PCCP

Accepted Manuscript

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/pccp

Applicability of Optimal Functional Tuning in Density Functional Calculations of Ionization Potentials and Electron Affinities of Adenine-Thymine Nucleobase Pairs and Clusters

Haitao Sun*, Shian Zhang and Zhenrong Sun*

State Key Laboratory of Precision Spectroscopy, Department of Physics,
East China Normal University
Shanghai 200062, P. R. China

Corresponding Author E-mail: htsun@phy.ecnu.edu.cn ; zrsun@phy.ecnu.edu.cn

Abstract

The intrinsic properties of DNA and RNA nucleic acid bases (NABs) such as ionization potentials (IPs) and electron affinities (EAs) are crucial to reveal various biochemical mechanisms. Successful applications of density functional theory (DFT) using nonempirically tuned long-range corrected (LC) functionals for calculations of vertical ionization potentials (VIPs) and electron affinities (VEAs) of various adenine-thymine (AT) nucleobase pairs and clusters are demonstrated. We employ tuning method by enforcing an asymptotically correct exchange-correlation potential adjusted to give frontier orbital energies ($-\varepsilon_{HOMO}$ and $-\varepsilon_{LUMO}$) representing IPs and EAs and assess the quality of prediction which are comparable to high-level EOM-IP-CCSD/CCSD methods. The delocalization error by different DFT functionals is quantified by calculations with fractional electron numbers. The cooperative effect of H-bonding and π-stacking on the IPs of AT clusters, as well as the reactivity parameters (global hardness and electrophilicity), are quantitatively characterized using the tuned LC functionals. The present work aims to provide a reliable and efficient theoretical tool for the prediction of the related electron donor and acceptor abilities of the NAB systems.

1. Introduction

It is well-known that DNA and RNA nucleic acid bases (NABs), containing the fundamental biological importance, have attracted both theoretical and experimental interest in life science,[1-4] biochemical and medical technologies[5] and NABs-based materials for nanoelectronics.[6] Ionization potentials (IPs) and electron affinities (EAs) are the important intrinsic properties of NABs allowing a deep understanding of the specific mechanisms, such as various radiation-induced phenomena of genetic materials,[7, 8] charge-transfer process along the DNA strand,[9] and reactivity and aromaticity of NABs in an external electric field.[10] The reliable experimental determination of IPs and EAs of NABs is relatively limited due to the various base pairing and stacking models,[11, 12] complicated surrounding environment and tautomerization in the gas phase.[13] The uncertainties in experimental determination of EAs are within a broad range from negative to positive values, extending up to several electron volts (eVs).[14] Recently, numerous theoretical studies based on high-level wave-function-theory methods (WFT) could produce reasonably accurate IPs and EAs with respect to experimental measurements. Serrano-Andrés et al. studied the vertical and adiabatic IPs and EAs of individual DNA bases using the multi-configurational perturbation methods (CASPT2)[14, 15] to guide the assignment of features in the experimental photoelectron spectra. Krylov et al. calculated the ionization energies of various isomers of NAB dimers by the equation-of-motion coupled-cluster method (for ionization potential) with single and double substitutions (EOM-IP-CCSD).[11, 12, 16] Pal et al. reported a benchmark investigation of electron affinities of individual NABs with an analogue of EOM-EA-CCSD.[17] However, the expensive computational expense of those WFT methods greatly limits the size of systems that can be explored and a potential long DNA strand is apparently impossible to deal with. Therefore, there is a great need to establish reliable, instructive, and

computationally efficient theoretical tools to describe the related electron donor and acceptor abilities of the NABs. For most application scenarios, density functional theory (DFT) can offer a desired compromise between accuracy of the results and computational cost for large geometries.[18-22] However, the high dependence on the exchange-correlation functional makes it a difficult task to choose the best functional.[23-26] Moreover, many of the available studies also confirm that the predictive power of DFT in NAB area leaves a lot to be desired.[16, 26] The most frequently applied functionals, such as generalized gradient approximations (GGA) and global hybrid GGA functionals, affording an incorrect asymptotic behavior of the potential, large delocalization error (DE),[27] self-interaction error (SIE),[28, 29] and lack of derivative discontinuity (DD),[30, 31] tend to overestimate delocalization effect of the hole and electron. Mantz et al. found that in stacked NAB dimers, eliminating the SIE is essential to reasonably describe the hole delocalization.[28] Further, Autschbach and Srebro systematically investigated the significant impact of the DE in Kohn-Sham theory (KST) calculations of molecular ground-state and response properties. [27] For instance, the Koopmans IPs and EAs based on conventional functionals such as PBE and B3LYP are greatly underestimated by the negative highest occupied molecular orbital (HOMO) and overestimated by the negative lowest unoccupied molecular orbital (LUMO), respectively, resulting in dramatically lower (up to several eVs) HOMO-LUMO orbital gap.[16, 32] The Minnesota and its derivative functionals with a fixed amount of exact-exchange (eX) such as M06-2X[33] have been shown to provide improved description of ionized states of long adenine/guanine stacks.[26] Brédas et al. indicated that the closely related electronic coupling (or transfer integral) parameters are very sensitive to the amount of eX.[34]

Recently, the so-called long-range corrected (LC) or range-separated density functionals[31, 35, 36] switching from DFT to HF at large interelectronic distances have attracted great interest due to

the significantly improved predictive power in π-conjugated systems.[13, 37, 38] LC functionals represent a particular form of generalized Kohn-Sham (GKS)[39] and can produce a correct asymptotic behavior of the potential due to the eX from HF theory used for the long-range part. The range-separation parameter ω in LC functionals can be considered as a function of the density[36, 40] and has been shown to be strongly system-dependent.[32, 41] It has been pointed out that an optimal ω may also depend on the property of interest and the criteria used to decide what "optimal" means.[27] Baer, Kronik (BK), and collaborators have considered a nonempirical criterion to determine the optimal tuned ω based on LC functionals (See Computational Methodology section for the details of tuning procedure). Our early work has applied this optimal tuning approach to a series of molecular ground-state and response properties and found that is obviously superior to the conventional functionals. For example, successful applications of long-range corrected functionals using tuning method for the hyperpolarizabilities of π-conjugated "push-pull" systems and the IPs, EAs and fundamental gap of π-conjugated oligomers and polymers are demonstrated.[32, 37] Foster and Wong found that tuning can accurately predict both fundamental and excitation gaps of isolated NAB molecules (i.e. guanine, adenine, cytosine, thymine and uracil).[42] However, it is obvious that to completely understand the electronic properties of NAB, not only the monomers but the larger fragments need to be investigated. Different interactions between individual NAB monomers and the formation of NAB network significantly affect the respective electron attachment/detachment energies[23] and play a vital role in the enhancement of photostability along the DNA strand.[43, 44] Therefore, computational characterization on the effect of various interactions of NAB pairs and clusters is meaningful for understanding a biologically relevant process based on a realistic model.[16]

