J|A|C|S
ARTICLES

Published on Web 03/03/2009

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, Ohio 43210

Received November 17, 2008; E-mail: herbert@chemistry.ohio-state.edu

Abstract: Vertical electronic excitations in model systems representing single- and double-stranded B-DNA are characterized using electronic structure theory, including both time-dependent density functional theory (TD-DFT) and correlated wave function techniques. Previous TD-DFT predictions of charge-transfer (CT) states well below the optically bright $^1\pi\pi^*$ states are shown to be artifacts of the improper long-range behavior of standard density-functional exchange approximations, which we rectify here using a long-range correction (LRC) procedure. For nucleobase dimers (hydrogen-bonded or π-stacked), TD-LRC-DFT affords vertical excitation energies in reasonable agreement with the wave function methods, not only for the $^1n\pi^*$ and $^1\pi\pi^*$ states but also for the CT states, and qualitatively reproduces well-known base-stacking effects on the absorption spectrum of DNA. The emergence of $^1\pi\pi^*$ Frenkel exciton states, localized on a single strand, is clearly evident, and these states (rather than low-energy CT states) are primarily responsible for the fact that DNA's absorption spectrum exhibits a red tail that is absent in monomer absorption spectra. For B-DNA in aqueous solution, the low-energy tail of the CT band (representing both intra- and interstrand CT states) appears at energies comparable to those of the optically bright $^1\pi\pi^*$ exciton states. In systems with more than one base pair, we also observe the emergence of delocalized, interstrand CT excitations, whose excitation energies may be significantly lower than the lowest CT excitation in a single base pair. Together, these observations suggest that a single Watson−Crick base pair is an inadequate model of the photophysics of B-DNA.

I. Introduction

A detailed understanding of the photochemistry and photophysics precipitated by UV excitation of nucleobases, nucleotides, and oligonucleotides is crucial to elucidating not only the mechanisms that underlie DNA's intrinsically high photostability,[1] but also those mechanisms that do sometimes lead to DNA damage, including strand cleavage and photoisomerization products.[1,2]

UV excitation of the lowest $^1\pi\pi^*$ state of an isolated nucleobase in aqueous solution results in radiationless decay back to S_0 on a variety of time scales,[3−7] the shortest of which is subpicosecond,[3,4] suggesting the existence of one or more conical intersections near the point of initial excitation. In deoxyribo-oligonucleotides, the initially excited $^1\pi\pi^*$ Frenkel exciton state decays on a time scale $\tau_1 \approx 0.2-0.4$ ps (as determined by time-resolved fluorescence experiments),[8−12] but femtosecond time-resolved absorption spectroscopy reveals that ground-state recovery occurs much more slowly, on a time scale $\tau_2 \gtrsim 100$ ps.[13−19] Kohler and co-workers[14−18] attribute this discrepancy between absorption and fluorescence measurements to the existence of optically dark "trap" states that are populated from an excitonic bright state on a time scale τ_1, but survive for a time τ_2 before decaying back to S_0. Recently, long-lived

(1) Crespo-Hernández, C. E.; Cohen, B.; Hare, P. M.; Kohler, B. *Chem. Rev.* **2004**, *104*, 1977.
(2) Cadet, J.; Vigny, P. The photochemistry of nucleic acids. In *Photochemistry and the Nucleic Acids*; Morrison, H., Ed.; Wiley: New York, 1990; Vol. 1.
(3) Pecourt, J.-M. L.; Peon, J.; Kohler, B. *J. Am. Chem. Soc.* **2000**, *122*, 9348.
(4) Pecourt, J.-M. L.; Peon, J.; Kohler, B. *J. Am. Chem. Soc.* **2001**, *123*, 10370.
(5) He, Y.; Wu, C.; Kong, W. *J. Phys. Chem. A* **2004**, *108*, 943.
(6) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *J. Phys. Chem. B* **2006**, *110*, 18641.
(7) Hare, P. M.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 435.
(8) Markovitsi, D.; Sharonov, A.; Onidas, D.; Gustavsson, T. *ChemPhysChem* **2003**, *4*, 303.
(9) Markovitsi, D.; Onidas, D.; Gustavsson, T.; Talbot, F.; Lazzarotto, E. *J. Am. Chem. Soc.* **2005**, *127*, 17130.
(10) Markovitsi, D.; Talbot, F.; Gustavsson, T.; Onidas, D.; Lazzarotto, E.; Marguet, S. *Nature* **2006**, *441*, E7.
(11) Miannay, F.-A.; Bányász, A.; Gustavsson, T.; Markovitsi, D. *J. Am. Chem. Soc.* **2007**, *129*, 14574.
(12) Onidas, D.; Gustavsson, T.; Lazzarotto, E.; Markovitsi, D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 5143.
(13) Crespo-Hernández, C. E.; Kohler, B. *J. Phys. Chem. B* **2004**, *108*, 11182.
(14) Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2005**, *436*, 1141.
(15) Crespo-Hernández, C. E.; Cohen, B.; Kohler, B. *Nature* **2006**, *441*, E8.
(16) Takaya, T.; Su, C.; de La Harpe, K.; Crespo-Hernández, C. E.; Kohler, B. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105*, 10285.
(17) Crespo-Hernández, C. E.; de La Harpe, K.; Kohler, B. *J. Am. Chem. Soc.* **2008**, *130*, 10844.
(18) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217.
(19) Buchvarov, I.; Wang, Q.; Raytchev, M.; Trifonov, A.; Fiebig, T. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 4794.

decay components have been observed also in time-resolved fluorescence experiments.[20,21]

On the basis of the similarity between decay transients obtained for single- versus double-stranded oligonucleotides,[14] as well as the correlation between ground-state recovery time constants τ_2 and charge-transfer (CT) propensities (as estimated from the ionization energies and electron affinities of the constituent nucleobases),[16] Kohler and co-workers hypothesize that the aforementioned trap states are excimers characterized by intrastrand CT between adjacent nucleobases.[14−18] This experimental identification of intrastrand (rather than interstrand) CT states is especially interesting in view of substantial evidence that the lowest $^1\pi\pi^*$ state of the guanine−cytosine base pair (G:C) decays via CT coupled to proton transfer, in both the gas phase[22−24] and aqueous solution.[25−28] (In particular, gas-phase calculations explain the experimental observation that the Watson−Crick tautomer of G:C exhibits a much shorter S_1 lifetime than do other tautomers.)[29] Together, these observations suggest that the excited-state dynamics of the nucleobases changes qualitatively in the presence of π-stacking interactions.

Ab initio quantum chemistry stands to play an important role in unraveling the complicated electronic structure of oligonucleotides, although for systems containing multiple nucleobases, ab initio calculations are primarily limited to time-dependent density functional theory (TD-DFT). While computationally feasible in rather large molecules or clusters, TD-DFT calculations are often plagued by spurious, low-energy CT states.[30−38] Consequently, standard TD-DFT is problematic (at best) for calculations involving large molecules, especially those containing multiple chromophores,[37,38] and is essentially useless if one's goal is to identify real CT states.

Fortunately, recently developed long-range correction (LRC) procedures,[39−44] which asymptotically incorporate nonlocal Hartree−Fock (HF) exchange into existing density functionals, have been shown to push spurious CT states out of the spectral region occupied by valence $n\pi^*$ and $\pi\pi^*$ states.[37,44,45] The LRC introduces one adjustable parameter (the length scale on which the theory switches from density-functional exchange to HF exchange), and CT excitation energies are exquisitely sensitive to the value of this parameter.[37] Careful parametrization, however, can afford a functional that is just as accurate for CT excitations energies as it is for localized excitation energies.[44]

In the present work, we examine the low-lying excited states of a variety of nucleic acid systems, primarily using TD-DFT, after first verifying that TD-DFT excitation energies compare reasonably well to results obtained at the CIS(D), CC2, and CASPT2 levels, for nucleobase dimers. In larger systems, we find that TD-DFT calculations using standard (non-LRC) density functionals are saddled with pervasive CT contamination, especially if one attempts to include the sugar/phosphate backbone, solvent molecules, or counter-ions in the TD-DFT calculation. The LRC procedure, however, removes those CT states that are obviously spurious (e.g., $PO_4^- \rightarrow Na^+$ CT states at ~4 eV above the ground state), yet both intra- and interstrand CT states remain, at energies comparable to, or only slightly above, the $^1\pi\pi^*$ excitation energies. Previous reports of CT states well below the $^1\pi\pi^*$ and $^1n\pi^*$ states[46−49] appear to be artifacts. On the other hand, uncorrelated configuration-interaction singles (CIS) calculations[50] place the CT states much too high in energy, well above the $^1\pi\pi^*$ states.

