Bimodal, dimetallic lanthanide complexes that bind to DNA: the nature of binding and its influence on water relaxivity

Jennifer E. Jones, Angelo J. Amoroso, I. Mihaela Dorin, Giacomo Parigi, Benjamin D. Ward, Niklaas J. Buurma* and Simon J.A. Pope*

Supplementary Information

Docking studies

Analysis of DNA base pair and base-pair-step parameters for structures carrying an intercalated ligand

For the generation of a target structure with an intercalation site for our docking studies, structural information on typical intercalation sites is required. The required structural information was collected both manually and through the 3DNA Landscapes[1] website.

Manual search

DNA structures containing intercalated ligands were identified in the Nucleic Acids Database[2] through the "NDB integrated search", selecting the "biomolecule contains" option. The following selection criteria were used: DNA "Y", RNA "N", Hybrid "N", Protein "N", Ligand "N", Drug "Y", Type of NA Drug Binding "INTERCALATION". All search criteria were combined using the AND operator, yielding 121 results. The results were further analyzed, excluding structures containing mismatched base pairs, "novel binding modes" and metal ions, and requiring structures to be B DNA duplex structures with at least one paired base pair above and below the intercalation site. For the structures remaining after this selection process, the base pair parameters for the base pairs above and below[3] the intercalation site and the base-pair-step parameters corresponding to the intercalation site were collected.[4]

3DNA Landscapes search

Base pair steps likely corresponding to intercalated ligand structures were identified on the 3DNALandscapes[1] website through the following approach. First, a search was carried out applying the following "structure filter": Experimental method "All", Molecular Content: Protein "No", Water "Either", Metal ion "No" and Drug/small molecule "Yes"; Conformational characteristics "All types of DNA structures", X-ray resolution cut-off "Better than 2 Ångstrom", NMR model content "First model". This search resulted in 276 PDB structures. On the Parameter Panel, base-pair-step parameters (Shift, Slide, etc.) were selected. The following constraints were then specified: Nucleotide unit "dimer", Chemical modifications "No", WC or non-WC base pairs "Watson-Crick", Base-pair-step conformational context "All", resulting in selection of 1354 sets of parameters.

The histogram for the rise parameter (Figure S1) shows two clearly separate distributions, one centred around approximately 3.3 Å and one centered around approximately 7.0 Å.

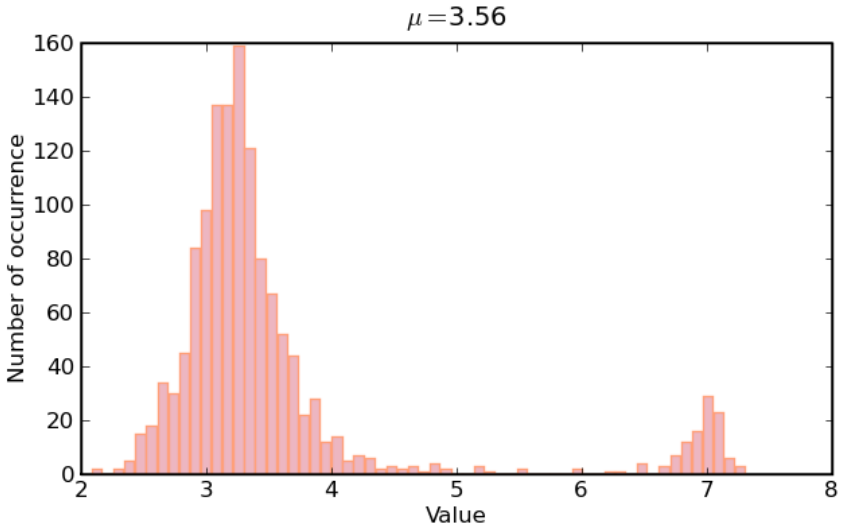


Figure S1: Distribution of the rise parameter for DNA structures with a bound drug/small molecule. Figure generated using 3DNA Landscapes, the x-axis scale corresponds to the base pair step rise in Å.

We attribute the peak around a rise of 7.0 Å to base-pair-steps containing an intercalator.

All 1354 sets of base-pair-step parameters were downloaded and sorted by the rise parameter (Microsoft Excel). All sets of base-pair-step parameters for which rise > 6.4 Å were selected and then divided by base pair step (AC/GT, AG/CT, *etc.*). A remarkable difference in the number of occurrences of the different base pair steps in the sub set of base-pair-step parameters for which rise > 6.4 Å was found (Table S1).

Table S1: Number of occurrences of the different base pair steps among base pair steps for which rise > 6.4 Å.

base pair step[a]	number of occurrences	base pair step	number of occurrences
AC/GT	1	CT/AG	1
AG/CT	1	GC/GC	1
AT/AT	1	GG/CC	2
CA/TG	13	GT/AC	1
CC/GG	1	TG/CA	11
CG/CG	70		

a) base pair steps in 3DNA notation

The CG/CG base pair step is clearly the most frequently encountered base pair step for which rise > 6.4 Å and therefore appears to be the site where intercalation is most likely. Because CG/CG is the base pair step for which most information is available, this step allows the most meaningful statistical analysis of base-pair-step parameters for intercalation. The CG/CG step was therefore selected as the intercalation site in our target DNA structure for our docking studies (*vide infra*).

DNA structural parameters for intercalation sites

Table S2 shows averages and standard deviations for base-pair-step parameters for DNA structures with an intercalated small molecule binder, obtained as described above. Table S2 also shows the base-pair-step parameters for the target base pair step of the unmodified PREDICTOR-generated duplex structure for d(ATCGAGACGTCTCGAT)$_2$ (the "starting structure" in Table S2) and the base-pair-step parameters for the final target structure with an intercalation gap (the "final structure" in Table S2) (*vide infra*).

Table S2: Base-pair-step parameters for DNA structures with an intercalated small molecule binder and base-pair-step parameters for the target base pair step of the PREDICTOR-generated duplex structure for d(ATCGAGACGTCTCGAT)$_2$ and the base-pair-step parameters for the final target structure

Manual search	3DNALandscapes search			
	Rise > 6.4 Å	Rise > 6.4 Å and CG/CG	starting structure	final structure

shift	0.594±0.901	-0.037±1.053	-0.033±1.071	-0.06	-0.06
slide	0.693±0.649	0.876±0.500	0.911±0.408	-0.25	-0.25
rise	6.987±0.178	6.969±0.169	6.971±0.125	3.51	6.97
tilt	0.491±4.724	-0.192±4.689	0.233±2.474	-0.15	-0.15
roll	-1.051±2.739	-1.073±4.252	-1.210±4.024	7.97	-1.21
twist	29.976±10.093	30.347±6.164	30.845±5.597	31.51	30.85

Table S2 shows close correspondence between the base-pair-step parameters for all structures with rise > 6.4 Å and the subset of structures where rise > 6.4 Å and the base pair step is CG/CG, as expected considering data for CG/CG steps dominates the full data set. The correspondence between parameter values from the manual search and the 3DNALandscapes search confirms the validity of our data collection approaches.

