

**Electronic supporting information for:**  
**Exploring ultraviolet photoinduced charge-transfer dynamics in**  
**a model dinucleotide of guanine and thymine**

Marta Duchi,<sup>1</sup> Michael P. O'Hagan,<sup>1</sup> Rhea Kumar,<sup>1‡</sup> Simon J. Bennie,<sup>1</sup>  
M. Carmen Galan,<sup>1</sup> Basile F. E. Curchod,<sup>2</sup> and Thomas A. A. Oliver<sup>1\*</sup>

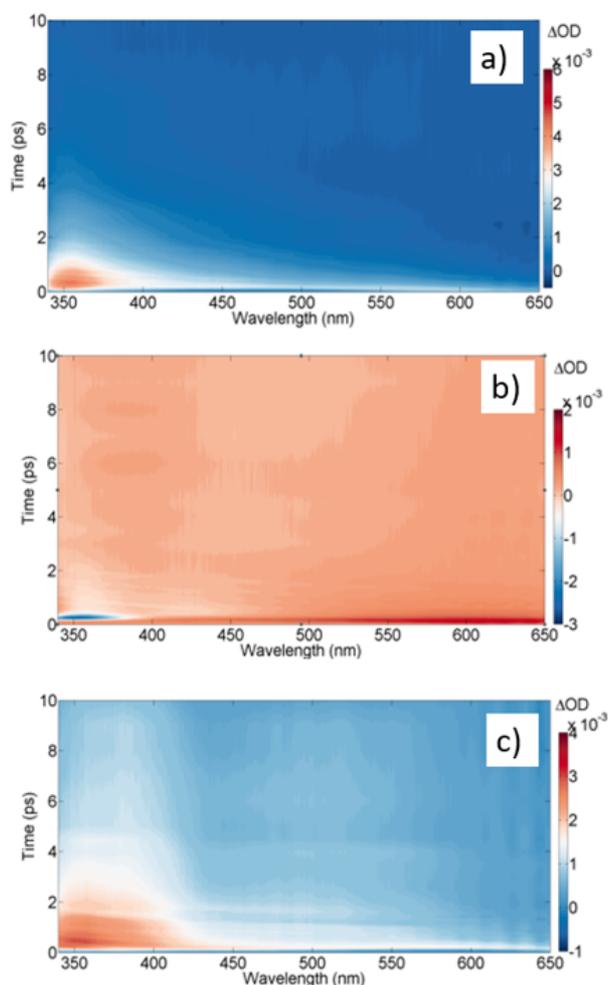
<sup>1</sup>School of Chemistry, Cantock's University of Bristol, Bristol, BS8 1TS, UK.

<sup>2</sup>Department of Chemistry, Durham University, South Road, Durham, DH1 3LE, UK.

\*Author for correspondence: [tom.oliver@bristol.ac.uk](mailto:tom.oliver@bristol.ac.uk)

‡ Current address: Department of Chemistry, Imperial College London, London, SW7 2AZ,  
UK.

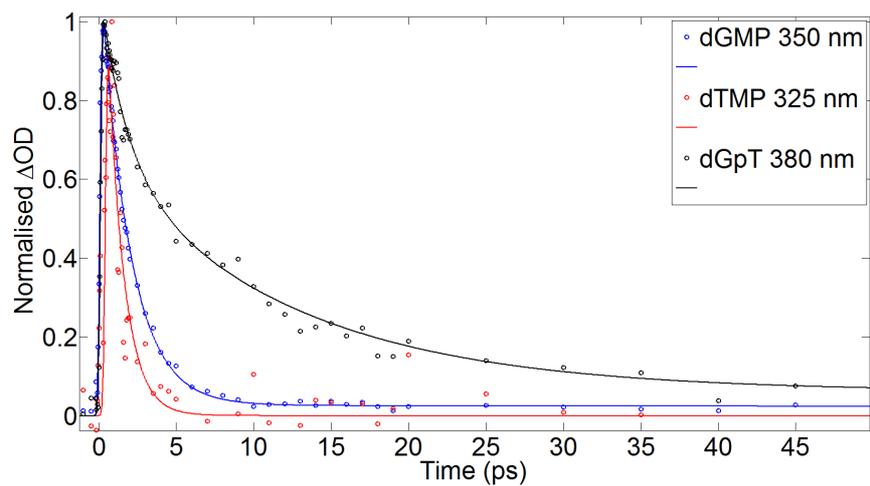
## Transient Absorption Spectroscopy



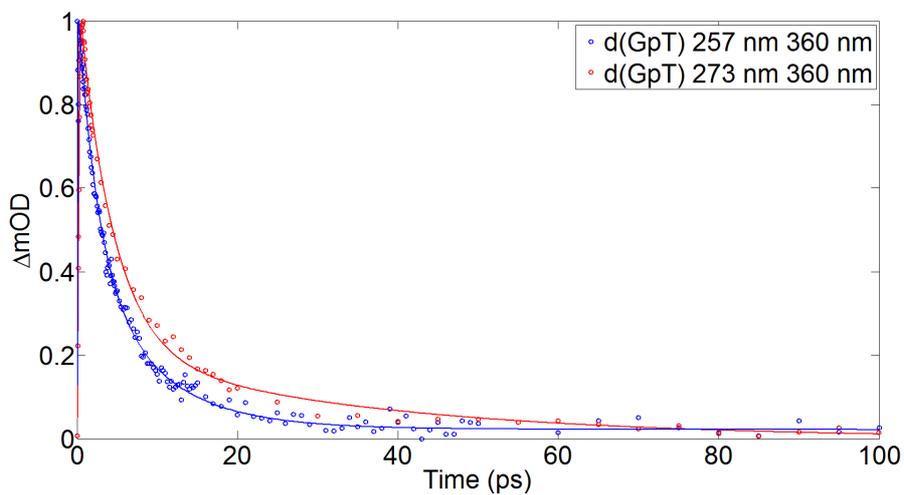
**Fig. S1** Transient absorption spectra for aqueous solutions of (a) dGMP, (b) dTMP and (c) d(GpT) using 257 nm photoexcitation.

**Table S1** Time constants obtained from fits to specific probe wavelengths of d(GpT), dGMP and dTMP samples.

Sample	Pump / nm	Probe / nm	$\tau_1$ / ps	$\tau_2$ / ps	$\tau_3$ / ns	$R^2$
d(GpT)	257	380	$3 \pm 2$	$9 \pm 7$	> 1 ns	0.99
	273	360	$2.1 \pm 0.7$	$11 \pm 2$	> 1 ns	0.99
dGMP	257	350	$1.78 \pm 0.07$	> 1 ns	-	0.98
	273	360	$2.1 \pm 0.4$	> 1 ns	-	0.99
dTMP	257	550	$0.46 \pm 0.04$	> 1 ns	-	0.97
	273	300	$1.4 \pm 0.3$	> 1 ns	-	0.94

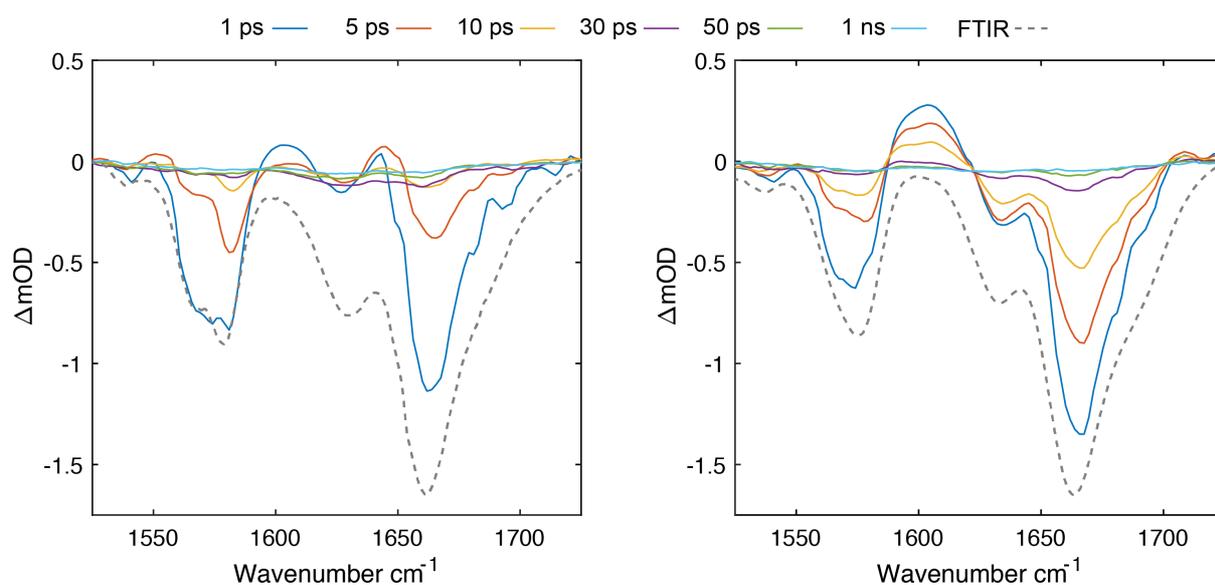


**Fig. S2** Comparison of normalised kinetics for the displayed wavelengths of d(GpT), dTMP and dGMP TA spectra under 257 nm irradiation.

