Supporting Information for: "Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states"

Adrian W. Lange and John M. Herbert[*]

Department of Chemistry, The Ohio State University, Columbus, OH 43210

February 3, 2009

1 TD-DFT basis set dependence

We examine basis set dependence of the vertical excitation energies of $\pi\pi^*$ and charge transfer (CT) states at the LRC-ωPBE level in Table S1. For A:T and A_2, the $\pi\pi^*$ excitation energies all tend to decrease, by roughly the same amount, as basis set size increases. As such, the order and energy spacing of the $\pi\pi^*$ states remain nearly constant for all basis sets. The CT states of A_2 show a similar trend as the $\pi\pi^*$ states except for the 6-311(2+,2+)G** basis. The $\pi\pi^*$ states of A_2 do not exhibit much stabilization in moving from the 6-311+G* basis to the 6-311(2+,2+)G** basis, whereas the CT states are more sensitive to this move. On the other hand, the CT state of A:T appears to experience only a small decrease in energy as a result of increasing basis set.

Table S1 also presents excitation energies for A_2 embedded in a cluster of DFT water molecules representing its first solvation shell. (This geometry is one particular snapshot from a molecular dynamics simulation.) The $\pi\pi^*$ exciton states of A_2 all appear to experience a solvent red shift of ~ 0.1 eV, independent of basis set. For the CT states, the basis-set dependence of the solvent shift is only slightly more pronounced, amounting to a 0.3–0.4 eV blue shift for the 3′-adenine $\rightarrow$

---

[*]herbert@chemistry.ohio-state.edu

Excited state	6-31G*	6-311G*	6-31+G*	6-311+G*	6-311(2+,2+)G**
		A:T base pair			
Thy $\pi\pi^*$	5.47 (0.17)	5.41 (0.17)	5.33 (0.18)	5.30 (0.18)	5.28 (0.18)
Ade $\pi\pi^*$ (W)	5.67 (0.02)	5.64 (0.02)	5.60 (0.08)	5.57 (0.09)	5.56 (0.09)
Ade $\pi\pi^*$ (B)	5.91 (0.37)	5.85 (0.38)	5.74 (0.33)	5.71 (0.32)	5.70 (0.31)
Ade $\rightarrow$ Thy CT	6.50 (0.03)	6.45 (0.02)	6.46 (0.04)	6.44 (0.04)	6.43 (0.05)
		A_2 π-stacked dimer			
$\pi\pi^*$ (W$-$)	5.70 (0.00)	5.66 (0.01)	5.54 (0.03)	5.51 (0.05)	5.49 (0.03)
$\pi\pi^*$ (W$+$)	5.72 (0.00)	5.68 (0.00)	5.63 (0.04)	5.61 (0.24)	5.59 (0.05)
$\pi\pi^*$ (B$-$)	5.82 (0.04)	5.74 (0.04)	5.66 (0.01)	5.64 (0.20)	5.63 (0.01)
$\pi\pi^*$ (B$+$)	6.00 (0.44)	5.93 (0.43)	5.79 (0.42)	5.77 (0.03)	5.75 (0.40)
$3'$-Ade $\rightarrow$ $5'$-Ade CT	6.35 (0.01)	6.29 (0.01)	6.13 (0.01)	6.12 (0.00)	6.01 (0.03)
$5'$-Ade $\rightarrow$ $3'$-Ade CT	6.55 (0.00)	6.48 (0.00)	6.37 (0.00)	6.35 (0.04)	6.16 (0.00)
		A_2 π-stacked dimer in a water cluster			
$\pi\pi^*$ (W$-$)	5.60 (0.05)	5.53 (0.07)	5.45 (0.07)	5.43 (0.07)	
$\pi\pi^*$ (W$+$)	5.63 (0.01)	5.58 (0.01)	5.55 (0.03)	5.53 (0.03)	
$\pi\pi^*$ (B$-$)	5.68 (0.01)	5.62 (0.01)	5.59 (0.00)	5.57 (0.00)	
$\pi\pi^*$ (B$+$)	5.85 (0.41)	5.77 (0.47)	5.67 (0.49)	5.65 (0.47)	
$3'$-Ade $\rightarrow$ $5'$-Ade CT	6.62 (0.00)	6.60 (0.02)	6.53 (0.06)	6.52 (0.07)	
$5'$-Ade $\rightarrow$ $3'$-Ade CT	5.89 (0.09)	5.85 (0.02)	5.83 (0.01))	5.81 (0.01)	

Table S1: Basis set dependence of vertical excitation energies (in eV) of simple nucleobase systems, calculated at the LRC-ωPBE level. Oscillator strengths are shown in parentheses.

$5'$-adenine CT, and a 0.4–0.6 eV red shift in the the $5'$-adenine $\rightarrow 3'$-adenine CT state. (As noted in the manuscript, the configurationally-averaged solvent shift is toward lower CT energies in aqueous solution, but individual solvent configurations can exhibit blue shifts, as is seen for the one of the two CT states in the present example.)

