Supporting Information for:

Interactions of a Platinum-Modified Perylene Derivative with the Human Telomeric G-Quadruplex

Lu Rao, Joshua D. Dworkin, William E. Nell, and Ulrich Bierbach*

Contents **page**

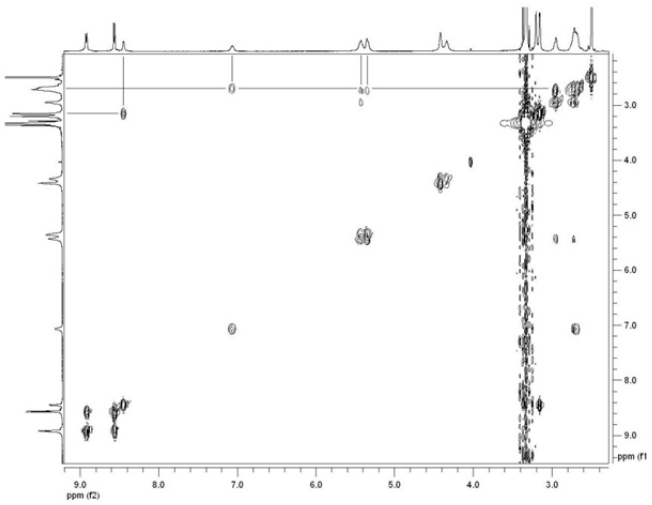

Figure S1. 2-D COSY spectrum (500 MHz, DMSO-*d6*) of compound **7**.

