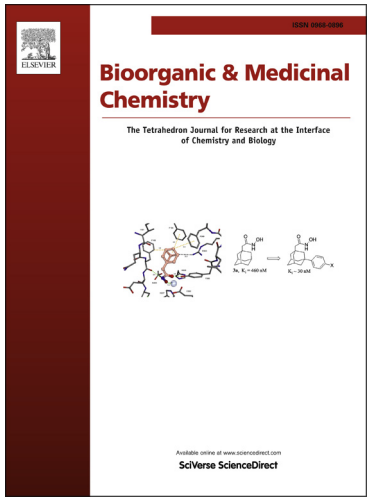


The influence of positional isomerism on G-quadruplex binding and anti-proliferative activity of tetra-substituted naphthalene diimide compounds

Sheila Mpima, Stephan A. Ohnmacht, Maria Barletta, Jarmila Husby, Luke C. Pett, Mekala Gunaratnam, Stephen T. Hilton, Stephen Neidle

PII: S0968-0896(13)00463-X
DOI: http://dx.doi.org/10.1016/j.bmc.2013.05.027
Reference: BMC 10853

To appear in: *Bioorganic & Medicinal Chemistry*

Please cite this article as: Mpima, S., Ohnmacht, S.A., Barletta, M., Husby, J., Pett, L.C., Gunaratnam, M., Hilton, S.T., Neidle, S., The influence of positional isomerism on G-quadruplex binding and anti-proliferative activity of tetra-substituted naphthalene diimide compounds, *Bioorganic & Medicinal Chemistry* (2013), doi: http://dx.doi.org/10.1016/j.bmc.2013.05.027

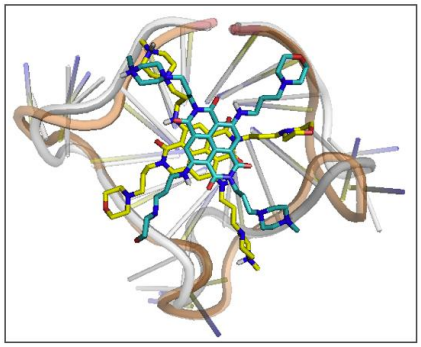
Graphical Abstract

The influence of positional isomerism on G-quadruplex binding and anti-proliferative activity in tetra-substituted naphthalene diimides

Leave this area blank for abstract info.

Sheila Mpima, Stephan A. Ohnmacht, Maria Barletta, Jarmila Husby, Luke C. Pett, Mekala Gunaratnam, Stephen T. Hilton and Stephen Neidle

The School of Pharmacy, University College London, London WC1N 1AX, UK



ELSEVIER

Bioorganic & Medicinal Chemistry
journal homepage: www.elsevier.com

The influence of positional isomerism on G-quadruplex binding and anti-proliferative activity of tetra-substituted naphthalene diimide compounds

Sheila Mpima, Stephan A. Ohnmacht, Maria Barletta, Jarmila Husby, Luke C. Pett, Mekala Gunaratnam, Stephen T. Hilton and Stephen Neidle*

The School of Pharmacy, University College London, 29-39 Brunswick Square, London WC1N 1AX, UK

ARTICLE INFO

Article history:
Received
Received in revised form
Accepted
Available online

Keywords:
G-quadruplex
naphthalenediimide
pancreatic cancer
molecular dynamics

ABSTRACT

The synthesis together with biophysical and biological evaluation of a series of tetra-substituted naphthalenediimide (ND) compounds, are presented. These compounds are positional isomers of a recently-described series of quadruplex-binding ND derivatives, in which the two N-methyl-piperidine-alkyl side-chains have now been interchanged with the positions of side-chains bearing a range of end-groups. Molecular dynamics simulations of a pair of positional isomers are in accord with the quadruplex stabilization and biological data for these compounds. Analysis of structure-activity data indicates that for compounds where the side-chains are not of equivalent length then the positional isomers described here tend to have improved cell proliferation potency and in some instances, superior quadruplex stabilization ability.

1. Introduction

Quadruplex DNAs and RNAs are four-stranded structures formed by the stabilization of repetitive G (guanine) tracts in duplex or single-stranded nucleic acids by means of Hoogsteen hydrogen bonds to form stable G-quartets.[1-3] The targeting of quadruplex DNA elements in telomeres and in oncogene promoter regions[4,5] is emerging as a useful approach to anticancer therapy, especially now that the existence of quadruplex DNAs as structural nucleic acid motifs distinct from double-stranded DNA has been validated at the genomic level.[6,7] The concept has several distinct consequences, in particular.

- Telomeric DNA length is maintained in the majority of human cancers by means of the reverse transcriptase activity of the telomerase enzyme complex, whose expression is up-regulated, contrasting with the lack of telomerase expression in normal somatic cells. Single-stranded telomeric DNA is the substrate for the enzyme and quadruplex formation or ligand-induced stabilization results in telomerase inhibition, with consequential telomere shortening, cellular senescence and selective cell death as a consequence of DNA damage induction.[4]

- The existence of quadruplex-forming sequences within promoter regions, and close to transcription start sites, that can be stabilized by appropriate small molecules, has been demonstrated for a number of oncogenes and other genes involved in cellular growth proliferation, such as *c-myc*, *c-kit*, *HIF*, *VEGF* and *HSP90*. Quadruplex stabilization leads to inhibition of transcription for that particular gene, which may have significant therapeutic consequences.[8] In addition, there is evidence that this can be accompanied by a DNA damage response, augmenting the consequences for affected cells.[9]

The task of designing small molecules that have selectivity for quadruplexes in general over duplex DNA affinity, has been effectively explored,[10,11a] with a large number of small molecules binding G-quadruplexes now described in the literature. The majority are synthetic compounds sharing the common features of a planar heteroaromatic core together with one or more cationic side-arms. Some have metallo- containing groups.[11b] However, the goal of designing ligands to target a single quadruplex is altogether more demanding, not least because only a small fraction of the potential quadruplex landscape has been examined. In reality this may not be a therapeutic disadvantage as a quadruplex-focused poly-targeting approach may be especially useful for those genetically highly aberrant solid cancers where multiple oncogene pathways may be dis-regulated.

A number of naphthalene diimide compounds have been described, some with high quadruplex affinity and cellular potency.[12-17] In general these have straight forward chemical accessibility and potential for ready analogue development. We have previously disclosed several series of tetra-substituted naphthalene diimide compounds,[18-25] a number of which have exceptional telomeric quadruplex-binding ability and associated potent cellular proliferation properties in a panel of cancer cell lines. Most recently we have used structure-based design methods to optimize the pharmacological properties of several analogues with four side-chains each terminating in a N-methyl-piperazine group.[25] Those with di-N-methyl-piperazine, di-

morpholine end-groups have superior (by up to 10-fold) cellular potency, notably against a panel of pancreatic and non-small-cell lung cancer cell lines.

This paper reports on a study to examine the effects on quadruplex binding and cell growth inhibition, of interchanging the positions of the extended side-chain substituents around the naphthalene diimide core, for these most recent active compounds,[25] in order to ensure that these features have been optimized. Thus, earlier compounds **1a**-**5a** are compared with their isomers **1**-**5**, whose synthesis and evaluation is reported here. Three quadruplex sequences have been used as examples of diverse quadruplex types and because of their potential involvement in the mode of action of tetra-substituted naphthalene diimides;[18-25] a well-studied human telomeric sequence and two promoter quadruplexes from the HSP90 gene.[22,26] The senescent cellular responses to the lead compound identified in a recent study (**2a**) suggests that it targets telomeres and elicits and DNA damage/stress response following telomeric quadruplex interaction.[25] Computational molecular modeling has been used to rationalize the effects of positional isomerism, and is based on the recent co-crystal structures of several naphthalene diimide analogues with human telomeric quadruplexes.[23, 25]

2. Results and Discussion

2.1 Chemistry

Synthesis of compounds **1**-**5** followed our recently-developed general procedure.[25] Dibromoisocyanuric acid was reacted with anhydride (I) in sulfuric acid to give the bis-brominated intermediate (II), followed by two sequential microwave reactions with selected primary amines, which afforded the crude tetra-substituted naphthalene diimides. Purification on a C18 reversed phase preparative HPLC system produced the final compounds.

