Chemical Physics 515 (2018) 434–440

Contents lists available at ScienceDirect

Chemical Physics

journal homepage: www.elsevier.com/locate/chemphys

The origin of fluorescence in DNA thio-analogues

Tolga N.V. Karsili*, Mushir Thodika, Linh Nguyen, Spiridoula Matsika*

Department of Chemistry, Temple University, Philadelphia, PA 19122, USA

ARTICLE INFO

Keywords:
Luminescent probes
Fluorescence
Nucleobase analogues
Photophysics
Conical intersections

ABSTRACT

Fluorescent probes are vital molecular tools for DNA biognostics – often aiding in the detection of malignant tumors. Biomimetic nucleobase analogues are amongst the most widely used probes due to their versatility and biocompatibility. Thieno[3,4-*d*]pyrimidin-4-amine (TPA) and isothiazolo[4,3-*d*]pyrimidin-7-amine (iTPA) are two recently developed sulphur-containing heteroaromatic fluorescent nucleobase analogues. Their inherent structures give rise to superior biomimetic properties that ensure their optimal interaction with double-helix DNA. Using high-level electronic structure methods, we simulate the absorption and emission spectra in order to better understand the origins of the strong absorption and emission profiles observed experimentally. Radiationless decay pathways are also computed and compared with those of the canonical DNA nucleobases, in order to unravel the geometric and electronic properties that make the thio-analogues fluorescent.

1. Introduction

Fluorescence is a spin-allowed radiative decay process, in which light emission occurs *via* relaxation of excited state population back to the ground electronic state, following molecular electronic excitation. Fluorescence typically occurs on a nanosecond timescale, i.e. much longer than internal nuclear motions and thus after molecular relaxation [1]. Bands in emission spectra typically show a characteristic shifted mirror image of the absorption profile. The shift in absorption/emission is a molecule specific quantity termed the Stokes shift [2]. In general, when designing fluorescent analogues the desired properties are: sufficiently red-shifted absorption compared to the absorption of the natural bases, so that the analogue can be selectively excited when it is incorporated in nucleic acids, increased quantum yield for fluorescence, and some dependence of the fluorescence on the environment, so that the environment can be probed based on fluorescence techniques [3,4].

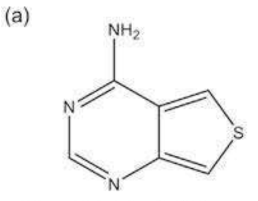
To date, many organic fluorophores have served as efficient luminescent probes for bioimaging [5–7]. DNA/RNA nucleobase analogues represent an important class of organic fluorophores that contain enhanced biocompatibility and high fluorescence quantum yields [8–10]. Their biocompatibiliy arises *via* their structural similarities to the canonical DNA/RNA nucleobases – manifesting in favorable intercalation with bulk double helix DNA and single-helix coiled RNA. The biologically relevant nucleobases are known to have ultrashort excited-state lifetimes, rapidly dissipating the excess energy imparted by electronic excitation and reforming the ground state parent structure (*via* ultrafast internal conversion) with almost unity quantum yields [11–15]. Understanding the mechanisms responsible for the ultrashort lifetimes of biologically relevant nucleobases provides ideas on how to modify the bases in order to increase the lifetimes of their electronically excited states and enhance their fluorescence quantum yields, while maintaining optimal interactions with the biological environment. This is an important criterion of any luminescent probe [3].

In the present manuscript we explore the photophysics of two recently developed nucleobase analogues, Thieno[3,4-*d*]pyrimidin-4-amine (henceforth TPA, see Fig. 1(a) for molecular structure) and isothiazolo[4,3-*d*]pyrimidin-7-amine (henceforth iTPA, see Fig. 1(b) for molecular structure) in order to unravel the electronic and geometric reasons for their enhanced fluorescence quantum yields [16,17]. These analogues were developed by Tor and coworkers as part of their efforts to create an emissive RNA alphabet [16], and show some analogies with canonical thio-analogues [18–20]. Initially analogues of all four RNA bases were synthesized with thiophene constituting the five membered ring, and TPA is part of this alphabet. In order to maintain some functionality of natural nucleobases the second series of analogues introduced a nitrogen heteroatom to the five-membered ring, giving rise to iTPA [17]. We have used high-level multi-reference electronic structure methods, as well as density functional theory, to compute the absorption and emission spectra of TPA and iTPA. In order to examine radiationless decay pathways we have also optimized conical intersections (CIs), that could mediate internal conversion to the ground state, and the associated excited state energy profiles leading to them. Even though CIs can be found in fluorescent molecules, fluorescence is not

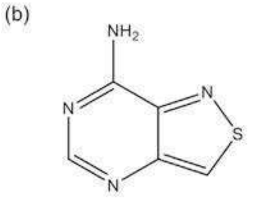
---

* Corresponding authors.
E-mail address: smatsika@temple.edu (S. Matsika).

https://doi.org/10.1016/j.chemphys.2018.08.049
Received 18 May 2018; Accepted 29 August 2018



thieno[3,4-*d*]pyrimidin-4-amine



isothiazolo[4,3-*d*]pyrimidin-7-amine

Fig. 1. Molecular structures associated with (a) TPA and (b) iTPA.

quenched since the CIs cannot be easily accessed, as will be shown in this work [3].

2. Computational and theoretical methodology

The ground state minimum energy geometry of TPA and iTPA were optimized (both in the gas phase and in aqueous solution) using the Becke, 3-parameter, Lee-Yang-Parr (B3LYP) [21,22] functional of Density Functional Theory (DFT), coupled to Pople's 6-31G(d) [23] basis set. The vibrational frequencies were also computed to confirm the optimized minima. The excited S_1 state minimum geometry for both molecules was also computed at the time-dependent DFT (TDDFT)/6-31G(d) level using the B3LYP and the CAM-B3LYP functionals, with and without solvation effect.

