THE JOURNAL OF PHYSICAL CHEMISTRY A

pubs.acs.org/JPCA Article

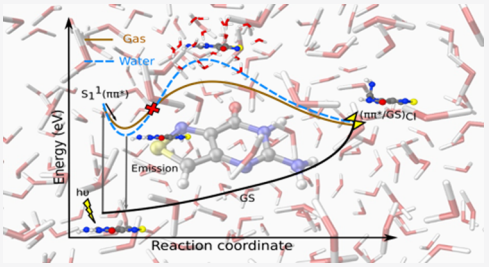
Solvation Structures and Deactivation Pathways of Luminescent Isothiazole-Derived Nucleobases: ^{tz}A, ^{tz}G, and ^{tz}I

Adalberto Vasconcelos Sanches de Araújo,[§] Danillo Valverde,[§] Sylvio Canuto, and Antonio Carlos Borin*

Cite This: J. Phys. Chem. A 2020, 124, 6834–6844 Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The photophysical relaxation pathways of ^{tz}A, ^{tz}G, and ^{tz}I luminescent nucleobases were investigated with the MS-CASPT2 quantum-chemical method and double-ζ basis sets (cc-pVDZ) in gas and condensed phases (1,4-dioxane and water) with the sequential Monte Carlo/CASPT2 and free energy gradient (FEG) methods. Solvation shell structures, in the ground and excited states, were examined with the pairwise radial distribution function ($G(r)$) and solute−solvent hydrogen-bond networks. Site-specific hydrogen bonding analysis evidenced relevant changes between both electronic states. The three luminescent nucleobases share a common photophysical pattern, summarized as the lowest-lying $^1(\pi\pi^*)$ bright state that is populated directly after the absorption of radiation and evolves barrierless to the minimum energy structure, from where the excess of energy is released by fluorescence. From the $^1(\pi\pi^*)_{min}$ region, the conical intersection with the ground state ($(\pi\pi^*/GS)_{CI}$) is not accessible due to the presence of high energetic barriers. By combining the present results with those reported earlier by us for the pyrimidine fluorescent nucleobases, we present a comprehensive description of the photophysical properties of this important class of new fluorescent nucleosides.



INTRODUCTION

Living organisms have been challenged by the harmful effects of ultraviolet (UV) radiation that affect the adjacent nucleobase bonding in DNA, which means that protein-building enzymes will not be able to read the DNA code at that point. This is a serious issue because DNA nucleobases have a large UV absorption cross-section. Fortunately, our genetic information is mostly preserved by sophisticated mechanisms such as, for instance, DNA intrinsic protection by dissipating the absorbed UV energy as heat, avoiding breaking the chemical bonds responsible for keeping the DNA strands together.

The self-protecting DNA photophysical property is intrinsically related to the canonical nucleobase capacity to release the excess of energy via remarkably fast and efficient radiationless mechanisms, which have been widely investigated.[1−8] The consensus is that conical intersections are the key for their understanding,[4,5,9] resulting in very stable photochemical canonical nucleobases with extremely low fluorescence quantum yields ($\Phi_F < 10^{-4}$). The distinguished short-lived and nonemissive canonical nucleobase properties are of paramount importance for life on Earth.

Conversely, experimental DNA structural investigations have faced considerable hurdles. The fluorescence spectroscopic technique,[10,11] a powerful experimental method to explore structural properties, cannot be applied efficiently because the canonical nucleobases are nonfluorescent. To overcome this shortcoming, scientists have focused on synthesizing surrogate fluorescent nucleobases to be used as molecular probes when incorporated into the DNA/RNA strands.[12−17]

Different compounds have been used as fluorescent nucleobase analogues over the years, such as 2-aminopurine[18−20] and compounds obtained by the addition of groups at a specific position as 8-vinylpurines.[21,22] More recently, a promising approach has used isomorphic purine/pyrimidine moieties such as 4-cyanoindole as the starting materials, which has been proposed as a universal probe for DNA studies[23] lately. Other examples, related to different classes of compounds, can be found in the literature.[24−28]