Table S2 also shows the base-pair-step parameters for the PREDICTOR-generated structure for d(ATCGAGACGTCTCGAT)$_2$ which served as the basis for the target DNA structure with an intercalation gap as well as the final structure with the intercalation gap (*vide infra*). The parameters shift, slide, and tilt were kept as predicted because the values from the statistical analysis do not differ significantly from the values from the PREDICTOR structure. The value for the rise parameter was changed to correspond to the average value for the results of the 3DNALandscapes search. The values for roll and twist were changed to the average value found for CG/CG base pairs for which rise > 6.4 Å.

Table S3 shows averages and standard deviations for base pair parameters for the base pairs above and below the intercalation site. For comparison, average base pair parameters for C·G and G·C base pairs for PREDICTOR generated structures (for Fox' HexA and HexB sequences[5, 6]) are given. The C·G and G·C base pairs were chosen for comparison as these will dominate the statistics as they did for the base-pair-step parameters and because these are the base pairs above and below the intercalation site in our target structure (*vide infra*). PREDICTOR-generated structures were chosen for comparison because our target structure is generated on the basis of a PREDICTOR-generated structure as well.

Table S3: Average base pair parameters for the base pair above and the base pair below intercalation sites and average base pair parameters for C·G and G·C base pairs.

	manual search[a]		predictor structures	
	bp above	bp below	C·G	G·C
Shear	0.179±0.18	-0.164±0.182	0.166±0.049	-0.166±0.047
Stretch	-0.239±0.39	-0.178±0.124	-0.244±0.035	-0.245±0.033
Stagger	-0.047±0.224	-0.167±0.245	0.083±0.079	0.081±0.079
Buckle	7.242±7.038	-12.222±7.001	0.328±2.563	-0.287±2.547
propeller	-0.473±5.19	0.555±3.423	-10.216±1.387	-10.217±1.433
Opening	-2.263±19.596	-0.045±2.252	-1.289±0.688	-1.247±0.707

a) The terms "above" and "below" are defined by their sequence in the 3DNA output.

The similarity between shear, stretch, stagger and opening for the base pair above the intercalation site and the same parameters for C·G, taken together with the similarity between shear, stretch, stagger and opening for the pair below the intercalation site and the same parameters for G·C, is in line with the observation that data for CG/GC base pair steps dominates the data. The different values for buckle and propeller are as anticipated for intercalation; a buckle away from the intercalation site generates more space for the intercalator, while reducing the propeller twist reduces steric hindrance while optimising stacking interactions between base pairs and flat intercalators.

Based on the data for DNA structures with intercalated small molecules or drugs and the average structures generated by PREDICTOR, the structure parameters for d(ATCGAGACGTCTCGAT)$_2$ were edited to generate parameters for open-d(ATCGAGA**CG**TCTCGAT)$_2$, the target structure for our docking studies (*vide infra*) as shown in Table S4.

Table S4:

d(ATCGAGACGTCTCGAT)$_2$		open-d(ATCGAGA**CG**TCTCGAT)$_2$	
bp above (C·G)	bp below (G·C)	bp above (C·G)	bp below (G·C)

Shear	0.25	-0.25	0.17	-0.17
Stretch	-0.26	-0.26	-0.26	-0.26
Stagger	0.08	0.08	0.08	0.08
Buckle	3.75	-3.74	9.00	-9.00
propeller	-11.96	-11.95	0.00	0.00
Opening	-1.02	-1.02	-1.02	-1.02

The values for stretch, stagger and opening were kept as in the original PREDICTOR-generated structure without the intercalation gap. The values for shear were changed to 0.17 and -0.17 as these values are more in line both with the averages for all structures with intercalators and with all PREDICTOR-generated structures without intercalators. Buckle was changed to 9.00 and -9.00 as the average of the (absolute) average values of buckle above and below intercalation sites. Propeller twist was similarly reduced to 0.

Construction of the target oligonucleotide structure for docking studies

A random self-complementary hexadecameric sequence, *viz*. ATCGAGACGTCTCGAT, with a central C-G step was generated on the basis of a series of random numbers generated by a portable very-long-period random number generator.[7] The choice for a central C-G step was inspired by the observation that this is the preferred binding site for the majority of intercalators as evidenced by the fact that 70 out of the total of 103 base pair steps for which rise > 6.4 Å are C-G steps (*vide supra*) and the (related) observation that intercalation in a C-G step tends to lead to stronger binding than intercalation in other steps.[8] The length of the sequence was chosen based on preliminary docking studies which suggested that for a minor-groove binding mode an oligonucleotide length of at least six base pairs is required. Thus, a structure offering both an intercalation gap and a sufficiently long stretch of undisturbed DNA, allows for unbiased binding mode screening.

A duplex structure for d(ATCGAGACGTCTCGAT)$_2$ was generated using PREDICTOR.[9] The PREDICTOR-generated structure was analyzed using 3DNA,[10-12] yielding the base pair and base-pair-step parameters. The central base-pair-step parameters and the base pair parameters for the base pairs flanking the central C-G step were edited to correspond to typical values for intercalation sites (*vide supra*) resulting in the input file for 3DNA reproduced below.

16 base-pairs
0 ***local base-pair & step parameters***

	Shear	Stretch	Stagger	Buckle	Prop-Tw	Opening	Shift	Slide	Rise	Tilt	Roll	Twist
A-T	0.04	-0.18	0.11	0.08	-15.13	1.57	0.00	0.00	0.00	0.00	0.00	0.00
T-A	-0.06	-0.19	0.08	0.25	-14.72	1.63	-0.06	-0.66	3.20	-0.01	-2.89	36.34
C-G	0.20	-0.28	0.09	-3.06	-11.11	-1.34	-0.03	-0.25	3.19	-0.00	1.99	39.46
G-C	-0.20	-0.28	0.09	3.05	-11.10	-1.36	0.06	-0.08	2.94	0.16	8.32	31.91
A-T	0.10	-0.20	0.05	-0.14	-12.70	0.49	-0.16	-0.26	3.19	-0.38	2.29	39.54
G-C	-0.13	-0.27	0.10	4.32	-10.58	-0.24	0.26	-0.17	3.09	0.32	2.40	34.22
A-T	0.12	-0.21	0.05	0.34	-14.10	1.49	-0.29	-0.23	3.22	-0.63	2.17	39.16
C-G	0.17	-0.26	0.08	9.00	00.00	-1.02	0.38	-0.65	3.21	1.49	0.28	34.20
G-C	-0.17	-0.26	0.08	-9.00	00.00	-1.02	-0.06	-0.25	6.97	-0.15	-1.21	30.85
T-A	-0.12	-0.21	0.05	-0.34	-14.12	1.50	-0.25	-0.65	3.21	-1.21	0.36	34.29
C-G	0.13	-0.27	0.10	-4.31	-10.58	-0.24	0.29	-0.23	3.22	0.64	2.18	39.17
T-A	-0.10	-0.20	0.05	0.13	-12.70	0.50	-0.39	-0.17	3.09	-0.62	2.25	34.19
C-G	0.20	-0.28	0.09	-3.07	-11.12	-1.35	0.29	-0.27	3.19	0.77	2.28	39.44
G-C	-0.20	-0.28	0.09	3.05	-11.10	-1.35	-0.19	-0.08	2.94	-0.44	8.20	31.92
A-T	0.06	-0.19	0.08	-0.24	-14.72	1.63	0.16	-0.44	3.19	0.32	1.86	39.00
T-A	-0.04	-0.18	0.11	-0.09	-15.14	1.57	-0.07	-0.86	3.19	-0.39	-2.06	35.59

Using the edited parameters, a new structure for d(ATCGAGACGTCTCGAT)$_2$ was generated using 3DNA. In the new structure, phosphate ester linkages between the deoxyribose units at the intercalation site were optimized (AMBER ff99SB as incorporated in UCSF Chimera[13]). Similarly, phosphate units at all base pair steps were optimized. Optimized ester P-O bond lengths were all between 1.60 and 1.67 Å. The resulting PDB file for our target DNA structure, which we call open-d(ATCGAGA**CG**TCTCGAT)$_2$, was used to prepare the PDBQT file (*vide infra*).