The polarizable continuum model, using the integral equation formalism variant (IEFPCM) as implemented in Gaussian [24], was used for water solvent [25]. The optimizations of the ground and excited states allowed for equilibrium solvation. The single point calculations used for the absorption spectra (see below) utilized the linear-response IEFPCM method and did not allow for solvent relaxations. This was done since absorption occurs too fast for complete solvent relaxation to occur.

The vertical excitation energies and oscillator strengths, associated with electronically excited states, were computed with several approaches. TDDFT, exploiting both the B3LYP/6-31G(d) and CAM-B3LYP/6-31G(d) functional/basis set, was used both in the gas phase as well as using the PCM model for aqueous solutions. In addition the gas phase results were tested using several high level electronic structure approaches. The (single-state) complete active space second-order perturbation theory (CASPT2) [26], coupled to Dunning's correlation-consistent basis set of double-ζ quantity (cc-pVDZ) [27] was used in the gas phase. The CASPT2/cc-pVDZ computations were based on a six state averaged complete active space self-consistent field (SA6-CASSCF) reference wavefunction and comprised an active space of twelve electrons in eleven orbitals (12,11). An imaginary level shift of 0.5 E_H was used to aid convergence and to mitigate the involvement of intruder state effects. This value has been shown to work well in related systems [28–30]. The equation of motion coupled cluster with singles and doubles (EOM-CCSD) with the 6-311G(d) and 6-311 + G(d) basis sets was also used to compute the excitation energies in gas phase.

The absorption and emission spectra were computed using our in-house SArCASM (Simulating AbsoRption spectra of Complex and Solvated Molecules) program and were based on the B3LYP/6-31G(d) and TD-B3LYP/6-31G(d) (or TD-CAM-B3LYP/6-31G(d)) ground and excited state equilibrium geometries, respectively. In SArCASM, 100 initial ground (for absorption) or excited (for emission) state geometries are prepared using a Wigner distribution [31] (using a locally modified version of the Newton-X sub-program [32]) based on the normal-modes associated with the global ground or excited state geometry. At each Wigner geometry, vertical excitation and emission energies and transition dipole moments (μ_{ij}) were computed using either the TD-B3LYP/PCM or the CASPT2 approach. The gas phase B3LYP minima were used for the CASPT2 spectra while the aqueous B3LYP minima (except the iTPA emission case where CAM-B3LYP was used) were used for the aqueous spectra. The excitation energy dependent photoabsorption cross section $P(E)$ was then obtained using Eq. (1).

$$P(E) = \frac{\pi e^2}{2 m_e c \, \epsilon_0} \sum_{j=1}^{1(\text{or } 4)} \left[\frac{1}{N_{TOT}} \sum_{N=1}^{N_{TOT}} f_{ij}^N \, g(E-\Delta E_{ij}^N, \delta) \right] \tag{1}$$

where g is a Lorentzian line shape function given by Eq. (2),

$$g(E-\Delta E_{ij}^N, \delta) = \frac{\hbar\delta}{2\pi} \left((E-\Delta E_{ij}^N)^2 + \left(\frac{\delta}{2}\right)^2 \right)^{-1}. \tag{2}$$

f_{ij} is the oscillator strength given by Eq. (3)

$$f_{ij}^N = \frac{2}{3} (\Delta E_{ij}^N) \sum_{\alpha=x,y,z} |\mu_{ij}^N|_\alpha^2 \tag{3}$$

and $\Delta E_{ij}^N = (E_j^N - E_i^N)$. m_e and e are the mass and charge of electron, respectively, while c is the speed of light. The internal sum in Eq. (1) is expressed over the set of total Wigner geometries ($N_{TOT} = 100$) while the external sum includes transitions from the initial state i (*i.e.*, the S_0 state) to final state j (*i.e.* S_1, S_2, S_3 and S_4 in the present modelling) with respective oscillator strengths f_{ij}^N as given by Eq. (3). δ is a broadening factor, which is arbitrarily set to 0.2 eV for each of the calculated absorption and emission profiles reported herein.

CIs, between the ground and first electronically excited state, were optimized at the SA2-CASSCF/6-31G(d) level using an active space of six electrons in six orbitals (6,6). Single point energies using the CASSCF optimized geometries were obtained at the CASPT2 and B3LYP levels. In that case the energies of the S_0 and S_1 states are not exactly degenerate so we used the average of the two energies as the energy at the CI.

Potential energy profiles between the ground state equilibrium geometry as the starting structure and the CI geometries as final structures were computed by means of linear interpolation in internal coordinates (LIIC) using TD-B3LYP/PCM and CASPT2.

All optimizations and DFT calculations were carried out in Gaussian 09 [24] while all CASPT2 calculations were carried out in Molpro 2015 [33]. Q-Chem [34] was used for the EOM-CCSD computations.

3. Results and discussion

3.1. Ground and excited state geometries and excitation energies

The ground state minimum energy geometries of TPA and iTPA in the gas phase and in solution phases are displayed in Fig. 2. The atomic numberings given on the molecular structures will be used henceforth in the remainder of this manuscript. All cartesian coordinates are reported in the Supporting Information (SI). S_0 minima in solution do not differ much from the gas phase minima. In both TPA and iTPA, all ring-centered atoms are in a common ring plane – highlighting the favorably conjugated π-system in the electronic ground state. The amino substituent at the C6 position displays a pyramidal geometry with respect to the ring-plane in TPA.