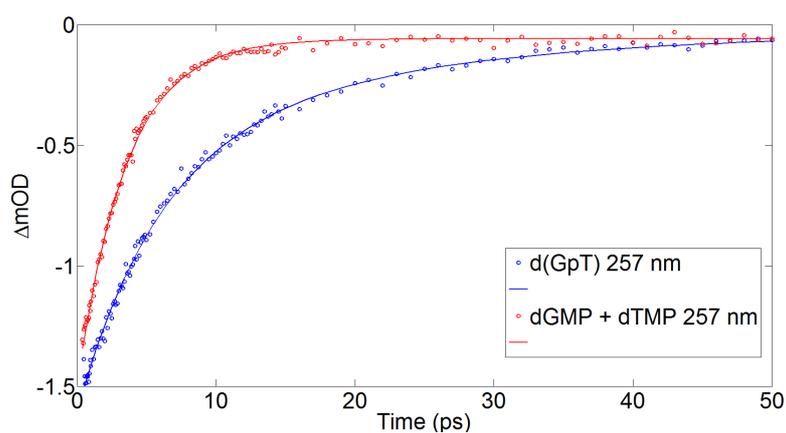


**Fig. S3** Comparison of normalised kinetics for d(GpT) samples photoexcited at 273 nm and 257 nm.

## Time Resolved Infrared Spectroscopy



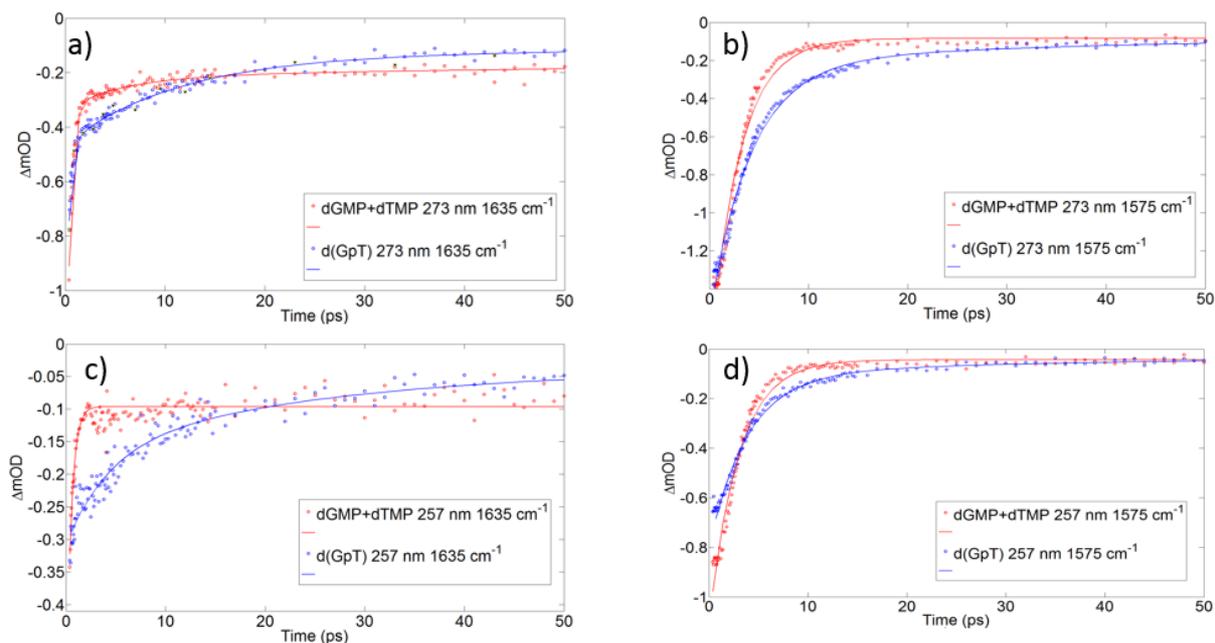
**Fig. S4** (a) TRIR spectra at stated values of  $t$ , obtained with 257 nm excitation of (a) dGMP + dTMP and (b) d(GpT) solutions.



**Fig. S5** Comparison of d(GpT) and dTMP + dGMP kinetics for  $\bar{\nu} = 1666 \text{ cm}^{-1}$  using 257 nm excitation. All fit parameters are given in Table S2.

**Table S2** Amplitudes and time constants returned from fitting TRIR data of d(GpT) and dGMP + dTMP solutions for 257 nm pump,  $1666 \text{ cm}^{-1}$  probe data.

d(GpT)		dGMP + dTMP	
$A_1 = 0.45$	$\tau_1 = 5.8 \pm 0.4 \text{ ps}$	$A_1 = 0.97$	$\tau_1 = 3.47 \pm 0.03 \text{ ps}$
$A_2 = 0.51$	$\tau_2 = 19 \pm 2 \text{ ps}$	$A_2 = 0.03$	$\tau_2 > 1 \text{ ns}$
$A_3 = 0.04$	$\tau_3 > 1 \text{ ns}$		$R^2 = 0.99$
	$R^2 = 0.99$		



**Fig. S6** TRIR kinetics for stated UV excitation and probe frequencies. Fit parameters are given in Table S3.

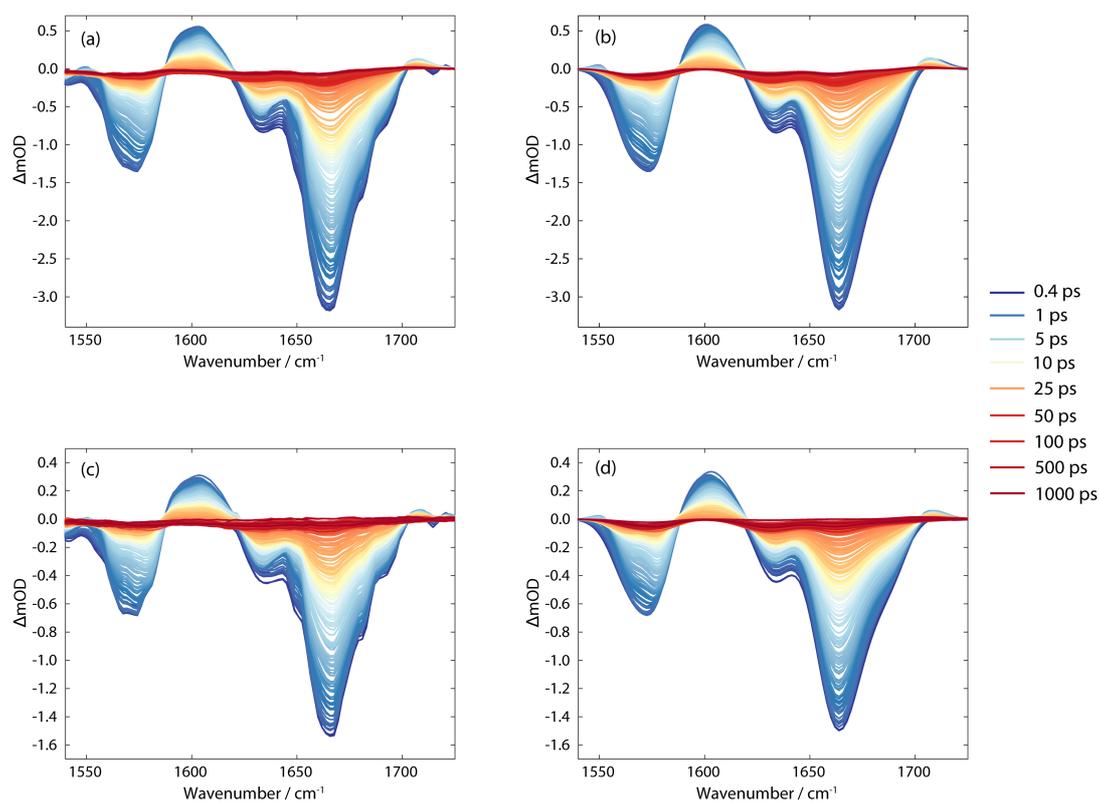
**Table S3** Kinetic fit parameters for 273 and 257 nm TRIR data of d(GpT) and a mixture of mononucleotides.