Construction of the ligand structure for docking studies

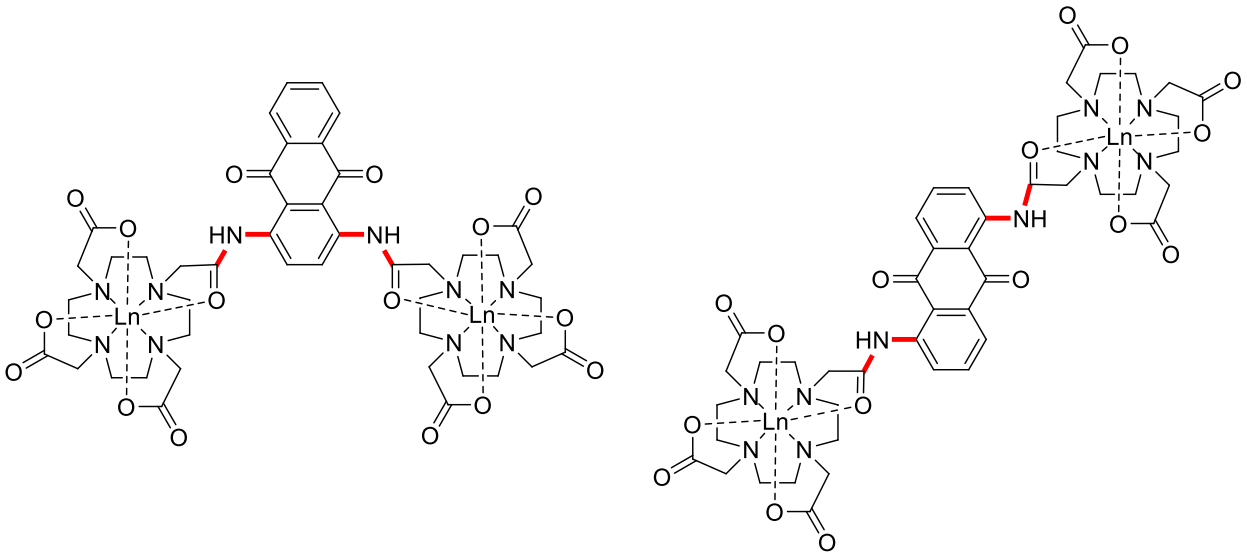
Structures for **M2-L1** and **M2-L2** were generated by combining the three structural fragments, *i.e.* the anthraquinone fragment with two copies of the macrocyclic-metal chelate complex. The required anthraquinone fragment was based on the DFT-optimized (B3PW91) structures for **3qm** and **4qm**. The macrocyclic-metal chelate complex was based on the crystal structure of the potassium complex of tris(1,1-dimethylethyl)10-{2—methoxy-6-nitrophenyl)amino]-2-oxoethyl}-1,4,7,10-tetraazacyclododecane-1,4,7-triacetate (Andrews *et al., Helv. Chim. Acta*, **2009**, *92*, 2159). The individual fragments to be combined were obtained by pruning both the anthraquinone and macrocyclic-metal chelate complex structures to obtain the required fragment, but each still with its own copy of the linking amide. The individual fragments were combined by precisely overlapping the amide functional groups (ACDLabs 12.0).

Following their generation, the PDBQT files were edited, replacing K for Fe(III) and adjusting the charges on the carboxylate-oxygens to be identical. The coordination environment of the metal centers was left unaltered. The choice for Fe(III) instead of Gd(III) was to take full advantage of existing parameterization in AutoDock Vina and maintain overall charge neutrality for the DNA binding complex.

Docking studies

Most docking programs incorporate parameterization of interactions as optimized for systems involving proteins and enzymes. Nevertheless, docking studies involving DNA have been carried out, see for example References [14-17]. Of particular interest here is the protocol developed by Ricci and Netz[18] involving docking using AutoDock 4.0 for situations where target DNA structures and binding modes are unknown. Ricci and Netz generated a DNA structure with a pre-created intercalation site. This structure was based on a canonical B DNA structure and optimized for intercalated ellipticine. For ellipticine, netropsin, aflatoxin B1 *exo*-8,9-epoxide and two Tröger bases, Netz and co-worker demonstrated that docking studies involving this optimized B DNA structure with intercalation gap can be used to predict whether a compound is a groove binder or an intercalator.

Here, AutoDock Vina[19] was used for docking studies. The required PDBQT (input) files for both ligand and DNA were generated using AutoDockTools 1.5.4.[20] The DNA structure was kept rigid and polar hydrogens were added. The grid box dimensions (40 × 40 × 60 Å) were determined by visual inspection so that the grid box encompassed the complete DNA structure, but also provided additional space to allow for maximum flexibility in ligand orientations. The ligand structure was kept rigid with the exception of the Ar-N and N-C bonds corresponding to the two amide links between the anthraquinone moiety and the metal-chelate complexes (indicated in red in Scheme S1).



Ln2-L1 **Ln2-L2**

Scheme S1

For docking, an *exhaustiveness* of 200 was chosen, together with *num_modes* of 10, with AutoDock Vina running on a dedicated computer.

Docking studies – results

The calculated affinities for the top three poses are summarised in Table S5.

Table S5: Top three ranked poses for Ln2-L1 and Ln2-L2 interacting with open-d(ATCGAGA**CG**TCTCGAT)$_2$.

	calculated affinity for pose / kcal mol^{-1}	
pose ranking	**Ln2-L1**	**Ln2-L2**
1	-13.4	-14.2
2	-13.4	-14.1
3	-13.1	-13.9

Table S5 clearly shows that the predicted binding affinity of **Ln2-L2** for open-d(ATCGAGA**CG**TCTCGAT)$_2$ than the predicted affinity of **Ln2-L1** for open-d(ATCGAGA**CG**TCTCGAT)$_2$. The corresponding binding modes are shown in Figure S2.

