

# Conformational control of benzophenone-sensitized charge transfer in dinucleotides

Thomas Merz,<sup>†‡</sup> Matthias Wenninger,<sup>¶‡</sup> Michael Weinberger,<sup>&</sup> Eberhard Riedle,<sup>\*¶</sup>  
Hans-Achim Wagenknecht,<sup>\*&</sup> and Martin Schütz<sup>\*,†</sup>

<sup>†</sup> *Institute of Physical and Theoretical Chemistry, University of Regensburg,  
Universitätsstraße 31, D-93040 Regensburg, Germany*

<sup>&</sup> *Karlsruhe Institute of Technology, Department of Chemistry, Fritz-Haber-Weg 6, Campus  
Süd, Geb. 30.42, 76131 Karlsruhe, Germany*

<sup>¶</sup> *Lehrstuhl für BioMolekulare Optik, Ludwig-Maximilians-Universität München,  
Oettingenstraße 67, 80538 Munich, Germany*

<sup>‡</sup> These authors contributed equally.

<sup>\*</sup> To whom correspondence should be addressed

*E-mail: martin.schuetz@chemie.uni-regensburg.de  
Riedle@physik.uni-muenchen.de  
Wagenknecht@kit.edu*

## Supporting Information

### Content

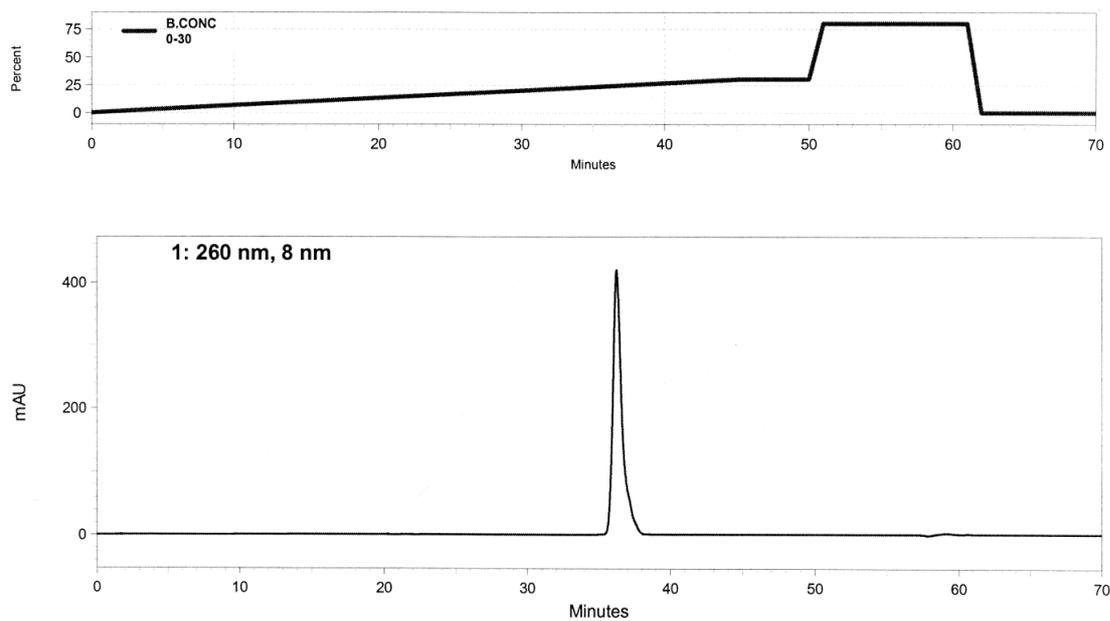
1.	Preparation of dinucleotides	2
2.	Transient spectroscopic methods and raw data	7
3.	Modelling of the ET dynamics with Marcus Theory	13
4.	Computational Work	17
5.	Additional references	34

## 1. Preparation of dinucleotides

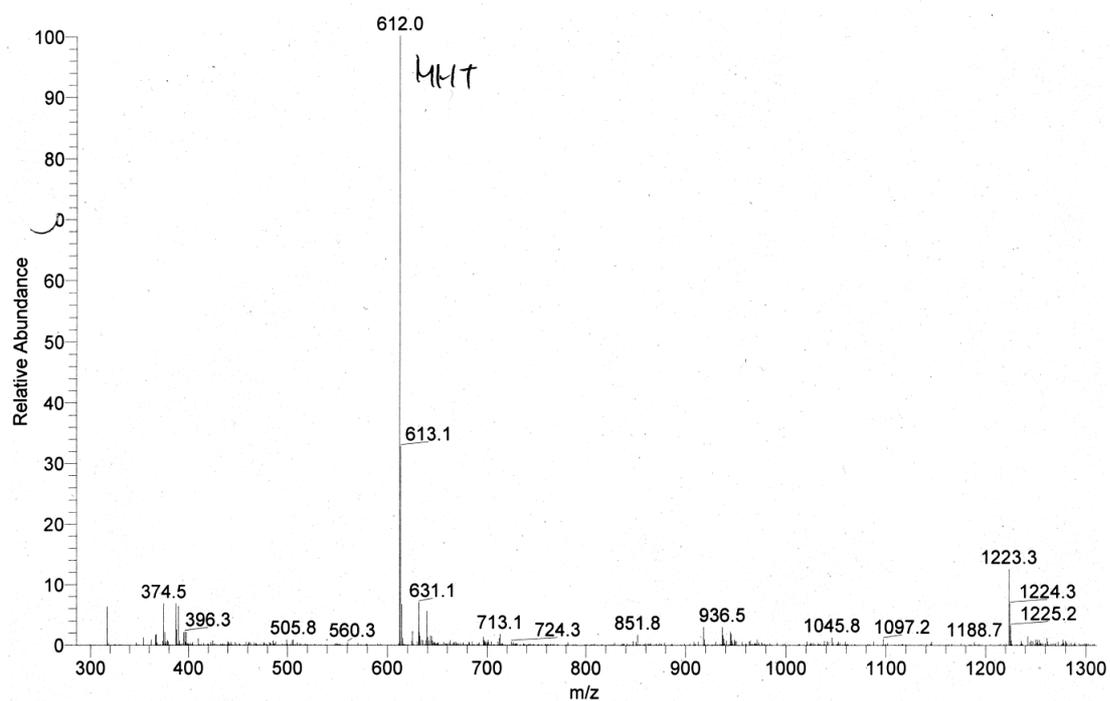
The dinucleotides **1X** were prepared on a DNA synthesizer using standard phosphoramidite chemistry. The synthesis of the phosphoramidite as DNA building block of **1** was described recently.<sup>24</sup> Reagents and controlled pore glass (CPG) (1  $\mu$ mol) were purchased and used without further purification. Dinucleotide synthesis was performed using standard coupling conditions. The concentration of the artificial building block was increased to 0.1 M (MeCN). After preparation, the oligonucleotides were cleaved from the resin and deprotected by treatment with conc.  $\text{NH}_4\text{OH}$  at 50 °C for 24 h. The oligonucleotides were purified by semi-preparative HPLC column (RP-18), using the following gradient, 0-15 % B over 50 min for oligomers (A: 50 mM  $\text{NH}_4\text{OAc}$ -buffer, B: acetonitrile), and identified by MS (ESI).

**Table SI-1** Summary of mass spectrometry analysis and extinction coefficient of dinucleotides **1X**.

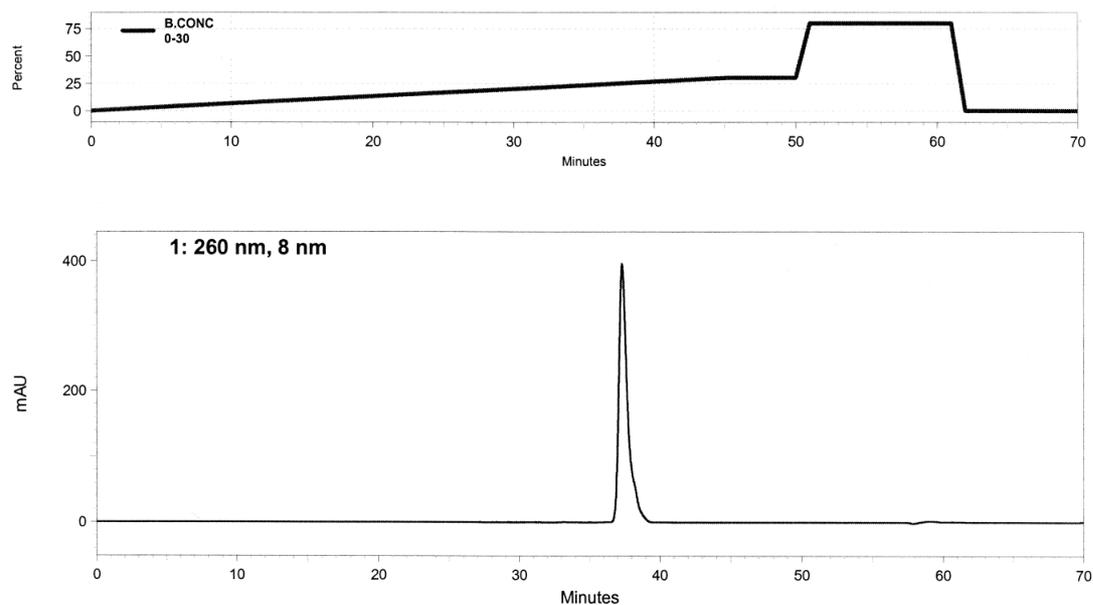
Dinucleotide	calc. mass [g/mol]	exp. mass [g/mol]	$\epsilon_{260 \text{ nm}}$ [ $\text{mM}^{-1} \text{ cm}^{-1}$ ]
<b>1A</b>	611.2	612.0 $[\text{M}+\text{H}]^+$	34.4
<b>1T</b>	602.2	603.1 $[\text{M}+\text{H}]^+$	28.4
<b>1G</b>	627.2	628.1 $[\text{M}+\text{H}]^+$	31.0
<b>1C</b>	587.2	588.0 $[\text{M}+\text{H}]^+$	27.1



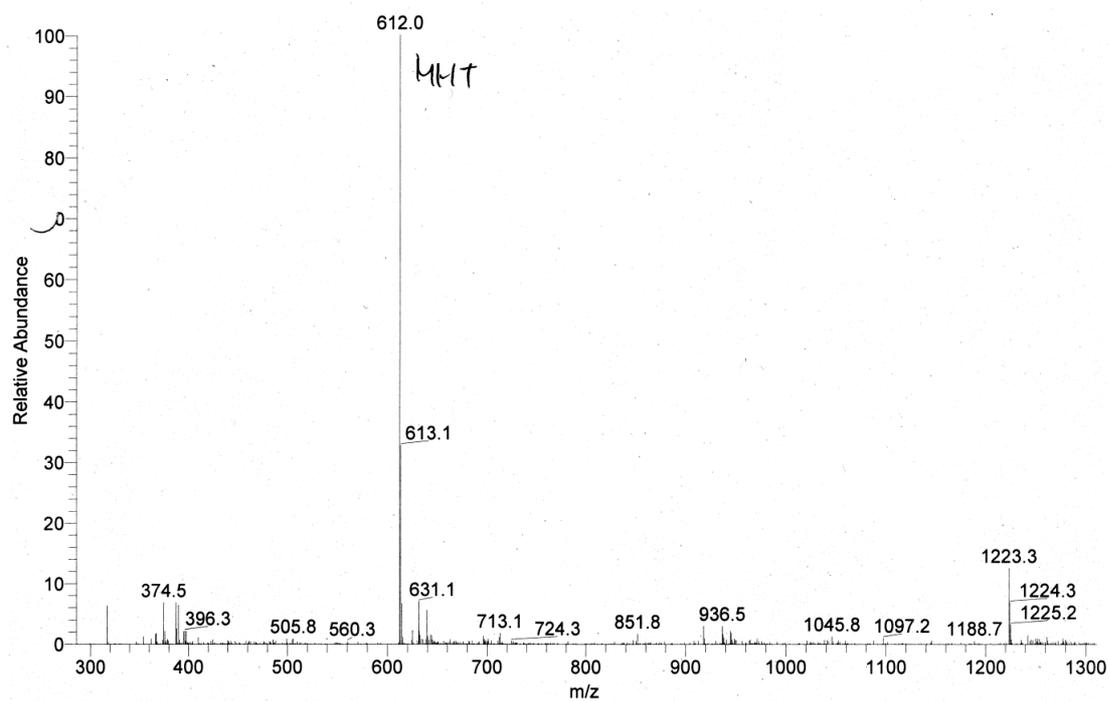
**Figure SI-1** Image of analytical HPLC run of **1A** including the gradient on top.



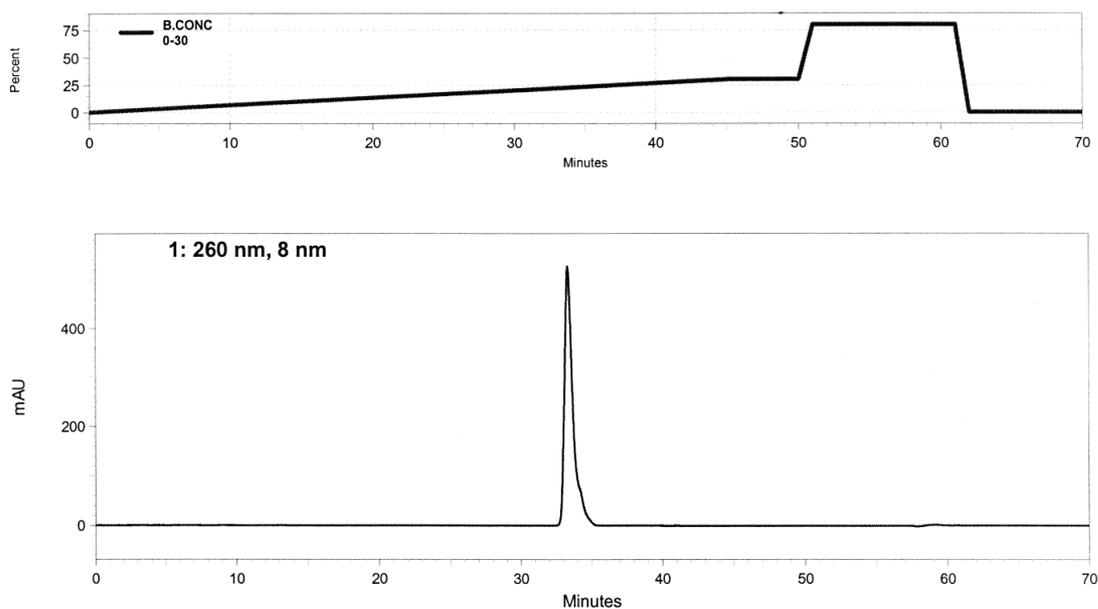
**Figure SI-2** Image of ESI mass spectrum of **1A**.



