THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

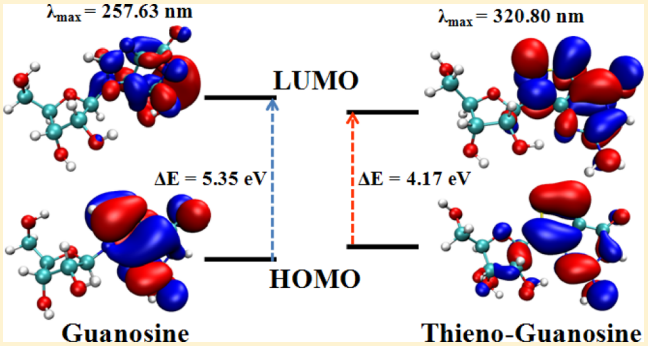
Thieno Analogues of RNA Nucleosides: A Detailed Theoretical Study

Pralok K. Samanta,[†] Arun K. Manna,[†] and Swapan K. Pati*[,†,‡]

[†]Theoretical Sciences Unit and [‡]New Chemistry Unit, Jawaharlal Nehru Center for Advanced Scientific Research, Jakkur P.O., Bangalore 560064, India

Supporting Information

ABSTRACT: We use first-principles density functional theory calculations to investigate the structural, energetic, bonding aspects, and optical properties of recently synthesized thieno-analogues of RNA nucleosides. The results are compared against the findings obtained for both the natural nucleosides as well as available experimental data. We find that the modified nucleosides form the hydrogen bonded Watson–Crick (WC) base pairing with similar H-bonding energy as obtained for the natural nucleosides. We have calculated and compared the charge transfer integrals for the H-bonded natural and thieno-modified nucleosides. We find that the thieno modification of these nucleosides strongly affects the charge transfer integrals due to the difference in extent of orbital delocalization in these two types of nucleosides. We also find that the degree of reduction of charge transfer integrals is larger for the H-bonded A–U pair than in the G–C pair. We also focus on the optical absorption properties of these thieno-modified nucleosides and their WC H-bonded base pairs in gas phase as well as with implicit water. Our calculated results show that the low energy peaks in the absorption spectra mainly arise because of the $\pi-\pi^*$ electronic transition for both the nucleosides, and the observed red shift for thieno-nucleosides compared to natural nucleosides are consistent with the calculated decrease in electronic gaps. Our results demonstrate that the thieno modification of natural nucleosides significantly modifies their electronic and optical properties, although the basic structural and bonding aspects remained the same. It also gives a microscopic understanding of the experimentally observed optical behaviors.



1. INTRODUCTION

Amino acid residues and nucleic acid residues are not fluorescence active. This makes long chain polymer molecules, like proteins and nucleic acids, fluorescence inactive. Modified amino and nucleic acid residues, possessing fluorescent properties in the UV region are very important for biomedical applications.[1] Deoxyribonucleic acid (DNA) and ribonucleic acid (RNA) are considered the building blocks of all life.[2−7] DNA sequences coding for specific genes are transcribed in the cell nucleus to yield heterogeneous nuclear ribonucleic acids (RNAs). The resulting primary RNA transcripts are processed, yielding mature RNAs, which are exported to the cytoplasm, where ribosome-based protein synthesis takes place. The transport, localization, stability, and translation efficiency of individual mRNA molecules are all well regulated. Additionally, a recent development suggests significant roles for noncoding RNA sequences in cellular regulatory process, adding to the multifaceted and intricate roles of these biomolecules in the cell.[8] Natural RNA nucleosides consist of four types of nucleobases: guanine (G), cytosine (C), adenine (A), and uracil (U). Since these nucleobases do not show fluorescence in the UV region, modification of nucleobases is important to improve the biomedical applications. For example, replacing the natural nucleosides with established fluorophore, typically polycyclic aromatic hydrocarbons (PAH), yields an unusual family of chromophoric fluorescence based analogues.[9] These PAH analogues can be used for the investigation of enzyme–substrate recognition.[10] The size-expanded nucleobases also show high fluorescence properties, which might play an important role in bioimaging and biomedical treatments.[11,12] Furthermore, various small fluorescent active noble metal clusters are designed using the DNA template for diagnostic applications.[13−16] Again, when these modified nucleosides link via phosphodiester bonds to form an oligomeric structure, these nucleosides yield unique water-soluble fluorophores.

Pteridines are naturally occurring, highly emissive heterocyclic organic compounds whose structures are related to that of the purines. Their intense and visible fluorescence and long excited state lifetime results from an isolated absorption band above 300 nm.[17] The modification can also be done by extending the conjugation of the natural nucleobases by fusing additional aromatic rings onto the pyrimidine and purine nuclei. This generates diverse expanded nucleobases and most of them retain their Watson–Crick (WC) hydrogen bonding (H-bonding) between the base pairs. Having an extended aromatic surface typically results in favorable photophysical