In this work, we performed DFT calculations on the IPs and EAs of various isomers of NAB pairs and clusters with H-bonding and π-stacking interactions (molecular structures shown in Figure 1), along with the optimally tuned LC functional. Firstly, with tuning method, the energy of the HOMO should be equal to the negative IP and the energy of the LUMO optimally close to the negative EA, and then the HOMO-LUMO energy gap $\Delta\varepsilon$ is by construction close to the fundamental gap ΔE_F (IP − EA). Therefore the MO energies from various DFT functionals can then be compared directly with the IP/EA obtained by the difference of total energies of neutral and charged systems using WFT methods. The present work only considers vertical IPs and EAs excluding the effects from structure relaxation. Further, the delocalization error for various DFT functionals is numerically quantified by the curvature of energy versus fractional electron numbers for two representative systems. Finally, the cooperative effects of H-bonding and π-stacking interactions on the ionization potentials of various NAB dimers and clusters, as well as the reactivity parameters, are accordingly quantitatively analyzed.

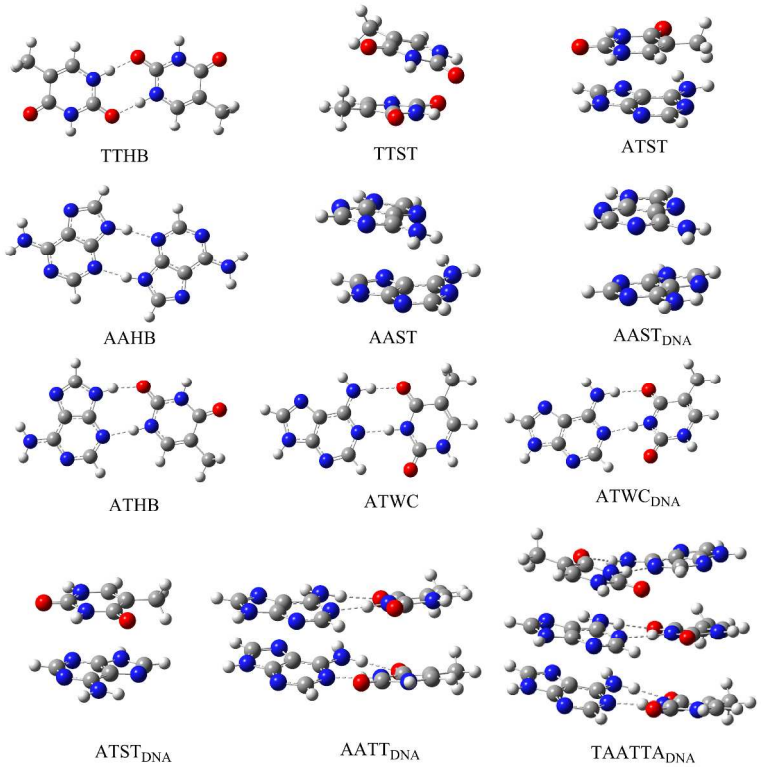

Figure 1. Molecular structures of various isomers of NAB pairs and clusters (T: Thymine, A: Adenine; HB: H-bonded, ST: stacked, WC: Watson-Crick, DNA: X-ray structure of DNA).

2. Computational methodology

Much of this work is concerned with the application and parameter-optimization of LC functionals. The following three-parameter separation of the interelectronic distance r_{12} was used for the range separation of the interelectronic distance in the exchange[31]

$$\frac{1}{r_{12}} = \frac{1-[\alpha+\beta\mathrm{erf}(\omega r_{12})]}{r_{12}} + \frac{\alpha+\beta\mathrm{erf}(\omega r_{12})}{r_{12}} \quad (1)$$

The "erf" term represents the standard error function. The first term on the right hand side is used for the short-range DFT component of the exchange, and the second term is used for the long-range eX component. The parameter α quantifies the fraction of eX in the short-range limit, while α + β gives the fraction of eX in the long-range limit. The optimal tuning approach empirically requires an asymptotically correct functional, i.e. α + β = 1, indicating a correct long-range behavior.[32] In Equation (1), ω is the range separation parameter, representing the inverse distance with the interelectronic separation around which the exchange functional switches from dominantly DFT exchange to dominantly eX. Optimal tuning is based on the finding that in exact Kohn-Sham (KS) and GKS theory,[45] that is, for an N-electron system the HOMO energy $\varepsilon_H(N)$ is negative IP exactly.[46] With conventional functionals, however, the differences can be very large. An optimal ω for LC functionals is nonempirically determined to satisfy the exact condition by minimizing the following equation:

$$J^2 = \sum_{i=0}^{1}[\varepsilon_H(N+i) + IP(N+i)]^2 \quad (2)$$

which optimally tunes both the HOMO of neutral system (N) and the orbital corresponding to LUMO. Since Janak's theorem does not explicitly relate the EA to the negative of the LUMO,

Equation (2) further includes the $N + 1$ system, which indirectly tunes the LUMO of the N electron system. As expected, the optimally tuned functionals typically result in $IP(N) \approx -\varepsilon_H(N)$ and $EA(N) \approx -\varepsilon_L(N)$. For more details and examples using "tuning" approach, see references 27, 40, 47.

All the gas-phase ground-state geometries of NAB pairs and clusters are shown in Figure 1. In order to keep a consistent comparison with previously published benchmark calculations, identical molecular geometries were taken from references [11, 12, 16] (See electronic supplemental information, ESI† for details and all the geometries as documented). All the single point calculations were carried out for the N and $N \pm 1$ systems using default SCF convergence criteria in Gaussian 09.[48] The behavior of delocalization error (DE) was investigated by calculating energies of NAB pairs as a function of fractional electron numbers using the developmental version of the NWChem package.[49] Optimal ω values were determined for all the NAB pairs with the long-range corrected ωPBE functional (LC-ωPBE)[50] with a polarized valence triple-ζ basis set (cc-pVTZ). Hereafter we refer the optimally tuned version as LC-ωPBE*. Foster and Wong found that a larger aug-cc-pVTZ basis set decreased the optimal ω values by roughly 0.01 $Bohr^{-1}$ for individual NABs which had little impact on the calculated parameters of properties and do not change the overall results.[42] The cc-pVTZ basis set has been proved and employed in NAB systems at reasonable computational cost without losing accuracy.[13, 51, 52] The ω values for each NAB pair and cluster are reported in Table 1 and plotted in Figure S1, ESI†. The optimal ω values ranges from 0.185 to 0.279 $Bohr^{-1}$ compared to a default 0.400 $Bohr^{-1}$, indicating the necessity of tuning approach. Recently, the equation-of-motion coupled cluster method for ionization potentials with single and double substitutions (EOM-IP-CCSD) has been successfully applied for the description of ionized states, especially for hole delocalization of ionized non-