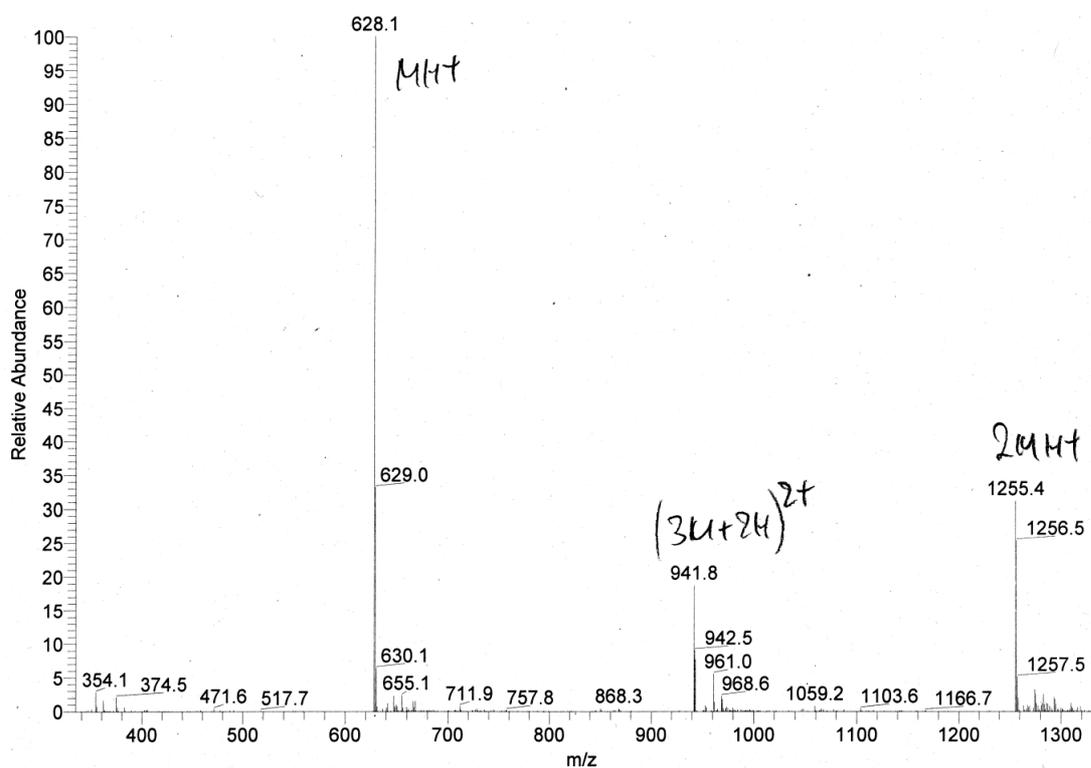
**Figure SI-3** Image of analytical HPLC run of **1T** including the gradient on top.



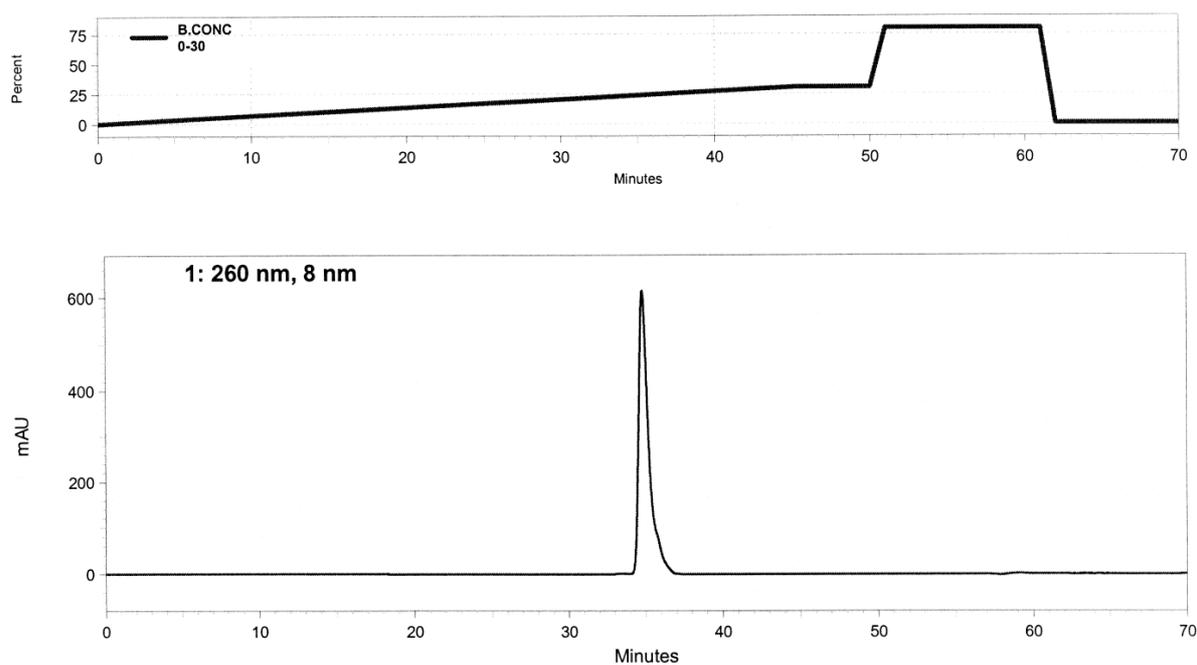
**Figure SI-4** Image of ESI mass spectrum of **1T**.



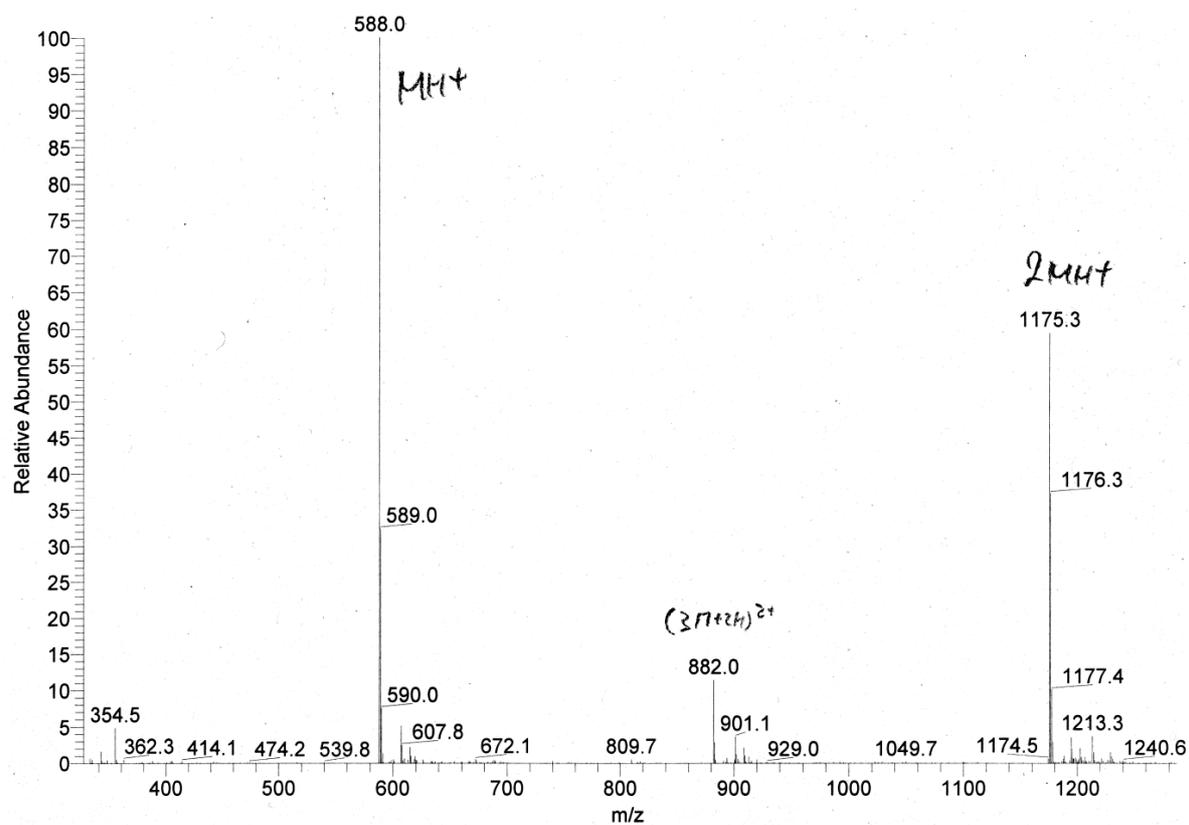
**Figure SI-5** Image of analytical HPLC run of **1G** including the gradient on top.



**Figure SI-6** Image of ESI mass spectrum of **1G**.



**Figure SI-7** Image of analytical HPLC run of **1C** including the gradient on top.



**Figure SI-8** Image of ESI mass spectrum of **1C**.

## 2. Transient spectroscopic methods and raw data

The femtosecond broadband pump-probe setup has been described in detail.<sup>43</sup> Briefly, a Ti:sapphire amplifier system (CPA 2001, Clark-MXR) with a repetition rate of 1 kHz was used to pump a noncollinear optical parametric amplifier (NOPA) which provides pulses with a center wavelength of 710 nm. After compression and frequency doubling in a BBO-crystal (1 mm thickness) we achieved pump pulses with a center wavelength of 355 nm and a pulse length of 70 fs. The beam with a pulse energy of 600 nJ was focused to a spot size of 100  $\mu\text{m}$  FWHM in a 1 mm cuvette containing 100-200 nmol of the sample diluted in 100-150  $\mu\text{L}$  of solvent of spectroscopic grade. The optical density at the excitation wavelength was below 0.05. A supercontinuum (290-750 nm) was generated and used as probe by focusing another part of the Ti:sapphire laser into a rotating  $\text{CaF}_2$  (4 mm thickness) disk. The polarizations of the pump and probe pulses were set to the magic angle ( $54.7^\circ$ ). A computer controlled delay line was used to set pump-probe delays up to 2 ns. After the interaction in the sample, the probe light was dispersed with a fused silica prism and detected with a 524 pixel CCD. The chirp of the white light was corrected for prior to data analysis, and the resulting temporal resolution was better than 100 fs. To obtain reliable kinetics from the TA data in view of the limited signal to noise ratio due to the limited amount of sample we used a global data analysis based on the formalism developed by Fita et. al.<sup>54</sup> It extracts decay associated difference spectra (DADS) from the transient data<sup>43</sup>. Before applying the global fit routine to our TA data, the wavelength scale was rescaled to a scale linear in  $\Delta E$  to avoid the enhancement of the weight of the long-wave components. The points of time showing effects of temporal pump-probe overlap were clipped. The global fit analysis provides a global set of time constants  $\sigma_i$  and the corresponding DADS  $A_i(\lambda)$ .

$$\Delta\text{OD}(\lambda, t) = \sum_i A_i(\lambda) \cdot \exp\left(-t/\tau_i\right) \quad (5)$$

Further data processing is based on a sequential rate model, shown in equation 6. We assume that the rate  $k_{\text{isc}}$  of ISC is fast in comparison to the remaining  $S_1$  deactivation processes.

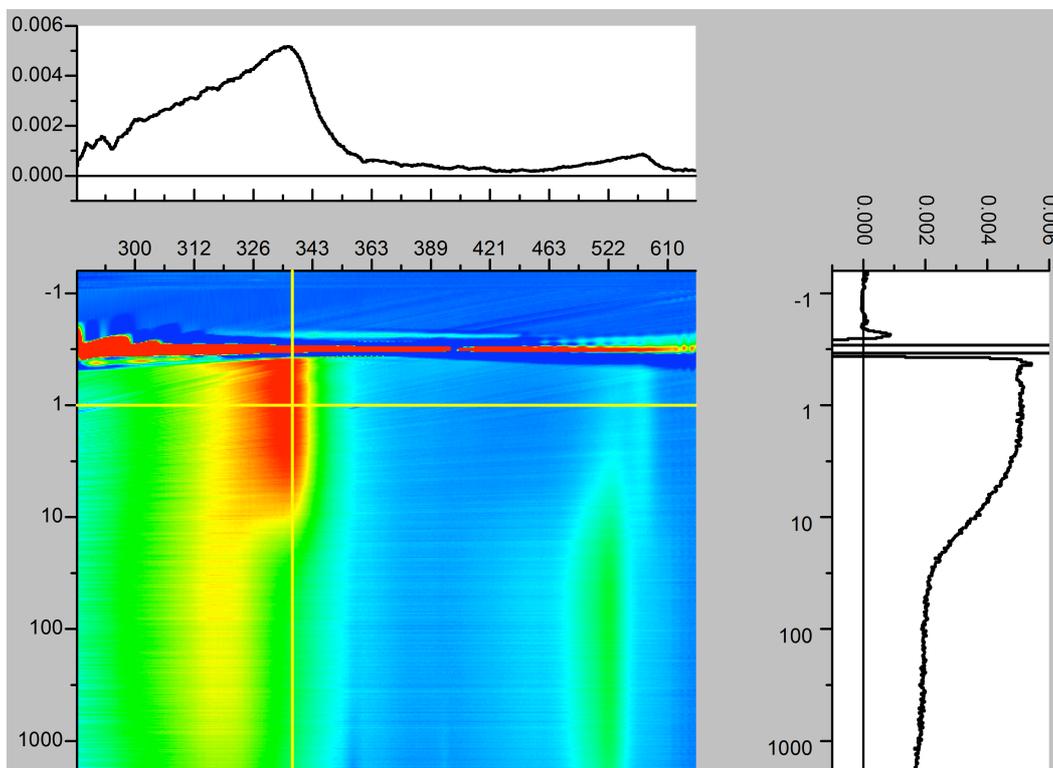
$$\begin{pmatrix} \dot{n}_{S1} \\ \dot{n}_{T1} \end{pmatrix} = \begin{pmatrix} -k_{\text{ISC}} & 0 \\ +k_{\text{ISC}} & -k_T \end{pmatrix} \cdot \begin{pmatrix} n_{S1} \\ n_{T1} \end{pmatrix} \quad (6)$$

By solving the differential equation system 6 we obtain solutions for  $n(t)$ , which are linear combinations of  $\exp(-t/\tau_i)$ . By inserting  $n(t)$  into equation 7, the species associated difference

