

# Imide Condensation as a Strategy for the Synthesis of Core-Diversified G-Quadruplex Ligands with Anticancer and Antiparasitic Activity\*\*

Steven T. G. Street,<sup>[a, b]</sup> Pablo Peñalver,<sup>[c]</sup> Michael P. O'Hagan,<sup>[a]</sup> Gregory J. Hollingworth,<sup>[d]</sup> Juan C. Morales,<sup>\*[c]</sup> and M. Carmen Galan<sup>\*[a]</sup>

**Abstract:** A facile imide coupling strategy for the one-step preparation of G-quadruplex ligands with varied core chemistries is described. The G-quadruplex stabilization of a library of nine compounds was examined using FRET melting experiments, and CD, UV-Vis, fluorescence and NMR titrations, identifying several compounds that were capable of stabilizing G-quadruplex DNA with interesting selectivity profiles. The best G4 ligand was identified as compound **3**, which was based on a perylene scaffold and exhibited 40-fold selectivity for a telomeric G-quadruplex over duplex DNA. Surprisingly, a tetra-substituted flexible core, compound **11**, also exhibited

selective stabilization of G4 DNA over duplex DNA. The anticancer and antiparasitic activity of the library was also examined, with the lead compound **3** exhibiting nanomolar inhibition of *Trypanosoma brucei* with 78-fold selectivity over MRC5 cells. The cellular localization of this compound was also studied via fluorescence microscopy. We found that uptake was time dependant, with localization outside the nucleus and kinetoplast that could be due to strong fluorescence quenching in the presence of small amounts of DNA.

## Introduction

G-quadruplexes (G4s) are a class of secondary nucleic acid structures formed from guanine-rich sequences. These motifs have garnered significant attention in recent years, as a wide variety of potential applications and functions continue to emerge.<sup>[1]</sup> Since the discovery of the first G4 DNA stabilizing ligand by Neidle and Hurley *et al.* in 1997<sup>[2]</sup> the field of G4 interactive compounds has experienced exponential growth.<sup>[3]</sup>

Advances in our understanding of G4 biology have driven new discoveries of the potential biomedical applications of G4 ligands in the therapeutic targeting of numerous diseases.<sup>[4]</sup> Whilst much attention has focused on the use of this strategy for the treatment of cancer,<sup>[5]</sup> recently their potential to be deployed as antiparasitic agents has come to light.<sup>[6]</sup> The stabilization of key G4-forming sequences in gene promoters is proposed to prevent the binding of transcription factors and lead to downstream gene silencing and toxic effects. Induction of the G4 motif at chromosome telomeres is also demonstrated to induce toxicity via inhibition of telomerase, a known anticancer target.

Whilst numerous small molecules have been reported to interact with G4 DNA for use as therapeutics,<sup>[7]</sup> only a handful have been investigated clinically<sup>[8]</sup> due to challenging syntheses, poor bioavailability, and poor target specificity.<sup>[7c]</sup> Whilst significant progress has been made, it is clear that efficient methods to rapidly generate diverse libraries of potential G4 ligands are needed to provide us with a better understanding of the parameters that govern G4 stabilization, destabilization, and target selectivity, which are key requirements for the development of clinically viable G4 interactive small molecules.

Recent efforts in our groups have been focused on the development of a series of G4 targeting compounds derived from the naphthalene diimide (NDI) scaffold.<sup>[9]</sup> We found that di-substituted diimides, particularly NDI-glycoconjugates, displayed interesting selectivity profiles towards G4 structures with lead compounds **1** and **2** (Figure 1, A) stabilizing particular G4 topologies with good discrimination. Moreover, compound **1** exhibited submicromolar activity against cancerous cell lines. Though NDI cores have been extensively studied as the basis of

[a] Dr. S. T. G. Street, Dr. M. P. O'Hagan, Prof. M. C. Galan  
School of Chemistry, University of Bristol  
Cantocks Close, Bristol, BS8 1TS (UK)  
E-mail: m.c.galan@bristol.ac.uk

[b] Dr. S. T. G. Street  
Department of Chemistry, University of Victoria  
Dr. S. T. G. Street  
Victoria BC V8P 5C2 (Canada)

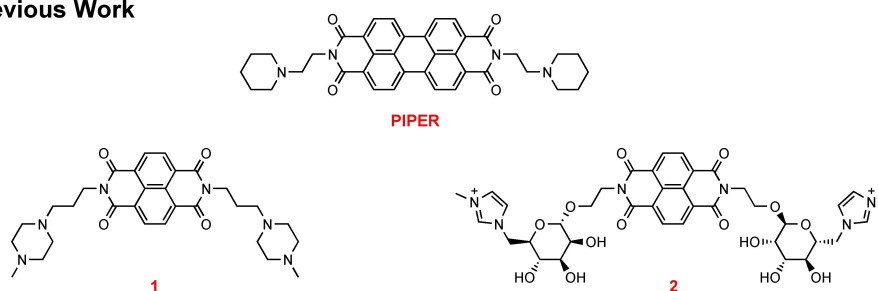
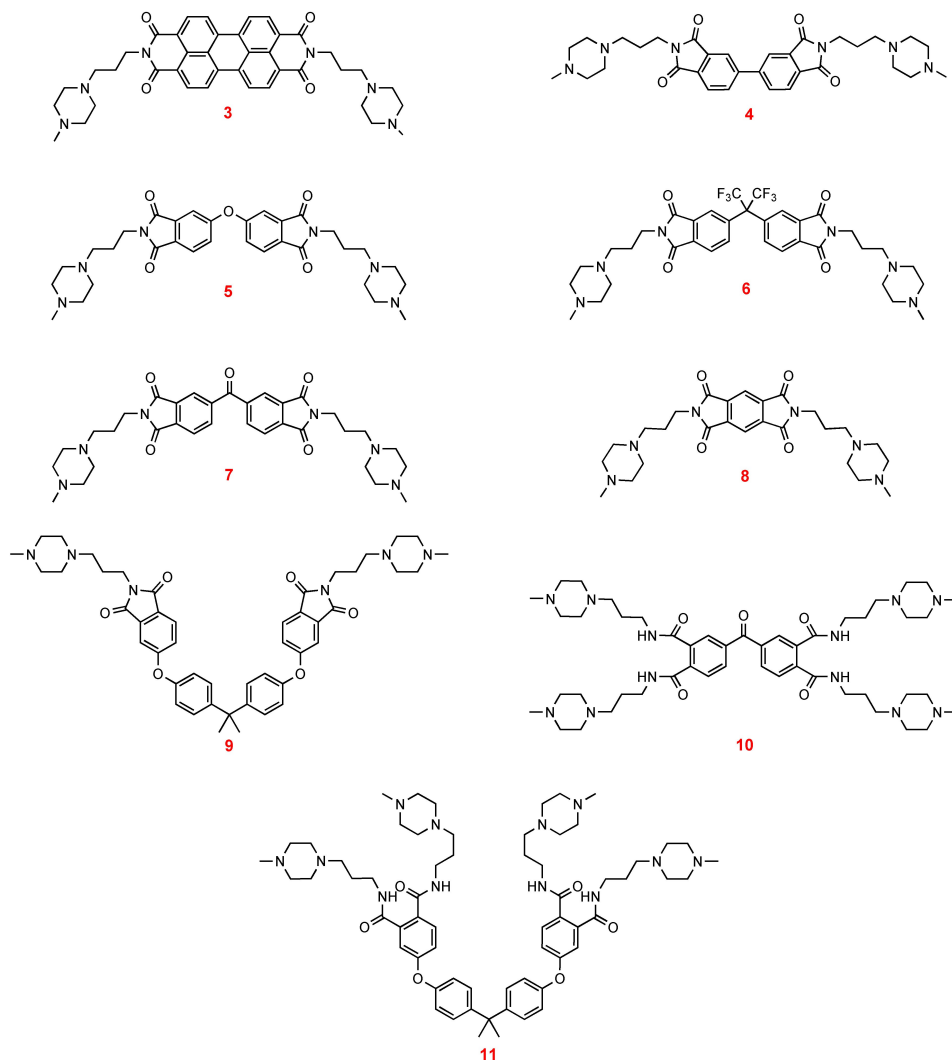
[c] Dr. P. Peñalver, Dr. J. C. Morales  
Instituto de Parasitología y Biomedicina López Neyra  
CSIC, PTS Granada, Avenida del Conocimiento, 17  
18016, Armilla, Granada (Spain)  
E-mail: jcmorales@ipb.csic.es

[d] Dr. G. J. Hollingworth  
Novartis Institutes for Biomedical Research  
Novartis Campus, 4002, Basel (Switzerland)

[\*\*] A previous version of this manuscript has been deposited on a preprint server (<https://doi.org/10.26434/chemrxiv.11955870.v1>).