In aqueous B-DNA, our calculations invariably predict both inter- and intrastrand CT states at comparable excitation energies, regardless of how the LRC is parametrized. This suggests that a realistic model of the excited states of DNA must consider both base-stacking and base-pairing interactions. A single Watson−Crick base pair fails to capture the complexity of DNA's electronic structure.

II. Computational Details

In this work, we confine our attention to nucleic acid systems composed of adenine and thymine bases. Unless otherwise noted, the geometries of all DNA oligomers correspond to the canonical B-DNA conformation.[51] The sugar/phosphate backbone is removed in most cases, and the glycosidic terminus capped with hydrogen, as this modification alters the excitation energies by only about 0.1 eV.

We model solvation through a mixed quantum mechanics/molecular mechanics (QM/MM) approach, using geometries obtained from molecular dynamics simulations at 298 K. In most cases, the nucleobases in these simulations were held rigid in their canonical B-DNA geometries, in which case the simulation amounts to an average over solvent geometries. (The sole exception is the simulation of adenine dinucleotide discussed in section V.C, in which the adenine monomers are allowed to relax.) Ten configurational snapshots, each separated by at least 5 ps, were extracted from the simulation for use in the QM/MM calculations. All water

(20) Kwok, W.-M.; Ma, C.; Phillips, D. L. *J. Am. Chem. Soc.* **2006**, *128*, 11894.
(21) Schwalb, N. K.; Temps, F. *Science* **2008**, *322*, 243.
(22) Schultz, T.; Samoylova, E.; Radloff, W.; Hertel, I. V.; Sobolewski, A. L.; Domcke, W. *Science* **2004**, *306*, 1765.
(23) Sobolewski, A.; Domcke, W. *Phys. Chem. Chem. Phys.* **2004**, *6*, 2763.
(24) Sobolewski, A. L.; Domcke, W.; Hättig, C. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 17903.
(25) Langer, H.; Doltsinis, N. L.; Marx, D. *ChemPhysChem* **2005**, *6*, 1734.
(26) Marwick, P. R. L.; Doltsinis, N. L.; Schlitter, J. *J. Chem. Phys.* **2007**, *126*, 045104.
(27) Marwick, P. R. L.; Doltsinis, N. L. *J. Chem. Phys.* **2007**, *126*, 175102.
(28) Schwalb, N. K.; Temps, F. *J. Am. Chem. Soc.* **2007**, *129*, 9272.
(29) Abo-Riziq, A.; Grace, L.; Nir, E.; Kabelac, M.; Hobza, P.; de Vries, M. S. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 20.
(30) Bernasconi, L.; Sprik, M.; Hutter, J. *J. Chem. Phys.* **2003**, *119*, 12417.
(31) Bernasconi, L.; Sprik, M.; Hutter, J. *Chem. Phys. Lett.* **2004**, *394*, 141.
(32) Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007.
(33) Magyar, R. J.; Tretiak, S. *J. Chem. Theory Comput.* **2007**, *3*, 976.
(34) Neugebauer, J.; Louwerse, M. J.; Baerends, E. J.; Wesołowski, T. A. *J. Chem. Phys.* **2005**, *122*, 094115.
(35) Neugebauer, J.; Gritsenko, O.; Baerends, E. J. *J. Chem. Phys.* **2006**, *124*, 214102.
(36) Lange, A.; Herbert, J. M. *J. Chem. Theory Comput.* **2007**, *3*, 1680.
(37) Lange, A. W.; Rohrdanz, M. A.; Herbert, J. M. *J. Phys. Chem. B* **2008**, *112*, 6304.
(38) Pollet, R.; Brenner, V. *Theor. Chem. Acc.* **2008**, *121*, 307.
(39) Iikura, H.; Tsuneda, T.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2001**, *115*, 3540.
(40) Song, J.-W.; Hirosawa, T.; Tsuneda, T.; Hirao, K. *J. Chem. Phys.* **2007**, *126*, 154105.
(41) Vydrov, O. A.; Scuseria, G. E. *J. Chem. Phys.* **2006**, *125*, 234109.
(42) Henderson, T. M.; Janesko, B. G.; Scuseria, G. E. *J. Chem. Phys.* **2008**, *128*, 194105.
(43) Chai, J.-D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 084106.
(44) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. *J. Chem. Phys.* **2009**, *130*, 054112.
(45) Tawada, Y.; Tsuneda, T.; Yanagisawa, S.; Yanai, T.; Hirao, K. *J. Chem. Phys.* **2004**, *120*, 8425.
(46) Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 37.
(47) Jean, J. M.; Hall, K. B. *Biochemistry* **2002**, *41*, 13152.
(48) Santoro, F.; Barone, V.; Improta, R. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 9931.
(49) Improta, R. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2656.
(50) Hardman, S. J. O.; Thompson, K. C. *Biochemistry* **2006**, *45*, 9145.
(51) Canonical B-DNA geometries were generated using the Nucleic program, which is part of the Tinker software package (Tinker, v. 4.2, http://dasher.wustl.edu/tinker).

molecules within 2.5 Å of any nucleobase atom were included explicitly in the QM region, while water molecules up to 10 Å away were included as TIP3P point charges. For calculation of the absorption spectrum, this procedure appears to be converged with respect to the size of the QM region. Additional simulation details, including tests of various QM/MM models, can be found in the Supporting Information.

TD-DFT calculations are performed using two LRC variants of the generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE),[52] which we denote as LRC-ωPBE (as implemented in ref 53) and LRC-ωPBEh (as implemented in ref 44). These two functionals employ somewhat different range-separation procedures, and while this distinction is important for the simultaneous description of both ground-state thermochemistry and TD-DFT excitation energies, it makes little difference in excitation energies alone.[44] The important difference between these two functionals is that the former contains no short-range HF exchange and employs a Coulomb attenuation parameter $\omega = 0.3a_0^{-1}$ that is optimal for excitation energies, in the absence of short-range HF exchange.[44,53] The LRC-ωPBEh functional contains 20% short-range HF exchange and uses $\omega = 0.2a_0^{-1}$. Here, the "h" indicates that the functional is a hybrid, even at short-range, and both ω and the fraction of short-range HF exchange have been optimized for TD-DFT excitation energies (and also for ground-state thermochemistry).[44]

Benchmark studies for a variety of molecules indicate that both TD-LRC-ωPBEh and TD-LRC-ωPBE calculations, with the parameters specified above, exhibit mean errors of 0.2−0.3 eV for both valence and CT excitation energies (see the Supporting Information as well as ref 44). In the Supporting Information, we also study the ω-dependence of TD-LRC-ωPBE and TD-LRC-ωPBEh excitation energies for the π-stacked adenine dimer, which reveals that these functionals slightly overestimate localized excitation energies, while LRC-ωPBEh slightly underestimates CT excitation energies. These errors will be taken into account when we present solution-phase absorption spectra in section V.

For comparison, we will also present results obtained from the widely used PBE0 hybrid functional[54] (also known as PBE1PBE).[55] TD-PBE0 calculations afford mean errors of ∼0.3 eV for localized valence excitation energies,[44,56] but severely underestimate CT excitation energies.[44,57]

Although basis set superposition error (BSSE) is extremely important in calculations of dimerization energies for π-stacked systems, simple estimates of the excited-state BSSE in the π-stacked cytosine dimer suggest that the S_0 and S_1 BSSEs differ by less than 0.03 eV,[58] and therefore vertical excitation energies are hardly affected. Moreover, the aug-cc-pVDZ and aug-cc-pVTZ basis sets afford virtually identical results for excitation energies, suggesting that complete-basis extrapolation (whereupon BSSE should vanish) would not greatly affect the excitation energies. As we are interested in excitation energies rather than binding energies, no attempt is made here to correct for BSSE.