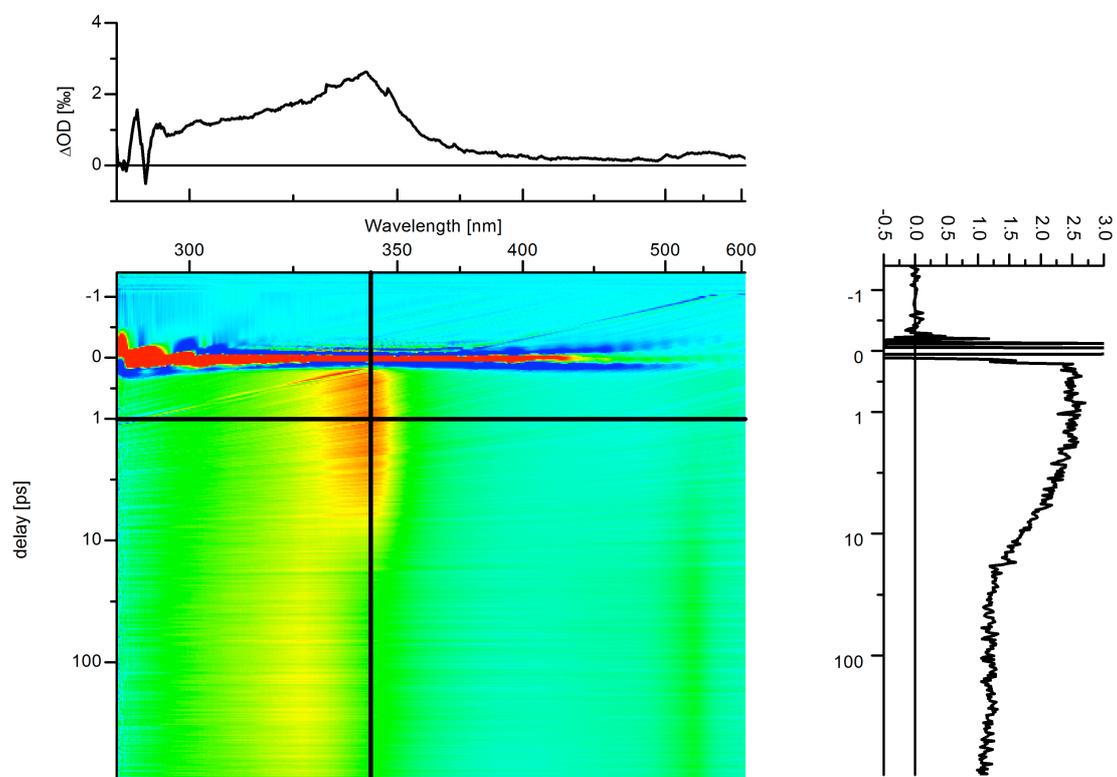
spectra  $\sigma_i(\lambda)$  can be extracted from the DADS.

$$\Delta OD(\lambda, t) = n_{S1}(t) \cdot \sigma_{S1}(\lambda) + n_{T1}(t) \cdot \sigma_{T1}(\lambda) \quad (7)$$

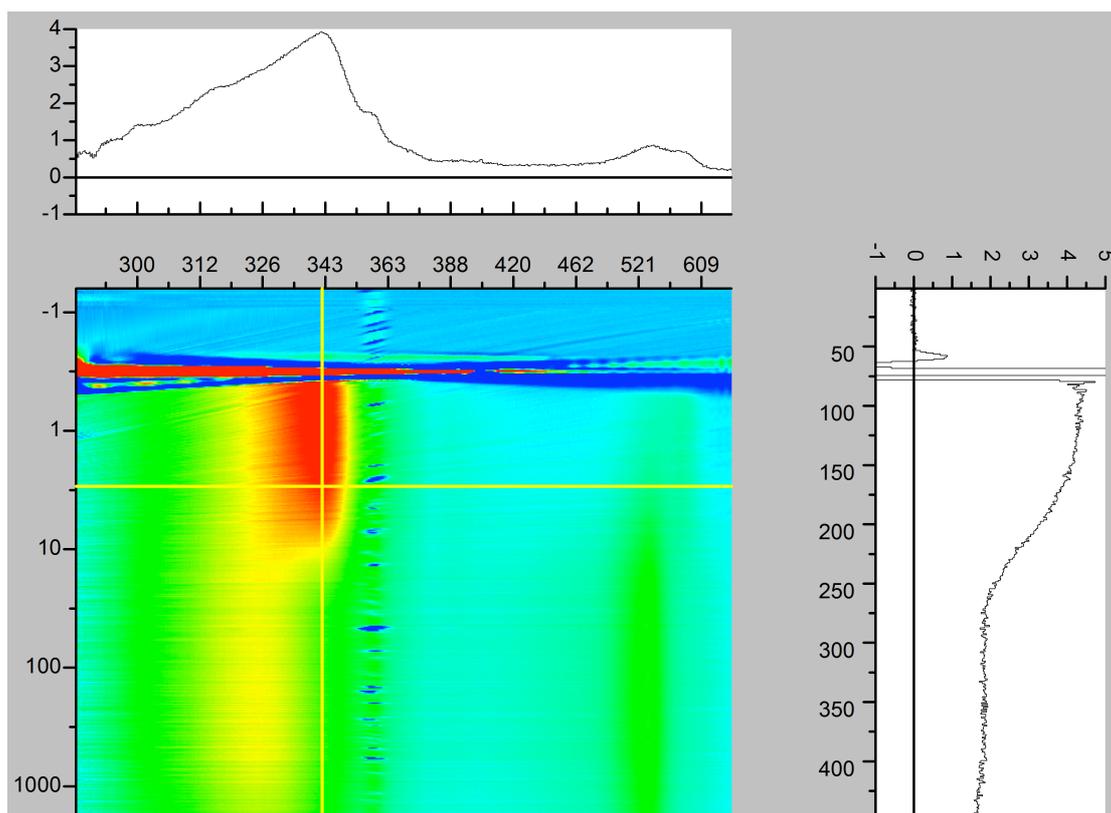
Instead of analyzing the temporal behavior just for one selected wavelength we also analyzed the temporal behavior of the integrated transient absorption signal of a selected wavelength region. This enhances the signal to noise ratio and minimizes the influence of spectral diffusion caused for instance by solvation or vibrational cooling. Within the selected wavelength region we added the transient signal of the selected data points in spectral direction for every point of the time scale. Then we used a standard fit routine to describe the temporal behavior of the integrated transient absorption by a sum of exponential functions.



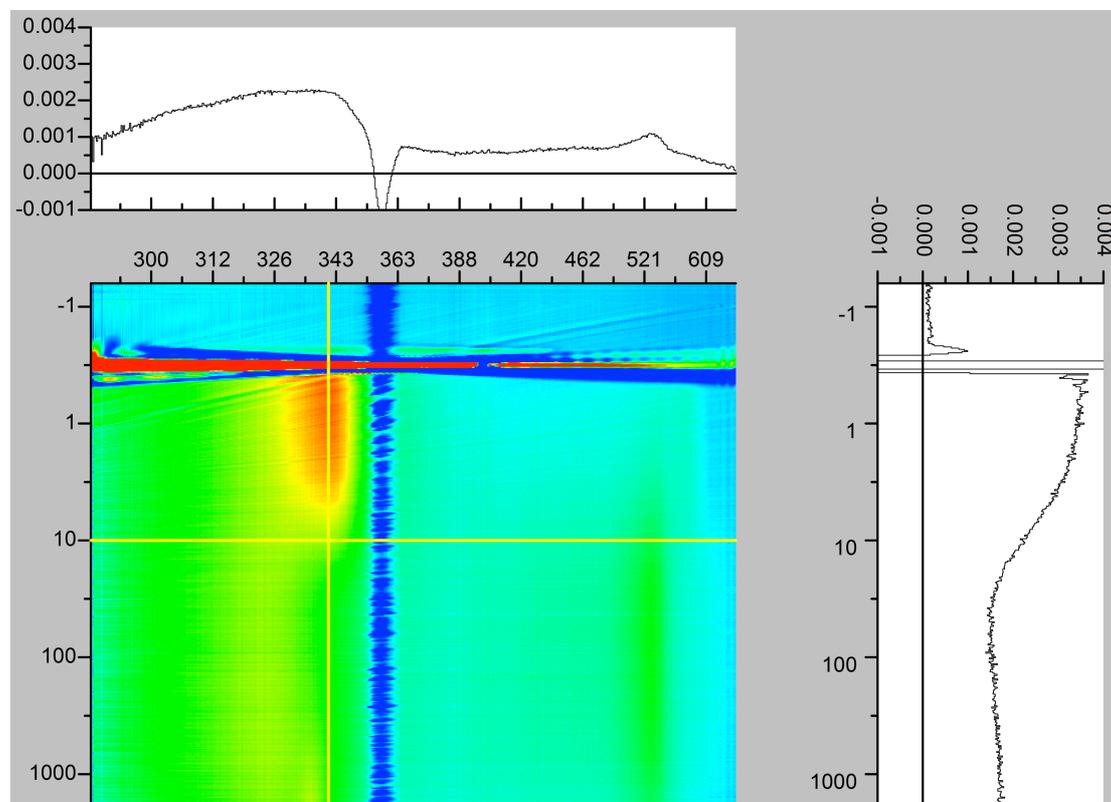
**Figure SI-9** Transient absorption of **BP** in MeOH



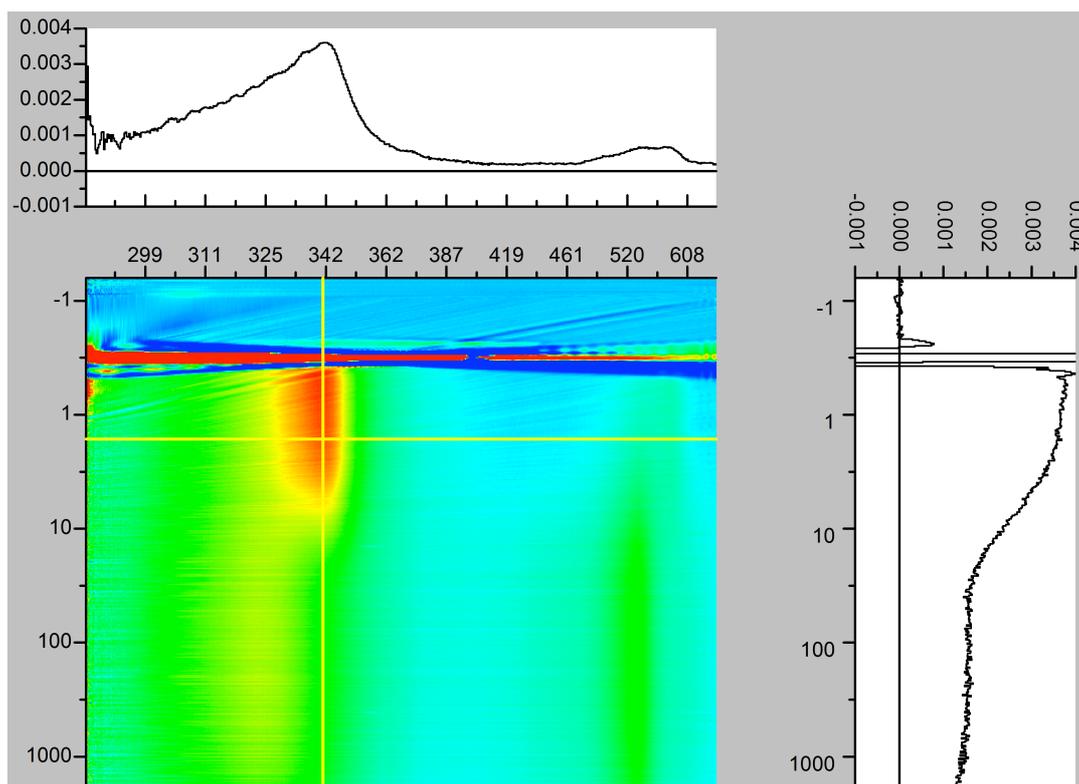
**Figure SI-10** Transient absorption of **1** in MeOH



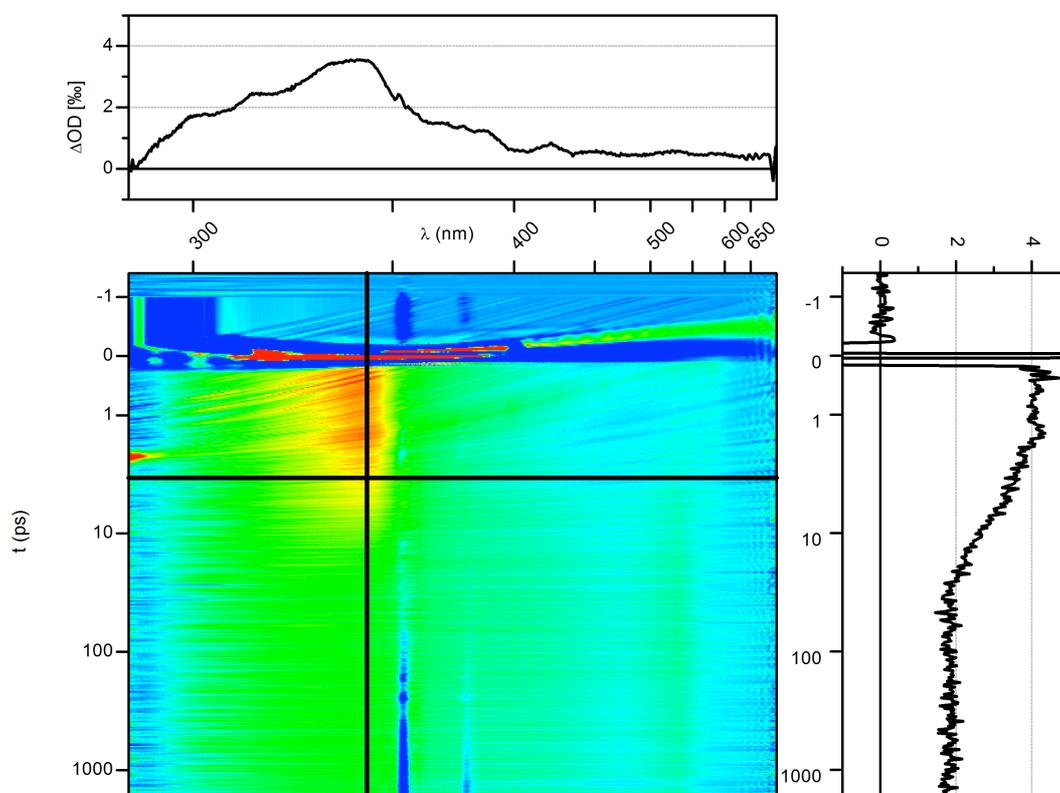
**Figure SI-11** Transient absorption of 1A in MeOH