Supporting information for this article is available on the WWW under <https://doi.org/10.1002/chem.202100040>

© 2021 The Authors. Chemistry - A European Journal published by Wiley-VCH GmbH. This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**A Previous Work****B This Work**

**Figure 1.** Compounds studied in this work, utilizing imide-based coupling of a variety of commercially available building blocks. (A) Lead compounds 1 and 2 from previous work,<sup>[9]</sup> and control G4 ligand PIPER.<sup>[12]</sup> (B) New compounds 3–11 studied in this work, probing the role of core flexibility, aromaticity, and electronics on G-quadruplex binding.

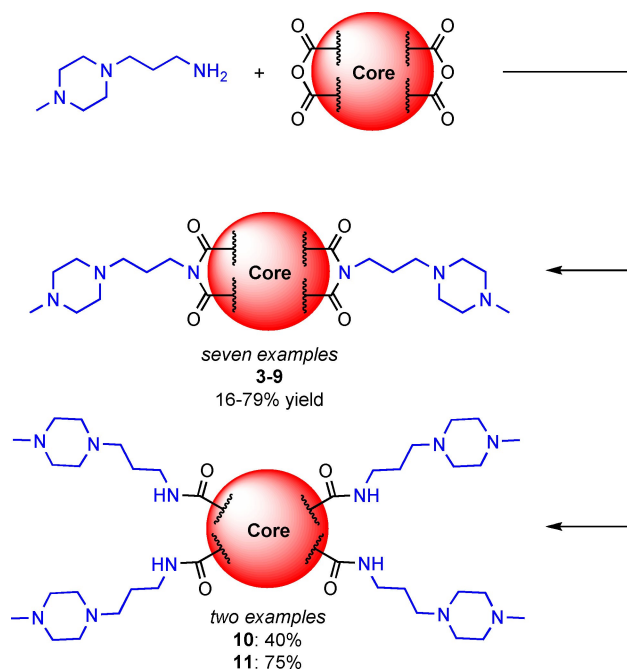
G4-targeting molecules,<sup>[10]</sup> to date there has been no systematic exploration of different diimide scaffolds to probe the role of the ligand core in driving G4 binding, and to identify new core-forming motifs. In this work, we utilize the simplicity of imide coupling chemistry to develop, in one step, a library of functional, amphiphilic small molecules with a variety of differ-

ent cores from commercially available materials. The work herein probes the role of core flexibility, structure, and electronics in facilitating G4 binding, and examines whether ligand G4 stabilisation is correlated with antiparasitic or anticancer activity *in vitro*.

## Results and Discussion

Previous work on G4 ligands<sup>[11]</sup> indicates that 1-(3-amino-propyl)-4-methylpiperazine is an effective side chain that is capable of conferring G4 recognition properties and aqueous solubility to aromatic cores. We thus rationalized that coupling this readily-available motif with a range of commercially available anhydride cores would yield a library of interesting amphiphiles in one step, allowing for fast diversification and generation of novel G4 binding motifs (Scheme 1). As a proof of concept, we started by examining seven different commercially available dianhydrides, which together probe a variety of features important for G4 binding, such as the flexibility of the central scaffold, the role of electronic and steric effects, conformational freedom, and the number and presentation of the aromatic rings. Nine new compounds were produced in total (Figure 1, B). The imide coupling reaction is compatible with a wide range of solvents; good results were obtained in toluene for many of the dianhydrides examined, whilst imidazole was used as the solvent for the synthesis of perylene diimide (PDI) **3**, due to the limited solubility of perylene-3,4,9,10-tetracarboxylic dianhydride. Purification of the resulting compounds by reverse-phase flash chromatography (water + 0.1% trifluoroacetate (TFA)/MeCN, 9:1 to 5:95) was sufficient to protonate the methylpiperazine rings, yielding the final compounds as water soluble TFA salts.

Compounds **3–9** were prepared in yields ranging from 16–79% as di-substituted bisimides, while tetra-substituted compounds **10** and **11** were obtained serendipitously in the presence of an excess of amine in 40 and 75% yield,



**Scheme 1.** The one-step synthesis of the nine compounds studied in this work. Conditions were: toluene, 145 °C for **4–8** and **10–11**; imidazole, 145 °C for **3** and DMSO, NEt<sub>3</sub>, 80 °C for **9**. For experimental details, see Supporting Information.

respectively (Scheme 1). PIPER was also synthesized according to the reported procedure,<sup>[12]</sup> and used as a positive control. Curiously, the dianhydrides used in **10** and **11** were the only cores where tetra-substitution was observed. Further investigation of LC-MS results from the isolated fractions of **11** revealed the presence of tri-substituted species containing an imide and two amides as a minor product (MH<sub>2</sub><sup>2+</sup> = 479). There was no evidence of tri-substituted species containing three amides and one carboxylic acid (expected MH<sub>2</sub><sup>2+</sup> = 488). Analysis of LC-MS data from the isolated fractions of **10** revealed the presence of multiple species with different core substitution patterns, amongst them the tri-substituted species with an imide and two amides (MH<sub>2</sub><sup>2+</sup> = 380) as well as the di-substituted bisimide, and the bis(amide/carboxylic acid). Again, no evidence of the tri-substituted species containing three amides and one carboxylic acid was found (expected MH<sub>2</sub><sup>2+</sup> = 389). The presence of multiple species with a lower degree of substitution is in accordance with the lower yield obtained for compound **10** (40%) compared to **11** (75%) and indicates that this reaction proceeded to a lesser degree of completion. These results suggest that tetra-substitution of **10** and **11** goes through the bisimide intermediate, which is presumably ring-opened by further nucleophilic addition. The variation in isolated yields can be explained by several factors. Comparing the synthesis of PDI **3** (yield 39%) with PIPER (yield 77%), it is clear that water is a better solvent than molten imidazole for the synthesis of amphiphilic PDIs. The low yields for compounds **7** (yield 33%) and **9** (yield 16%) can be explained by competition between the desired di-substituted product and undesired tri- and tetra-substituted side products, which are removed during purification. Also in the case of **9**, it would appear that DMSO/NEt<sub>3</sub> are not favourable reaction conditions for imide condensation. The remaining variation in yield likely results from interplay between the sterics and electronics of the cores, though no obvious trends stand out.

With the library in hand, we began our investigation into G4 stabilization by conducting a Förster resonance energy transfer (FRET) melting assay to assess DNA stabilization using the procedure outlined by De Cian *et al.*<sup>[13]</sup> Briefly, compounds **1** and **3–11** were tested against four model G4-forming oligonucleotides and a hairpin duplex sequence to assess G4: Duplex selectivity as well as G4:G4 selectivity. The sequences examined (at 200 nM) were the human telomeric G4 in potassium buffer (F21T-K<sup>+</sup>, mixed parallel/hybrid G4)<sup>[14]</sup> and sodium buffer (F21T-Na<sup>+</sup>, antiparallel G4),<sup>[15]</sup> the c-Myc promoter G-quadruplex (FmycT-K<sup>+</sup>, parallel G4),<sup>[16]</sup> a polymorphic G4 found in *T. brucei* (FEBr1T-K<sup>+</sup>, a mixed G4 topology),<sup>[6b]</sup> and a hairpin duplex DNA sequence (F10T-K<sup>+</sup>). Together, these sequences represent targets for therapeutic intervention to treat parasitic infections,<sup>[6b]</sup> or cancer.<sup>[1b]</sup> The topology of each sequence was confirmed independently via circular dichroism (CD, Figure S1). Initially, all compounds were tested at 10 μM concentration (Table S1 and Figure S2), with active ligands also being tested at 5 μM (Table 1) and 1 μM (Table S2). Results revealed that only compounds **3**, **9** and **11** were capable of stabilizing G4 structures (Table 1 and Figure S2), although compound **9** only exhibited stabilization at the highest concentration of 10 μM.