Ideally, fluorescent nucleoside analogues used as probes should minimally disturb the structure and function of nucleic acids composed of the canonical nucleobases. In the quest for synthesizing isomorphic and isofunctional fluorescent nucleoside analogues, Tor and co-workers synthesized a new RNA alphabet using isothiazole-[4,3-*d*]pyrimidine as a common core,[29] resulting in the following compounds: ^{tz}A (isothiazolo-adenosine analogue), ^{tz}G (isothiazolo-guanosine analogue), ^{tz}C (isothiazolo-cytidine analogue), ^{tz}U (isothiazolo-uridine analogue), and ^{tz}I (isothiazolo-inosine analogue); the latter has

Received: April 16, 2020
Revised: July 30, 2020
Published: August 3, 2020

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

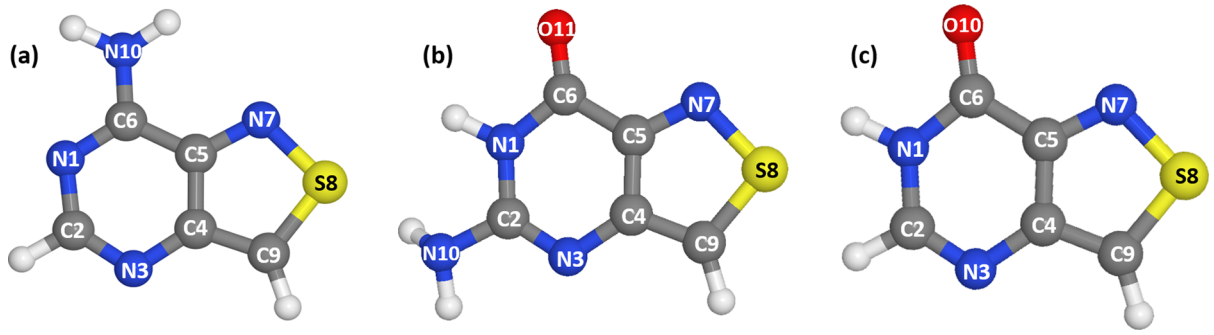


Figure 1. Schematic structures of the fluorescent purine analogues (a) ^{tz}**A**, (b) ^{tz}**G**, and (c) ^{tz}**I**, with the numbering scheme.

been derived from ^{tz}**A** by enzymatic oxidation in an effort to prove the biological isofunctionality of adenine. Afterward, the same authors[30] expanded the class of emissive and isofunctional RNA nucleoside analogues to their noncanonical compounds. Although their photophysical properties have been determined experimentally by Tor et al., the fundamental aspects of the nature of their electronic states and fluorescence processes have been barely investigated from the theoretical point of view.

Some aspects about the electronic structure and hydrogen bond (HB) pairing have been studied[31−34] in the scope of the density functional theory (DFT), including solvent effects with the polarizable continuum model (PCM).[35] We carried out a detailed study on the fluorescent nucleoside analogues, scrutinizing the nature of specific solute−solvent interactions and the fundamental aspects behind the absorption spectra;[36] the photophysical relaxation mechanisms of the ^{tz}**C** and ^{tz}**U** pyrimidine analogues were also investigated by us recently.[37]

Following our previous successful approach, it is now timely and relevant to examine the photophysical properties of the isolated purine analogues ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** (Figure 1), in gas and condensed phases, describing details about the molecular nature of their excited states, emission spectra, Stokes shifts, and photophysical deactivation pathways, to unravel the electronic and geometric features behind their enhanced fluorescence quantum yields. The physical−chemical characterization of isolated systems in gas and condensed phases is the startup process to investigate their behavior in conditions that mimic real environments, as for example, in a DNA-like structure.

In this contribution, the excited state potential energy hypersurfaces were characterized by means of state minima, conical intersections, minimum energy paths (MEPs), and linear interpolations in internal coordinate calculations using, whenever possible, the state-of-the-art complete active space second-order perturbation theory (CASPT2) method, taking into account solvent effects via the sequential quantum mechanics/molecular mechanics (QM/MM) protocol using the Monte Carlo and CASPT2 methods.[38,39] Our results thus contribute to a better understanding of the photophysical properties of this important class of fluorescent nucleoside analogues.

