JCTC Journal of Chemical Theory and Computation

pubs.acs.org/JCTC Article

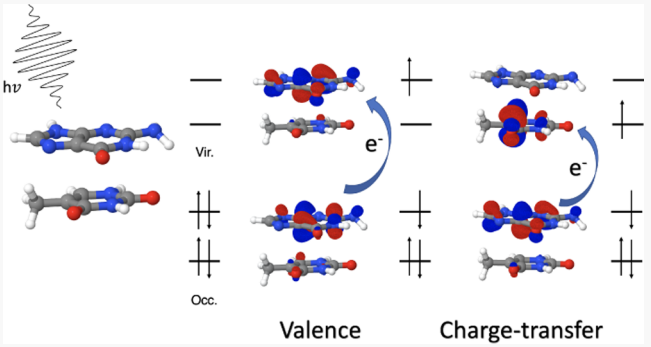
Benchmarking Quantum Mechanical Methods for the Description of Charge-Transfer States in π-Stacked Nucleobases

Camilo Zuluaga, Vincent A. Spata, and Spiridoula Matsika*

Cite This: J. Chem. Theory Comput. 2021, 17, 376−387 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Charge-transfer (CT) states are of special interest in photochemical research because they can facilitate chemical reactions through the rearrangement of electrons and subsequently chemical bonds in a molecular system. Of particular importance to this research is the transfer of electrons between π-stacked nucleobases in DNA because they play an important role in its photophysics and photochemistry. Computational methods are paramount for the study of CT states because of the current inability of experimental methods to easily detect such states. However, many ab-initio wavefunction-based and density functional theory (DFT) methods fail to accurately describe these CT states. Here, we benchmark how 40 different quantum mechanical methods describe the excited states of a guanine−thymine π-stacked nucleobase dimer system, both in 5′-TG-3′ and 5′-GT-3′ conformations. We find that the distance between the nucleobases plays a major role in the energy of the CT state and in the difference of the dipole moments between the CT and ground state. There is a larger range of values (and errors) for the energies of CT states compared to those of states localized on one nucleobase. Wavefunction-based methods have similar errors for the CT and localized valence states, while DFT methods are very sensitive to the amount of Hartree−Fock exchange. Long-range-corrected functionals with a careful balance of the Hartree−Fock exchange included can predict very accurate CT states and a balanced description with the localized states.



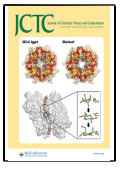
INTRODUCTION

Over the last 50 years, the field of quantum chemistry has been revolutionized by the advent of modern computing. Although ground-state properties can mostly be computed very accurately, nowadays, progress still needs to be made for excited states, particularly for systems of more than a dozen atoms. The accuracy of the description of excited states depends on their character, and a group of particularly problematic states involves states with charge-transfer (CT) character.

The focus of this work is on CT states in π-stacked nucleobase dimer systems. Generally, CT states can facilitate chemical reactions by the movement of electronic charges from one fragment to another. In DNA, this can happen as both intra-strand and inter-strand CT, where the charge moves in the same strand of DNA or the charge moves across DNA strands.[1,2] Experimental evidence suggests that CT can occur via inter-strand, but π-stacked intra-strand CT dominates overall.[3] Time-resolved infrared spectroscopy has allowed experiments to detect the importance and dynamics of CT states,[3−8] but excitation energies (EEs) are difficult to measure. Theoretical studies prove to be a valuable tool in locating and describing CT states, but only with proper benchmarking.

When DNA nucleobases absorb light, they can become damaged if unable to safely rid themselves of excess energy. Damaged DNA can present itself in many ways in the organism. One of such ways involves the formation of thymine−thymine cyclobutane-pyrimidine dimers (CPDs).[9] CPDs have been linked to uncontrolled cell replication, which can ultimately result in cancer. CT states are one of such proposed mechanisms that have been shown to provide alternative decay pathways for the repair of DNA damage. Experimental and computational studies have shown that CT states play a role in several ways to repair CPD dimers: CT is the first step in the repair of the lesion via photolyase;[10] another way involves the direct transfer of an electron from the readily oxidized nucleobase guanine or oxoguanine to the adjacent thymine in a TT-CPD dimer.[11] Alternatively, another mechanism involves an excimer complex between adjacent guanine and adenine nucleobases, which ultimately repairs a CPD lesion via the CT state from the adenine to the CPD dimer.[12] CT from adjacent bases to thymine has also been

Received: September 20, 2020
Published: December 21, 2020



https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

found to prevent the formation of CPD by providing alternative pathways.[13,14] The CT states also play an important role in the photophysics of oligonucleotides as well, as has been revealed by many experimental and theoretical studies.[4−6,15] They have been detected by UV pump/IR probe experiments and have been found responsible for longer lived signals in oligonucleotides.[3−8]

Computationally, the study of CT states is challenging. Although the role of CT states in DNA has been theoretically investigated, a systematic study of the performance of different methods has not been made. Several studies have focused on benchmarking CT states using time-dependent density functional theory (TD-DFT),[16−21] but benchmarks for CT states using wavefunction-based methods are limited,[22−25] and they have not been done for DNA because of the greater computational cost associated with more than one DNA nucleobase. The simple configuration interaction with single (CIS) excitations is well known to overestimate the energies for all excited states but more so for CT states, with errors greater than 1 eV.[22] A recent benchmark study on wavefunction methods studying CT states using several small molecules found that coupled cluster with single and double excitations (CCSD) methods tend to consistently overestimate CT states by about 0.30 ± 0.08 eV, whereas the algebraic diagrammatic construction approach through second order (ADC(2)) consistently underestimates them by about 0.36 ± 0.24 eV compared to a high-level theory CCSDT-3.[25] DFT-based methods, on the other hand, have issues well known in the literature: CT states in DFT methods when the adiabatic approximation and approximate exchange−correlation density functionals are used consistently underestimate the CT energies and may lead to a near-continuum of predicted CT states.[17,18,20] The root of the DFT problem is well established as the self-interaction error (SIE),[26,27] which leads to fragments of the system having fractional charges whose CT energies are over-stabilized and do not follow a Coulombic nature of $1/R$. A promising workaround the SIE is the inclusion of Hartree−Fock (HF) exchange in the xc functional,[28] leading to several more hybrid functionals that have a varying degree of HF exchange at the short and long ranges.[29,30]

This work focuses on using a variety of methods to benchmark the energies and dipole moments of CT states in 5′-GT-3′ and 5′-TG-3′ (thereafter called just GT and TG, respectively) π stacked dimers. We focus on these dimers because of their higher proclivity in stabilizing the CT states (because of their redox properties), as the description of the CT states is our primary interest here. This makes them a more ideal system to test for CT state description. We focus on an extensive study of methods, making it harder to extend the variation of systems as well. Nevertheless, it seems reasonable to expect that the results found here are representative of CT in nucleobases. Both single reference and multireference methods (MRMs) will be tested in order to establish the effect of both dynamical and static correlation. In addition, several DFT functionals will be benchmarked. We will focus on the inclusion of long-range exchange (LRE) as a means to allow DFT functionals to perform in a comparable manner to high-level methods, such as ADC(3).[31]

COMPUTATIONAL DETAILS

In order to probe the excited states of the dimer system, a trimer 5′-TGT-3′ system was built using Chimera.[32] Furthermore, single bases thymine and guanine were each individually optimized at a CCSD level with a 6-31G(d) basis set. Subsequently, the single bases were overlaid on top of the trimer system using the root-mean-square deviation algorithm implemented in Chimera, which uses a least-square fitting to calculate the distance between two separate atoms. The atoms used for fitting were the aromatic rings of the bases of the trimer system to the aromatic rings of the individual basis. The result is two dimer systems 5′-TG-3′ and 5′-GT-3′ that conserve the orientation of B-DNA. Figure 1 describes both of the DNA dimers of concern in this paper, with the direction shown and the separation vector from the center of mass of each nucleobase depicted.

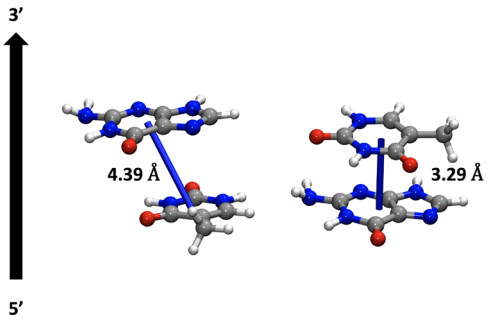


Figure 1. Relative orientations of the two dimer systems 5′-TG-3′ and 5′-GT-3′, respectively, with the separation vector drawn, which represents the distance between the nucleobases' center of mass. The length of the vector in Å is also shown along with the helical orientation.