For nucleobase dimers, we compare TD-DFT excitation energies to those obtained using various correlated wave function models, including CC2,[59] CIS(D),[60] and a spin-component-scaled (SCS) version of CIS(D).[61] In all three cases, we employ the resolution-of-identity (RI) version of the method.[62,63] In principle, the highest level method employed here is CC2, and for single-reference problems with closed-shell ground states, benchmark studies indicate that CC2 excitation energies are typically within 0.3 eV of full configuration interaction,[64,65] CASPT2,[66] or experiment.[67] In particular, CC2 results for each of the nucleobase monomers lie within 0.3 eV of CASPT2 results, for all and states within 7 eV of the ground state, and these CC2 excitation energies also compare favorably to "best estimates" obtained either from experiment or from higher-level ab initio calculations.[66] A recent study of 4-(dimethylamino)benzonitrile found CC2 results for CT states to be in excellent agreement with higher-level methods that include triple excitations.[68] In comparison, CIS(D) excitation energies are of slightly lower quality, and typically fall within 0.2−0.5 eV of experiment,[61,69] with slightly better performance for SCS-CIS(D).[61]

All TD-DFT calculations employ the Tamm−Dancoff approximation,[70] and only singlet excitations are considered. The LRC-DFT functionals have been implemented in a locally modified version of Q-Chem,[71] which we also use for the CIS(D) calculations. CC2 calculations were performed using Turbomole v. 5.9.[72] To characterize the excited states, we examine attachment/detachment densities,[73] natural transition orbitals (NTOs),[74] and (unrelaxed) excited-state Mulliken charges.

For the CIS(D) and CC2 calculations, we employ the largest basis sets that are feasible given our available hardware, 6-311+G* for CIS(D) and TZVP in the case of CC2. For the TD-DFT calculations, a careful study of the basis-set dependence (as reported in Table S1 of the Supporting Information) reveals that enlarging the basis set from 6-31G* to 6-311(2+,2+)G** tends to reduce the excitation energies by 0.1−0.4 eV but maintains both the ordering of the states and their relative oscillator strengths. As our interest in this work lies in the relative positions of the $\pi\pi^*$ and CT states, in large systems with multiple nucleobases and explicit solvent, most of our TD-DFT calculations employ the 6-31G* basis set.

III. Assessment of TD-DFT

A. Failure of Standard Functionals. Contemporary GGAs, and also hybrid functionals with less than 100% HF exchange, do not reproduce the correct asymptotic distance dependence for CT excitation energies.[75] In TD-DFT calculations on large molecules,[32,33] clusters,[34−37] or liquids,[30,31] this manifests as a near-continuum of spurious, low-energy CT states. Anomalously low CT excitation energies can be found between donor and acceptor orbitals that are localized on two molecules connected by a hydrogen bond,[36] that is, on length scales much shorter than the 3.4 Å base-stacking distance.

(52) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865.
(53) Rohrdanz, M. A.; Herbert, J. M. *J. Chem. Phys.* **2008**, *129*, 034107.
(54) Adamo, C.; Barone, V. *J. Chem. Phys.* **1999**, *110*, 6158.
(55) Ernzerhof, M.; Scuseria, G. E. *J. Chem. Phys.* **1999**, *110*, 5029.
(56) Adamo, C.; Scuseria, G. E.; Barone, V. *J. Chem. Phys.* **1999**, *111*, 2889.
(57) Peach, M. J. G.; Benfield, P.; Helgaker, T.; Tozer, D. J. *J. Chem. Phys.* **2008**, *128*, 044118.
(58) Santoro, F.; Barone, V.; Improta, R. *J. Comput. Chem.* **2008**, *29*, 957.
(59) Christiansen, O.; Koch, H.; Jørgensen, P. *Chem. Phys. Lett.* **1995**, *243*, 409.
(60) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. *Chem. Phys. Lett.* **1994**, *219*, 21.
(61) Grimme, S.; Izgorodina, E. I. *Chem. Phys.* **2004**, *305*, 223.
(62) Rhee, Y. M.; Head-Gordon, M. *J. Phys. Chem. A* **2007**, *111*, 5314.
(63) Hättig, C.; Weigend, F. *J. Chem. Phys.* **2000**, *113*, 5154.
(64) Koch, H.; Christiansen, O.; Jørgensen, P.; Olsen, J. *Chem. Phys. Lett.* **1995**, *244*, 75.
(65) Christiansen, O.; Koch, H.; Jørgensen, P.; Olsen, J. *Chem. Phys. Lett.* **1996**, *256*, 185.
(66) Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110.
(67) Kohn, A.; Hättig, C. *J. Chem. Phys.* **2003**, *119*, 5021.
(68) Kohn, A.; Hättig, C. *J. Am. Chem. Soc.* **2004**, *126*, 7399.
(69) Grimme, S.; Neese, F. *J. Chem. Phys.* **2007**, *127*, 154116.
(70) Hirata, S.; Head-Gordon, M. *Chem. Phys. Lett.* **1999**, *314*, 291.
(71) Shao, Y.; et al. *Phys. Chem. Chem. Phys.* **2006**, *8*, 3172.
(72) Alrichs, R.; Bär, M.; Häser, M.; Horn, H.; Kölmel, C. *Chem. Phys. Lett.* **1989**, *162*, 165.
(73) Head-Gordon, M.; Graña, A. M.; Maurice, D.; White, C. A. *J. Phys. Chem.* **1995**, *99*, 14261.
(74) Martin, R. L. *J. Chem. Phys.* **2003**, *118*, 4775.
(75) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. *J. Chem. Phys.* **2003**, *119*, 2943.

Table 1. Vertical Excitation Energies (in eV) for the Low-Lying Singlet Excited States of Adenine (A), Thymine (T), and Dimers Thereof, in Their Canonical B-DNA Geometries

	method					
excited state	PBE0[a]	LRC-ωPBE[a]	LRC-ωPBEh[a]	CIS(D)[a]	SCS-CIS(D)[a]	CC2[b]
Adenine Monomer						
nπ*	5.19	5.26	5.37	5.87	5.80	5.34
$\pi\pi$* (W,L_b)	5.46	5.62	5.64	5.55	5.35	5.46
$\pi\pi$* (B,L_a)	5.56	5.75	5.70	5.67	5.36	5.58
Thymine Monomer						
nπ*	5.26	4.93	4.95	4.89	4.75	4.84
$\pi\pi$*	5.59	5.34	5.36	5.43	5.30	5.31
A:T Base Pair						
Thy nπ*	4.87	5.10	5.11	5.27	5.18	4.94
Ade nπ*	5.35	5.51	5.58	5.54	5.50	5.54
Thy $\pi\pi$*	5.14	5.30	5.32	5.37	5.23	5.21
Ade $\pi\pi$* (W)	5.41	5.57	5.58	5.46	5.25	5.40
Ade $\pi\pi$* (B)	5.52	5.71	5.66	5.66	5.39	5.47
Ade → Thy CT	4.45	6.44	6.06	6.98	6.96	6.04
A_2 π-Stacked Dimer						
5′ nπ*	5.15	5.22	5.33	5.42	5.33	5.26
3′ nπ*	5.16	5.24	5.35	5.42	5.34	5.27
$\pi\pi$* (W−)	5.30	5.51	5.50	5.40	5.22	5.39
$\pi\pi$* (W+)	5.43	5.61	5.62	5.57	5.46	5.41
$\pi\pi$* (B−)	5.51	5.64	5.64	5.68	5.58	5.42
$\pi\pi$* (B+)	5.56	5.77	5.72	5.76	5.74	5.55
3′-Ade → 5′-Ade CT	4.95	6.12	5.88	6.28	6.28	6.19
5′-Ade → 3′-Ade CT	5.09	6.35	6.00	6.51	6.48	6.32
Mean Absolute						
MAD (nπ* and $\pi\pi$*)[c]	0.11	0.13	0.13	0.16	0.13	
MAD (CT)[d]	1.35	0.16	0.22	0.41	0.39	

[a] 6-311+G* basis. [b] TZVP basis. [c] Mean absolute deviation with respect to CC2, for the 16 valence excitations. [d] Mean absolute deviation with respect to CC2, for the 3 CT excitations.