**Table 1.** DNA stabilization of PIPER and **1**, **3**, **9** and **11** assessed via FRET Melting Assay at 5  $\mu\text{M}$  ligand and 200 nM DNA concentration. ( $\Delta T_m$  reported to 1 °C using  $\Delta T_{\text{max}}$ , error is reported as  $\sigma$ ). Buffers composition was 10 mM KCl, 90 mM LiCl and 10 mM Li cacodylate for F21T-K<sup>+</sup>, Febr1T-K<sup>+</sup> and F10T-K<sup>+</sup>; 1 mM KCl, 99 mM LiCl and 10 mM Li cacodylate for FmycT-K<sup>+</sup>; and 100 mM NaCl and 10 mM Li cacodylate for F21T-Na<sup>+</sup>. Data for **1** against non-parasitic G4s has been previously reported and is reproduced with permission.<sup>[9]</sup>

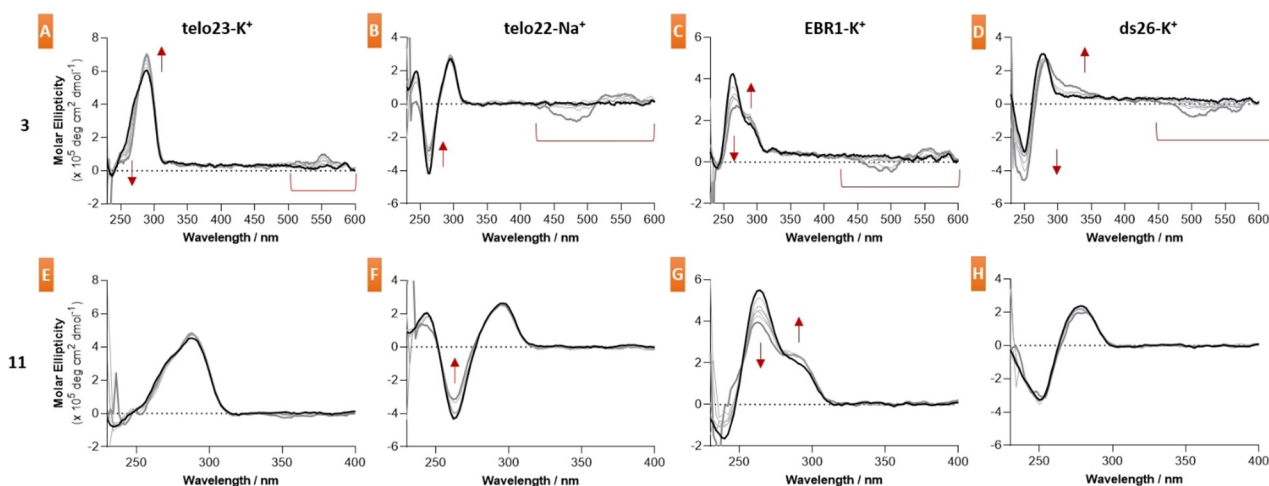
| Compound                | F21T-K <sup>+</sup> | F21T-Na <sup>+</sup> | FmycT-K <sup>+</sup> | Febr1T-K <sup>+</sup> | F10T-K <sup>+</sup> |
|-------------------------|---------------------|----------------------|----------------------|-----------------------|---------------------|
| PIPER <sup>[a]</sup>    | 37 ± 2              | 17 ± 2               | 21 ± 2               | 33 ± 1                | 1 ± 1               |
| <b>1</b>                | 14 ± 2              | -2 ± 1               | 7 ± 2                | 9 ± 1                 | 1 ± 1               |
| <b>3</b> <sup>[a]</sup> | 34 ± 1              | 35 ± 4               | 37 ± 2               | 33 ± 1                | 6 ± 2               |
| <b>9</b>                | -1 ± 1              | -4 ± 1               | -1 ± 2               | -4 ± 1                | 0 ± 1               |
| <b>11</b>               | 7 ± 2               | 1 ± 1                | 5 ± 1                | 6 ± 1                 | 0 ± 1               |

[a] Spectral overlap between the ligand and the FAM/TAMRA FRET pair may interfere with these results. Additionally, the oligonucleotide did not completely unfold under these conditions. These results are therefore not reliable indications of DNA Stabilization.

Whilst PDI based ligands PIPER and **3** exhibited large shifts in  $\Delta T_m$  for all G4 structures and minimal shifts in  $\Delta T_m$  for duplex structures (Figure S3–4), spectral overlap between the PDI core and the FAM/TAMRA FRET pair in the 450–650 nm region prevents definitive conclusions about DNA stabilization from FRET melting experiments. Despite this, a significant difference was observed between the apparent  $\Delta T_m$  for G4 sequences and the duplex sequence. Compounds **9** and **11** also displayed some mild stabilization of G4 DNA, with minimal stabilization of the F10T duplex (Figure S5–6). Tetra-substituted ligand **11** had superior stabilization over di-substituted **9**, in agreement with the number of basic residues being a primary driver for G4 stabilization.<sup>[9,10d,17]</sup> Lead compound **1** also exhibited modest stabilization of the parasitic Febr1T-K<sup>+</sup> G<sub>4</sub> (Figure S7). The lack of G4 stabilization of compounds **4–8** and **10** can likely be attributed to the increased conformational freedom and reduced hydrophobic aromatic surfaces present in many of the

structures. Thus, our data suggests that aromatic  $\pi$ - $\pi$  stacking interactions and structural rigidity of the core motif are key driving forces for G-quadruplex stabilization, and should be considered as important factors alongside the number and location of basic residues when designing optimum G-quadruplex interactive compounds. More rigid cores not only provide hydrophobic and  $\pi$ - $\pi$  stacking interactions with terminal G-tetrads, but also facilitate the correct orientation of side chains into the G4 grooves, leading to a lower entropic penalty for adopting the optimal conformation compared to more flexible cores. Interestingly, our approach also reveals that flexible cores with enough positively charged groups, such as in compound **11**, can show G4 binding ability with a certain selectivity over duplex DNA.

To further probe the G4 and duplex DNA stabilization potential of ligands **1**, **3**, **9**, and **11**, Circular Dichroism (CD) studies were conducted on four different oligonucleotide structures (Figure 2 and S8–9). These were: telo23-K<sup>+</sup> (which forms a hybrid G<sub>4</sub>),<sup>[14]</sup> telo22-Na<sup>+</sup> (which forms an antiparallel G<sub>4</sub>),<sup>[15]</sup> EBR1-K<sup>+</sup> (which forms a mixed G<sub>4</sub>)<sup>[6b]</sup> and ds26-K<sup>+</sup> (which forms a self-complementary duplex). Together, these sequences cover a range of DNA topologies with therapeutic potential. The EBR1-K<sup>+</sup> G<sub>4</sub> is specific to *T. brucei*, whilst the telomeric G<sub>4</sub>s (telo23-K<sup>+</sup> and telo22-Na<sup>+</sup>) are common to both human and *T. brucei* genomes. Oligonucleotide concentrations were 5  $\mu\text{M}$  in all cases, whilst ligand concentration was varied from 1 to 10 equivalents. Results from the titration of **1** with the EBR1-K<sup>+</sup> G<sub>4</sub> and ds26-K<sup>+</sup> duplex revealed perturbation of both structures upon binding (Figure S8). This is similar to the G<sub>4</sub> disruption previously observed for the perturbation of **1** with telo23-K<sup>+</sup> and telo22-Na<sup>+</sup> G<sub>4</sub>'s.<sup>[9]</sup> These results are also in accordance with previous observations of the G<sub>4</sub>:duplex selectivity of **1**, where a significant reduction in G<sub>4</sub> stabilization was observed via FRET melting experiments when competitor duplex DNA was present.<sup>[9]</sup> Titration of **3** with each oligonucleotide revealed