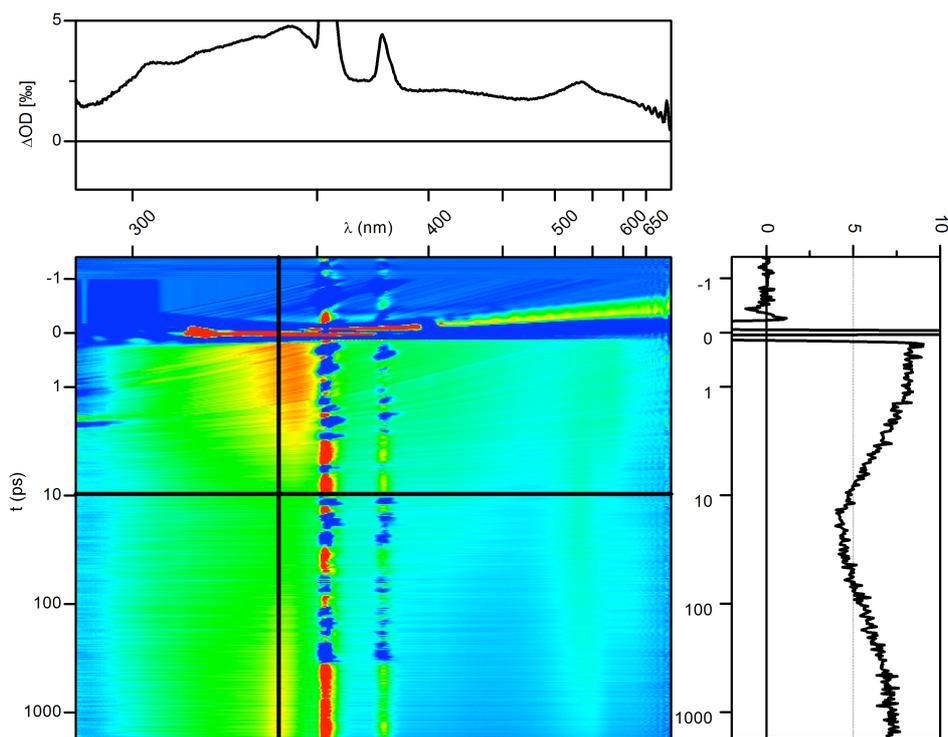
**Figure SI-12** Transient absorption of 1G in MeOH



**Figure SI-13** Transient absorption of **1C** in MeOH



**Figure SI-14** Transient absorption of **1C** in H<sub>2</sub>O



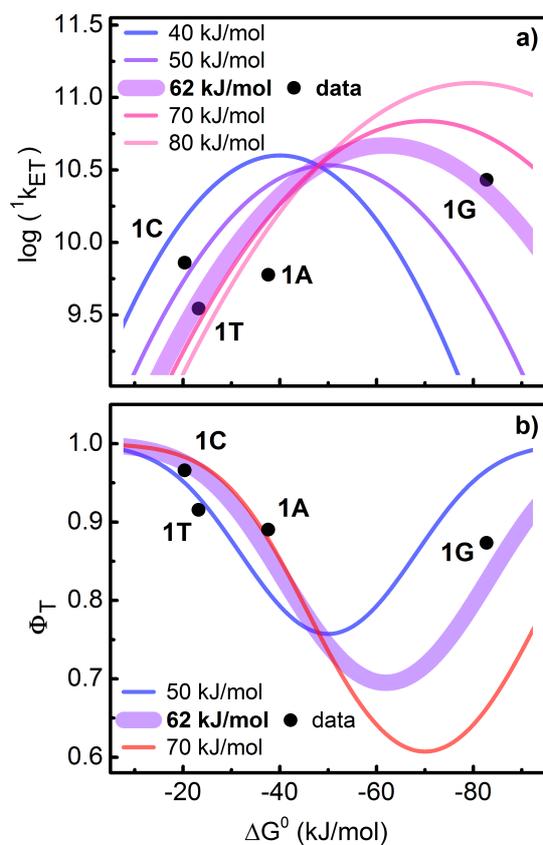
**Figure SI-15** Transient absorption of **1G** in H<sub>2</sub>O

### 3. Modelling of the ET dynamics with Marcus Theory

We apply Marcus theory to discuss the dependence of the ET rate on the standard oxidation potential  $E^0$  of the nucleobase. In Marcus theory the ET rate  $k_{\text{ET}}$  depends on the change in Gibbs free energy  $\Delta G^0$  according to

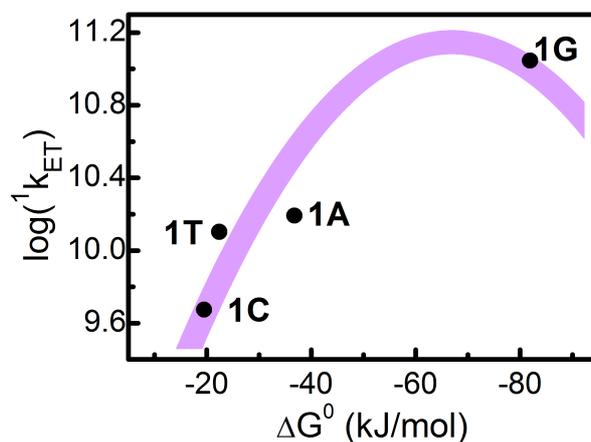
$$k_{\text{ET}} = \frac{2\pi}{\hbar} |H_{\text{AB}}|^2 \frac{1}{\sqrt{2\pi\lambda k_{\text{B}}T}} \exp\left(\frac{1}{k_{\text{B}}T} \cdot \frac{(\lambda + \Delta G^0)^2}{4\lambda}\right) \quad (\text{SI-2})$$

Formally we can fit the two unknown parameters  $H_{\text{AB}}$  and  $\lambda$  to the four measured values of  $^1k_{\text{ET}}$ . We therefore fitted  $H_{\text{AB}}$  for selected values of  $\lambda$  in a meaningful range from 40 to 80 kJ/mol. Since the observed rate represents an average of the electron transfer dynamics in an ensemble of conformations of the dinucleotide, the effective value of  $H_{\text{AB}}$  obtained in the fit cannot yet be directly interpreted as electronic coupling matrix element. The resulting functional dependence is displayed in Fig. SI-16 a). For each curve an effective value of  $H_{\text{AB}}$  is associated to the chosen  $\lambda$ . A close inspection shows that only  $\lambda$  values around 62 kJ/mol allow a good reproduction of the experimental observations. The resulting fit values are  $\lambda = 62$  kJ/mol (0.64 eV) and  $H_{\text{AB}} = 1.49$  meV (12  $\text{cm}^{-1}$ ).



**Figure SI-16** a) Logarithm of  $^1k_{ET}$  with fit curves according to eq. versus the change of free energy  $\Delta G^0$  in MeOH. b) Benzophenone triplet yield in MeOH determined by TA spectroscopy (dots) and calculated yield using eq. 3 (lines).

We also applied the Marcus equation for an evaluation of the dependence of  $^1k_{\text{ET}}$  on  $\Delta G^0$  in  $\text{H}_2\text{O}$ . The reorganization energy of  $\lambda = 67 \text{ kJ/mol}$  ( $0.70 \text{ eV}$ ) is close to the value found for  $\text{MeOH}$  while the fitted electronic coupling strength is  $H_{\text{AB}} = 2.64 \text{ meV}$  ( $21 \text{ cm}^{-1}$ ). Again, the parameter  $H_{\text{AB}}$  is just an effective value and will be later discussed in connection with the distribution of donor acceptor distances.



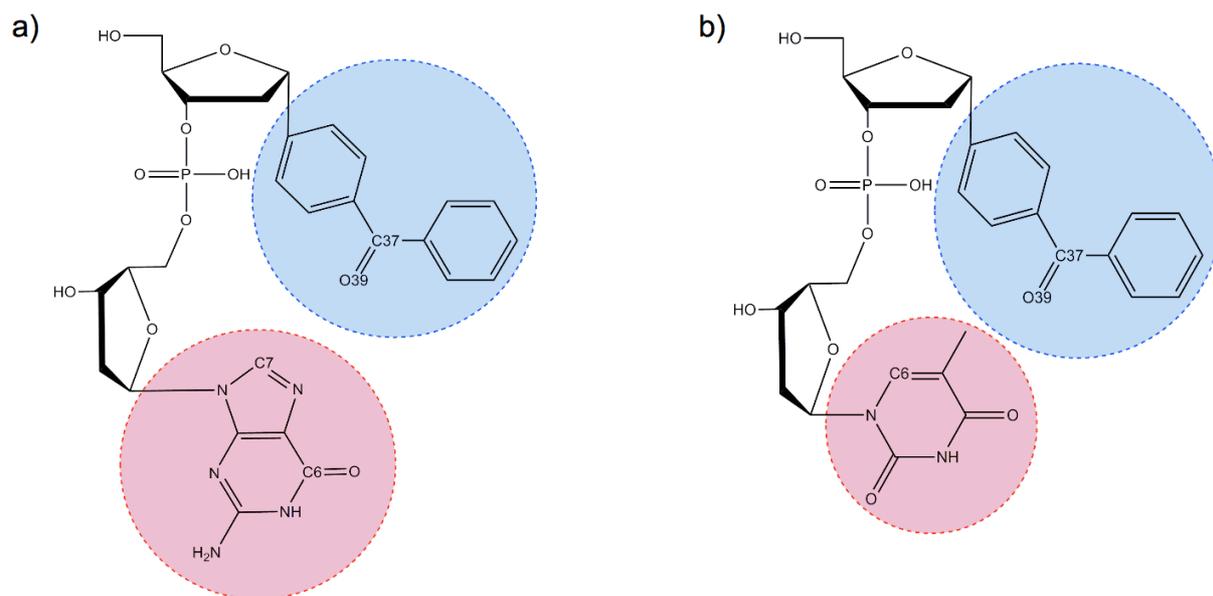
**Figure SI-17** Logarithm of  $^1k_{\text{ET}}$  in  $\text{H}_2\text{O}$  with fit curve according to eq. 1 with parameters  $\lambda = 67 \text{ kJ/mol}$  ( $0.70 \text{ eV}$ ) and  $H_{\text{AB}} = 2.64 \text{ meV}$  ( $21 \text{ cm}^{-1}$ ) vs. change of free energy  $\Delta G^0$ .

## 4. Computational Work

The original dinucleotide system was neutralized with an additional H atom at the  $\text{PO}_4^-$  group for all QM calculations, for the MD simulation (so no cation was necessary to balance the total system to a charge of zero) and for the QM/MM calculations.

The CC2 calculations for the non-dissolved molecule were performed with standard auxiliary basis sets<sup>57</sup> and Density Fitting approximation<sup>58</sup>. Excited state properties were calculated using TD-DFT (B3LYP functional<sup>32,33</sup>), and TD-CC2<sup>30,34-36</sup> linear response.

The classical molecular mechanics (MM) simulations, carried out to generate adequate starting structures for subsequent QM/MM optimizations in different solvent environments, utilize the Amber force field<sup>59</sup>. Appropriate force field parameters for **1G** and **1T** were created on the basis of the GAFF parameters<sup>60</sup> and types (see Fig. SI-23 and SI-24, Tab. SI-3, SI-4 and Tab. SI-5), which were assigned with Amber tools. Moreover, the types and parameters of the phosphate group had to be modified for the case of an uncharged molecule. The charges were obtained from a restrained electrostatic potential (RESP) fit on top of a HF gas phase calculation in the 6-31G\* basis<sup>61</sup> at the folded gas phase geometry. The parameters for the two solvents investigated in this work, i.e. water and methanol, are available in the Amber force field. For solvation of the unfolded and the folded geometries, rectangular boxes of size  $54 \times 48 \times 48 \text{ \AA}^3$  and  $54 \times 58 \times 55 \text{ \AA}^3$ , were used. This resulted in 3146 and 4583 TIP3P water molecules and 2009 and 1408 MeOH molecules, respectively. The initial solvated gas phase geometries were first minimized at the level of the force field. Afterwards, the resulting **1G** geometries were restrained with a weak restraint (10 kcal/mol), and a 20 ps constant volume MD simulation was carried out, increasing the temperature from 20 to 300 K. Thereafter a unrestrained constant pressure MD for 100 ps was performed to equilibrate the system. This was followed by a 10 ns MD production run at constant pressure (1 atm) and temperature (300 K). From this production run individual geometries were then selected for a subsequent QM/MM treatment. The QM region only comprises the benzophenone (BP) and the guanidine (G) subunits, the linking part of the molecule (like sugar and phosphate groups) is treated at the MM level (cf. discussion in section “Theoretical elucidation of conformations and molecular dynamics in water and methanol”). The QM part is schematically depicted in Figure SI-18. During the QM/MM geometry optimizations all atoms (including the whole set of solvent molecules) were fully relaxed.



**Figure SI-18** Molecular structure of the **1G** dinucleotide (a) and **1T** dinucleotide (b). The benzophenone (BP) is highlighted in light blue and the guanine (G) and thymine (T) in light red. These parts of **1G** were also used as QM-region if QM/MM scheme was used.

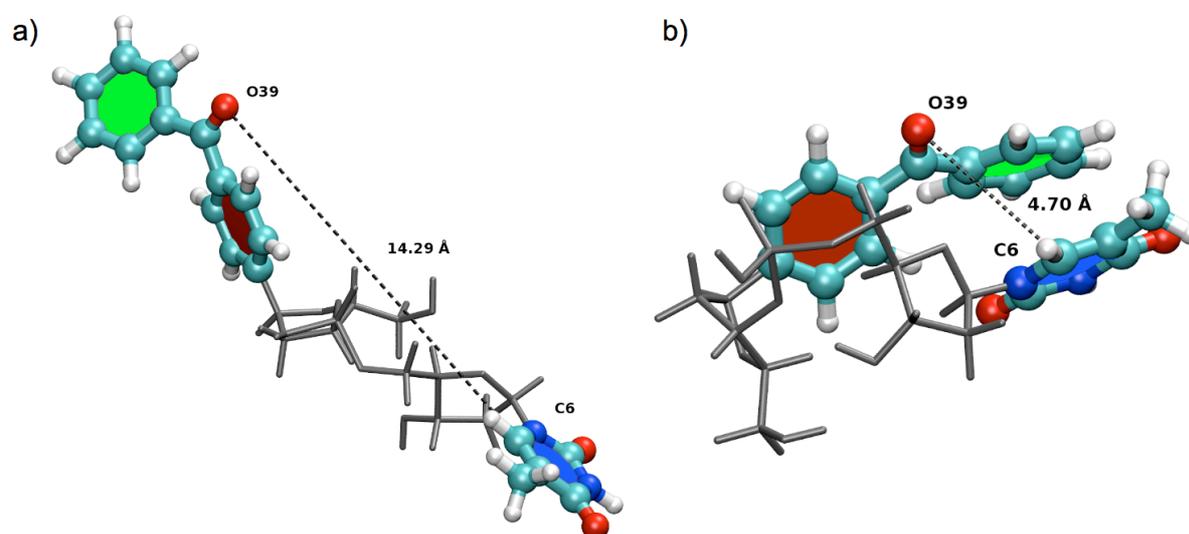
For the QM/MM calculations the ChemShell interface<sup>62</sup> together with TURBOMOLE<sup>63</sup> (to calculate energies and gradients of the QM part) was employed. For the MM part the DL-POLY molecular dynamics module<sup>64</sup> of ChemShell was used.