LRC density functionals attempt to rectify this problem by smoothly introducing HF exchange (and attenuating GGA exchange) in the asymptotic limit. This switching is accomplished by a Ewald-type partition of the Coulomb operator and governed by a range-separation parameter, ω, where $1/\omega$ represents the length scale for attenuation of GGA exchange. On the basis of thermochemical benchmarks, values of $\omega \approx 0.3-0.4a_0^{-1}$ have been proposed for this parameter.[39−43,53] For TD-LRC-ωPBE, $\omega = 0.3a_0^{-1}$ affords the best performance for excitation energies[44] and is therefore the value used in most of our calculations.

Table 1 compares vertical excitation energies, at several levels of theory, for the gas-phase adenine and thymine monomers as well as the Watson−Crick A:T base pair and the π-stacked adenine dimer, A_2, in its B-DNA geometry. We label the two low-energy $^1\pi\pi$* states of adenine as either "bright" (B) or "weak" (W), according to their relative oscillator strengths. In A_2, these states are "Frenkel excitons", that is, symmetric ("+" in Table 1) or antisymmetric ("−") linear combinations of the localized $^1\pi\pi$* excitations on either monomer.

For the monomers, the excitation energies at all levels of theory are 0.3−0.5 eV higher than those obtained using multireference calculations,[76−80] most of which are in reasonable agreement with experimental band maxima in the gas phase.[81,82] Differences between these multireference calculations and the excitation energies reported in Table 1 reflect both differences in the monomer geometries (we have not optimized the geometries for the gas phase), as well as the comparatively lower quality of our methods. The important comparison for our purposes is how TD-DFT compares to CIS(D) and CC2 at the geometries used here.

Table 2. Vertical Excitation Energies (Using the 6-311G* Basis) for the Lowest Intermolecular CT State between Two Adenine Monomers (MP2/6-311++G** Geometries) Separated by 20 Å and Given a Twist Angle of 36°, as in B-DNA

method	ΔE_{CT}/eV
TD-PBE0	5.55
TD-LRC-ωPBE	8.81
TD-LRC-ωPBEh	8.35
CIS(D)	8.36
SCS-CIS(D)	8.45
eq 1, experimental data	7.19
eq 1, ab initio data	6.74

For the $\pi\pi$* and nπ* excitations, PBE0 is in good agreement with the wave function methods, but this functional places CT states in A_2 more than 1 eV below what is predicted by wave function methods. In A:T, PBE0 places CT states more than 2 eV below what the wave function methods predict. This seems suspicious and leads us to consider the lowest intermolecular CT excitation energy of two gas-phase adenine monomers separated by $R = 20$ Å. The lowest intermolecular CT excitation energy, ΔE_{CT}, is bounded below according to[75]

$$\Delta E_{CT} > \mathrm{IP} + \mathrm{EA} - 1/R \qquad (1)$$

(in atomic units), where IP and EA denote the ionization potential and electron affinity of adenine, respectively. We can thus estimate ΔE_{CT} using either experimental data (IP = 8.45 eV,[83] EA = −0.54 eV[84]) or high-level ab initio data (IP = 8.37 eV,[85] EA = −0.91 eV[86]) for gas-phase adenine. These estimates for ΔE_{CT} are presented in Table 2, alongside the results of ab initio calculations on A_2 at $R = 20$ Å.

Comparison to CIS(D) calculations reveals that eq 1, which overestimates the Coulomb stabilization of the ionized monomers, is not a particularly close lower bound; nevertheless, the TD-PBE0 result for adenine → adenine CT lies substantially below the minimum value established by eq 1. The two LRC functionals, in contrast, provide fairly good agreement with CIS(D) results for ΔE_{CT}.

In larger DNA oligomers, where TD-DFT is the only feasible ab initio method, CT contamination proliferates rapidly as the length of the oligomer increases, when standard functionals such as PBE0 are employed. This is illustrated in Figure 1, which plots the number of CT states appearing below the brightest $\pi\pi$* state in the first absorption band, for a sequence of single-

(76) Fülscher, M. P.; Serrano-Andrés, L.; Roos, B. O. *J. Am. Chem. Soc.* **1997**, *119*, 6168.
(77) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691.
(78) Mburu, E.; Matsika, S. *J. Phys. Chem. A* **2008**, *112*, 12485.
(79) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Am. Chem. Soc.* **2005**, *127*, 6257.
(80) Perun, S.; Sobolewski, A. L.; Domcke, W. *J. Phys. Chem. A* **2006**, *110*, 13238.
(81) Clark, L. B.; Peschel, G. G.; Tinoco, I., Jr. *J. Mol. Biol.* **1962**, *4*, 500.
(82) Clark, L. B.; Peschel, G. G.; Tinoco, I., Jr. *J. Phys. Chem.* **1965**, *69*, 3615.
(83) Kim, S. K.; Lee, W.; Herschbach, D. R. *J. Phys. Chem.* **1996**, *100*, 7933.
(84) Aflatooni, K.; Gallup, G. A.; Burrow, P. D. *J. Phys. Chem. A* **1998**, *102*, 6205.
(85) Roca-Sanjuán, D.; Rubio, M.; Merchán, M.; Serrano-Andrés, L. *J. Chem. Phys.* **2006**, 125.
(86) Roca-Sanjuán, D.; Merchán, M.; Serrano-Andrés, L.; Rubio, M. *J. Chem. Phys.* **2008**, *129*, 095104.

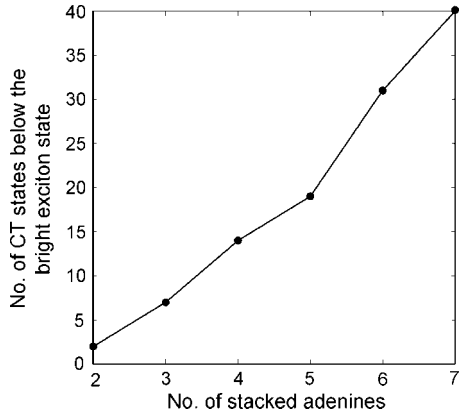

Figure 1. Plot of the number of CT states below the brightest $^1\pi\pi^*$ exciton state, computed at the TD-PBE0/6-31G* level for a sequence of single-stranded adenine multimers in their canonical B-DNA geometries, with backbone atoms removed.

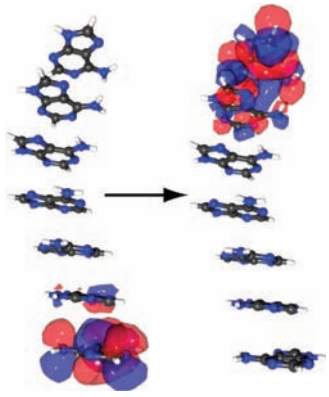

Figure 2. Natural transition orbitals of a spurious end-to-end CT state appearing at 5.6 eV above the ground state, 0.1 eV below the brightest $\pi\pi^*$ state. The calculation is performed at the TD-PBE0/6-31G* level, for π-stacked A_7 multimer. This particular particle/hole pair represents 99.9% of the transition density for the state in question.

stranded A_n homologues of increasing length. At the TD-PBE0 level, the number of low-energy CT states increases as $\sim n^2$, whereas the LRC-ωPBEh functional, for example, predicts no CT states below the bright exciton state, in any of these A_n systems. The growth in CT states predicted by PBE0 can be understood by noting that most of the predicted CT states, regardless of the functional that is employed, involve donor and acceptor orbitals that are localized on individual nucleobases. If there are n nucleobases, then there are $n(n-1)$ ways to transfer an electron from one base to another, and a close inspection of Figure 1 reveals that the TD-PBE0 method puts essentially all monomer-to-monomer CT states below the bright $\pi\pi^*$ state. This includes states such as the one depicted in Figure 2, in which an electron in A_7 is transferred from one end of the strand to the other.

Although these calculations correspond to gas-phase A_n, Improta and co-workers[48,49] have examined A_n multimers at the TD-PBE0 level, using a polarizable continuum model of aqueous solvation. They, too, report low-lying CT states between nonadjacent nucleobases, and they raise the possibility that these states might be artifacts of the method. The absence of such states in TD-LRC-DFT calculations confirms this suspicion.