2.2 Fluorescence Resonance Energy Transfer (FRET) studies

The extent of G-quadruplex stabilization by compounds **1**-**5** and **1a**-**5a**, as measured with a high-throughput FRET (Fluorescence Resonance Energy Transfer) technique,[27] has been used to compare these naphthalene diimide derivatives with their positional isomers (Table 1). In general differences in melting temperatures (ΔT_m values) between individual isomers are greatest when the respective side chains are of different linker length. Both compounds in the pair **2** and **2a** have almost identical melting temperatures to each other (and to all the quadruplexes) for all three quadruplexes; it is notable that in both compounds the side chains contain three -CH_2- groups. The stabilizing effects of compounds **3**, **3a** and **3b** are also closely similar for quadruplex binding. However differences in duplex binding are apparent, with the two compounds **3** and **3a** showing some duplex DNA stabilization even at a low DNA concentration.

Significant differences in melting temperatures between pairs of positional isomers are apparent when side chains are less than two carbon atoms long. Thus comparing pairs of compounds **4** and **4a**, and **5** and **5a**,the melting temperature differences are up to 21°C at a 1 μM quadruplex concentration (**4** vs **4a** with the F21T sequence) and up to 26°C at 2 μM (**5**vs **5a** with the HSP90A sequence). Differences in melting temperatures of up to 7°C were observed between**1** and **1a**, even though their side chains contain equal numbers of -CH_2-groups. These particular differences may be due to the acyclic nature of the OMe- group, with its less bulky and less rigid framework, compared to the cyclic 6-membered N-methyl-piperazine moiety. This result is particularly interesting as it suggests that having an oxygen atom (without a further available nitrogen atom, as in the morpholino-analogues) within the side chain, can provide significant stabilization in the FRET assay, although the cellular assay results suggest that the biological profile of this compound is sub-optimal (see below).

2.3 Short term SRB and telomere length assays

All compounds were screened for their anti-proliferative activities against several cancer cell lines (breast, renal (x2), pancreatic (x2), lung) in a short-term 96 hr sulphorhodamine (SRB) assay (Table 2). A general trend is apparent, that the majority of the new positional isomers, compounds **1**-**5**, are generally more active in these assays than their previously-synthesized counterparts (compounds **1a**-**5a**). It is notable that the change in substituent positions has improved the biological profile of compounds **4a**and **5a**, with their positional isomers, compounds **4**and **5**, having sub-micromolar potencies in the majority of cancer cell lines in the panel. The trend of particular activity in the pancreatic and lung carcinoma cell lines MIA PaCa2 and A549, observed for compounds in the earlier **1a**-**5a** series,[25] is still apparent for the positional isomers. Data is now reported as well (where available) for the chemo-resistant PANC1 pancreatic carcinoma[28] cell line, which shows that the isomer of the previously most active compound **2a**, is five-fold more active in PANC1 than in MIA PaCa2 cells, with a large therapeutic window compared to its activity in the normal human fibroblast cell line WI38.

Possible effects on telomere length were also examined. MIA PaCa2 cells were exposed to compounds **2** and **2a** at various concentrations in the range 1-7 nM for a period of three to five weeks. Cells were treated twice a week and counted at the end of each week. Cell samples were taken at weekly intervals and Southern blot analysis was performed on extracted DNA. No changes in telomere length were found with either compound at any of the time points (blots not shown here).

2.4 Molecular modeling studies

Molecular modeling studies were performed with the most potent positional isomers, the compounds, **2** and **2a**. Figure 2 shows a schematic view of these two simulated quadruplex (G4) ligand complexes; (a) the original X-ray structure of the 21-mer complex with compound **2a** (PDB id 3UYH), and (b) the modeled complex with compound **2** and the 21-mer quadruplex, based on this X-ray structure, obtained by interchanging the N-methyl piperazine and morpholino side chains attached to the naphthalene diimide core. The core itself remained unchanged, which is clearly visible in the structural alignment (c) of the two complexes. An overview of the 50 ns MD simulations is given in Table 3. RMSD values for the G4 backbone atoms as a function of the simulation time were used as a measure of stabilization of the two models (G4-**2a**, G4-**2**), comparing both initial reference, and time-averaged structures of the models (with the latter providing superior insight into the structures reaching the plateau). The trajectories of the G4-ligand complexes were stabilized at around ~3.1 Å (and around ~1.1 Å with respect to the time-averaged structure obtained from the MD) for the G4-**2a** complex at around ~3.0 Å (and at around ~1.9 Å for the time-averaged structure) for the G4-**2** complex. This suggests that **2a** has improved G4-stabilizing properties comparing to its isomer, compound **2**. The trajectory of the native G4 21-mer, which was

used as a control structure, was stabilized at around ~2 Å (and around ~1.2 Å respectively).

Cluster analysis was applied to the large amount of data generated in order to provide a statistical description of the dynamics of the G4-ligand complexes. This analysis used the 1920 frames (~48 ns) extracted from the MD trajectories at time interval of 25 ps for the matrix construction, with a 2.0 Å RMSD cut-off applied for the neighbour search. The resulting clusters are then represented by middle structures, the conformations sampled during the simulation at 300 K. While the G4-**2a** complex was represented by one cluster only (i.e ~ 100% of the conformational space sampled), the first seven clusters were required to cover the total ensemble of sampled conformational space for the G4-**2** complex, and even this only included~85% of the space. Cluster 1 was always the predominant conformer; Figure 3 shows the predominant conformations which were adopted by these two G4-ligand complexes through the simulation time, with respect to their starting structures in Figure 2 (prior to the MD simulations).

In Figure 3(a) the original conformation adopted by compound **2a** in the X-ray structure (yellow) has only slightly changed through the 50 ns MD run, as represented by one predominant conformation (green). One morpholino side chain has moved from its position, becoming closer to the quadruplex loops (the rmsd between the two structures is ~ 1.8 Å); Figures 3 (b-d). In the case of the complex with compound **2**, progressive anti-clockwise rotation of the ligand takes place until the ligand has moved by 90° from its starting structure position. This is represented by seven predominant clusters (sampling ~85% conformational space). Only the three most significant changes/cluster representations are shown here, which follow ligand rotation as the simulation progresses. Calculation of binding free energies for **2** and **2a** has given very similar values.

3. Conclusions

We have generated a small focused library of tetra-substituted naphthalene diimides that are positional isomers of the previously-disclosed series,[25] using a short three-step synthesis utilizing microwave irradiation. Several of the isomers showed very potent stabilization of several representative quadruplexes, with ΔT_m values of up to 34°C at a 1 μM concentration. They compare well with the behaviour of their positional isomers, with one new compound showing superior stabilizing abilities for all three quadruplexes compared to the behavior of its previously-reported isomer, together with negligible duplex stabilization. This compound (**5**), with two pyranose end-groups, is also active in the cancer cell line panel, in contrast to its positional isomer. Several of these compounds are potent inhibitors of cell growth in pancreatic and lung cancer cell lines, with IC_{50} values for two compounds (**2** and **2a**) against the chemo-resistant PANC1 line, of 2-3 nM. Long-term telomere length studies showed no change in mean telomere length, indicating that inhibition of telomerase is not involved in the anti-proliferative response to these agents. Further studies will aim to elucidate the mode of action of these compounds in responsive cell lines, following the observations that compound **2a** produces senescence in affected cells, accompanied by changes in the expression of a number of cellular stress/DNA damage response genes.