Received: February 22, 2012
Revised: May 9, 2012
Published: June 7, 2012

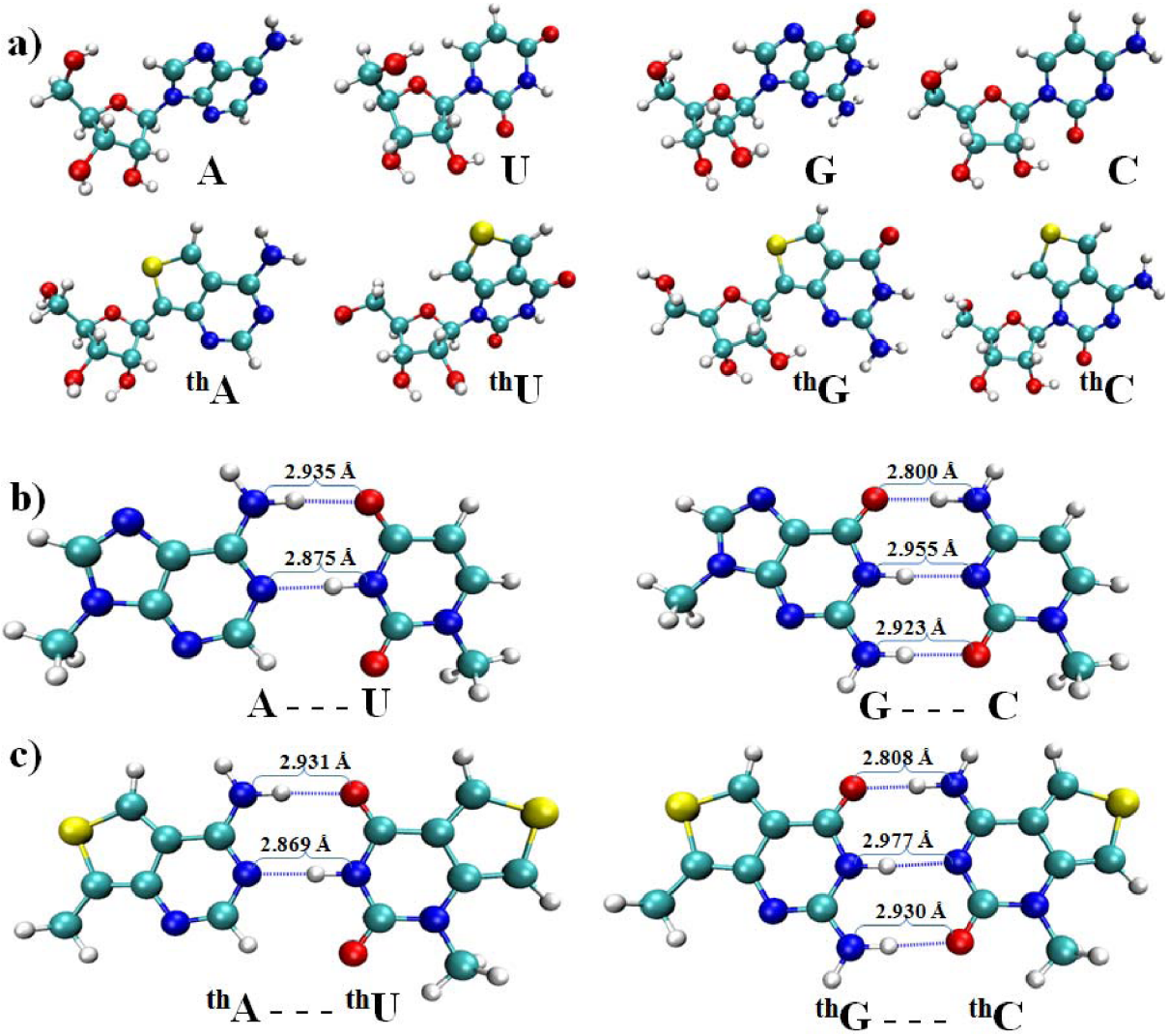


Figure 1. B3LYP/6-31++G (d, p) level optimized structures of natural RNA nucleosides and thieno-nucleosides (thieno[3,4-*d*]pyridmidine nucleosides) (a) and H-bonding between A−U, G−C (b) and thA−thU, thG−thC (c). Atom color code: blue (N), cyan (C), white (H), red (O), and yellow (S).

properties, with red-shifted absorption bands compared to their natural counterparts.[18−23] Extended fluorescence nucleoside analogues are distinguished by fluorescent moieties, which are linked or conjugated to the natural nucleobases via either flexible or rigid linkers. Connecting known chromophores via an electronically nonconjugating linker yields a nucleoside analogue with photophysical features, which are normally very similar to that of the parent fluorophore. Extending the purine and pyrimidine nucleobases by electronically conjugating them to additional aromatic moieties typically generates a new chromophore with unique photophysical characteristics.[24−26]

However, nucleoside modification can be realized by replacing the natural nucleobase by its isomorphic nucleobase analogues. Isomorphic nucleobase analogues are heterocyclic organics compound that closely resemble the corresponding natural nucleobases with respect to their overall dimensions.[27,28] A clear advantage of these analogues is their strong similarity to the native nucleosides and minimally perturbing nature. Since favorable photophysical characteristics (e.g., red-shifted absorption and high emission quantum efficiencies) are typically associated with significant structural perturbation and extended conjugation, isomorphic fluorescent nucleosides are the most challenging to design. The advancement of solid-phase oligonucleotide synthesis facilitates the incorporation of modified nucleotide into DNA/RNA oligomers, which show lots of promise in biological and technological applications. Moreover, several theoretical studies have been reported in the literature on natural as well as modified base pairs of nucleobases in order to explore their structural stability and the electrical conductivity. These include various metal modified WC base pairs[29−32] and artificial base pairs,[33−37] analogues of natural nuclobases. These motivate chemists to be interested in exploring new analogues of nucleotides. Recently, Shin et al. have synthesized four thieno modified RNA nucleosides; namely, thA, thU, thG, and thC (see Figure 1a) and have reported their photophysical properties.[38]

Motivated by the experimental work by Shin et al., here, we explore the structural, energetic, and optical properties of these thieno modified ribonucleosides (thA, thU, thG, and thC) along with their natural analogues (A, U, G, and C) using quantum chemical calculations. We analyze and compare the structures and the propensity of WC base pairs formation in between two complementary nucleosides considering both natural and modified nucleobases. We also focus on the optical absorption

dx.doi.org/10.1021/jp301752k | *J. Phys. Chem. B* 2012, 116, 7618−7626

Table 1. Calculated Stabilization Energies (ΔE_{stab}) and Its Basis Set Superposition Error (BSSE) Corrected Values for Various WC Base Pair of Nucleobases with Varying Exchange and Correlation Energy Functionals

	ΔE_{stab} (kcal mol^{-1})			ΔE_{stab} + BSSE (kcal mol^{-1})			ΔE_{stab} + BSSE + ZPE[a] (kcal mol^{-1})		
systems	B3LYP	M06-2X	WB97XD	B3LYP	M06-2X	WB97XD	B3LYP	M06-2X	WB97XD
A−U	−12.88	−15.25	−16.48	−12.17	−14.46	−15.76	−10.91	−13.50	−14.60
thA−thU	−12.80	−15.16	−16.48	−12.10	−14.37	−15.77	−11.24	−13.53	−14.60
G−C	−25.98	−28.40	−30.47	−24.99	−27.33	−30.47	−23.38	−26.11	−28.87
thG−thC	−23.99	−26.25	−28.40	−23.01	−24.39	−27.42	−21.60	−23.30	−25.77

[a]ZPE is the zero-point energy.

characteristics in nucleosides as well as in their WC pairs and find reasonably good agreement with the experimental results. Moreover, we provide a microscopic origin of the low-energy peaks in the absorption spectra as observed experimentally.

2. COMPUTATIONAL DETAILS

The geometry of various RNA nucleosides and their thieno modified analogues are optimized using density functional theory (DFT), and the electronic absorption spectra are calculated using time-dependent density functional theory (TD-DFT) methods as implemented in the Gaussian09 package.[39] All the calculations are performed using B3LYP hybrid exchange and correlation functional,[40−42] with the 6-31++G(d,p) basis set for all atoms. For the H-bonded WC base pairs, we also invoke M06-2X[43] and WB97XD[44,45] exchange and correlation functional in order to take into account the weak noncovalent interactions, such as H-bonding and van der Waals interaction. Note that the structure and energetics of the H-bonded nucleobase pairs are strongly dependent on the calculation levels used for their computations. The robustness of the calculation level used in the present study has been already proved by several studies in weakly bound molecular complexes.[46−54] The calculations are performed both in gas phase and with water surrounding, using polarizable continuum model (PCM)[55] to include the solvent effect. After geometry optimization, frequency calculations are performed to remove any vibrational unstable mode. The convergence criterion for the Self-Consistent-Field (SCF) is set to "Tight," and the "UltraFine" grid is used for numerical integration in DFT as implemented in Gaussian09 sets of codes.[39]