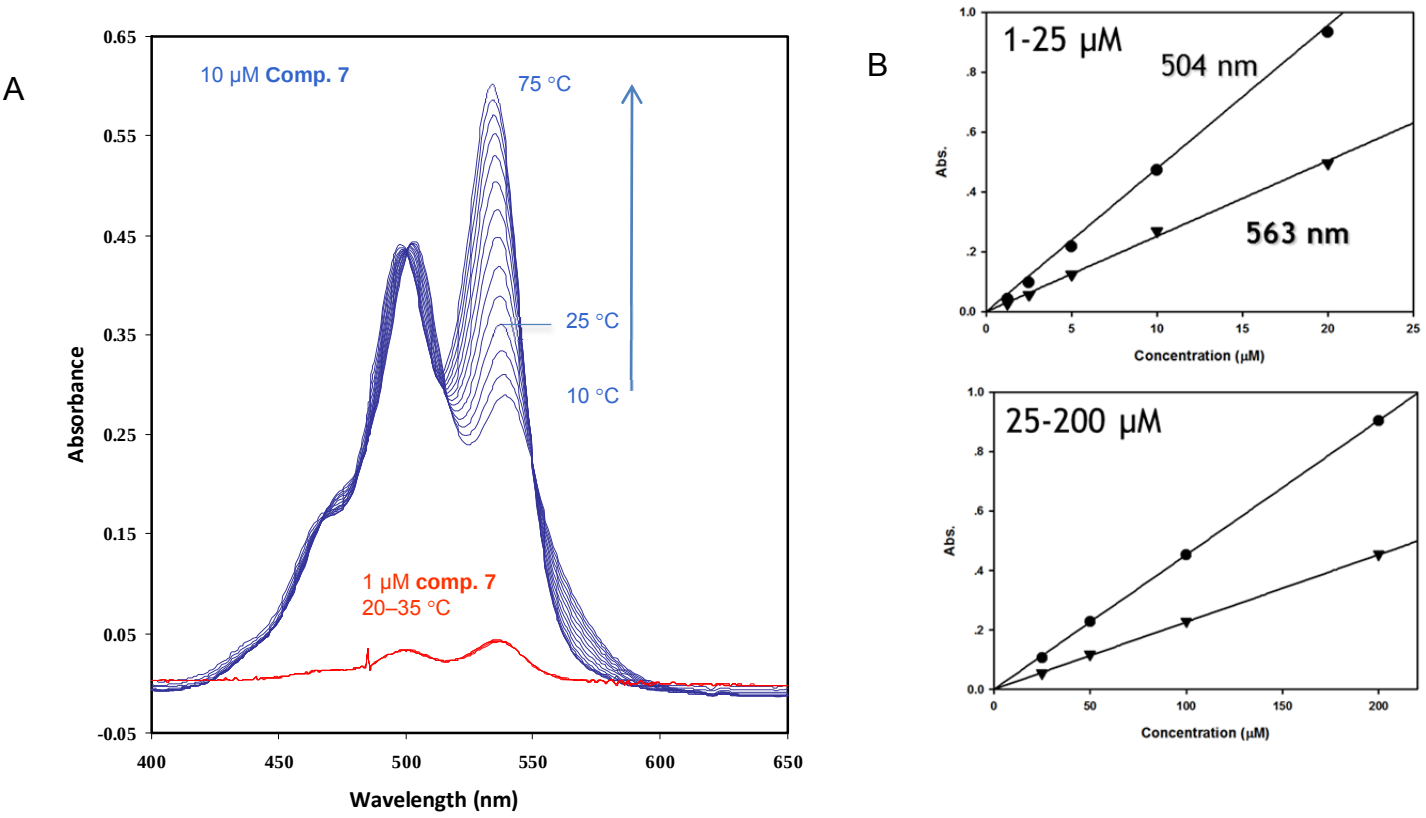

Figure S2. (A) Variable-temperature UV-visible spectra recorded for 1 µM and 10 µM solutions of compound **7** in 20 mM potassium phosphate buffer (35 mM K^+). The blue traces show that at a concentration of 10 µM compound 7 exists in a temperature-dependent equilibrium of monomers and dimers. By contrast, no hyperchromicity of the 550 nm band is observed for the red traces with an increase in temperature, suggesting that at a concentration of 1 µM the compound has dissociated into monomers. (B) Plots of absorbance vs. ligand concentration for 1–25 µM and 25–200 µM solutions of compound **7** in the same buffer. Solutions were prepared by serial dilution and spectra recorded in 10-mm or 1-mm quartz cuvettes, respectively. The data suggests that compound **7** exists as dimers over a wide concentration range but does not undergo extensive aggregation, in which case a distinct decrease in absorptivity/absorbance would be expected with an increase in ligand concentration.

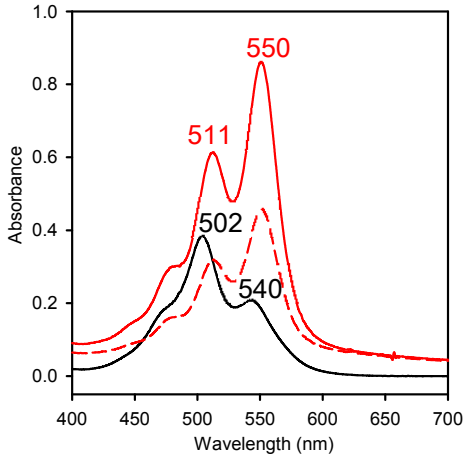

Figure S3. UV-visible spectra (ligand region) of compound **7** titrated into a solution containing the quadruplex formed by the sequence **dG-24** (200 μM nt, 20 mM Tris/20 mM KCl, pH 7.2, 25 °C) at 1:1 (dashed red trace) and 2:1 (solid red trace) ligand-to-DNA ratios. The black trace is the UV-visible spectrum recorded for compound **7** in the absence of DNA.