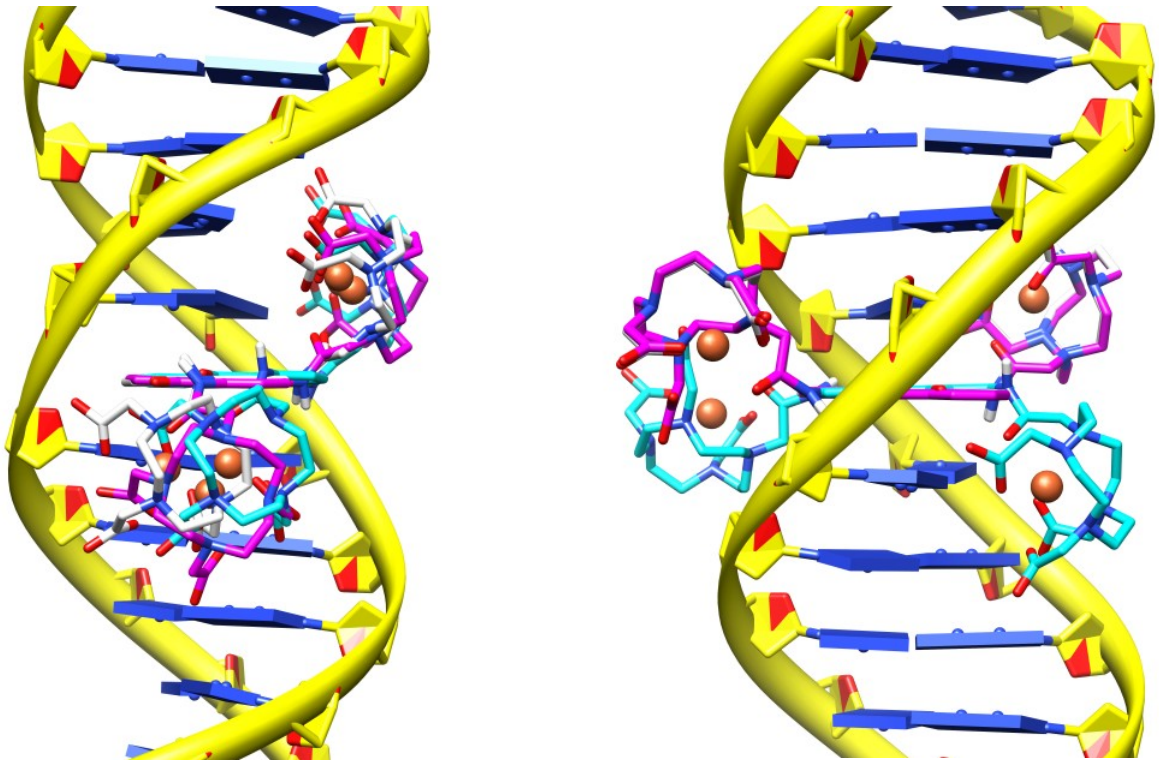


Figure S2: The top three poses for **Ln2-L1** interacting with open-d(ATCGAGA**CG**TCTCGAT)$_2$ (*Left*) and the top three poses for **Ln2-L2** interacting with open-d(ATCGAGA**CG**TCTCGAT)$_2$ (*Right*).

Figure S2 shows that the top poses for **Ln2-L2** interacting with open-d(ATCGAGA**CG**TCTCGAT)$_2$ all correspond to the spearing binding mode. The top three poses for **Ln2-L1** with open-d(ATCGAGA**CG**TCTCGAT)$_2$ corresponds to intercalation with both macrocyclic-metal chelate complexes located in the major groove.

Experimental

General physical measurements

Luminescence Spectroscopy

All photophysical data were obtained on a JobinYvon-Horiba Fluorolog spectrometer fitted with a JY TBX picosecond photodetection module and a Hamamatsu R5509-73 detector (cooled to –80 °C using a C9940 housing) for near-IR measurements. For the near-IR lanthanide ion lifetimes the pulsed laser source was a Continuum Minilite Nd:YAG configured for 355 nm output. For the ligand-based fluorescence lifetimes 372 and 459 nm NanoLEDs (operating at 1 MHz) were utilised. All lifetime data were obtained using the JY-Horiba FluoroHub single photon counting module and the lifetimes obtained from data fits using the provided DAS9 software. IR spectra were recorded on an ATR equipped Varian 7000 FT-IR. LR mass spectra were obtained using a Bruker MicroTOF LC. UV-vis data were recorded as solutions on a Jasco 570 spectrophotometer.

DNA Binding Studies

For DNA binding the pH of aqueous solutions was determined with a Hanna Instruments pH 210 pH meter equipped with a VWR 662-1759 glass electrode. UV-visible spectra were recorded using a Jasco V-630Bio or Jasco V-650 UV-visible spectrophotometer coupled with a Peltier temperature controller at 25 °C. Circular dichroism spectra were recorded on a Chirascan CD Spectrophotometer. Water was purified using an ELGA option-R 7BP water purifier. All experiments were carried out in buffer (50 mM NaCl, 25 mM MOPS, titrated with aqueous NaOH to pH 7.0). Calf thymus DNA was obtained from Sigma-Aldrich as a lyophilised solid sodium salt. Calf thymus DNA solutions were prepared by suspending in CT buffer followed by sonication for a total of 30 min while maintaining the samples in ice. Resulting DNA solutions were dialysed extensively against buffer using a 3.5 kDa MW cut-off dialysis membrane. DNA concentrations were determined spectrophotometrically using ε(260nm) = 12824 $M(bp)^{-1}cm^{-1}$.[26] Purity of DNA solutions was checked by UV-visible spectroscopy from the ratio A260 nm / A280 nm, which was found to be 1.92, suggesting that solutions are essentially free of protein. Solutions of **Gd2-L1** and **Gd2-L2** in buffer were prepared so that the absorbance at lmax (*A*lmax) was approximately 0.8. Quartz cuvettes (1.00 cm pathlength, HELLMA Suprasil) were filled with 2500 ml of the solutions of **Gd2-L1** and

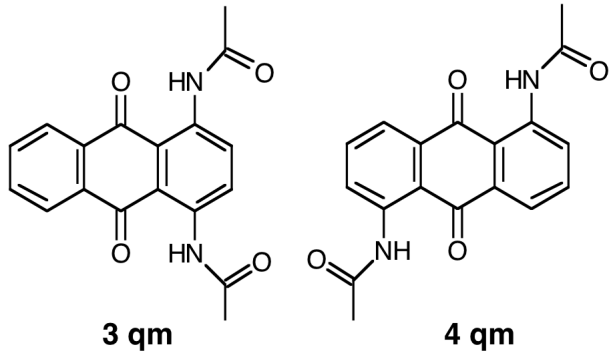
Gd2-L2. Concentrations of **Gd2-L1** and **Gd2-L2** were determined using ε (440 nm) = 1770 $dm^3mol^{-1}cm^{-1}$ and ε (405 nm) = 3200 $dm^3mol^{-1}cm^{-1}$, respectively.

UV-vis titrations were carried out by adding aliquots of a concentrated stock solution of calf thymus DNA in buffer to the ligand solution in the cuvette (with a total of 615 ml of a 8.9 mM solution of CT DNA solution added to **Gd2-L2** and a total of 460 ml of a 6.1 mM CT DNA added to **Gd2-L1**). UV-visible spectra in the range 200-900 nm were recorded after each addition of DNA. Absorptions at selected wavelengths as a function of DNA and DNA-binder concentrations were extracted from the UV-visible spectra and analysed globally in terms of a multiple independent binding sites model, correcting for ligand dilution, using Origin 7.5.