The DL-FIND<sup>65</sup> optimizer, also integrated into ChemShell, was driving the QM/MM geometry optimisations. The electrostatic interaction of the MM with the QM part was described by including all MM point charges in the one electron part of the QM Hamiltonian. For the separation of the MM from the QM part the link atom approach<sup>66</sup> and the charge shift scheme were employed.

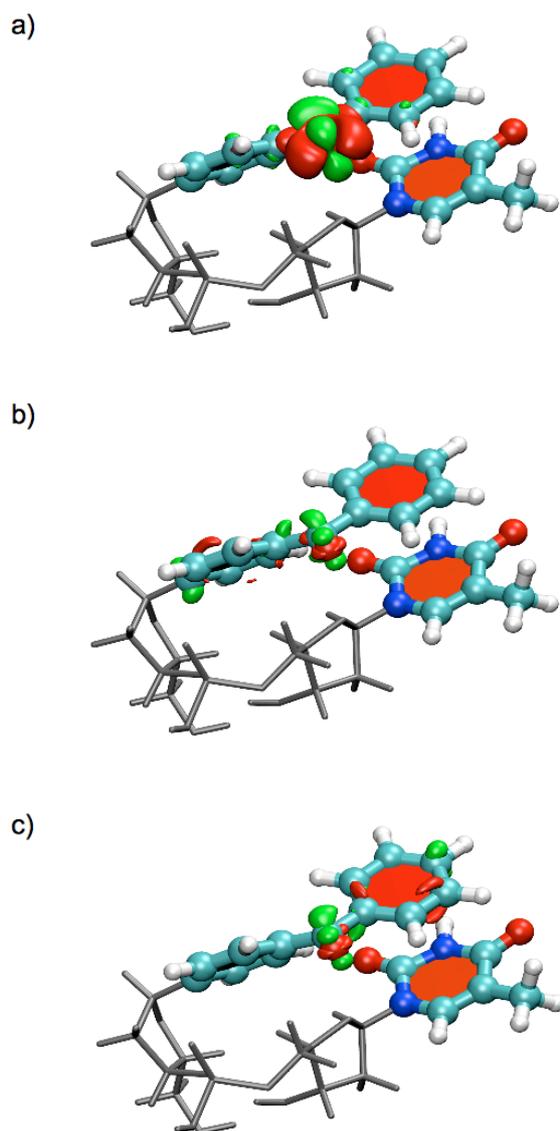
For the QM/MM study all individual MD snapshots were first minimized in the electronic ground state using B3LYP-D/def-SVP (DFT, B3LYP functional with dispersion correction, def-SVP AO basis). At the resulting geometries excitation energies and first-order properties (dipole moments, density differences) were calculated with the TD-CC2/aug-cc-pVDZ method (TD-CC2 response in the aug-cc-pVDZ AO basis). Next, the geometry of the  $S_1$  state was optimized. The calculations on the isolated molecule indicate that for folded geometries a low-lying CT state might be present. In fact, TD-DFT with the B3LYP functional artificially places this CT state always below all other excited states. In order to avoid contamination of the  $S_1$  state by the artificially low-lying CT state another hybrid functional with a larger

fraction of nonlocal exchange than B3LYP, the BHLYP functional<sup>67</sup> was employed instead for the geometry optimizations on the  $S_1$  and CT state surface. This leads to a better separation of the  $S_1$  ( $n\pi^*$ ) and CT states (for the B3LYP functional the CT state is always the lowest and spoils the ( $n\pi^*$ ) state by admixing, such that a proper geometry optimization of the ( $n\pi^*$ ) state is not possible).

These calculations were performed in the def-SVP AO basis. At the optimized  $S_1$  state geometries again excitation energies and first-order properties were calculated with TD-CC2/aug-cc-pVDZ. Finally, also geometry optimizations on the surface of the CT state were carried out (TD-DFT, BHLYP/def-SVP), when present, with subsequent excitation energy and first-order property calculations using TD-CC2/aug-cc-pVDZ.



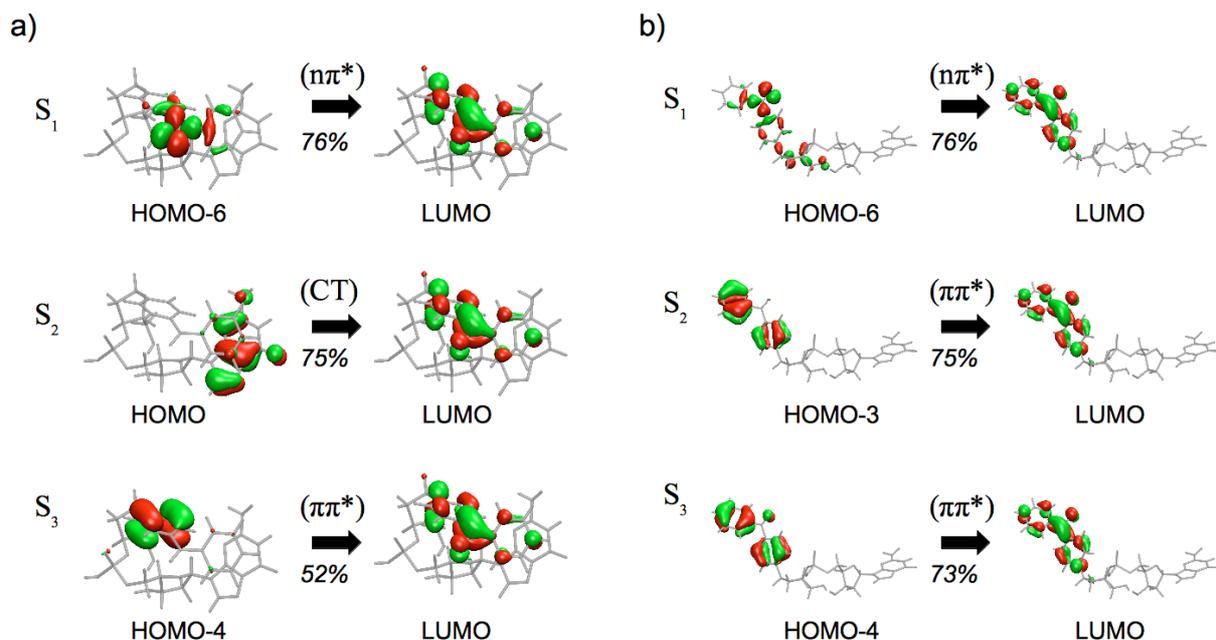
**Figure SI-19** Optimized structures of 1T in the gas phase. The five- and six-membered rings are colored in red/green and blue for better recognition. Also the distance  $d(\text{C6-O39})$ . (a) unfolded structure computed with DFT/BP in the def-SVP basis.  $d(\text{C6-O39})=15.59 \text{ \AA}$  and  $d(\text{C6-C3'})=16.09 \text{ \AA}$  (b) folded structure computed with CC2/def-SVP;  $d(\text{C6-O39})=6.92 \text{ \AA}$  and  $d(\text{C6-C37})=6.10 \text{ \AA}$



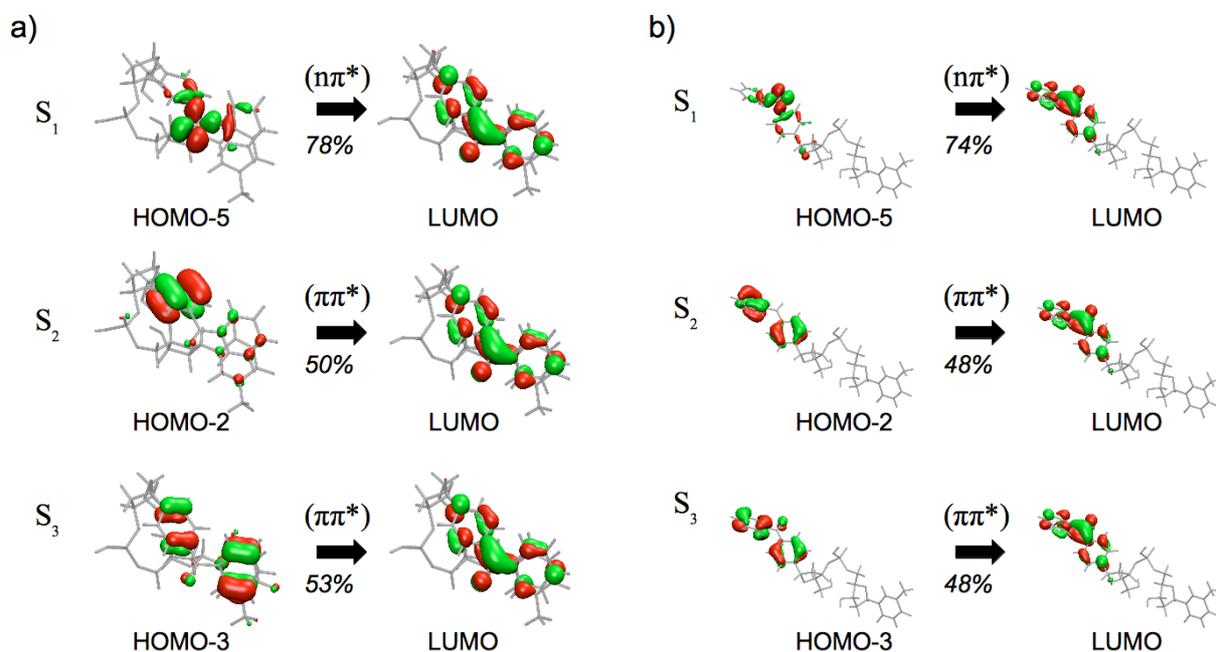
**Figure SI-20** Electron density difference plots for the first three excited states of the folded geometry of **1T**, calculated with TD-CC2 in def-SVP basis. Isosurfaces are plotted for  $\pm 0.004$  a.u., red refers to a decrease, green to an increase in the density upon excitation. (a)  $S_1$ : ( $n\pi^*$ ) state;  $|\Delta\mu| = 1.31$  D (b)  $S_2$ : ( $\pi\pi^*$ ) type state;  $|\Delta\mu| = 1.15$  D (c)  $S_3$ : ( $\pi\pi^*$ ) type state;  $|\Delta\mu| = 1.92$  D

**Table SI-2** Results for excitation energies of the folded and unfolded **1G** molecule at the optimized ground state for comparison; TD-B3LYP/def-SVP, TD-CC2/def-SVP and TD-CC2/aug-cc-pVDZ;

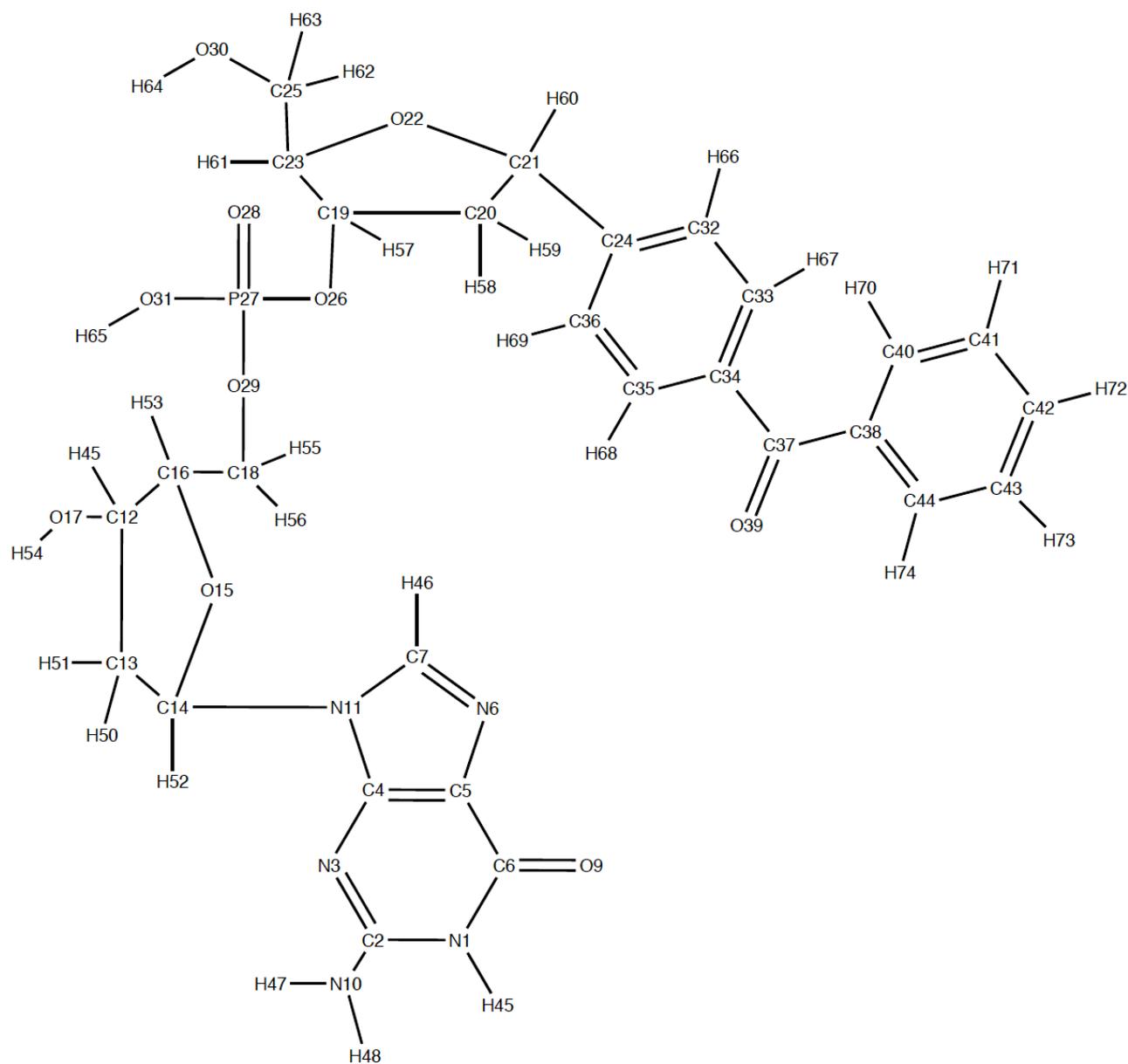
method	folded geometry					unfolded geometry				
	singlet			triplet		singlet			triplet	
	type	state	$\omega$ [eV]	state	$\omega$ [eV]	type	state	$\omega$ [eV]	state	$\omega$ [eV]
TD-B3LYP	(n $\pi^*$ )	S <sub>1</sub>	3.36	T <sub>1</sub>	2.92	(n $\pi^*$ )	S <sub>1</sub>	3.50	T <sub>1</sub>	2.85
	CT	S <sub>2</sub>	3.54	T <sub>2</sub>	3.23	CT	S <sub>2</sub>	3.56	T <sub>2</sub>	3.29
	( $\pi\pi^*$ )	S <sub>3</sub>	4.42	T <sub>3</sub>	3.28	( $\pi\pi^*$ )	S <sub>3</sub>	4.45	T <sub>3</sub>	3.37
TD-CC2	(n $\pi^*$ )	S <sub>1</sub>	3.74	T <sub>1</sub>	3.37	(n $\pi^*$ )	S <sub>1</sub>	3.82	T <sub>1</sub>	3.31
def-SVP	CT	S <sub>2</sub>	4.59	T <sub>2</sub>	3.88	( $\pi\pi^*$ )	S <sub>2</sub>	4.88	T <sub>2</sub>	3.92
	( $\pi\pi^*$ )	S <sub>3</sub>	4.78	T <sub>3</sub>	4.00	( $\pi\pi^*$ )	S <sub>3</sub>	4.96	T <sub>3</sub>	4.00
TD-CC2	(n $\pi^*$ )	S <sub>1</sub>	3.63	T <sub>1</sub>	3.30	(n $\pi^*$ )	S <sub>1</sub>	3.63	T <sub>1</sub>	3.27
a-VDZ	CT	S <sub>2</sub>	4.33	T <sub>2</sub>	3.83	( $\pi\pi^*$ )	S <sub>2</sub>	4.69	T <sub>2</sub>	3.87
	( $\pi\pi^*$ )	S <sub>3</sub>	4.66	T <sub>3</sub>	3.93	( $\pi\pi^*$ )	S <sub>3</sub>	4.78	T <sub>3</sub>	3.96