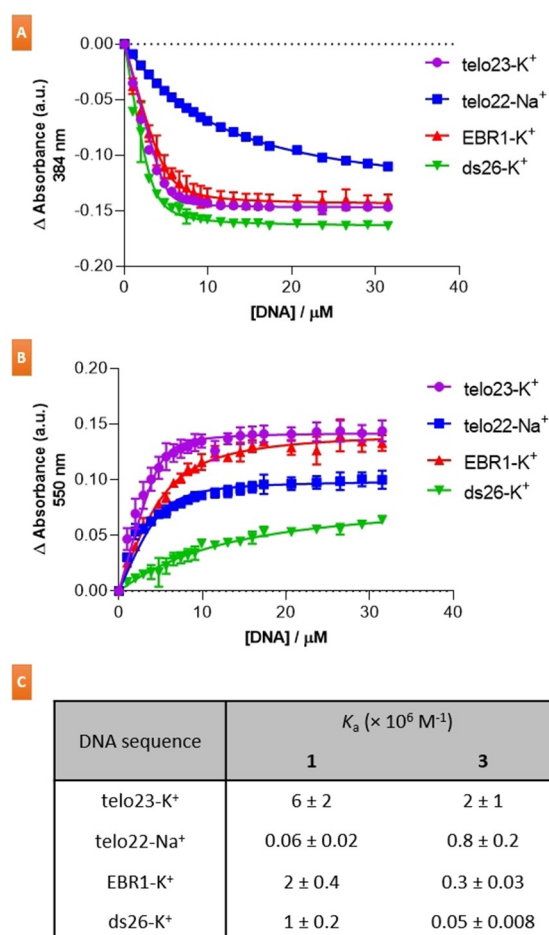


**Figure 2.** CD titrations of compounds **3** and **11** with telo23-K<sup>+</sup>, telo22-Na<sup>+</sup>, EBR1-K<sup>+</sup> (G-quadruplex) and ds26-K<sup>+</sup> (duplex) DNA structures, probing the effects of the compounds on DNA topology. (A–D) Effects of **3** on the (A) telo23-K<sup>+</sup>, (B) telo22-Na<sup>+</sup>, and (C) EBR1-K<sup>+</sup> G-quadruplex structures, and (D) the ds26-K<sup>+</sup> duplex. (E–H) Effects of **11** on the (E) telo23-K<sup>+</sup>, (F) telo22-Na<sup>+</sup>, and (G) EBR1-K<sup>+</sup> G-quadruplex structures, and (H) the ds26-K<sup>+</sup> duplex. Red arrows denote changes in molar ellipticity observed, whilst the red underlined region denotes induced CD from **3**. Oligonucleotide concentration was 5  $\mu\text{M}$ , with ligand concentration varied up to 10 equivalents (50  $\mu\text{M}$ ). For full details, see Supporting Information.



perturbations in the observed CD signal for every structure, and were more significant with the EBR1-K<sup>+</sup> G4 and ds26-K<sup>+</sup> duplex. It is interesting to note the presence of an induced CD band at 400–600 nm that is either due to induced CD from interaction with the oligonucleotide, or from the ligand itself, as PDIs are known to aggregate into helical structures with CD bands in this region.<sup>[18]</sup> However a control titration of the ligand into buffer in the absence of DNA revealed no CD signal (data not shown). The induced CD signal can therefore be attributed to the binding of ligand **3** to the chiral DNA structures. Di-substituted ligand **9** exhibited no effect on any DNA structure examined via CD (Figure S9). In contrast, tetra-substituted ligand **11** modified the CD signal of the telo22-Na<sup>+</sup> G4 and the EBR1-K<sup>+</sup> G4 upon binding, showing no change when binding to the telo23-K<sup>+</sup> G4 or the ds26-K<sup>+</sup> duplex structure.

To obtain quantitative information on the binding of NDI **1** and PDI **3** to the previously assessed G4 and duplex DNA structures, ultraviolet-visible absorbance spectroscopy (UV-Vis) titrations were conducted (Figure 3 and S10). In all cases the ligand concentration was kept constant at 10 μM, whilst oligonucleotide was titrated in up to a final concentration of 30 μM. The observed binding isotherms were fitted to an independent/equivalent sites binding model, and binding constants ( $K_a$ ) were determined. Results from the titration of telo23-K<sup>+</sup>, telo22-Na<sup>+</sup>, EBR1-K<sup>+</sup> and ds26-K<sup>+</sup> oligonucleotides into a solution of NDI **1** (Figure 3 and S10) revealed hypochromicity, and a red-shift in the absorbance of the ligand, which is indicative of end-stacking interactions with the G4 sequences. The titration with telo23-K<sup>+</sup> and telo22-Na<sup>+</sup> yielded a  $K_a$  of  $6 \times 10^6$  and  $6 \times 10^4$ , respectively. The  $K_a$  determined for the binding of **1** to the telo23-K<sup>+</sup> G4 via UV-Vis is in agreement with the  $K_a$  determined previously via isothermal titration calorimetry (ITC).<sup>[9]</sup> The 100-fold selectivity observed for the hybrid form of the telomeric G4 over the anti-parallel form is also in agreement with previous observations from FRET melting assays,<sup>[9]</sup> providing quantitative information on the unique interactions **1** has with different topologies of the telomeric G4. Examples of G4 ligands with this level of discrimination between different G4 topologies are comparatively rare, and further examination of this apparent selectivity might yield useful insights into the future developments of compounds with interesting G4:G4 discrimination. Compound **1** also exhibited significant binding to the EBR1-K<sup>+</sup> G4 and ds26-K<sup>+</sup> duplex ( $K_a = 2 \times 10^6$  and  $1 \times 10^6$ , respectively). These results confirm the significant binding to duplex DNA observed for compound **1**, which is similar to its affinity for the EBR1-K<sup>+</sup> G4, and indicate that it is not as selective for G4 structures over duplex structures as FRET measurements suggest. Attempts to fit the observed binding isotherms to a 1:1 binding model were unsuccessful, resulting in poor fitted curves, whilst a 2:1 model described the data more convincingly. This is in agreement with the 2:1 stoichiometry observed via ITC for the telo23-K<sup>+</sup> G4 observed previously for compound **1**,<sup>[9]</sup> and is also in agreement with potential end-stacking of the ligand on terminal G-tetrads. Results from the titration of the various oligonucleotides into a solution of PDI **3** revealed a similar affinity for the telo23-K<sup>+</sup> G4 ( $K_a = 2 \times 10^6$ ), but a much better selectivity profile, with 40-fold



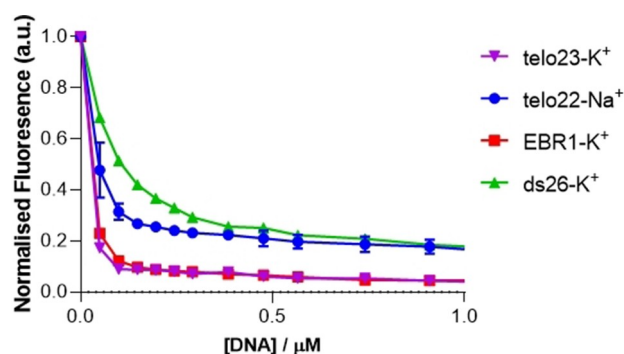
**Figure 3.** UV/Vis binding isotherms and binding constants ( $K_a$ ) for the association between compounds **1** (A) and **3** (B) with telo23-K<sup>+</sup>, telo22-Na<sup>+</sup>, EBR1-K<sup>+</sup>, and ds26-K<sup>+</sup> DNA structures, following the change in ligand absorbance at 384 nm for **1** (A) or 550 nm for **3** (B). (C) Binding constants determined for the association of **1** and **3** with each of the DNA structures examined, fitted using an independent/equivalent sites binding model, with 2:1 ligand:DNA stoichiometry. Ligand concentration was 10 μM, with oligonucleotide concentration varied up to 30 μM. For full details, see Supporting Information.

selectivity over the ds26-K<sup>+</sup> duplex ( $K_a = 5 \times 10^4$ ) and ~7-fold selectivity over the EBR1-K<sup>+</sup> G4 ( $K_a = 3 \times 10^5$ ) versus 6-fold and 3-fold, respectively, for **1**.