Density Functional Theory

Density functional theory calculations were performed on the Gaussian 03 program.[21] Geometry optimizations were carried out on **3-qm** and **4-qm** (see Scheme S2) without geometry constraints, using the B3PW91 functional and a 6-31G(d,p) basis set for all atoms. All optimizations were followed by frequency calculations to ascertain the nature of the stationary point (minimum or saddle point). Molecular images were created using the Molekel program.[22]

Scheme S2: The 1,4- (**3-qm**) and 1,5-amido-substituted AQ (**4-qm**) molecular models investigated using DFT methods.



Relaxometry

The $1/T_1$ NMRD profiles were obtained on a Stelar Spinmaster FFC-2000 relaxometer, typically covering magnetic fields from 2.35 x 10^{-4} to 0.71 T (corresponding to a proton Larmor frequency range 0.01 - 30 MHz). A Spinmaster Variable Temperature Controller allowed the setting and monitoring of the sample temperature ranging in these experiments between 20 – 37 °C with a resolution up to 0.1 °C. For field

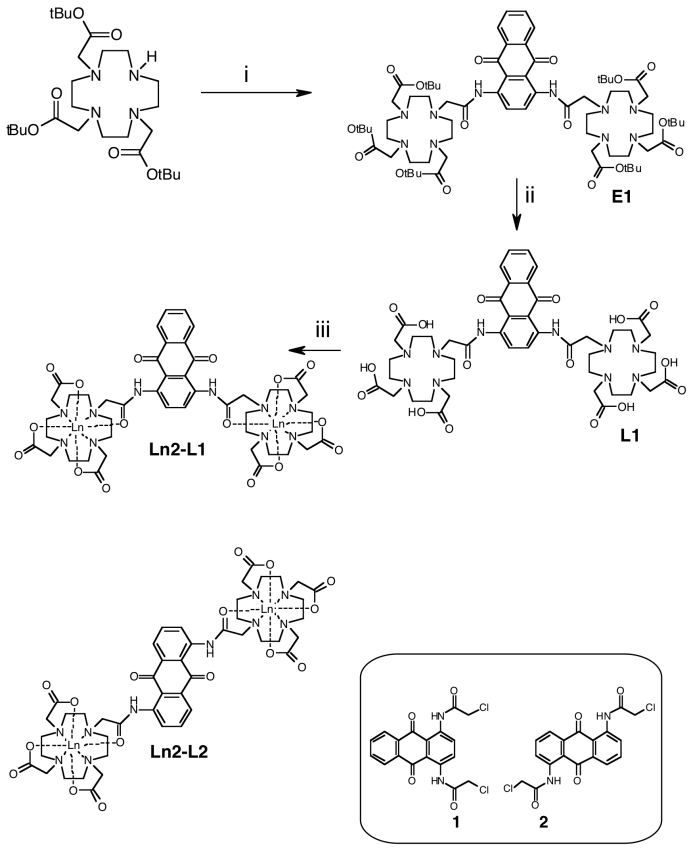
strengths less than 10 MHz a pre-polarising field sequence was used with the sample polarised at 30 MHz before being allowed to relax at the field of interest. All relaxation profiles were treated as a mono-exponential decay and fitted with an error of less than 1%. For the DNA binding experiments complex relaxivity (r_1) at 30 MHz was calculated by correcting for the contribution of the water/DNA sample at 30 MHz (0.39 and 0.28 at 20 and 37 °C respectively). The relaxation profiles for the complexes have been reproduced assuming one water molecule (in accordance with the lifetime data) regularly coordinated to Gd^{III}, *i.e.* two protons at 3.1 Å. Because of the low relaxation rates in the absence of DNA, the distance of closest approach of diffusing water protons was fixed to 4.5 Å (instead of the most widely used 3.6 Å). The diffusion coefficients were fixed to 2.3 and 3.5×10^{-5} cm^2/s at 20 and 37°C, respectively.

Materials and procedures

1,4-diaminoanthraquinone (90 %, technical grade) and 1,5-diaminoanthraquinone (97 %) were purchased from Alfa Aesar and used as received. Although all new reported compounds are air stable both in the solid state and solution, the free ligands and complexes are hygroscopic if left to the open atmosphere for prolonged periods.

Syntheses of precursors, ligands and corresponding lanthanide complexes

Scheme S3. Synthesis of the ligands and corresponding dimetallic complexes: i) bis-chloroacetamide (**1** or **2**) 1eq. KI, $^{i}Pr_2EtN$, PhMe, Δ; ii) TFA, DCM; iii) 2 eq. $Ln(OTf)_3$, MeOH.



Synthesis of 1: Excess chloroacetylchloride (0.5 ml, 6.4 mmol) in MeCN (5ml) was added dropwise to a solution of 1,4-diaminoanthraquinone (0.500 g, 2.1 mmol) and 1.1 eq. NEt_3 (0.34 mL, 2.42 mmol) in MeCN (10 mL) at 0 ^{o}C. The reaction was allowed to warm to room temperature and after 48 hours the solvent was removed *in vacuo*. CH_2Cl_2 (50 mL) was then added and the solution extracted with water (2 x

25 mL). The organic phase was collected, dried over $MgSO_4$, filtered and solvent removed *in vacuo*, yielding a red-brown powder (0.119 g, 14 % based on technical grade 1,4-diaminoanthraquinone). ^{1}H NMR (400 MHz, $CDCl_3$): δ_H = 9.10 (2H, s, Ar*H*), 8.35 (2H, dd, J_{HH} = 3.3 and 2.5 Hz, Ar*H*), 7.75 (2H, dd, J_{HH} = 3.3 and 2.5 Hz, Ar*H*), 4.20 (s, 4H, CH_2) ppm. Insufficient solubility prevented spectroscopic analysis by ^{13}C NMR. IR ν(C=O) (solid): 1681, 1583 cm^{-1}. EI MS found *m/z* = 390 $\{M\}^{\cdot+}$ (assuming ^{35}Cl). HRMS (EI) found *m/z* 390.0168 $\{M\}^{\cdot+}$; $C_{18}H_{12}N_2O_4Cl_2$ requires 390.0169. Melting Pt. 291-295°C.

Synthesis of 2: Excess chloroacetylchloride (0.25 mL, 3.16 mmol) in MeCN (5 mL) was added dropwise to a solution of 1,5-diaminoanthraquinone (0.250 g, 1.05 mmol) and 2 eq. NEt_3 (0.3 mL, 2.15 mmol) in MeCN (10 mL) at 0 °C. The reaction and work-up was carried out in an identical manner above for **1**. The product was obtained as an orange powder (0.402 g, 98 %). Insufficient solubility prevented spectroscopic analysis by ^{1}H NMR. IR ν(C=O) (solid): 1683, 1579 cm^{-1}. ES$^+$ MS found *m/z* = 390 $\{M\}^{\cdot+}$ (assuming ^{35}Cl). HRMS (EI) found *m/z* 390.0169 $\{M\}^{\cdot+}$; $C_{18}H_{12}N_2O_4Cl_2$ requires 390.0169. Melting Pt. >300°C.