Probe frequency / $\text{cm}^{-1}$	d(GpT) (273 nm)	d(GpT) (257 nm)	Mixture (273 nm)	Mixture (257 nm)
1575 $\text{cm}^{-1}$ (G)	$\tau_1 = 4.1 \pm 0.8$ ps $\tau_2 = 33 \pm 17$ ps $\tau_3 > 1$ ns $R^2 = 0.99$	$\tau_1 = 3.1 \pm 0.8$ ps $\tau_2 = 27 \pm 7$ ps $\tau_3 > 1$ ns $R^2 = 0.97$	$\tau_1 = 2.9 \pm 0.1$ ps $\tau_2 > 1$ ns $R^2 = 0.98$	$\tau_1 = 2.91 \pm 0.07$ ps $\tau_2 > 1$ ns $R^2 = 0.98$
1635 $\text{cm}^{-1}$ (T)	$\tau_1 < 1$ ps $\tau_2 = 11.4 \pm 0.7$ ps $\tau_3 > 1$ ns $R^2 = 0.98$	$\tau_1 = 4.2 \pm 0.8$ ps $\tau_2 = 26 \pm 5$ ps $\tau_3 > 1$ ns $R^2 = 0.95$	$\tau_1 < 1$ ps $\tau_2 > 1$ ns $R^2 = 0.98$	$\tau_1 < 1$ ps $\tau_2 > 1$ ns $R^2 = 0.95$

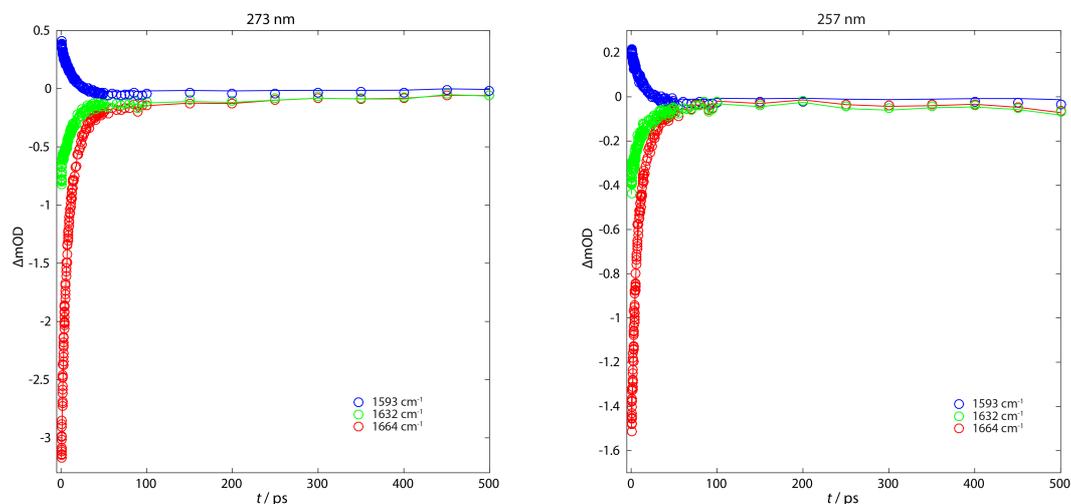
### Gaussian Fitting Analysis

A Gaussian fitting analysis was undertaken to analyse the d(GpT) TRIR data. These analyses modelled the experimental data with 8 time-dependent Gaussian basis functions. To model the 5 negative bleach features, Gaussian functions with the central frequency and FWHMs fixed to the values returned from fitting linear FTIR spectra. An additional two positive Gaussians were chosen to correspond to hot  $S_0$  vibrational molecules at  $\sim 1550$  and  $1590 \text{ cm}^{-1}$ . A third positive feature centred at  $1708 \text{ cm}^{-1}$  was used to model the  $G^+$  cation formation. As the parameters for

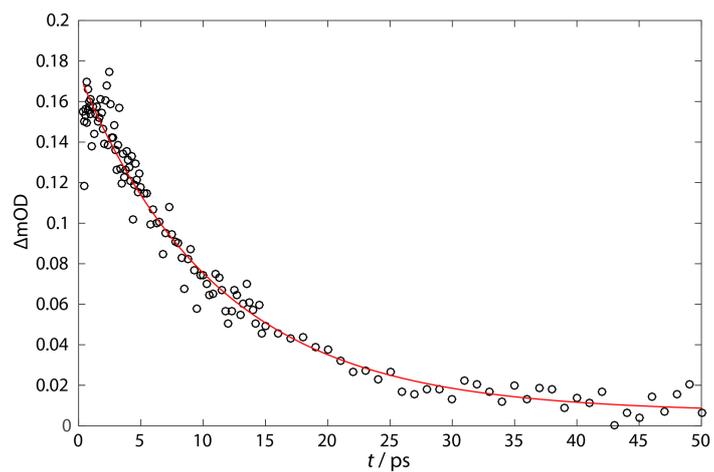
these positive features (central frequency or peak width) are unknown, these parameters were allowed to float alongside the amplitudes of the negative features. As we note in the main manuscript and many prior studies have shown, the vibrational cooling timescale of nucleotides is mode dependent, and thus the amplitudes of Gaussians associated with bleach features were fitted independently. Fig. S7 shows the TRIR data (left hand-side) and results of global analysis (right hand-side). Example fits to specific probe wavenumbers are displayed in Fig. S8.



**Fig. S7** LHS: TRIR data, RHS, global analysis for (a,b) 273 nm and (c,d) 257 nm excitation.

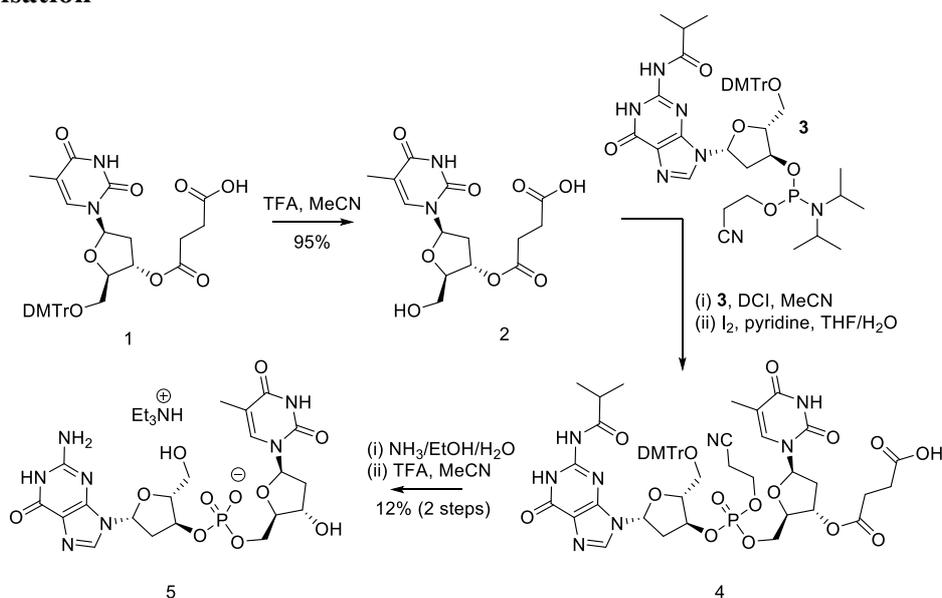


**Fig. S8** Examples of kinetic fits (solid lines) returned by the Gaussian fitting analysis to data (open circles) for the stated probe frequencies for (a) 273 nm and (b) 257 nm irradiation.