The charge transfer integrals between the nucleobases within a WC base pair are calculated by DFT method using the fragment orbital approach, as implemented in the Amsterdam density functional program (ADF).[56] Within this approach, the orbitals of a WC base pair can be expressed in terms of the molecular orbital of the individual nucleobases.[57,58] The charge transfer integral and the site energies involved in hole transport are directly obtained as the off-diagonal and diagonal matrix elements of the Kohn−Sham Hamiltonian matrix, $H_{KS} = SCEC^{-1}$, where S is the intermolecular overlap matrix, C is the molecular orbitals coefficient, and E is the molecular orbital energy. This method provides an accurate estimation of charge transfer integral due to the consideration of the spatial overlap between the nucleobases. For our calculation, we use asymptotically corrected exchange correlation potential SAOP (statistical average of orbital potentials)[59] with triple-ζ double polarization (TZ2P)[60] basis sets. This methodology has been applied earlier to describe charge transport properties of stacks of triphenylene molecules[61] and α-Ologofurans[62] and to explain the effect of fluctuation in base pair level in time of charge transfer in DNA[63,64] and site-selective photo-oxidation in DNA.[57,58]

3. RESULTS AND DISCUSSION

3.1. Structure and Energetic of the WC Base Pairs. In WC base pairs of natural nucleobases, A and T are connected by two H-bonds (N−H···N + N−H···O) while G and C are connected by three H-bonds (N−H···N + 2N−H···O). Here, we focus our attention on the nature of the H-bonding (WC base pairing) pattern in thieno-nucleobases and compare the results against their natural nucleobases. In order to model the H-bonding pattern in these nucleosides within out limited computational resource, we replace the sugar group by a methyl group, which does not impose major structural changes in nucleobases. This has been successfully used in previous studies concerning the base pairing of nucleobases.[29,31,57,65,66] We find that these modified fluorescence active nucleobases form H-bonding similar to the WC base pairing in natural nucleobases. The fully relaxed structures of all nucleosides and WC base pairs formed by both natural as well as thieno-nucleobases are shown in Figure 1b,c.

We find that the overall distance between H-bond donor and acceptor is slightly decreased for the A−U base pair, while there is an increment in separation for the G−C pair upon thieno modification. This can be clearly seen from the indicated distance values for the WC pairs in Figure 1. The small change in the H-bonding distance is associated with the changes in N−H bond length in individual nucleobases present in the WC H-bonded complex, which is caused by the difference in partial charge distributions in these two types of nucleobases. To compare the strength of H-bonding interactions between the WC base pairs of modified nucleobases and natural nucleobases, we calculate net stabilization energy (ΔE_{stab}) due to H-bonding in nucleobase pairs using the following equation:

$$\Delta E_{stab} = E_{(A-U/G-C)} - E_{(A/G)} - E_{(U/C)}$$

where $E_{(A-U/G-C)}$ is the optimized total energy of WC base pair A−U (or G−C), and $E_{(A-U/G-C)}$ is the optimized energy of individual nucleobases (A/G/U/C). The calculated stabilization energies of thA−thU (A−U) and thG−thC (G−C) are −12.80 kcal mol^{-1} (−12.88 kcal mol^{-1}) and −23.99 kcal mol^{-1} (−25.98 kcal mol^{-1}), respectively, obtained with B3LYP/6-31++g(d,p) level of calculations (see Table 1). The present findings are consistent with the previously reported theoretical results[36] for the natural nucleobases. We find that there is no significant change in the stabilization energy due to the WC H-bonded base pair formation between thieno-nucleobases compared to the natural nucleobases. Note that, although there is a very small change (~0.08 kcal mol^{-1}) in ΔE_{stab} for thA−thU compared to natural A−U, we find about 2 kcal mol^{-1} change in stabilization energy while going from thG−thC to G−C with

dx.doi.org/10.1021/jp301752k | J. Phys. Chem. B 2012, 116, 7618−7626

the later one showing a greater extent of stabilization. The small decrease in stabilization energy is consistent with the changes in N−H bond lengths of associated nucleobases. To test the reliability of the employed B3LYP exchange and correlation functional in calculating ΔE_{stab} for these H-bonded nucleobases, we also have computed the stabilization energy using two different energy functional; namely, M06-2X and WB97XD using the same basis set 6-31++g (d, p) as mentioned above. We find that although ΔE_{stab} is increased slightly, the overall trend in stabilization energy remains the same for these WC pairs of nucleobases. The small increase (3−5 kcal mol^{-1}) in ΔE_{stab} with respect to the usage of B3LYP functional is well consistent with the ability of incorporating medium and long-range van der Waals interactions by these two energy functionals. We also have calculated the ΔE_{stab} corrected with basis set superposition error (BSSE)[67] and zero point energy (ZPE), as tabulated in Table 1 for all these exchange and correlation functional used in the present study. We find a very small change in the stabilization energy, and interestingly, the overall findings still remain unchanged. Moreover, it is interesting to ask what happens to the H-bonding energy profile when the base pairs are modified by these present nucleobases. To shed light on this, we calculate an H-bonding energy profile with varying distance between the H-atom and the H-donor/acceptor atom of individual nucleobases. Earlier studies predict that the conventional H-bonding energy profile in WC base pairs shows a double-well potential with the bridged hydrogen (H) atom, forming a transitional state between two minimal energy states (D-H···A + D···H-A).[28,68]

For quantitative estimation of the H-bonding profiles in the present (modified) base pairs, we plot the energy profiles for the N−H···N bond for thA−thU and thG−thC and compare the same with the results obtained for A−U and G−C in Figure 2.

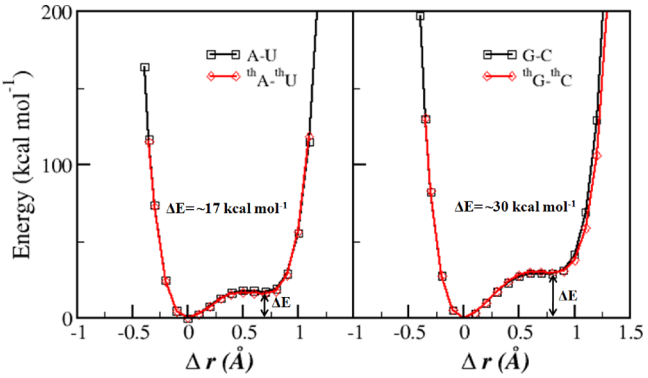


Figure 2. Potential energy profiles for the A−U, thA−thU, G−C, and thG−thC. The energies are scales so as to make the minimal energy configurations as zero in energy (corresponding to N−H···N for the gas phase). The proton position (*X*-axis) for the minimum energy state corresponds to Δr = 0, and movement to either side involves positive and negative Δr.

The nature of the double-well H-bonding energy profile, i.e., characteristic two energy minima, for the natural WC base pairs remain the same after theino modification. As shown in Figure 2, we find that the two energy minima are separated by ∼17 kcal mol^{-1} and ∼30 kcal mol^{-1} for A−U/thA−thU and G−C/thG−thC pairs, respectively. These results suggest that the thieno-modified base pairs can also form WC base pairing to form duplex DNA/RNA through the stable H-bonding interaction as found for natural nucleobases.