In general the ΔT_m values for quadruplex stabilization overall do not correlate with the IC_{50} anti-proliferative data. However in one instance (compound **5)**, it appears that the longer side chain amino functionality attached to the 6-membered diimide moieties of the naphthalene diimide core, results in superior anti-proliferative activity and higher ΔT_m values across the three quadruplexes, compared to compound **5a**. The same pattern of behaviour was not observed for compounds **4** and **4a**, suggesting that the one-carbon atom difference between the furanose and pyranose rings in the two pairs of compounds is sufficient to produce differences in quadruplex stabilization and biological response, assuming that the latter is a consequence of the former. This suggests a concept of '*virtual overall elongation'* by one side chain atom. Having the longer side-chain group (in this case n=3 N-methyl-piperazine) connected via the diimide functionality possibly provides an overall net increase in quadruplex stabilization as the optimal chain length may not necessarily be n=3, but actually is between n=2 and n=3, which is better achieved by having the shorter side chains in the optimal position to interact with the quadruplex structure (Figure 4).

The molecular dynamics simulation studies on the telomeric quadruplex complexes with the positional isomers **2** and **2a** have revealed that an anti-clockwise rotation of ~90° in the G-quartet plane for compound **2**, so that it ends up in the same position as the experimental pose for compound **2a** (Figure 3e). This is in accord with the very similar quadruplex stabilization data from the FRET assay (Table 2) as well as closely similar inhibition of cell proliferation. This equivalence, not yet evaluated by simulation, appears to be the case only when the side-chains are all of equal length, and possibly with the end-groups having similar steric features, as in the N-methyl-piperazine and morpholine groups of **2** and **2a**.

This study has shown that the positional isomer (compound **2**) of the previously-reported[25] lead compound **2a**, does not show any advantages in terms of either quadruplex binding or inhibition of cell proliferation, at least in the lines evaluated to date, and the former compound is currently being evaluated in tumor xenografts. On the other hand, the present study has revealed that the positional isomers of more asymmetrically-shaped compounds such as **3a** and **4a** are significantly more potent and thus may provide new directions for lead compound generation. The fact that clogP values are identical for both members of an isomer pair indicates that differences in cellular behavior are not due to uptake differences but to more subtle differences in binding to their cellular targets.

4. Experimental procedures

4.1 Chemistry

All chemicals, reagents and solvents were purchased from Sigma-Aldrich, Alfa Aesar and Lancaster Synthesis and used without further purification. Solvents were supplied by Aldrich, VWR and Fisher scientific. Microwave reactions were performed in an Initiator microwave (Biotage, Sweden). Column chromatography was performed using BDH silica gel (BDH 153325P). HPLC analysis was carried out with a Gilson apparatus combining a 322 PUMP and an Agilent 1100 SERIES detector, using a C18 5μ (100 x 4.6 mm) column (41622271 (W), YMC, Japan), at a flow of 1 mL/min. Preparative HPLC was carried out with a Gilson apparatus combining a 322 PUMP and a UV/VIS-155 detector with detection at 254 nm, using a C18 5μ (100 x 20 mm) column (201022272) (W), YMC, Japan, at a flow of 20 mL/min. Water and methanol with 0.1 % formic acid were used as solvents for HPLC. NMR spectra were recorded at 400 MHz Bruker spectrometer in $CDCl_3$ (with 0.05 % TMS, Cambridge Isotope Laboratories, USA). NMR spectra were analyzed with MestReC 4.5.6.0 with chemical shifts using TMS as a standard (δ = 0.00 ppm). NMR multiplicity abbreviations are s (singlet), bs (broad singlet), d (doublet), t (triplet), 4q (quartet),

5q (quintet), and m (multiplet). Coupling constants *J* are reported as observed in Hertz (Hz). High Resolution Mass spectra (HRMS) were measured on a Micromass Q-TTOF Ultima Global tandem mass spectrometer run under electrospray ionisation (ESI), and processed using the MassLab 3.2 software. The general procedure for the synthesis of compounds **1-6** has been previously published and is available online.[25] Compounds **1a**-**5a** have been synthesized according to this previously published procedure.

Compound **1**

^{1}H NMR ($CDCl_3$ 400MHz)δ9.43 (2H, t,*J*= 5.5Hz, 2x NH), 8.20 (2H, s, ArH), 4.26 (4H, t, *J*= 7.4Hz, 2x CH_2), 3.66 (4H, q, *J*=6.4, 4.0Hz, 2x CH_2), 3.59 (4H, t, *J*=5.8Hz, 2x CH_2), 3.42 (6H, s, 2x OMe), 2.54(18H, m, 9x CH_2), 2.25 (6H, s, 2x CH_3), 2.07 (4H, m, 2x CH_2), 1.99-1.92 (6H, m,3xCH_2), ^{13}C NMR ($CDCl_3$ 100 MHz) δ 166.2 (2x quat), 163.1 (2x quat), 149.3 (2x quat), 125.9 (2x quat), 121.2 (2x quat), 118.3 (2x CH), 102.1 (2x quat),70.1 (4x CH_2), 58.8 (4x CH_2), 56.0 (2x CH_2), 55.1 (4x CH_2), 52.9 (4x CH_2), 45.9 (2x CH_3), 40.5 (2x CH_2), 38.8 (2x CH_2), 29.5 (2x CH_3), 25.2 (2x CH_2).HRMS (ES^+) calculated for $C_{38}H_{56}N_8O_6$ $[M+H]^+$721.4069, found 721.4401.

Compound **2**

^{1}H NMR($CDCl_3$ 400MHz) δ 9.43 (2H, t, *J*=5.5Hz, 2x NH), 8.20 (2H, s, ArH), 4.26 (4H, t, *J*=7.5Hz, 2x CH_2), 3.80(8H, t, *J*=4.7Hz , 4x CH_2), 3.64 (4H, q, *J*= 6.5, 2.0Hz, 2x CH_2), 2.97 (8H, s, 4x CH_2), 2.79 (8H, s, 4x CH_2), 2.68-2.60(22H, m, 11x CH_2),2.05-1.95(8H, m, 4x CH_2).^{13}C NMR($CDCl_3$100MHz) δ166.1 (2x quat), 163.1 (2x quat), 149.2 (2x quat), 125.8 (2x quat), 121.3 (2x quat), 118.5 (2x CH), 101.9 (2x quat), 66.6 (4x CH_2), 56.1 (2x CH_3), 55.1 (2x CH_2), 53.6 (4x CH_2), 53.3 (4x CH_2), 50.3 (4x CH_2), 43.8 (2x CH_2), 41.3 (2x CH_2), 38.4 (2x CH_2), 26.0 (2x CH_2), 24.6 (2x CH_2).HRMS (ES^+) calculated for $C_{44}H_{66}N_{10}O_6$ $[M + H]^+$ 831.5203, found 831.5203.

Compound **3**

^{1}H NMR ($CDCl_3$ 400MHz)δ 9.42 (2H, t, *J*=5.5Hz, 2x NH), 8.17(2H, s, ArH), 4.35(4H, t, *J*=7.00Hz, 2x CH_2), 3.76 (8H, t, *J*=5.5Hz, 4x CH_2), 3.59 (4H, dd, *J*=12.5,6.6Hz, 2x CH_2) 2.77 (18H, m, 9x CH_2), 2.52 (20H, m, 2x CH_3, 7 x CH_2) 2.06-1.97(4H, m, 2x CH_2). ^{13}C NMR ($CDCl_3$100MHz) δ 166.0 (2x quat), 163.1 (2x quat), 149.3 (2x quat), 125.8 (2x quat), 121.3 (2x quat), 118.5 (2x quat), 101.9 (2x quat),66.8 (4x CH_2), 56.2 (2x CH_3), 55.2 (2x CH_2), 53.9 (4x CH_2), 53.8 (2x CH_2), 51.4 (2x CH_2), 44.2 (2x CH_2), 41.3 (2x CH_2), 37.3 (2x CH_2), 34.4 (2x CH), 26.2 (2x CH_2). HRMS (ES^+) calculated for$C_{42}H_{62}N_{10}O_6$ $[M+H]^+$ 803.4934, found 803.4940.