METHODOLOGY

The ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** photophysics and relaxation mechanisms were investigated in the gas phase, 1,4-dioxane, and water, following as close as possible the methodology used previously by us.[36,37] In this section, we summarize the main aspects of our computational procedure; further details can be found in our aforementioned papers.

Electronic Structure Calculations. Geometries were optimized in the gas phase, 1,4-dioxane, and water using the multistate complete active space second-order multiconfigurational perturbation theory (MS-CASPT2),[40] with a zeroth-order state-average complete active space self-consistent field (SA-CASSCF)[41] wave function, obtained by averaging over the two lowest-lying electronic states and the cc-pVDZ double-ζ atomic basis sets.[42] ^{tz}**A** and ^{tz}**I** geometry optimizations were carried out with active spaces built by all π, π^* orbitals and electrons (CAS(12,10)); ^{tz}**G** geometry optimizations were performed with an active space encompassing eleven π, π^* orbitals and fourteen electrons (CAS(14,11)). Minimum energy crossing points, between the ground (GS) and lowest $^1(\pi\pi^*)$ states ($(\pi\pi^*/GS)_{CI}$), were optimized by means of the restricted Lagrange multiplier technique,[43] using the same level of theory and active space of the aforementioned optimization procedures.

Excitation energies in all environments were computed with the cc-pVDZ double-ζ atomic basis sets and the single-state CASPT2 (SS-CASPT2)[44−46] method, using an active space comprising 10 π-type orbitals, three nonbonding (n) orbitals associated with nitrogen (n_N) and/or oxygen (n_O) atoms, depending on the structure of the molecules (CAS(18,13)). It is worth mentioning that for all optimized MS-CASPT2 minimum energy crossing points, the off-diagonal elements are small (at about 0.001 au) and the degeneracy computed at the SS-CASPT2 level is about the same as that computed at the MS-CASPT2 level, with differences smaller than at about 0.01 eV. Further details about the solvation techniques can be found in the next section.

Gas-phase minimum energy paths (MEPs) from the Franck−Condon (FC) region, along the corresponding lowest-lying $^1(\pi\pi^*)$ state potential energy hypersurface, were computed at the SA(2)-CASSCF/cc-pVDZ level with the full-π active space, following the steepest descent path technique.[43] At each point of the MEP, the dynamical correlation was included recomputing the energies at the SS(6)-CASPT2-(18,13)/cc-pVDZ level, following the so-called CASPT2//CASSCF protocol.[5] Whenever MEP calculations were not feasible, energetic profiles connecting minimum energy crossing points and optimized geometries were investigated with the linear interpolation using the internal coordinate technique (LIIC), computing the vertical energies at each point as described above.

In all of the cases, CASPT2 calculations were carried out using the imaginary shift technique[47] (0.2 Hartree); no IPEA shift (ionization potential-electron affinity)[48] correction was

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

employed, as suggested recently.[49] Electronic structure calculations were carried out using MOLCAS 8 software.[50]

Solvation Effects. Solvent effects were taken into account with the ASEC−FEG method,[51] combining the average solvent electrostatic configuration (ASEC) and the free energy gradient (FEG) methods. ASEC is a single configuration that represents the statistical average and leads to a nonhomogeneous solvent distribution that is particularly important for protic solvents.[52]

In the ASEC−FEG methodology, a classical molecular mechanics (MM) simulation is performed to sample the structure of the liquid under the given thermodynamic conditions. Based on statistical criteria, a set of hundreds of statistically uncorrelated configurations of the solvent is chosen to be used in the QM calculation, on which they are grouped in a single configuration assuming an average solvent electrostatic configuration embedding for the solvent, the ASEC. Solute geometry optimizations, in the presence of the solvent, can be performed by incorporating the free energy gradient (FEG)[53,54] method, which allows one to investigate the free energy hypersurface and, consequently, the minima and saddle points. For further technical details, we refer the reader to the original papers and to our most recent publications in which the ASEC−FEG method was used to describe the solvent effects on the electronic structure and photophysics of related fluorescent nucleobases.[36,37] Classical Monte Carlo simulations were carried out with the DICE package.[55]