Tables 1 and 2 list the vertical excitation energies for the first two electronically excited singlet states of TPA and iTPA obtained at all the theoretical levels used in this work. Fig. 3 displays the corresponding orbital promotions associated with the electronically excited states at the B3LYP level. Because of the energetic proximity of the first two excited states their ordering depends on the level of theory. In TPA, all

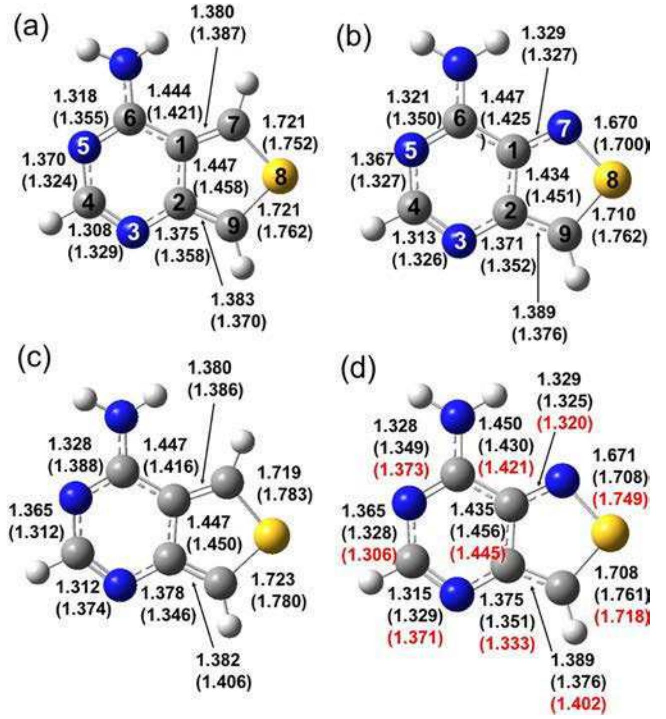


Fig. 2. Ground state minimum energy geometry of (a) gas phase TPA, (b) gas phase iTPA, (c) aqueous TPA and (d) aqueous iTPA. The displayed bond distances in Å are given for the B3LYP ground state equilibrium geometry. Corresponding bond lengths of the equilibrium geometry of the first electronically excited state at the B3LYP level are given in parenthesis. For the iTPA S_1 minimum both the B3LYP minimum and the CAM-B3LYP (in red) geometries are given in (d). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

methods except CASPT2 predict the first excited state to have $\pi\pi^*$ character and S_2 to have $n\pi^*$ character. Adding solvation at the TDDFT/PCM level maintains this ordering. The experimental absorption and emission maxima are also given in Table 1. B3LYP/PCM gives vertical excitation energies in better agreement with experimental absorption maxima, 0.1–0.2 eV higher than the experimental values [16]. On the other hand, the other methods overestimate the excitation energy to the $\pi\pi^*$ excited state by 0.4–0.6 eV. The $\pi\pi^*$ transition is accompanied by a strong oscillator strength manifesting from the appreciable overlap between the initial and final orbitals involved in the electronic excitation. In contrast, $n\pi^*$ state is accompanied by a comparatively weak oscillator strength manifesting from the poorer spatial overlap between the initial and final orbitals involved in the electronic excitation. In iTPA, most methods predict the two S_1 and S_2 states to be almost degenerate, and the energetic ordering of the $\pi\pi^*$ and $n\pi^*$ states is even more sensitive on the level of theory and the environment. B3LYP/PCM provides the vertical excitation energies with the best agreement to the experimental absorption maximum. Overall, it appears that the B3LYP/PCM model provides the best energies for the $\pi\pi^*$ states for both molecules, but it may not describe the $n\pi^*$ states as well. In all cases CAM-B3LYP destabilizes the $n\pi^*$ states. This is expected since these states have increased charge transfer character compared to the $\pi\pi^*$ states and CAM-B3LYP is designed to correct the stabilization of charge transfer states present in B3LYP. As a result B3LYP will give accurate absorption spectra but there may be some limitations when describing the emission properties. We will see this in more detail below. The CASPT2 and EOM-CCSD methods give somewhat overestimated excitation energies for the $\pi\pi^*$ states compared to the experimental spectra, but they provide a more balanced description between the different character states. The better agreement of the B3LYP energies to experiment compared to the higher level CASPT2 and CCSD is most likely a fortuitous cancellation of errors. Finally, solvation effects (which were only calculated at the TDDFT level) are very small for the $\pi\pi^*$ states while they blueshift the $n\pi^*$ states by about 0.2–0.3 eV. A blueshift for these states is expected.

The ordering of $\pi\pi^*$ and $n\pi^*$ states plays an important role on emission since it may determine the character of the S_1 state at its minimum. This is clearly seen in this work. We have optimized the S_1 state at the B3LYP and CAM-B3LYP levels, both in the gas phase and solution. For TPA CAM-B3LYP predicts a $\pi\pi^*$ minimum in both gas phase and solution while B3LYP predicts a $\pi\pi^*$ minimum in solution but an $n\pi^*$ minimum in the gas phase. Since we are modeling spectra in solution it is clear that a $\pi\pi^*$ minimum is responsible for the emission. Table 1 shows the vertical emission energies at the S_1 minima. Again B3LYP/PCM gives the best agreement with the experimental emission maximum in TPA.

In iTPA both B3LYP and CAM-B3LYP predict an $n\pi^*$ minimum in gas phase. In solution however B3LYP predicts an $n\pi^*$ minimum while CAM-B3LYP predicts a $\pi\pi^*$ minimum. Modeling the experimental spectrum, as will be discussed in the next section, indicates that the minimum corresponds to a $\pi\pi^*$ transition, so in solution we expect a $\pi\pi^*$ minimum to be leading to fluorescence, although an $n\pi^*$ minimum may also be present leading to some fluorescence quenching.

Fig. 2 displays the equilibrium geometry of the first electronically excited state (S_1) of TPA and iTPA in the gas phase and aqueous solution. The gas phase minimum for both molecules at the B3LYP level corresponds to an $n\pi^*$ state, and this is evident in the distortions going from the S_0 to S_1 minimum being similar for both molecules. The S_1 minimum in aqueous solution however has $\pi\pi^*$ character, exhibiting different distortions. In general, the double bonds elongate more for the $\pi\pi^*$ minimum compared to the $n\pi^*$ minimum.

The experimental Stokes shift for TPA is 0.68 eV, in good agreement with the TDDFT/PCM value of 0.77 eV (calculated as the difference between absorption and emission vertical excitation energies).

Table 1
Vertical absorption (calculated at S_0 minimum) and emission (calculated at S_1 minimum) energies in eV and oscillator strengths f (in parenthesis) for TPA. *Indicates reversal of the character between S_1 and S_2 states. Experimental values are taken from Ref. [16].