Compound **4**

^{1}H NMR ($CDCl_3$ 400MHz)δ 9.54(2H, t, *J*=5.3Hz, 2x NH), 8.11(2H, s, ArH), 4.30-4.23 (6H, m, 3x CH_2), 4.04-3.99 (2H, m, 2x CH), 3.91-3.85 (2H, m, 2x CH), 3.71-3.66 (2H, m, 2x CH), 3.56-3.50 (2H, m, 2x CH), 2.80-2.50 (14H, m, 7x CH_2), 2.61(4H, t, *J*=7.1Hz, 2x CH_2), 2.50 (6H, s, 2x NMe), 2.21-2.13 (2H, m, 2x CH), 2.07-1.94 (8H, m, 4x CH_2), 1.80-1.72(2H, m, 2x CH). ^{13}C NMR ($CDCl_3$ 100MHz)δ 166.0 (2x quat) , 163.0 (2x quat), 149.2 (2x quat), 125.3 (2x quat), 121.2 (2x quat), 118.4 (2x quat), 102.0 (2x quat), 68.5 (2x CH_2), 55.2 (2x CH_3), 53.4 (2x CH_2), 50.1 (4x CH_2), 47.3 (2x CH_2), 43.9 (2x CH_2), 38.5 (2x CH), 29.3 (2x CH_2), 25.9 (4x CH_2), 24.6 (2x CH_2).HRMS (ES^+) calculated for $C_{44}H_{64}N_8O_6$ [M + H]$^+$745.4378, found 745.4401.

Compound **5**

^{1}H NMR ($CDCl_3$, 400MHz) δ 9.46 (2H, t, *J*= 5.6Hz, 2x NH), 8.13 (2H, s, 2x ArH), 4.23 (4H,t, *J*= 7.3Hz, 2x CH_2), 4.04 (4H, dd, *J*= 11.5, 3.4 Hz,2x CH_2), 3.48-3.41 (8H, m, 4x CH_2), 2.87-2.79 (8H, m, 4x CH_2), 2.64 (8H, m, 4x CH_2), 2.55(4H, t, *J*= 7.2Hz, 2x CH_2), 1.92 (6H, s, NMe), 1.96-1.87 (6H, m, 2 x CH_3), 1.55 (4H, d, *J*= 12.8 Hz,2x CH_2), 1.45 (4H, m, 2x CH_2).^{13}C NMR ($CDCl_3$, 100MHz) δ 166.2 (2x quat), 163.0 (2x quat), 149.4 (2x quat), 125.8 (2x quat), 121.2 (2x quat), 118.4 (2x quat), 102.0 (2x quat), 67.6 (4x CH_2), 55.1 (2x CH_3), 53.2 (2x CH_2), 50.3 (4x CH_2), 49.1 (2xCH_2), 43.7 (2x CH_2), 38.4 (2x NMe), 35.2 (4x CH_2), 30.9 (4x CH_2). HRMS (ES^+) calculated for $C_{42}H_{60}N_8O_6$ $[M+H]^+$ 773.4714, found 773.4713.

Analytical data for **1a-5a** has previously been reported (see reference 25, with assigned compound numbers **2a**, **2c**, **2d**, **2e** and **2h**)

4.2 FRET assay

The following oligonucleotide sequences, all purchased from Eurofins, were used: F21T: (5'-FAM-GGG TTAGGGTTAGGGTTAGGG-TAMRA-3'), Hsp90A: (5'-FAM-GGGCCAAAGGGAAGGGGTGGG-TAMRA-3'), Hsp90B: (5'-FAM-GGGCGGGCCAAAGGGAAGGGG-TAMRA-3')T-Loop: (5'-FAM-TATAGCTATATTTTTTATAGCTATA-TAMRA-3'). TAMRA (6-carboxytetramethylrhodamine) is the acceptor fluorophore, and FAM (6-carboxyfluorescein) is the donor fluorophore. From 50 μM stock solutions, 400 nM solutions in FRET buffer (60 mM potassium cacodylate pH 7.4) were prepared. The nucleotides were annealed by heating the samples to 95 °C for 10 min and allowing them to cool down to RT within 3 h. 10 mM solutions of the compounds in deionised water were prepared and diluted to double of the required concentrations with FRET buffer. In RT-PCR 96 well plates (MJ Research, Waltham, MA), each well was loaded with 50 μL of nucleotide solution and 50 μL of drug solution.

Drug concentrations of 0.1, 0.5, 1, 2, 3, 4 and 5 were used, and every drug concentration was repeated three times. Measurements were made on a DNA Opticon Engine (MJ Research) with excitation at 450 – 495 nm and detection at 515 – 545 nm. The fluorescence was read at intervals of 0.5 °C in the range 30 – 100 °C. Before each reading the temperature was held constant for 30 s. The raw data were processed using Origin (Version 7.0, OriginLab Corp.). The graphs were smoothed using a 10-point running average and normalized. The melting temperatures were obtained by determining the maxima of the first derivative of the smooth melting curves. The ΔT_m value is the melting temperature difference between the oligonucleotide with ligand and the negative control native sequence in the absence of ligand.

4.3 Cell culture

The cell lines MCF7, A549, MIA PaCa2, PANC1, RCC4, 786-O and WI38, purchased from American Type Cell Culture (ATCC), were maintained in monolayer culture in 75 cm^2 flasks (TPP, Switzerland) under a humidified 5% CO_2 atmosphere at 37 °C. Incubations were also done under these conditions, unless specified otherwise. For the cell lines MCF7 and A549, Dulbecco's MEM medium was used (GIBCO 21969, Invitrogen, UK) supplemented with L-glutamine (2 mM, GIBCO 25030, Invitrogen, UK), essential amino acids (1 %, GIBCO 11140, Invitrogen, UK), foetal calf serum (10 %, S1810, Biosera, UK) and hydrocortisone (0.5 μg/mL, Acros Organics, UK). For MIA PaCa2 and PANC1 cells, Dulbecco's MEM medium was also used, supplemented with L-glutamine (2 mM) and foetal calf serum (10 %). MEM medium (M2279, Sigma, UK) with added L-glutamine (2 mM), essential amino acids (1 %) and foetal calf serum (10 %) was used for the WI-38cell line. RPMI medium1640 (GIBCO 31870, Invitrogen, UK), supplemented with L-glutamine (2 mM), foetal calf serum (10 %) and sodium

pyruvate (1 mM, 11360-039, Invitrogen, UK) was used for the 786-Ocell line. To passage the cells, they were washed with PBS (GIBCO 14040, Invitrogen, UK), treated with trypsine (GIBCO 25300, Invitrogen, UK), and re-seeded into fresh medium, resulting in an initial cell density of approximately $1x10^4$ cells/ mL medium. Cells were counted using a Neubauer haemocytometer (Assistant, Germany) by microscopy on a suspension of cells obtained by washing with PBS, trypsinisation, centrifugation at 8 °C at 8000 rpm for 3 min, and re-suspension in fresh medium.

4.4 Sulforhodamine B (SRB) assay

Cells were counted and diluted to the required concentration in 20 mL medium. For the MCF7, A549, MIA PaCa2 and PANC1 cell lines, 1000 - 4000 cells with 160 µL media were seeded into each well of a 96 well plate (Nunc, Denmark). After incubation for 24 hrs, the compounds to be tested, dissolved in 40 µL of medium were added in a range of concentrations, and the cells incubated for 96 hrs. The medium was then removed and the cells fixed by incubation with TCA (10 %, Sigma-Aldrich, UK) in water for 30 min at 4 °C. After removal of the TCA, the cells were washed with deionised water five times and dried at 60 °C for 1 hr. The cells were then incubated with sulforhodamine B (80 µL, 0.4 % in 1 % acetic acid, Acros Organics, UK) for 15 min at RT. The SRB was removed, the wells washed with 1 % acetic acid (200 µL), and dried at 60 °C for 1 hr. Tris-base (100 µL, 10 mM, Acros Organics, UK) solution was added to each well, and the plates were gently shaken for 5 min. The absorbance at 540 nm was measured with a plate reader (Spectrostar Omega, BMG Labtech, Germany). The data were normalized to the value of 100 for the control experiment (untreated cells), and the IC_{50} values were obtained by interpolation from a plot with Origin (Version 7.0, OriginLab Corp.), as the concentration required for a reduction in absorbance intensity of 50%.