Given the large number of methods used in this work, Table 1 is used to enumerate the methods depicted in the figures that will follow. In addition, Table 2 documents both the ω parameter and the percentage of short-range exchange (SRE) used in the functionals. Several long-range corrected (LRC) functionals were chosen with varying degrees of the aforementioned ω parameter (to be defined below). Typical global hybrid functionals were also chosen to showcase the failing of regular DFT functionals. The complete list is shown in Table 1. The following wavefunction single reference methods (SRMs) were chosen: CIS[50] and perturbative correction of the CIS method, CIS(D);[51] algebraic diagrammatic construction scheme for the polarization propagator through second and third order, ADC(2) and ADC(3);[52] the equation of motion coupled cluster with singles and doubles (EOM-CCSD),[53,54] as well as the efficient coupled-cluster (CC) alternative, similarity transformed equations of motion CCSD (STEOM-CCSD).[55] The STEOM-CCSD method was run in ORCA using the default procedure to choose the active space, which is from state-averaged CIS natural transition orbitals (NTOs) with an occupation number threshold of 0.001. ADC(3) will be used in this work to benchmark against the other methods. ADC(3) is the excited-state analog of MP3, with the single and double excitations treated consistently through third and first order, respectively.[31,56,57] Although ADC(3) contains double excitations, it has been shown to often perform better than CCSD,[31,56,57] so it has been used as benchmark in the literature.[58−62] MRMs are included as well, specifically the complete active space self-consistent field (CASSCF) and n-electron valence state perturbation theory (NEVPT2).[63,64] Several active spaces were tested in the calculations, and these are indicated as (n, m), where n is the

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

Table 1. Numbering of Each Method Used in This Study[a]

1	ADC(3)	11	NEVPT2 (10,10)	21	ωB97X-D	31	ωB97X-rV
2	ADC(2)	12	NEVPT2 (12,12)	22	ωM05-D	32	ωM06-D3
3	EOM-CCSD	13	B3LYP	23	ωB97X-D3	33	CAM-QTP01
4	STEOM-CCSD	14	PBE0	24	M11	34	rCAM-B3LYP
5	CIS	15	M06	25	CAM-QTP00	35	ωB97
6	CIS(D)	16	M06-2X	26	LRC-ωPBE	36	revM11
7	CASSCF (8,8)	17	MN12-SX	27	ωB97M-V	37	LC-VV10
8	CASSCF (10,10)	18	N12-SX	28	ωB97M-rV	38	LC-rVV10
9	CASSCF (12,12)	19	SOGGA11-X	29	ωB97X	39	LC-ωPBE08
10	NEVPT2 (8,8)	20	LRC-ωPBEh	30	ωB97X-V	40	LRC-BOP

[a]Long-ranged corrected range separated functionals (20−40) are ordered in increasing ω parameter. For functionals with matching ω, they are ordered in increasing order of short-range HF exchange included.

Table 2. Long-Ranged Corrected Methods Used in This Work Shown with the Respective ω Parameter and SRE Included in the Functional[a]

label	method	ω [bohr^{-1}]	SRE (%)
20	LRC-ωPBEh[33]	0.20	20.0
21	ωB97X-D[34]	0.20	22.2
22	ωM05-D[35]	0.20	37.0
23	ωB97X-D3[36]	0.25	19.6
24	M11[37]	0.25	42.8
25	CAM-QTP00[38]	0.29	54.0
26	LRC-ωPBE[39]	0.30	0
27	ωB97M-V[40]	0.30	15.0
28	ωB97M-rV[40,41]	0.30	15.0
29	ωB97X[42]	0.30	15.8
30	ωB97X-V[43]	0.30	16.7
31	ωB97X-rV[41,43]	0.30	16.7
32	ωM06-D3[36]	0.30	27.2
33	CAM-QTP01[44]	0.31	23.0
34	rCAM-B3LYP[45]	0.33	18.4
35	ωB97[42]	0.40	0
36	revM11[46]	0.40	22.5
37	LC-VV10[47]	0.45	0
38	LC-rVV10[41,47]	0.45	0
39	LC-ωPBE08[48]	0.45	0
40	LRC-BOP[49]	0.47	0

[a]All functionals have 100% LRE, except rCAM-B3LYP and CAM-QTP00.

number of electrons and m is the number of orbitals. Only π orbitals were included in the active spaces, and pictures of the orbitals can be found in Supporting Information. The details of how many states were averaged in CASSCF can also be found in Supporting Information. For SRM, typically the number of guess vectors included in the reference wavefunction was greater than that of the number of states requested in order to capture CT states that might otherwise be lost to the reorganization of states from the HF reference. The basis set used in all calculations was cc-pVDZ, except the MRMs which used def2-SVP and EOM-CCSD which used 6-31G(d). In addition, to probe how different basis sets affect the CT states, we tested two methods, ADC(2) and M11, with larger basis sets. ADC(2) energies were calculated using Turbomole[65−67] with the resolution of identity approximation using the cc-pVTZ basis set. The ADC(2) results with the cc-pVDZ basis set were obtained both with and without the resolution of identity (using Turbomole and QChem respectively) and then were compared with one another. The energies were identical, while the dipole moments had differences of 0.5−1 debye. M11 energies were calculated with both cc-pVTZ and cc-pVQZ basis sets, in addition to cc-pVDZ.

In order to characterize CT states, for each SRM excited state calculated, NTOs were requested: NTOs are derived from the one-electron transition density matrix, and a single-valued decomposition is performed, which yields eigenvalues for the transitions.[68] The negative eigenvalues represent the "holes" left by the electron upon excitation, and the positive eigenvalues represent where the electron is after excitation. For MRM instead, the difference density was used to characterize transitions, that is, the electron density for both the ground state and the excited state are calculated, and the excited-state density subtracted from the ground-state density is plotted. The idea is that the difference densities showcase the movement of an electron from the ground state to the excited state. For both SRM and MRM, ground- and excited-state dipole moments were calculated and used to calculate difference dipole moments, $\Delta\vec{\mu}$, where the ground-state dipole is subtracted from the excited-state dipole for each transition. The difference dipole moment is the difference density in the vector form. We use difference dipole moments to describe CT states as these are used to describe the magnitude of CT in Stark spectroscopy experiments,[69] and in general, their magnitude and direction provide a robust way to characterize them.

To study the excited states, the programs QCHEM[70] and Gaussian 16[71] for SR methods and ORCA[72] for MRM and STEOM-CCSD were used to calculate a number of excited states. Chimera[32] was used to construct the dimers, and QCHEM was used to obtain the optimized structures.

RESULTS AND DISCUSSION

Characteristics of Excited States. Figure 2 shows the excited states for each monomer, thymine and guanine, as well as the two dimers studied in this work, 5′-TG-3′ and 5′-GT-3′. The states are color-coded for easy identification. Red lines correspond to states localized on thymine, blue lines to states localized on guanine, and green lines to CT states. It is clear that the monomer states exist in the dimers as localized states, although they are shifted because of the environment of the second base. These shifts will be discussed further below. The CT states are new states that appear in the dimers but are not present in the monomers. Their location is sensitive to the sequence, as well as the methodology, and will be the focus of much of the discussion below.

The identification of the character of states is obtained by using NTOs, difference densities, and difference dipole moments. Figure 3 showcases the ways used to characterize

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

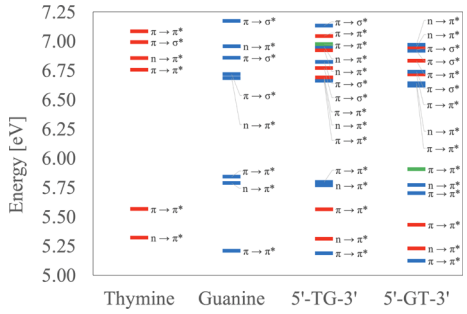


Figure 2. Energy level diagram of EEs at the ADC(3)/cc-pVDZ level for thymine, guanine, 5′-TG-3′, and 5′-GT-3′. Red levels indicate thymine local transitions, blue guanine local transitions, and green indicate a CT state.

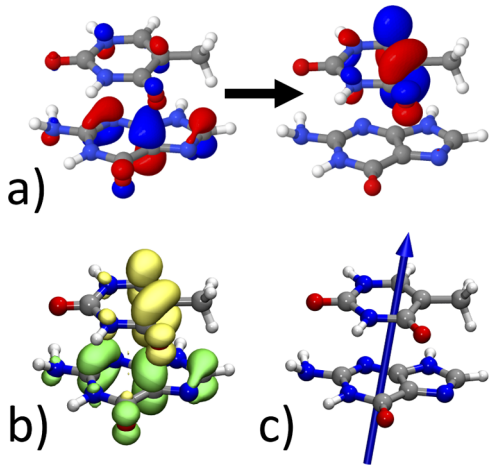


Figure 3. (a) NTOs, (b) difference densities, and (c) transition dipole moment vector $\Delta\vec{\mu}$ for the lowest-lying CT state in GT. NTOs and $\Delta\vec{\mu}$ shown here were calculated at the ADC(3)/cc-pVDZ level, while the difference density was performed at the CASSCF(12,12)/def2-SVP level. In the NTOs, the "hole" orbital is to the left of the arrow and the "particle" orbital is to the right. For difference densities, green is the ground-state density and yellow is the excited-state density. Isosurface values are 0.003 for the densities and 0.045 for the NTOs.