Geometry	$S_{0,min}$	$S_{1,min}(n\pi^*)$		$S_{1,min}(\pi\pi^*)$
Electronic state	S_1 ($\pi\pi^*$)	S_2 ($n\pi^*$)	S_1	S_1
B3LYP/6-31G(d)	3.84 (9.9E−2)	3.93 (1.2E−2)	2.78 (4.7E−3)	
CAM-B3LYP/6-31G(d)	4.13 (0.14)	4.40 (2.69E−3)		3.33 (0.11)
EOM-CCSD/6-311G(d)	4.43 (0.17)	4.62 (5.0E−3)		
EOM-CCSD/6-311+G(d)	4.32 (0.18)	4.60 (4.58E−3)		
CASPT2/cc-pvdz	4.02* (6.1E−3)	4.31* (3.7E−2)		
B3LYP/PCM/6-31G(d)	3.83 (0.14)	4.17 (1E−3)		3.06 (0.18)
CAM-B3LYP/PCM/6-31G(d)	4.11 (0.18)	4.65 (2E−3)		3.27 (0.24)
exp	3.63			2.95

Table 2
Vertical absorption (calculated at S_0 minimum) and emission (calculated at S_1 minimum) energies in eV and oscillator strengths f (in parenthesis) for iTPA. *Indicates reversal of the character between S_1 and S_2 states. Experimental values are taken from Ref. [17].

Geometry	$S_{0,min}$		$S_{1,min}(n\pi^*)$	$S_{1,min}(\pi\pi^*)$
Electronic state	S_1 $(n\pi^*)$	S_2 $(\pi\pi^*)$	S_1	S_1
B3LYP/6-31G(d)	3.62 (3.0E − 4)	3.81 (0.11)	2.52 (4.0E − 3)	
CAM-B3LYP/6-31G(d)	4.12 (7.0E − 4)	4.13 (0.15)	2.96 (7.0E − 4)	
EOM-CCSD/6-311G(d)	4.39 (7.1E − 4)	4.40 (0.19)		
EOM-CCSD/6-311 + G(d)	4.31* (0.20)	4.37* (9.9E − 4)		
CASPT2/cc-pvdz	4.33* (0.27)	4.36* (7.8E − 4)		
B3LYP/PCM/6-31G(d)	3.81 (6.0E − 3)	3.81 (0.14)	2.68 (4.7E − 3)	
CAM-B3LYP/PCM/6-31G(d)	4.11* (0.20)	4.30* (4.0E − 4)		3.29 (0.15)
exp	3.67			3.02

Similarly, for iTPA, the experimental Stokes shift of 0.65 eV (when using the CAM-B3LYP $\pi\pi^*$ minimum) compares well with the theoretical of 0.52 eV.

3.2. Simulated absorption and emission spectra

The simulated absorption and emission spectra for TPA and iTPA in aqueous solution are displayed in Figs. 4(a) and (b), respectively. The experimental spectra taken from Ref. [16,17] are superimposed with those that were simulated. The absorption profiles were simulated by means of sampling 100 Wigner geometries around the ground state minimum energy geometry. The simulated spectra for both TPA and iTPA, in aqueous solution, agree remarkably well with the experimentally measured absorption maxima, including minor peak details [16,17]. More specifically the absorption spectrum for TPA shows a maximum around 330 nm and two smaller peaks at shorter wavelengths in agreement with the spectrum taken by Tor and coworkers [16]. The absorption spectrum of iTPA shows the main peak and a smaller peak at shorter wavelength, again in agreement with the experimental spectrum [17]. For completeness the gas phase absorption spectra, computed using CASPT2, are displayed in the SI (Figs. S7, S8). These spectra do not agree as well with experiment highlighting the importance of solvation as well as methodology. In addition, the small number of states included in the CASPT2 calculations did not allow for any description of peaks at shorter wavelengths.

The emission profiles based on a Wigner distributions of the S_1 minima are also shown in Fig. 4. The emission spectrum for TPA is also in very good agreement with experiment [16]. There is some overall blueshift but in general the spectrum agrees well with experiment. In the case of iTPA we should first remember that the B3LYP and CAM-B3LYP methods gave different minima. The former gave an $n\pi^*$ minimum while the latter gave a $\pi\pi^*$ minimum. Fig. S9 shows the emission spectrum obtained from the $n\pi^*$ minimum. It is clear that this minimum does not reproduce the experimental spectrum. It has very low intensity and the spectrum is quite broad because of some mixing with the S_2 $\pi\pi^*$ state. Even a small amount of mixing leads to a big effect in the spectrum since it is a mixing of an intense band to a very weak band. This actually is the case for the CASPT2 emission spectra (shown in Figs. S6, S7) which also correspond to $n\pi^*$ minima. The spectrum obtained from the $\pi\pi^*$ minimum however is in much better agreement with experiment. Since B3LYP gives better excitation energies for the $\pi\pi^*$ states we have calculated that spectrum using the CAM-B3LYP minimum and normal modes but evaluating the excitation energies at each point using B3LYP. As is shown in Fig. 4b the spectrum