Proceeding from the smallest to largest basis sets in Table S1, the energy gap between the brightest $\pi\pi^*$ state and the adenine $\rightarrow$ thymine CT state in A:T changes by only 0.1 eV. The bright state/CT energy gap in A_2 shows a similar dependence on basis set, both in the gas phase and in the water cluster.

The 6-31G* basis set thus yields energy difference between excited states that are comparable to those obtained with much larger basis sets. Since our main focus is the relative ordering of the $\pi\pi^*$ and CT excitation energies, in rather large systems with explicit solvent molecules, 6-31G* is our basis of choice for most of the calculations described in the manuscript.

2 Benchmarking the LRC functionals

Rohrdanz *et al.*[1] have recently analyzed the performance of various LRC density functionals for TD-DFT excitation energies. In particular, they examined a set of benchmark (CC2 and CASPT2) excitation energies[2] that includes both localized ($n\pi^*$ and $\pi\pi^*$) and CT excitation energies. Table S2 here is analogous to Table II of Rohrdanz *et al.*,and presents excitation energies for this set of molecules, calculated at the TD-LRC-ωPBE/cc-pVDZ and TD-LRC-ωPBEh/cc-pVDZ levels. In the notation of Rohrdanz *et al.*, the functional that we call LRC-ωPBEh corresponds to the parameters $C_{\mathrm{HF}} = 0.2$ (*i.e.*, 20% short-range Hartree–Fock exchange) and $\omega = 0.3\ a_0^{-1}$; this is the recommended parameter set from that study. The functional that we call LRC-ωPBE corresponds to $C_{\mathrm{HF}} = 0$ and $\omega = 0.3\ a_0^{-1}$, a parameter set that was not considered in detail by Rohrdanz *et al.*

Both functionals afford root-mean-square errors (RMSEs) of 0.3 eV, for both localized and CT excitation energies. (For comparison, TD-PBE0 affords RMSEs of 0.3 eV and 3.0 eV, respectively, for the localized and CT excitation energies in this data set.[1,2]) When the signs of the errors are taken into account, however, a more nuanced picture emerges.

For localized excitations, the two LRC functionals afford excitation energies that are in reasonable agreement with one another, and with the benchmarks. In most cases, the LRC functionals

Molecule	Excitation	Type	Benchmark[a]	LRC-ωPBE[b]	LRC-ωPBEh[c]
dipeptide	$n_1 \to \pi_2^*$	CT	8.07	8.02	7.75
	$\pi_1 \to \pi_2^*$	CT	7.18	7.03	6.93
	$n_1 \to \pi_1^*$	L	5.62	5.57	5.65
	$n_2 \to \pi_2^*$	L	5.79	5.83	5.91
β-dipeptide	$n_1 \to \pi_2^*$	CT	9.13	8.83	8.45
	$\pi_1 \to \pi_2^*$	CT	7.99	8.33	8.01
	$n_1 \to \pi_1^*$	L	5.40	5.55	5.64
	$n_2 \to \pi_2^*$	L	5.10	5.69	5.77
tripeptide	$\pi_1 \to \pi_2^*$	CT	7.01	7.02	6.92
	$\pi_2 \to \pi_3^*$	CT	7.39	7.23	7.15
	$\pi_1 \to \pi_3^*$	CT	8.74	9.22	8.72
	$n_1 \to \pi_3^*$	CT	9.30	9.31	8.88
	$n_2 \to \pi_3^*$	CT	8.33	8.50	8.22
	$n_1 \to \pi_2^*$	CT	8.12	7.97	7.69
	$n_1 \to \pi_1^*$	L	5.74	5.58	5.67
	$n_2 \to \pi_2^*$	L	5.61	5.84	5.91
	$n_3 \to \pi_3^*$	L	5.91	5.92	6.00
N-phenyl pyrrole	$1\ ^1B_2$	L	4.85	5.14	5.14
	$2\ ^1A_1$	L	5.13	5.30	5.25
	$2\ ^1B_2$	CT	5.47	5.55	5.36
	$3\ ^1A_1$	CT	5.94	6.47	6.05
DMABN	^1B	L	4.25	4.75	4.75
	^1A	CT	4.56	5.02	4.99
RMSE for L excitations				0.29	0.31
RMSE for CT excitations				0.28	0.31
MSE for L excitations				0.18	0.23
MSE for CT excitations				0.10	-0.16

[a] From Peach *et al.*[2]
[b] $C_{\mathrm{HF}} = 0, \omega = 0.3\ a_0^{-1}$
[c] $C_{\mathrm{HF}} = 0.2, \omega = 0.2\ a_0^{-1}$

Table S2: Vertical excitation energies (in eV) for a set of localized (L) and charge-transfer (CT) excitations, including root-mean-square errors (RMSEs) and mean signed errors (MSEs), the latter defined as the benchmark value minus the TD-DFT value.

overestimate the excitation energies (corresponding to positive errors, in our sign convention), and the mean *signed* error (MSE) for either functional is +0.2 eV for the localized excitations.