**Fig. S9** Kinetics associated with the  $G^+$  feature extracted from Gaussian fitting analysis. Data were fit to an exponential decay function with  $\tau = 11.1 \pm 0.3$  ps

## 2'-deoxyguanosine 3'-monophosphate 5'-thymidine synthetic procedure and characterisation



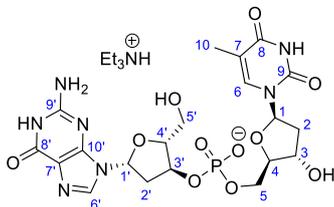
**Scheme S1:** Nucleotide coupling and deprotection procedure for the preparation of d(GpT) (**5**)

### General information

Commercial starting materials were purchased and used without further purification. Dry solvents were obtained by passage through a column of anhydrous alumina using equipment from Anhydrous Engineering (University of Bristol) based on the Grubbs' design and transferred anhydrously. Reactions requiring anhydrous conditions were performed under N<sub>2</sub> using oven-dried glassware. Thin-layer chromatography analysis was carried out using Merck, Kieselgel 60 F<sub>254</sub> plates and visualised under UV light. Flash column chromatography was performed on silica gel (Merck, 230-400 mesh, 40-63 μm particle size). Fractions determined to contain the desired product were combined and concentrated *in vacuo*. Preparative reversed-phase HPLC was performed on a Grace Discovery Sciences Reveleris Prep System with a Phenomenex Luna 5 μm C18(2) 100 Å AXIA packed (250 × 21.2 mm) column. The flow rate of the mobile phase was 14 mL / min. For compound detection, the instrument was set to monitor the 220 nm, 254 nm and 280 nm UV absorbance. Infra-red spectra were recorded in the range 2000-1300 cm<sup>-1</sup> on a Perkin Elmer Spectrum as a solution in D<sub>2</sub>O. Absorbance intensities are reported as strong (s), medium (m) and weak (w). NMR spectra were recorded on Varian 400 MHz or Varian 500 MHz spectrometers at 25.0 °C. Chemical shifts are quoted in ppm with spectra referenced to the residual proton of the deuterated solvent. Other abbreviations used are: br (broad), s (singlet), d (doublet), t (triplet), q (quartet), m (multiplet) and app. (apparent). Assignments of <sup>1</sup>H NMR and <sup>13</sup>C NMR signals were made using COSY,

HSQC and HMBC experiments. Mass spectra were determined by the University of Bristol mass spectrometry service using nanospray ionisation in negative ion mode.

### 2'-deoxyguanosyl-(3',5')-thymidine phosphate triethylammonium salt (**5**)

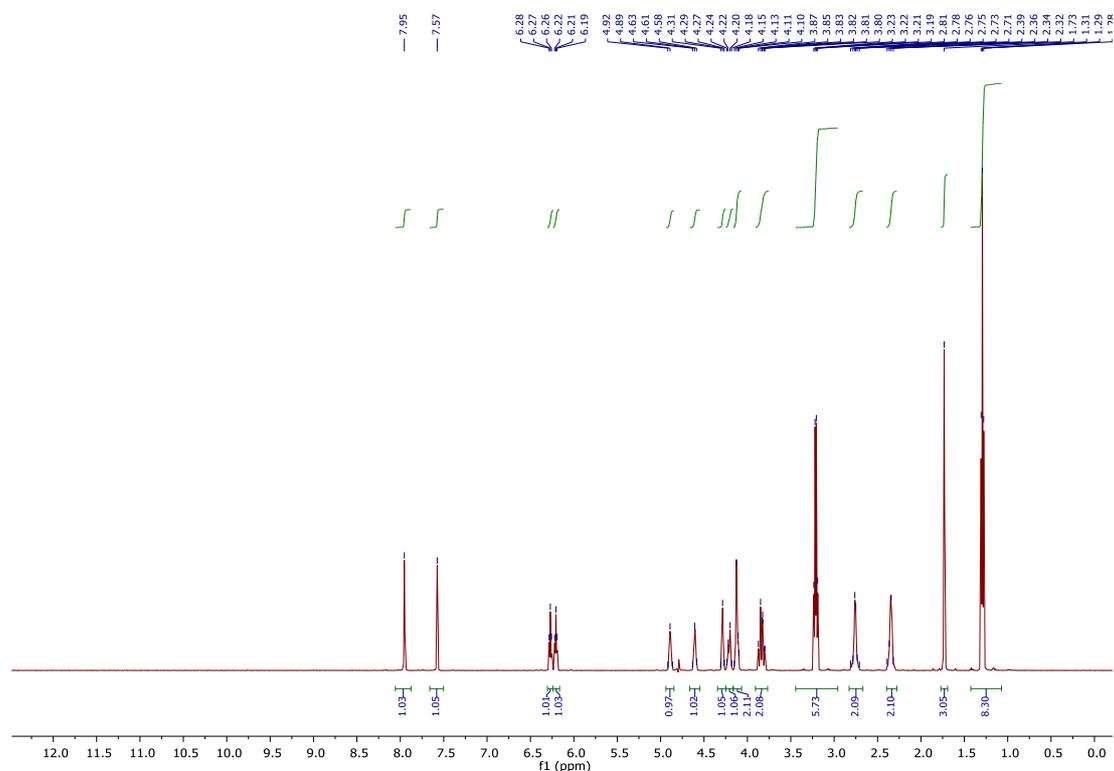


To a solution of 5'-*O*-DMT-thymidine-3'-*O*-succinic acid **1** (2.07 g, 3.21 mmol) in MeCN (60 mL) was added 3% v/v trifluoroacetic acid in MeCN (40 mL). After stirring for 30 min, TLC analysis indicated complete consumption of the starting material. The deep red mixture was concentrated *in vacuo* and the resultant oil was purified by flash column chromatography, eluting with 15% MeOH in CH<sub>2</sub>Cl<sub>2</sub>, to afford the 5'-deprotected nucleotide **2** as a colourless oil (1.04 g, 95%) which was used immediately in the following coupling step. Compound **2** (1.04 g, 3.04 mmol), phosphoramidite **3** (3.25 g, 3.87 mmol) and 4,5-dichloroimidazole (3.78 g, 32.0 mmol) were dried under high vacuum for 30 minutes. Anhydrous MeCN (15 mL) was added and the mixture stirred at rt for 90 min, then quenched with EtOH (200  $\mu$ L). After 5 min, pyridine (1.2 mL) was added, followed by a solution of I<sub>2</sub> (0.1 M in 2:1 THF:H<sub>2</sub>O) until the yellow-brown colour persisted following stirring for 5 min. The reaction was quenched by addition of 5% w/v aqueous NaHSO<sub>3</sub> (75 mL) and extracted with CHCl<sub>3</sub> (2 x 100 mL). The combined organics were dried (MgSO<sub>4</sub>), filtered and concentrated *in vacuo* to afford the crude protected dinucleotide **4** which was subjected to global deprotection without further purification. **4** was transferred to a reaction tube in EtOH (5 mL), aqueous ammonia (20 mL) was added and the tube sealed. The reaction was heated to 55 °C and stirred for 16 h, then concentrated *in vacuo*. The residue was dissolved in MeCN (60 mL) and 3% v/v trifluoroacetic acid in MeCN (40 mL) added. The reaction was stirred for 2 min and neutralised with aqueous NH<sub>4</sub>HCO<sub>3</sub> solution (40 mL). The mixture was partitioned between CH<sub>2</sub>Cl<sub>2</sub> (100 mL) and H<sub>2</sub>O (100 mL). The layers were separated and the organic back-extracted with H<sub>2</sub>O (50 mL). The combined aqueous extractions were lyophilised to a white solid. This was re-suspended in H<sub>2</sub>O (20 mL) and the insoluble solid removed by filtration and discarded after <sup>31</sup>P NMR indicated absence of the desired product. The remaining filtrate was purified by reversed-phase HPLC, using a mobile phase of 100 mM triethylammonium acetate in a gradient of 5% MeCN in H<sub>2</sub>O to 80% MeCN in H<sub>2</sub>O. Fractions determined to contain the desired product by <sup>31</sup>P NMR were