Synthesis of E1: Compound **1** (0.050 g, 0.13 mmol), 1,4,-tris(*tert*-butoxycarbonylmethyl)-1,4,7,10-tetraazacyclododecane (0.155 g, 0.26 mmol), KI (0.05 g, 0.3 mmol) and EtN^iPr_2 (4 mL) were added to toluene (10 mL) under inert conditions. The reaction was heated at 70 °C for 5 hours, the reaction monitored by thin layer chromatography (CH_2Cl_2:MeOH, 9:1) until completion. The reaction was allowed to cool before removal of the solvent *in vacuo*. Ethyl acetate was added and washed with water (2 x 20 mL), dried over magnesium sulphate, filtered and solvent removed *in vacuo* yielding a brown oil (0.105 g, 61 %). ^{1}H NMR (400MHz, $CDCl_3$): δ_H = 9.15 (2H, s, Ar*H*), 8.2 (2H, dd, J_{HH} = 3.4 and 2.5 Hz, Ar*H*), 7.75 (2H, dd, J_{HH} = 3.4 and 2.5 Hz, Ar*H*), 3.4-2.5 (48H, br m, -NCH_2CH_2N, -NCH_2CO and -CH_2), 1.3 (54H, m, CCH_3) ppm. ^{13}C-{^{1}H} NMR (101 MHz, $CDCl_3$): δ_C = 171.9, 170.8, 169.6, 168.7, 136.8, 134.2, 133.8, 128.5, 127.6, 126.2, 81.1, 80.6, 57.8, 55.9, 55.0, 54.6, 51.9, 51.1, 50.3, 50.1, 48.1, 46.5, 27.2, 26.9 ppm. IR ν(C=O) (solid): 1734, 1700, 1653, 1576 cm^{-1}. ES$^+$ MS found *m/z* = 1369 $\{M+Na\}^+$.

Synthesis of E2: Compound **2** (0.050 g, 15.6 mmol), 1,4,-tris(tert-butoxycarbonylmethyl)-1,4,7,10-tetraazacyclododecane (0.155 g, 31.2 mmol), KI (excess) and EtN^iPr_2 (4 mL) were added to toluene (10 mL) under inert conditions as for compound **E2** above. Following work-up an orange solid was isolated (0.138 g, 80%). ^{1}H NMR (400MHz, $CDCl_3$): δ_H = 9.15 (2H, d, ^3J = 8.4 Hz, Ar*H*), 8.0 (2H, d, J_{HH} = 8.3 Hz, Ar*H*), 7.2 (2H, app. t, Ar*H*), 3.4-2.8 (48H, bm, -NCH_2CH_2N, -NCH_2CO and -CH_2), 1.4 (54H, m, CCH_3) ppm. ^{13}C-{^{1}H}

NMR (101 MHz, $CDCl_3$): δ_C = 172.2, 169.8, 168.8, 168.4, 140.0, 134.9, 133.8, 125.2, 121.8, 116.8, 76.2, 75.7, 57.2, 55.6, 54.8, 51.6, 51.1, 50.2, 48.5, 46.4, 44.7, 39.1, 27.3, 26.8 ppm. IR ν(C=O) (solid): 1734, 1700, 1653, 1576 cm^{-1}. ES^+ MS found *m/z* = 1369 $\{M+Na\}^+$.

Synthesis of L1: **E1** (0.096 g, 0.07 mmol) was dissolved in CH_2Cl_2 (3 mL) and trifluoroacetic acid (3 mL) was slowly added dropwise. The reaction was stirred at room temperature for 48 hr. The solvents were removed *in vacuo*, and the residue repeatedly re-dissolved in methanol and dried under reduced pressure (procedure repeated four times). The resultant residue was dissolved in a minimum volume of ethanol and added dropwise to stirring diethyl ether at 0°C, producing a red precipitate (0.063 g, 84%). ^{1}H NMR (400MHz, D_2O): δ_H = 8.6 (2H, s, Ar*H*), 7.95 (2H, br s, Ar*H*), 7.75 (2H, br s, Ar*H*), 3.95-2.75 (48H, br overlapping m, -NC*H*$_2$C*H*$_2$N, -NC*H*$_2$CO and -C*H*$_2$) ppm. IR ν(C=O) (solid): 1684, 1653, 1576 cm^{-1}. HRMS (ES^+) found *m/z* 1011.4417 $\{M + H\}^+$; $C_{46}H_{63}N_{10}O_{16}$ requires 1011.4418. Calculated: C 43.15; H 4.69 %; N 9.32 % for $C_{46}H_{62}N_{10}O_{16}\bullet 4C_2HO_2F_3\bullet 2H_2O$. Found: C 43.35 %; H 5.12 %; N 9.72 %. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 302 (6100), 454 (3100) nm.

Synthesis of L2: The same procedure was followed for the deprotection of **E2** (0.096 g, 0.07 mmol), albeit using a 72 hr reaction time (as determined by ^{1}H NMR analysis of the reaction mixture), yielding a yellow precipitate as a first crop (0.058 g, 55 %). Further crops were obtained by storing the ethanol/diethyl ether solution in a fridge freezer. ^{1}H NMR (400MHz, D_2O): δ_H 8.45 (2H, bd, $^3J_{HH}$ = 8.1 Hz, Ar*H*), 7.6 (2H, m, Ar*H*), 7.5 (2H, m, Ar*H*), 3.8-3.0 (48H, bm, -NC*H*$_2$C*H*$_2$N, -NC*H*$_2$CO and -C*H*$_2$) ppm. IR ν(C=O) (solid): 1684, 1653, 1578 cm^{-1}. HRMS (ES^+) found *m/z* 1011.4409 $\{M + H\}^+$; $C_{46}H_{63}N_{10}O_{16}$ requires 1011.4418. Calculated: C 44.21 %; H 4.53 %; 9.55 N % for $C_{46}H_{62}N_{10}O_{16}\bullet 4C_2HO_2F_3$. Found: C 44.90 %; H 5.21 %; N 10.23 %. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 268 (37200), 426 (11300) nm.

General Procedure for the formation of lanthanide complexes

A mixture of the ligand and $Ln(OTf)_3$ was stirred in MeOH at 50 °C for 24 h. Chromatographic methods were used to assess that no free ligand remained. The solution was then concentrated *in vacuo* and added dropwise to stirring diethyl ether at 0 °C. The precipitate was filtered, washed with diethyl ether and dried *in vacuo*. This procedure was then repeated twice to give the desired lanthanide complex. The hygroscopic nature of the complexes prevented reproducible elemental analyses from being obtained. However, the homogeneity of each isolated complex was assessed for absence of free metal (xylenol orange) and assessed chromatographically for absence of free chromophoric ligand.

Nd2-L1 was isolated as an orange powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 262 (29100), 302 (9000), 433 (3600) nm. IR ν(C=O) (solid): 1653, 1577, 1559 cm^{-1}. MS (MALDI): *m/z* 1293 $\{M + H\}^+$.