To compare and contrast the changes in aromatic character of individual aromatic rings (5-membered and 6-membered) present in natural nucleobases and in their thieno-modified analogues, we calculate nucleolus-independent chemical shift (NICS)[69] for all individual nucleobases as well as for their WC pairs. The NICS calculations provide a measure to account the electronic effects on the stability of the systems. It is expected that purine bases (adenine and guanine) and pyrimidine bases (uracil and cytosine) show distinct aromatic behavior. It should be mentioned that, in a set of earlier reports, NICS has been successfully used to measure the strength of aromatic character in a large number of organic molecules.[69−75] The NICS values are calculated using the gauge independent atomic orbital (GIAO) method.[75] For the calculations, probes, (Bq's) are placed at the center of the systems (NICS (0)) and latter, 1 Å distance above the molecular plane (NICS (1)).[73] We have verified our findings with results earlier reported for natural nucleobases (A, U, G, and C).[70,74] We find that thieno-modified bases show a decrease in the NICS value of the 6-membered ring of purine and perimidine bases, whereas it increases the NICS value of the 5-membered rings of purine bases (see Table 2). This indicates that, in the 6-membered ring

Table 2. Calculated NICS(0) and NICS(I) Values (in ppm) of Both Natural Nucleobases and Thieno-Nucleobases and Their WC Base Pairs[a]

	6-membered ring		5-membered ring	
nucleobases	NICS(0)	NICS(I)	NICS(0)	NICS(I)
A	−6.72	−8.40	−11.68	−9.89
thA	−3.43	−6.26	−13.94	−11.16
U	−1.58	−2.15		
thU	−0.42	−1.37	−12.66	−9.40
G	−3.16	−3.76	−11.49	−9.37
thG	−0.60	−2.27	−12.21	−9.46
C	−1.73	−3.42		
thC	−0.35	−2.55	−12.96	−9.53
A−U	−6.14 (A)	−7.71 (A)	−11.58 (A)	−9.77 (A)
	−1.51 (U)	−2.17 (U)		
thA−thU	−3.00 (thA)	−5.57 (thA)	−13.80 (thA)	−10.86 (thA)
	−0.41 (thU)	−1.41 (thU)	−12.57 (thU)	−9.24 (thU)
G−C	−3.46 (G)	−3.88 (G)	−11.23 (G)	−9.06 (G)
	−1.81 (C)	−1.81 (C)		
thG−thC	−1.21 (thG)	−2.70 (thG)	−12.06 (thG)	−9.07 (thG)
	−0.62 (thC)	−2.37 (thC)	−13.07 (thC)	−9.66 (thC)

[a]The values within brackets correspond to the results for the nucleobase present in the WC base pair.

of the thieno-nucleobases (thA, thU, thG, and thC), electrons are less delocalized in comparison to their natural analogues, whereas the extent of electrons delocalization is more in the 5-membered ring of thA and thG compared to their natural partners. We also calculate the NICS values for the WC base pair of natural as well as modified nucleobases to find the effect of H-bonding on the aromaticity of individual rings. We find that for the H-bonded pairs, A−U and thA−thU, NICS values decrease for both 5- and 6-membered rings compared to the free bases indicating less aromatic character, whereas, for the G−C and thG−thC pairs, the aromaticity of the 5- and 6-membered rings increases (decreases) for pyrimidine (purine) bases. For a better comparison, all the NICS values are provided in Table 2.

dx.doi.org/10.1021/jp301752k | J. Phys. Chem. B 2012, 116, 7618−7626

Table 3. Hole Transfer Integral, J, and Generalized Charge Transfer Integral, J', between WC and $\pi-\pi$ Stacked Base Pairs of Thieno-Nucleobases and Natural Nucleobases (Values within Bracket)

		J (eV)	J' (eV)
along transverse direction through WC base pair	thA–thU (A–U)	0.026 (−0.066)	0.007 (−0.023)
	thG–thC (G–C)	−0.065 (−0.069)	−0.025 (−0.025)
along strand direction through $\pi-\pi$ stacked base pair	thA–thA (A–A)	0.039 (0.267)	0.010 (0.093)
	thU–thU (U–U)	−0.375 (−0.164)	−0.131 (−0.049)
	thG–thG (G–G)	−0.402 (0.173)	−0.143 (0.064)
	thC–thC (C–C)	0.362 (0.277)	0.131 (0.112)

3.2. Site Energies and Generalized Charge Transfer Integrals. To estimate and compare the extent of charge transfer, particularly hole, through these DNA base pairs, we evaluate the charge transfer integral between the nucleobases of a WC base pair considering both the natural and thieno-nucleobases. Moreover, since the majority of charge carriers move through the π-stacked nucleobase channels along the strand direction compared to the transverse H-bonded nucleobase pair, we also consider calculating the charge transfer integrals in different stacked nucleobase pairs for both the natural and modified nucleobases. All the pairs of stacked nucleobases are optimized using fully dispersion corrected density functional wb97xd with the 6-31++g(d,p) basis set. The fully relaxed geometries of both the natural and thieno-modified π-stacked nucleobases are shown in Figure S3, Supporting Information. Here, we should mention that the extent of the transfer integrals is strongly dependent on the DNA/RNA base pair conformational dynamics.[63] The charge transfer through the H-bonded DNA base pairs is represented by the tight-binding Hamiltonian, which is given by, $H = \sum_i \varepsilon_i a_i^+ a_i + \sum_{i,j(i\neq j)} J_{i,j}\ (a_i^+ a_j + h.c.)$, where a_i^+ and a_i are the creation and annihilation operators of a charge at the i^{th} nucleobase, and J_{ij} is the charge transfer integral between the i^{th} and j^{th} nucleobases. The generalized charge transfer integral is calculated using the formula $J_{ij}' = J_{ij} - S_{ij}\ (\varepsilon_i + \varepsilon_j)/2$, where ε_i and ε_j are the site energies of the two nucleobases (i^{th} and j^{th} position) of a WC base pair, and S_{ij} is the overlap between them.

It was predicted that the site energies for the hole of isolated nucleobases increase in the order G < A < C < T.[57] We also observe the same trend for both natural nucleobases and thieno-modified nucleobases, where T is replaced by U. The site energies (ε_i) are 8.706 eV, 9.367 eV, 10.358 eV, and 10.432 eV for G, A, C, and U, respectively, whereas the same for the thieno-modified nucleobases are 8.520 eV, 9.206 eV, 9.883 eV, and 9.929 eV for thG, thA, thC, and thU, respectively. We find that, although the results follow similar trend in modified nucleobases, the magnitude is significantly reduced compared to the values obtained for natural nucleobases. This results from the destabilization of the highest occupied molecular orbital (HOMO) of nucleobases changes upon thieno modification. The calculated generalized hole transfer integrals between the nucleobases within WC base pairs are J' = −0.023 eV for A–U, J' = −0.025 eV for G–C, J' = −0.007 eV for thA–thU, and J' = −0.025 eV for thG–thC (see Table 3). It is clear from these values that the modified nucleobases are slightly less prominent to transfer a hole through the H-bonded WC base pairs. It also shows that the reduction in hole transfer integrals is quite significant in the case of thA–thU than in thG–thC when compared with the results obtained for natural nucleobase pairs. Interestingly, we find a larger amount of hole transfer integrals for π-stacked nucleobase pairs along the strand direction than in WC base pairs. This can be ascribed to the fact of large spatial overlaps between the two interstrand π-stacked nucleobases in comparison to the insignificant molecular overlap between the interstrand H-bonded WC base pairs.