When backbone atoms are introduced, TD-PBE0 calculations lose all semblance of plausibility. As an example, we consider the dinucleotide $(ApA)^-$, where the "p" denotes the phosphate/sugar backbone, again in its canonical B-DNA geometry. TD-PBE0 predicts 18 CT states below the first bright state in this system (see Table 3), 14 of which involve a significant amount of $PO_4^- \rightarrow$ adenine CT. The situation is even worse when nearby water molecules and a Na^+ counter-ion are included in the calculation [which we accomplish using a $Na^+(ApA)^-(H_2O)_{47}$ cluster], in which case each of the first 20 excited states is a CT state in which a water molecule on the surface of the cluster either donates or accepts an electron. (In this example, the solvent stabilizes the phosphate orbitals to such an extent that $PO_4^- \rightarrow$ adenine CT states are not observed among the first 20 excited states.)

Table 3. Vertical Excitation Energies (in eV) for Adenine (Ade) Dinucleotide Computed at the TD-DFT/6-31G* Level

excited state	PBE0	LRC-ωPBE	LRC-ωPBEh
Gas-Phase $(ApA)^-$			
brightest $\pi\pi^*$	5.72	5.93	5.86
lowest $PO_4^- \rightarrow$ Ade CT	3.93	6.11	5.77
lowest Ade $\rightarrow$ Ade CT	5.17	6.41	6.12
no. CT states below bright state	18	0	1
$Na^+(ApA)^-(H_2O)_{47}$			
brightest $\pi\pi^*$	>5.26	5.83	5.75
lowest Ade $\rightarrow$ Ade CT	4.78	6.03	5.70
no. CT states below bright state	>20	1	2

As we have shown previously,[36] spurious CT states in cluster calculations can be reduced in number, although not totally eliminated, by addition of MM point charges beyond the QM region, and presumably also by polarizable continuum models, for the same reason. However, such techniques do not rectify the underlying problem, a qualitatively incorrect description of long-range CT by standard density functionals. Standard functionals cannot be used to explore the complex DNA systems of interest here.

B. TD-LRC-DFT. Examining Table 1, we see that both LRC functionals afford TD-DFT excitation energies in good agreement with CC2 theory (in principle, the highest level of calculation reported here), with an average deviation from CC2 of just more than 0.1 eV. For the n$\pi^*$and $\pi\pi^*$ states, the TD-PBE0 method is in similar agreement with the CC2 results, but this method underestimates each of the CT excitation energies by more than 1 eV. For the LRC-ωPBEh functional, we note that the largest deviations occur for the two CT states in A_2, which are each underestimated (as compared to CC2) by about 0.3 eV, although the TD-LRC-ωPBE results are in excellent agreement with CC2. In a recent study of the LRC-ωPBEh functional,[44] the statistical error in vertical excitation energies was found to be about 0.3 eV (with respect to CASPT2 and other high-level benchmarks), for both valence excitations and CT excitations. The present results are in line with this, if we accept the accuracy of the CC2 benchmarks, and thus it seems safe to proceed to larger systems using these functionals. We note that in complex systems such as $Na^+(ApA)^-(H_2O)_{47}$, the LRC functionals move essentially all of the CT states to energies comparable to, or higher than, those of the states comprising the first absorption band (see Table 3). The same is true in π-stacked A_n.

IV. Base-Stacking and Base-Pairing Effects

A. Absorption Spectra. Consistent with other calculations,[76−79] we find that the first absorption band of adenine monomer is comprised of two closely spaced $^1\pi\pi^*$ states, with the higher-energy state ("^1L_a") possessing the larger oscillator strength. In

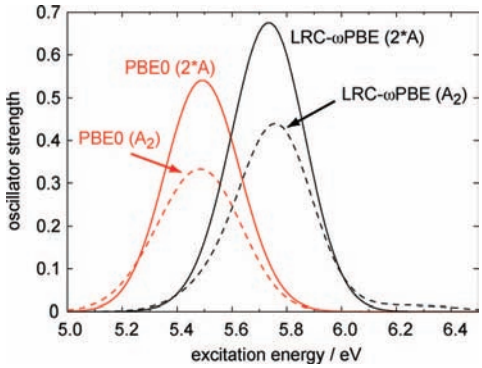

Figure 3. Absorption spectra for A_2 (broken curves) and for adenine monomer (solid curves), computed by applying a 0.3 eV Gaussian broadening to the gas-phase vertical excitation energies, weighted by their respective oscillator strengths. The monomer spectra are labeled as "2*A" to indicate that these oscillator strengths are weighted by an additional factor of 2. Both calculations employ the 6-31G* basis set and use canonical B-DNA geometries.

oligomers, the coupling between localized $^1\pi\pi^*$ excitations on individual nucleobases leads to delocalized "excitonic" states, and Platt's L_a/L_b notation[87] loses it meaning. When we need a nomenclature for the $^1\pi\pi^*$ exciton states, we will label them as simply "bright" (B) or "weak" (W), in reference to their relative oscillator strengths.

Figure S4 in the Supporting Information shows how the gas-phase excitation spectra evolve in A_n, T_n, and A_n:T_n, from $n = 1$ to $n = 4$, when the monomers are assembled in the B-DNA configuration. As the number of stacked bases increases, the higher-energy bright state shifts to slightly higher energy, while the weaker, lower-lying $\pi\pi^*$ states shift slightly to the red. The net result is a blue-shifted absorption spectrum (relative to the monomer's spectrum) with a red tail, consistent with experimental steady-state absorption measurements.[13] These effects arise due to coupling introduced by base stacking,[89−91] which causes higher-energy, symmetric combinations of localized $\pi\pi^*$ excitations to exhibit larger oscillator strengths than lower-energy, antisymmetric combinations, leading to the observed red tail.

In fact, all of the qualitative changes in the absorption spectrum of A_n, relative to the monomer spectrum, can be rationalized from these crude gas-phase calculations, by applying a Gaussian broadening to the gas-phase stick spectra and weighting the excitation energies according to their oscillator strengths. The resulting spectra for adenine monomer and dimer are shown in Figure 3. The aforementioned blue shift in the peak absorption intensity is clearly evident, as is the (very slight) red tail and also the hypochromic effect,[88] that is, the decrease in absorption intensity engendered by π-stacking.

Spectra are computed in Figure 3 at both the TD-PBE0 and the TD-LRC-ωPBE levels. Apart from an overall solvatochromatic red shift of ~0.2 eV, the TD-PBE0 spectra are quite similar to those computed previously (at the same level of theory) for 9-methyladenine and its π-stacked dimer,[48,49] using a polarizable continuum model of aqueous solvation in conjunction with the same broadening procedure employed here. In these earlier studies, the low-energy tail of the A_2 spectrum was attributed to weakly absorbing CT states, but our own TD-PBE0 calculations put the oscillator strengths of the CT excitations at <0.005, and as such these states do not contribute to the spectra shown in Figure 3. (At the TD-LRC-ωPBE level, the CT oscillator strengths are only 0.01, as compared to oscillator strengths as large as 0.41 for the $\pi\pi^*$ states.) This demonstrates that it is not necessary to invoke low-energy CT states to explain stacking-induced changes in the absorption spectrum; excitonic coupling suffices.

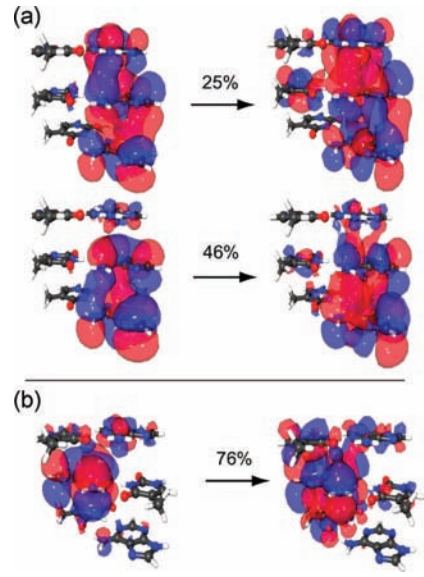

Figure 4. Natural transition orbitals (NTOs) corresponding to the excited state with largest oscillator strength in (a) A_3:T_3 and (b) ATA:TAT. In (a), the exciton is localized almost entirely on the adenine strand and consists of two significant NTO particle/hole pairs, whereas the bright state in ATA:TAT is mostly a localized monomer-like excitation.