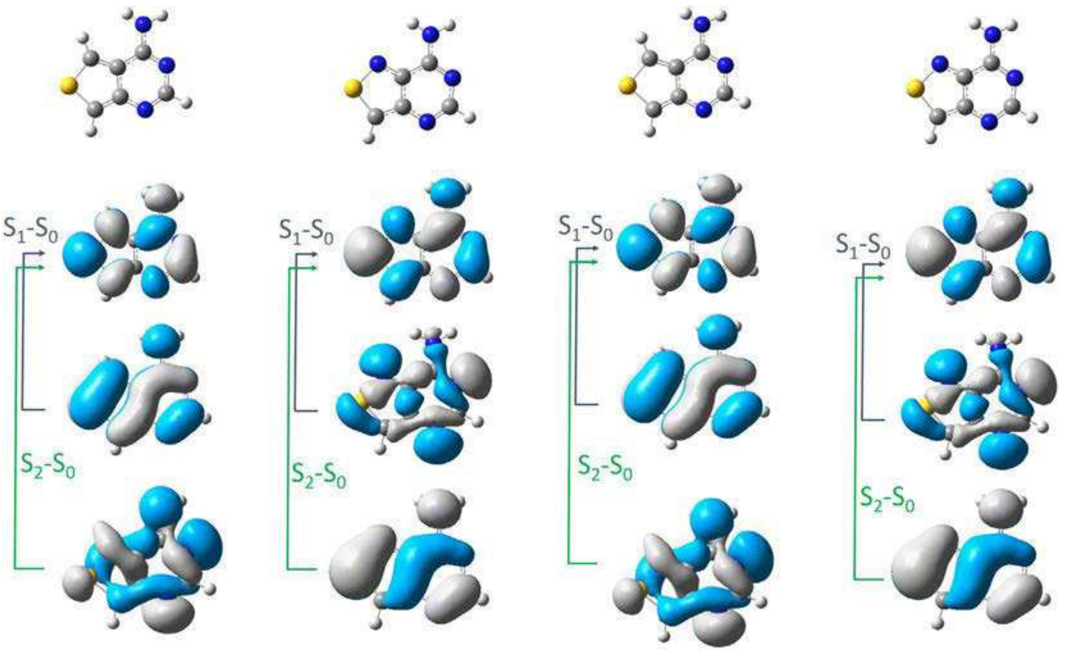


Fig. 3. Orbitals and orbital promotions associated with vertical absorption transitions for (from left to right) $TPA_{(g)}$, $iTPA_{(g)}$, $TPA_{(aq)}$ and $iTPA_{(aq)}$. All results taken from TD-B3LYP calculations.

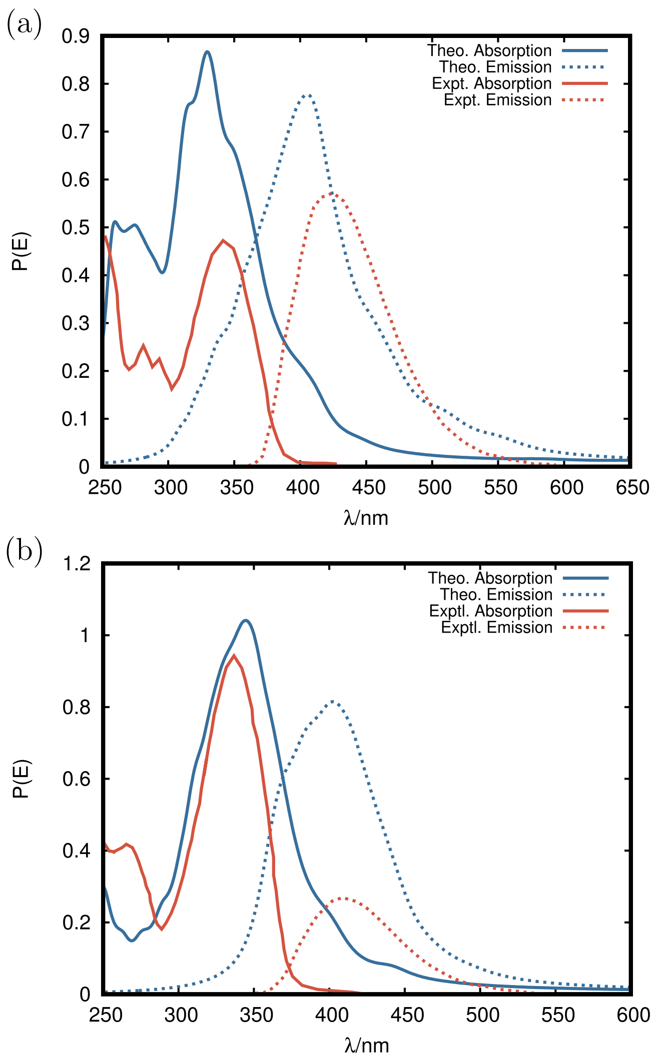


Fig. 4. Absorption and emissions spectra for (a) TPA and (b) iTPA. The absorption and emission spectra are shown in solid and dotted lines, respectively. These spectra were computed at the TD-B3LYP/6-31G(d) level of theory with solvation effects included using the PCM model (see methodology for details). Both absorption and emission profiles were based on the Wigner geometries. The emission spectrum for iTPA was computed using the CAM-B3LYP optimized minimum and normal modes (see main text). The experimental spectra are also shown in red lines taken from Ref. [16,17] The experimental spectra have different scaling for the intensity of absorption and emission and are taken directly from the figures in Ref. [16,17].

has a maximum very close to the experimental maximum and is somewhat broader than the experimental spectrum. It should be noted that our spectrum is based only on the radiative decay but the experimental spectrum is affected by all the nonradiative mechanisms as well. Furthermore, the Wigner distribution is based on a harmonic well of a single surface so any vibronic couplings with the other near by states is not included. This is especially true for iTPA since the S_2 state is near the S_1 at the S_1 minimum, so coupling is expected.

3.3. Mechanism of fluorescence

In order to better understand the fluorescence behavior of these analogues we need to examine radiationless relaxation mechanisms. In particular we need to understand the role of CIs, which have been found to be abundant in natural nucleobases, and how they differ in the current systems. We start by considering the well-known photophysics of 9H-adenine. In the isolated gas phase, the vertically excited S_1 and S_2 states of 9H-adenine are, respectively, of $n\pi^*$ and $\pi\pi^*$ character. The energy of the S_1 state is reduced by about 0.4 eV upon geometry relaxation [13]. The optimized S_1/S_0 CI is about 1 eV more stable than the Franck-Condon (FC) geometry and ca. 0.5 more stable that the S_1 optimized geometry. The path connecting the FC geometry, through the S_1 minimum, and en route to the S_1/S_0 CI is barrierless. The extremely small quantum yield for fluorescence in adenine is a direct manifestation of the barrierless transit from the FC region to the CI, motion through which promotes non-radiative $S_1 \leftarrow S_0$ internal conversion.