CT states. Figure 3a shows the NTOs, a much better descriptor than typical orbital analysis, because it takes into account all transitions as a whole. It clearly shows that in CT states, electron density moves from one base to the other (here from guanine to thymine). Figure 3b shows difference dipole moment ($\Delta\vec{\mu}$). Typical valence transitions do not have a large magnitude of $\Delta\vec{\mu}$ ($|\Delta\vec{\mu}|$ will be denoted $\Delta\mu$), around 1−4 debye for most valence transitions. Rydberg states and CT states on the other hand show a much greater magnitude $\Delta\mu$, often around or greater than 10 debye. The direction of the vector is also an indication of CT because it is perpendicular to the bases, moving from one to the other. Figure 3c shows difference densities, in which the excited-state density and the ground-state density are shown in different colors to picture the CT transition.

CT States. Figure 4 depicts all of the EEs for the lowest-lying CT state involving a transition from the guanine to the thymine nucleobase for every method included in this work when compared to ADC(3). At first glance, it is clear that the EEs of CT states for all methods of the TG strand are blue-shifted when compared to GT orientation. Most methods

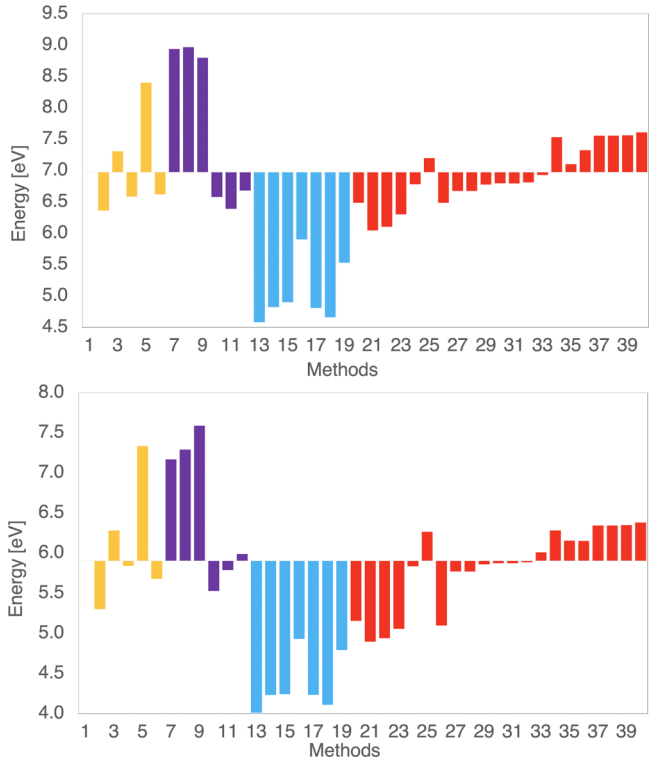


Figure 4. Energies of CT states for all methods compared to ADC(3) for TG (top) and for GT (bottom). Yellow corresponds to SRMs, purple to MRMs, blue to DFT methods that are not LRC, and red corresponds to LRC functionals.

reproduce around a 1 eV blue shift, except in the cases of some hybrid functionals, which show a smaller shift, and CASSCF which show larger shifts. We attribute this blue shift to both the larger separation distance between the nucleobases in the TG orientation and lesser overlap between the stacked nucleobase in the TG orientation. Figure 1 depicts the dimers used in this work, showcasing the relative orientation, and marks the separation distance between the center of mass of each respective nucleobase. As shown in that figure, the distance between the center of masses of the two bases is 4.39 Å in TG and 3.29 Å in GT, more than 1 Å shorter. This substantial difference in the distances causes the blue shift in the CT energies of TG compared to GT, which is universal for all methods. The difference in the distances between the two orientations has been observed in previous molecular dynamics simulations.[13] More specific observations for the various methods will be given below.

Figure 5 shows the $\Delta\mu$ magnitude for all of the methods reported. With the MRM, we used the CASSCF description to obtain the NEVPT2 character; as such, $\Delta\mu$ is the same for both CASSCF and NEVPT2. Similarly, for CIS(D) and STEOM-CCSD SRM methods the wavefunction was taken from preliminary CIS calculations and $\Delta\mu$ is replicated in the figure for consistency. As a result, there are a smaller number of wavefunction-based methods we can compare to assess $\Delta\mu$ accuracy. It is also not clear which $\Delta\mu$ is more accurate among every method because we need accurate wavefunctions for both the ground state and the excited state. Because we do not have NEVPT2 results, we can compare the effects of static correlation and dynamical correlation separately, through CASSCF and ADC(3), respectively. Qualitatively, the $\Delta\mu$ consistently show a larger magnitude for the TG orientation

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

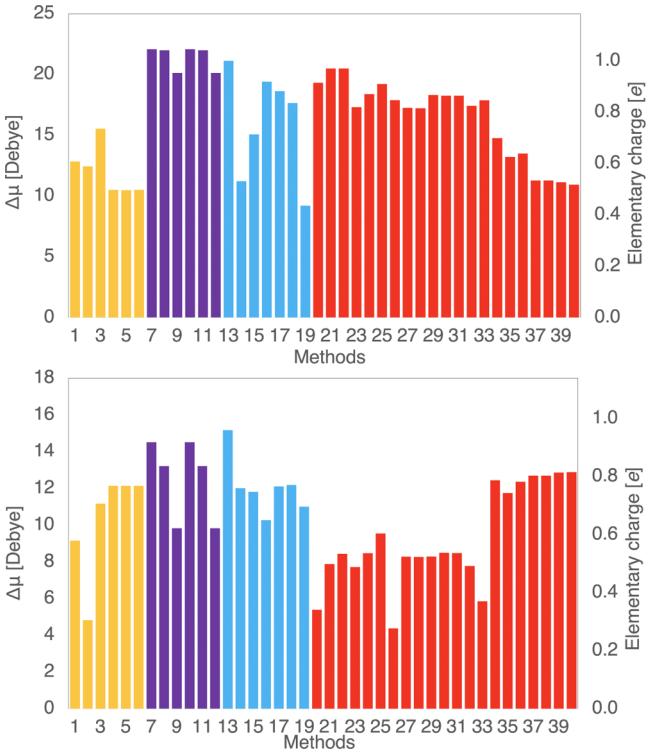


Figure 5. Left axis: $\Delta\mu$ of the lowest-lying CT state for all methods for TG (top) and for GT (bottom). Right axis: Estimated elementary charge transferred (e^-) for the lowest-lying predicted CT in both orientations.

than that of the GT orientation. Considering that the TG bases are farther when compared to GT, the larger dipole is expected. In fact, most methods do reproduce the larger $\Delta\mu$ expected for TG, although there are exceptions.

It may be more informative then to remove the distance dependence and examine how much charge is transferred during the excitation. Through $\Delta\mu$, we are able to estimate the amount of elementary charge that is transferred[73,74] between the two fragments, that is, the nucleobases: we accomplish this by using the classical definition of the dipole as the product between distance and charge, where the distance is the separation vector from each fragment's center of mass. Figure 5 depicts the predicted amount of charge transferred for all the methods. The closer this value is to 1 unit of e^-, the purer the CT state will be as far as having one electron transfer. We verify our method for calculating the CT from $\Delta\mu$ in the Supporting Information with Bader analysis[75,76] performed on the CASSCF difference densities of the CT states. The results from Bader analysis are somewhat smaller than this approach but the trends are very similar. Generally, what we see is that more charge is predicted to be transferred in the TG orientation than GT, although again not all methods predict this trend.

Comparing the energies to the charge transferred, it seems as though a smaller distance can red-shift the CT energy, while the larger $\Delta\mu$ points to more charge being transferred between fragments, irrespective of the larger distance between the fragments in question. The red shift of the energy is obvious because of the Coulombic $1/R$ dependence. The amount of charge transferred is likely affected by increased state mixing and orbital overlap at closer distances. When the CT and excitonic states mix at closer distance, the amount of charge transferred is reduced, but at larger distances, the CT states are purer. Below, we describe in more detail the performance of wavefunction-based methods and DFT-based methods separately.