RESULTS AND DISCUSSION

Equilibrium Geometries. As observed experimentally by Rovira et al.[29] and by us at the CASPT2/cc-pVDZ level,[36] the lowest bright excited state of ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** can be populated after irradiation with UV light, in both gas and condensed phases (experimental results show the peaks of absorption at 3.63, 3.66, and 3.94 eV in 1,4-dioxane, respectively[29]). According to our previous study,[36] the main bright state in the Franck−Condon (FC) region is derived from the ground state by a single electron excitation from the highest occupied molecular orbital (HOMO) to the lowest unoccupied molecular orbital (LUMO), describing a state characterized by a $\pi \rightarrow \pi^*$ configuration. ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** HOMO and LUMO exhibit similar contributions from the π orbitals localized on the purine skeleton (Figure 2).

In this contribution, ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** bright $^1(\pi\pi^*)$ state geometries were optimized in the gas phase, 1,4-dioxane, and water using the ASEC−FEG approach. In the gas phase, ^{tz}**A** and ^{tz}**I** $^1(\pi\pi^*)$ minima show completely planar structures. As to the ^{tz}**G** $^1(\pi\pi^*)$ excited state minimum structure, it was found to have a planar skeleton, with the amino group (NH_2) hydrogen atoms out of the ring plane by 25°. When compared to the gas-phase ground state geometries (obtained at the MP2/cc-pVDZ level[36]) the main differences of the excited state gas-phase minima structures are the bond distances and the pyramid angle of the amino group (α angle), as depicted in the superposition of both gas-phase structures (Figure 3).

Solvent effects (1,4-dioxane and water) on selected bond lengths, in relation to the gas-phase values, of the $^1(\pi\pi^*)$ minima are displayed in Figure 4. As it can be noticed (positive values mean larger bond lengths than the gas phase), both solvents influence the bond lengths likewise for most of the selected parameters, except for the C_4-C_5 and C_6-C_{10} bond lengths in ^{tz}**G** and ^{tz}**I**, respectively. In these cases, the effect of

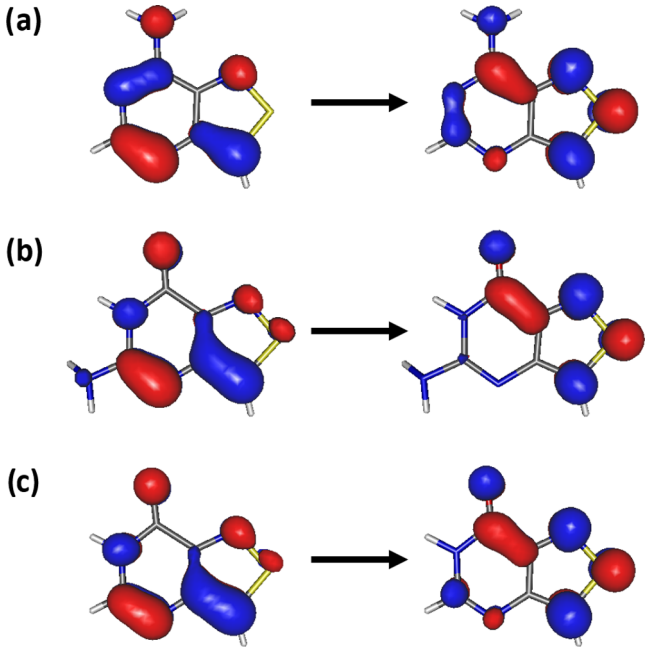


Figure 2. Dominant electronic transition for describing the lowest $^1(\pi\pi^*)$ excited state of (a) ^{tz}**A**, (b) ^{tz}**G**, and (c) ^{tz}**I**. The orbitals were obtained from the averaged valence natural orbitals, computed at the SA(6)-CASSCF(18,13)/cc-pVDZ level in the gas phase.