Gd2-L1 was isolated as an orange powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 261 (13100), 312 (3300), 439 (1500) nm. IR ν(C=O) (solid): 1653, 1583, 1559 cm^{-1}. MS (MALDI): *m/z* 1321 $\{M + H\}^+$.

Er2-L1 was isolated as an orange powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 261 (11400), 306 (3100), 436 (1300) nm. IR ν(C=O) (solid): 1653, 1592 br cm^{-1}. MS (MALDI): *m/z* 1339 $\{M + H\}^+$.

Yb2-L1 was isolated as an orange powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 261 (29500), 306 (6700), 440 (2900) nm. IR ν(C=O) (solid): 1653, 1589, 1559 cm^{-1}. MS (MALDI): *m/z* 1351 $\{M + H\}^+$.

Nd2-L2 was isolated as a red powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 226 (19900), 269 (23200), 407 (4500) nm. IR ν(C=O) (solid): 1636, 1594, 1578 cm^{-1}. MS (MALDI): *m/z* 1293 $\{M + H\}^+$.

Gd2-L2 was isolated as a red powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 228 (21000), 269 (31300), 407 (4800) nm. IR ν(C=O) (solid): 1636, 1577 br cm^{-1}. MS (MALDI): *m/z* 1321 $\{M + H\}^+$.

Er2-L2 was isolated as a red powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 227 (17500), 270 (19200), 407 (4000) nm. IR ν(C=O) (solid): 1636, 1597, 1578 cm^{-1}. MS (MALDI): *m/z* 1339 $\{M + H\}^+$.

Yb2-L2 was isolated as a red powder. UV-vis. λ_{max} (ε / mol^{-1} dm^3 cm^{-1}): 226 (21400), 269 (25800), 403 (4900) nm. IR ν(C=O) (solid): 1636, 1602, 1581 cm^{-1}. MS (MALDI): *m/z* 1352 $\{M + H\}^+$.

Additional Table

Table 1. Visible absorbance maxima and near-IR luminescence lifetimes for the complexes.

Complex	**λ_{abs} (nm)**	**λ_{det} (nm)**	**τ (H_2O) / ns**	**τ (D_2O) / ns**	***q***
Nd2-L1	433	1055	76	245	0.8
Nd2-L2	407	1055	77	242	0.8
Er2-L1	436	1535	-	850	-
Er2-L2	407	1535	-	1044	-
Yb2-L1	440	980	610	6407	1.3
Yb2-L2	403	980	756	7901	1

Errors on lifetimes are ±10 %; errors on *q* calculations are ±20%.

Additional Figures

Figure S3. Pictorial representations of the calculated HOMO (bottom) and LUMO (top) of the model chromophores **3-qm** (left) and **4-qm** (right).

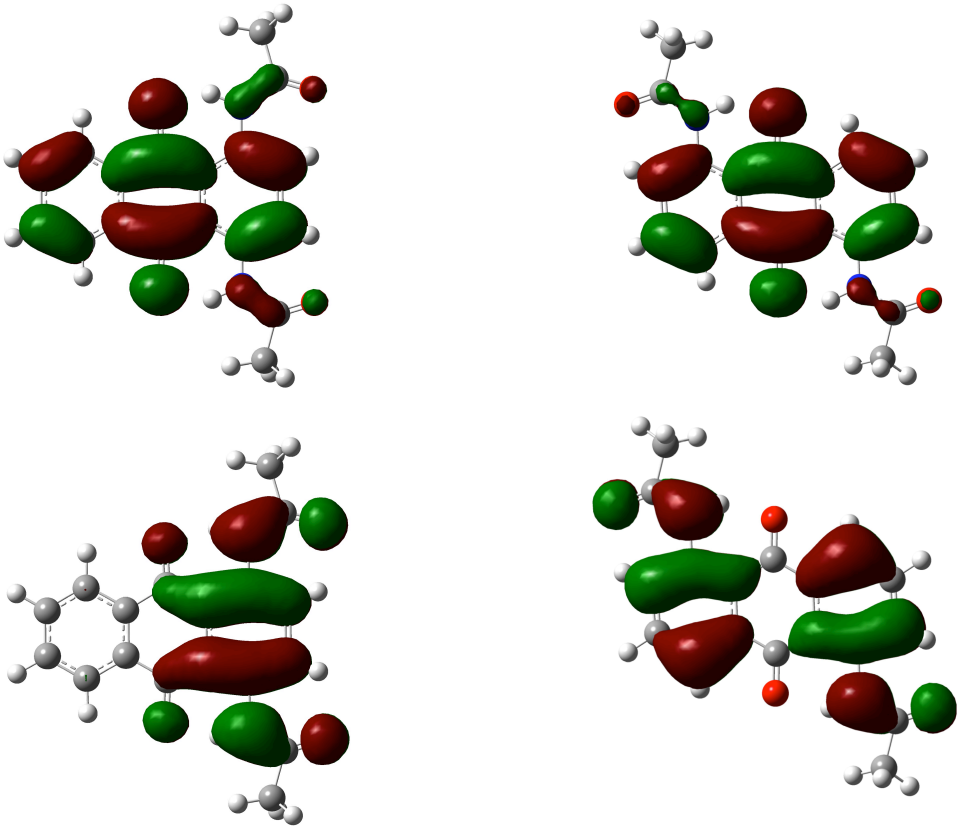


Figure S4. Steady state emission spectra for **Ln2-L2** (red-Yb; blue-Nd; green-Er) measured in D_2O using λ_{ex} = 400 nm. Spectra offset for clarity. Inset: 77 K emission spectra for **Gd2-L1** (black) and **Gd2-L2** (grey) measured as a methanol: ethanol (1:9) glass.