For the CT states, the LRC-ωPBE functional affords errors of either sign, with about equal frequency, and has a MSE of +0.1 eV (corresponding to overestimation, in the mean). The LRC-ωPBEh functional, on the other hand, underestimates the CT excitation energies in 10/13 cases. Moreover, in two of the cases where LRC-ωPBEh *overestimates* a CT excitation energy. The only case where LRC-ωPBEh significantly overestimates a CT excitation energy is the ^1A state of N,N-(dimethyl)benzonitrile (DMABN), a state that actually exhibits very little CT character in the gas phase.[2] The MSE for CT states in the case of LRC-ωPBEh is +0.2 eV, *i.e.*, the signed error is in the opposite direction as it is in the case of LRC-ωPBE. (To one significant digit, the MSEs for CT states do not change if we throw out the questionable DMABN data point.)

3 Supplementary energies and oscillator strengths
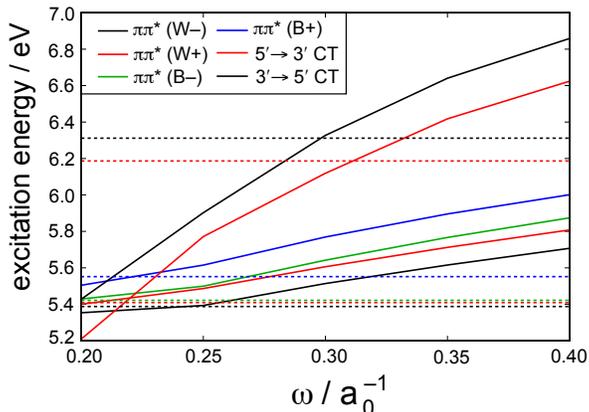
Figure S1: Vertical excitation energies for the $^1\pi\pi^*$ and CT states of A$_2$, computed at the LRC-ωPBE/6-311G* level as a function of the LRC range parameter, ω. CC2/TZVPP results are shown as horizontal dotted lines.

CT excitation energies are especially sensitive to the value of ω, as shown in Fig. S1 for a π-stacked adenine dimer, A$_2$. The $n\pi^*$ and $\pi\pi^*$ excitation energies are much less affected by the value of ω than the CT excitations. Reasonable agreement with the SCS-CIS(D) benchmarks, for both

valence and CT states, is obtained using $\omega = 0.3\ a_0^{-1}$. Because the CT excitation energies decrease rapidly as ω decreases, and because our hypothesis is that non-LRC density functionals substantially overstabilize the CT states, we also investigate the smaller value $\omega = 0.2\ a_0^{-1}$. Results presented below indicate that only values in the range $0.2\ a_0^{-1} < \omega < 0.3\ a_0^{-1}$ are acceptable for the systems considered here.

Included below are the vertical excitation energies and oscillator strengths of the simple nucleobase systems (Table S3), the ATATA system (Table S4), and the ATA:TAT system (Table S5).
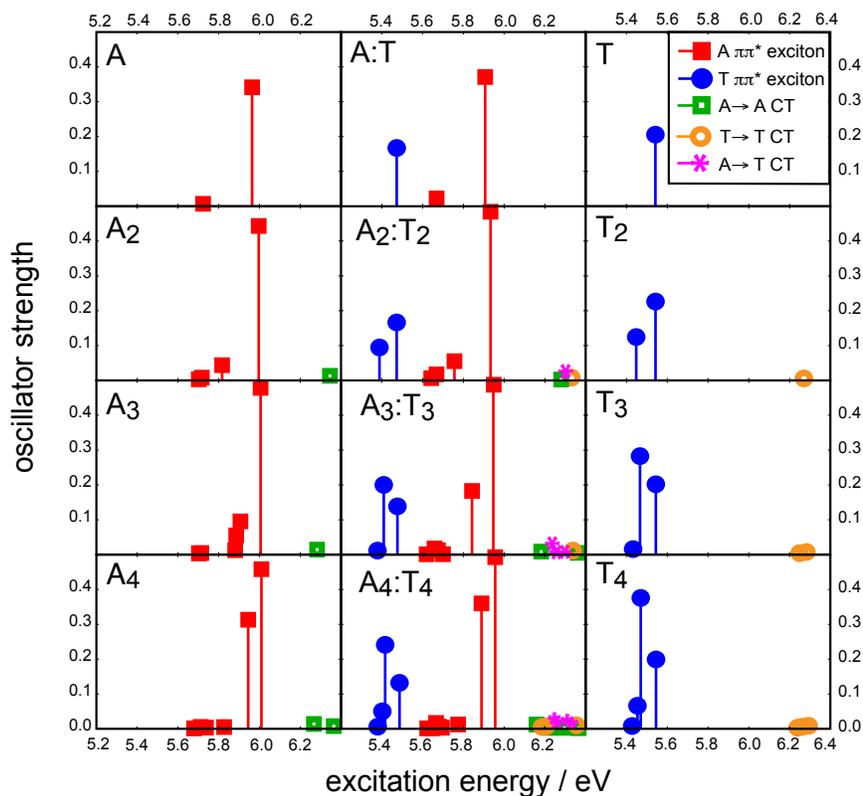
Figure S2: Stick spectra of various gas-phase A_n, T_n, and A_n:T_n multimers composed of nucleobases arranged in the canonical B-DNA geometries. All calculations are performed at the LRC-ωPBE/6-31G* level. To avoid congestion, $n\pi^*$ excitations are omitted from these spectra. (Rydberg states are largely absent due to the omission of diffuse basis functions.) The lowest adenine $\rightarrow$ thymine CT states of A:T appear at 6.5 eV, out of the range depicted here.