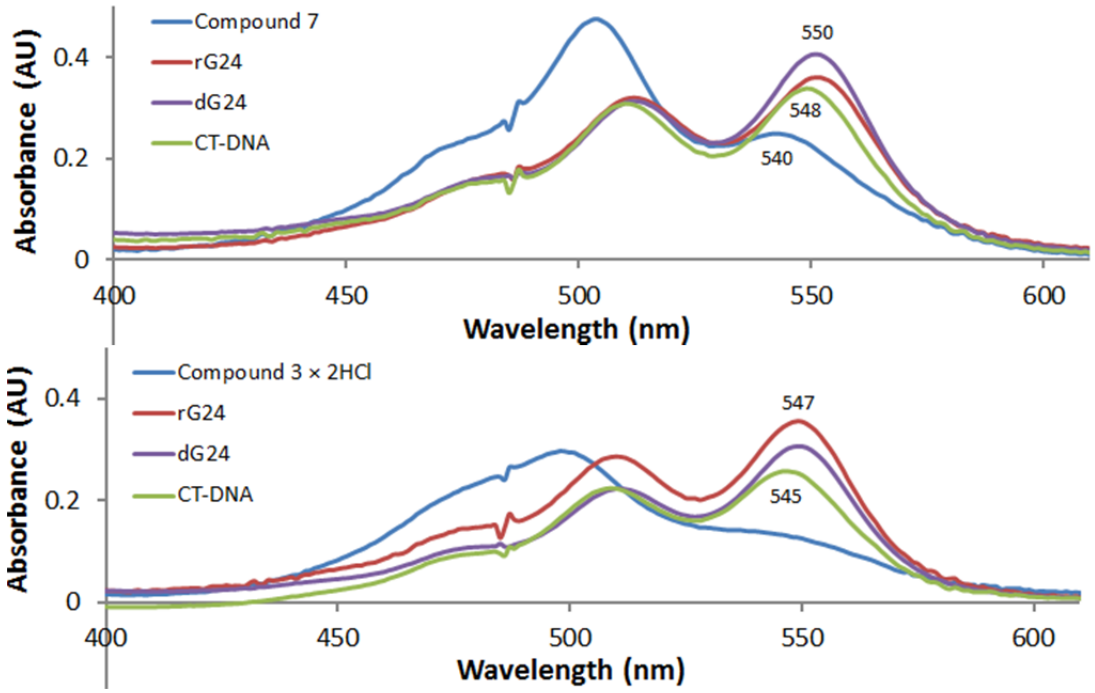

Figure S4. Relative hyperchromicities monitored in the 550 nm region in UV-visible spectra (ligand region) for titrations of 10 μM solutions of compounds **7** (top) and **3**x2HCl (bottom) with one equivalent of **dG-24** and **rG-24** (10 μM strand, corresponding to 240 μM n.t.) and calf thymus (CT) DNA (240 μM n.t.), in 20 mM potassium phosphate, pH 7.2, 25 °C. Note: the greatly reduced absorptivity of free **3**x2HCl compared to free compound **7** (blue traces) in the 500–550 nm range caused by the higher extent of aggregation of the platinum-free perylene (assay based on: Kern, J. T. et al, *Biochemistry*, **2002**, 41, 11379-11389.

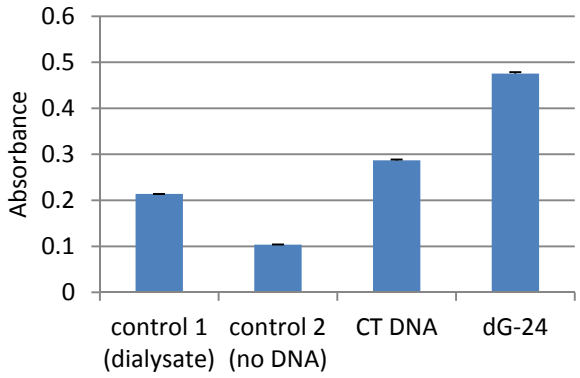

Figure S5. Results of the equilibrium dialysis experiment performed with compound **7** in 20 mM Tris-HCl buffer (pH 7.2) containing 20 mM KCl in a 28-well microdialyzer Solutions containing 200 □M (n.t.) DNA, along with a DNA-free control, were equilibrated against a circulating 7.5 mM solution of compound **7** at 25 °C for 72 h. Absorbances were determined at 504 nm. Error bars indicate standard errors for averages of three individual determinations. Note that control 2 shows inefficient equilibration with the dialysate after 72 h of incubation, which is most likely caused by aggregation of compound 7 on the dialysis membranes. Because of this effect, which was found to be less severe for generated cellulose than for cellulose ester membranes, the assay likely underestimates the amount of DNA-associated ligand.

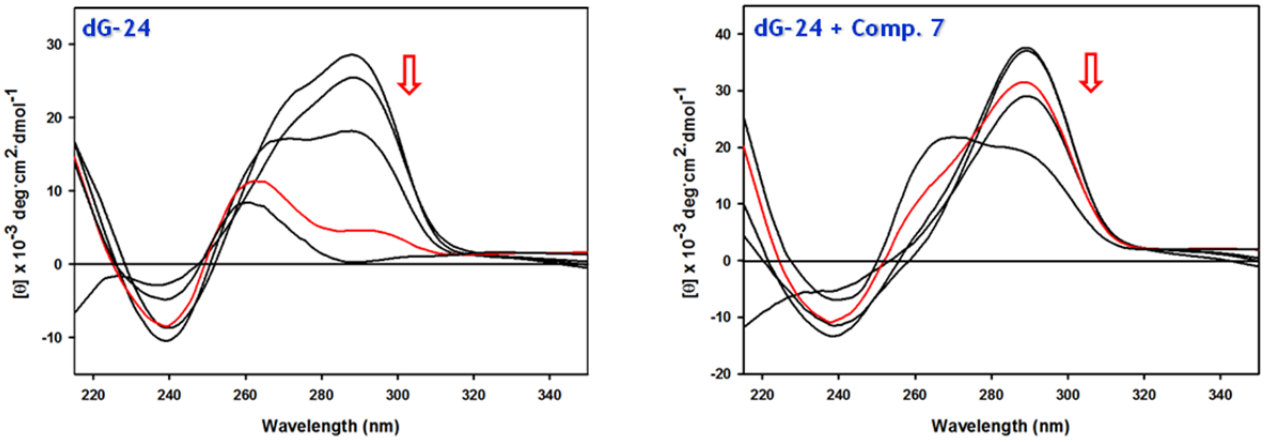

Figure S6. CD spectra recorded for the DNA region of **dG-24** (left; 200 μM nt, 20 mM Tris/20 mM KCl, pH 7.2, 25 °C) and **dG-24** in the presence of one equivalent of compound **7** (right) at 25, 40, 50, 60 (red traces), and 70 °C (traces recorded at 90 °C not shown). Red arrows indicate that the intensity of the positive band at 290 nm decreases as the temperature is increased. Note that, at 60 °C, **dG-24** mainly exists in its denatured, random-coil form (weak positive band at 260 nm), whereas at the same temperature the quadruplex structure persists in the presence of compound **7**. Thus, compound **7** increases the thermal stability of the G-quadruplex formed by the human telomeric DNA. Unfortunately, compound **7** decomposes at temperatures exceeding 70 °C rendering the thermal denaturation process irreversible and precluding the extraction of meaningful ΔT_m values.