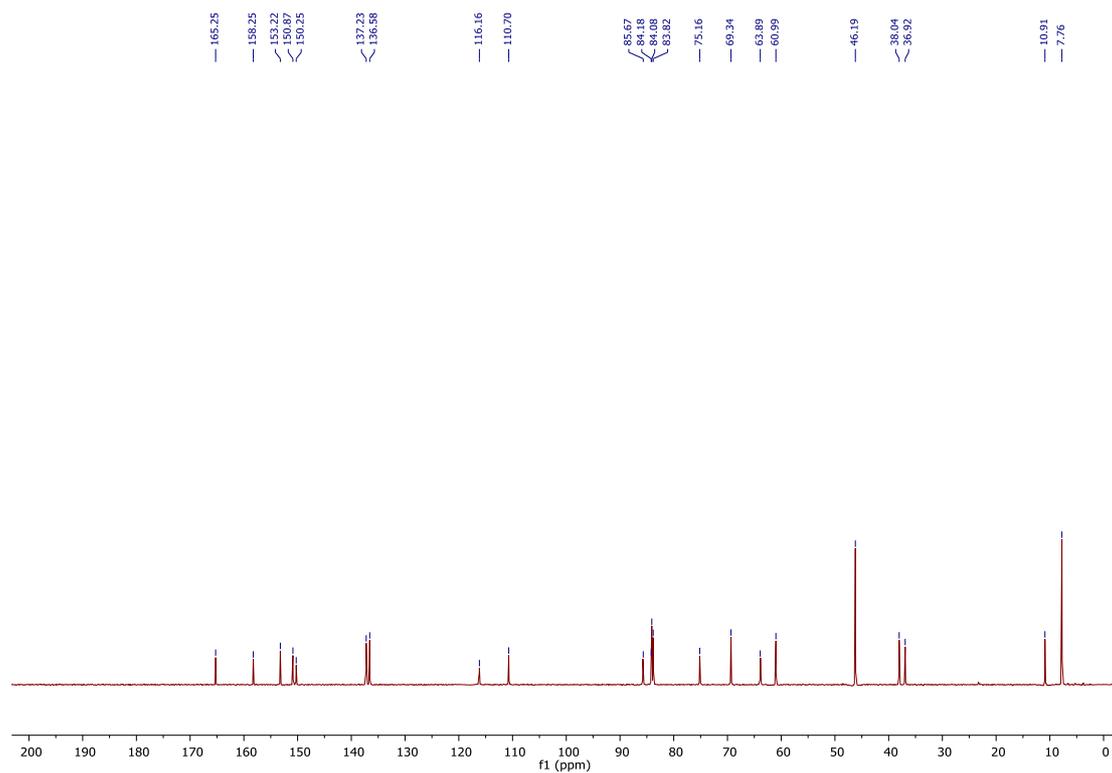
combined and lyophilised to afford triethylammonium salt **5** as a white powder (252 mg, 12% over two steps).

**<sup>1</sup>H NMR** (500 MHz, D<sub>2</sub>O) δ 7.95 (1H, s, 6'-CH), 7.57 (1H, s, 6-CH), 6.27 (1H, t, *J* = 6.7 Hz, 1-CH), 6.21 (1H, t, *J* = 6.6 Hz, 1'-CH), 4.96 – 4.83 (1H, m, 3'-CH), 4.66 – 4.54 (1H, m, 3-CH), 4.34 – 4.25 (1H, m, 4'-CH), 4.25 – 4.16 (1H, m, 5-CHH), 4.16 – 4.06 (2H, m, 5-CHH and 4-CH), 3.91 – 3.77 (2H, m, 5'-CH<sub>2</sub>), 3.21 (6H, q, *J* = 7.3 Hz, Et<sub>3</sub>NH<sup>+</sup>), 2.89 – 2.64 (2H, m, 2'-CH<sub>2</sub>), 2.43 – 2.28 (2H, m, 2-CH<sub>2</sub>), 1.73 (3H, s, 10-CH<sub>3</sub>), 1.29 (9H, t, *J* = 7.3 Hz, Et<sub>3</sub>NH<sup>+</sup>). **<sup>13</sup>C NMR** (126 MHz, H<sub>2</sub>O) δ 165.3 (8-CO), 158.2 (8'-CO), 153.2 (9'-CO), 150.9 (9-CO), 150.3 (10'-C), 137.2 (6'-CH), 136.6 (6-CH), 116.2 (7'-C), 110.7 (7-C), 85.7 (4'-CH), 84.2 (4-CH), 84.1 (1'-CH), 83.8 (1-CH), 75.2 (3'-CH), 69.3 (3-CH), 63.9 (5-CH<sub>2</sub>), 61.0 (5'-CH<sub>2</sub>), 46.2 (Et<sub>3</sub>NH<sup>+</sup>), 38.0 (2-CH<sub>2</sub>), 36.9 (2'-CH<sub>2</sub>), 10.9 (10-CH<sub>3</sub>), 7.8 (Et<sub>3</sub>NH<sup>+</sup>). **<sup>31</sup>P NMR** (162 MHz, H<sub>2</sub>O) δ -0.87. **v<sub>max</sub> / cm<sup>-1</sup>** (thin film) 1663 (s), 1633 (m), 1575 (m), 1475 (m), 1458 (m), 1406 (w), 1357 (w). **HRMS** (nanospray) calculated for C<sub>20</sub>H<sub>25</sub>N<sub>7</sub>O<sub>11</sub>P<sup>-</sup>: 570.1350; found: 570.1341.