Wavefunction-Based Methods. The most basic method to describe excited states is CIS.[50] This method may give qualitative results for excited states but greatly overestimates their energy. As shown in Figure 4 and as has been demonstrated before,[22] it overestimates CT states even more. CIS truncates the excitation operator for only single excitations. Perturbative corrections to CIS, such as CIS(D), add dynamical correlation and improve the performance considerably by perturbatively adding higher-order, double excitations.[51]

Other very popular methods to describe excited states are the CC[77] methods and the ADC[52,78] methods. CC methods recover correlation by using a HF reference wavefunction and constructing a transformed Hamiltonian $\bar{H} = e^{-T}H\,e^{T}$, where the cluster operator T is then truncated, much like CIS, to allow for approximate methods. CCSD includes singly and doubly excited determinants in the CC operator. We consider the EOM-CCSD[54] and similarity transformed equations of motion coupled cluster (STEOM-CCSD) variants. STEOM-CCSD[55] performs another similarity transformation to the Hamiltonian in addition to the exponential operator transformation, which strongly reduces the coupling between singly excited determinants and more highly excited configurations. EEs are then extracted by diagonalizing the twice transformed Hamiltonian in the space of single excitations only.

ADC(n)[52] is a series of size-extensive excited-state methods based on perturbation theory. They rely on the two-particle Green's function, specifically the propagator function, which has a Fourier transform. Essentially, ADC(n) methods extend the MPn formalism into excited states. ADC(3)[31,56,57] was used to benchmark the remaining methods described here, as discussed above. Although ideally, we would prefer to include triple excitations to CCSD, this is not possible for these systems. The scaling of ADC(3) and previous performance makes it the best choice as our benchmark, as has been shown in previous work.[58−62]

CASSCF is used to treat static correlation by choosing an active space and optimizing simultaneously the orbitals and the coefficients of the expansion over Slater determinants or configuration state functions. NEVPT2 is based on the CASSCF wavefunction and provides additional dynamical correlation perturbatively up to doubly excited configurations using a Dyall Hamiltonian which includes all excitations both into and out of the orbitals in the active space.[64]

These methods were chosen because they are used extensively to study organic molecules such as the nucleobases and CT states. In particular, ADC(2) has been used extensively in our group to examine CT states in oligonucleotides,[13,14,73,74] so a better understanding of their performance is important.

Figures 4 and 5 show how the rest of the methods compare with ADC(3) for CT EEs and $\Delta\mu$, respectively. It is expected that CIS and CASSCF over-estimate all CT EE. Usually, after a treatment of perturbation theory, we expect these energies to decrease. Evidently, this is the case for both TG and GT orientations, where CIS and CASSCF both overestimate EE when compared to ADC(3) by as much as 2 eV, while CIS(D) and NEVPT2 underestimate them by 0.1−0.6 eV (except one case in NEVPT2(12,12) which is slightly higher than ADC(3)). We used three different active spaces in CASSCF,

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

and this allows us to examine the sensitivity of the results to the active space. The CASSCF energies change by 0.17 eV in TG and 0.42 eV in GT when the active space changes, while the NEVPT2 results are varied by 0.29 eV in TG and 0.46 eV in GT. The results are more sensitive in GT, likely because of the increased mixing of orbitals in this case when the bases are closer to each other. One has to be careful in designing the active space in order to avoid greater variations in the results. We have designed the active spaces to be similar in the two systems, as discussed in Supporting Information. This was more challenging for GT because the closer proximity of the nucleobases lead to increased mixing between the orbitals.

EOM-CCSD overestimates the CT states by about 0.3−0.4 eV compared to ADC(3), while ADC(2) underestimates them by about 0.6 eV. ADC(2), STEOM-CCSD, CIS(D), and NEVPT2 all behave similarly in describing the energies and predict CT states about 0.1−0.6 eV below ADC(3). Recent benchmarks on different systems are consistent with our results, and they have shown that ADC(2) underestimates CT states by 0.36 ± 0.24 eV, whereas EOM-CCSD overestimates them 0.30 ± 0.08 eV when compared to CCSDT-3.[25]

Describing $\Delta\mu$ is more limited because several of the methods we used calculate the properties at a lower level than the energies. NEVPT2 properties use CASSCF wavefunctions, so the dipoles are at the CASSCF level. Similarly, STEOM-CCSD and CIS(D) are at the CIS level. ADC(3), ADC(2), EOM-CCSD, and CASSCF predict the magnitude of $\Delta\mu$ for TG to be larger than that of GT as we would expect, while CIS predicts the opposite. $\Delta\mu$ at the CASSCF level are much larger than ADC(3), but they agree on the qualitative trend. The ratio of $\Delta\mu$ for TG over $\Delta\mu$ for GT is between 1.4 and 2.0 for ADC(3), EOM-CCSD, and CASSCF. ADC(2) predicts a larger ratio of 2.6. Although ADC(3) has more dynamical correlation, the multireference nature of CASSCF is expected to be important when the wavefunction and properties of each state are considered. Dipole as an one-electron property will depend less on dynamical correlation, but it will be sensitive on the correct representation of the different states. Thus CASSCF should provide an accurate description of the dipoles because of the optimization of the active space for both the excited states and ground state. The importance of static correlation on the description of $\Delta\mu$ is obvious when comparing CIS and CASSCF. While the energies obtained by these two methods were similar, the dipoles are very different, showing the sensitivity of the wavefunction on the mixing.

The calculated amount of CT can be used as a metric for CT state character description. For TG, CASSCF(12,12) describes almost pure CT with values 0.95 e^-. The single reference wavefunction-based methods all predict CT amounts which are lower with values below 0.61 e^- except for EOM-CCSD which has a value of 0.74 e^-. In the GT structure both CASSCF and SRM give similar values (between 0.6 and 0.9 e^-), except for ADC(2) which has a CT value of only 0.32 e^-.

DFT Methods. DFT methods, as opposed to wavefunction methods, have had their shortcomings with respect to CT states described extensively in the literature. The core of the issue stems from the well-known SIE. The SIE leads to not only drastically underestimated CT EE but also to a spurious near continuum of predicted CT states when considering explicitly solvent molecules around a chromophore.[18−20] Additionally, CT states in DFT often fail to reproduce the expected $1/R$ potential energy surface along the separation of fragments.[79]

The failure of TD-DFT, when the adiabatic approximation and approximate exchange−correlation density functionals are used, which has been reviewed by Dreuw and Head-Gordon,[18] can be understood from the eigenvalue equation used to solve for the EEs, ΔE.

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^* & \mathbf{A}^* \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} = \Delta E \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \end{pmatrix} \tag{1}$$

The matrix elements in their general form are given by

$$A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) + (ia|jb) - c_{\mathrm{HF}}(ij|ab) + (1 - c_{\mathrm{HF}})(ia|f_{\mathrm{xc}}|jb) \tag{2}$$

$$B_{ia,jb} = (ia|bj) - c_{\mathrm{HF}}(ib|aj) + (1 - c_{\mathrm{HF}})(ia|f_{\mathrm{xc}}|bj) \tag{3}$$

where i, j are ground-state occupied orbitals, a, b are virtual orbitals, ϵ_a, ϵ_i represent orbital energies of single-electron orbitals, and f_{xc} is the kernel of the exchange−correlation functional. The two-electron integrals are given in Mulliken notation, while the spin component is neglected here. From these equations, several methods can be derived. In the limiting cases, where $c_{\mathrm{HF}} = 1$, the equations reduce to TD-HF. By neglecting the **B** matrix as well, we get the CIS equations. When we neglect the **B** matrix in TD-DFT, it results in the commonly implemented Tamm−Dancoff approximation to TD-DFT.

In the simplest case of a CT of two fragments, X and Y, that are very far away from each other so that there is zero overlapping orbitals, an electron in an occupied orbital of fragment X is moved to a virtual orbital in fragment Y. In that case, all integrals with orthogonal orbitals will vanish, the **B** block will vanish, and the expression for the **A** matrix elements is $A_{ia,jb} = \delta_{ij}\delta_{ab}(\epsilon_a - \epsilon_i) - c_{\mathrm{HF}}(ij|ab)$. In the case of a pure functional with $c_{\mathrm{HF}} = 0$, the only remaining terms in **A** will be the difference in energy of ϵ_a and ϵ_i, which is the main contributor to the EE. This difference is at the heart of the problem of the SIE because ϵ_a is the energy that takes into account all Coulombic interactions, including with orbital i, which becomes unoccupied in the CT state. That means that the transferred electron in orbital a experiences the electrostatic repulsion with itself still being in orbital i. This self-interaction effect is canceled in TD-HF by the HF exchange term. Hybrid DFT functionals can also cancel some of the SIE and recuperate some of the $1/R$ dependence by subtracting the exact exchange correction $c_{\mathrm{HF}}(ij|ab)$ and eliminating the DFT exchange dependence on f_{xc}, but this is not enough. Global hybrid functionals include some HF exchange, but they do not describe CT states well.