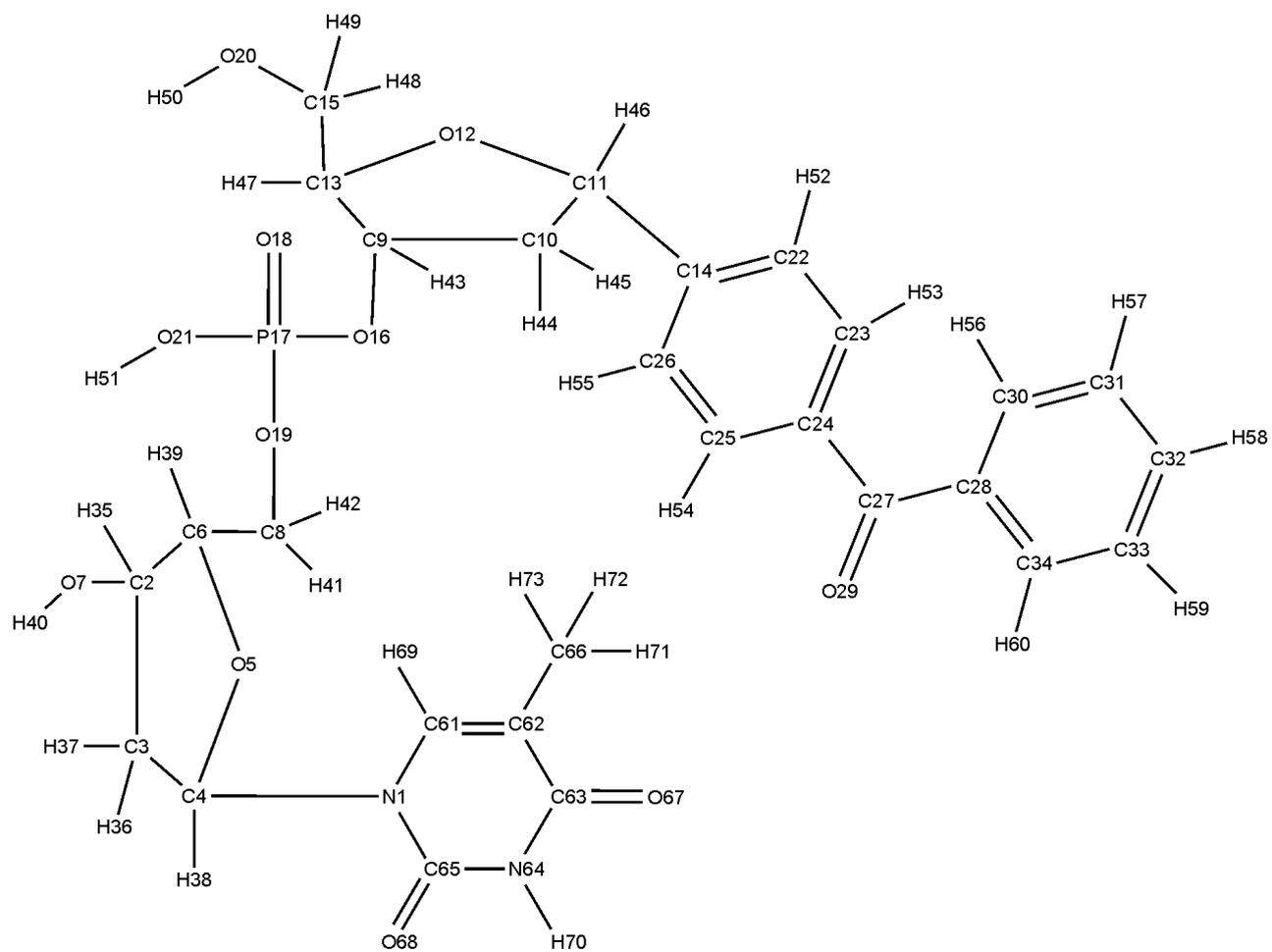
**Figure SI-21** Relevant molecular orbitals (HF/def-SVP) involved in the TD-CC2 excitations of folded (a) and unfolded (b) 1G calculated at the ground state geometry. Negative and positive values are depicted in red and green (isovalue 0.04); percentage specifies the weight in the singles vector.



**Figure SI-22** Relevant molecular orbitals (HF/def-SVP) involved in the TD-CC2 excitations of folded (a) and unfolded (b) 1T calculated at the ground state geometry. Negative and positive values are depicted in red and green (isovalue 0.04); percentage specifies the weight in the singles vector.



**Figure SI-23** Atom names and numbers for **1G**.



**Figure SI-24** Atom names and numbers for **1T**.

**Table SI-3** Types and charges for **1G** applied in the AMBER force field (see also Figure SI 17).

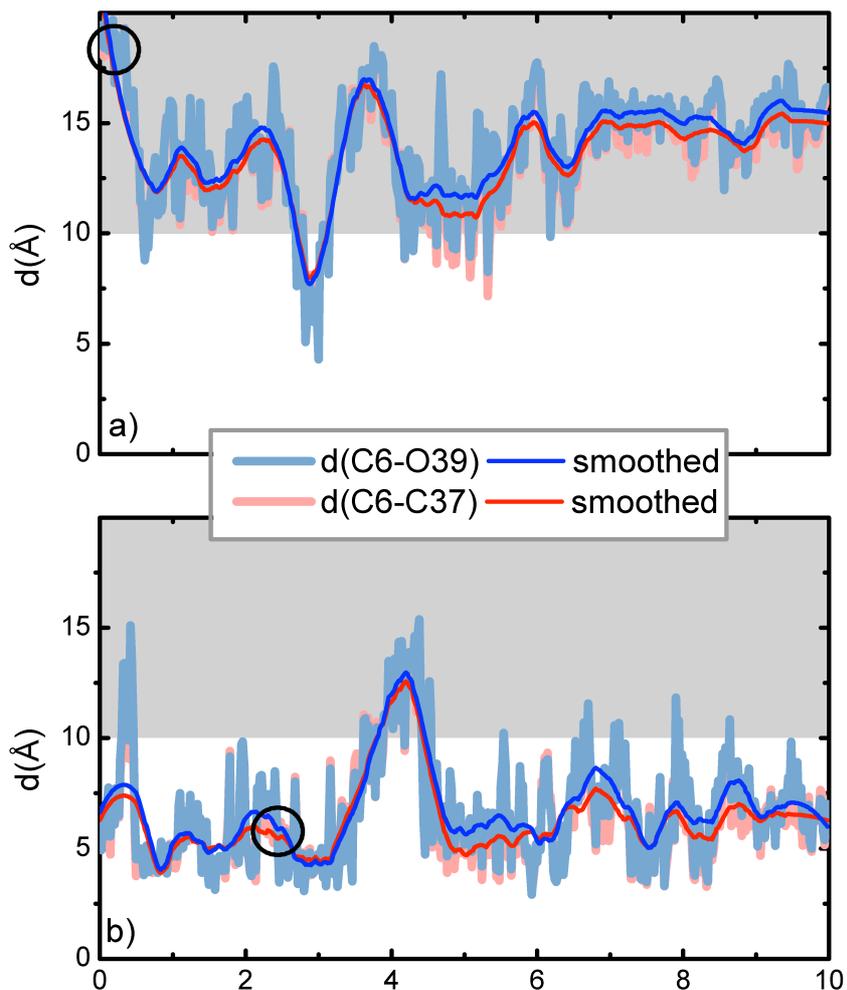
atomnumber	type	charge	atomnumber	type	charge
N1	NA	-0.666700	C38	ca	-0.008430
C2	CA	0.767410	O39	o	-0.534260
N3	NC	-0.688020	C40	ca	-0.054700
C4	CB	0.300820	C41	ca	-0.209040
C5	CB	0.282460	C42	ca	-0.124470
C6	C	0.528430	C43	ca	-0.166420
C7	CK	0.056260	C44	ca	-0.147130
N8	NB	-0.587410	H45	H	0.417700
O9	O	-0.571560	H46	H5	0.193450
N10	N2	-0.933060	H47	H	0.432150
N11	N*	-0.126000	H48	H	0.411720
C12	CT	0.269570	H49	H1	0.059050
C13	CT	-0.189890	H50	HC	0.098050
C14	CT	0.268530	H51	HC	0.082720
O15	OS	-0.500570	H52	H2	0.119340
C16	CT	0.263830	H53	H1	0.052470
O17	OH	-0.698290	H54	HO	0.402920
C18	CT	-0.128090	H55	H1	0.138760
C19	CT	0.209570	H56	H1	0.180610
C20	CT	-0.503530	H57	H1	0.137590
C21	CT	0.246090	H58	HC	0.166380
O22	OS	-0.533130	H59	HC	0.194410
C23	CT	0.361430	H60	H1	0.103950
C24	ca	-0.010880	H61	H1	0.094870
C25	CT	-0.186370	H62	H1	0.165940
O26	OS	-0.487470	H63	H1	0.097250
P27	P	1.227500	H64	HO	0.454020
O28	O2	-0.670720	H65	HO	0.488570
O29	OS	-0.474240	H66	ha	0.165100
O30	OH	-0.597390	H67	ha	0.109760
O31	OH	-0.700360	H68	ha	0.199300
C32	ca	-0.245970	H69	ha	0.150070
C33	ca	-0.074890	H70	ha	0.092850
C34	ca	0.035560	H71	ha	0.157660
C35	ca	-0.269360	H72	ha	0.156080
C36	ca	-0.059090	H73	ha	0.153900
C37	c	0.485100	H74	ha	0.168070

**Table SI-4** Types and charges for **1G** applied in the AMBER force field (see also Figure SI-?).

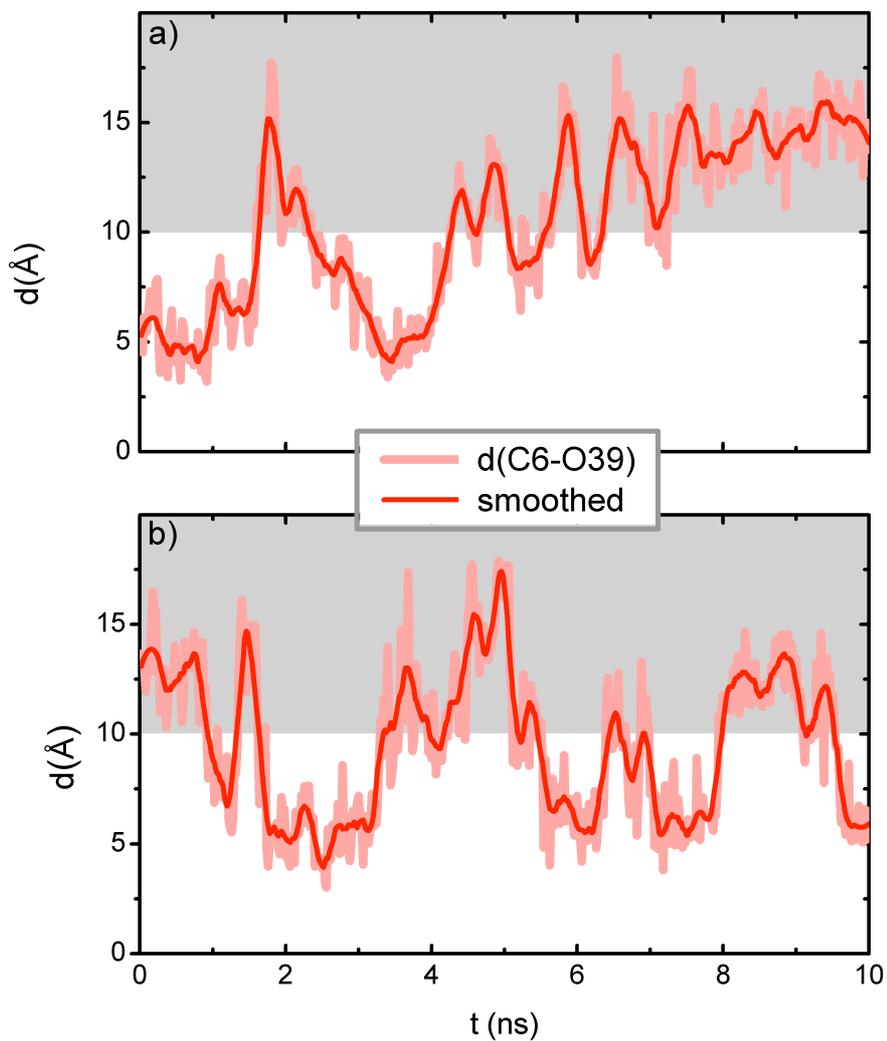
atomnumber	type	charge	atomnumber	type	charge
N1	N*	0.002450	H38	H2	0.145830
C2	CT	0.326270	H39	H1	0.111940
C3	CT	-0.338670	H40	HO	0.308050
C4	CT	0.108930	H41	H1	0.133860
O5	OS	-0.385510	H42	H1	0.116920
C6	CT	0.083870	H43	H1	0.125420
O7	OH	-0.662360	H44	HC	0.183280
C8	CT	0.036800	H45	HC	0.188490
C9	CT	0.136080	H46	H1	0.094000
C10	CT	-0.578800	H47	H1	0.055840
C11	CT	0.384290	H48	H1	0.105040
O12	OS	-0.607400	H49	H1	0.069310
C13	CT	0.310120	H50	HO	0.432560
C14	ca	0.051910	H51	HO	0.491330
C15	CT	0.078980	H52	ha	0.169100
O16	OS	-0.203580	H53	ha	0.174770
P17	P	1.138750	H54	ha	0.188690
O18	O2	-0.676280	H55	ha	0.234640
O19	OS	-0.457090	H56	ha	0.149430
O20	OH	-0.700340	H57	ha	0.168190
O21	OH	-0.706290	H58	ha	0.166640
C22	ca	-0.208800	H59	ha	0.175000
C23	ca	-0.127010	H60	ha	0.176290
C24	ca	-0.037850	C61	CM	-0.365710
C25	ca	-0.072930	C62	CM	0.206630
C26	ca	-0.382600	C63	C	0.579150
C27	c	0.359620	N64	NA	-0.599810
C28	ca	0.199670	C65	C	0.624020
O29	o	-0.555830	C66	CT	-0.582060
C30	ca	-0.169250	O67	O	-0.579470
C31	ca	-0.218140	O68	O	-0.608660
C32	ca	-0.118180	H69	H4	0.289330
C33	ca	-0.192340	H70	H	0.408470
C34	ca	-0.190300	H71	HC	0.171900
H35	H1	0.058270	H72	HC	0.174100
H36	HC	0.168230	H73	HC	0.155230
H37	HC	0.107570			