Unlike NDI **1**, PDI **3** displayed appreciable binding to the telo22-Na<sup>+</sup> G4 ( $K_a = 8 \times 10^5$ ), indicating that the specific selectivity observed for **1** is not applicable to the larger PDI core of **3** (Figure 3 and S10). Like compound **1**, **3** was also best described by a 2:1 binding model (A) over a 1:1 model. The initial absorbance profile of PDI **3** was consistent with literature reports of aggregated perylene bisimide dyes, and inconsistent with the absorbance spectra of monomeric PDIs, indicating that **3** exists in an aggregated form under these conditions.<sup>[18,19]</sup> Additionally, we observed the presence of a shoulder at 550 nm, which is associated with the presence of face-to-face  $\pi$ - $\pi$  stacks of the PDI core (*H*-aggregates)<sup>[18,19a]</sup> and comparable to the charge-transfer exciton band reported for the perylene analogue perylenetetracarboxylic dianhydride,<sup>[20]</sup> indicating that

the PDI cores are in close proximity to each other. The aggregation of **3** under aqueous conditions was further reinforced by comparing the NMR spectra of **3** in water with that in  $\text{CDCl}_3$  (Figure S11). In  $\text{CDCl}_3$ , the spectrum is well resolved, with observable fine structure. In water, significant line-broadening is observed, with a complete loss in fine structure which is in accordance with aggregated species.<sup>[21]</sup> Upon addition of DNA, a hyperchromic shift in absorbance is observed, with the absorbance profile of **3** exhibiting well-resolved vibronic structure, resembling that of monomeric PDIs.<sup>[18,19,22]</sup> This data demonstrates that PDI **3** undergoes disaggregation in the presence of the DNA sequences examined here, leading to monomeric PDIs that are bound to G4 or duplex DNA. These results imply that the  $K_a$  for the binding of **3** to the G4 and duplex sequences examined here must be higher than the  $K_a$  for the self-association of **3** which leads to aggregation. Thus, the presence of the G4/duplex sequences examined is sufficient to induce the disaggregation of PDI **3**. This observation is in agreement with previous work by Kern et al.<sup>[19a,22]</sup> who showed that the binding of PDIs to G4 and duplex DNA is mediated through ligand disaggregation. Unfortunately, the absorbance of compounds **9** and **11** overlapped with the absorbance from the oligonucleotide ( $\lambda_{\text{abs}} < 300 \text{ nm}$ ), preventing analysis via UV-Vis titration.

To further investigate the effects of DNA binding on PDI **3**, fluorescence titrations were conducted on the previously assessed G4 and duplex DNA structures (telo23- $\text{K}^+$ , telo22- $\text{Na}^+$ , EBR1- $\text{K}^+$  and ds26- $\text{K}^+$ ). The concentration of **3** was kept constant at  $1 \mu\text{M}$ , and the concentration of oligonucleotide was varied up to  $1 \mu\text{M}$ . The results (Figure 4 and S12) reveal that **3** undergoes significant fluorescence quenching in the presence of DNA. Interestingly, a high level of fluorescence quenching was observed at very low levels of DNA, with a ligand:DNA ratio of 5:1 for ds26- $\text{K}^+$  and between 7:1 to 10:1 for G4 DNA. This phenomenon has been previously observed for the binding of PDI ligands to G4 and duplex DNA.<sup>[19a]</sup> The stoichiometry of  $> 5:1$  is in stark contrast to the stoichiometry observed via UV-Vis and implies that phenomena other than specific binding to DNA may be responsible for the observed fluorescence quenching. Considering that PDI **3** has been shown to undergo

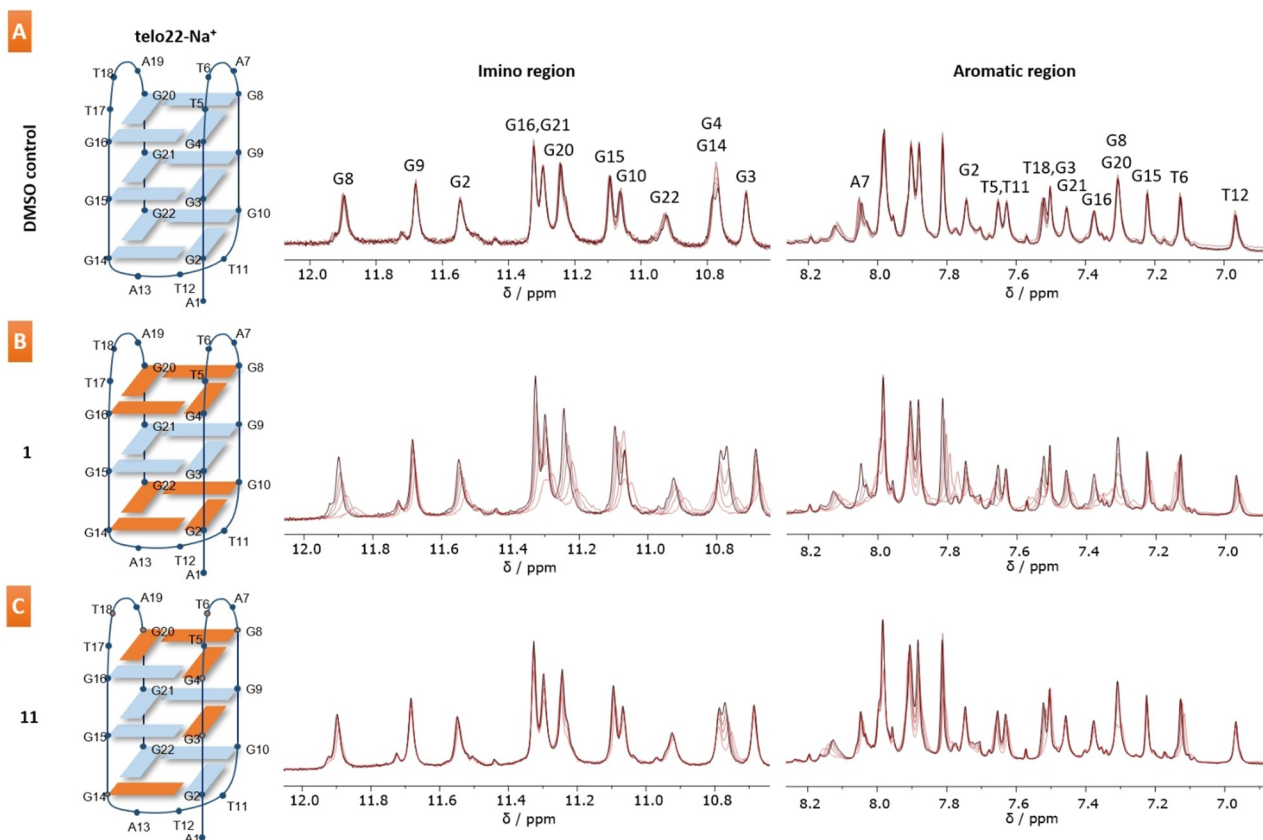


**Figure 4.** Binding isotherms for association of **3** ( $1 \mu\text{M}$ ) with up to  $1 \mu\text{M}$  of (A) telo23- $\text{K}^+$ , (B) telo22- $\text{Na}^+$ , and (C) EBR1- $\text{K}^+$  G-quadruplex and (D) ds26- $\text{K}^+$  DNA, following the change in fluorescence of **3** at  $550 \text{ nm}$ . For full details, see the Supporting Information.

aggregation in aqueous conditions, we hypothesize that the reduction in fluorescence is likely due to the disaggregation phenomena, which may result in increased fluorescence quenching. The sharpest fluorescence quenching was observed for telo23- $\text{K}^+$ , which is in accordance with UV-Vis data that suggested **3** has the strongest interaction with this structure. The lowest fluorescence quenching was observed for ds26- $\text{K}^+$ , which is also in agreement with UV-Vis data that suggested **3** has the weakest interaction with duplex DNA. Whilst the EBR1- $\text{K}^+$  G4 also displayed a similar fluorescence quenching profile to telo23- $\text{K}^+$ , telo22- $\text{Na}^+$  induced a lower level of fluorescence quenching, similar to that of ds26- $\text{K}^+$ , which is in contrast to the binding observed via UV-Vis. Overall, fluorescence titrations revealed that **3** undergoes significant fluorescence quenching in the presence of G4 DNA, and a lower degree of fluorescence quenching in the presence of duplex DNA.