In order to improve the performance of TD-DFT to describe CT states, more recent functionals introduce exact exchange into the functional form via range separation. This has led to a new brand of functionals known as range separated hybrid functionals (RSH), which work by splitting the Coulombic operator using the error function erf (and error complementary function, erfc) into a short range (SR) and long range (LR) part, with an added ω parameter that dictates the distance at which the SR and LR exchange terms are activated or deactivated

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

$$\frac{1}{r_{12}} = \frac{\mathrm{erfc}(\omega r_{12})}{r_{12}} + \frac{\mathrm{erf}(\omega r_{12})}{r_{12}} \quad (4)$$

The first term on the right hand describes the SR and the second the LR exchange included. The left term decays rapidly as $1/\omega$. As with many DFT methods, the lack of an exact f_{xc} functional has led to several variants in adding different amounts of HF exchange with different parameterization. When the contribution of the right term is set to 1, these are LRC RSH functionals. In tuned range-separated hybrid functionals, the ω parameter is tuned for the system of interest, leading to more specific functionals but increasing their cost.[30]

In this work, we have separated the list of functionals used into two groups: methods 13−19 are hybrid functionals which are not LRC. They are global hybrid functionals except for N12-SX and MN12-SX, which are range separated but include 0% long range HF exchange (so they use screened exchange). Methods 20−40 are LRC RSH functionals. Methods 20−40 are arranged in order of increasing ω and SR exchange, that is, when the methods have the same ω parameter, they are arranged in terms of increasing SR HF exchange. Table 2 shows the exact ω parameter and SRE used in each functional.

In Figure 4, it is clear that the LRC functionals have a very different behavior compared to the non-LRC. All non-LRC underestimate the CT energies considerably, by almost 2 eV in most cases. Among them, the M06-2X and SOGGA11X are performing somewhat better. The LRC functionals on the other hand perform much better. The energy of the CT state increases as ω and as SRE increase. For the functionals at the end of the list where SRE is not present and $\omega > 0.4$, the CT states are actually overestimated when compared to ADC(3). The functionals 24, 27−33 in both TG and GT orientations perform the best when compared to ADC(3). These are mostly the ωB97-X variants, with ω around 0.3. M11 is also performing very well (with $\omega = 0.25$). Figure 6 also clearly shows the trend of systematically blue-shifted EEs as ω increases, and it highlights that for all functionals, the energy of TG is about 1 eV blue-shifted compared to that of GT.

The dipole moments exhibit a very different behavior. The $\Delta\mu$ magnitudes for most of the non-LRC hybrid functionals generally agree qualitatively with wavefunction methods. The ratio of $\Delta\mu$ for TG over GT has similar values to the wavefunction methods, with values between 1.3 and 1.9 in most cases of hybrid non-LRC functionals, exceptions being the PBE0 and SOGGA11-X functionals. The agreement with CASSCF is very good throughout the non-LRC hybrids for the GT geometry where there is less separation between the bases than compared to TG. The behavior of LRC functionals is quite interesting, and different from the wavefunction methods. $\Delta\mu$ for the LRC functionals in the TG orientation shows a decreasing trend as ω increases, whereas GT results show generally the opposite behavior. The functionals with the largest ω ($\omega > 0.32$) show a distinct different behavior compared to the previous ones, which is obvious in the bottom panel of Figure 6. While functionals 20−33 in the list predict $\Delta\mu$ for TG to be more than twice as large as that of GT (often 20 debye vs 8 debye and a ratio between 2 and 4), the functionals 34−40 ($\omega > 0.32$) predict similar results for $\Delta\mu$ in both orientations (ratio TG/GT 0.8−1.2) and the magnitudes match those from CIS. The inverse trend is a result of increasing addition of HF exchange at shorter distances, as more HF exchange is included in the operator, $\Delta\mu$ resembles

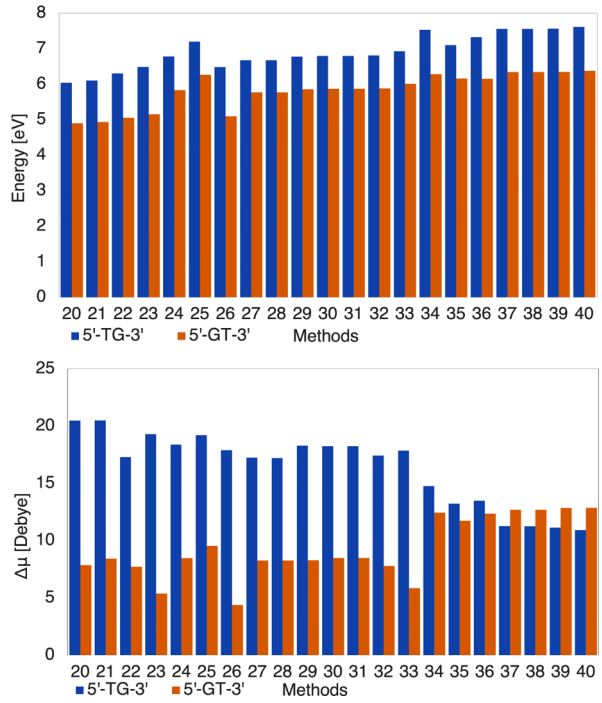


Figure 6. Results for TD-DFT LRC functionals ordered based on ω: EEs (top) and $\Delta\mu$ (bottom).

the CIS dipole. It seems that for functionals 34−40, the higher degree of HF exchange leads to worse performance for the dipole moments. We expect that the $\Delta\mu$ for TG is larger than that of GT based on distance. However, this is not reproduced by these functionals. They overestimate the amount of charge transferred when the bases are close to each other. In that case, orbital overlap should also be important, and this may not be described well by them. Overall, inclusion of HF exchange should be kept at a reasonable level if we want a proper description of the physical properties.

To that end, we found that the best LRC functionals that fit both the criteria of reproducing the 1 eV gap between conformations follow the CASSCF dipole trend and have similar energy to an ADC(3) level are M11, ωM06-D3, and ωB97 variants. The optimum value of ω based on these functionals is 0.3.

We have also tried to see if dispersion corrections have any effect on the CT energies and dipole moments. Previous work had shown that dispersion corrected functionals are accurate for valence states, but can be inaccurate for excitations to Rydberg states, or more generally states where the nature of the electron density is significantly different from the ground state.[80] According to our results, the functionals that include dispersion do not show any noticeable differences compared to the other ones. The ω parameter seems to dominate the behavior of CT states.

Local Excited States. Ideally, in order for electronic structure methods to be useful in photophysics and photochemistry, they should give a balanced description between CT and localized excited states. In this section, we examine the behavior of local excited states and how the methods perform compared to ADC(3). We chose the first $\pi\pi^*$ excited state localized in guanine in the TG dimer as a representative local state. Figure 7 depicts the EEs for the first bright transition localized in guanine base for all methods considered when compared to ADC(3). Experimentally, there are several

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

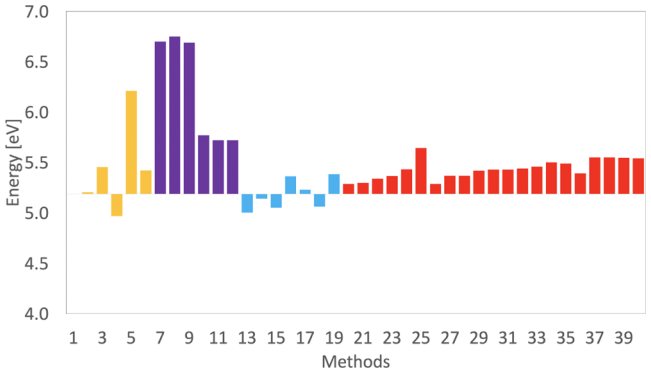


Figure 7. Energies of the first valence $\pi\pi^*$ excited state localized in guanine in TG obtained using all methods and compared to the ADC(3) value.

tautomers of guanine present in the gas phase, so it is hard to determine the correct experimental value. A very accurate EOM-CCSD(T)/aug-cc-pVTZ calculation predicts the EE to be 4.86 eV.[15] All the methods, including ADC(3) overestimate the EE. Our aim in this work is not to predict the best value for this property, and because we have used a much smaller basis set, it is expected that we cannot reproduce the best value. Our aim is rather to find the range of predicted values in the set of methods used. The EE at the ADC(3) level is 5.21 eV. The typical hybrid non-LRC functionals show the best results. STEOM-CCSD also shows very good values. The non-LRC functionals are within 0.2 eV of the ADC(3) energy, whereas LRC functionals overestimate the energy by 0.1−0.5 eV. While the LRC functionals improve the performance of CT states, they do so at the expense of the local valence excited states, so the balance between all types of excited states is hard to achieve. Among the wavefunction-based methods, ADC(2) agrees the best with ADC(3), followed by the other SRM except CIS. CIS and CASSCF perform the worse, as expected. NEVPT2 corrects much of the overestimated energies, but it is still about 0.5 eV too high.