**Table SI-5** Modifications applied to the standard GAFF parameter of AMBER force field.

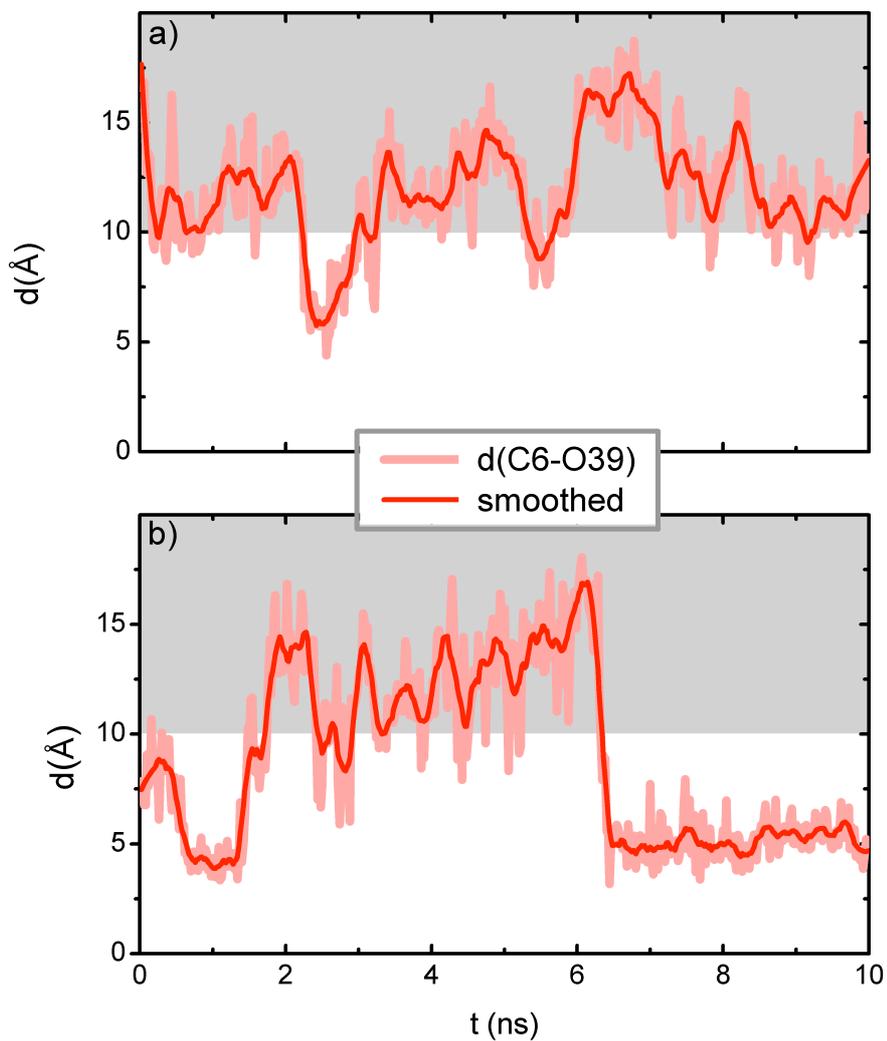
MASS				
c	12.01	0.616		
ca	12.01	0.360		
ha	1.008	0.135		
o	16.00	0.434		
BOND				
CT-ca	323.5	1.513		
ca-ca	478.4	1.387		
ca-ha	344.3	1.087		
c -ca	349.7	1.487		
c -o	648.0	1.214		
ANGL				
CT-CT-ca	62.5	114.61		
CT-ca-ca	63.8	120.63		
ca-CT-OS	67.7	110.51		
ca-CT-H1	46.8	110.95		
ca-ca-ca	67.2	119.97		
ca-ca-ha	48.5	120.01		
c -ca-ca	64.6	120.14		
ca-c -ca	62.8	119.53		
ca-c -o	68.7	123.44		
O2-P -OH	150.0	108.23		
O2-P -OS	250.0	108.23		
OS-P -OS	250.000	102.46		
HO-OH-P	150.000	110.14		
P-OS-CT	250.000	118.00		
DIHEDRAL				
CT-CT-ca-ca	6	0.000	0.000	2.000
ca-ca-ca-CT	4	14.500	180.000	2.000
CT-ca-ca-ha	4	14.500	180.000	2.000
OS-CT-ca-ca	6	0.000	0.000	2.000
ca-ca-ca-ca	4	14.500	180.000	2.000
ca-ca-ca-ha	4	14.500	180.000	2.000
H1-CT-ca-ca	6	0.000	0.000	2.000
ca-ca-ca-c	4	14.500	180.000	2.000
ca-c -ca-ca	4	14.500	180.000	2.000
o -c -ca-ca	4	14.500	180.000	2.000
c -ca-ca-ha	4	14.500	180.000	2.000
ha-ca-ca-ha	4	14.500	180.000	2.000
IMPROPER				
ca-ca-ca-ha	1.1	180.0	2.0	
c -ca-ca-ca	1.1	180.0	2.0	
ca-ca-c -o	10.5	180.0	2.0	
NONBON				
o	1.6612	0.2100		
ca	1.9080	0.0860		
c	1.9080	0.0860		
ha	1.4590	0.0150		



**Figure SI-25** Trajectories of MD run; (a) shows **1G** in methanol environment and (b) **1G** in water; the distances  $d(\text{C6-C37})$  and  $d(\text{C6-O39})$  in Å are plotted against the simulation time in ns. The grey area is the region, in which the molecule is unfolded. The cycles depict the chosen snapshot (a) MD of **1G** MeOH, unfolded (0.2 ns;  ${}^b\mathbf{1G}_{\text{MeOH}}$ ) starting geometry (b) MD of **1G** water, folded (2.4 ns;  ${}^b\mathbf{1G}_{\text{H}_2\text{O}}$ ) starting geometry



**Figure SI-26** Trajectories of MD run; (a) shows **1T** in methanol environment and folded starting structure and (b) **1T** in water and unfolded starting structure; the distances and  $d(\text{C6-O39})$  in Å are plotted against the simulation time in ns. The grey area is the region, in which the molecule is unfolded.

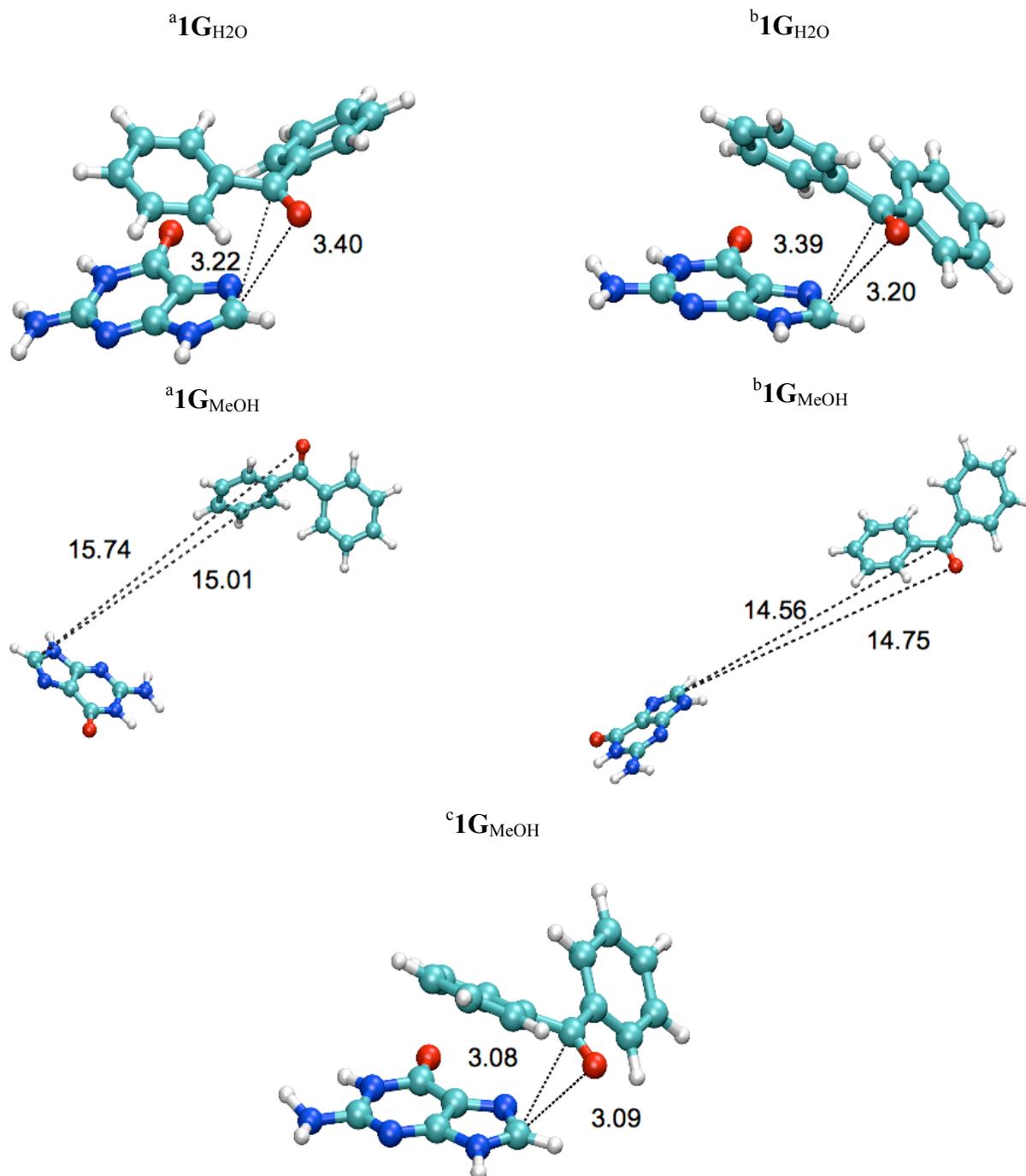


**Figure SI-27** Trajectories of MD run; (a) shows **1T** in methanol environment and unfolded starting structure and (b) **1T** in water and folded starting structure; the distances and  $d(\text{C6-O39})$  in Å are plotted against the simulation time in ns. The grey area is the region, in which the molecule is unfolded.

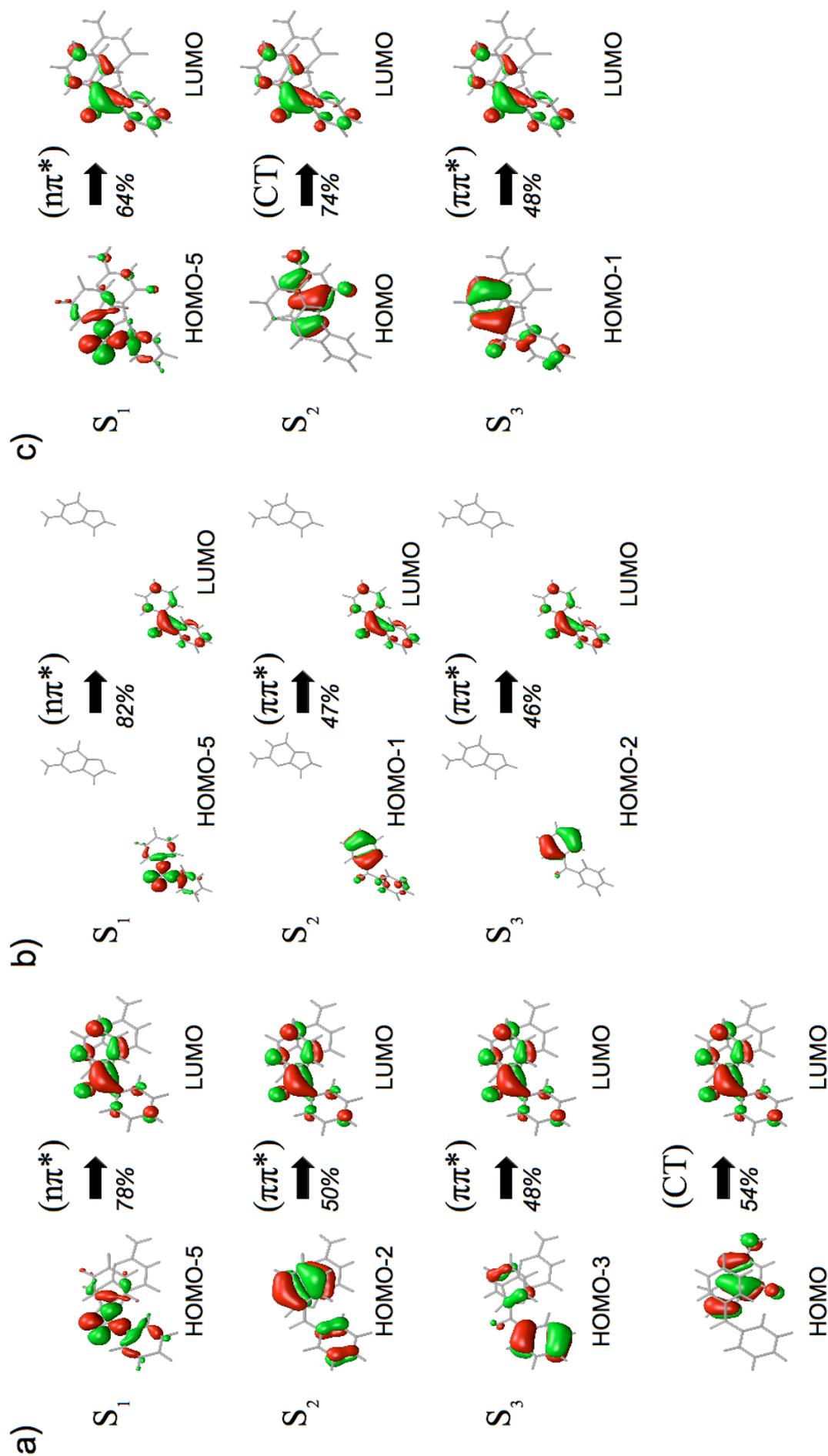
The trajectories of the MD simulation of **1T** in methanol (starting from the folded geometry), and of the MD simulation of **1T** in water (starting from the unfolded geometry), are displayed in Figure SI-22(a) and Figure SI-22(b). The trajectories of the converse starting geometries are displayed in Figure SI-23(a) and Figure SI-23(b), respectively. The degree of folding is measured by the distances  $d(\text{C6-O39})$  (cf. Figure SI-16(b)), which reflect the spatial separation of the BP and T subunits. For the isolated molecule these distances amount to about 5 Å for the folded, and 14 Å for the unfolded form (see Figure SI-17). As for the MD simulations for **1G** we consider geometries with distances above 10 Å as unfolded (grey shaded area), and geometries with distances below 10 Å as folded forms. Compared to the MD simulations of **1G**, the **1T** dinucleotide in water is much more flexible. One can see, that the  $\pi$ -stacking between the BP and T is not that pronounced, compared to **1G**. This is not that surprising due to the fact, that the purines are much more hydrophobic than the pyrimidines. In the less polar methanol solvent, BP and T undergo almost no  $\pi$ -stacking and can mostly be found as unfolded molecules during the simulation time.