covalent dimers.[16] Herein, EOM-IP-CCSD data from references 11, 12 were then employed as benchmark of vertical ionization potentials in this work. Due to the lack of reliable benchmark for vertical electron affinities, we then performed CCSD/aug-cc-pVDZ calculations which have been shown to provide reasonable vertical electron affinities of individual NABs.[14] In addition, results using the popular pure GGA functional PBE, global hybrid B3LYP and meta-GGA M06-2X are also reported.

Table 1. Optimal ω values (Bohr^{-1}) using LC-ωPBE/cc-pVTZ for various isomers of NAB pairs and clusters.

System	Optimal ω	System	Optimal ω
Default	0.400	ATST	0.242
AAHB	0.263	A_{DNA}	0.279
AAST	0.238	$AAST_{DNA}$	0.232
TTHB	0.265	$ATWC_{DNA}$	0.263
TTST	0.245	$ATST_{DNA}$	0.251
ATHB	0.250	$AATT_{DNA}$	0.213
ATWC	0.252	$TAATTA_{DNA}$	0.185

3. Results and discussion

Calculated vertical IPs and EAs, as well as the HOMO and LUMO energies by various DFT functionals are assessed at first. The results mainly focus on the numbers obtained from the optimally tuned LC-ωPBE*. A comparison with the IPs and EAs of EOM-IP-CCSD/CCSD follows. The behavior of delocalization error (DE) is then investigated. The cooperative effects of H-bonding and π-stacking on the IPs of NAB pairs and clusters are quantitatively discussed. Finally, the reactivity parameters such as electrophilicity and global hardness are computationally analyzed in an applied external electric field.

3.1 Orbital energies ($-\varepsilon_{HOMO}$, $-\varepsilon_{LUMO}$), IP and EA

Table 2 collects orbital energies ($-\varepsilon_{HOMO}$ and $-\varepsilon_{LUMO}$), as well as IPs and EAs, for three representative NAB pairs and cluster (ATWC, ATST and $AATT_{DNA}$). Numerical data for other

systems are provided in Table S1 in ESI.† Among various DFT functionals, the calculated orbital energy gap $\Delta\varepsilon$ ($\varepsilon_{LUMO} - \varepsilon_{HOMO}$) ranges from 3.29 to 9.77 eV for ATWC, from 3.54 to 10.00 eV for ATST1, and 2.98 to 9.41 eV from $AATT_{DNA}$. The fundamental gap ΔE_F (IP − EA) varies much less than $\Delta\varepsilon$, indicating its less functional-dependence.[16] Interestingly, for the optimally tuned LC-ωPBE*, $\Delta\varepsilon$ is in excellent agreement with the ΔE_F. The corresponding negative HOMO and LUMO energies also agree well with the vertical IPs and EAs, respectively. Further, as shown in Table 2, the LC-ωPBE* results of both $-\varepsilon_{HOMO}$(IPs) and $-\varepsilon_{LUMO}$(EAs) are in good agreement with high-level EOM-IP-CCSD/CCSD data. For ATWC pair the differences between the EOM-IP-CCSD/CCSD data and negative orbital energies ($-\varepsilon_{HOMO}$/$-\varepsilon_{LUMO}$) using various DFT functionals, as well as the corresponding IPs/EAs are shown as in Figure 2a. The mean absolute deviation (MAD) of molecular orbitals and vertical IPs/EAs for all the NAB pairs and clusters are reported in Table 2. As shown in Figure 2b, the negative HOMO energies using tuned LC-ωPBE* functional are in excellent agreement with EOM-IP-CCSD benchmark with a high statistical correlation of R^2=0.98 and a MAD of only ~0.06 eV. The MAD values for EAs using LC- ωPBE* are ~0.12 eV which is still acceptable but one should also note that the level of CCSD/aug-cc-pVDZ is not the best for benchmark of VEAs. Overall, the results show LC-ωPBE* significantly outperform other conventional functionals. It could reasonably be explained that optimal tuning obey the exact KS theory to a great extent, producing $IP(N) \approx -\varepsilon_H(N)$ and $EA(N) \approx -\varepsilon_L(N)$. For the IPs and EAs calculated from total energy differences, all the DFT functionals (except pure GGA PBE) agree well with the benchmark data, indicating the importance of including a fraction of eX from HF. However, for PBE and B3LYP the negative orbital energies are far from their corresponding IPs and EAs. For example, PBE and B3LYP produce very low $-\varepsilon_{HOMO}$ and very high $-\varepsilon_{LUMO}$, and then a much smaller $\Delta\varepsilon$, just as indicated in

the Introduction. It can be seen that M06-2X provides a significantly improved description for the orbital energies with respect to PBE and B3LYP. However, similar to PBE and B3LYP, $-\varepsilon_{LUMO}$ from M06-2X are qualitatively wrong, predicting the incorrect sign of the electron affinity as shown in Table 2. It is also interesting to point out that the calculated orbital energies of optimally tuned LC-ωPBE* show an obvious improvement compared to those of the nontuned LC-ωPBE. The above results demonstrate that $-\varepsilon_{HOMO}$ and $-\varepsilon_{LUMO}$ from the conventional functionals fail to predict the vertical ionization potentials and electron affinities. Overall, the nonempirically tuned LC-DFT functionals can both quantitatively and qualitatively describe the IPs and EAs. It is important to mention that the computational scaling of CCSD or EOM-CCSD formalism as N^6, currently preventing routine calculations on larger systems such as in this work. The combination of low cost with high accuracy using the optimal "tuning" approach again confirms the clear benefit from the perspective of computational cost.

Table 2. Calculated negative HOMO energy $-\varepsilon_{HOMO}$, LUMO energy $-\varepsilon_{LUMO}$, IP, EA, orbital energy gap $\Delta\varepsilon$ and fundamental gap ΔE_F for ATWC, ATST and $AATT_{DNA}$ using various DFT methods.[a] All units are in eVs.