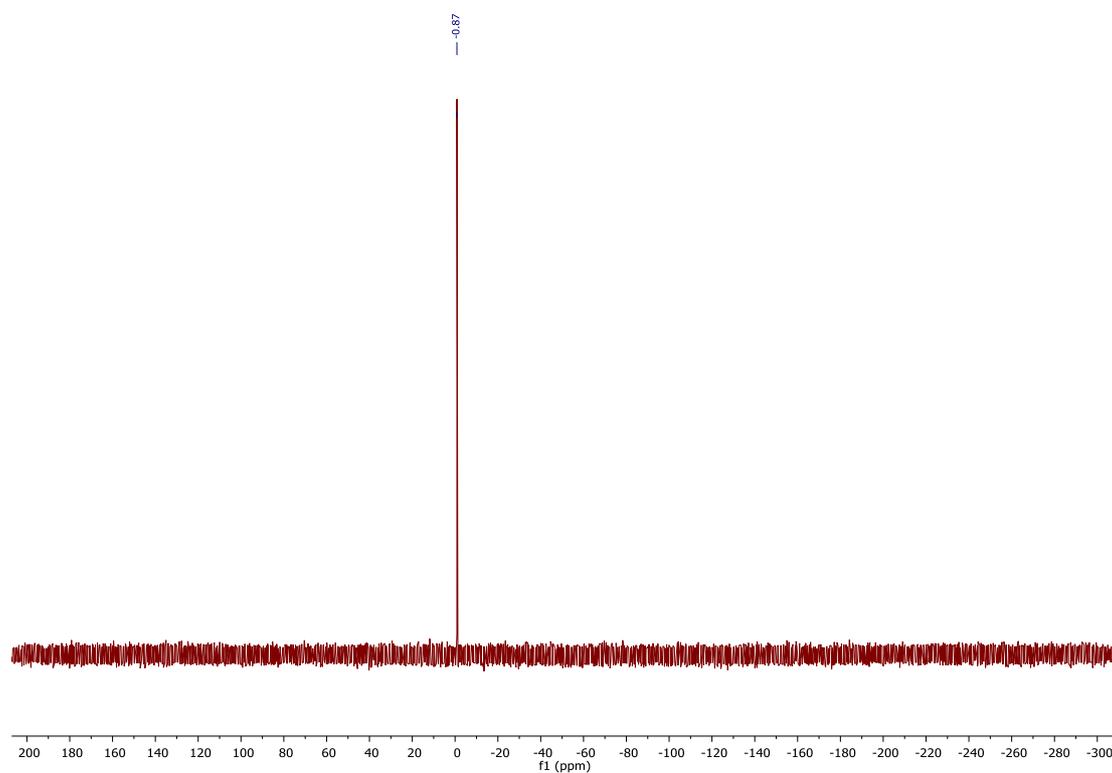
## Nuclear Magnetic Resonance



**Fig. S10** <sup>1</sup>H-NMR 500 MHz spectrum of d(GpT) in D<sub>2</sub>O.



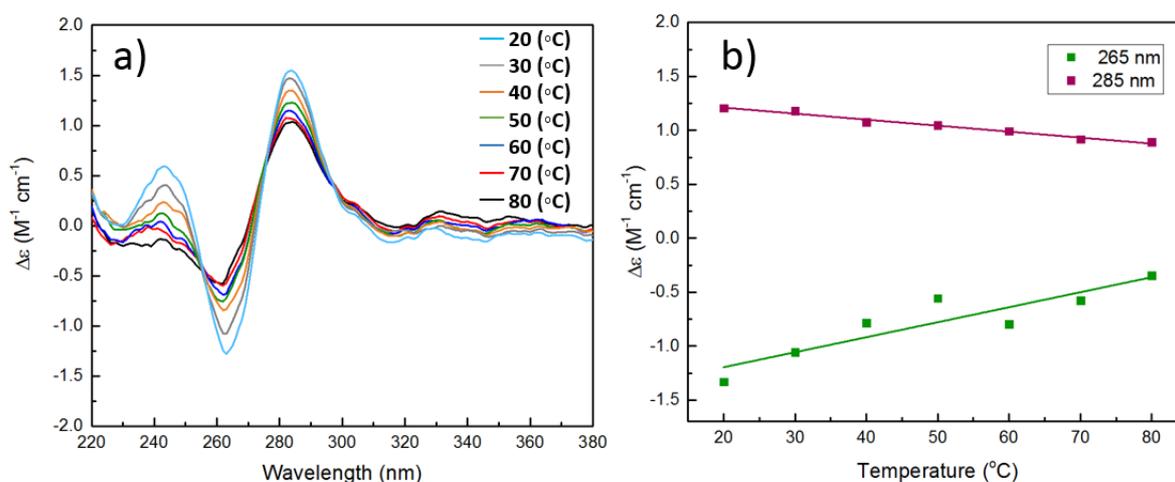
**Fig. S11**  $^{13}\text{C}$ -NMR 126 MHz spectrum of d(GpT) in  $\text{H}_2\text{O}$



**Fig. S12**  $^{31}\text{P}$ -NMR 400 MHz spectrum of d(GpT) in  $\text{H}_2\text{O}$

## Temperature Dependent Circular Dichroism Spectroscopy

Circular dichroism (CD) is a well known probe of  $\pi$ -stacking in chiral structures such as DNA.<sup>1</sup> Temperature dependent CD measurements were used to determine whether any d(GpT) molecules are  $\pi$ -stacked at room temperature in a 1  $\mu$ M d(GpT) buffered D<sub>2</sub>O solution. As the temperature was ramped from 20 to 80 °C, CD spectra showed a decreased molar ellipticity associated with peaks at 265 and 285 nm (see Fig. S14); the additional thermal energy allows molecules to explore a wider range of the conformational free energy surface reducing the total number of  $\pi$ -stacked molecules in the ensemble.



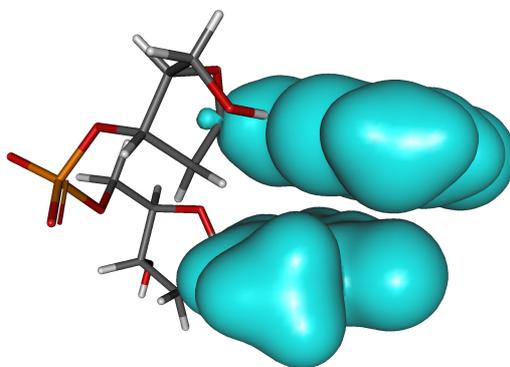
**Fig. S13** (a) d(GpT) CD spectra between for the  $20 \leq T \leq 80$  °C range, and (b) temperature dependence of the maxima and minima observed in (a). Note all spectra are given in units of molar ellipticity per residue.

## Computational Methods

Classical molecular dynamics (MD) simulations of solvated d(GpT) were performed using the Amber 14 software, utilising standard Amber force field parameters for d(GpT). The simulation was composed of a single d(GpT) molecule (initially in an open conformation), surrounded by 2356 TIP3P water molecules and a sodium cation in an orthorhombic box. The energy of the initial configuration was first minimized for 2500 cycles. The system was heated to 300 K for 25 ps in an NVT ensemble, followed by 20 ns of equilibration in an NpT ensemble. A production run was performed for 40 ns (NpT ensemble), from which representative geometries were selected. Each selected geometry was re-optimised at the DFT/ $\omega$ B97X-D/6-31G\*/IEFPCM (water) level of theory, as described in the main manuscript. LR-TDDFT was

used to describe  $\pi$ -stacked nucleobases as motivated by previous studies.<sup>2-7</sup> All calculations in this study used the long-range and dispersion corrected functional  $\omega$ B97X-D.

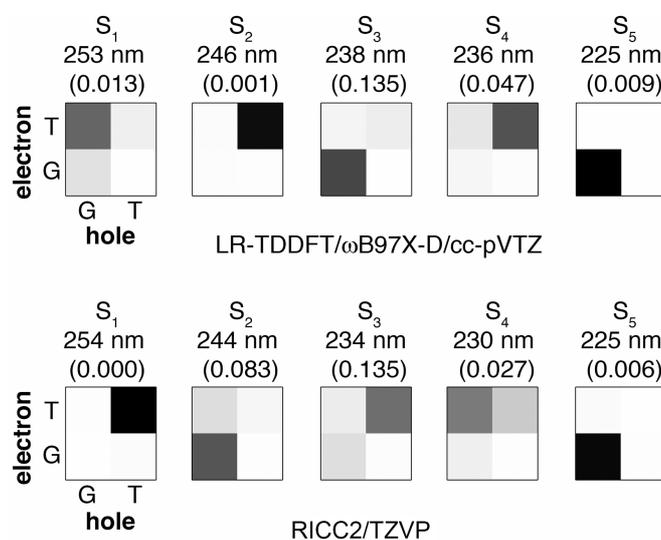
To validate our computational approach and the choice of  $\omega$ B97X-D exchange-correlation function, we performed a benchmark calculation of a reduced model of the isolated dinucleotide. Specifically, the reduced system was comprised solely from the guanine and thymine moieties in the geometry of the  $d(\text{GpT})_s$  conformer, deleting the sugar-phosphate backbone and terminating the ring N atoms with hydrogen atoms. We henceforth refer to this structure as  $\text{GT}_i$ . The vertical electronic transitions of  $\text{GT}_i$  calculated at the LR-TDDFT/ $\omega$ B97X-D/cc-pVTZ level (in Gaussian09<sup>8</sup>) were compared with those obtained with RICC2/TZVP (performed with Turbomole<sup>9</sup>). The phosphate and sugar groups removed in our reduced model system do not contribute significantly to the low-lying vertical transitions, as attested by previous work,<sup>2</sup> the analysis of the transition-density matrices (Fig. 3 in main text), and additional projector-embedding calculations.<sup>10-12</sup> Projector embedding calculations for  $d(\text{GpT})_s$  required an active region of 144 electrons (Fig. S14) out of a total 298 to ensure proper convergence (within  $< 0.1$  eV) of configuration interaction singles (CIS) in Hartree-Fock calculations for the first 5 excited states.