Excited State	Method					
	PBE0	LRC-ωPBE		CIS(D)	SCS-	CC2
		$\omega = 0.2\ a_0^{-1}$	$\omega = 0.3\ a_0^{-1}$		CIS(D)	
Adenine Monomer						
$n\pi^*$	5.19 (0.00)	4.86 (0.00)	5.26 (0.00)	5.87 (0.01)	5.80 (0.01)	5.34 (0.00)
$\pi\pi^*$ (W)	5.46 (0.19)	5.39 (0.02)	5.62 (0.06)	5.55 (0.36)	5.35 (0.36)	5.46 (0.00)
$\pi\pi^*$ (B)	5.56 (0.10)	5.44 (0.29)	5.75 (0.30)	5.67 (0.07)	5.36 (0.07)	5.58 (0.32)
Thymine Monomer						
$n\pi^*$	5.26 (0.00)	4.56 (0.00)	4.93 (0.00)	4.89 (0.00)	4.74 (0.00)	4.84 (0.00)
$\pi\pi^*$	5.59 (0.27)	5.05 (0.16)	5.34 (0.22)	5.43 (0.41)	5.30 (0.41)	5.31 (0.20)
A:T Base Pair						
Thy $n\pi^*$	4.87 (0.00)	4.70 (0.00)	5.10 (0.00)	5.27 (0.00)	5.18 (0.00)	4.94
Ade $n\pi^*$	5.35 (0.00)	5.08 (0.00)	5.51 (0.00)	5.54 (0.00)	5.50 (0.00)	5.54
Thy $\pi\pi^*$	5.14 (0.15)	5.00 (0.13)	5.30 (0.18)	5.37 (0.35)	5.23 (0.35)	5.21
Ade $\pi\pi^*$ (W)	5.41 (0.20)	5.35 (0.17)	5.57 (0.09)	5.46 (0.37)	5.25 (0.37)	5.40
Ade $\pi\pi^*$ (B)	5.52 (0.12)	5.51 (0.11)	5.71 (0.32)	5.66 (0.14)	5.39 (0.14)	5.47
Ade $\rightarrow$ Thy CT	4.45 (0.00)	5.29 (0.05)	6.44 (0.04)	6.98 (0.00)	6.96 (0.00)	6.04
A$_2$ π-Stacked Dimer						
$5'$ $n\pi^*$	5.15 (0.00)	4.82 (0.00)	5.22 (0.00)	5.42 (0.00)	5.33 (0.00)	5.26 (0.00)
$3'$ $n\pi^*$	5.16 (0.00)	4.83 (0.00)	5.24 (0.00)	5.42 (0.00)	5.34 (0.00)	5.27 (0.00)
$\pi\pi^*$(W$-$)	5.30 (0.04)	5.35 (0.01)	5.51 (0.03)	5.40 (0.05)	5.22 (0.05)	5.39 (0.05)
$\pi\pi^*$(W$+$)	5.43 (0.17)	5.40 (0.08)	5.61 (0.04)	5.57 (0.24)	5.46 (0.24)	5.41 (0.00)
$\pi\pi^*$(B$-$)	5.51 (0.03)	5.43 (0.18)	5.64 (0.01)	5.68 (0.20)	5.58 (0.20)	5.42 (0.00)
$\pi\pi^*$(B$+$)	5.56 (0.16)	5.50 (0.11)	5.77 (0.41)	5.76 (0.03)	5.74 (0.03)	5.55 (0.40)
$3'$-Ade $\rightarrow$ $5'$-Ade CT	4.95 (0.00)	5.21 (0.02)	6.12 (0.01)	6.28 (0.00)	6.28 (0.00)	6.19
$5'$-Ade $\rightarrow$ $3'$-Ade CT	5.09 (0.00)	5.43 (0.18)	6.33 (0.01)	6.50 (0.04)	6.48 (0.04)	6.32

Table S3: Vertical excitation energies (in eV) and oscillator strengths for low-lying singlet excited states of simple nucleobase systems. All methods used the 6-311+G* basis set except CC2, where the TZVP basis was used. Geometries correspond to canonical B-DNA. Some of the CC2 relaxed oscillator strengths failed to converge and are therefore not reported here. For the CIS(D) and SCS-CIS(D) excitation energies, the oscillator strengths are CIS values.