One should note that the calculated values of charge transfer integrals are method dependent and have been used in the present study only for a comparison purposes. However, our finding fairly matches with the available theoretical results reported in previous studies.[35,57] Our results also suggest that thieno modification of nucleobases enhances the extent of charge transfer integrals through π-stacked base pairs, highlighting its potential use as an efficient hole transport material.

3.3. Optical Absorption. The absorption spectra of the natural nucleosides and modified thieno-nucleosides are calculated using TD-DFT formalism both in gas phase as well as in the presence of solvent (water). We plot the calculated absorption spectra of natural nucleosides and thieno-nucleosides in Figure 3 for the purpose of detailed comparison.

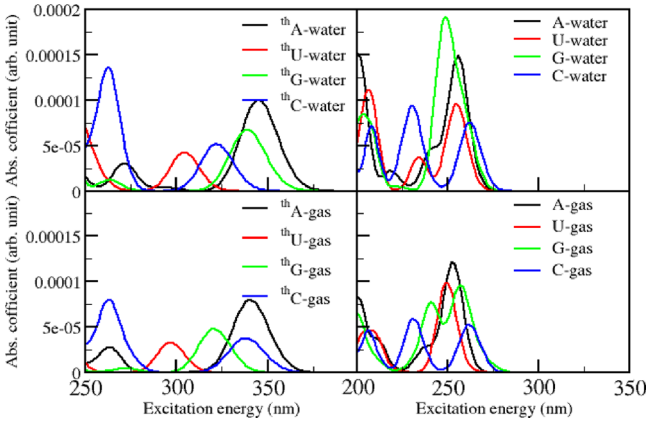


Figure 3. Absorption spectra of natural nucleosides and thieno-nucleosides.

We find the lowest energy absorption peaks at 345.09 nm, 304.47 nm, 338.92 nm, and 322.22 nm for thA, thU, thG, and thC, respectively, in water, which are in good agreement with the experimentally reported absorption peak values at 341 nm, 304 nm, 321 nm, and 320 nm for thA, thU, thG, and thC, respectively, in aqueous solvent. For comparison, we also calculate the absorption spectra for the natural RNA nucleosides, which show absorption peaks at 255.84 nm, 255.70 nm, 259.18 nm, and 262.27 nm for A, U, G, and C, respectively, which are at lower wavelength than the thieno-nucleosides. The earlier reported results[76,77] support our calculated absorption spectra for natural nucleosides. The lowest energy excitations and corresponding oscillator strengths for all the studied systems are given in the Table 4.

dx.doi.org/10.1021/jp301752k | J. Phys. Chem. B 2012, 116, 7618–7626

Table 4. HOMO−LUMO Gap (ΔE_{HL}) and Lowest Excitation Energies of the Natural Nucleosides and Thieno-Nucleosides

			lowest excitation energy and corresponding oscillator strength (os. str.)					
	gas phase	solvent phase (water)	calcd in gas		calcd in water		experimentally in water	
name of nucleoside	ΔE_{HL} (eV)	ΔE_{HL} (eV)	eV	os. str.	eV	os. str.	eV	nm
A	5.30	5.29	4.90	0.243	4.85	0.297	4.77	259.7
U	5.42	5.43	4.97	0.198	4.85	0.144		
G	5.24	5.35	4.81	0.153	4.78	0.149	4.91	252.7
C	5.36	5.38	4.73	0.105	4.73	0.150	4.58	270.7
thA	3.99	4.02	3.64	0.160	3.59	0.200	3.64	341.0
thU	4.71	4.64	4.17	0.066	4.07	0.085	4.08	304.0
thG	4.35	4.17	3.86	0.096	3.66	0.134	3.86	321.0
thC	4.17	4.40	3.66	0.075	3.85	0.103	3.87	320.0
A−U	4.61	4.84	4.86	0.266	4.84	0.342		
			4.96	0.101	4.91	0.131		
thA−thU	3.82	3.98	3.64	0.134	3.67	0.076		
			4.02	0.087	3.95	0.091		
G−C	3.83	4.60	4.75	0.087	4.64	0.126		
			4.80	0.093	4.84	0.154		
thG−thC	2.92	3.67	3.44	0.084	3.46	0.083		
			3.80	0.114	3.81	0.127		

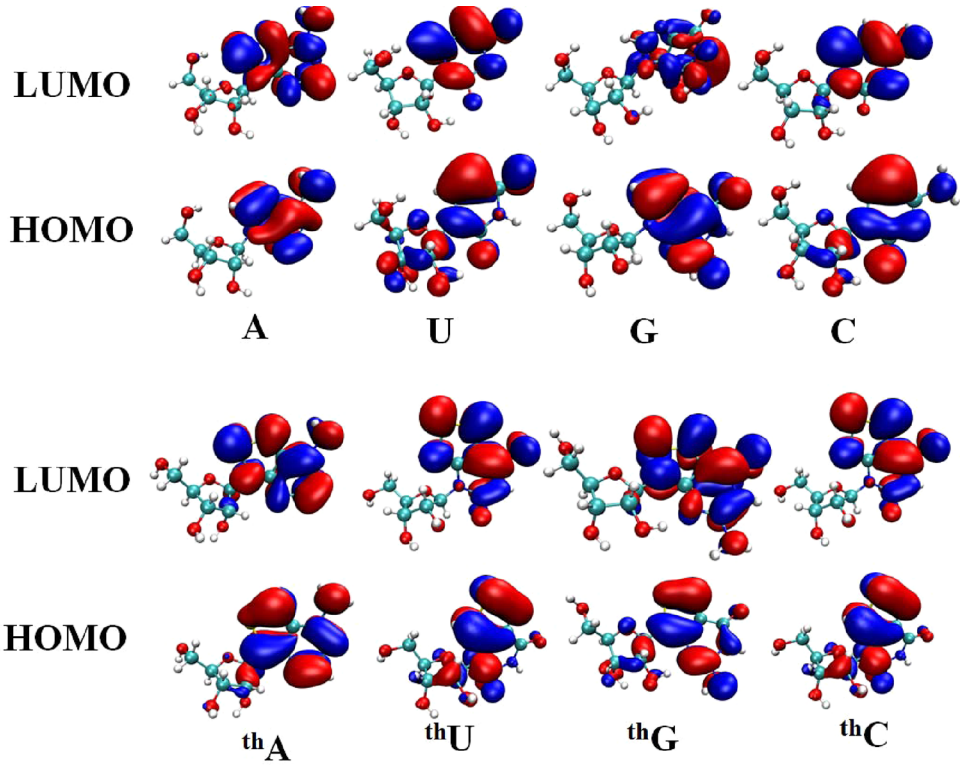


Figure 4. Calculated highest occupied (HOMO) and lowest unoccupied (LUMO) molecular orbital of natural nucleosides and thieno-nucleosides (in water).