	PBE	B3LYP	M06-2X	LC-ωPBE	LC-ωPBE*	EOM-IP-CCSD/ CCSD[b]
				ATWC		
$-\varepsilon_{HOMO}$	5.22	6.00	7.32	8.62	7.95	
IP	7.37	7.72	8.17	8.15	7.99	8.01
$-\varepsilon_{LUMO}$	1.93	1.27	0.23	-1.15	-0.68	
EA	-0.19	-0.39	-0.59	-0.52	-0.62	-0.66
$\Delta\varepsilon$	3.29	4.73	7.09	9.77	8.63	
ΔE_F	7.56	8.11	8.76	8.67	8.61	
				ATST		
$-\varepsilon_{HOMO}$	5.54	6.36	7.68	8.96	8.22	
IP	7.77	8.09	8.49	8.47	8.27	8.26
$-\varepsilon_{LUMO}$	2.00	1.34	0.33	-1.04	-0.53	
EA	-0.19	-0.32	-0.46	-0.44	-0.50	-0.59
$\Delta\varepsilon$	3.54	5.02	7.35	10.00	8.75	
ΔE_F	7.96	8.41	8.95	8.91	8.77	
				$AATT_{DNA}$		
$-\varepsilon_{HOMO}$	4.95	5.71	7.00	8.30	7.40	

IP	6.76	7.14	7.63	7.87	7.38	7.46
$-\varepsilon_{LUMO}$	1.97	1.30	0.29	-1.11	-0.43	
EA	0.19	-0.08	-0.40	-0.35	-0.46	-
$\Delta\varepsilon$	2.98	4.41	6.71	9.41	7.83	
ΔE_F	6.57	7.22	8.03	8.22	7.84	
All NAB pairs in Figure 1						
$MAD_{\varepsilon H}$[c]	2.76	1.94	0.61	0.69	0.06	
MAD_{IP}[c]	0.72	0.39	0.15	0.27	0.06	
$MAD_{\varepsilon L}$[c]	2.51	1.84	0.83	0.52	0.12	
MAD_{EA}[c]	0.48	0.34	0.20	0.18	0.13	

[a]The orbital energies are calculated with cc-pVTZ basis set. [b]EOM-IP-CCSD values are calculated with 6-311+G(d,p) basis set, taken from references 12, 16. The CCSD values are calculated with aug-cc-pVDZ basis set in this work. [c]The mean absolute deviation (MAD) values are summarized for the ten (seven) NAB pairs for IP (EA) with respect to the EOM-IP-CCSD(CCSD) values in Table 2 and Table S1.

Table 3. Calculated Negative HOMO energy $-\varepsilon_{HOMO}$ and LUMO energy $-\varepsilon_{LUMO}$ for various isomers of NAB pairs and clusters using optimally tuned LC-ωPBE*/cc-pVTZ. All units are in eVs.

	AAHB	AAST	TTHB	TTST	ATHB	ATWC	ATST
$-\varepsilon_{HOMO}$	8.26	8.09	8.83	8.77	8.33	7.95	8.22
$IP_{EOM\text{-}IP\text{-}CCSD}$[a]	8.23	8.16	8.88	8.91	8.36	8.01	8.26
$-\varepsilon_{LUMO}$	-0.91	-0.77	-0.58	-0.31	-0.62	-0.68	-0.53
EA_{CCSD}[b]	-1.02	-0.37	-0.62	-0.13	-0.64	-0.66	-0.59
	A_{DNA}	$AAST_{DNA}$	$ATWC_{DNA}$	$ATST_{DNA}$	$AATT_{DNA}$	$TAATTA_{DNA}$	
$-\varepsilon_{HOMO}$	8.29	7.76	8.10	8.07	7.40	7.12	
$IP_{EOM\text{-}IP\text{-}CCSD}$[a]	8.21	7.81	8.01	-	7.46	-	
$-\varepsilon_{LUMO}$	-1.23	-0.94	-0.78	-0.70	-0.43	-0.42	

[a]EOM-IP-CCSD values are calculated with 6-311+G(d,p) basis set, taken from references 12, 16. [b]The vertical EA values by CCSD are calculated with aug-cc-pVDZ basis set in this work.

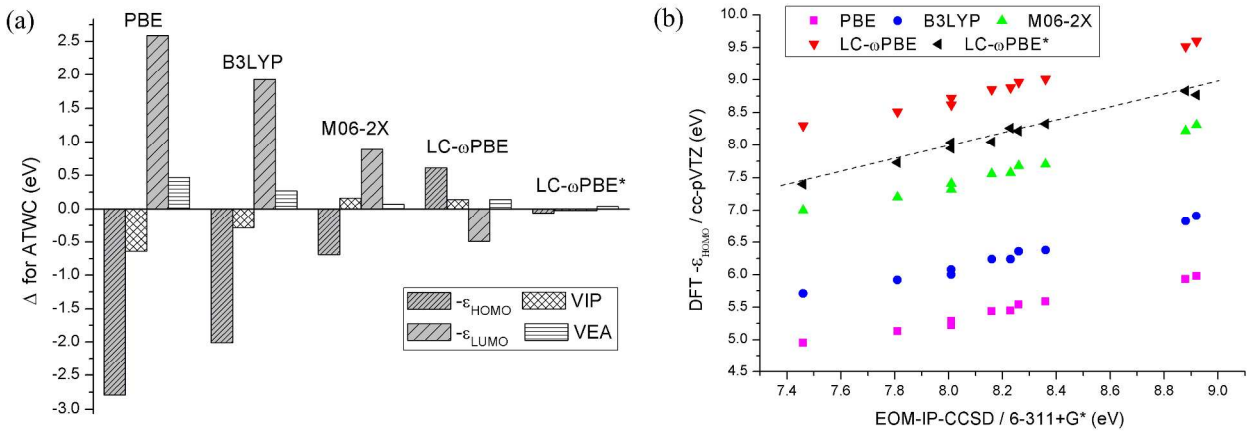

Figure 2. (a) Differences between the orbital energies ($-\varepsilon_{HOMO}$, $-\varepsilon_{LUMO}$) and calculated vertical IP and EA with respect to the EOM-IP-CCSD/CCSD benchmark data for ATWC. (b) Comparison between DFT

orbital energies $-\varepsilon_{HOMO}$ and EOM-IP-CCSD values for all the 10 NAB pairs in Figure 1. The diagonal dashed line represents an ideal 100% agreement with the EOM-IP-CCSD numbers.