Excited State	VEE/ eV	Oscillator Strength	Character
1	4.92	0.00	Thy 2 $n\pi^*$
2	4.93	0.00	Thy 4 $n\pi^*$
3	5.25	0.00	Ade 3 $n\pi^*$
4	5.25	0.00	Ade 5 $n\pi^*$
5	5.26	0.00	Ade 1 $n\pi^*$
6	5.41	0.09	Thy 4 $\pi\pi^*$
7	5.42	0.06	Thy 2 $\pi\pi^*$
8	5.65	0.00	Ade 3 →Thy4 CT
9	5.67	0.00	Ade 1 $\pi\pi^*$
10	5.71	0.01	Ade 5 $\pi\pi^*$
11	5.84	0.19	Ade 3 $\pi\pi^*$
12	5.86	0.01	Ade 3 →T4 CT
13	5.87	0.00	Ade 1 $n\pi^*$
14	5.87	0.00	Ade 5 $n\pi^*$
15	5.89	0.02	Ade 3 →Thy4 CT
16	5.89	0.28	Ade 1 $\pi\pi^*$
17	5.92	0.27	Ade 3 →Thy4 CT (+ Ade5 $\pi\pi^*$)
18	5.97	0.01	Ade 1 →Thy2 CT
19	6.18	0.00	Ade 1 $n\pi^*$
20	6.19	0.00	Ade 3 $n\pi^*$
21	6.21	0.00	Ade 5 $n\pi^*$
22	6.21	0.00	Ade 3 →Thy2 CT

Table S4: LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATATA. The monomers are labeled by number in 5′ to 3′ order.

Excited State	VEE/ eV	Oscillator Strength	Character
1	5.12	0.00	Thy 3β $n\pi^*$
2	5.13	0.00	Thy 1β $n\pi^*$
3	5.19	0.00	Thy 2α $n\pi^*$
4	5.35	0.03	Thy 2α $\pi\pi^*$(+ Ade 2β →Thy 3β CT)
5	5.38	0.15	Ade 2β →Thy 3β CT
6	5.45	0.06	Thy 1β $\pi\pi^*$
7	5.51	0.00	Ade 1α $n\pi^*$
8	5.52	0.00	Ade 2β $n\pi^*$
9	5.54	0.00	Ade 3α $n\pi^*$
10	5.57	0.01	Ade 2β →Thy 3β CT
11	5.63	0.01	Ade 1α →Thy 2α CT
12	5.64	0.01	Ade 3α $\pi\pi^*$
13	5.75	0.11	Ade 2β →Thy 3β CT
14	5.77	0.39	Ade 2β $\pi\pi^*$(+ Ade 3α $\pi\pi^*$)
15	5.83	0.20	Ade 3α $\pi\pi^*$(+ Ade 1α $\pi\pi^*$)
16	5.87	0.15	Ade 1α $\pi\pi^*$(+ Ade 2β $\pi\pi^*$)
17	5.92	0.03	Ade 1α →Thy 2α CT
18	5.97	0.00	Ade 2β $n\pi^*$(+ Ade 2α $n\pi^*$)
19	5.98	0.00	Ade 3α $n\pi^*$
20	5.98	0.00	Ade 1β $n\pi^*$(+ Ade 2α $n\pi^*$
21	6.04	0.00	Ade 2β →Thy 1β CT
22	6.12	0.01	Ade 3α →Thy 1α CT
23	6.22	0.00	Thy 3β $n\pi^*$
24	6.22	0.00	Thy 2α $n\pi^*$
25	6.31	0.00	Thy 1β $n\pi^*$
26	6.32	0.00	Ade 1α $n\pi^*$
27	6.34	0.00	Ade 2β $n\pi^*$
28	6.40	0.02	Ade 3α →Thy 1β CT
29	6.41	0.00	Ade 3α $n\pi^*$
30	6.50	0.02	Ade 2β →Thy 2α CT

Table S5: LRC-ωPBE/6-31G* vertical excitation energies (VEEs) and oscillator strengths of ATA:TAT. The monomers are labeled by number in 5$'$ to 3$'$ order and additionally according to their strand (α = ATA, β = TAT).