To understand the origin of the absorption peaks, we analyze the relevant frontier molecular orbitals (FMOs) of the thieno-nucleoside as well as of the natural nucleoside involved in the lowest energy transitions. The highest occupied (HOMO) and the lowest unoccupied (LUMO) molecular orbital are shown in the Figure 4. The orbital diagram clearly shows that HOMO and LUMO are localized mainly on the base part of the nucleosides, depending on the nature of the nucleobases (purine or pyrimidine). We find that the lowest energy (highest wavelength, λ) transitions for thieno-nucleosides correspond to HOMO−LUMO transitions ($\pi-\pi^*$). Inclusion of water as the solvent within the PCM model reduces the HOMO−LUMO gap (see Table 2). This is related with exciton stabilization in the presence of water. Therefore, the λ_{max} value in water is higher compared to λ_{max} values obtained in the gas phase for lowest energy transition.

Here, we focus on the absorption spectra of WC H-bonded natural and thieno-modified nucleobases to focus on the H-bonding effect on the transition energy if there is any. Since the relevant molecular orbital of nucleosides are mainly localized on the nucleobases, it is expected that nucleobases will have the same absorption spectra as found in nucleosides. To confirm

dx.doi.org/10.1021/jp301752k | J. Phys. Chem. B 2012, 116, 7618−7626

this, we have also calculated the absorption spectra of nucleobases where the sugar group of the nucleoside is replaced by a methyl group. We find that the calculated absorption spectra for nucleosides and nucleobases are quite similar and that the changes in peak positions are negligible (see in Figure S1, Supporting Information). We consider four different WC base pairs (A−U, G−C, thA−thU, and thG−thC) of nucleobases and calculate the absorption spectra both in the gas phase and in the aqueous phase. The simulated absorption spectra are shown in the Figure 5. We find that the natural WC

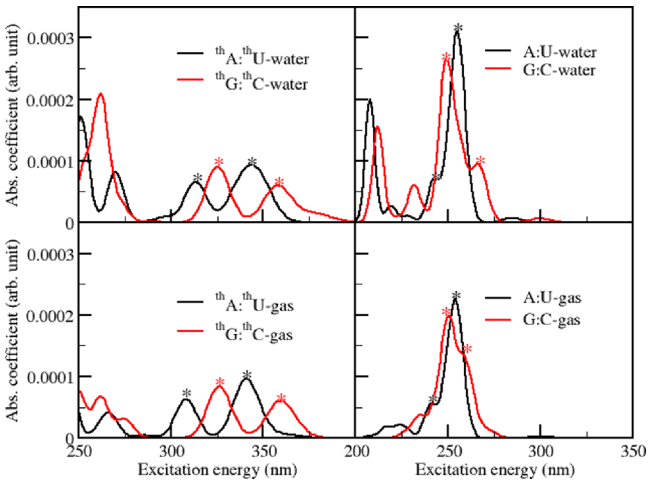


Figure 5. Absorption spectra of Watson−Crick (WC) base pairs of natural nucleobases and thieno-nucleobases. The asterisk (*) indicates the low-energy peaks in the absorption spectra.

base pairs (A−U and G−C) show the lowest energy absorption peaks near 250 nm as like for individual A, U, G, and C. Interestingly, our calculated results show that there are two new peaks for each thieno-nucleobase pairs (thA−thU and thG−thC) in the range 300−400 nm. The thA−thU (thG−thC) WC pairs show low energy absorption peaks at 347.76 and 313.50 nm (379.93 and 325.28 nm), whereas thA, thU, thG, and thC show low energy absorption peak at 345.09 nm, 304.47 nm, 338.92 nm, and 322.22 nm, respectively (Figure 3 and 5), in the presence of solvent water. We find significant red shift (∼10−20 nm) of the low energy absorption peaks of thieno-nucleobases due to the weak H-bonding interaction mediated through WC base pairing. This effect is termed as the fingerprint of hydrogen bonding in the optical spectra of natural DNA and has also been observed in the case of H-bonded size-expanded analogues of natural nucleobases.[37,78,79] It is expected that the weak H-bonding interactions modulate the electron density distribution in the frontier molecular orbitals of WC base pairs when compared to free nucleobases. We also analyze the origin of these low-energy peaks in terms of FMOs analysis relevant to these low-energy electronic transitions. The relevant FMOs of the WC base pairs of both the natural and thieno-modified nucleobases are shown in Figure S2, Supporting Information. We find that for both thA−thU and A−U nucleobase pairs, the low-energy absorption occur due to the HOMO → LUMO + 1 orbital transition. A closer look at these FMOs reveal that both HOMO and LUMO + 1 are localized on the A part of A−U, whereas, in the case of thA−thU nucleobase pair, the HOMO is localized on the thA, and the LUMO + 1 is distributed over both of the nucleobases, resulting in a low-energy shift of $\pi-\pi^*$ optical absorption. In the case of both G−C and thG−thC, we find that the low-frequency optical absorption occurs because of the electronic transitions from the HOMO − 1 to LUMO. Moreover, as can be seen from Figure S2, Supporting Information, the electron density distribution of these two FMOs is similar and is mainly localized on the C (thC) part of G−C (thG−thC) base pairs. Also note that, the red-shift in the optical absorption spectrum for thA−thU is larger compared to the low-energy shift obtained for the thG−thC base pair when compared with their natural analogues. This is because of the larger extent of LUMO + 1 and LUMO orbital distribution found for thA−thU than in thG−thC WC base pairs, respectively. Here, we also point out that the red-shift in optical absorption was also observed for the size-expanded analogues of natural nucleobases studied earlier.[37]

Additionally, we have also calculated the infrared spectra of these H-bonded WC base pairs of natural as well as thieno-modified nucleobases. Particularly, we focus on the stretching frequency of N−H bonds involved in forming WC H-bonds between complementary nucleobase pairs. The calculated N−H stretching frequency values are provided in Table 5 for various

Table 5. Calculated Normal Mode Vibrational Frequencies of N−H Bond Stretching Directly Involved in WC H-Bonding

	stretching frequency (cm^{-1})		stretching frequency (cm^{-1})	
types of H-bonds	thG−thC	G−C	thA−thU	A−U
O···H−N	3165.29	3150.57		
N−H···N	3236.82	3219.35	2943.77	2970.89
N−H···O	3423.15	3390.73	3386.81	3404.30

H-bonding environments. The calculated N−H stretching frequency values for WC pairs of natural nucleobases are in accordance with the previously reported theoretical and experimental results.[33,80,81] For the case of the A−U pair, the N−H stretching frequency in the N−H···N hydrogen bond is red-shifted by ∼27 cm^{-1} and the N−H···O hydrogen bond by ∼17 cm^{-1}, respectively, by the thieno modification of the nucleobases. For the G−C pair, the N−H stretching frequency in the N−H···N hydrogen bond is blue-shifted by ∼17 cm^{-1} and the same in the two N−H···O hydrogen bonds by ∼32 cm^{-1} and ∼14 cm^{-1}. The changes in N−H stretching frequencies are consistent with the changes in bond lengths due to the formation of WC H-bonds.