Last, we note that excitons in A_n:T_n oligomers tend be be localized on a single strand, as shown for example in Figure 4a, due to the energy-gap dependence of the excitonic coupling[91] and the mismatch between adenine and thymine monomer excitations. This result, which arises naturally from TD-DFT calculations, is consistent with results obtained from exciton model Hamiltonians.[91,92]

B. Charge-Transfer States. For the two LRC functionals discussed in this work, the lowest adenine → thymine CT state in gas-phase A:T appears at 6.50 eV above the ground state, while the lowest thymine → adenine interstrand CT state is not among the first 20 excited states, placing it at least 7.8 eV above the ground state. These observations can be rationalized on the basis of the IPs and EAs of adenine and thymine,[83−86] which strongly favor adenine as the electron donor and thymine as the aceeptor. We do not observe thymine → adenine CT in the energy range of interest here.

In duplex A_n:T_n systems, we observe interstrand CT states in which the particle and/or the hole is delocalized over several nucleobases, as depicted in Figure 5. In certain cases, including those shown in the figure, such states may involve adenine → thymine CT between two bases that are not hydrogen bonded to one another [i.e., from the kth adenine on one strand to the $(k \pm 1)$st thymine on the opposite strand]. In larger multimers, delocalization appears to be the norm; for $n \geq 3$, we rarely observe localized, interstrand CT states.

(87) Platt, J. R. *J. Chem. Phys.* **1949**, *17*, 484.
(88) Tinoco, I., Jr. *J. Am. Chem. Soc.* **1960**, *82*, 4785.
(89) Ritze, H.-H.; Hobza, P.; Nachtigallová, D. *Phys. Chem. Chem. Phys.* **2007**, *9*, 1672.
(90) Emanuele, E.; Markovitsi, D.; Millié, P.; Zakrzewska, K. *ChemPhysChem* **2005**, *6*, 1387.
(91) Czader, A.; Bittner, E. R. *J. Chem. Phys.* **2008**, *128*, 035101.
(92) Bittner, E. R. *J. Photochem. Photobiol., A* **2007**, *190*, 328.

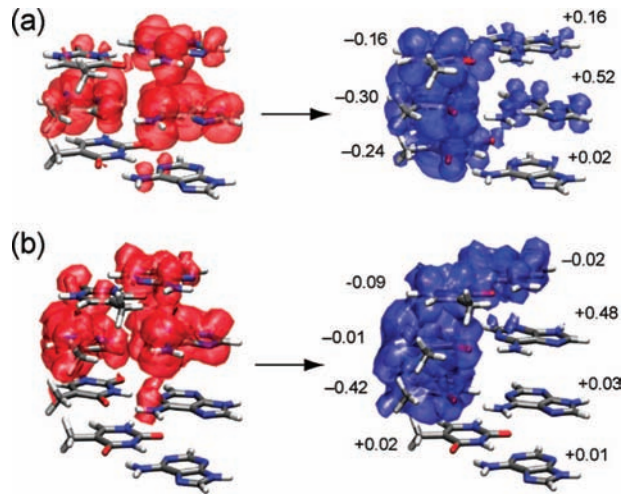

Figure 5. Attachment densities (in blue) and detachment densities (in red) for delocalized, interstrand CT states in A_3:T_3 and A_4:T_4, in which there is significant CT between nucleobases that are not hydrogen bonded. Each calculation was performed at the TD-LRC-ωPBE/6-31G* level, and the surfaces shown encapsulate 90% of the densities. Excited-state Mulliken charges on each monomer are also provided. (The monomers are essentially neutral in the ground state.)

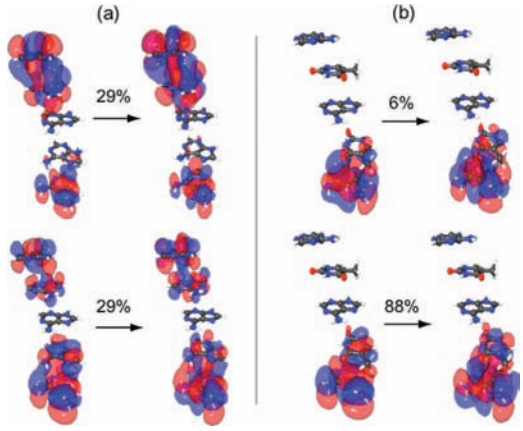

Figure 6. Natural transition orbitals (NTOs) corresponding to the state with the largest oscillator strength in (a) A_5 and (b) ATATA. The two particle/hole NTO pairs with largest amplitude are shown in either case. In A_5, the exciton state couples $\pi\pi^*$ excitations that are four bases removed in sequence, and furthermore displays a characteristic nodal structure resulting from a linear combination of localized $\pi\pi^*$ excitations. As such, there are several significant excitation amplitudes, even in the NTO basis. In contrast, the bright state in ATATA is predominantly localized on a single adenine monomer and is well-described by a single NTO particle/hole pair.

C. Alternating Base Sequences. The discussion so far has been limited to multimers constructed from homopolymers of adenine and thymine. Here, we briefly investigate the excited states found in alternating sequences of adenine and thymine. Because of energy mismatch between the monomer $\pi\pi^*$ excitations of the two nucleobases, the strong excitonic coupling observed in A_n and T_n is absent in $(AT)_n$, and consequently the bright states in the latter system are localized excitations. This is illustrated for ATATA in Figure 6. Intrastrand CT states in this single-stranded heteropolymer are much closer in energy to the brightest adenine $\pi\pi^*$ states than was seen in A_n. (Excitation energies for ATATA are available in Table S4 in the Supporting Information.)

Hybridization of ATA with TAT, to form the duplex ATA:TAT, leads to some coupling between nucleobases on opposite strands, as evident from the NTOs depicted in Figure 4b, but the coupling is weak enough that it does not induce formation of delocalized interstrand CT states, as were observed in A_n:T_n. Interstrand adenine → thymine CT on the central base pair of ATA:TAT appears at 6.5 eV, about 0.3 eV higher than in A_3:T_3, and nearly equivalent to the CT excitation energy in A:T. This observation underscores the stability of delocalized, interstrand CT in A_n:T_n, which results primarily from the delocalized character of the virtual orbitals along the thymine strand. In ATA:TAT, energy mismatch of the monomer orbitals precludes such delocalization, and in ATA:TAT we see no evidence of delocalized, interstrand CT states within 6.7 eV of the ground state. Intrastrand CT states, on the other hand, appear at energies just below the adenine absorption peak (see Table S5 in the Supporting Information), as observed also in ATATA.

Table 4. CIS(D)/6-311+G* Vertical Excitation Energies (in eV) for the Low-Energy Excited States of A_2:T_2

excited state	CIS(D)	SCS-CIS(D)
Thy $\pi\pi^*$ (−)	5.19	5.05
Thy $\pi\pi^*$ (+)	5.34	5.21
Ade $\pi\pi^*$ (W−)	5.28	5.10
Ade $\pi\pi^*$ (W+)	5.38	5.18
Ade $\pi\pi^*$ (B−)	5.55	5.27
Ade $\pi\pi^*$ (B+)	5.58	5.45
Ade → Ade CT	6.00	6.00
Thy → Thy CT	6.34	6.40
Ade → Thy CT (localized)	6.41	6.36
Ade → Thy CT (delocalized)	6.10	6.20

The experimental results of Crespo-Hernández et al.[14] suggest that CT states are formed in high yield in $(AT)_{n/2}$:$(TA)_{n/2}$, just as they are in single-stranded A_n and double-stranded A_n:T_n; however, the lifetime of these states is a factor of 2 shorter in the case of alternating sequences.[14] In addition, the experimental results of Buchvarov et al.[19] suggest that $(AT)_{n/2}$:$(TA)_{n/2}$ oligomers do not form delocalized excitons to the same extent as in the homopolymers. Our data directly support the latter suggestion and are at least consistent with the observations of Crespo-Hernández et al. Weak excitonic coupling between adenine and thymine prevents the formation of delocalized CT states in ATA:TAT.