4   MD simulations and QM/MM models

We model solvation through a mixed quantum mechanics/molecular mechanics (QM/MM) approach with water clusters derived from molecular dynamics (MD) simulations. DNA oligomers were added to a pre-equilibrated, cubic box ($L = 49.323$ Å) of water molecules at a density of 0.99 g/cm^3. MD simulations were run at constant temperature (298 K) under periodic boundary conditions with the AMBER99[4,5] and TIP3P[6] force fields, as implemented in the Tinker program.[7] In some cases, the nucleobases and/or backbone were rigidly constrained during the MD simulations to retain the same B-DNA conformation as in the gas-phase calculations. After equilibration, water molecules within a specified radius (see below) of any nucleobase atom were retained as explicit QM waters in the excited-state calculation. Ten configurations, each separated by at least 5 ps, were extracted from the MD simulations for QM/MM calculations. Initially, we considered three different QM/MM models.

1. All water molecules within 2.5 Å of any nucleobase atom are included in the QM region, and all remaining water molecules are completely omitted. We call this the "Cluster Model".

2. A second model includes the aforementioned cluster as the QM region, but also includes TIP3P point charges for water molecules within 10 Å of the QM region. We call this "QM/MM Model 1".

3. QM/MM Model 2 is analogous to Model 1, but the QM region extends to 4.0 Å around the nucleobases and the MM region extends out another 14.0 Å.

For the aqueous A:T and A_2 systems discussed below, both the cluster model and QM/MM Model 1 include about 15 QM water molecules, whereas QM/MM Model 2 includes approximately 50 QM water molecules. QM/MM Models 1 and 2 include ~ 550 and ~ 700 MM water molecules, respectively.

Figure S3 displays absorption spectra for hydrated A:T and A_2.

Finally, we present a comparison between CIS(D) and SCS-CIS(D) absorption spectra for aqueous A:T and aqueous A_2 (QM/MM Model 1). This is shown in Fig. S4.

4.1   TD-LRC-DFT results for A_2:T_2

The largest aqueous system considered here is A_2:T_2. The absorption spectrum for this system, calculated at the TD-LRC-ωPBE level, is shown in Fig. S5. On average, the $\pi\pi^*$ exciton energies
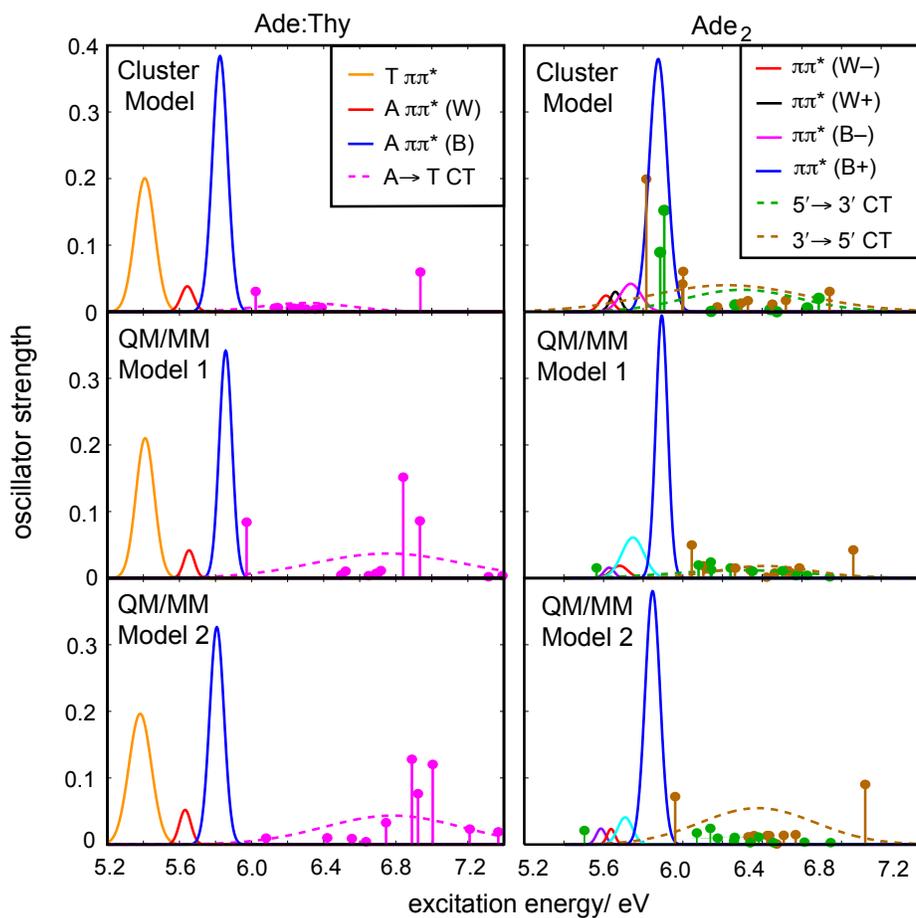
Figure S3: Absorption spectra for aqueous A:T and A$_2$ at the LRC-ωPBE/6-31G* level. (To avoid congestion, the optically-weak $n\pi^*$ states are omitted.) Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well.