4. CONCLUSIONS

In summary, we have shown that the thieno-nucleosides form similar H-bonded WC base pairs as found in natural nucleosides using DFT calculations. We find the similar structural and bonding aspects resulting in comparable H-bonding energy among various WC base pairs in both natural as well as in modified nucleosides. Moreover, the calculated charge transfer integrals in different WC base pairs formed by various thieno-modified nucleobases are significantly reduced compared to their natural analogue, and the reduction is larger for the A−U than G−C pair. Interestingly, we find a greater extent of transfer integrals for the $\pi-\pi$ stacked modified nucleobases along the strand direction when compared with the values obtained for their natural analogues.We also have focused on the optical absorption characteristics of individual nucleosides as well as their H-bonded WC base pairs consisting

dx.doi.org/10.1021/jp301752k | J. Phys. Chem. B 2012, 116, 7618−7626

of both natural nucleosides and their thieno-analogues, in order to better understand the recently reported experimental findings. We find that all the thieno-nucleosides absorb light at lower energy (>300 nm) corresponding to the $\pi-\pi^*$ transition in comparison to their natural analogues, and the observed trend in the red shift is well consistent with the calculated HOMO−LUMO gaps. Therefore, we believe that our results have demonstrated the various aspects of WC base pairs formed by the natural as well as thieno-nucleosides present in RNA and have provided a microscopic understanding of the experimentally observed optical properties.

■ ASSOCIATED CONTENT

Supporting Information

Absorption spectra of natural and thieno modified nuclecosides and nucleobases, molecular orbital diagram of WC base pairs, and optimized structures of π-stacked WC base pairs. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pati@jncasr.ac.in.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

P.K.S. and A.K.M. thank CSIR for SRF, and S.K.P. acknowledges CSIR and DST, Govt. of India, for research grants.

■ REFERENCES

(1) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579−2619.
(2) de La Harpe, K.; Kohler, B. *J. Phys. Chem. Lett.* **2011**, *2*, 133−138.
(3) Gustavsson, T.; Improta, R.; Markovitsi, D. *J. Phys. Chem. Lett.* **2010**, *1*, 2025−2030.
(4) Ju, Y. S.; Kim, J.; Kim, S.; Hong, D.; Park, H.; Shin, J.-Y.; Lee, S.; Lee, W.-C.; Kim, S.; Yu, S.-B.; Park, S.-S.; Seo, S.-H.; Yun, J.-Y.; Kim, H.-J.; Lee, D.-S.; Yavartanoo, M.; Kang, H. P.; Gokcumen, O.; Govindaraju, D. R.; Jung, J. H.; Chong, H.; Yang, K.-S.; Kim, H.; Lee, C.; Seo, J.-S. *Nat. Genet.* **2011**, *43*, 745−747.
(5) Kohler, B. *J. Phys. Chem. Lett.* **2010**, *1*, 2047−2053.
(6) Markovitsi, D.; Gustavsson, T.; Vayaḷ, I. *J. Phys. Chem. Lett.* **2010**, *1*, 3271−3276.
(7) Van Atta, R. B.; Bernadou, J.; Meunier, B.; Hecht, S. M. *Biochemistry* **1990**, *29*, 4783−4789.
(8) Bartel, D. P. *Cell* **2009**, *136*, 215−233.
(9) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265−4274.
(10) Matray, T. J.; Kool, E. T. *Nature* **1999**, *399*, 704−708.
(11) Zhang, L.; Chen, X.; Liu, H.; Han, L.; Cukier, R. I.; Bu, Y. *J. Phys. Chem. B* **2010**, *114*, 3726−3734.
(12) Zhang, L.; Li, H.; Chen, X.; Cukier, R. I.; Bu, Y. *J. Phys. Chem. B* **2009**, *113*, 1173−1181.
(13) Gwinn, E. G.; O'Neill, P.; Guerrero, A. J.; Bouwmeester, D.; Fygenson, D. K. *Adv. Mater.* **2008**, *20*, 279−283.
(14) Kim, J. Y.; Lee, J. S. *Nano Lett.* **2009**, *9*, 4564−4569.
(15) Sharma, J.; Yeh, H.-C.; Yoo, H.; Werner, J. H.; Martinez, J. S. *Chem. Commun.* **2010**, *46*, 3280−3282.
(16) Yu, J.; Choi, S.; Dickson, R. M. *Angew. Chem., Int. Ed.* **2009**, *48*, 318−320.
(17) Hawkins, M. E. *Cell Biochem. Biophys.* **2001**, *34*, 257−281.
(18) Lee, A. H. F.; Kool, E. T. *J. Org. Chem.* **2005**, *70*, 132−140.
(19) Lee, A. H. F.; Kool, E. T. *J. Am. Chem. Soc.* **2006**, *128*, 9219−9230.
(20) Liu, H. B.; Gao, J. M.; Kool, E. T. *J. Org. Chem.* **2005**, *70*, 639−647.
(21) Scopes, D. I.; Barrio, J. R.; Leonard, N. J. *Science* **1977**, *195*, 296−298.
(22) Secrist, J. A., III; Barrio, J. R.; Leonard, N. J. *Science* **1972**, *175*, 646−647.
(23) Secrist, J. A., III; Barrio, J. R.; Leonard, N. J.; Weber, G. *Biochemistry* **1972**, *11*, 3499−506.
(24) Seela, F.; Zulauf, M. *Chem.—Eur. J.* **1998**, *4*, 1781−1790.
(25) Seela, F.; Zulauf, M.; Sauer, M.; Deimel, M. *Helv. Chim. Acta* **2000**, *83*, 910−927.
(26) Zhao, Y.; Knee, J. L.; Baranger, A. M. *Bioorg. Chem.* **2008**, *36*, 271−277.
(27) Rudy, B.; Gitler, C. *Biochim. Biophys. Acta* **1972**, *288*, 231−236.
(28) Saxena, R.; Shrivastava, S.; Chattopadhyay, A. *J. Phys. Chem. B* **2008**, *112*, 12134−12138.
(29) Brancolini, G.; Di Felice, R. *J. Chem. Phys.* **2011**, *134*, 205102.
(30) Liu, H.; Li, G.; Ai, H.; Li, J.; Bu, Y. *J. Phys. Chem. C* **2011**, *115*, 22547−22556.
(31) Miyachi, H.; Matsui, T.; Shigeta, Y.; Hirao, K. *Phys. Chem. Chem. Phys.* **2009**, *12*, 909−917.
(32) Shapir, E.; Brancolini, G.; Molotsky, T.; Kotlyar, A. B.; Di Felice, R.; Porath, D. *Adv. Mater.* **2011**, *23*, 4290−4294.
(33) Frey, J. A.; Leist, R.; Leutwyler, S. *J. Phys. Chem. A* **2006**, *110*, 4188−4195.
(34) Khakshoor, O.; Wheeler, S. E.; Houk, K. N.; Kool, E. T. *J. Am. Chem. Soc.* **2012**, *134*, 3154−3163.
(35) Migliore, A.; Corni, S.; Varsano, D.; Klein, M. L.; Di Felice, R. *J. Phys. Chem. B* **2009**, *113*, 9402−9415.
(36) Sponer, J.; Jurecka, P.; Hobza, P. *J. Am. Chem. Soc.* **2004**, *126*, 10142−10151.
(37) Varsano, D.; Garbesi, A.; Di Felice, R. *J. Phys. Chem. B* **2007**, *111*, 14012−14021.
(38) Shin, D.; Sinkeldam, R. W.; Tor, Y. *J. Am. Chem. Soc.* **2011**, *133*, 14912−14915.
(39) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H. P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2009.
(40) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652.
(41) Lee, C. T.; Yang, W. T.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785−789.
(42) Miehlich, B.; Savin, A.; Stoll, H.; Preuss, H. *Chem. Phys. Lett.* **1989**, *157*, 200−206.
(43) Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215−241.
(44) Chai, J. D.; Head-Gordon, M. *Phys. Chem. Chem. Phys.* **2008**, *10*, 6615−6620.
(45) Chai, J. D.; Head-Gordon, M. *J. Chem. Phys.* **2008**, *128*, 84106.
(46) Antony, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2008**, *10*, 2722−2729.
(47) Gu, J.; Wang, J.; Leszczynski, J.; Xie, Y.; Schaefer, H. F., III. *Chem. Phys. Lett.* **2009**, *473*, 209−210.
(48) Hobza, P. *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* **2011**, *107*, 148−168.
(49) Ivanova, B. B.; Spiteller, M. *J. Mol. Struct.* **2011**, *1003*, 1−9.
(50) Mackie, I. D.; DiLabio, G. A. *J. Phys. Chem. A* **2008**, *112*, 10968−10976.