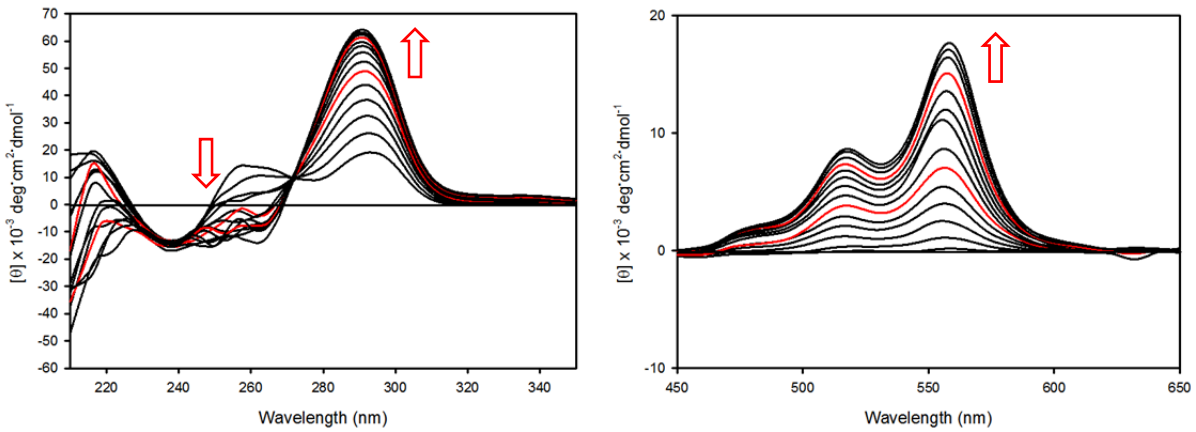

Figure S7. CD spectra in the DNA region (left) and ligand ICD region (right) recorded for the titration of the sequence **dG-24** (200 μM nt, 20 mM Tris, pH 7.2, 25 °C) with 2.6 equivalents of compound **7** in K^+- and Na^+-free buffer. The red traces represent spectra recorded at 1:1 and 2:1 ligand-to-DNA ratios. The red arrows indicate spectral changes with increasing amounts of compound **7**. Note the saturation that occurs at a 2:1 stoichiometry. (Note: the partial conversion of the sequence from random-coil to folded G-quadruplex in the absence of added monocation and compound **7**, based on the CD signal observed in the DNA region at ~290 nm, is most likely caused by residual ammonium ion in the HPLC-purified, dialyzed DNA sample.)

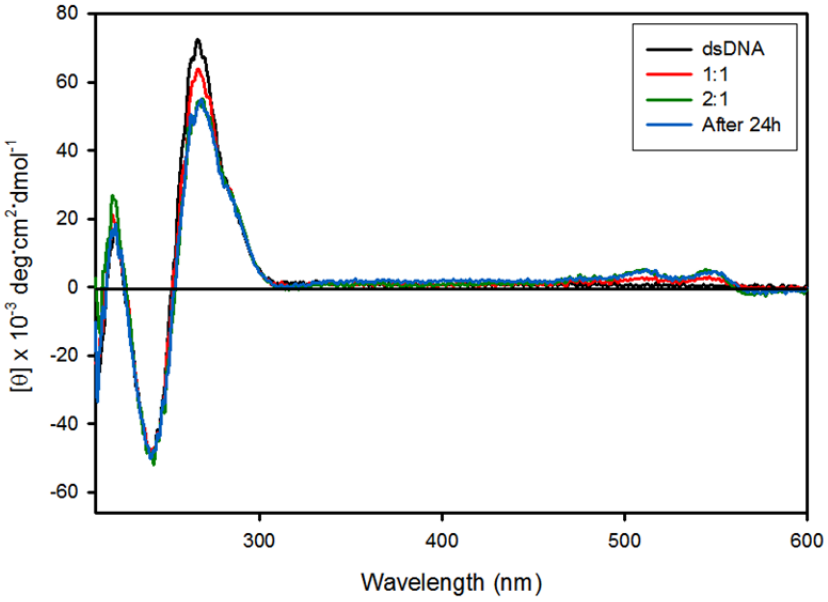

Figure S8. Interaction of compound **7** with the double-stranded telomeric sequence **ds-22** monitored by CD spectropolarimetry (20 mM Tris, 20 mM KCl, pH 7.2). Note the weak negative long-wavelength ICD band observed for the 2:1 incubation ratio, suggesting aggregation of the two ligand molecules on the duplex.