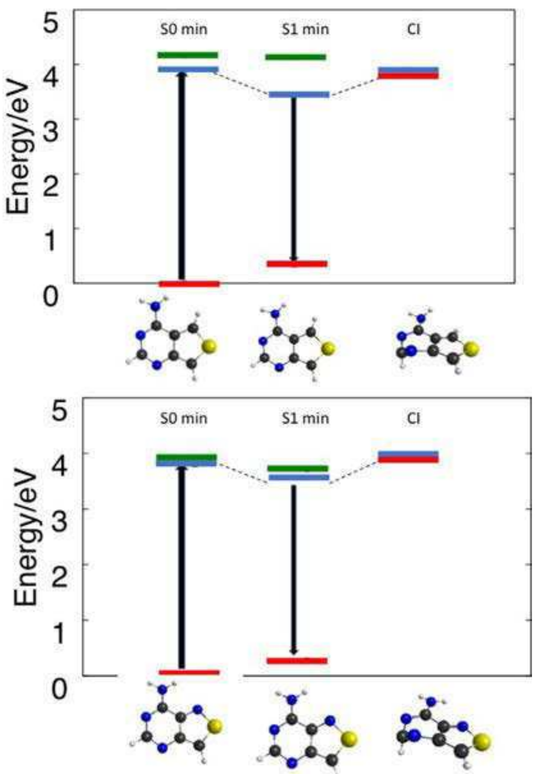


Fig. 5. Energies of important points at the TDDFT/PCM level for (top) TPA and (bottom) iTPA. A degenerate energy equal to the average of the S_1 and S_0 energies at the CASSCF level optimized geometry is used as the energy at the conical intersections.

Fig. 5 displays the energy level diagram associated with the geometric evolution of the FC geometry to that of the optimized S_1/S_0 CI geometry for (a) TPA and (b) iTPA using B3LYP/PCM. LIIC paths connecting the S_0 geometry to the CI are reported in SI. The optimized geometry of the S_1/S_0 CI is analogous to that in 9H-adenine. In both TPA and iTPA, the optimized geometry of the S_1 state is below that of the optimized S_1/S_0 CI, driving the thesis that the majority of the pre-excited population relaxes to the S_1 minimum. This process will be accelerated in bulk solution through vibrational energy transfer to the bulk solution. The remaining population at the S_1 minima is likely to relax to S_0 via fluorescence, accounting for the non-zero quantum yield for fluorescence measured by Tor and co-workers.

The experimental quantum yield for fluorescence in iTPA is smaller than that for TPA. According to our calculations this likely originates from the close proximity of the $\pi\pi^*$ and $n\pi^*$ states in iTPA. These states are almost degenerate upon absorption and this can lead to a bifurcation of population leading to a smaller population reaching the $\pi\pi^*$ minimum. It is possible that the $n\pi^*$ minimum traps some of the population leading to fluorescence quenching.

Similar energy level diagrams at the CASPT2 level are shown in SI. In that case the energy of the CI is not much higher than the S_1 minimum. Furthermore, the nature of the S_1 minimum in this case is an $n\pi^*$ transition. Both of these changes indicate that in the gas phase it is possible that the fluorescence of these systems will be further quenched.

4. General discussion and conclusions

In our present manuscript we have highlighted the most plausible reaction paths associated with the photophysics of two DNA-base analogues: TPA and iTPA. In both cases, their enhanced biocompatibility ensures optimal binding to bulk cellular DNA, while their geometric properties improve their radiative quantum yields when compared to the purine nucleobases from which they derive. Although the heavy S-atom is likely to promote spin-orbit coupling, we have not considered intersystem crossing, since we are at foremost only interested in the energetics of the singlet excited states from which fluoresence is observed.

In comparing TPA and iTPA with 9H-adenine (i.e. their closest molecular relative), the six-membered ring center remains unchanged in all molecules. As such the energetic position of the CI in TPA/iTPA is similar to that of 9H-adenine (ca. 4 eV) [13]. This is unsurprising, given the geometric similarity of the six-membered ring system. In contrast, the five-membered imidazole moiety in 9H-adenine is replaced by a thiophene and isothiazole moiety in TPA and iTPA, respectively. The inclusion of a 3rd-row heteroatom leads to a dramatic change in the excitation energy – specifically a bathochromic shift in the absorption is observed in TPA/iTPA when compared to that of 9H-adenine. This is a direct consequence of the S-heteroatom, in which the $3P_X$ electrons conjugate poorly with the proximal C-centered $2P_X$ electrons within the thiophene moiety. This weakened conjugation leads to a poorly conjugated π-system, thus leading to less stable π orbitals (cf. 9H-adenine). The destabilization of the π orbitals is directly linked to the stabilization of the π^* orbitals – reducing the energy gap between π and π^* orbitals. This ultimately leads to a bathochromic shift in the vertical excitation, placing the vertically excited states at the isoelectronic limit (at their highest). This effect is observed most clearly in iTPA in which the CI is above that of the vertically excited S_1 and S_2 states. In both TPA and iTPA the optimized S_1 energy is substantially lower than that of the CI, driving the thesis that the vertically excited population relaxes to the S_1 rather than relaxing to S_0 via $S_1 \leftarrow S_0$ internal conversion at the CI. In 9H-adenine, the vertically excited states and the S_1 minimum energy geometry is energetic higher than that of the CI, proving no sufficiently bound region to which the pre-excited population can relax on S_1. Instead, the pre-excited population undergoes rapid internal conversion back to S_1 via the CI - with almost unit quantum yield. Our results however indicate that the situation may be different in the gas phase.

The calculated absorption spectra for both systems are in very good agreement with the experimental spectra, providing confidence about our computational approach. The emission spectra are also in relatively good agreement, although it should be highlighted that the level of theory and the role of the environment were crucial in reproducing these spectra.