4.5 Computational studies

The crystal structure[25] of an intramolecular human telomeric DNA G-quadruplex complexed with the naphthalene diimide compound **2a** (pdb id 3UYH; 1.95 Å resolution) was used as a starting point for the two quadruplex-ligand complex molecular dynamics simulations. (Although reported as a 22-mer the flanking 5' terminal residue was not clearly observed in the crystal structure, hence the truncated 21-mer was used here). Consecutive K^+ ions vertically aligned within the central core of the quadruplex mid-way between each G-quartet were retained at their respective crystallographic positions, while the K^+ ion outside of the central core of the G-tetrad (at the 5'-site) was removed. The ligand, compound **2a**, was located at the 3'-terminal G-quartet site of the 21-mer G-quadruplex.

The structure of the positional isomer, compound **2**, was obtained by interchanging the side-chains using the Discovery Studio Visualizer program(www.accelrys.com). The overall net charge of both ligands was assigned as +2, and geometry was optimized by a short cycle of energy minimization employing the fast Dreiding force field within the Accelrys Discovery Studio.

In total, three 50 ns molecular dynamics (MD) simulations were performed; (1) on the 21-mer native quadruplex[29] (as a reference), (2) and (3) on 21-mer complexes of **2** and **2a** (with the ligand binding poses in both cases corresponding to the crystal structure). All full-atom simulations were performed with the GROMACS v 4.5.3 program,[30] employing the *parmbsc0* force field[31] previously ported into GROMACS. The topologies and other parameters for the small-molecule fragments were obtained via the ACPYPE tool, employing the ANTECHAMBER module of the AMBER[32] program with the GAFF force field. All MD protocols were kept identical for consistency of the results. Explicit solvent simulations were performed at T=300K with a time constant for coupling of 0.1 ps under the control of a velocity rescaling thermostat, and isotropic constant-pressure boundary conditions controlled by the Parinello-Rahman algorithm of pressure coupling. Long-range electrostatics were calculated using the PME algorithm with grid spacing of 1.17 Å, and the LINCS algorithm was employed to constrain all bonds. Non-bonded van der Waals interactions were treated in terms of Lennard-Jones 12-6 potential with a 10.0 Å cutoff. The solute was soaked in a triclinic box of TIP3P water with a minimal clearance of 20.0 Å between periodic images for the starting configurations. Additionally, positively-charged K^+ counter-ions were included in the systems to neutralize the negative net-charge on the DNA backbone. In each of the MD runs, there were two temperature-coupling groups; DNA with the structural K^+ ions (and ligand, when present), and water with counter-ions. The systems were then subjected to 5,000 steps of potential energy minimization, followed by 200 ps of molecular dynamics at 200K while keeping the solutes constrained, and a further 100 ps of MD during which the systems were slowly heated to 300K and further equilibrated prior to unconstrained 50 ns production-level MD trajectory calculations. The time-step applied was 2.0 fs with coordinates saved every 5.0 ps. All MD simulations were computed on in-house Linux 64-bit Intel Core-i7 workstations, with efficient parallel scaling and double-precision calculations to prevent any energy conservation and stability issues. Trajectories were analyzed employing the programs in the GROMACS 4.5.3 suite package and visualized by means of the VMD[33] and PyMol[34] programs. The initial 2 ns were rejected for the analysis of the MD trajectories.

Cluster analysis methods previously described[35] were applied here, to identify clusters of structures in a trajectory by means of the GROMOS agglomerative clustering algorithm (as implemented in the GROMACS clustering utility), with a rmsd cutoff distance of 2.0 Å applied for two structures considered to be neighbors.

The MM/PB(GB)SA method[36] was used to compute the relative free energies of binding, employing the thermodynamic cycle that combines molecular mechanical (MM) energies with the implicit solvent methods. This method takes advantage of multiple snapshots from a trajectory, to provide an average of energies. The change in free energy for the molecules upon complex formation was calculated (for each of the snapshots) as the difference in free energy between their bound and unbound states, as given in equation 1, free energy of each term was then calculated according to equation 2:

$$\Delta G_{bind} = G_{(G4\text{-}LIG\ complex)} - G_{(G4)} - G_{(LIG)} \quad (1)$$

$$\Delta G = \Delta E_{MM} + \Delta G_{SOL} - T\Delta S \quad (2)$$

where the molecular mechanics energy (E_{MM}) term is a sum of the internal energy (bonds, angles and dihedrals), electrostatic energy and van der Waals term; while the G_{SOL} term accounts for the solvation energy, comprising both polar and nonpolar component. The polar part of the solvation term accounts for the electrostatic contribution to solvation and was calculated using both the Poisson-Boltzmann (PB) model, and the generalized-Born (GB) model[37] (igb=2; model GB^{OBC1}) with rescaled effective Born radii, accounting for interstitional spaces between atom spheres. Solvent probe radius, dielectric constants grid-spacing parameters were kept at default values. The entropy term

(TΔS) was not included in our simulations as different conformations and binding poses of a ligand were explored. The free-energy calculations on both quadruplex complexes were performed employing the single-trajectory approach over 48 ns, representing 1920 frames of the simulation.

Acknowledgments and notes

Sheila Mpima and Stephan A. Ohnmacht contributed equally to this work. We are grateful for support to Cancer Research UK (Program Grant No. C129/A4489 to SN), and the Pancreatic Cancer Research Fund (project grant to SN).