D. CIS(D) Results for A_2:T_2. Although we are unable to extend the CC2 calculations beyond two nucleobases, the CIS(D) calculations can be extended to slightly larger systems, and in Table 4 we list the low-energy excited states of A_2:T_2 computed at the CIS(D) level. Consistent with the results for A_2 and A:T (Table 1), we find intrastrand adenine → adenine CT states 0.4−0.5 eV above the brightest excitonic states, with interstrand CT states a bit higher. [Note that the comparison to CC2 in Table 1 indicates that CIS(D) may overestimate the adenine → thymine CT excitation energies.] Perhaps the most interesting feature of these A_2:T_2 calculations is the existence of an adenine → thymine CT state that is delocalized over all four nucleobases, which appears only 0.1−0.2 eV above the adenine → adenine CT state and well below another adenine → thymine CT state that is localized on a single base pair. This example demonstrates how the electronic structure may change significantly as the model system is extended beyond two nucleobases.

The overall picture that emerges from these CIS(D) calculations is consistent with that obtained from TD-LRC-DFT. For A_2:T_2, both methods predict small red shifts, relative to A_2, in the adenine-localized $\pi\pi^*$exciton energies as well as the lowest adenine → adenine CT excitation. At the CIS(D) level, base

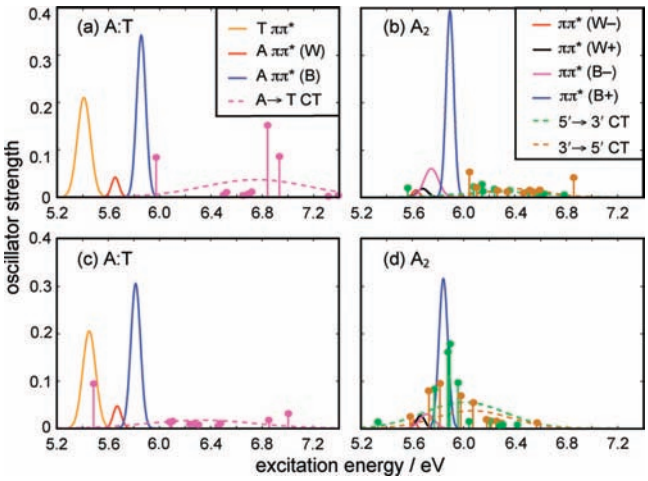

Figure 7. Absorption spectra for aqueous A:T and A_2 obtained from a TD-DFT/6-31G* QM/MM calculation. The LRC-ωPBE functional is used in (a) and (b), whereas the LRC-ωPBEh functional is used in (c) and (d). To avoid congestion, the optically weak $^1n\pi^*$ states are omitted. Gaussian distributions are obtained from averages over solvent configuration; for the CT states, the stick spectra are shown as well. The CT states around 6.9 eV borrow intensity from the second $\pi\pi^*$ absorption band, which is not shown.

stacking stabilizes the localized interstrand adenine → thymine CT state by 0.6 eV relative to the unstacked A:T base pair and stabilizes the delocalized adenine → thymine CT state to an even greater extent. In the gas phase, and using the canonical B-DNA geometry, it appears that the CT states are significantly higher in energy than the $\pi\pi^*$ exciton states comprising the first absorption band. It is also apparent, however, that whenever interstrand CT states are found, intrastrand CT states are always present at comparable energies.

V. Multimers in Aqueous Solution

Up to this point, we have considered only gas-phase multimers, to identify the electronic effects of base stacking and pairing, as distinct from solvent interactions. We next consider solution-phase multimers, using a QM/MM model of aqueous solvation that treats the first solvation shell with DFT and bulk water with MM point charges, as detailed in the Supporting Information.

A. TD-LRC-DFT Results for A:T and A_2. Figure 7 displays TD-LRC-DFT absorption spectra for aqueous A:T and A_2, obtained by averaging over solvent configurations. The duplex A_2:T_2 is examined in the Supporting Information. (Apart from a ~0.2 eV reduction in the adenine → thymine CT excitation energies, the A_2:T_2 spectrum is largely a superposition of the A_2 and A:T spectra examined here.)

Each excitonic $\pi\pi^*$ state in Figure 7 is represented by a gaussian centered at the mean excitation energy and weighted by the mean oscillator strength, with a width (standard deviation) obtained from configurational averaging. The two LRC functionals examined here afford nearly identical results for the $\pi\pi^*$ states, which are red-shifted by about 0.2 eV relative to gas-phase calculations, in good agreement with experimentally measured solvatochromatic shifts.[82] The narrow widths of the gaussian distributions indicate that these states are largely unaffected by solvent fluctuations.

The CT states are far more sensitive to solvent configuration and span a range of more than 1 eV in both A:T and A_2. To emphasize this fact, stick spectra for the CT states are included in Figure 7, along with gaussian fits as described above. For A_2, in particular, the CT states sometimes overlap with the $\pi\pi^*$ states (or, less frequently, fall slightly below the $\pi\pi^*$ states), but there is also a broad tail out to significantly higher energies, 0.5−1.0 eV above the $\pi\pi^*$ absorption band.

On average, the CT states lie about 0.1 eV below the corresponding gas-phase values, although this number belies the breadth of the CT distribution in solution. As a check of our averaging procedure, however, we also calculated the solvatochromatic shift for the lowest CT state of A_2, using the SS(V)PE solvation model[93] at the LRC-ωPBE/6-31G* level. [Because SS(V)PE is not yet available for TD-DFT, the CT excitation energy was instead calculated by using the maximum overlap method[94] to find an excited-state self-consistent field solution corresponding to charge transfer.] This procedure also predicts a 0.1 eV solvatochromatic red shift, in perfect agreement with the average QM/MM result. A more conventional TD-DFT/polarizable continuum calculation of the solvent shift predicts a very small red shift of <0.1 eV.[49] We take these results as an affirmation of the validity of the QM/MM averaging procedure.

In addition to being sensitive to solvent configuration, the CT excitation energies are also far more sensitive to the choice of LRC functional than are the $\pi\pi^*$ states. Consistent with gas-phase results, the LRC-ωPBEh functional predicts a systematic 0.3−0.4 eV reduction in the CT excitation energies, as compared to those obtained using LRC-ωPBE. For aqueous A_2, this is enough to move the CT band from being just above the $\pi\pi^*$ band [in the case of LRC-ωPBE, Figure 7b] into a spectral region that largely overlaps the $\pi\pi^*$ band [for LRC-ωPBEh, Figure 7d], leading to substantial configuration mixing and intensity borrowing in the latter case. In A:T, both functionals predict that the CT states mostly lie above the $\pi\pi^*$ states.

These differences between functionals prompted us to re-examine the benchmark TD-LRC-DFT data of Rohrdanz et al.[44] When compared to a set of high-level ab initio benchmarks,[57] both functionals afford the same root-mean-square error, 0.3 eV, for both localized ($n\pi^*$ and $\pi\pi^*$) excitations and CT excitations, as summarized in Table S2 of the Supporting Information. The LRC-ωPBEh functional, however, almost always overestimates the CT excitation energies, while LRC-ωPBE exhibits positive and negative errors with approximately equal frequency. For CT excitation energies, the mean signed errors for these two functionals are −0.2 eV (for LRC-ωPBEh) and +0.1 eV (for LRC-ωPBE). For $n\pi^*$ and $\pi\pi^*$ excitations, both functionals exhibit a mean signed error of +0.2 eV, indicating that localized excitation energies are overestimated, on average.

Assuming that the benchmarks in ref 57 are representative, this suggests shifting the CT states in Figure 7 by −0.1 and +0.2 eV for LRC-ωPBE and LRC-ωPBEh, respectively, while shifting the $\pi\pi^*$ states by −0.2 eV in both cases. This modification removes the discrepancy between the two LRC functionals, leaving a spectrum in which the CT band is centered just above the $\pi\pi^*$ band, with only a few low-energy outliers that overlap the $\pi\pi^*$ band. The shifted spectra may be found in Figure S7 in the Supporting Information. (Note that oscillator strengths in the shifted spectra are not reliable, especially for the CT states, due to intensity borrowing when TD-DFT predicts CT/$\pi\pi^*$ quasi-degeneracies.)

(93) Chipman, D. M. *J. Chem. Phys.* **2000**, *112*, 5558.
(94) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. *J. Phys. Chem. A* **2008**, *112*, 13164.