This study represents a detailed exploration of two recently proposed fluorescent molecular sensors that are biocompatible with DNA. Through a detailed understanding of the photophysics of such analogues, a geometric and electronic account can be devised and compared with the known photophysics of natural nucleobases. DNA/RNA nucleobases are remarkably photostable, in that they rapidly dissipate the excess energy provided by photoexcitation in the form of heat, reforming the parent ground state molecule. Since DNA is already a naturally biocompatible molecular system, understanding the geometric modifications that enhance the excited state lifetimes and thus improve the fluorescence quantum yield is vital for designing more effective fluorescent analogous for use in DNA/RNA biodiagnostics. It is therefore our future aim to continue in the design of modified DNA/RNA nucleobases, by differing the heteroatom positions and types, whilst retaining the biocompatibility with DNA. Computations of the properties of such modified nucleobase analogues are vital in guiding future experiments since the thermochemistry, dynamics and absorption/emission spectra can be modelled prior to the more expensive production chemistry step. We therefore expect many more such studies in the future.

Acknowledgements

The authors are grateful to the National Science Foundation (Grant No.: CHE-1465138 and CHE-1800171) for funding.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.chemphys.2018.08.049.

References

[1] A. Jablonski, Efficiency of anti-stokes fluorescence in dyes, Nature 131 (1933) 839–840.

[2] J.R. Lakowicz, Principles of Fluorescence Spectroscopy, Plenum Press, New York, 2006.

[3] S. Matsika, Modified Nucleobases, Vol. 355, Springer, Berlin-Heidelberg, 2015, pp. 209–243. https://doi.org/10.1007/12-82014-532.

[4] J. Liu, C. Liu, W. He, Fluorophores and their applications as molecular probes in living cells, Curr. Org. Chem. 17 (6) (2013) 564–579.

[5] E.A. Specht, E. Braselmann, A.E. Palmer, A critical and comparative review of fluorescent tools for live-cell imaging, Anal. Rev. Physiol. 79 (2017) 93–117.

[6] T. Terai, T. Nagano, Small-molecule fluorophores and fluorescent probes for bioimaging, Pflügers Arch. 465 (2013) 347–359.

[7] J. Li, D. Yim, W.-D. Jang, J. Yoon, Recent progress in the design and applications of fluorescence probes containing crown ethers, Chem. Soc. Rev. 46 (2017) 2437–2458, https://doi.org/10.1039/C6CS00619A.

[8] M.J. Rist, J.P. Marino, Fluorescent nucleotide base analogs as probes of nucleic acid structure, dynamics and interactions, Curr. Org. Chem. 6 (9) (2002) 775–793.

[9] J.N. Wilson, E.T. Kool, Fluorescent dna base replacements: reporters and sensors for biological systems, Org. Biomol. Chem. 4 (2006) 4265–4274.

[10] L.M. Wilhelmsson, Fluorescent nucleic acid base analogues, Q. Rev. Biophys. 43 (2) (2010) 159–183.

[11] R. Improta, F. Santoro, L. Blancafort, Quantum mechanical studies on the photophysics and the photochemistry of nucleic acids and nucleobases, Chem. Rev. 116 (2016) 3540–3593.

[12] C.E. Crespo-Hernandez, B. Cohen, P.M. Hare, B. Kohler, Ultrafast excited-state dynamics in nucleic acids, Chem. Rev. 104 (2004) 1977–2019.

[13] B. Marchetti, T.N.V. Karsili, M.N.R. Ashfold, W. Domcke, A ‘bottom up’, ab initio computational approach to understanding fundamental photophysical processes in nitrogen containing heterocycles, dna bases and base pairs, Phys. Chem. Chem. Phys. 18 (2016) 20007–20027.

[14] S. Ullrich, A.C. Borin, M. Barbatti (Eds.), Topics in Current Chemistry – Photoinduced Phenomena in Nucleic Acids I, vol. 355, Springer, Berlin-Heidelberg, 2015.

[15] S. Ullrich, A.C. Borin, M. Barbatti (Eds.), Topics in Current Chemistry – Photoinduced Phenomena in Nucleic Acids II, Vol. 356 Springer, Berlin-Heidelberg, 2015.

[16] D. Shin, R.W. Sinkeldam, Y. Tor, Emissive RNA Alphabet.

[17] A.R. Rovira, A. Fin, Y. Tor, Chemical mutagenesis of an emissive RNA alphabet, J. Am. Chem. Soc. 137 (46) (2015) 14602–14605.

[18] S. Arslancan, L. Martinez-Fernandez, I. Corral, Photophysics and photochemistry of canonical nucleobases’ thioanalogs: from quantum mechanical studies to time resolved experiments, Molecules 22 (2017) 998.

[19] M. Sholokh, R. Improta, M. Mori, R. Sharma, C. Kenfack, D. Shin, K. Voltz, R.H. Stote, O.A. Zaporozhets, M. Botta, Y. Tor, Y. Mély, Tautomers of a fluorescent g surrogate and their distinct photophysics provide additional information channels, Angew. Chem. Int. Ed. 55 (28) (2016) 7974–7978.

[20] M. Pollum, L. Martnez-Fernandez, C.E. Crespo-Hernandez, Modified Nucleobases, Vol. 355, Springer, Berlin - Heidelberg, 2015, p. 245. https://doi.org/10.1007/128-2014-532.

[21] A.D. Becke, Density-functional exchange-energy approximation with correct asymptotic behavior, Phys. Rev. A 38 (1988) 3098–3100.

[22] development of the Colle-Salvetti correlation-energy formula into a functional of the electron density.

[23] R. Ditchfield, W.J. Hehre, J.A. Pople, Self consistent molecular orbital methods. IX. An extended gaussian type basis for molecular orbital studies of organic molecules, J. Chem. Phys. 54 (2) (1971) 724–728.

[24] 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, J.E. Peralta, F. Ogliaro, M. Bearpark, J.J. Heyd, E. Brothers, K.N. Kudin, V.N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J.C. Burant, S.S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. 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, Farkas, J.B. Foresman, J.V. Ortiz, J.