References

1. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res*. **2006**, *34*, 5402-5415.
2. Phan, A. T. *FEBS J*. **2010**, *277*, 1107-1117.
3. Bochman, M. L.; Paeschke, K.; Zakian, V. A. *Nat. Rev. Genetics* **2012**, *13*, 770-780.
4. De Cian, A..; Lacroix, L..; Douarre, C.; Temime-Smaali, N.; Trentesaux, C., Riou, J.-F.; Mergny, J.-L. *Biochimie* **2008**, *90*, 131-155.
5. Balasubramanian, S.; Hurley, L. H.; Neidle, *S. Nat. Rev. Drug Discovery* **2011**, *10*, 261-275.
6. Müller, S.; Kumari, S.; Rodriguez, R.; Balasubramanian S. *Nat. Chem*. **2010**, 2, 1095-1098.
7. Biffi, G.; Tannahill, D; McCafferty, J.; Balasubramanian, S. *Nat. Chem*. **2013**, *5*, 182-186.
8. Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593-11598.
9. Rodriguez, R.; Miller, K. M.; Forment, J. V.; Bradshaw, C. R.; Nikan, M.; Britton, S.; Oelschlaegel, T.; Xhemalce B.; Balasubramanian, S.; Jackson, S. P. *Nat. Chem. Biol*. **2012**, *8*, 301-310.
10. Neidle, S. *Therapeutic Applications of Quadruplex Nucleic Acids*. Academic Press, San Diego, **2011.**
11. (a) Monchaud, D.; Teulade-Fichou, M.-P. *Org. Biomol. Chem*. **2008**, *6*, 627-636. (b) Georgiades, S. N.; Abd Karim, N. H.; Suntharalingam, K.; Vilar, R. *Angew. Chem. Int. Ed. Engl*. **2010** *49*, 4020-4034; Ralph, S. F. *Curr. Top. Med. Chem*. **2011**, *11*, 572-590.
12. Di Antonio, M.; Doria, F.; Richter, S. N.; Bertipaglia, C.; Mella, M.; Sissi, C.; Palumbo, M.; Freccero, M. *J. Am. Chem. Soc*. **2009**, *131*, 13132-13141
13. Nadai, M.; Doria, F.; Di Antonio, M.; Sattin, G.; Germani, L.; Percivalle, C.; Palumbo, M.; Richter, S. N.; Freccero, M. *Biochimie* **2011**, *93*, 1328-1340
14. Peduto, A.; Pagano, B.; Petronzi, C.; Massa, A.; Esposito, V.; Virgilio, A.; Paduano, F.; Trapasso, F.; Fiorito, F.; Florio, S.; Giancola, C.; Galeone, A.; Filosa, R. *Bioorg. Med. Chem*. **2011**, *19*, 6419-6429.
15. McKnight, R. E.; Reisenauer, E.; Pintado, M. V.; Polasani, S. R.; Dixon, D. W. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 4288-4291.
16. Doria, F.; Nadai, M.; Folini, M.; Di Antonio, M.; Germani, L.; Percivalle, C.; Sissi, C.; Zaffaroni, N.; Alcaro, S.; Artese, A.; Richter, S. N.; Freccero, M. *Org. Biomol. Chem.* **2012**, *10*, 2798-2806.
17. Doria, F.; Nadai, M.; Sattin, G.; Pasotti, L.; Richter, S. N.; Freccero, M. *Org. Biomol. Chem*. **2012**, *10*, 3830-3840.
18. Cuenca, F.; Greciano, O.; Gunaratnam, M.; Haider, S.; Munnur, D.; Nanjunda, R.; Wilson, W. D.; Neidle S. *Bioorg. Med. Chem. Lett*. **2008**, *18*, 1668-1673.
19. Gunaratnam, M.; Cuenca, F.; Neidle S. Naphthalene diimide compounds interacting with G-quadruplex regions in DNA. PCT/GB08/51131.
20. Gunaratnam, M.; Swank, S.; Haider, S. M.; Galesa, K.; Reszka, A. P.; Beltran, M.; Cuenca, F.; Fletcher, J. A.; Neidle, S. *J. Med. Chem*. **2009**, *52*, 3774-3783.
21. Hampel, S. M.; Sidibe, A.; Gunaratnam, M.; Riou, J.-F.; Neidle, S. *Bioorg. Med. Chem. Lett*. **2010**, *20*, 6459-6463.
22. Gunaratnam, M.; de la Fuente, M.; Hampel, S. M.; Todd, A. K.; Reszka, A. P.; Schatzlein, A.; Neidle, S. *Bioorg. Med. Chem.* **2011**, *19*, 7151-7157.
23. Collie, G. W.; Promontorio, R.; Hampel, S. M.; Micco, M.; Neidle, S.; Parkinson, G. N. *J. Am. Chem. Soc*. **2012**, *134*, 2723-2731.
24. Hampel, S. M.; Pepe, A.; Greulich-Bode, K. M.; Malhotra, S. V.; Reszka, A. P.; Veith, S.; Boukamp, P.; Neidle, S. *Molec. Pharmacol*. **2013**, *83*, 470-480.
25. Micco, M.; Collie, G. W.; Dale, A. G.; Ohnmacht, S. A.; Pazitna, I.; Gunaratnam, M.; Reszka, A. P.; Neidle, S. *J. Med. Chem*. **2013**, DOI: 10.1021/jm301899y.
26. Ohnmacht, S. A.; Micco, M.; Petrucci, V.; Todd, A. K.; Reszka, A. P.; Gunaratnam, M.; Carvalho, M. A.; Zloh, M.; Neidle, S. *Bioorg. Med. Chem. Lett.***2012**, *22*, 5930-5935.
27. Guyen, B.; Schultes, C. M.; Hazel, P.; Mann J.; Neidle, S. *Org. Biomol. Chem* .**2004**, *2*, 981-988.
28. Fryer, R. A.; Barlett, B.; Galustian, C.; Dalgleish, A. G. *Anticancer Res*. **2011**, *31*, 3747-3756.
29. Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876-880.
30. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; van der Spoel, D.; Hess, B.;Lindahl, E. *Bioinformatics* **2013**, *29*, 845-854.
31. Pérez, A.; Marchán, I.; Svozil, D.; Sponer, J.; Cheatham, T. E.; Laughton, C. A.; Orozco, M. *Biophysical J*. **2007**, *92*, 3817-3829.
32. Case, D. A.; Darden, T. A.; Cheatham, T. E. III.; Simmerling, C. L; Wang, J.; Duke, R. E.; Luo, R.; Walker, R .C.; Zhang, W.; Merz, K. M.; Roberts, B.; Hayik, S.; Roitberg, A.; Seabra, G.; Swails, J.; Goetz, A. W.; Kolossváry, I.; Wong, K. F.; Paesani, F.; Vanicek, J.; Wolf, R. M.; Liu, J.; Wu, X.; Brozell, S. R.; Steinbrecher, T.; Gohlke, H.; Cai, Q.; Ye, X.; Wang, J.; Hsieh, M.-J.; Cui, G.; Roe, D. R.; Mathews, D. H.; Seetin, M. G.; Salomon-Ferrer, R.; Sagui, C.; Babin, V.; Luchko, T.; Gusarov, S.; Kovalenko, A.; Kollman, P. A. (2012), AMBER 12, University of California, San Francisco, USA.http://ambermd.org
33. Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33-38.
34. DeLano, W. L. 2008, The PyMOL Molecular Graphics System. DeLano Scientific LLC, Palo Alto, CA, USA. http://www.pymol.org
35. Husby, J.; Todd, A. K.; Haider, S. M.; Zinzalla, G.; Thurston, D. E.; Neidle, S. *J. Chem. Inf. Model.* **2012**, *52*, 1179-1192.
36. Gohlke, H.; Case, D. A. *J. Comput. Chem.* **2004**, *25*,238-250.
37. Onufriev, A.; Bashford, D.;Case, D. A. *Proteins* **2004**, *55*, 383-394.

Table 1. ΔT_m values (°C) for a human telomeric quadruplex sequence (F21T), two HSP90 quadruplex sequences, and a representative duplex DNA sequence (T loop), at 1 and 2 μM ligand concentration (with 60 mM K^+ at pH 7.4). Esds are ± 0.5°C, estimated from multiple readings. Data for compounds **1a**-**5a** has been taken from reference 25.

Compound	F21T		Hsp90 A		Hsp90 B		T loop	
	1μM	2 μM	1 μM	2 μM	1 μM	2 μM	1 μM	2 μM
1	<2	16	22	24	22	26	0	0
1a	<2	22	29	32	24	29	0	0
2	29	34	34	36	30	34	5	9
2a	27	32	33	36	29	33	5	9
3	26	31	33	36	28	32	2	8
3a	27	31	34	36	29	34	0	1
3b	28	33	32	36	31	34	8	12
4	<2	15	21	23	13	17	0	0
4a	23	26	27	30	21	28	0	0
5	<2	22	2	27	3	23	0	0
5a	<2	<2	<2	<2	<2	<2	0	0

Table 2. Growth inhibition data for compounds **1**-**5** and **1a**-**5a** for a panel of cancer cell lines and a normal human fibroblast line (WI-38), measured by the SRB assay for 96 h exposure and expressed as IC_{50} values in μM. Data for compounds **1a**-**5a** has been previously reported,[25] apart from the response in the PANC1 cell line The panel of cancer carcinoma cell lines used includes: A549 (lung), RCC4 and 786-0 (renal), MIA PaCa2 and PANC1(pancreatic), and MCF7 (breast). Data for the two most active compounds are highlighted in mauve.