Overall, the errors and deviations between the methods for the valence state are much smaller than the errors we saw above for the CT states. We use Table 3 to highlight this difference in performance. In that table, we use the ranges of EE of all the methods to compare the descriptions of $\pi\pi^*$ states to that of the CT states. The CT states show a much larger degree of variability. The range from the highest energy to the lowest energy predicted for CT states for all methods considered is around 3.03 and 2.37 eV, when excluding CIS and CASSCF energies, for TG and GT, respectively. Yet, for $\pi\pi^*$, the same ranges are 0.80 eV for TG and 0.93 eV for GT (again excluding CIS and CASSCF). The range of values for the CT states when excluding CIS, CASSCF, and non-LRC hybrid functionals reduces to 1.57 and 1.48 eV for TG and GT, respectively, which means the range for CT states is almost halved, another strong indication that the non-LRC functionals are incorrect to describe CT. The improvement of the CT state energies between DFT methods when excluding the non-LRC functionals provides a strong case for the use of balanced LRC functionals when describing CT states.

Table 3. Comparison between the Range of EE of the First CT State and First Bright Guanine Transition throughout All the Methods Excluding CIS and CASSCF Only (Top)[a]

	CT EE [eV]			$\pi\pi^*$ EE [eV]		
orientation	min	max	range	min	max	range
Excluding CIS, CASSCF						
TG	4.58	7.61	3.03	4.97	5.77	0.80
GT	4.01	6.38	2.37	4.84	5.77	0.93
Excluding CIS, CASSCF, Non-LRC Functionals						
TG	6.04	7.61	1.57	4.97	5.77	0.80
GT	4.90	6.38	1.48	4.84	5.77	0.93

[a]Same as top but excluding poor performing functionals (bottom).

Finally, we have calculated the EE of monomer guanine in order to get information about the shift of the local EE of the monomer when it is π stacked with thymine. The results are given in Supporting Information. The shifts for TG are very small, almost negligible red shifts predicted by all methods. The shifts for GT are larger, and this is expected because the bases are closer. They are close to 0.1 eV for the SRM, the non-LRC functionals and for the LRC with very large values of ω. However, they show different behavior for the LRC functionals with intermediate values of ω, where they often predict a blue shift.

Basis Set Effects. To gauge the effects of the basis set size, we calculated both the lowest lying CT states for both conformers, and the first bright state for guanine in the TG conformation in two larger basis sets for M11 and a larger basis set for ADC(2). The results are shown in Table 4. The overall effect of increasing the basis set is a stabilization of the CT state, for both ADC(2) and M11. This effect is also seen for the guanine bright state. When going from double zeta to triple zeta, the CT states are stabilized by around 0.2 eV for ADC(2) and around 0.1 eV for M11. Furthermore, increasing to the quadruple zeta case in M11 further stabilized it by another 0.05 eV. This is almost exactly reproduced in the bright state, where it is first stabilized by around 0.1 eV from double to triple zeta and then 0.05 eV from triple to quadruple zeta. Given that the effect of increasing the size of the basis set shows the same degree of change to both the bright local and the CT state, this suggests that the basis set will not produce an imbalanced description between the two types of states. We can also see that the effect of increasing the basis sets is not the same for different methods, so a careful basis set dependence should be specific to the electronic structure used.

The magnitudes $\Delta\mu$ are also shown in Table 4 for the different basis sets. $\Delta\mu$ obtained using M11 is relatively stable with respect to basis set size. The values change by about 1 debye as the basis set increases, although the change is not monotonic, so convergence is not reached. The behavior of ADC(2) is even more problematic. The values of $\Delta\mu$ change by 4−5 debye when the basis set increases from DZ to TZ. This change highlights that the dipole representing the wavefunction and degree of CT is very sensitive to basis set size, much more than the effect of the energies.

CONCLUSIONS

CT states in π-stacked complexes of thymine and guanine have been studied with a variety of both wavefunction-based methods and DFT. The EE from all methods were compared to those of ADC(3) which served as the benchmark method. The errors of the various methods in reproducing CT states were also compared to errors in describing local $\pi\pi^*$ states. A general feature reproduced by all methods in this work is that distance and overlap between the π-stacked nucleobases

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

Table 4. Effect of Larger Basis Sets for Wavefunction Method ADC(2) and Functional M11[a]

method	basis set	5′-TG-3′ E/eV	5′-GT-3′ E/eV	G $\pi\pi^*$ (TG) E/eV	5′-TG-3′ $\Delta\mu$/debye	5′-GT-3′ $\Delta\mu$/debye
ADC(2)	DZ	6.36	5.30	5.21	13.1	4.2
ADC(2)	TZ	6.10	5.09	5.05	17.4	9.2
M11	DZ	6.78	5.84	5.43	18.4	8.5
M11	TZ	6.68	5.76	5.34	19.4	10.1
M11	QZ	6.64	5.71	5.30	18.3	9.4

[a]EEs and $\Delta\mu$ are shown.

generate a blue shift in CT EE when going from the GT to the TG orientation.

Moreover, wavefunction-based methods provide a more general picture: all SRM except CIS and EOM-CCSD underestimate the CT states, whereas EOM-CCSD overestimates it by 0.3−0.4 eV and CIS by 1.4 eV. ADC(2) underestimates the CT energies by 0.6 eV compared to ADC(3). CIS(D) and STEOM-CCSD compare similarly to each other and predict energies red-shifted by about 0.1−0.4 eV compared to the benchmark ADC(3). In the case of the MRMs, increasing the active space does not seem to systematically improve results by recovery of static correlation, while addition of dynamic correlation with the NEVPT2 method leads to much more accurate energies. NEVPT2 in most cases underestimates the CT energies by 0.1−0.6 eV, while in one case, it slightly overestimates it by 0.1 eV. For wavefunction methods, the errors of the valence local $\pi\pi^*$ states are similar in magnitude to those of the CT states, but in most cases, the energies are overestimated compared to ADC(3) while the energies of CT states are underestimated. Exceptions are EOM-CCSD and STEOM-CCSD which have similar errors for the CT and local states, making them good choices for a balanced description, with EOM-CCSD having the most balanced description.

Given the ubiquity and diversity of DFT methods, special consideration was taken when analyzing the parameters that give rise to the LRC flavor of functionals, namely, the ω parameter and inclusion of SR HF exchange. The trend clearly shows that as ω increases, the energies become more blue-shifted. This is the case in both the local bright states and the CT states studied here. However, with the proper functional, one can obtain near ADC(3) and MRM level results at a fraction of the computational cost. We found that the best-performing functionals are the following: M11, ωM06-D3, and several ωB97 variants because they reproduce the CT energetics and reproduce the CASSCF dipole magnitudes and quantitative CT. The optimum ω parameter for these systems is around 0.3. The non-LRC functionals perform better than the LRC functionals for the valence local $\pi\pi^*$ state, but they have very large errors for the CT states, so they cannot provide a balanced description.

In addition to the energies, we have also examined the $\Delta\mu$ vector which gives an estimate of the amount of charge transferred during excitation to a CT state. $\Delta\mu$ shows much more variability to the choice of both wavefunction and DFT methods. Most wavefunction-based methods predict $\Delta\mu$ for TG to be larger than that of GT, and their ratio is between 1.3 and 1.9, while ADC(2) is particularly problematic for the GT dimer predicting a very small $\Delta\mu$. On the other hand, CIS shows the opposite trend, with a larger $\Delta\mu$ magnitude for GT compared to TG. Hybrid non-LRC DFT functionals reproduce the trend of a greater TG dipole magnitude compared to the GT dipole magnitude. The LRC functionals show an unpredictable behavior. For smaller values of ω, the ratio of TG versus GT $\Delta\mu$ is too large while when $\omega > 0.4$ the trend is reversed and the ratio is around 1, that is, $\Delta\mu$ is larger for GT or they are very similar. These results indicate that inclusion of more HF exchange can worsen the results because of less correlation from the DFT f_{xc} functional in the Coulombic operator, so there is a limit to the appropriate HF exchange correction that can be included.

We have finally examined the effect of basis sets on two methods, ADC(2) and M11. This effect is similar for CT and local states, so it does not lead to an imbalance between the two types of states. Overall, even though we only studied two geometries, the errors in the energies are found to be systematic between them, while the errors in the dipoles are more variant.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jctc.0c00973.

Details on MRMs; Bader analysis for charge distributions; results for all methods; additional results on the local excited states; and cartesian coordinates of the two geometries (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Spiridoula Matsika − *Department of Chemistry, Temple University, Philadelphia, Pennsylvania, United States;* orcid.org/0000-0003-2773-3979; Email: smatsika@temple.edu

Authors

Camilo Zuluaga − *Department of Chemistry, Temple University, Philadelphia, Pennsylvania, United States*

Vincent A. Spata − *Department of Chemistry, Temple University, Philadelphia, Pennsylvania, United States*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jctc.0c00973

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors acknowledge support by the National Science Foundation under grant CHE-1800171. This research includes calculations carried out on HPC resources supported in part by the National Science Foundation through major research instrumentation grant number 1625061 and by the US Army Research Laboratory under contract number W911NF-16-2−

0189. Computational work was also performed using the Extreme Science and Engineering Discovery Environment (XSEDE), which is supported by National Science Foundation grant no. ACI-1548562.