1,4-dioxane is negligible. It is also interesting to observe that the largest variation observed in all compounds (^{tz}**A**, ^{tz}**G**, and ^{tz}**I**) is for the N_1-C_2 bond length, which is larger at the excited state in both solvents. Nonetheless, in ^{tz}**I**, it is noticed in water, while for both ^{tz}**A** and ^{tz}**G**, it is observed in 1,4-dioxane. For ^{tz}**I** and ^{tz}**G** in 1,4-dioxane and water, the N_7-S_8 bond length exhibits large contraction, while in ^{tz}**A**, it is elongated in both solvents. The S_8-C_9 bond length, the other bond related to the sulfur atom, changes very little in both solvents in ^{tz}**G**. However, in ^{tz}**A** and ^{tz}**I**, it is shorter in 1,4-dioxane and water, mainly in ^{tz}**I**.

Solvent Structure and Hydrogen-Bond Networks. Solvation shell structures around the ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** $^1(\pi\pi^*)_{min}$, in 1,4-dioxane and water, were analyzed with the aid of the minimum-distance pairwise distribution function $G_{MD}(r)$[56] (Figure 5). For completeness, these distribution functions were also compared to those reported by us for the ground state previously.[36]

For analysis of the number of hydrogen bonds (HB, X···H−Y), we used the procedure described by us previously.[36] In short, the amount of hydrogen bond was derived from an ensemble of snapshots extracted from the Monte Carlo sampling, with the aid of the following variables[57−59] (thresholds in parentheses): (i) R_{XY} ($R_{XY} \leq 3.75$Å), defined by the first valley of the pairwise radial distribution function ($G_{MD}(r)$) between X and Y atoms; (ii) the angle θ_{XY} ($\theta_{XY} \leq 40°$), defined by the X, H, and Y atoms; and (iii) the binding energy, E_{HB} ($E_{HB} \leq -0.01$ kcal/mol). It cannot be assured that all nearby solvent molecules are indeed making the solute−solvent hydrogen bond. For this reason, the energetic criterion is needed.[38] It is also worth mentioning that only the set of atomic charges in the solute was refined in the optimization procedure, with the same Lennard-Jones parameters for the ground and excited states; hence, the differences reported on the radial distribution functions may be attributed to the

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

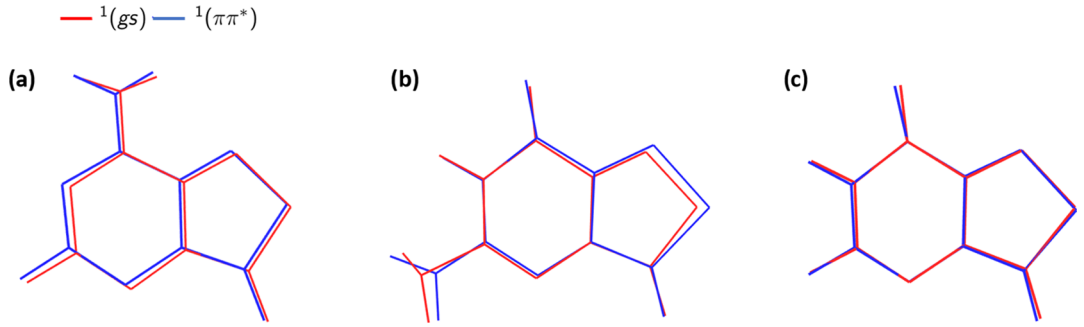


Figure 3. Superposition of the optimized ground (red) and $^1(\pi\pi^*)$ (blue) state structures of the (a) ^{tz}**A**, (b) ^{tz}**G**, and (c) ^{tz}**I** fluorescent nucleobases optimized in the gas phase. The geometries in the ground state were optimized at the MP2/cc-pVDZ level, while the MS(2)-CASPT2/cc-pVDZ level was adopted for $^1(\pi\pi^*)$ excited state.