**Table SI-6** TD-BHLYP QM/MM singlet excitation energies at the respective minima, calculated in def-SVP basis.

structure	state	FC-point			(n $\pi^*$ ) state			CT state		
		(optimised with B3LYP-D/def-SVP)			(optimised with BHLYP/def-SVP)			(optimised with BHLYP/def-SVP))		
		type	$\omega$ /eV	$\omega$ /nm	type	$\omega$ /eV	$\omega$ /nm	type	$\omega$ /eV	$\omega$ /nm
<sup>a</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	S <sub>1</sub>	(n $\pi^*$ )	4.25	291	(n $\pi^*$ )	3.45	359	CT	2.58	4.81
	S <sub>2</sub>	CT	4.87	254	CT	4.63	268	(n $\pi^*$ )	3.85	3.22
	S <sub>3</sub>	( $\pi\pi^*$ )	5.00	247	( $\pi\pi^*$ )	4.91	253	( $\pi\pi^*$ )	4.66	266
<sup>b</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	S <sub>1</sub>	(n $\pi^*$ )	4.21	295	(n $\pi^*$ )	3.61	343	CT	2.32	534
	S <sub>2</sub>	CT	4.75	261	CT	4.62	267	(n $\pi^*$ )	3.48	323
	S <sub>3</sub>	( $\pi\pi^*$ )	4.94	251	( $\pi\pi^*$ )	4.82	257	( $\pi\pi^*$ )	4.28	290
<sup>a</sup> <b>1G</b> <sub>MeOH</sub>	S <sub>1</sub>	(n $\pi^*$ )	4.13	300	(n $\pi^*$ )	3.58	346			
	S <sub>2</sub>	( $\pi\pi^*$ )	5.05	245	( $\pi\pi^*$ )	4.49	249		not available	
	S <sub>3</sub>	( $\pi\pi^*$ )	5.23	237	( $\pi\pi^*$ )	5.00	248			
<sup>b</sup> <b>1G</b> <sub>MeOH</sub>	S <sub>1</sub>	(n $\pi^*$ )	4.21	295	(n $\pi^*$ )	3.62	342			
	S <sub>2</sub>	( $\pi\pi^*$ )	4.94	251	( $\pi\pi^*$ )	4.74	261		not available	
	S <sub>3</sub>	( $\pi\pi^*$ )	5.07	244	( $\pi\pi^*$ )	4.83	257			
<sup>c</sup> <b>1G</b> <sub>MeOH</sub>	S <sub>1</sub>	(n $\pi^*$ )	4.35	285				CT	2.14	580
	S <sub>2</sub>	CT	4.37	284		not available		(n $\pi^*$ )	3.92	316
	S <sub>3</sub>	( $\pi\pi^*$ )	4.98	249				( $\pi\pi^*$ )	4.13	300



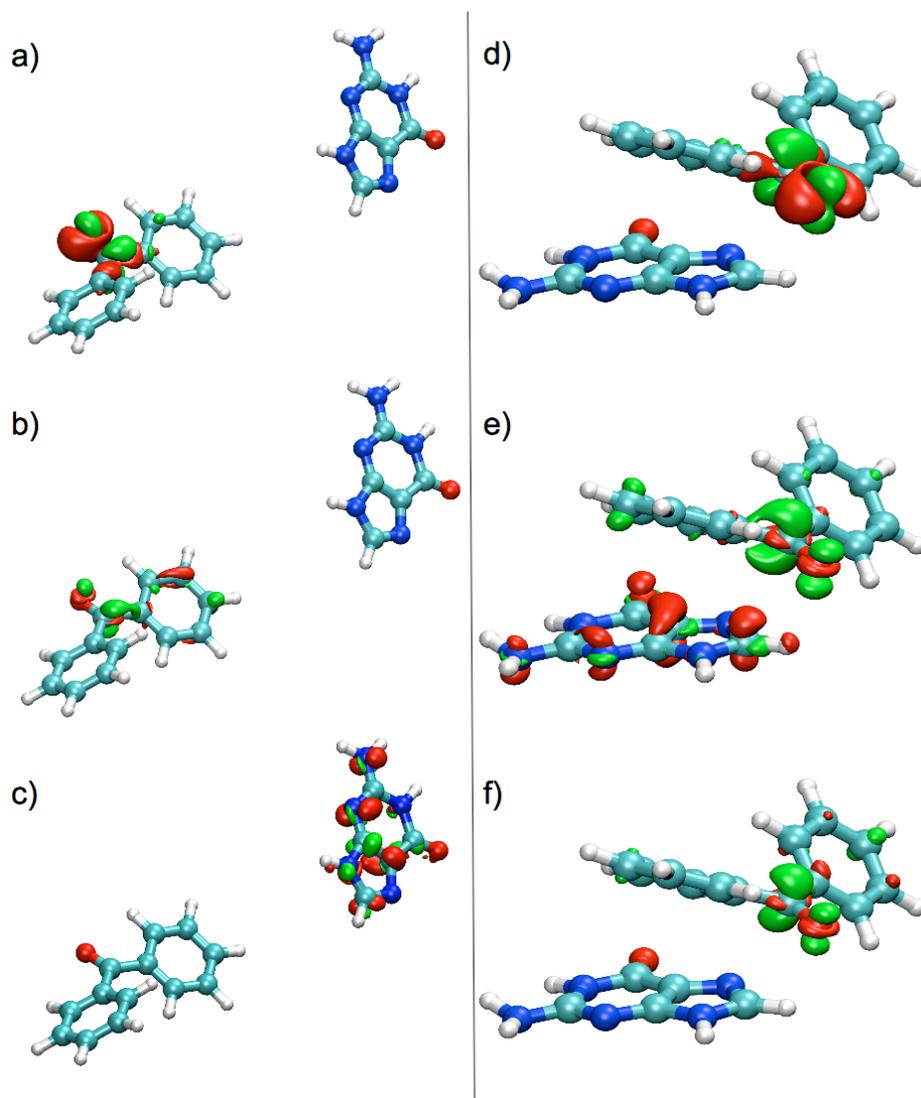
**Figure SI-28** Geometries of the chosen snapshots at the B3LYP-D/def-SVP optimized ground state with corresponding distances  $d(C7-O29)$  and  $d(C7-C37)$ ; only QM-region is shown.



**Figure SI-29** Relevant molecular orbitals (HF/def-SVP) involved in the TD-CC2 excitations of  ${}^1\mathbf{IG}_{\text{H}_2\text{O}}$  (a),  $(b)$   ${}^1\mathbf{IG}_{\text{MeOH}}$  and  $(c)$   ${}^1\mathbf{IG}_{\text{MeOH}}$  calculated at the ground state geometry. Negative and positive values are depicted in red and green (isovalue 0.04); percentage specifies the weight in the singles vector.

**Table SI-7** TD-CC2 QM/MM singlet and triplet excitation energies at the respective FC minima for comparison, calculated in def-SVP basis.

structure	type	state	E [eV]	state	E [eV]
	(nπ*)	S <sub>1</sub>	4.25	T <sub>1</sub>	3.61
<sup>a</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	(ππ*)	S <sub>2</sub>	4.87	T <sub>2</sub>	3.91
	(ππ*)	S <sub>3</sub>	5.00	T <sub>3</sub>	4.14
	(nπ*)	S <sub>1</sub>	4.21	T <sub>1</sub>	3.53
<sup>b</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	(ππ*)	S <sub>2</sub>	4.75	T <sub>2</sub>	3.92
	(ππ*)	S <sub>3</sub>	4.94	T <sub>3</sub>	4.07
	(nπ*)	S <sub>1</sub>	3.78	T <sub>1</sub>	3.50
<sup>a</sup> <b>1G</b> <sub>MeOH</sub>	(ππ*)	S <sub>2</sub>	4.70	T <sub>2</sub>	3.98
	(ππ*)	S <sub>3</sub>	4.82	T <sub>3</sub>	4.10
	(nπ*)	S <sub>1</sub>	3.84	T <sub>1</sub>	3.56
<sup>b</sup> <b>1G</b> <sub>MeOH</sub>	(ππ*)	S <sub>2</sub>	4.58	T <sub>2</sub>	3.92
	(ππ*)	S <sub>3</sub>	4.65	T <sub>3</sub>	4.12
	(nπ*)	S <sub>1</sub>	3.96	T <sub>1</sub>	3.66
<sup>c</sup> <b>1G</b> <sub>MeOH</sub>	CT	S <sub>2</sub>	4.28	T <sub>2</sub>	3.99
	(ππ*)	S <sub>3</sub>	4.61	T <sub>3</sub>	4.00



**Figure SI-30** Electron density difference plots for the first three excited singlet states of  ${}^a\mathbf{1G}_{\text{MeOH}}$ , (left side) and  ${}^c\mathbf{1G}_{\text{MeOH}}$  (right side) calculated with TD-CC2 in the aug-cc-pVDZ basis in the QM/MM framework, at the TD-DFT/B3LYP-D FC point. Isosurfaces are plotted for  $\pm 0.005$  a.u., red refers to a decrease, green to an increase in the density upon excitation; only QM-region is shown.

(a)  $S_1$ : ( $n\pi^*$ ) state;  $|\Delta\mu| = 1.45$  D

(b)  $S_2$ : ( $\pi\pi^*$ ) state;  $|\Delta\mu| = 2.94$  D

(c)  $S_3$ : ( $\pi\pi^*$ ) state;  $|\Delta\mu| = 3.36$  D

(d)  $S_1$ : ( $n\pi^*$ ) state;  $|\Delta\mu| = 1.31$  D

(e)  $S_2$ : CT state;  $|\Delta\mu| = 12.44$  D

(f)  $S_3$ : ( $\pi\pi^*$ ) state;  $|\Delta\mu| = 1.96$  D

**Table SI-8** Distances  $d(\text{C7-O39})$  and  $d(\text{C6-O39})$  in Å for all chosen snapshots at the respective QM/MM optimized structures.

structure	optimized state	$d(\text{C7-O39})$ in Å	$d(\text{C6-O39})$ in Å
<sup>a</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	FC-point (B3LYP-D/def-SVP)	3.22	5.34
	( $n\pi^*$ )-state (TD-BHLYP/def-SVP)	3.21	5.41
	CT-state (TD-BHLYP/def-SVP)	2.73	5.15
<sup>b</sup> <b>1G</b> <sub>H<sub>2</sub>O</sub>	FC-point (B3LYP-D/def-SVP)	3.39	6.03
	( $n\pi^*$ )-state (TD-BHLYP/def-SVP)	3.69	6.14
	CT-state (TD-BHLYP/def-SVP)	2.93	5.54
<sup>a</sup> <b>1G</b> <sub>MeOH</sub>	FC-point (B3LYP-D/def-SVP)	14.75	17.91
	( $n\pi^*$ )-state (TD-BHLYP/def-SVP)	14.77	17.86
	CT-state (TD-BHLYP/def-SVP)	-	-
<sup>b</sup> <b>1G</b> <sub>MeOH</sub>	FC-point (B3LYP-D/def-SVP)	15.74	15.29
	( $n\pi^*$ )-state (TD-BHLYP/def-SVP)	15.41	15.03
	CT-state (TD-BHLYP/def-SVP)	-	-
<sup>c</sup> <b>1G</b> <sub>MeOH</sub>	FC-point (B3LYP-D/def-SVP)	3.09	5.82
	( $n\pi^*$ )-state (TD-BHLYP/def-SVP)	-	-
	CT-state (TD-BHLYP/def-SVP)	2.85	5.75

#### 4. Additional references

56. P. Fita, E. Luzina, T. Dziembowska, C. Radzewicz and A. A. Grabowska, *J. Chem. Phys.*, 2006, **125**, 184508.
57. F. Weigend, M. Häser, H. Patzelt and R. Ahlrichs, *Chem. Phys. Lett.*, 1998, **294**, 143.
58. R. Ahlrichs, *Phys. Chem. Chem. Phys.*, 2004, **6**, 5119.
59. D. A. Case, T. A. Darden, T. E. Cheatham, III, C. L. Simmerling, J. Wang, R. E. Duke, R. Luo, M. Crowley, R. C. Walker, W. Zhang, K. M. Merz, B. Wang, S. Hayik, A. Roitberg, G. Seabra, I. Koloss, K. F. Wong, F. Paesani, J. Vanicek, X. Wu, S. R. Brozell, T. Steinbrecher, H. Gohlke, L. Yang, C. Tan, J. Mongan, V. Hornak, G. Cui, D. H. Mathews, M. G. Seetin, C. Sagui, V. Babin and P.A. Kollman, *AMBER 10*, 2008, University of California, San Francisco.
60. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
61. C. I. Bayly, P. Cieplak, W. Cornell and P. A. Kollman, *J. Phys. Chem.*, 1993, **97**, 10269.
62. P. Sherwood, *J. Mol. Struct. (Theochem)*, 2003, **632**, 1.
63. R. Ahlrichs, M. Bär, M. Häser, H. Horn and C. Kölmel, *Chem. Phys. Lett.*, 1989, **162**, 165.
64. W. Smith, C. W. Yong and P. M. Rodger, *Mol. Simul.*, 2002, **28**, 385.
65. J. Kästner, J. M. Carr, T. W. Keal, W. Thiel, A. Wander and P. Sherwood, *J. Phys. Chem. A.*, 2009, **113**, 11856.
66. A. H. de Vries, P. Sherwood, S. J. Collins, A. M. Rigby, M. Rigutto and G. J. Kramer, *J. Phys. Chem. B.*, 1999, **103**, 6133.
67. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 1372.