To further assess how compounds **1**, **3**, **9** and **11** interacted with G4 DNA, NMR titrations were conducted with the telo22- $\text{Na}^+$  G4, as it forms a single, well-defined topology – the antiparallel basket (Figure 5 and S13). The oligonucleotide concentration was kept constant at  $185 \mu\text{M}$ , and 0.25, 0.5, 1, or 2 equivalents of ligand were added. The spectrum for the oligonucleotide in aqueous buffer was indicative of an antiparallel basket (Figure 5A) and addition of DMSO as a control ligand barely changed the NMR signals. Results for the titration with **1** revealed significant perturbations of all protons in the imino and aromatic regions and a minor shift for the G9 imino proton (Figure 5B). This could be consistent with **1** interacting with the terminal G-tetrads and a stoichiometry of 2:1, if, at the same time, it is inducing significant topological changes to the quadruplex upon binding that also affects the internal G4 tetrad. In fact, end-stacking would mean significantly affecting the lateral and diagonal loops configuration since they may prevent the NDI core motif from efficiently stacking on the terminal G tetrad. These results are also consistent with previous molecular docking studies of **1**, which also suggested an end-stacking binding mode.<sup>[9]</sup> Furthermore, this scenario may also provide a plausible explanation for the low binding constant observed for **1** against the telo22- $\text{Na}^+$  G4 due to this potential induced-fit binding mechanism.

In contrast, PDI **3** had no discernible change on the chemical shift of the telo22- $\text{Na}^+$  G4 despite the interaction being observable via CD, UV-Vis, and fluorescence (Figure S13B). This effect has been observed for other G4 ligands such as Neidle's diphenyl triazole ligands when binding to telo22- $\text{Na}^+$  or HIVPRO1 G4's and has been attributed to unspecific binding.<sup>[23]</sup> In case of compound **3** which shows concentration-dependant aggregation, the high concentrations used in the NMR titration could alter the equilibrium between the PDI aggregates and binding to G4 DNA, favouring PDI aggregation and a weak association with the telo22- $\text{Na}^+$  G4. Similarly, di-substituted **9** also exhibited minimal effects on the telo22- $\text{Na}^+$  G4 (Figure S13C), which is in accordance with the minimal topological change upon binding revealed by the CD titrations, and also points to potential unspecific binding. In contrast, tetra-substituted **11** exhibited small specific perturbations of imino protons G4 and G14 as well as aromatic protons both in



**Figure 5.** (A) NMR titration of the telo22-Na<sup>+</sup> G-quadruplex with DMSO as control ligand. (B–C) NMR titrations of (B) **1** and (C) **11** with the telo22-Na<sup>+</sup> G-quadruplex. Telo22-Na<sup>+</sup> concentration was 185  $\mu$ M in all cases, with 0 (black trace), 0.25, 0.50, 1, and 2 equivalents of ligand added (red traces). Orange residues in the schematics denote interactions between the ligand and quadruplex. Solutions of DNA were prepared in 25 mM sodium phosphate buffer, pH 7.0 containing 70 mM sodium chloride and 10% D<sub>2</sub>O. For full details, see Supporting Information.

the tetrads (G3, G8 and G20) and in the loops (T5, T6, T11 and T18, Figure 5C). This indicates binding interactions predominantly located at the upper tetrad and on one of the grooves, with a clear different mode to that observed for compounds **1**, **3** and **9**, possibly due to its tetra-substitution with four positive charged groups in an octopus-like arrangement.

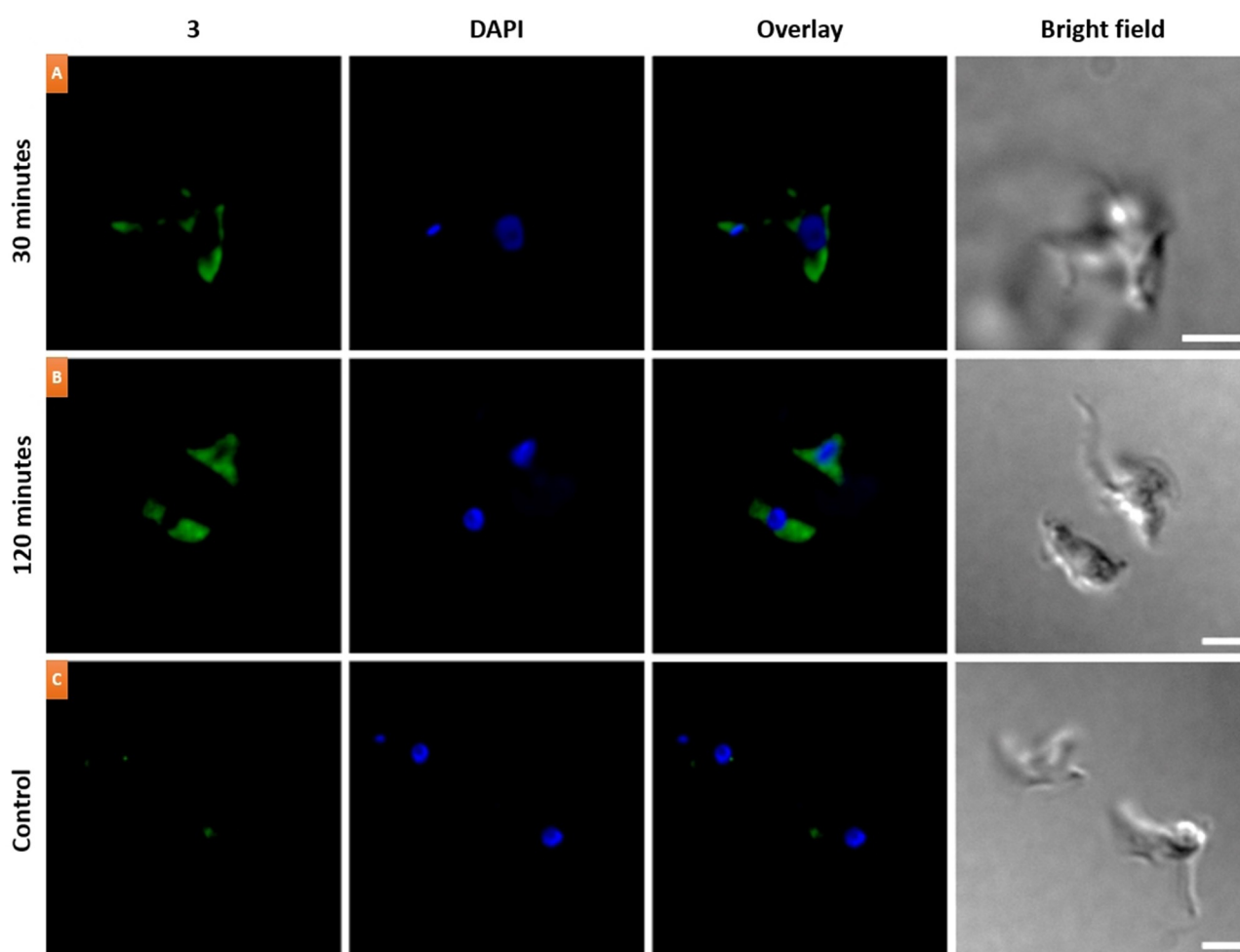
To investigate the potential anticancer and antiparasitic activity of compounds **1** and **3–11**, cytotoxicity assays were conducted on MRC5 non-tumoral cells and on HeLa cervical carcinoma cells, as well as *T. brucei* and *L. major* parasites. Doxorubicin, suramin and miltefosine were used as control compounds.<sup>[6b,9]</sup> After incubation with varied concentrations of the compounds for 72 h, cell viability was measured using either an alamarBlue™ assay (for MRC5, HeLa and *T. brucei*) or MTT assay (for *L. major*). Absolute EC<sub>50</sub> values and selectivity indexes were then calculated from dose-response curves (Table 2 and S14). The results revealed that G4 interactive compounds **1**, **3**, **9**, and **11** all inhibited cell viability, with the highest selectivity index for HeLa cells over MRC5 cells obtained by compound **1** at 10-fold. The most active compounds against parasites were **1**, **3** and **11**, which also exhibited the highest G4 stabilization. All three were particularly effective against *T. brucei*, with **3** exhibiting the best EC<sub>50</sub> value in the nanomolar range. Compounds **3** and **11** showed the highest selectivity for

*T. brucei* over MRC5, with **3** displaying a 78-fold selectivity. This represents a large potential therapeutic window, since considerably lower toxicity is observed towards non-tumoral MRC5 cells (EC<sub>50</sub> = 2  $\mu$ M). The efficacy and selectivity of **1**, **3** and **11** for *T. brucei* over MRC5 was higher than for other typical G4 interactive compounds such as Pyridostatin and BRACO-19 (EC<sub>50</sub> values of 5.5 and 7.8  $\mu$ M, and selectivity index (SI) of 0.7 and 1.5, respectively).<sup>[6b]</sup> In the case of the activity towards *L. major*, only compounds **1** and **11** displayed relevant toxicity (EC<sub>50</sub> values of 0.47 and 18  $\mu$ M, respectively). Whilst no correlation between G4 stabilization and anticancer activity was observed for these compounds, cytotoxicity towards *T. brucei* was correlated with G4 stabilization, with **1**, **3** and **11** all exhibiting antiparasitic activity. Certainly, other mechanisms of action must also be considered and will be studied in the future.