(51) Martiniano, H. F. M. C.; Cabral, B. J. C. *J. Mol. Struct.* **2010**, *946*, 26−32.
(52) Salzner, U.; Aydin, A. *J. Chem. Theory Comput.* **2011**, *7*, 2568−2583.
(53) Solomon, G. C.; Vura-Weis, J.; Herrmann, C.; Wasielewski, M. R.; Ratner, M. A. *J. Phys. Chem. B* **2010**, *114*, 14735−14744.
(54) Zhao, Y. D.; Liao, Y. *Adv. Mater. Res.* **2011**, *322*, 120−124.
(55) Scalmani, G.; Frisch, M. J. *J. Chem. Phys.* **2010**, *132*, 114110.
(56) Velde, G. T.; Bickelhaupt, F. M.; Baerends, E. J.; Guerra, C. F.; Van Gisbergen, S. J. A.; Snijders, J. G.; Ziegler, T. *J. Comput. Chem.* **2001**, *22*, 931−967.
(57) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Lewis, F. D.; Berlin, Y. A.; Ratner, M. A.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2005**, *127*, 14894−14903.
(58) Senthilkumar, K.; Grozema, F. C.; Guerra, C. F.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Am. Chem. Soc.* **2003**, *125*, 13658−13659.
(59) Chong, D. P.; Gritsenko, O. V.; Baerends, E. J. *J. Chem. Phys.* **2002**, *116*, 1760.
(60) Snijders, J. G.; Vernooijs, P.; Baerends, E. J. *At. Data Nucl. Data Tables* **1981**, *26*, 483−509.
(61) Senthilkumar, K.; Grozema, F. C.; Bickelhaupt, F. M.; Siebbeles, L. D. A. *J. Chem. Phys.* **2003**, *119*, 9809−9817.
(62) Mohakud, S.; Alex, A. P.; Pati, S. K. *J. Phys. Chem. C* **2010**, *114*, 20436−20442.
(63) Mallajosyula, S. S.; Gupta, A.; Pati, S. K. *J. Phys. Chem. A* **2009**, *113*, 3955−3962.
(64) Mallajosyula, S. S.; Pati, S. K. *J. Phys. Chem. Lett.* **2010**, *1*, 1881−1894.
(65) Jurecka, P.; Nachtigall, P.; Hobza, P. *Phys. Chem. Chem. Phys.* **2001**, *3*, 4578−4582.
(66) Zhang, H. Y.; Calzolari, A.; Di Felice, R. *J. Phys. Chem. B* **2005**, *109*, 15345−15348.
(67) Boys, S. F.; Bernardi, F. *Mol. Phys.* **1970**, *19*, 553−566.
(68) Lowdin, P. O. *Rev. Mod. Phys.* **1963**, *35*, 724−732.
(69) Schleyer, P. V.; Maerker, C.; Dransfeld, A.; Jiao, H. J.; Hommes, N. *J. Am. Chem. Soc.* **1996**, *118*, 6317−6318.
(70) Cysewski, P. *J. Mol. Struct.* **2005**, *714*, 29−34.
(71) Mohan, P. J.; Datta, A.; Mallajosyula, S. S.; Pati, S. K. *J. Phys. Chem. B* **2006**, *110*, 18661−18664.
(72) Rehaman, A.; Datta, A.; Mallajosyula, S. S.; Pati, S. K. *J. Chem. Theory Comput.* **2006**, *2*, 30−36.
(73) Stanger, A. *J. Org. Chem.* **2006**, *71*, 883−893.
(74) Sun, G.; Nicklaus, M. C. *Theor. Chem. Acc.* **2007**, *117*, 323−332.
(75) Wolinski, K.; Hinton, J. F.; Pulay, P. *J. Am. Chem. Soc.* **1990**, *112*, 8251−8260.
(76) Improta, R.; Barone, V. *Theor. Chem. Acc.* **2008**, *120*, 491−497.
(77) Onidas, D.; Markovitsi, D.; Marguet, S.; Sharonov, A.; Gustavsson, T. *J. Phys. Chem. B* **2002**, *106*, 11367−11374.
(78) Varsano, D.; Di Felice, R.; Marques, M. A. L.; Rubio, A. *J. Phys. Chem. B* **2006**, *110*, 7129−7138.
(79) Wesolowski, T. A. *J. Am. Chem. Soc.* **2004**, *126*, 11444−11445.
(80) Krishnan, G. M.; Kuhn, O. *Chem. Phys. Lett.* **2007**, *435*, 132−135.
(81) Yang, M.; Szyc, L.; Rottger, K.; Fidder, H.; Nibbering, E. T. J.; Elsaesser, T.; Temps, F. *J. Phys. Chem. B* **2011**, *115*, 5484−5492.