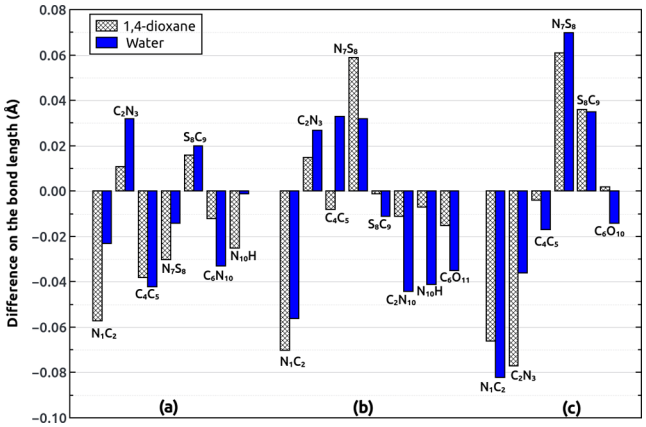


Figure 4. Solvation effects, in relation to the gas phase, on selected bond lengths for the $^1(\pi\pi^*)_{min}$-optimized structures of the (a) ^{tz}**A**, (b) ^{tz}**G**, and (c) ^{tz}**I** fluorescent nucleobases optimized in 1,4-dioxane and water at the MS(2)-CASPT2/cc-pVDZ level. The full-π space was adopted as an active space. The positive value means a larger bond length in the gas phase.

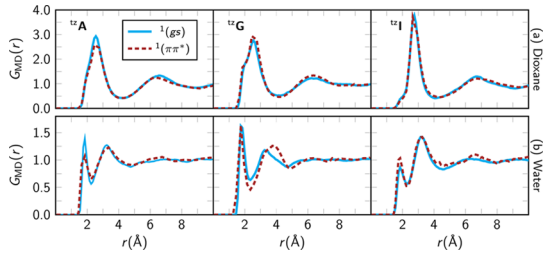


Figure 5. Minimum-distance pairwise distribution function $G_{MD}(r)$ for ground and $^1(\pi\pi^*)$ states of emissive RNA molecules ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** in (a) 1,4-dioxane and (b) water.

redistribution of atomic charges of the solute after absorption of UV light.

The minimum-distance pairwise distribution functions for the excited states in 1,4-dioxane do not exhibit significant changes in comparison with those for the ground state.[36] The 1,4-dioxane $G_{MD}(r)$ around the three $^1(\pi\pi^*)_{min}$ structures is very similar (Figure 5), characterized by two well-defined solvation shells. The shoulder on the left side of the first peak at about 2 Å may be interpreted as the solvent molecules that make hydrogen bonds (HB) with the solute. Nonetheless, not all molecules in this region are taking part in hydrogen bonds. For a realistic description of the effective number of solute−solvent hydrogen bonds, further arguments on angular and energy criteria were taken into account, as discussed above.

The first solvation shell ranges from 1.40 to 4.20 Å for the three modified purine analogues, enfolding 15 molecules of 1,4-dioxane for ^{tz}**A** and ^{tz}**I**, and 17 for ^{tz}**G**. We notice a greater number of solvent moieties in the first solvation shell for ^{tz}**I** and ^{tz}**G** excited states in comparison to those obtained by us[36] for the ground states (14 and 15 1,4-dioxane molecules, respectively); no change is observed in relation to the ground state for ^{tz}**A**.

The aqueous solvation shells for the excited states differ more from those observed for the ground state. The sharp peak around 2 Å can be associated with solute−solvent hydrogen bonds, better defined than that observed in 1,4-dioxane. The first solvation shell starts at about 2.5 Å, with a peak of around 3.5 Å. The most notable difference is the decrease in the intensity of the first peak observed for the ^{tz}**A** excited state, suggesting a smaller probability for encountering water molecules in this range. For ^{tz}**G**, a broader first hydration shell is observed, which leads to the maximum number of water molecules on the shell when compared to the other molecules.

Based on the above-mentioned criteria, the excited state hydrogen-bonding shell includes 6 water molecules for ^{tz}**A**, 9 for ^{tz}**G**, and 5 for ^{tz}**I**. As to the first complete solvation shell, it contains 49 water molecules for ^{tz}**A**, 57 for ^{tz}**G**, and 43 for ^{tz}**I**. For the sake of comparison, in the ground state, the first complete solvation shell includes 49 for ^{tz}**A**, 49 for ^{tz}**G**, and 44 for ^{tz}**I**, which indicates that the ^{tz}**G** excited state is more susceptible to solvent effects than the others.