**Fig. S14** Electron density plot with an iso-surface value of  $0.01 \text{ e}^-/\text{au}^3$  for the 144-electron active region determined for accurate embedded excitation energies.

Calculations of the isolated  $\text{GT}_i$  structure using RICC2/TZVP and LR-TDDFT/ $\omega$ B97X-D/cc-pVTZ returned similar electronic characters for the low-lying states as detailed in Fig. S15. The ordering of the first two excited states is inverted between the two methods, however it is important to note that in both cases the states are very close in energy ( $< 0.2$  eV). Importantly,

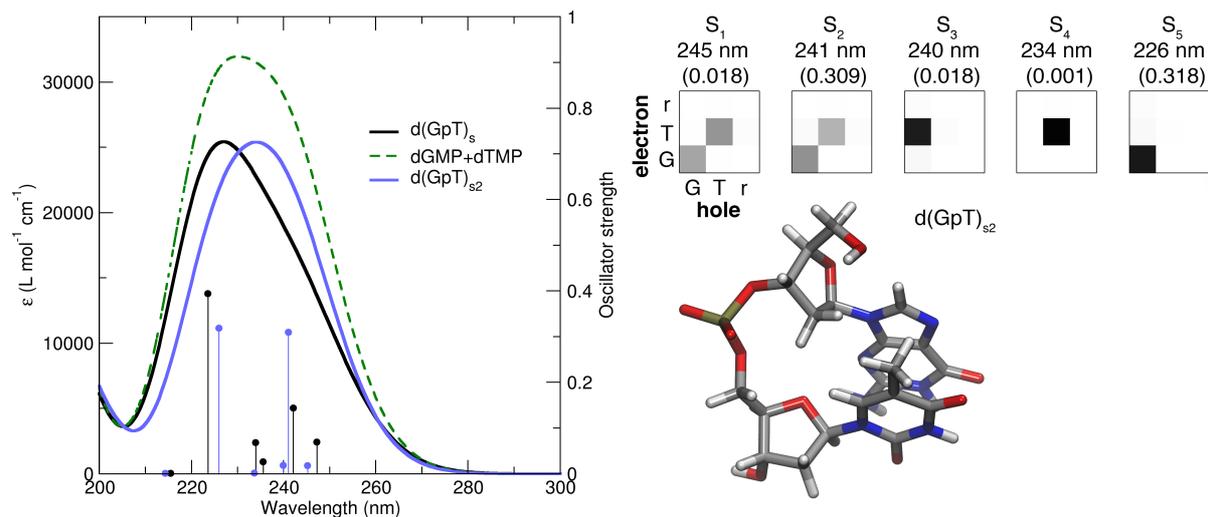
RICC2 confirms the presence of states with a significant charge-transfer character, in agreement with  $\omega$ B97X-D calculations.



**Fig. S15** LR-TDDFT/ $\omega$ B97X-D/cc-pVTZ (upper panel) and RICC2/TZVP (lower panel) vertical transitions for the isolated and reduced model GT<sub>i</sub> system (see text). Analysis of the transition-density matrices for the first five low-lying vertical excitation of  $\pi$ -stacked GT. The horizontal axis represents the location of the hole, the vertical axis gives the position of the electron. The grey scale is used to depict the importance of a certain electron-hole correlation to each displayed transition (black being the most important).

An additional  $\pi$ -stacked structure was selected from the classical molecular dynamics simulation, optimised with DFT, and the associated vertical electronic transitions calculated—see Fig. S16. This structure, labelled d(GpT)<sub>s2</sub>, is similar to d(GpT)<sub>s</sub> but the guanine nucleobase is rotated by 180° (compare inset of Fig. S16 with Fig. 1a).  $\omega$ B97X-D/cc-pVTZ/IEFPCM// $\omega$ B97X-D/6-31G\*/IEFPCM calculations of d(GpT)<sub>s</sub> and d(GpT)<sub>s2</sub> revealed that the ground state energy of the d(GpT)<sub>s</sub> is more stable by 34.4 kJ/mol (0.36 eV).

The vertical electronic transitions of the d(GpT)<sub>s2</sub> geometry have a higher percentage of localised character compared to the respective transitions in d(GpT)<sub>s</sub>. With reference to the Fig. 3(c) and the right-hand panel of Fig. S16: the S<sub>1</sub> and S<sub>2</sub> states of d(GpT)<sub>s2</sub> show coupled excitation of both the guanine and thymine nucleobases, whereas S<sub>3</sub> is almost purely G → T charge-transfer in character. This coupling results in the S<sub>0</sub>–S<sub>2</sub> electronic transition having a larger associated oscillator strength and shifts the maximum of the theoretical spectra towards the red compared to d(GpT)<sub>s</sub> (compare the left panel of Fig. S16 to Fig. 3(a)).



**Fig. S16** LR-TDDFT/ $\omega$ B97X-D/cc-pVTZ/IEFPCM vertical transitions of  $d(\text{GpT})_{s2}$ . Left panel: calculated absorption spectra. Right panel: analysis of the transition-density matrices for the first five low-lying vertical excitation of  $d(\text{GpT})_{s2}$ . The horizontal axis represents the location of the hole, the vertical axis gives the position of the electron. The grey scale is used to depict the importance of a certain electron-hole correlation to each displayed transition (black being the most important). The structure of  $d(\text{GpT})_{s2}$  is displayed in the bottom right.

## Cartesian co-ordinations of the minimum energy d(GpT)<sub>0</sub> geometry

H	2.925399	-1.201502	-1.141492
O	1.990490	-1.377033	-1.385668
C	1.246528	-1.411116	-0.189794
H	1.556429	-2.244442	0.459115
H	0.197708	-1.573618	-0.462774
C	1.337319	-0.118051	0.609157
H	0.610094	-0.147767	1.425787
O	2.649220	0.000290	1.184140
C	3.257794	1.204828	0.780099
H	3.159920	1.967656	1.561741
N	4.683640	0.978692	0.624348
C	5.687529	1.703725	1.231546
H	5.451023	2.521502	1.899205
N	6.886945	1.302745	0.912742
C	6.668604	0.245927	0.050859
C	7.607849	-0.607007	-0.617399
O	8.829377	-0.592125	-0.591053
N	6.930427	-1.578589	-1.393114
H	7.537811	-2.230889	-1.874907
C	5.574159	-1.715020	-1.512330
N	5.106769	-2.743162	-2.253280
H	5.695091	-3.197092	-2.935401
H	4.113913	-2.750678	-2.439725
N	4.721223	-0.918734	-0.901499
C	5.313342	0.030288	-0.136632
C	1.114896	1.154534	-0.227919
H	0.566359	0.946722	-1.149381
C	2.535109	1.632448	-0.495850
H	2.949587	1.099701	-1.354874
H	2.590144	2.707374	-0.671773
O	0.434687	2.143324	0.537413
P	-1.192271	2.311734	0.332091
O	-1.521512	2.258855	-1.128004
O	-1.603868	3.435482	1.226242
O	-1.742421	0.874150	0.935937
C	-1.952940	0.731776	2.329403
H	-2.530645	1.578589	2.711738
H	-0.993027	0.703511	2.864424
C	-2.705495	-0.560005	2.593887
H	-2.799472	-0.675182	3.680607
O	-4.021745	-0.484739	2.039293
C	-4.204381	-1.479206	1.049451
H	-4.816165	-2.292799	1.440589
N	-4.969845	-0.910180	-0.055282
C	-4.510251	0.226606	-0.693906
H	-3.565899	0.608431	-0.316668
C	-5.169862	0.837221	-1.699048
C	-4.662350	2.070147	-2.385034
H	-4.566411	1.898848	-3.463003
H	-5.362949	2.902759	-2.255056
H	-3.684755	2.354403	-1.985977
C	-6.440953	0.273876	-2.132881
O	-7.148213	0.714831	-3.028811
N	-6.833655	-0.863725	-1.427895
H	-7.715346	-1.282403	-1.698094
C	-6.176297	-1.497392	-0.394676
O	-6.633451	-2.490819	0.149653
C	-2.063257	-1.823910	1.988084
H	-0.981565	-1.703613	1.855700
C	-2.805414	-1.964868	0.670113
H	-2.359451	-1.307258	-0.079130
H	-2.807719	-2.992105	0.301873
O	-2.352204	-2.977498	2.758888
H	-1.918981	-2.882032	3.616987

