput. Chem.*, 2010, **31**, 455.