The inherent fluorescence of PDI **3** enabled us to track its localization in both cells and parasites. Preliminary absorbance and fluorescence experiments indicated that **3** exists in *H*-aggregated form in PBS at 25  $\mu$ M (Figure S15). The ability of PDI **3** to undergo internalization into both HeLa (human cervical carcinoma) cells and *T. brucei* parasites was then assessed via fluorescence microscopy (Figure 6 and Figure S16). After incubating HeLa cells with **3** (5  $\mu$ M) for 30 minutes, limited uptake was observed (Figure S16A), however greater uptake was

**Table 2.** Cellular cytotoxicity and antiparasitic activity after 72 h incubation with **1** and **3–11** as well as **Doxorubicin** (MRC5/HeLa positive control), **Suramin** (*T. brucei* positive control) and **Miltefosine** (*L. major* positive control). Reported as absolute EC<sub>50</sub> values measured in  $\mu\text{M}$  using alamarBlue™ fluorescence (for MRC5, HeLa and *T. brucei*) or MTT absorbance (for *L. major*) to assess cellular metabolism, with error represented as  $\sigma$ . SI = selectivity index. The best results are highlighted in bold. Some data for **1** and **Doxorubicin** have been previously reported.<sup>[9]</sup>

| Compound    | MRC5             | HeLa              | <i>T. brucei</i>    | <i>L. major</i>    | SI (MRC5/HeLa) | SI (MRC5/ <i>T. brucei</i> ) | SI (MRC5/ <i>L. major</i> ) |
|-------------|------------------|-------------------|---------------------|--------------------|----------------|------------------------------|-----------------------------|
| Doxorubicin | –                | 0.40 ± 0.05       | –                   | –                  | –              | –                            | –                           |
| Suramin     | –                | –                 | 0.044 ± 0.006       | –                  | –              | –                            | –                           |
| Miltefosine | –                | –                 | –                   | 5.8 ± 0.01         | –              | –                            | –                           |
| <b>1</b>    | <b>3.9 ± 0.5</b> | <b>0.5 ± 0.05</b> | <b>0.61 ± 0.3</b>   | <b>0.47 ± 0.07</b> | <b>9.5</b>     | <b>6.4</b>                   | <b>8.3</b>                  |
| <b>3</b>    | <b>2.1 ± 0.6</b> | <b>53 ± 6.2</b>   | <b>0.027 ± 0.01</b> | > 100              | <b>0.04</b>    | <b>78</b>                    | –                           |
| <b>4</b>    | > 100            | > 100             | 57 ± 10             | > 100              | –              | –                            | –                           |
| <b>5</b>    | > 100            | > 100             | 67 ± 18             | > 100              | –              | –                            | –                           |
| <b>6</b>    | > 100            | > 100             | 36 ± 2              | > 100              | –              | –                            | –                           |
| <b>7</b>    | > 100            | > 100             | 70 ± 13             | > 100              | –              | –                            | –                           |
| <b>8</b>    | > 100            | > 100             | 52 ± 7              | > 100              | –              | –                            | –                           |
| <b>9</b>    | 88 ± 10          | 29 ± 3.5          | 33 ± 24             | > 100              | 3.0            | 2.7                          | –                           |
| <b>10</b>   | > 100            | > 100             | 55 ± 3              | > 100              | –              | –                            | –                           |
| <b>11</b>   | <b>59 ± 37</b>   | <b>14 ± 1</b>     | <b>2.8 ± 2</b>      | <b>18 ± 1</b>      | <b>4.2</b>     | <b>21</b>                    | <b>3.2</b>                  |



**Figure 6.** Fluorescence microscopy images of *T. brucei* parasites after incubation with **3** (5  $\mu\text{M}$ ) for (A) 30 min or (B) 120 min respectively. Control parasites are displayed in (C). Nuclear staining was performed using DAPI (blue fluorescence). Excitation was performed with the 350–450 and 492–518 filters for DAPI and **3**, respectively. A triple filter (437–474, 508–550 and 595–670 nm) was used to detect the fluorescence emission of both DAPI and **3**. Scale bar = 5  $\mu\text{m}$ .

detected after 120 minutes (Figure S16B). The lack of colocalized fluorescence with DAPI staining suggested limited presence of compound **3** in the nucleus, although it can be

observed in the rest of the cell. A similar scenario occurred for **3** when assayed in *T. brucei*. Uptake of **3** was observed after just 30 minutes incubation (Figure 6A), and remained after 120



minutes (Figure 6B). DAPI staining of the nucleus and kinetoplast also revealed no overlap with the fluorescence emission of **3**. In both cases, the lack of fluorescence emission from **3** in the nucleus or kinetoplast can be expected given that **3** exhibits significant fluorescence quenching in the presence of small quantities of DNA. In any case, further experiments would be needed to rule out the possibility that **3** is unable to localize into the nucleus and the kinetoplast.

## Conclusion

In conclusion, we have developed a practical strategy for the preparation of a diverse array of amphiphilic compounds using simple yet versatile imide coupling chemistry, examining their potential use as G4 ligands and as therapeutic agents. A library of nine compounds with different core motifs were produced in one step, from cheap, commercially available starting materials. We have examined the role of the core motif in facilitating G4 binding, highlighting the importance of core rigidity such as in the case of compounds **1** and **3**. Most compounds with a flexible or non-fused core did not induce G4 stabilization. However, it is interesting to note that compound **11**, containing a flexible core and four positively charged groups arranged in an octopus-like pose, is capable of G4 binding thus opening a potential alternative strategy for developing G4 binders. In total, four compounds showed binding to G4 DNA using FRET and their interactions with a variety of G4 and duplex DNA structures was assessed via CD, UV-Vis, Fluorescence, and NMR titrations. The results revealed that previous NDI **1** exhibits appreciable binding to duplex DNA, though **1** also displayed 100-fold selectivity for the mixed parallel/hybrid telo23-K<sup>+</sup> G4 over the antiparallel telo22-Na<sup>+</sup> G4. Compound **1** also seems to interact with the telo22-Na<sup>+</sup> G4 through end-stacking, which corroborates the 2:1 binding stoichiometry previously determined. The analogous PDI, **3**, exhibited significant binding to G4 DNA, and superior selectivity for G4 DNA over duplex DNA, with a 40-fold selectivity for the telo23-K<sup>+</sup> G4 over the ds26-K<sup>+</sup> duplex. The core-diversification yielded a new linear, flexible core with four non-fused aromatic rings and a much greater degree of conformational freedom, that marks a departure from the fused polycyclic aromatic NDI core on which it was based. Di-substitution of this core (compound **9**) was insufficient to induce DNA stabilization, however the tetra-substituted analogue **11** exhibited selective stabilization of G4 DNA over duplex DNA, and interacted with specific residues at the top tetrad and on one of the grooves of the telo22-Na<sup>+</sup> G4. The anticancer and antiparasitic activity of the library was also examined, with PDI **3** displaying potent toxicity towards *T. brucei*, and 78-fold selectivity over control MRC5 cells. The potent antiparasitic activity of **3** correlated with strong stabilization of G4 DNA, although other modes of action could be operating. Compound **3** was internalized into both HeLa cells and *T. brucei*, and can be imaged at therapeutically relevant concentrations (5 μM) with uptake that is time dependent. No evidence for nuclear or kinetoplast uptake of **3** was found, which could be due to the fluorescence quenching of **3** in the

presence of DNA. We hope this approach to diversify the core of future G4 ligands may inspire other library-based convenient approaches to G4 interactive compounds that could help identify potent and selective G4 ligands with promising therapeutic activity.