Table 1 shows the number of HBs in the ground[36] and excited states. As in the ground state, the molecule that presents the least number of hydrogen bonds is ^{tz}**I**, with an average of 0.20 in 1,4-dioxane and 4.48 in water, suggesting more hydrogen bonds in the excited state in both solvents. These changes are similar for ^{tz}**A**, for which the number of HBs increases from 1.63 in the ground state to 1.81 in the excited state in 1,4-dioxane, and from 6.14 to 6.66 in water. The most significant variation occurs for ^{tz}**G**, for which the hydrogen bonds in 1,4-dioxane decrease from 2.45 (ground state) to 1.88 (excited state); on the other hand, in water, we observe more HBs in the excited state (9.12) than in the ground state (8.26).

In the excited state, both ^{tz}**A** and ^{tz}**G** make preferably 2 hydrogen bonds (79.4 and 46.5% of the configurations, respectively) in 1,4-dioxane. For ^{tz}**I**, most of the configurations reveal no HB (80.1%). In water, the solute molecules show more configurations that make 7 (36.8%, ^{tz}**A**), 9 (47.2%, ^{tz}**G**),

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

Table 1. Percentage Statistics of Solute−Solvent Hydrogen Bonds for Modified Purines in 1,4-Dioxane and Water for the Ground State (GS) and the Excited State ($\pi\pi^*$)[a]

		^{tz}A		^{tz}G		^{tz}I	
medium	no. of HBs	GS	$\pi\pi^*$	GS	$\pi\pi^*$	GS	$\pi\pi^*$
dioxane	0	3.5	1.0			97.1	80.1
	1	31.4	18.7	9.0	33.8	2.9	19.4
	2	64.1	79.4	37.5	46.5		0.5
	3	1.0	1.0	52.8	17.7		
	4			0.7	1.9		
	⟨HB⟩	1.63	1.81	2.45	1.88	0.03	0.20
water	2					2.3	1.7
	3	0.2	0.3			15.6	16.5
	4	3.8	1.9			40.4	37.3
	5	21.8	11.6	0.4		29.9	31.0
	6	39.5	28.5	4.0	0.1	9.8	11.2
	7	26.1	36.8	18.0	2.7	1.8	1.9
	8	7.5	17.5	36.0	19.5	0.2	0.2
	9	0.9	3.2	29.5	47.2		
	10	0.1	0.2	10.2	25.1		
	11			1.7	5.0		
	12			0.2	0.5		
	⟨HB⟩	6.14	6.66	8.26	9.12	4.35	4.48

[a]⟨HB⟩ is the average number of hydrogen bonds. The ground state is taken from our previous work.[36]

and 4 (37.3%, ^{tz}I) hydrogen bonds. The average number of hydrogen bonds in both solvents follows the following trend: ^{tz}I → ^{tz}A → ^{tz}G, in agreement with the tendency observed for the ground state solvent structures.

Although changes in hydrogen bonds are not significant when a global molecular analysis is carried out, a closer look at specific sites reveals interesting details. The site-specific pairwise distribution function $g(r)$ for some relevant sites on the solute molecules is displayed in Figure 6, and the corresponding site-specific hydrogen bond analysis is displayed in Table 2.

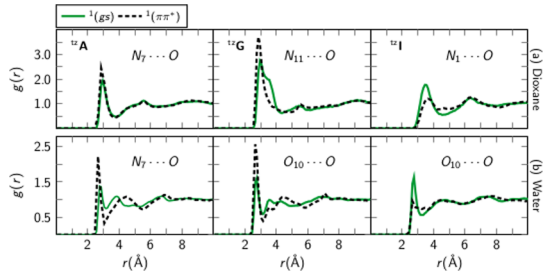


Figure 6. Site-specific pairwise distribution function $g(r)$ for the most electronegative atoms on the ground and $^1(\pi\pi^*)$ states of the emissive RNA molecules ^{tz}A, ^{tz}G, and ^{tz}I in (a) 1,4-dioxane and (b) water.