REFERENCES

(1) *Topics in Current Chemistry—Photoinduced Phenomena in Nucleic Acids I*; Ullrich, S., Borin, A. C., Barbatti, M., Eds.; Springer: Berlin, Heidelberg, 2015; Vol. *355*.
(2) *Topics in Current Chemistry—Photoinduced Phenomena in Nucleic Acids II*; Ullrich, S., Borin, A. C., Barbatti, M., Eds.; Springer: Berlin, Heidelberg, 2015; Vol. *356*.
(3) Bucher, D. B.; Schlueter, A.; Carell, T.; Zinth, W. Watson-Crick Base Pairing Controls Excited-State Decay in Natural DNA. *Angew. Chem., Int. Ed.* **2014**, *53*, 11366−11369.
(4) Middleton, C. T.; de La Harpe, K.; Su, C.; Law, Y. K.; Crespo-Hernández, C. E.; Kohler, B. DNA Excited-State Dynamics: From Single Bases to the Double Helix. *Annu. Rev. Phys. Chem.* **2009**, *60*, 217−239.
(5) Bucher, D. B.; Pilles, B. M.; Carell, T.; Zinth, W. Charge separation and charge delocalization identified in long-living states of photoexcited DNA. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 4369−4374.
(6) Schreier, W. J.; Gilch, P.; Zinth, W. Early Events of DNA Photodamage. *Annu. Rev. Phys. Chem.* **2015**, *66*, 497−519.
(7) Martinez-Fernandez, L.; Zhang, Y.; de La Harpe, K.; Beckstead, A. A.; Kohler, B.; Improta, R. Photoinduced long-lived charge transfer excited states in AT-DNA strands. *Phys. Chem. Chem. Phys.* **2016**, *18*, 21241−21245.
(8) Kufner, C.; Zinth, W.; Bucher, D. UV-Induced Charge-Transfer States in Short Guanosine-Containing DNA Oligonucleotides. *ChemBioChem* **2020**, *21*, 2306.
(9) O'Neil, L. L.; Grossfield, A.; Wiest, O. Base Flipping of the Thymine Dimer in Duplex DNA. *J. Phys. Chem. B* **2007**, *111*, 11843−11849.
(10) Zhang, M.; Wang, L.; Zhong, D. Photolyase: Dynamics and electron-transfer mechanisms of DNA repair. *Arch. Biochem. Biophys.* **2017**, *632*, 158−174. , Flavoproteins: Beyond the Classical Paradigms
(11) Zhang, Y.; Dood, J.; Beckstead, A. A.; Li, X.-B.; Nguyen, K. V.; Burrows, C. J.; Improta, R.; Kohler, B. Efficient UV-induced charge separation and recombination in an 8-oxoguanine-containing dinucleotide. *Proc. Natl. Acad. Sci. U.S.A.* **2014**, *111*, 11612−11617.
(12) Bucher, D. B.; Kufner, C. L.; Schlueter, A.; Carell, T.; Zinth, W. UV-Induced Charge Transfer States in DNA Promote Sequence Selective Self-Repair. *J. Am. Chem. Soc.* **2016**, *138*, 186−190.
(13) Lee, W.; Matsika, S. Role of charge transfer states into the formation of cyclobutane pyrimidine dimers in DNA. *Faraday Discuss.* **2019**, *216*, 507−519.
(14) Lee, W.; Matsika, S. QM/MM studies reveal pathways leading to the quenching of the formation of thymine dimer photoproduct by flanking bases. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9927−9935.
(15) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **2016**, *116*, 3540−3593.
(16) Dreuw, A.; Weisman, J. L.; Head-Gordon, M. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.* **2003**, *119*, 2943.
(17) Dreuw, A.; Head-Gordon, M. Failure of Time-Dependent Density Functional Theory for Long-Range Charge-Transfer Excited States: The Zincbacteriochlorin−Bacteriochlorin and Bacteriochlorophyll−Spheroidene Complexes. *J. Am. Chem. Soc.* **2004**, *126*, 4007−4016.
(18) Dreuw, A.; Head-Gordon, M. Single-Reference ab Initio Methods for the Calculation of Excited States of Large Molecules. *Chem. Rev.* **2005**, *105*, 4009−4037.
(19) Lange, A.; Herbert, J. M. Simple Methods To Reduce Charge-Transfer Contamination in Time-Dependent Density-Functional Calculations of Clusters and Liquids. *J. Chem. Theory Comput.* **2007**, *3*, 1680−1690.
(20) Isborn, C. M.; Mar, B. D.; Curchod, B. F. E.; Tavernelli, I.; Martínez, T. J. The Charge Transfer Problem in Density Functional Theory Calculations of Aqueously Solvated Molecules. *J. Phys. Chem. B* **2013**, *117*, 12189−12201.
(21) Zhao, Y.; Truhlar, D. G. Density Functional for Spectroscopy: No Long-Range Self-Interaction Error, Good Performance for Rydberg and Charge-Transfer States, and Better Performance on Average than B3LYP for Ground States. *J. Phys. Chem. A* **2006**, *110*, 13126−13130.
(22) Subotnik, J. E. Communication: Configuration interaction singles has a large systematic bias against charge-transfer states. *J. Chem. Phys.* **2011**, *135*, 071104.
(23) Dutta, A. K.; Nooijen, M.; Neese, F.; Izsák, R. Exploring the Accuracy of a Low Scaling Similarity Transformed Equation of Motion Method for Vertical Excitation Energies. *J. Chem. Theory Comput.* **2018**, *14*, 72−91.
(24) Mewes, S. A.; Plasser, F.; Krylov, A.; Dreuw, A. Benchmarking Excited-State Calculations Using Exciton Properties. *J. Chem. Theory Comput.* **2018**, *14*, 710−725.
(25) Kozma, B.; Tajti, A.; Demoulin, B.; Izsák, R.; Nooijen, M.; Szalay, P. G. A New Benchmark Set for Excitation Energy of Charge Transfer States: Systematic Investigation of Coupled Cluster Type Methods. *J. Chem. Theory Comput.* **2020**, *16*, 4213−4225.
(26) Zhang, Y.; Yang, W. A challenge for density functionals: Self-interaction error increases for systems with a noninteger number of electrons. *J. Chem. Phys.* **1998**, *109*, 2604−2608.
(27) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E. Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals. *J. Chem. Phys.* **2006**, *125*, 194112.
(28) Ruzsinszky, A.; Perdew, J. P.; Csonka, G. I.; Vydrov, O. A.; Scuseria, G. E. Spurious fractional charge on dissociated atoms: Pervasive and resilient self-interaction error of common density functionals. *J. Chem. Phys.* **2006**, *125*, 194112.
(29) Dombroski, J. P.; Taylor, S. W.; Gill, P. M. W. KWIK: Coulomb Energies inO(N) Work. *J. Phys. Chem.* **1996**, *100*, 6272−6276.
(30) Kümmel, S. Charge-Transfer Excitations: A Challenge for Time-Dependent Density Functional Theory That Has Been Met. *Adv. Energy Mater.* **2017**, *7*, 1700440.
(31) Harbach, P. H. P.; Wormit, M.; Dreuw, A. The third-order algebraic diagrammatic construction method (ADC(3)) for the polarization propagator for closed-shell molecules: Efficient implementation and benchmarking. *J. Chem. Phys.* **2014**, *141*, 064113.
(32) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. UCSF Chimera?A visualization system for exploratory research and analysis. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(33) Rohrdanz, M. A.; Martins, K. M.; Herbert, J. M. A long-range-corrected density functional that performs well for both ground-state properties and time-dependent density functional theory excitation energies, including charge-transfer excited states. *J. Chem. Phys.* **2009**, *130*, 054112.
(34) Chai, J.-D.; Head-Gordon, M. Long-range corrected hybrid density functionals with damped atom-atom dispersion corrections. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(35) Lin, Y.-S.; Tsai, C.-W.; Li, G.-D.; Chai, J.-D. Long-range corrected hybrid meta-generalized-gradient approximations with dispersion corrections. *J. Chem. Phys.* **2012**, *136*, 154109.
(36) Lin, Y.-S.; Li, G.-D.; Mao, S.-P.; Chai, J.-D. Long-Range Corrected Hybrid Density Functionals with Improved Dispersion Corrections. *J. Chem. Theory Comput.* **2013**, *9*, 263−272.
(37) Peverati, R.; Truhlar, D. G. Improving the Accuracy of Hybrid Meta-GGA Density Functionals by Range Separation. *J. Phys. Chem. Lett.* **2011**, *2*, 2810−2817.