Compound	RCC4	786-O	MCF7	MIA PaCa2	A549	PANC1	WI38
1	7.60 ±0.01	16.10 ±0.63	17.50 ±0.98	7.05 ±0.83	5.75 ±0.51	2.50 ±0.34	>25
1a	10.51 ±0.14	7.17 ±0.41	5.62 ±0.15	2.79 ±0.09	2.54 ±0.01	n/a	3.32 ±0.50
2	0.61 ±0.07	0.44 ±0.07	0.014 ±0.001	0.050 ±0.002	0.007 ±0.001	**0.002 ±0.001**	0.30 ±0.023
2a	0.56 ±0.05	0.32 ±0.01	0.070 ±0.007	0.010 ±0.01	0.019 ±0.005	**0.003 ±0.0005**	0.23 ±0.01
3	0.30 ±0.016	0.16 ±0.021	0.021 ±0.001	0.005 ±0.001	0.02 ±0.001	0.005 ±0.001	0.18 ±0.002
3a	1.75 ±0.18	0.63 ±0.06	0.17 ±0.01	0.04 ±0.01	1.55 ±0.02	n/a	0.61 ±0.02
3b	0.28 ±0.06	n/a	0.03 ±0.01	0.01 ±0.01	0.023 ±0.006	n/a	2.46 ±0.02
4	3.00 ±0.17	0.39 ±0.02	0.56 ±0.034	0.095 ±0.003	0.16 ±0.002	1.50 ±0.17	2.3 ±0.08
4a	>10	>10	>10	5.65 ±0.93	5.57 ±0.87	n/a	>10
5	2.20 ±0.015	0.29 ±0.002	0.51 ±0.017	0.137 ±0.031	0.46 ±0.01	1.10 ±0.13	2.8 ±0.23
5a	8.38 ±0.5	1.20 ±0.03	3.12 ±0.13	2.50 ±0.01	2.92 ±0.01	n/a	12.65 ±0.11

Table 3. Summary of the simulations and MD simulation statistics together with calculated binding free energies for the lead compound **2** and its positional isomer **2a**.

Simulation system	Timescale	RMSD [Å]		Cluster analysis*	ΔG_{bind} [kcal/mol]	
		G4 vs. initial	G4 vs. avg		MM/PBSA	MM/GBSA
G4-**2a**	50 ns	3.09	1.09	1 cluster/ 100%	-50.48	-38.72
G4-**2**	50 ns	2.99	1.94	7 clusters/ 84%	-45.91	-37.59
Native G4 21-mer	50 ns	1.98	1.24	-	-	-

Figure legends

Scheme. Outline of the synthesis of tetra-substituted naphthalene diimides.

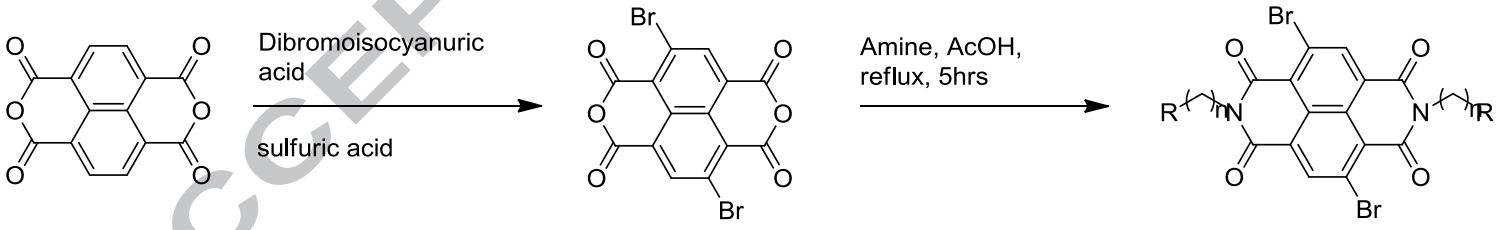
Figure 1. The group of tetra-substituted naphthalene diimide compounds reported and discussed in this study.

Figure 2. Structures of the two G4-ligand complexes, (a) G4-**2a**, (b) G4-**2**, and (c) structural alignment of the two complexes prior to the MD simulations. The DNA structure is shown in cartoon representation, and the ligands, **2** (purple) and **2a** (yellow) are in stick representation. Oxygen atoms are colored red, and nitrogen atoms are blue.

Figure 3. The predominant conformations adopted by the two G4-ligand complexes through the simulation time, with respect to their starting structures (prior to the MD simulations). (a) The G4-**2a** complex at the start of the simulation (yellow), and the sole predominant conformer formed during the simulation (green). (b-d) Show an anti-clockwise rotation of ~90° in the G-quartet plane for compound **2** (cyan) with respect to its starting-structure/conformation (purple). (b) The first 20-25 ns of the simulation time with the starting rotation highlighted by a red arrow; (c) at ~35 ns, a significant rotation of the ligand occurs, as shown, and (d) from ~45 ns onwards, ligand rotation is practically complete, with the morpholino side-chains replacing the N-methyl piperazine ones, and vice versa. (e) Shows the final rotated conformation of compound **2** (cyan) superimposed onto the initial, reference structure of compound **2a** (yellow). The cluster representations are always aligned and superimposed onto the starting structures, in order to retain the perspective of the conformational change. The DNA of the reference (ie initial) structure/conformation is colored orange, while the individual cluster representations are colored white (G4-DNA only).

Figure 4. Visualization of the differences in side-chain linker length, based on the substitution patterns around the ND scaffold - suggesting an increased circular shape (as opposed to elliptical) for the most biologically potent analogues, with equal-length side-chains.

Scheme



	R1	R2
1	NMePip (n=3)	OMe (n=3)
1a	OMe (n=3)	NMePip (n=3)
2	NMePip (n=3)	Morph (n=3)
2a	Morph (n=3)	NMePip (n=3)
3	NMePip (n=2)	Morph (n=3)
3a	Morph (n=2)	NMePip (n=3)
3b	Morph (n=3)	NMePip (n=2)
4	NMePip (n=3)	2-THF (n=1)
4a	2-THF (n=1)	NMePip (n=3)
5	NMePip (n=3)	4-THP (n=1)
5a	4-THP (n=1)	NMePip (n=3)