3.2 Behavior of delocalization error (DE)

Yang et al. indicated that problems with DFT are intimately connected to violations of basic conditions of constraints known to exist in an exact KS or GKS framework.[53] Conventional functionals with inappropriate approximations always over-/underestimate the delocalization of hole/electron and create the so-called delocalization error (DE). As we know, in exact KS theory, the energy of an atom or molecule as a function of electron number *E(N)* should afford straight-line segments between integers.[32] The curvature of *E(N)* is indicative of the DE.[53] Here, as shown in Figure 3, we examine the behavior of *E(N)* for ATWC and $AATT_{DNA}$ as representative examples. The results confirm the expectations: pure GGA PBE and global hybrid B3LYP produce large positive curvatures of *E(N)*, indicating that these functionals are too delocalized. Magnitude-wise, the meta-GGA M06-2X produce less pronounced DE, compared to conventional PBE and B3LYP. Interestingly, for LC functional the behavior of DE is much more improved. However, the electron-rich ($\Delta N > 0$) part by LC-ωPBE still afford an obvious negative curvature, indicating a somewhat HF-like character.[54] Overall, the optimally tuned LC-ωPBE* can yield the smallest DEs and the tuning procedure appears to be also successful for the larger $AATT_{DNA}$. In addition, the behavior of delocalization errors using various DFT functionals agree with their performance of calculations of molecular orbitals and IP/EA. The results suggest that the DEs can be potentially considered as an additional indicator to select parametrizations among a continuum of functionals.

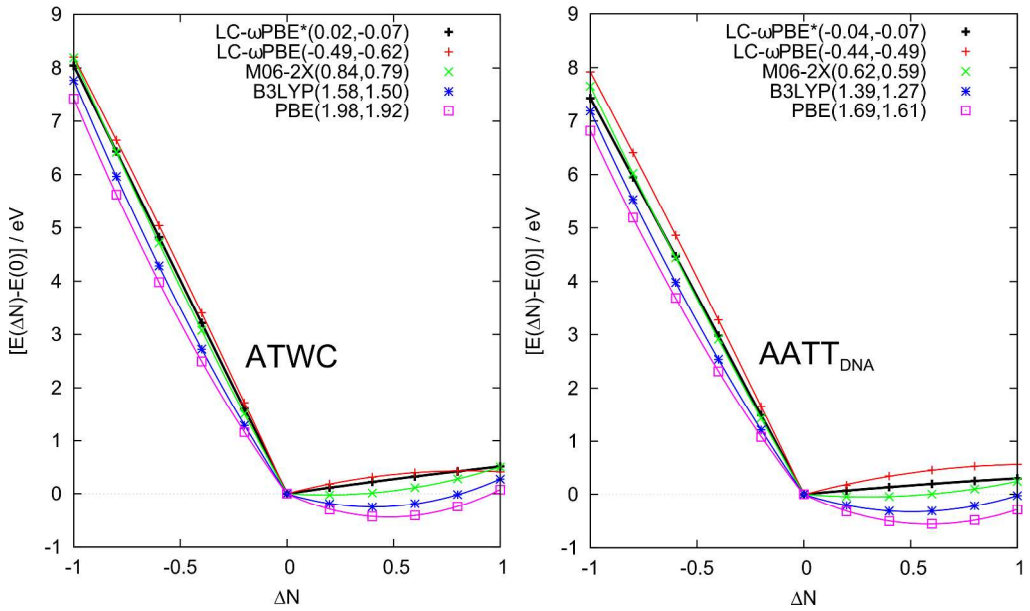

Figure 3. Energy of ATWC (left) and $AATT_{DNA}$ (right) as a function of fractional electron number, ΔN, relative to neutral system ($\Delta N = 0$). The numerical values in the plot correspond to the coefficients of $(\Delta N)^2$ of quadratic fits to $E(N)$ in the electron-deficient and electron-rich regime, respectively ($\Delta N < 0$, $\Delta N > 0$). The coefficients of $(\Delta N)^2$ indicate the presence of absence of curvature. Energies are calculated at DFT/def2-TZVP levels.

3.3 Characterization of cooperative effect on ionization potentials

Again, accurate determination (both quantitatively and qualitatively) of ionization potentials and electron affinities are important for a fundamental understanding of ionization and electron-impact radiation damage and repair and specific mechanisms of charge transport in the NAB clusters. The related electronic states are significantly affected by various interactions, including base pairing, H-bonding, π-stacking, electrostatic effect, etc. Crespo-Hernández and coworkers qualitatively found that base pairing via H-bonding reduces the ionization energies and oxidation potentials for NAB individuals.[23] Sugiyama and Satio indicated the lowering of IP depends on the increasing hole delocalization in π-stacked systems.[55] Apparently, a qualitative study is not enough but a quantitative prediction could be much more challenging. Krylov and coworkers performed high-level benchmark calculations using EOM-IP-CCSD for AT tetramer ($AATT_{DNA}$) and quantitatively studied the additive and cooperative effects of H-bonding and π-stacking on

the ionization energies.[16] Herein, due to the efficiency and reliability of the optimally tuned LC-ωPBE* functional, we take the opportunity to investigate the cooperative effect of H-bonding and π-stacking of a longer AT hexamer taken from DNA duplex of X-ray DNA structure. This model is more realistic for simulating charge-transfer process in DNA as high-level WFT methods are not possible for these systems.

As shown in Table 3, the effect of π-stacking interactions of AAST and ATST on the ionization potentials is more pronounced than that of H-bonding interactions of AAHB and non-WC ATHB. Their stacking geometries result in a lower IP and indicate a stronger delocalization of HOMO by π-stacking effect. The WC-form ATWC has an even lower IP compared to that of both non-WC ATHB and stacked ATST, indicating that the H-bonding interaction in WC pair is stronger than the stacking interaction.[51] We accordingly focus on the HOMOs and LUMOs of NAB pairs and clusters taken from X-ray DNA structure. Since adenine has a relatively lower IP than thymine, the HOMO of the $ATWC_{DNA}$ and $ATST_{DNA}$ dimers mainly localize on the adenine fragment as expected (see Figure S2, ESI†). For $AAST_{DNA}$, the HOMO is delocalized equally on the two adenine bases. For more complex and realistic systems ($AATT_{DNA}$ and $TAATTA_{DNA}$), the shape of the HOMO is similar to that of the $AAST_{DNA}$ dimer. Furthermore the addition of the TT stacked pair or AT H-bonded pair does not obviously affect the character of HOMOs as shown in Figure 4. As expected the LUMO of the AT_{DNA} dimers ($ATWC_{DNA}$ and $ATST_{DNA}$) consist of thymine's LUMO, that is, the electron is mainly populated on thymine piece (Figure S2, ESI†). However, for the LUMO of the TT_{DNA} dimer, the electron is only localized on one thymine base, unlike the case of AA_{DNA} whose HOMO equally distributing on both adenine fragments. This observation is also found in more complex systems ($AATT_{DNA}$ and

$TAATTA_{DNA}$) and confirm the fact that the LUMO energy of $TAATTA_{DNA}$ slightly changed compared to that of $AATT_{DNA}$ as shown in Table 3.