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

(38) Verma, P.; Bartlett, R. J. Increasing the applicability of density functional theory. IV. Consequences of ionization-potential improved exchange-correlation potentials. *J. Chem. Phys.* **2014**, *140*, 18A534.
(39) Rohrdanz, M. A.; Herbert, J. M. Simultaneous benchmarking of ground- and excited-state properties with long-range-corrected density functional theory. *J. Chem. Phys.* **2008**, *129*, 034107.
(40) Mardirossian, N.; Head-Gordon, M. ωB97M-V: A combinatorially optimized, range-separated hybrid, meta-GGA density functional with VV10 nonlocal correlation. *J. Chem. Phys.* **2016**, *144*, 214110.
(41) Mardirossian, N.; Ruiz Pestana, L.; Womack, J. C.; Skylaris, C.-K.; Head-Gordon, T.; Head-Gordon, M. Use of the rVV10 Nonlocal Correlation Functional in the B97M-V Density Functional: Defining B97M-rV and Related Functionals. *J. Phys. Chem. Lett.* **2017**, *8*, 35−40.
(42) Chai, J.-D.; Head-Gordon, M. Systematic optimization of long-range corrected hybrid density functionals. *J. Chem. Phys.* **2008**, *128*, 084106.
(43) Mardirossian, N.; Head-Gordon, M. ωB97X-V: A 10-parameter, range-separated hybrid, generalized gradient approximation density functional with nonlocal correlation, designed by a survival-of-the-fittest strategy. *Phys. Chem. Chem. Phys.* **2014**, *16*, 9904−9924.
(44) Jin, Y.; Bartlett, R. J. The QTP family of consistent functionals and potentials in Kohn-Sham density functional theory. *J. Chem. Phys.* **2016**, *145*, 034107.
(45) Cohen, A. J.; Mori-Sánchez, P.; Yang, W. Development of exchange-correlation functionals with minimal many-electron self-interaction error. *J. Chem. Phys.* **2007**, *126*, 191109.
(46) Verma, P.; Wang, Y.; Ghosh, S.; He, X.; Truhlar, D. G. Revised M11 Exchange-Correlation Functional for Electronic Excitation Energies and Ground-State Properties. *J. Phys. Chem. A* **2019**, *123*, 2966−2990.
(47) Vydrov, O. A.; Van Voorhis, T. Nonlocal van der Waals density functional: The simpler the better. *J. Chem. Phys.* **2010**, *133*, 244103.
(48) Weintraub, E.; Henderson, T. M.; Scuseria, G. E. Long-Range-Corrected Hybrids Based on a New Model Exchange Hole. *J. Chem. Theory Comput.* **2009**, *5*, 754−762.
(49) Song, J.-W.; Hirosawa, T.; Tsuneda, T.; Hirao, K. Long-range corrected density functional calculations of chemical reactions: Redetermination of parameter. *J. Chem. Phys.* **2007**, *126*, 154105.
(50) Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135−149.
(51) Head-Gordon, M.; Rico, R. J.; Oumi, M.; Lee, T. J. A doubles correction to electronic excited states from configuration interaction in the space of single substitutions. *Chem. Phys. Lett.* **1994**, *219*, 21−29.
(52) Dreuw, A.; Wormit, M. The Algebraic Diagrammatic Construction Scheme for the Polarization Propagator for the Calculation of Excited States. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2015**, *5*, 82−95.
(53) Stanton, J. F.; Bartlett, R. J. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.* **1993**, *98*, 7029−7039.
(54) Krylov, A. I. Equation-of-Motion Coupled-Cluster Methods for Open-Shell and Electronically Excited Species: The Hitchhiker's Guide to Fock Space. *Annu. Rev. Phys. Chem.* **2008**, *59*, 433.
(55) Nooijen, M.; Bartlett, R. J. A new method for excited states: Similarity transformed equation-of-motion coupled-cluster theory. *J. Chem. Phys.* **1997**, *106*, 6441−6448.
(56) Trofimov, A. B.; Stelter, G.; Schirmer, J. A consistent third-order propagator method for electronic excitation. *J. Chem. Phys.* **1999**, *111*, 9982−9999.
(57) Trofimov, A. B.; Stelter, G.; Schirmer, J. Electron excitation energies using a consistent third-order propagator approach: Comparison with full configuration interaction and coupled cluster results. *J. Chem. Phys.* **2002**, *117*, 6402−6410.
(58) Holland, D. M. P.; Seddon, E. A.; Trofimov, A. B.; Gromov, E. V.; Wormit, M.; Dreuw, A.; Korona, T.; de Oliveira, N.; Archer, L. E.; Joyeux, D. A study of the excited electronic states of normal and fully deuterated furan by photoabsorption spectroscopy and high-level ab initio calculations. *J. Mol. Spectrosc.* **2015**, *315*, 184−195.
(59) Plasser, F.; Dreuw, A. High-Level Ab Initio Computations of the Absorption Spectra of Organic Iridium Complexes. *J. Phys. Chem. A* **2015**, *119*, 1023−1036.
(60) Prlj, A.; Sandoval-Salinas, M. E.; Casanova, D.; Jacquemin, D.; Corminboeuf, C. Low-Lying ππ* States of Heteroaromatic Molecules: A Challenge for Excited State Methods. *J. Chem. Theory Comput.* **2016**, *12*, 2652−2660.
(61) Mewes, S. A.; Mewes, J.-M.; Dreuw, A.; Plasser, F. Excitons in Poly(para phenylene vinylene): a Quantum-Chemical Perspective Based on High-Level ab initio Calculations. *Phys. Chem. Chem. Phys.* **2016**, *18*, 2548−2563.
(62) Azarias, C.; Habert, C.; Budzák, Š.; Blase, X.; Duchemin, I.; Jacquemin, D. Calculations of n→π* Transition Energies: Comparisons Between TD-DFT, ADC, CC, CASPT2, and BSE/GW Descriptions. *J. Phys. Chem. A* **2017**, *121*, 6122−6134.
(63) Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. M. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.* **1980**, *48*, 157−173.
(64) Schapiro, I.; Sivalingam, K.; Neese, F. Assessment of n-Electron Valence State Perturbation Theory for Vertical Excitation Energies. *J. Chem. Theory Comput.* **2013**, *9*, 3567−3580.
(65) Hättig, C. Response Theory and Molecular Properties (A Tribute to Jan Linderberg and Poul Jørgensen); *Advances in Quantum Chemistry*; Jensen, H., Ed.; Academic Press, 2005; Vol. *50*; pp 37−60.
(66) *TURBOMOLE V7.0 2015*, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989−2007; TURBOMOLE GmbH, since 2007. available from http://www.turbomole.com.
(67) Hättig, C.; Hellweg, A.; Köhn, A. Distributed memory parallel implementation of energies and gradients for second-order Møller-Plesset perturbation theory with the resolution-of-the-identity approximation. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1159−1169.
(68) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775.
(69) Bublitz, G. U.; Boxer, S. G. STARK SPECTROSCOPY: Applications in Chemistry, Biology, and Materials Science. *Annu. Rev. Phys. Chem.* **1997**, *48*, 213−242.
(70) Shao, Y.; et al. Advances in molecular quantum chemistry contained in the Q-Chem 4 program package. *Mol. Phys.* **2015**, *113*, 184−215.
(71) Frisch, M. J.; et al. *Gaussian 16* Revision C.01; Gaussian Inc: Wallingford CT, 2016.
(72) Neese, F. The ORCA program system. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 73−78.
(73) Spata, V. A.; Matsika, S. Photophysical deactivation pathways in adenine oligonucleotides. *Phys. Chem. Chem. Phys.* **2015**, *17*, 31073−31083.
(74) Spata, V. A.; Matsika, S. Role of Excitonic Coupling and Charge-Transfer States in the Absorption and CD Spectra of Adenine-Based Oligonucleotides Investigated through QM/MM Simulations. *J. Phys. Chem. A* **2014**, *118*, 12021−12030.
(75) Tang, W.; Sanville, E.; Henkelman, G. A Grid-Based Bader Analysis Algorithm without Lattice Bias. *J. Phys.: Condens. Matter* **2009**, *21*, 084204.
(76) Yu, M.; Trinkle, D. R. Accurate and Efficient Algorithm for Bader Charge Integration. *J. Chem. Phys.* **2011**, *134*, 064111.
(77) Bartlett, R. J. Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry. *J. Phys. Chem.* **1989**, *93*, 1697−1708.
(78) Schirmer, J. Beyond the random-phase approximation: A new approximation scheme for the polarization propagator. *Phys. Rev. A* **1982**, *26*, 2395−2416.

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387

(79) Kozak, C. R.; Kistler, K. A.; Lu, Z.; Matsika, S. Excited-State Energies and Electronic Couplings of DNA Base Dimers. *J. Phys. Chem. B* **2010**, *114*, 1674−1683.
(80) Briggs, E. A.; Besley, N. A. Modelling excited states of weakly bound complexes with density functional theory. *Phys. Chem. Chem. Phys.* **2014**, *16*, 14455−14462.

https://dx.doi.org/10.1021/acs.jctc.0c00973
J. Chem. Theory Comput. 2021, 17, 376−387