## Cartesian co-ordinations of the minimum energy d(GpT)<sub>s</sub> geometry

H	1.616897	3.089911	0.343877
O	1.138473	3.860828	0.717369
C	0.368938	4.431861	-0.317164
H	1.005837	4.834309	-1.119165
H	-0.186618	5.268720	0.119182
C	-0.625999	3.460555	-0.942844
H	-1.302458	4.014398	-1.602881
O	0.090544	2.497216	-1.736648
C	-0.179687	1.190084	-1.277464
H	-0.940205	0.712692	-1.904418
N	1.011613	0.377371	-1.433254
C	1.108027	-0.774962	-2.184107
H	0.237072	-1.183632	-2.677434
N	2.304412	-1.295682	-2.188015
C	3.046793	-0.436128	-1.402192
C	4.424442	-0.501350	-1.012709
O	5.271345	-1.337737	-1.290474
N	4.762295	0.603275	-0.196567
H	5.713939	0.593026	0.150018
C	3.901916	1.566857	0.248079
N	4.375224	2.484515	1.121816
H	5.368428	2.640940	1.206279
H	3.785752	3.289579	1.282178
N	2.635186	1.615281	-0.109805
C	2.261831	0.598615	-0.923358
C	-1.441511	2.655622	0.079995
H	-1.502632	3.158416	1.047552
C	-0.674460	1.343137	0.153467
H	0.168156	1.440501	0.839891
H	-1.307780	0.510603	0.445407
O	-2.750831	2.466120	-0.449324
P	-3.912234	1.764483	0.495057
O	-5.210581	2.069788	-0.180682
O	-3.645723	2.059080	1.936565
O	-3.561790	0.156808	0.304644
C	-3.683352	-0.391481	-0.994451
H	-3.149698	0.228548	-1.725045
H	-4.737436	-0.435462	-1.295444
C	-3.114119	-1.794080	-1.018503
H	-3.306787	-2.224686	-2.008827
O	-1.700602	-1.753461	-0.792312
C	-1.358813	-2.812907	0.090471
H	-1.226716	-3.743575	-0.464438
N	-0.070487	-2.524270	0.691481
C	0.068406	-1.478969	1.581874
H	-0.860919	-1.014614	1.883005
C	1.247249	-1.035611	2.057308
C	1.377743	0.124614	2.996475
H	0.395328	0.513399	3.276948
H	1.910265	-0.164920	3.907891
H	1.947596	0.934650	2.527175
C	2.459513	-1.702837	1.605375
O	3.598567	-1.396555	1.929682
N	2.237713	-2.768608	0.736393
H	3.062259	-3.222716	0.362302
C	1.042343	-3.206268	0.208612
O	0.982372	-4.119978	-0.595005
C	-3.705418	-2.750841	0.038228
H	-4.596214	-2.325391	0.512595
C	-2.553965	-2.922096	1.029000
H	-2.575100	-2.102297	1.751158
H	-2.587338	-3.873572	1.563930
O	-4.018221	-3.960736	-0.632888
H	-4.407510	-4.563119	0.014095

## Cartesian co-ordinations of the minimum energy d(GpT)<sub>s2</sub> geometry

H	2.318332	-3.910911	1.581575
O	1.909094	-4.247506	0.774967
C	2.856133	-4.173284	-0.279423
H	2.523893	-4.886153	-1.038859
H	3.850750	-4.485684	0.064692
C	2.969388	-2.800511	-0.935998
H	3.781995	-2.845957	-1.668881
O	1.763684	-2.478907	-1.638487
C	1.093036	-1.395436	-1.024038
H	1.131713	-0.509359	-1.663058
N	-0.311872	-1.729209	-0.885285
C	-0.831491	-2.810909	-0.207109
H	-0.170518	-3.522480	0.272849
N	-2.136699	-2.860766	-0.238911
C	-2.499073	-1.754427	-0.983906
C	-3.795677	-1.249683	-1.317283
O	-4.903395	-1.672510	-1.006995
N	-3.681924	-0.079190	-2.098882
H	-4.563131	0.376143	-2.303492
C	-2.511633	0.533968	-2.468981
N	-2.618341	1.701369	-3.157199
H	-1.757354	2.029381	-3.570571
H	-3.448834	1.880839	-3.702403
N	-1.324523	0.075147	-2.151559
C	-1.379707	-1.047807	-1.399780
C	3.214058	-1.633043	0.027334
H	3.737618	-1.936333	0.939485
C	1.795108	-1.152300	0.308169
H	1.345128	-1.771702	1.087488
H	1.745931	-0.106781	0.599981
O	3.996749	-0.676464	-0.674767
P	4.372145	0.759542	0.039453
O	5.572081	1.278759	-0.687631
O	4.342014	0.624682	1.532232
O	3.033242	1.597603	-0.456421
C	2.510320	2.613577	0.376237
H	3.176312	3.488083	0.383849
H	2.402271	2.252889	1.406513
C	1.163858	3.030292	-0.177864
H	1.280826	3.295929	-1.235871
O	0.249842	1.930967	-0.087880
C	-0.984538	2.402818	0.402211
H	-1.629984	2.744663	-0.410513
N	-1.685763	1.292233	1.038838
C	-0.992734	0.383264	1.808755
H	0.070139	0.576180	1.872413
C	-1.564710	-0.672046	2.420100
C	-0.798761	-1.675508	3.226438
H	-1.206965	-1.758170	4.238899
H	-0.866823	-2.667761	2.767400
H	0.256170	-1.398503	3.300462
C	-3.003685	-0.853494	2.276047
O	-3.659925	-1.755734	2.774984
N	-3.630708	0.131544	1.512916
H	-4.627019	0.023628	1.365685
C	-3.056058	1.210130	0.876832
O	-3.712262	2.015407	0.234145
C	0.513213	4.216793	0.577662
H	1.230467	4.715572	1.241098
C	-0.612684	3.541677	1.346121
H	-0.239533	3.129232	2.289324
H	-1.440797	4.221394	1.552266
O	-0.077250	5.147526	-0.312179
H	0.628870	5.559805	-0.826444

## References

- 1 J. Kypr, I. Kejnovska, D. Renciuik and M. Vorlickova, *Nucleic Acids Res.*, 2009, **37**, 1713–1725.
- 2 F. Santoro, V. Barone and R. Improta, *Proc. Natl. Acad. Sci. U.S.A.*, 2007, **104**, 9931–9936.
- 3 F. Santoro, V. Barone and R. Improta, *J. Comput. Chem.*, 2008, **29**, 957–964.
- 4 A. E. Raeber and B. M. Wong, *J. Chem. Theory Comput.*, 2015, **11**, 2199–2209.
- 5 H. Sun, S. Zhang, C. Zhong and Z. Sun, *J. Comput. Chem.*, 2015, **37**, 684–693.
- 6 F. Plasser, A. J. A. Aquino, H. Lischka and D. Nachtigallová, in *Photoinduced Phenomena in Nucleic Acids II: DNA Fragments and Phenomenological Aspects*, eds. M. Barbatti, A. C. Borin and S. Ullrich, Springer International Publishing, Cham, 2015, pp. 1–37.
- 7 J. J. Nogueira, F. Plasser and L. González, *Chem. Sci.*, 2017, **8**, 5682–5691.
- 8 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, 2016.
- 9 TURBOMOLE V7.1 2016, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from <http://www.turbomole.com>.
- 10 F. R. Manby, M. Stella, J. D. Goodpaster and T. F. Miller, *J. Chem. Theory Comput.*, 2012, **8**, 2564–2568.
- 11 S. J. Bennie, M. Stella, T. F. Miller and F. R. Manby, *J. Chem. Phys.*, 2015, **143**, 024105–8.
- 12 S. J. Bennie, B. F. E. Curchod, F. R. Manby and D. R. Glowacki, *J. Phys. Chem. Lett.*, 2017, **8**, 5559–5565.