Cioslowski, D.J. Fox, Gaussian 09, Revision B.01, 2009.

[25] J. Tomasi, B. Mennucci, R. Cammi, Quantum mechanical continuum solvation models, Chem. Rev. 105 (2005) 2999–3093.

[26] K. Andersson, P.A. Malmqvist, B.O. Roos, A.J. Sadlej, K. Wolinski, Second-order perturbation-theory with a CASSCF reference function, J. Phys. Chem. 94 (1990) 5483.

[27] T.H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, J. Chem. Phys. 90 (2) (1989) 1007–1023.

[28] B. Marchetti, T.N.V. Karsili, M.N.R. Ashfold, W. Domcke, A bottom up ab initio computational approach to understanding fundamental photophysical processes in nitrogen containing heterocycles, dna bases and base pairs, Phys. Chem. Chem. Phys. 18 (2016) 20007–20027.

[29] T.N.V. Karsili, B. Marchetti, R. Moca, M.N.R. Ashfold, Uv photodissociation of pyrroles: symmetry and substituent effects, J. Phys. Chem. A 117 (46) (2013) 12067–12074.

[30] T.N.V. Karsili, A.M. Wenge, S.J. Harris, D. Murdock, J.N. Harvey, R.N. Dixon, M.N.R. Ashfold, OH bond fission in 4-substituted phenols: S1 state predissociation viewed in a Hammett-like framework, Chem. Sci. 4 (2013) 2434–2446.

[31] B. Mario, S. Kakali, Effects of different initial condition samplings on photodynamics and spectrum of pyrrole, Int. J. Quant. Chem. 116 (10) 762–771.

[32] M. Barbatti, G. Granucci, M. Persico, M. Ruckenbauer, M. Vazdar, M. Eckert-Maksic, H. Lischka, The on-the-fly surface-hopping program system Newton-X: application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems, J. Photochem. Photobio. A 190 (2007) 288.

[33] H.-J. Werner, P.J. Knowles, G. Knizia, F.R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K.R. Shamasundar, T.B. Adler, R.D. Amos, A. Bernhardsson, A. Berning, D.L. Cooper, M.J.O. Deegan, A.J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A.W. Lloyd, R.A. Mata, A.J. May, S.J. McNicholas, W. Meyer, M.E. Mura, A. Nicklass, D.P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A.J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, Molpro, version 2015.1, a package of ab initio programs, see, 2015.

[34] Y. Shao, Z. Gan, E. Epifanovsky, A.T.B. Gilbert, M. Wormit, J. Kussmann, A.W. Lange, A. Behn, J. Deng, X. Feng, D. Ghosh, M. Goldey, P.R. Horn, L.D. Jacobson, I. Kaliman, R.Z. Khaliullin, T. Kus, A. Landau, J. Liu, E.I. Proynov, Y.M. Rhee, R.M. Richard, M.A. Rohrdanz, R.P. Steele, E.J. Sundstrom, H.L. Woodcock III, P.M. Zimmerman, D. Zuev, B. Albrecht, E. Alguire, B. Austin, G.J.O. Beran, Y.A. Bernard, E. Berquist, K. Brandhorst, K.B. Bravaya, S.T. Brown, D. Casanova, C.-M. Chang, Y. Chen, S.H. Chien, K.D. Closser, D.L. Crittenden, M. Diedenhofen, R.A. DiStasio Jr., H. Do, A.D. Dutoi, R.G. Edgar, S. Fatehi, L. Fusti-Molnar, A. Ghysels, A. Golubeva-Zadorozhnaya, J. Gomes, M.W.D. Hanson-Heine, P.H.P. Harbach, A.W. Hauser, E.G. Hohenstein, Z.C. Holden, T.-C. Jagau, H. Ji, B. Kaduk, K. Khistyaev, J. Kim, J. Kim, R.A. King, P. Klunzinger, D. Kosenkov, T. Kowalczyk, C.M. Krauter, K.U. Lao, A.D. Laurent, K.V. Lawler, S.V. Levchenko, C.Y. Lin, F. Liu, E. Livshits, R.C. Lochan, A. Luenser, P. Manohar, S.F. Manzer, S.-P. Mao, N. Mardirossian, A.V. Marenich, S.A. Maurer, N.J. Mayhall, E. Neuscamman, C.M. Oana, R. Olivares-Amaya, D.P. O'Neill, J.A. Parkhill, T.M. Perrine, R. Peverati, A. Prociuk, D.R. Rehn, E. Rosta, N.J. Russ, S.M. Sharada, S. Sharma, D.W. Small, A. Sodt, T. Stein, D. Stueck, Y.-C. Su, A.J.W. Thom, T. Tsuchimochi, V. Vanovschi, L. Vogt, O. Vydrov, T. Wang, M.A. Watson, J. Wenzel, A. White, C.F. Williams, J. Yang, S. Yeganeh, S.R. Yost, Z.-Q. You, I.Y. Zhang, X. Zhang, Y. Zhao, B.R. Brooks, G.K.L. Chan, D.M. Chipman, C.J. Cramer, W.A. Goddard III, M.S. Gordon, W.J. Hehre, A. Klamt, H.F. Schaefer III, M.W. Schmidt, C.D. Sherrill, D.G. Truhlar, A. Warshel, X. Xu, A. Aspuru-Guzik, R. Baer, A.T. Bell, N.A. Besley, J.-D. Chai, A. Dreuw, B.D. Dunietz, T.R. Furlani, S.R. Gwaltney, C.-P. Hsu, Y. Jung, J. Kong, D.S. Lambrecht, W. Liang, C. Ochsenfeld, V.A. Rassolov, L.V. Slipchenko, J.E. Subotnik, T. Van Voorhis, J.M. Herbert, A.I. Krylov, P.M.W. Gill, M. Head-Gordon, Advances in molecular quantum chemistry contained in the Q-Chem 4 program package, Mol. Phys. 113 (2) (2015) 184–215.