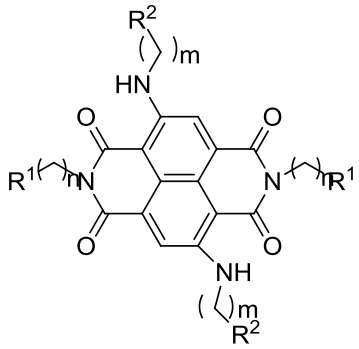


Figure 1

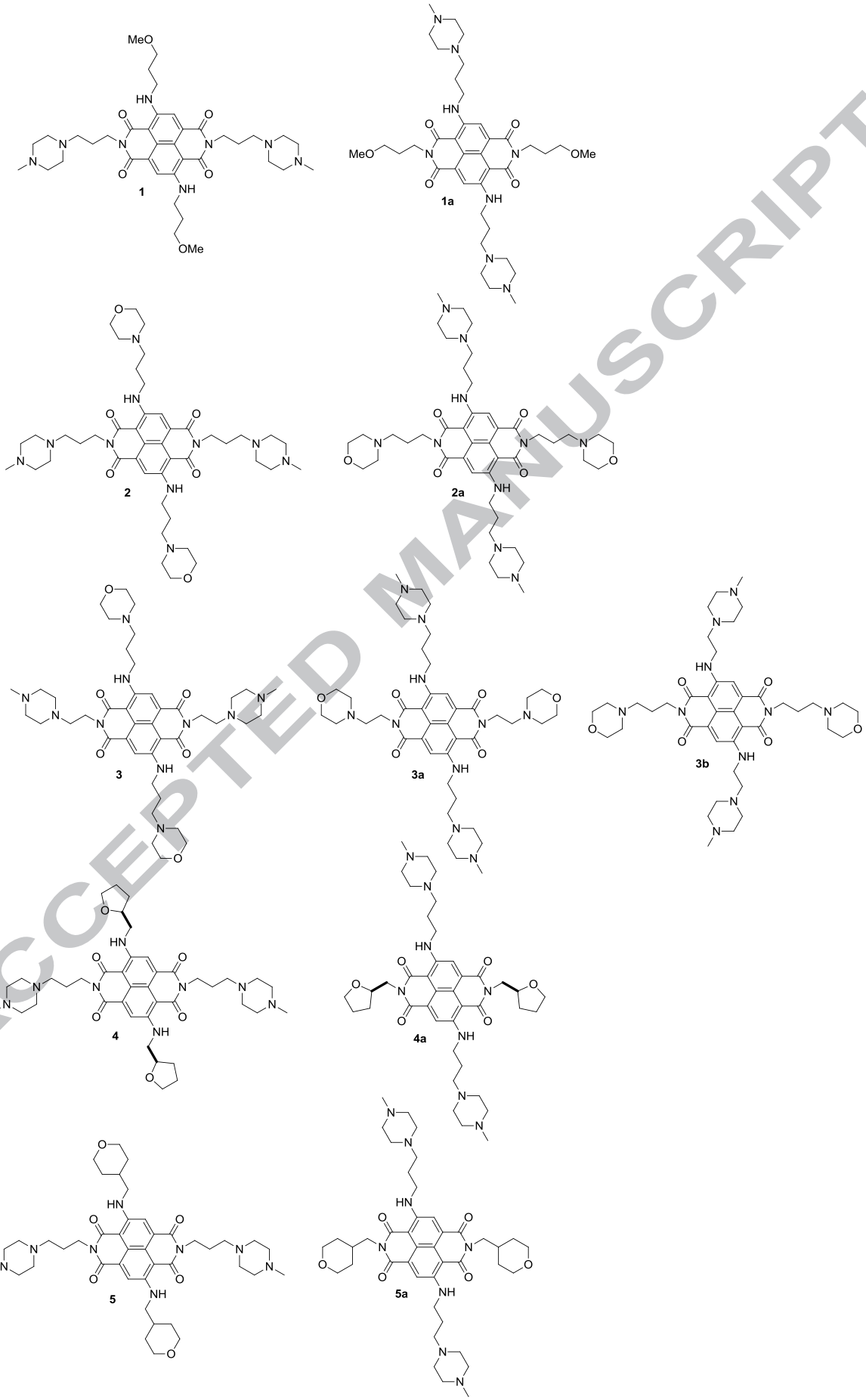


Figure 2

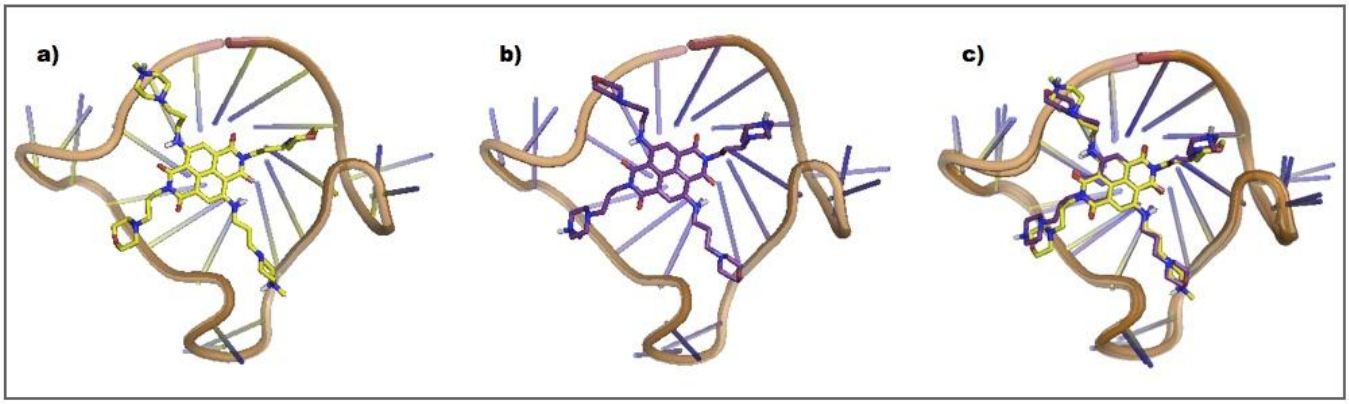


Figure 3

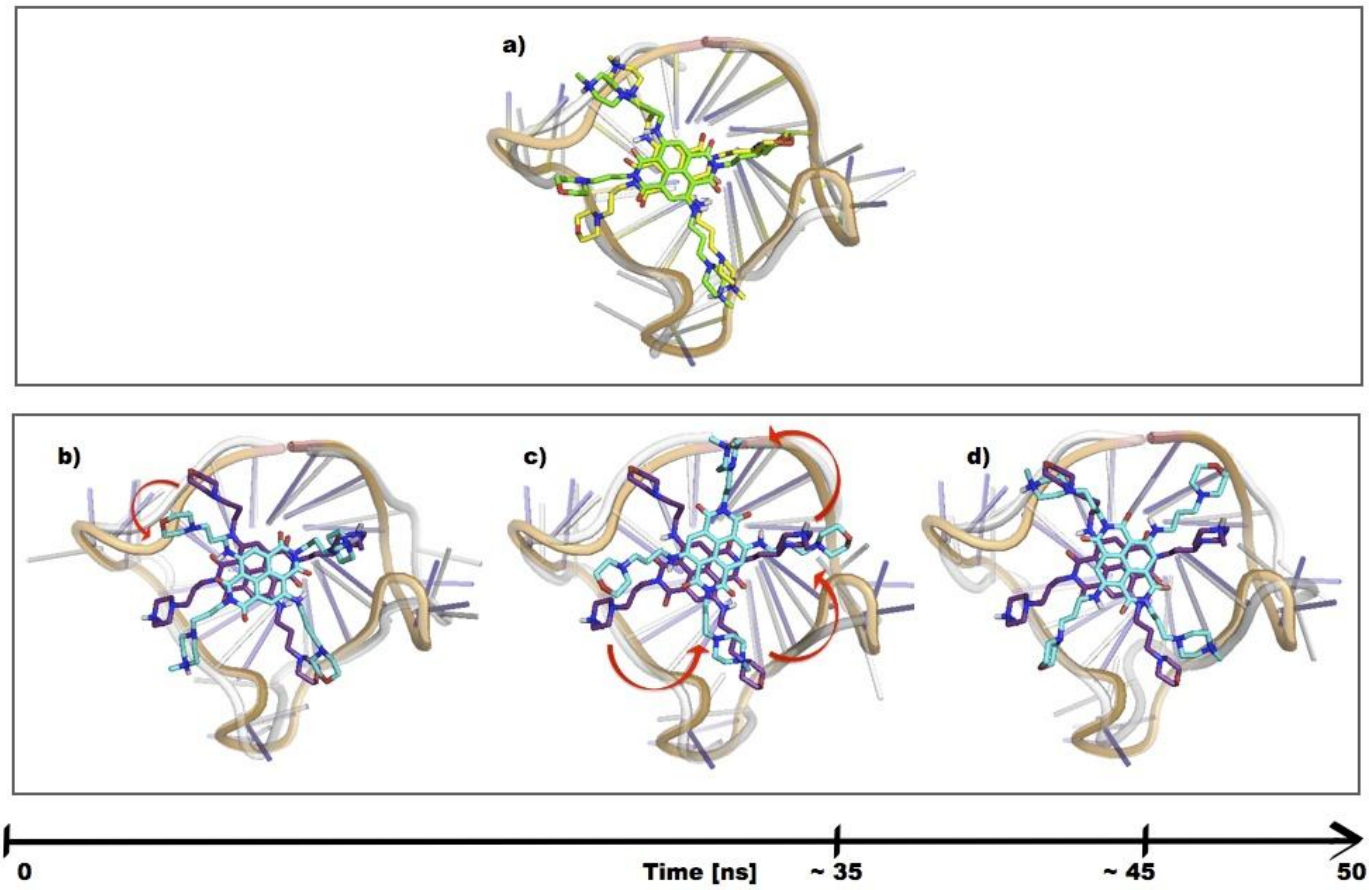


Figure 4