S11
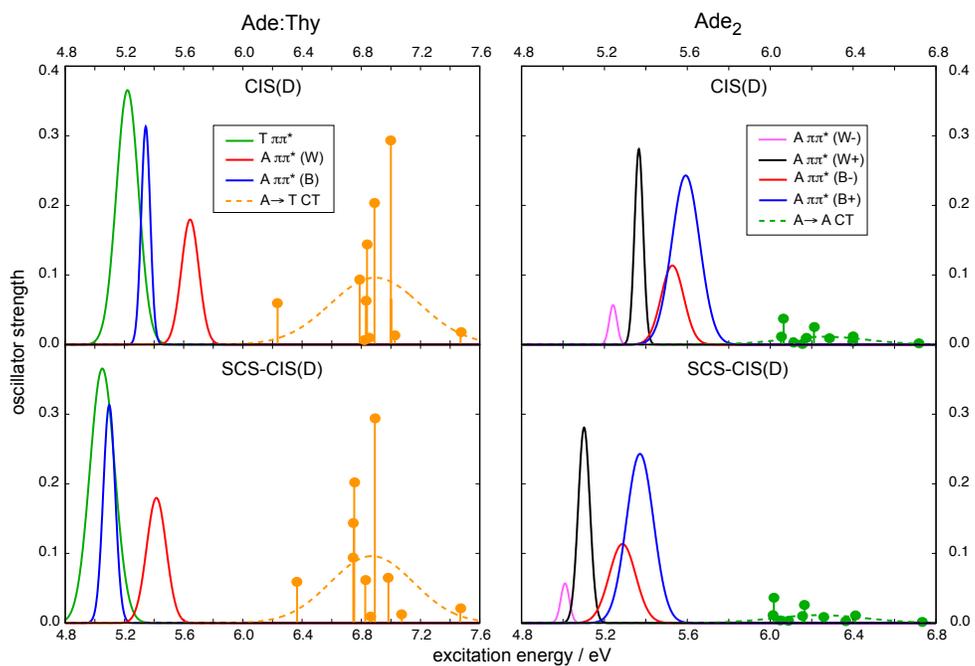
Figure S4: Absorption spectra for hydrated A:T (left panels) and hydrated A_2 (right panels), obtained at the CIS(D)/6-311+G* level using QM/MM Model 1. Gaussian distributions were obtained from averages over solvent configuration (using CIS oscillator strengths), though for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second $\pi\pi^*$ band and the CT states, lending significant oscillator strength to the latter.
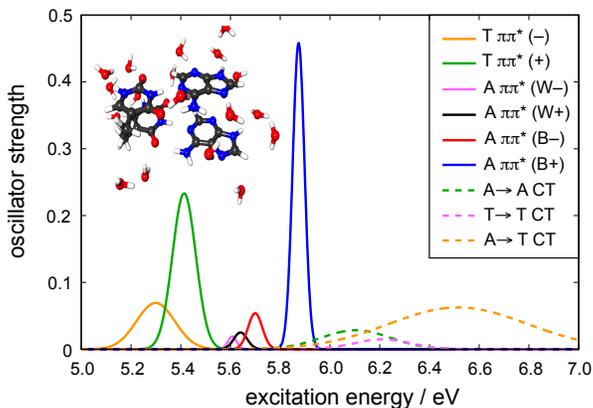
Figure S5: Absorption spectrum of A_2:T_2 obtained from a TD-LRC-ωPBE/6-31G* QM/MM calculation. The QM region is depicted in shown in the figure.

for this system are within 0.05 eV of those obtained for the analogous QM/MM model of A_2. (This close similarity arises in part because we constrain the nucleobases to their canonical B-DNA geometries during the molecular dynamics simulation; only the water degrees of freedom are subject to configurational averaging.) Compared to gas-phase A_2:T_2, the exciton states of both adenine and thymine exhibit solvatochromatic red shifts of ~ 0.1 eV.

Both intra- and interstrand CT states in this system are stabilized by 0.1–0.2 eV relative to the gas phase. As in aqueous A_2, the low-energy tail of the adenine $\rightarrow$ adenine CT band has a slight overlap with the most intense $\pi\pi^*$ band. Although the interstrand CT states are higher in energy, on average, than the intrastrand CT states (both adenine $\rightarrow$ adenine and thymine $\rightarrow$ thymine), there is significant overlap between the intra- and interstrand CT bands. Moreover, the interstrand CT states in aqueous A_2:T_2 are about 0.2 eV lower than those in aqueous A:T. In the larger system, delocalization of the virtual orbitals over adjacent π-stacked bases stabilizes the interstrand CT states.