Further, the additive or cooperative effects of H-bonding and π-stacking on the ionization potentials are also quantified in Figure 4, and some results are compared to the EOM-IP-CCSD numbers from Krylov's work.[16] Compared to the IP of isolated adenine (8.29 eV) stacking and H-bonding interaction result in the shifts of -0.53 and -0.19 eV, respectively, and a sum of -0.72 eV for both interactions. Comparing with a shift of -0.89 eV for the tetramer $AATT_{DNA}$, the extra -0.17 eV (versus -0.15 eV for EOM-IP-CCSD) can be attributed to the cooperative effect of H-bonding and π-stacking interactions, that is, cross interactions between non-H-bonded base fragments.[16] Further, we consider an AT hexamer case. As shown in Figure 4, the AT hexamer includes additional stacking interactions of A and T bases corresponding to the initial T and A bases. So the additivity of -0.22 eV resulting from AT stacking interaction is counted as shown in Figure 4. The total estimated shift including the three types of interactions is -0.94 eV, is to be compared with a computed shift of -1.17 eV for $TAATTA_{DNA}$. Therefore, -0.23 eV can be attributed to the cooperative effects of the AA stacking, AT stacking and AT H-bonding interactions. Compared to the cooperative effect (-0.17 eV) of $AATT_{DNA}$ tetramer, the additional AT stacking interaction in $TAATTA_{DNA}$ results in a stronger cooperative effect (-0.23 eV) which can enhance the stability of the whole system. A 3D plot of electrostatic potential of AT clusters (as shown in Figure S3, ESI†) could be helpful to get a more vivid view of understanding the cooperative effects resulting from the diagonal interactions. The stronger cooperative effect can also be indirectly reflected as more polarized (more red) oxygen atoms on $TAATTA_{DNA}$ compared to those on $AATT_{DNA}$. Overall, our main goal in this section is to prove that the HOMO energies using tuned LC-ωPBE* functional can both quantitatively and qualitatively

describe the effects of H-bonding and π-stacking on the ionization potentials of NAB clusters, assessing the quality of prediction which are comparable to the much more expensive EOM-IP-CCSD method.

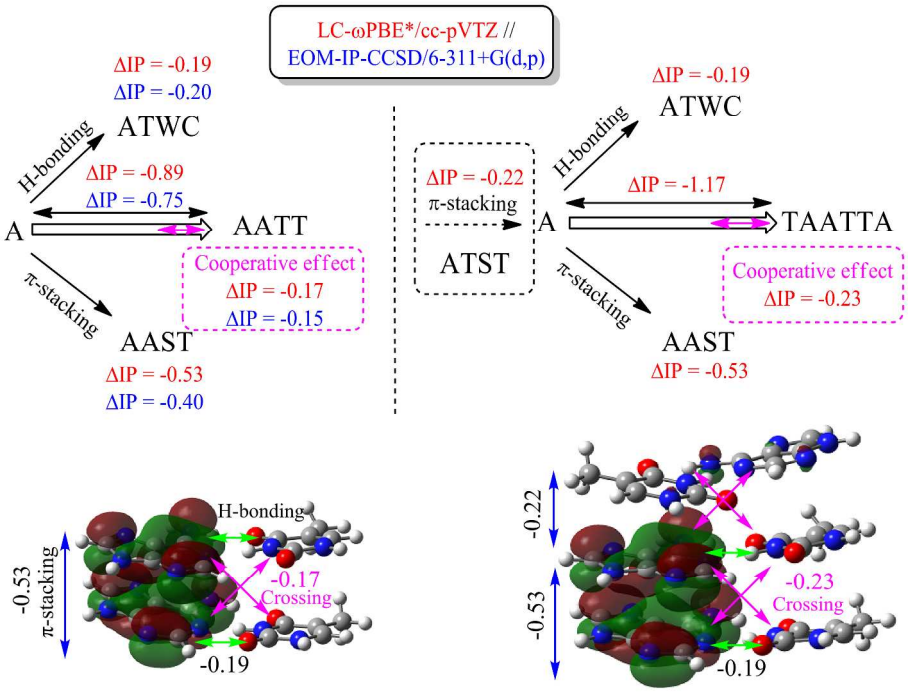

Figure 4. Schematic diagram of cooperative effect on the ionization potentials of the AT tetramer and hexamer ($AATT_{DNA}$ and $TAATTA_{DNA}$) compared to that of adenine monomer. The HOMOs (bottom) of the $AATT_{DNA}$ and $TAATTA_{DNA}$ are calculated using the optimally tuned LC-ωPBE*.

3.4 Reactivity parameters toward external electric field

It is well-known that the effect of the external electric field is common and important on the chemical reactions of biomolecules.[10] Recently, various reactivity descriptors, such as global hardness, chemical potential, electrophilicity and Fukui functions, have been theoretically estimated by density functional reactivity theory (DFRT).[20] Kar and Pal have investigated the electric field response of the molecular reactivity descriptors.[56] As shown in equation (3) they proposed a parameter of electrophilicity (ω) to measure the electrophilicity strength of a molecule and its propensity to soak up electrons.[57]

$$\omega = \frac{\mu^2}{2\eta} \quad (3)$$

where $\mu = (\varepsilon_{LUMO} + \varepsilon_{HOMO})/2$, representing the chemical potential and $\eta = (\varepsilon_{LUMO} - \varepsilon_{LUMO})/2$, representing the global hardness.[58] The definitions are based on the finite difference approximation and Koopmans's theorem. Herein, the ω, μ and η parameters of the representative ATWC pair are calculated as a function of electric field strength using various DFT functionals as shown in Figure 5 and detailed numbers are collected in Table S2, ESI†. The range of external electric field strength is chosen from 0.00 to ±0.01 au as suggested by reference [10] for NABs. It can be seen that the reactivity parameters (global hardness η and electrophilicity ω) are indeed affected by the presence of an external electric field and all the DFT functionals show a similar trend when the electric field strengths increase or decrease. First, the finding that the maximum value for η and minimum value for ω are found with the strength of +0.002 au, indicating a maximum stability induced by an applied electric field according to the maximum hardness principle (MHP), that is, maximum hardness results in maximum stability. Application of the growing strength of field along the $-x$ direction will decrease the global hardness and increase the electrophilicity. These results are well consistent with Dutta and Bhattacharyya 's findings.[10] However, it is observed that the much delocalized PBE functional significantly overestimate the ω values and underestimate the η values, especially in strong electric fields of ± 0.01 au. B3LYP can do slightly better but still has an ignorable deviation compared to the CCSD value. The meta-GGA M06-2X significantly improves the behavior of vibration of reactivity parameters. Overall, the tuned LC-ωPBE* can exhibit more reasonable behavior of reactivity parameters with respect to the change of strength of electric field, allowing a potential tool for the quantitative investigation of specific mechanism of NABs in an applied external electric field.