Ligand Binding Models for ITC Studies

dG-24

Initially, all binding equilibria were treated as sequential binding events, for which a fitting routine is available in the Origin software of the ITC program. The relevant equations are listed in the **rG-24** section below.

DNA-DNA Equilibrium Model:

$$M' \xleftrightarrow{K0} M$$
$$M + L \xleftrightarrow{K1} ML$$
$$ML + L \xleftrightarrow{K2} ML_2$$

To better describe the binding of the first two equivalents, a pre-binding equilibrium (K_0) of the two conformations of the telomeric G-quadruplex was included in the list of equations. M' is the disfavored structure of the two forms, while M represents the structure that compound **7** binds to. L represents compound **7**, and the two binding constants, K_1 and K_2, describe the first two binding events.

Calculations:

$$[M_0] = [M'] + [M] + [ML] + [ML_2]$$
$$[L_0] = [L] + [ML] + 2[ML_2]$$
$$K_0 = \frac{[M]}{[M']}$$
$$K_1 = \frac{[ML]}{[M][L]}$$
$$K_2 = \frac{[ML_2]}{[ML][L]}$$

$[M_0]$ is the total concentration of the DNA, and $[L_0]$ is the total concentration of compound **7**.

$$[M_0] = \frac{[M]}{K_0} + [M] + K_1[M][L] + K_1K_2[M][L]^2$$

$$[M] = \frac{[M_0]}{(\frac{1}{K_0} + 1) + K_1[L] + K_1K_2[L]^2} = \frac{\frac{1}{K_1K_2}[M_0]}{\frac{1}{K_1K_2}(\frac{1}{K_0} + 1) + \frac{1}{K_2}[L] + [L]^2}$$

$$[L_0] = [L] + K_1[M][L] + 2K_1K_2[M][L]^2$$

$$[L_0] = [L] + \frac{\frac{1}{K_2}[M_0][L]}{\frac{1}{K_1K_2}(\frac{1}{K_0} + 1) + \frac{1}{K_2}[L] + [L]^2} + \frac{2[M_0][L]^2}{\frac{1}{K_1K_2}(\frac{1}{K_0} + 1) + \frac{1}{K_2}[L] + [L]^2}$$

To simplify this equation, we set:

$$a = \frac{1}{K_1K_2}(\frac{1}{K_0} + 1)$$

$$b = \frac{1}{K_2}$$

$$x = [L]$$

$$[L_0] = x + \frac{b[M_0]x}{a + bx + x^2} + \frac{2[M_0]x^2}{a + bx + x^2}$$

This leads to a cubic equation for x.

$$x^3 + px^2 + qx + r = 0$$

$$p = b + 2[M_0] - [L_0]$$

$$q = a + b([M_0] - [L_0])$$

$$r = -a[L_0]$$

Q represents the total heat generated by the DNA–ligand interactions, and ΔQ is the heat generated in each addition.

$$Q = V_0(\Delta H_0([M] + [ML] + [ML_2]) + \Delta H_1([ML] + [ML_2]) + \Delta H_2[ML_2])$$

$$Q = V_0[M_0]\frac{ex^2 + dx + c}{x^2 + bx + a}$$

$$c = \frac{\Delta H_0}{K_1 K_2}$$
$$d = \frac{\Delta H_0 + \Delta H_1}{K_2}$$
$$e = \Delta H_0 + \Delta H_1 + \Delta H_2$$
$$\Delta Q = Q_i + \frac{dV_i}{V_0}\left(\frac{Q_i + Q_{i-1}}{2}\right) - Q_{i-1}$$

rG-24

Model:

$$2M' \xrightarrow{100\%} M$$
$$M + L \xleftrightarrow{K1} ML$$
$$ML + L \xleftrightarrow{K2} ML_2$$
$$ML_2 + L \xleftrightarrow{K3} ML_3$$
$$ML_3 + L \xleftrightarrow{K4} ML_4$$
$$ML_4 + L \xleftrightarrow{K5} ML_5$$
$$ML_5 + L \xleftrightarrow{K6} ML_6$$

The telomeric RNA G-quadruplex associates to form a dimer in solution (only species present), which is accounted for in the first equation. *M′* and *M* represent the monomeric and dimeric form of **rG-24**, respectively. The interaction between the ligand (*L*) and the dimer was treated as sequential binding events.