## Supporting Information

Experimental Details, including further FRET melting experiments, CD experiments, UV-Vis experiments, fluorescence experiments, NMR experiments, cytotoxicity studies and fluorescence microscopy is available in the supporting information.

## Acknowledgements

For funding, STGS thanks the EPSRC (EP/G036764/1 and EP/N509619/1) and NSERC, MPO thanks the EPSRC (EP/L015366/1), JCM/PP thank Spanish Ministerio de Ciencia Innovación y Universidades (Grants CTQ2015-64275-P and RTI2018-099036-B-I00') and MCG thanks the European Research Council (ERC-COG: 648239).

## Conflict of Interest

The authors declare no conflict of interest.

**Keywords:** Aggregation · Amphiphiles · Antiprotozoal agents · Drug design · G-quadruplexes

- [1] a) M. P. O'Hagan, J. C. Morales, M. C. Galan, *Eur. J. Org. Chem.* **2019**, 4995–5017; b) S. Balasubramanian, L. H. Hurley, S. Neidle, *Nat. Rev. Drug Discovery* **2011**, *10*, 261–275; c) J. Ren, T. Wang, E. Wang, J. Wang, *Analyst* **2015**, *140*, 2556–2572; d) J. L. Huppert, *Chem. Soc. Rev.* **2008**, *37*, 1375–1384.
- [2] D. Sun, B. Thompson, B. E. Cathers, M. Salazar, S. M. Kerwin, J. O. Trent, T. C. Jenkins, S. Neidle, L. H. Hurley, *J. Med. Chem.* **1997**, *40*, 2113–2116.
- [3] M. P. O'Hagan, J. L. Mergny, Z. A. E. Waller, *Biochimie* **2018**, *147*, 170–180.
- [4] a) N. Maizels, *EMBO Rep.* **2015**, *16*, 910–922; b) E. Ruggiero, S. N. Richter, *Nucleic Acids Res.* **2018**, *46*, 3270–3283; c) J. Spiegel, S. Adhikari, S. Balasubramanian, *Trends Chem.* **2020**, *2*, 123–136.
- [5] a) C. Nakanishi, H. Seimiya, *Biochem. Biophys. Res. Commun.* **2020**, *531*, 45–50; b) J. Zell, F. Rota Sperti, S. Britton, D. Monchaud *RSC Chem. Biol.* **2021**, *2*, 47–76.
- [6] a) D. Bhartiya, V. Chawla, S. Ghosh, R. Shankar, N. Kumar, *Genomics* **2016**, *108*, 224–231; b) E. Belmonte-Reche, M. Martínez-García, A. Guédin, M. Zuffo, M. Arévalo-Ruiz, F. Doria, J. Campos-Salinas, M. Maynadier, J. J. López-Rubio, M. Freccero, J.-L. Mergny, J. M. Pérez-Victoria, J. C. Morales, *J. Med. Chem.* **2018**, *61*, 1231–1240; c) M. P. O'Hagan, P. Peñalver, R. Fisher, J. C. Morales, M. C. Galan, *Chem. Eur. J.* **2020**, *26*, 6224–6233.
- [7] a) A. R. Duarte, E. Cadoni, A. S. Ressurreição, R. Moreira, A. Paulo, *ChemMedChem.* **2018**, *13*, 869–893; b) E. Ruggiero, S. N. Richter, *Nucleic Acids Res.* **2018**, *46*, 3270–3283; c) S. Neidle, *J. Med. Chem.* **2016**, *59*, 5987–6011.
- [8] S. Asamitsu, S. Obata, Z. Yu, T. Bando, H. Sugiyama, *Molecules* **2019**, *24*, 429.
- [9] S. T. G. Street, D. N. Chin, G. J. Hollingworth, M. Berry, J. C. Morales, M. C. Galan, *Chem. Eur. J.* **2017**, *23*, 6953–6958.

- [10] a) M. Arévalo-Ruiz, F. Doria, E. Belmonte-Reche, A. De Rache, J. Campos-Salinas, R. Lucas, E. Falomir, M. Carda, J. M. Pérez-Victoria, J.-L. Mergny, M. Freccero, J. C. Morales, *Chem. Eur. J.* **2017**, *23*, 2157–2164; b) M. Zuffo, A. Stucchi, J. Campos-Salinas, M. Cabello-Donayre, M. Martínez-García, E. Belmonte-Reche, J. M. Pérez-Victoria, J. L. Mergny, M. Freccero, J. C. Morales, F. Doria, *Eur. J. Med. Chem.* **2019**, *163*, 54–66; c) V. Pirota, M. Nadai, F. Doria, S. N. Richter, *Molecules* **2019**, *24*, 426; d) S. Neidle, *Therapeutic Applications of Quadruplex Nucleic Acids*, Academic Press, San Diego, **2012**.
- [11] S. M. Hampel, A. Sidibe, M. Gunaratnam, J.-F. Riou, S. Neidle, *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6459–6463.
- [12] C. Sissi, L. Lucatello, A. P. Krapcho, D. J. Maloney, M. B. Boxer, M. V. Camarasa, G. Pezzoni, E. Menta, M. Palumbo, *Bioorg. Med. Chem.* **2007**, *15*, 555–562.
- [13] A. De Cian, L. Guittat, M. Kaiser, B. Saccà, S. Amrane, A. Bourdoncle, P. Alberti, M.-P. Teulade-Fichou, L. Lacroix, J.-L. Mergny, *Methods* **2007**, *42*, 183–195.
- [14] K. N. Luu, A. T. Phan, V. Kuryavyi, L. Lacroix, D. J. Patel, *J. Am. Chem. Soc.* **2006**, *128*, 9963–9970.
- [15] Y. Wang, D. J. Patel, *Structure* **1993**, *1*, 263–282.
- [16] A. Ambrus, D. Chen, J. Dai, R. A. Jones, D. Yang, *Biochemistry* **2005**, *44*, 6, 2048–2058.
- [17] F. Cuenca, O. Greciano, M. Gunaratnam, S. Haider, D. Munnur, R. Nanjunda, W. D. Wilson, S. Neidle, *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1668–1673.
- [18] S. Ghosh, X. Q. Li, V. Stepanenko, F. Würthner, *Chem. Eur. J.* **2008**, *14*, 11343–11357.
- [19] a) J. T. Kern, S. M. Kerwin, *Bioorg. Med. Chem. Lett.* **2002**, *12*, 3395–3398; b) F. Würthner, Z. Chen, V. Dehm, V. Stepanenko, *Chem. Commun.* **2006**, *4*, 1188–1190.
- [20] a) Z. Shen, P. E. Burrows, S. R. Forrest, M. Ziari, W. H. Steier, *Chem. Phys. Lett.* **1995**, *236*, 129–134; b) V. Bulović, P. E. Burrows, S. R. Forrest, J. A. Cronin, M. E. Thompson, *Chem. Phys.* **1996**, *210*, 1–12.
- [21] S. R. Laplante, R. Carson, J. Gillard, N. Aubry, R. Coulombe, S. Bordeleau, P. Bonneau, M. Little, J. O'Meara, P. L. Beaulieu, *J. Med. Chem.* **2013**, *56*, 5142–5150.
- [22] a) J. T. Kern, P. W. Thomas, S. M. Kerwin, *Biochemistry* **2002**, *41*, 11379–11389.
- [23] a) A. D. Moorhouse, A. M. Santos, M. Gunaratnam, M. Moore, S. Neidle, J. E. Moses, *J. Am. Chem. Soc.* **2006**, *128*, 15972–15973; b) M. Arévalo-Ruiz, S. Amrane, F. Rosu, E. Belmonte-Reche, P. Peñalver, J. L. Mergny, J. C. Morales, *Bioorg. Chem.* **2020**, *99*, 103786.

---

Manuscript received: February 5, 2021

Accepted manuscript online: March 29, 2021

Version of record online: May 2, 2021