Interstrand CT states between non-hydrogen–bonded nucleobases were observed in only half of the cluster configurations and are thus omitted from Fig. S5. When such states are observed, they appear in the range of 6.3–7.0 eV, and in each case there exists a lower-energy adenine $\rightarrow$ thymine CT state localized on a hydrogen-bonded base pair. It is possible that—as observed in the gas phase—these delocalized interstrand CT states are stabilized in larger oligomers. We plan to investigate this

possibility in future work.
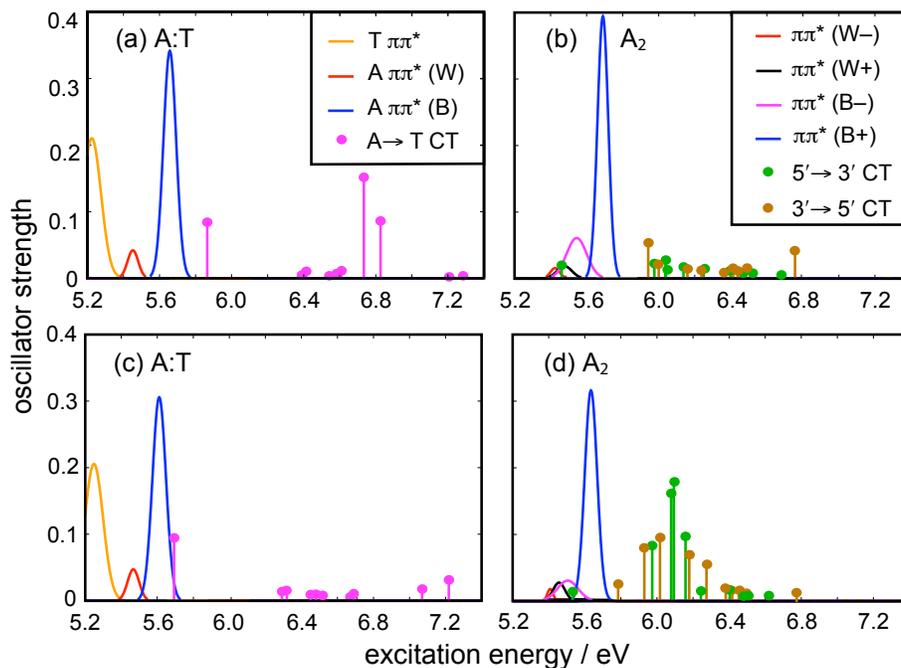
Figure S6: TD-LRC-DFT absorption spectra for aqueous A:T and A_2, in which individual excitation energies have been shifted according to estimated errors as described in the manuscript The LRC-ωPBE functional is used in (a) and (b), whereas the LRC-ωPBEh functional is used in (c) and (d). Oscillator strengths for the CT states are not reliable, as a result of intensity borrowing at the original, calculated excitation energies.

5 Complete Q-Chem reference

The complete citation for Ref. 69 in the paper is given below.[8]

References

[1] Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. *J. Chem. Phys.* **2009** (in press).

[2] Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.

[3] So, R.; Alavi, S. *J. Comput. Chem.* **2007**, *28*, 1776.

[4] Wang, J.; Cieplak, P.; Kollman, P. A. *J. Comput. Chem.* **2000**, *21*, 1049.

[5] Giudice, E.; Lavery, R. *Acc. Chem. Res.* **2002**, *35*, 350.

[6] Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D. *J. Chem. Phys.* **1983**, *79*, 926.

[7] Tinker, v. 4.2 (`http://dasher.wustl.edu/tinker`).

[8] Shao, Y.; Fusti-Molnar, L.; Jung, Y.; Kussmann, J.; Ochsenfeld, C.; Brown, S. T.; Gilbert, A. T. B.; Slipchenko, L. V.; Levchenko, S. V.; O'Neill, D. P.; DiStasio R. A., Jr.; Lochan, R. C.; Wang, T.; Beran, G. J. O.; Besley, N. A.; Herbert, J. M.; Lin, C. Y.; Van Voorhis, T.; Chien, S. H.; Sodt, A.; Steele, R. P.; Rassolov, V. A.; Maslen, P. E.; Korambath, P. P.; Adamson, R. D.; Austin, B.; Baker, J.; Byrd, E. F. C.; Dachsel, H.; Doerksen, R. J.; Dreuw, A.; Dunietz, B. D.; Dutoi, A. D.; Furlani, T. R.; Gwaltney, S. R.; Heyden, A.; Hirata, S.; Hsu, C.-P.; Kedziora, G.; Khaliullin, R. Z.; Klunzinger, P.; Lee, A. M.; Lee, M. S.; Liang, W.; Lotan, I.; Nair, N.; Peters, B.; Proynov, E. I.; Pieniazek, P. A.; Rhee, Y. M.; Ritchie, J.; Rosta, E.; Sherrill, C. D.; Simmonett, A. C.; Subotnik, J. E.; Woodcock III, H. L.; Zhang, W.; Bell, A. T.; Chakraborty, A. K.; Chipman, D. M.; Keil, F. J.; Warshel, A.; Hehre, W. J.; Schaefer III, H. F.; Kong, J.; Krylov, A. I.; Gill, P. M. W.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.