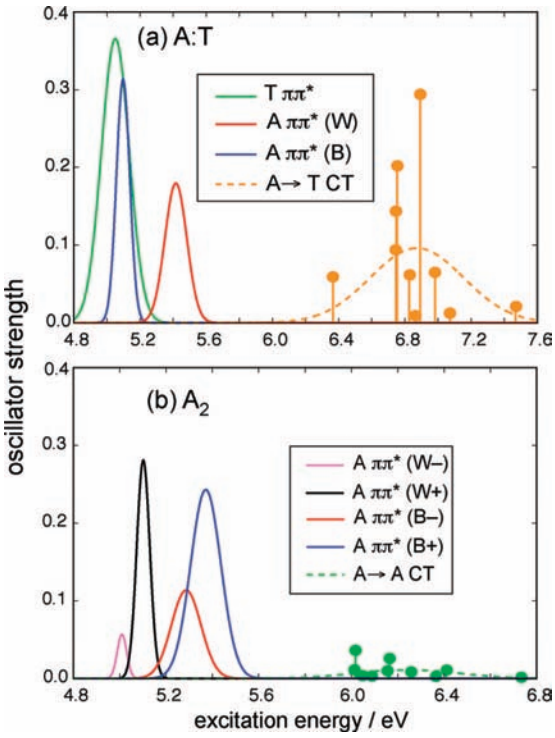

Figure 8. Absorption spectra for (a) hydrated A:T and (b) hydrated A_2, obtained from a SCS-CIS(D)/6-311+G* QM/MM calculation, using the same configurational snapshots used to generate Figure 7. Gaussian distributions were obtained from averages over solvent configuration (using CIS oscillator strengths), although for the CT states the stick spectra are shown as well. In A:T there is considerable mixing between the second $\pi\pi^*$ band (not shown) and the CT states, lending significant oscillator strength to the latter.

B. CIS(D) Results for A:T and A_2. Absorption spectra for A_2 and A:T, computed at the SCS-CIS(D) level, are shown in Figure 8. [Spectra obtained at the CIS(D) can be found in Figure S6 of the Supporting Information.] Like the TD-LRC-DFT calculations, this method predicts a 0.2 eV solvatochromatic red shift in the $\pi\pi^*$ excitation energies, although CT states at the CIS(D) and SCS-CIS(D) levels appear at slightly higher energies, as compared to TD-LRC-DFT predictions. Based on comparison to CC2 results in gas-phase A_2 and A:T, it appears that CIS(D) and SCS-CIS(D) slightly overestimate the intrastrand CT excitation energies and significantly overestimate interstrand CT excitation energies, whereas TD-LRC-DFT results are much closer to CC2.

C. TD-LRC-DFT Results for $(ApA)^-$. Last, we consider adenine dinucleotide, $(ApA)^-$, in aqueous solution, using the same QM/MM approach as above, except that we allow the adenine monomer geometries to relax during the molecular dynamics simulation. (The backbone atoms are still held rigid, to maintain π-stacking.) The TD-LRC-ωPBE absorption spectrum is shown in Figure 9, which includes the corrective shifts discussed above. Consistent with other calculations of backbone-induced shifts,[95] the bright exciton peaks (B±) are red-shifted by ~0.1 eV relative to those in A_2. Relaxation of the adenine monomer geometries broadens the absorption peaks and also lowers the CT excitation energies somewhat, resulting in more significant overlap between the two bands. As in $Na^+(ApA)^-(H_2O)_{47}$, the presence of water molecules pushes the PO_4^- → adenine CT states to at least 6.5 eV in solution, and it

(95) So, R.; Alavi, S. *J. Comput. Chem.* **2007**, *28*, 1776.

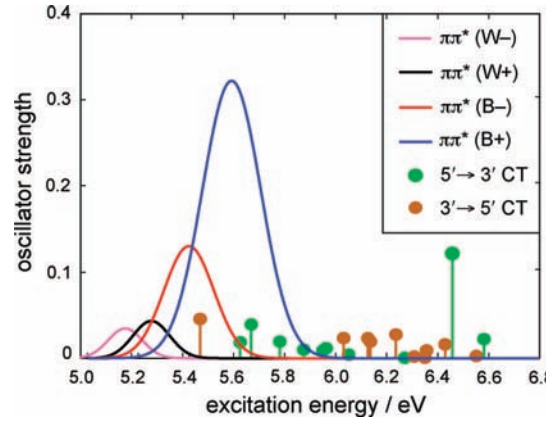

Figure 9. Absorption spectrum of aqueous $(ApA)^-$ calculated at the TD-LRC-ωPBE/6-31G* level, including corrective shifts as discussed in the text. Weakly absorbing nπ^* states are omitted for clarity.

therefore appears unlikely that the phosphate group plays a significant role in the low-energy photophysics of B-DNA.

The picture established by these solution-phase absorption spectra contrasts sharply with the results from previous TD-DFT studies of A_2 and related systems,[46−49] where CT states were found well below the bright states. In light of the foregoing discussion, we feel confident in ascribing these previous TD-DFT results to artifacts of standard TD-DFT's underestimation of CT excitation energies. In this context, we note that it is nearly impossible to repeat our aqueous $(ApA)^-$ calculations using PBE0, due to an inordinately large number of low-energy CT states, many of them involving the phosphate group and/or the solvent.

VI. Conclusions

Carefully calibrated TD-LRC-DFT calculations correct the severe underestimation (≳1 eV) of CT excitation energies exhibited by most standard density functionals, and provide excitation energies within ~0.3 eV of high-level benchmarks such as CC2 theory, for both valence (nπ^* and $\pi\pi^*$) and CT excitation energies.

In this work, we used TD-LRC-DFT to investigate the effects of base stacking and base pairing on DNA multimers composed of adenine and thymine, using a QM/MM model of aqueous solvation that includes a full solvation shell of QM water molecules. The location of CT excited states is exquisitely sensitive to the instantaneous configuration of the water molecules, and both the inter- and the intrastrand CT states span a range of more than 1 eV, as a function of solvent configuration. (Notably, the average CT excitation energy, as predicted by continuum solvation models, differs from the gas-phase value by no more than 0.1 eV.) Intrastrand adenine → adenine CT states are, on average, slightly lower in energy than interstrand adenine → thymine CT states, although there is significant overlap between the two bands. The low-energy tail of the intrastrand CT band overlaps the high-energy part of the brightest exciton state in the first $\pi\pi^*$ absorption band.

The errors in our calculated excitation energies could easily be ~0.3 eV, which (depending on the sign) could have the effect of shifting the CT states into greater overlap with the $\pi\pi^*$ states, or shifting the entirety of the CT band to a position just above the $\pi\pi^*$ band. We feel confident, however, that the CT states are neither well below, nor significantly above, the $\pi\pi^*$ band. As such, these results stand in marked contrast to previous TD-

DFT studies of π-stacked DNA multimers,[46−49] where CT states more than 1 eV below the $\pi\pi^*$ states were reported, using functionals such as B3LYP and PBE0 that are known to underestimate CT excitation energies. We regard these purported, low-energy CT states as artifacts. On the other hand, uncorrelated CIS calculations place the CT states 1−2 eV above the $\pi\pi^*$ states,[50] but this energy gap narrows to 0.5 eV or less when electron correlation is introduced.

The fact that both intra- and interstrand CT states appear at comparable energies suggests that both base-stacking and base-pairing interactions must be considered simultaneously in a reasonable simulation of the excited-state dynamics of DNA, whereas previous simulations have tended to focus on hydrogen-bonded base pairs. Additional work, including more realistic base sequences and a more thorough consideration of the effect of DNA dynamics, is needed. TD-LRC-DFT methods, with careful calibration of CT excitation energies, appear to be a promising way to build upon the simulations reported here.

Acknowledgment. We thank Dr. Mary Rohrdanz for extending the statistical analysis of ref 44 to the LRC-ωPBE functional employed here. This work was supported by an NSF CAREER award (CHE-0748448) and by the ACS Petroleum Research Fund. Calculations were performed at the Ohio Supercomputer Center under project no. PAS0291. Orbital and density plots were generated using MacMolPlt[96] and Visual Molecular Dynamics.[97]

Supporting Information Available: Additional computational details, benchmark excitation energies and oscillator strengths, and absorption spectra for some other nucleic acid systems, plus a complete citation for ref 71. This material is available free of charge via the Internet at http://pubs.acs.org.

JA808998Q

(96) Bode, B. M.; Gordon, M. S. *J. Mol. Graphics Modell.* **1998**, *16*, 133.
(97) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33.