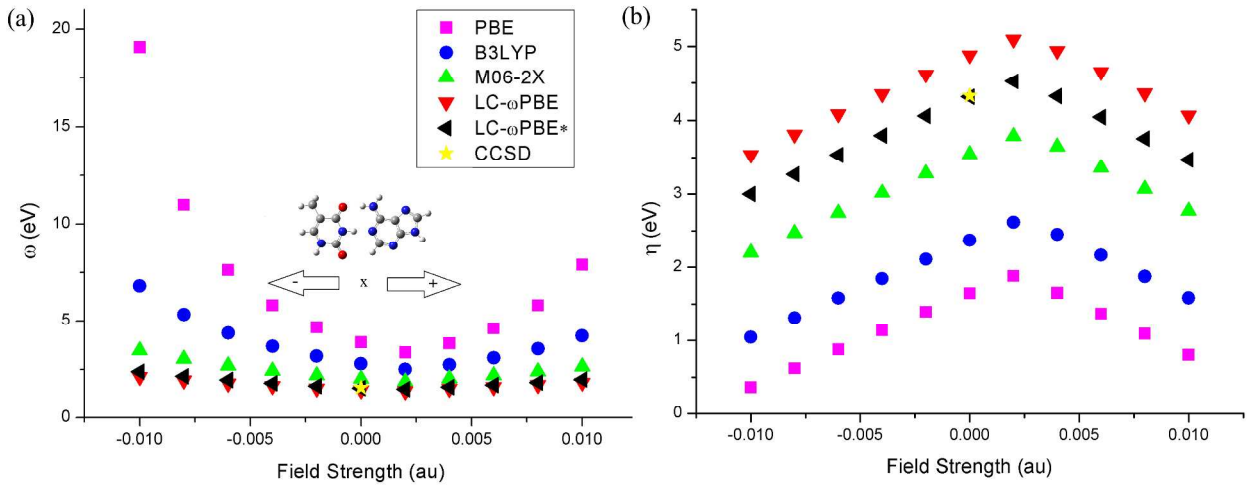

Figure 5. Calculated reactivity descriptors of (a) electrophilicity ω and (b) global hardness η of ATWC as a function of electric field strength along the x axe at different DFT levels with cc-pVTZ basis set. 1 au = 51.4 V/Å = 51.4 × 10^{10} V/m.

Conclusions

In conclusion, the vertical ionization potentials and electron affinities of various isomers of adenine-thymine nucleobase pairs and clusters were investigated using various DFT functionals. The cooperative effects of π-stacking and H-bonding interactions on the ionization potentials of AT clusters, as well as the reactivity parameters are studied. The main conclusions are as follows:

- The nonempirically tuned LC-DFT functionals can reasonably describe the IPs and EAs, which are in good agreement with the high-level EOM-IP-CCSD/CCSD methods. By construction, the negative HOMO and LUMO energies ($-\varepsilon_{HOMO}$ and $-\varepsilon_{LUMO}$) are equal to the vertical IPs and EAs calculated from total energy differences, respectively, and the HOMO-LUMO orbital gap $\Delta\varepsilon$ is a good approximation to the IP-EA fundamental gap. However, the conventional functionals fail to give quantitative and qualitative description of IPs and EAs.

- The delocalization error (DE) by different DFT functionals is quantified by calculations with fractional electron numbers. The optimally tuned LC-ωPBE* can almost eliminate the DEs, while conventional functionals in use (PBE and B3LYP) tend to produce pronounced DEs.
- Analysis of the relative magnitudes of the shift in ionization potentials caused by π-stacking and H-bonding interactions indicates the existence of an obvious cooperative effect. The cooperative effect on the AT tetramer was quantitatively reproduced by optimally tuned LC functionals and the cooperative effect on the AT hexamer has a stronger influence on the decreased IPs, indicating an enhancement of stability. In addition, the tuned LC functional can produce more reasonable reactivity parameters with respect to the strength of electric field, while the conventional functionals overly evaluate this behavior.
- Finally, this work provides an efficient and reliable theoretical tool to describe the electronic structures and related electron attach- and detachment abilities of NAB clusters and may be potentially explored to a long DNA strand. A parallel study on the excited states of NAB pairs is ongoing in our group.

Acknowledgments.

This work has been supported by the China Postdoctoral Science Foundation (2014M561435), National Natural Science Fund (Nos. 11004060 and 11027403), Shanghai Municipal Science and Technology Commission (No. 14JC1401500), and Shanghai Rising-Star Program (No. 12QA1400900). H.S. sincerely appreciates the valuable discussion with Prof. Jochen Autschbach (University at Buffalo) and thanks for the introduction of "tuning" concept during his visit.

References:

(1) P. Tumbale, J. S. Williams, M. J. Schellenberg, T. A. Kunkel and R. S. Williams, *Nature*, 2014, **506**, 111-115.
(2) B. N. G. Giepmans, S. R. Adams, M. H. Ellisman and R. Y. Tsien, *Science*, 2006, **312**, 217-224.
(3) C. E. Crespo-Hernández, B. Cohen and B. Kohler, *Nature*, 2005, **436**, 1141-1144.
(4) G. R. Fleming and G. D. Scholes, *Nature*, 2004, **431**, 256-257.
(5) K. Kawai, H. Kodera, Y. Osakada and T. Majima, *Nat. Chem.*, 2009, **1**, 156-159.
(6) J. C. Genereux and J. K. Barton, *Nat. Chem.*, 2009, **1**, 106-107.
(7) S. Steenken, J. P. Telo, H. M. Novais and L. P. Candeias, *J. Am. Chem. Soc.*, 1992, **114**, 4701-4709.
(8) A. O. Colson and M. D. Sevilla, *J. Phys. Chem.*, 1995, **99**, 3867-3874.
(9) S. Steenken, *Chem. Rev.*, 1989, **89**, 503-520.
(10) B. J. Dutta and P. K. Bhattacharyya, *J. Phys. Chem. B* 2014, **118**, 9573-9582.
(11) A. A. Golubeva and A. I. Krylov, *Phys. Chem. Chem. Phys.*, 2009, **11**, 1303-1311.
(12) K. B. Bravaya, O. Kostko, M. Ahmed and A. I. Krylov, *Phys. Chem. Chem. Phys.*, 2010, **12**, 2292-2307.
(13) L. Jensen and N. Govind, *J. Phys. Chem. A*, 2009, **113**, 9761-9765.
(14) D. Roca-Sanjuán, M. Merchán, L. Serrano-Andrés and M. Rubio, *J. Chem. Phys.*, 2008, **129**, 095104-095114.
(15) D. Roca-Sanjuán, M. Rubio, M. Merchán and L. Serrano-Andrés, *J. Chem. Phys.*, 2006, **125**, 084302-084308.
(16) K. B. Bravaya, E. Epifanovsky and A. I. Krylov, *J. Phys. Chem. Lett.*, 2012, **3**, 2726-2732.
(17) A. K. Dutta, T. Sengupta, N. Vaval and S. Pal, *ArXiv e-prints*, 2014, 2014arXiv1409.7266D.
(18) W. Kohn and L. Sham, *Phys. Rev. A*, 1965, **140**, 1133-1138.
(19) W. Kohn, *Rev. Mod. Phys.*, 1999, **71**, 1253-1266.
(20) R. G. Parr and W. Yang, *Density functional theory of atoms and molecules*. Oxford University Press: 1989.
(21) A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
(22) C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B*, 1988, **37**, 785-789.
(23) C. E. Crespo-Hernández, D. M. Close, L. Gorb and J. Leszczynski, *J. Phys. Chem. B*, 2007, **111**, 5386-5395.
(24) D. M. A. Vera and A. B. Pierini, *Phys. Chem. Chem. Phys.*, 2004, **6**, 2899-2903.
(25) X. Li, Z. Cai and M. D. Sevilla, *J. Phys. Chem. A*, 2002, **106**, 1596-1603.
(26) A. Kumar and M. Sevilla, *J. Phys. Chem. B*, 2011, **115**, 4990-5000.
(27) J. Autschbach and M. Srebro, *Acc. Chem. Res.*, 2014, **47**, 2592-2602.
(28) Y. Mantz, F. Gervasio, T. Laino and M. Parrinello, *J. Phys. Chem. A*, 2007, **111**, 105.
(29) Y. Paukku and G. Hill, *J. Phys. Chem. A*, 2011, **115**, 4804.
(30) D. J. Tozer, *J. Chem. Phys.*, 2003, **119**, 12697-12699.
(31) T. Yanai, D. P. Tew and N. C. Handy, *Chem. Phys. Lett.*, 2004, **393**, 51-57.
(32) H. Sun and J. Autschbach, *J. Chem. Theory Comput.*, 2014, **10**, 1035.
(33) D. Truhlar, *Theor. Chim. Acta*, 2008, **120**, 215241.

(34) C. Sutton, J. S. Sears, V. Coropceanu and J.-L. Brédas, *J. Phys. Chem. Lett*, 2013, **4**, 919-924.
(35) H. Iikura, T. Tsuneda, T. Yanai and K. Hirao, *J. Chem. Phys.*, 2001, **115**, 3540-3544.
(36) R. Baer and D. Neuhauser, *Phys. Rev. Lett.*, 2005, **94**, 043002-7.
(37) H. Sun and J. Autschbach, *ChemPhysChem*, 2013, **14**, 2450.
(38) U. Salzner and A. Aydin, *J. Chem. Theory Comput*, 2011, **7**, 2568-2583.
(39) A. Seidl, A. Gorling, P. Vogl, J. A. Majewski and M. Levy, *Phys. Rev. B*, 1996, **53**, 3764-3774.
(40) L. Kronik, T. Stein, S. Refaely-Abramson and R. Baer, *J. Chem. Theory Comput.*, 2012, **8**, 1515-1531.
(41) T. Körzdörfer, J. S. Sears, C. Sutton and J.-L. Brédas, *J. Chem. Phys.*, 2011, **135**, 204107.
(42) M. E. Foster and B. M. Wong, *J. Chem. Theory Comput.*, 2012, **8**, 2682-2687.
(43) D. Roca-Sanjuán, M. Merchán and L. Serrano-Andrés, *Chem. Phys. Lett.*, 2008, **349**, 188-196.
(44) T. Gustavsson, R. Improta and D. Markovitsi, *J. Phys Chem. Lett.*, 2010, **1**, 2025-2030.
(45) J. F. Janak, *Phys. Rev. B*, 1978, **18**, 7165-7168.
(46) M. Levy, J. P. Perdew and V. Sahni, *Phys. Rev. A*, 1984, **30**, 2745-2748.
(47) T. Körzdörfer and J.-L. Brédas, *Acc. Chem. Res.*, 2014, ASAP.
(48) M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, N. J. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox *Gaussian 09*, Gaussian, Inc.: Wallingford, CT, USA, 2009.
(49) M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus and W. A. de Jong, *Comput. Phys. Commun.*, 2010, **181**, 1477-1489.
(50) O. A. Vydrov and G. E. Scuseria, *J. Chem. Phys.*, 2006, **125**, 234109.
(51) H. Karabyk, R. Sevincek and H. Karabyk, *Phys. Chem. Chem. Phys.*, 2014, **16**, 15527-15538.
(52) I. Y. Zhang and X. Xu, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12554-12570.
(53) A. J. Cohen, P. Mori-Sánchez and W. Yang, *Science*, 2008, **321**, 792-794.
(54) B. Kirtman, S. Bonness, A. Ramirez-Solis, B. Champagne, H. Matsumoto and H. Sekino, *J. Chem. Phys.*, 2008, **128**, 114108.
(55) H. Sugiyama and I. Saito, *J. Am. Chem. Soc.*, 1996, **118**, 7063-7068.
(56) R. Kar and S. Pal, *Theor. Chem. Acc.*, 2008, **120**, 375-383.
(57) R. G. Parr, L. V. Szentpaly and S. Liu, *J. Am. Chem. Soc.*, 1999, **121**, 1922-1924.
(58) R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.*, 1983, **105**, 7512-7516.

TOC

Successful applications of optimally tuned long-range corrected functionals for calculations of vertical ionization potentials and electron affinities of various adenine-thymine nucleobase pairs and clusters are demonstrated.