Sequential binding sites calculations:

$$K_1 = \frac{[ML]}{[M][L]}$$
$$K_2 = \frac{[ML_2]}{[ML][L]}$$

$$K_3 = \frac{[ML_3]}{[ML_2][L]}$$

...

$$P = 1 + K_1[L] + K_1K_2[L]^2 + \cdots + K_n[L]^n$$

$$F_0 = \frac{1}{P}$$

$$F_1 = \frac{K_1[L]}{P}$$

$$F_2 = \frac{K_1K_2[L]^2}{P}$$

...

$$F_n = \frac{K_1K_2 \cdots K_n[L]^n}{P}$$

$$[L_0] = [L] + [M_0]\sum_{i=1}^{n} iF_i$$

$$Q = V_0[M_0](\Delta H_1F_1 + [\Delta H_1 + \Delta H_2]F_2 + \cdots + [\Delta H_1 + \Delta H_2 + \cdots + \Delta H_n]F_n)$$

$$\Delta Q = Q_i + \frac{dV_i}{V_0}(\frac{Q_i + Q_{i-1}}{2}) - Q_{i-1}$$

Experimental Details for Molecular Mechanics/Molecular Dynamics Simulations

Models of the G-quadruplexes formed by $d(TAGGGT)_2$ (**dG-12**) were generated using published atomic coordinates available for high-resolution structures of relevant quadruplexes. The parallel and antiparallel forms of $(\mathbf{dG\text{-}12})_2$ were generated from the crystal structure of the telomeric sequence $d(AGGG[TTAGGG]_3)$ (PDB ID: 1KF1)[1] and the NMR solution structure of $d(TAGGG[TTAGGG]_3)$ (PDB ID: 2JSM; residues G3–G11),[2] respectively. Both models, which displayed pseudo-two-fold symmetry, were built from appropriately truncated or extended segments of the experimental structures. A model of $[\{Pt(dien)\}_2(\mu\text{-}\mathbf{4}\text{-}S,S')](NO_3)_4$ ($LPt_2(NO_3)_4$, **7**) was built with Discovery Studio. Force field parameters and partial charges for the platinum-modified dimethylthioureaethyl moieties were derived from a published AMBER force field developed for PT-ACRAMTU,[3] which has been demonstrated to accurately reproduce solid state geometry of this pharmacophore.

The DNA starting geometries were energy minimized with the CHARMm27 force field[4] to a final RMS gradient of 1 kcal mol^{-1} $Å^{-1}$ using a distance-dependent dielectric to remove bad contacts and conformational strain prior to the molecular dynamics simulations. Minimizations were performed with four K^+ ions (+1 charge, 1.33 Å) placed along the quadruplex stem and distance contraints between Hoogsteen hydrogen bond donors and acceptors (2.95–3.00 Å, with force constants k_{min} and k_{max} of 100 and 200 kcal mol^{-1} $Å^{-2}$, respectively). The two external K^+ ions were then removed and two molecules of compound 7 (minimized, C_2 symmetry) were manually docked with the outermost G-tetrads of each G-quadruplex such that the perylene was at π-stacking distance and the platinum-containing side chains were positioned in opposite grooves. The ligand–DNA complexes were then energy minimized as described above.

Molecular dynamics (MD) simulations of the ligand–DNA complexes were performed in a TIP3P water box[5] containing 0.2 M KCl. Long-range electrostatics were treated using the Particle Mesh Ewald (PME) method, and hydrogen bonds were constrained with the SHAKE algorithm.[6,7] The pre-minimized systems (100–200 iterations) were then subjected to molecular dynamics simulations. A typical protocol consisted of 5 ps (1-fs steps) of heating from 50 K to 300 K, an equilibration period of 10 ps (1-fs steps) at 300 K and constant pressure, and a production stage of 1 ns (2-fs steps). The conformational space was sampled along the dynamics trajectory during the last 200 ps of the simulation with snapshots taken every 20 ps. Ten conformations of each ligand–DNA complex were minimized in the water box with all constraints gradually removed to a final maximum RMS derivative of 0.1 kcal mol^{-1} $Å^{-1}$. After the water box was removed, the free energy for each structure was calculated using the MM-PBSA method.[8]

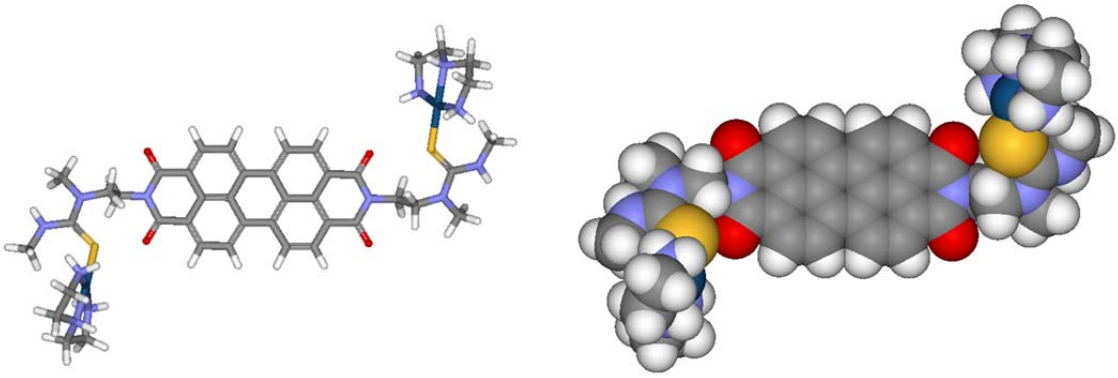

Figure S9. Stick (left) and CPK (left) models of compound **7** (lowest-energy conformation, C_2 symmetry) generated in Discovery Studio (Accelrys, San Diego) using the CHARMm27 force field. This minimized structure, in which the platinum-modified side chains are located on the same side of the perylene chromophore (*syn* conformation), was used in the docking experiments with the bimolecular G-quadruplexes formed by **dG-12**.