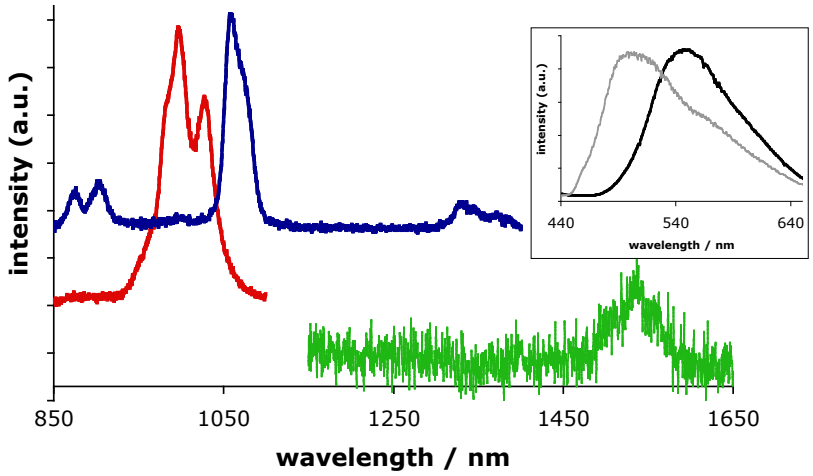
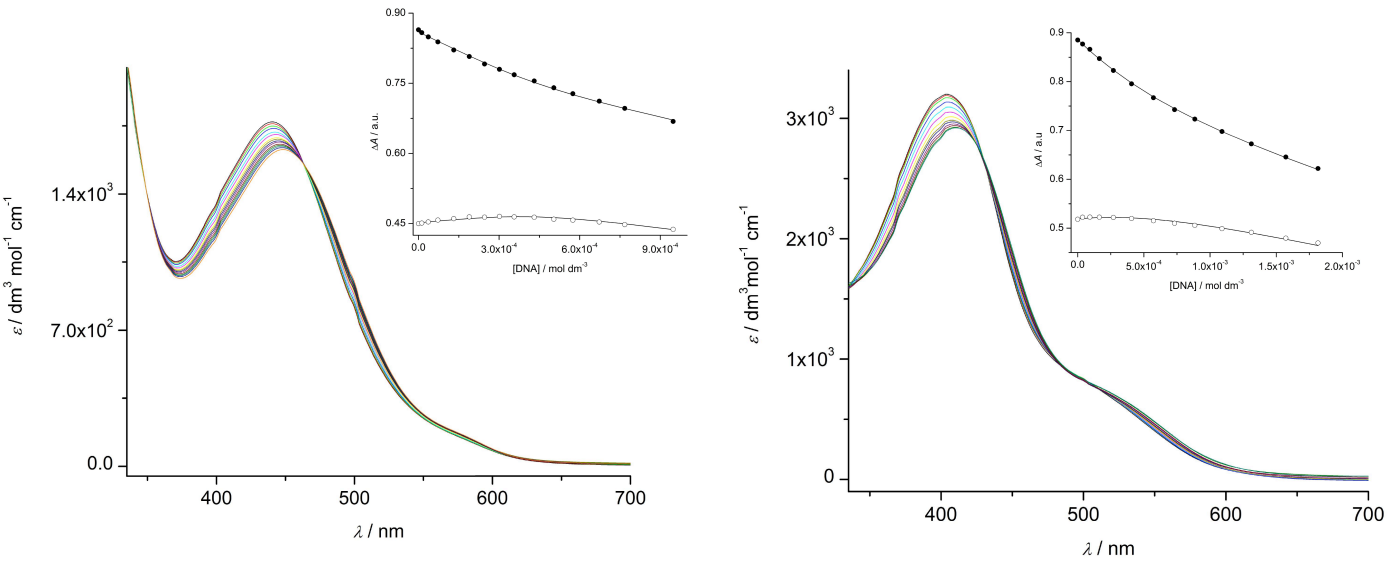


Figure S5. UV-vis extinction coefficients for **Gd2-L1** upon the addition of 0 – 0.95 mM calf thymus DNA (*left*) and for **Gd2-L2** upon the addition of 0 – 1.82 mM calf thymus DNA (*right*). Inset: (left) global fit (line) of the absorbance at 440 nm (•) and 495 nm (o) of a **Gd2-L1** as a function of added calf thymus DNA; (right) global fit (line) of the absorbance at 385 nm (•) and 450 nm (o) of a **Gd2-L2** as a function of added calf thymus DNA. All measurements in 25 mM MOPS, 50 mM NaCl, pH 7.0 at 25 °C.



References

1. Zheng, G. H.; Colasanti, A. V.; Lu, X. J.; Olson, W. K., *Nucleic Acids Res.* **38**, D267-D274.

2. Berman, H. M.; Olson, W. K.; Beveridge, D. L.; Westbrook, J.; Gelbin, A.; Demeny, T.; Hsieh, S. H.; Srinivasan, A. R.; Schneider, B., *Biophys. J.* 1992**, 63** (3), 751-759.

3. The terms above and below are defined as in their sequence in the 3DNA output

4. Full data tables are available from the authors.

5. Hampshire, A. J.; Fox, K. R., *Anal. Biochem.* 2008**, 374** (2), 298-303.

6. Hampshire, A. J.; Rusling, D. A.; Broughton-Head, V. J.; Fox, K. R., *Methods* 2007**, 42** (2), 128-140.

7. Marsaglia, G., Zaman, A., *Computers in Physics* 1994**, 8** (1), 117.

8. Chaires, J. B., *Methods Enzymol.* 2001**, 340**, 3-22.

9. Farwer, J.; Packer, M. J.; Hunter, C. A., *In Silico Biol.* 2007**, 7** (6), 595-600.

10. Lu, X. J.; Olson, W. K., *Nat. Protoc.* 2008**, 3** (7), 1213-1227.

11. Lu, X. J.; Olson, W. K., *Nucleic Acids Res.* 2003**, 31** (17), 5108-5121.

12. Zheng, G.; Lu, X.; Olson, W. K., *Nucleic Acids Res.* 2009**, 37** (SUPPL. 2).

13. Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E., *J. Comput. Chem.* 2004**, 25** (13), 1605-1612.

14. Evans, D. A.; Neidle, S., *J. Med. Chem.* 2006**, 49** (14), 4232-4238.

15. Holt, P. A.; Chaires, J. B.; Trent, J. O., *J. Chem. Inf.* 2008**, 48** (8), 1602-1615.

16. Holt, P. A.; Ragazzon, P.; Strekowski, L.; Chaires, J. B.; Trent, J. O., *Nucleic Acids Res.* 2009**, 37** (4), 1280-1287.

17. Anthony, N. G.; Huchet, G.; Johnston, B. F.; Parkinson, J. A.; Suckling, C. J.; Waigh, R. D.; Mackay, S. P., *J. Chem. Inf.* 2005**, 45** (6), 1896-1907.

18. Ricci, C. G.; Netz, P. A., *J. Chem. Inf.* 2009**, 49** (8), 1925-1935.

19. Trott, O.; Olson, A. J., *J. Comput. Chem.* **31** (2), 455-461.

20. Sanner, M. F., *J. Mol. Graphics Modell.* 1999**, 17** (1), 57-61.

21. Gaussian 03, Revision E.01, Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, Jr., J. A.; Vreven, T.; Kudin, K. N.; Burant, J.C.; Millam, J.M.; Iyengar, S.S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G.A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J.E.; Hratchian, H.P.; Cross, J.B.; Bakken, V.; Adamo, C.; Jaramillo, J.;

Gomperts, R.; Stratmann, R.E.; Yazyev, O.; Austin, A.J.; Cammi, R.; Pomelli, C.; Ochterski, J.W.; Ayala, P.Y.; Morokuma, K.; Voth, G.A.; Salvador, P.; Dannenberg, J.J.; Zakrzewski, V.G.; Dapprich, S.; Daniels, A.D.; Strain, M.C.; Farkas, O.; Malick, D.K.; Rabuck, A.D.; Raghavachari, K.; Foresman, J.B.; Ortiz, J.V.; Cui, Q.; Baboul, A.G.; Clifford, S.; Cioslowski, J.; Stefanov, B.B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R.L.; Fox, D.J.; Keith, T.; Al-Laham, M.A.; Peng, C.Y.; Nanayakkara, A.; Challacombe, M.; Gill, P.M.W.; Johnson, B.; Chen, W.; Wong, M.W.; Gonzalez, C.; Pople, J.A.; Gaussian, Inc., Wallingford CT, 2004.

22. Varetto, U. MOLEKEL 5.4; Swiss national Supercomputing Center: Manno (Switzerland).