(A) **dG-12** (P)

(B) **dG-12** (AP)

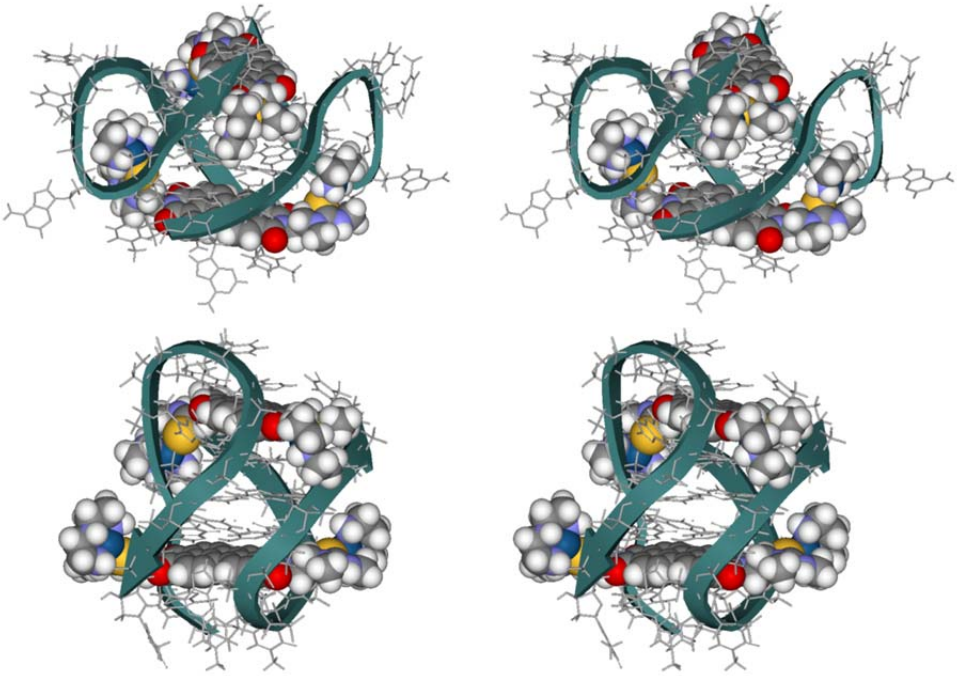

Figure S10. Stereoviews of the simulated structures of the 2:1 ligand–DNA complexes formed between compound 7 and the parallel (P, panel **A**) and antiparallel (AP, panel **B**) **dG-12** G-quadruplex, respectively. The DNA molecules are drawn as gray stick models, the DNA backbone is highlighted as green ribbons, and the ligand molecules are depicted in CPK style.

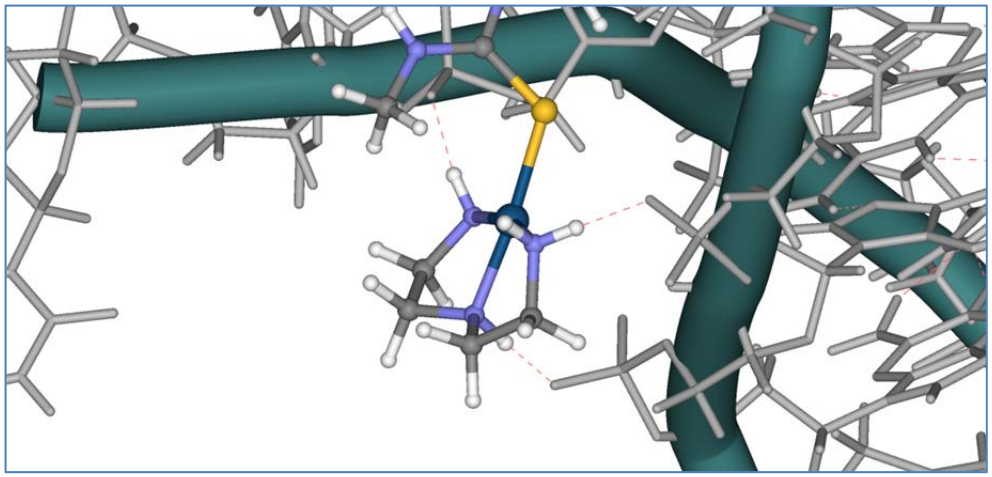

Figure S11. Hydrogen bonding interactions (D···A < 2.5 Å) between a [Pt(dien)] moiety and DNA phosphate observed in the complex formed by **dG-12** (AP).

Discussion of the Modeling Results and the Energy Calculations

Structural aspects

To study possible binding modes and structural details of compound **7** in complex with G-quadruplex DNA, molecular mechanics/molecular dynamics (MM/MD) modeling was performed. When generating the 2:1 complexes formed by compound **7** with bimolecular G-quadruplex DNA, two molecules of compound **7** were docked at van der Waals distance with the outermost G-quartets such that the long perylene dimensions were perpendicular to each other in order to avoid steric clashes in the grooves. In other words, each groove was allowed to only accommodate one platinum moiety. In this arrangement, the stack of three G-quartets of the G-quadruplex structure is sandwiched between two molecules of compound **7**. The sequence **dG-12** (parallel and antiparallel bimolecular G-quadruplexes) was used as model of the DNA secondary structure (Figure S10).

The distance between the two nitrogen atoms in the perylene diimine derivatives is approximately 11 Å, which is close to the dimensions of the G-quartet. Using simulated annealing protocols without distance restraints in molecular dynamics studies, it was demonstrated that the two molecules of compound **7** form stable non-covalent aggregates with the G-quadruplexes. No disruption of the Hoogsteen hydrogen bonding network was observed. The distance between the platinum centers is ~23 Å and the semi-rigid platinum-containing side chains in compound **7** adopted from PT-ACRAMTU are sufficiently flexible to insert into the grooves of the G-quadruplexes. Electrostatic attraction between the positively charged platinum moieties and the negatively charged phosphodiester backbone favor this binding mode. Several ligand–DNA hydrogen bonds involving the dien ligand on platinum and DNA phosphate are observed (Figure S11).

Energy calculations

In this study we demonstrated that compound **7** favors the antiparallel over the parallel G-quadruplex structure formed by the sequence **dG-12**. Hence, we were interested in determining the factors that contribute to the preference of one form of G-quadruplex over the other. To achieve this goal,

molecular dynamics simulations and energy minimizations were performed, and the resulting averaged structures of the ligand–DNA complexes were then used in energy calculations. Specifically, the Molecular Mechanics–Poisson-Boltzmann surface area (MM-PBSA[8]) method was used to calculate the free energy of the systems by combining classical bonding and non-bonding energies (MM) with solvation energies (PBSA). The latter is based on the solvent-exposed surface area of the models. The results of these calculations are summarized in Table S1. Both the solvation energy and non-polar terms seem to favor the antiparallel form. However, overall, the total free energy of the antiparallel and the parallel form are virtually the same, which does not support our experimental findings. One reason for the shortcomings of the MM-PBSA method is that it relies on the quality of the force field parameters and the charge distribution scheme. While the (force field independent) solvation energy seems to be in agreement with the experimental data, the non-optimized set of *ad hoc* parameters may be responsible for the limitations of the MM–PBSA method to predict the energetically favored DNA secondary structure in this system.

Table S1. Thermodynamic Data Extracted from MM–PBSA Calculations on the Two Simulated Ligand–DNA Complexes.

Model[a]	***Potential Energy***[b]	***Van der Waals Energy***	***Electrostatic Energy***	***Solvation Energy***	***Non-polar Term***	***Free Energy***
dG-12 (AP)	–2052.2 ± 26.1	–37.9 ± 5.9	–3224.4 ± 22.1	–2542.7 ± 21.2	28.6 ± 0.5	–4566.3 ± 9.6
dG-12 (P)	–2217.5 ± 32.2	–12.4 ± 8.2	–3398.4 ± 32.1	–2388.6 ± 30.7	30.1 ± 0.3	–4576.0 ± 13.4

[a] **dG-12 AP:** antiparallel structure of the **dG-12** G-quadruplex in complex with two molecules of compound **7**; **dG-12 P**: parallel structure of the **dG-12** G-quadruplex in complex with two molecules of compound **7**. [b]Energies are given in kcal/mol.

Cited Literature

(1) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876-880.
(2) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res* **2007**, *35*, 6517-6525.
(3) Baruah, H.; Wright, M. W.; Bierbach, U. *Biochemistry* **2005**, *44*, 6059-6070.
(4) MacKerell, A. D., Jr.; Banavali, N.; Foloppe, N. *Biopolymers* **2000**, *56*, 257-265.
(5) Takemura, K.; Kitao, A. *J Phys Chem B* **2007**, *111*, 11870-11872.
(6) Hauptman, H. A. *Methods Enzymol* **1997**, *277*, 3-13.
(7) Darden, T.; York, D.; Pedersen, L. *J Chem Phys* **1993**, *98*, 10089-10092.
(8) Fadrna, E.; Spackova, N.; Stefl, R.; Koca, J.; Cheatham, T. E., 3rd; Sponer, J. *Biophys J* **2004**, *87*, 227-242.