The analysis of the $g(r)$ (Figure 6) for ^{tz}A in 1,4-dioxane indicates that around the N_7 atom, the solvation structure is very similar for both states. On the other hand, the intensity of the first peak at the N_1 position diminishes for ^{tz}I in 1,4-dioxane.

Turning our attention to the ^{tz}G $N_{11}\cdots O$ $g(r)$ in 1,4-dioxane, a more pronounced difference is noted between the solvation shells for the ground and excited states. In the excited state, the shoulder around 3 Å disappears and a more intense peak is observed at about 2 Å. The ⟨HB⟩ also exhibits differences worth mentioning. While in the ground state the number of HB for the two hydrogen atoms at $N_{11}-H\cdots O$ are, respectively, 0.67 and 0.81, in the excited state they are decreased to 0.24 and 0.30, respectively.

Table 2. Site-Specific Hydrogen Bond Average (⟨HB⟩) for the Ground State (GS) and Excited State ($\pi\pi^*$) of the Modified RNA Purines in 1,4-Dioxane and Water

		1,4-dioxane		water	
	⟨HB⟩	GS	$\pi\pi^*$	GS	$\pi\pi^*$
^{tz}A	$N_{10}-H_{11}\cdots O$	0.71	0.89	0.86	0.97
	$N_{10}-H_{12}\cdots O$	0.91	0.92	1.00	1.00
	$N_{10}\cdots H-O$			0.10	0.01
	$N_1\cdots H-O$			1.48	0.74
	$N_3\cdots H-O$			1.77	1.74
	$N_7\cdots H-O$			0.56	1.94
	$S_8\cdots H-O$			0.02	0.26
^{tz}G	$N_{11}-H_{13}\cdots O$	0.67	0.24	0.97	1.00
	$N_{11}-H_{14}\cdots O$	0.81	0.33	0.95	1.00
	$N_1-H\cdots O$	0.97	0.30	1.01	0.99
	$N_{11}\cdots H-O$			0.12	0.00
	$N_1\cdots H-O$			0.01	0.01
	$N_3\cdots H-O$			1.72	0.19
	$N_7\cdots H-O$			0.80	1.02
	$S_8\cdots H-O$			0.33	0.85
	$O_{10}\cdots H-O$			2.27	3.05
^{tz}I	$N_1-H\cdots O$	0.03	0.20	0.12	0.13
	$N_1\cdots H-O$			0.00	0.00
	$N_3\cdots H-O$			1.15	0.53
	$N_7\cdots H-O$			0.65	0.79
	$S_8\cdots H-O$			0.17	0.47
	$O_{10}\cdots H-O$			2.10	2.50

The changes in aqueous medium are even more remarkable (Figure 6). For ^{tz}A and ^{tz}G, the peak around 2 Å for the $g(r)$ on the N_7 and O_{10} (respectively) sites increases. According to our analysis, in relation to the ground state, the number of HBs observed on the excited state of ^{tz}A increases from 0.56 to 1.94 on the N_7 site; for ^{tz}G, it goes from 2.27 to 3.05 on the O_{10} site. An opposite trend is found for the O_{10} site on the ^{tz}I excited state, which goes from 2.50 (ground state) to 2.27 (excited state). For larger distances, the $g(r)$ exhibits similar behaviors for both states.

Other findings are still worth mentioning. The number of HBs at the N_3 position for both ^{tz}G and ^{tz}I reduces by, on average, 1.53 and 0.62, respectively. The N_1 position of ^{tz}A also shows the same behavior, with a drop of 0.74 HBs. An interesting trend for the three molecules is the increase of the ⟨HB⟩ at the S_8 position, ranging in the interval between 0.24 and 0.52 in the excited state. In the ground state, these molecules have a relatively small number of hydrogen bonds at this position.

Emission Spectra and Stokes Shift. Vertical emission energies of ^{tz}A, ^{tz}G, and ^{tz}I computed at the SS(6)-CASPT2(18,13)/cc-pVDZ level in the gas phase, 1,4-dioxane, and water are displayed in Table 3.

The lowest-lying $^1(\pi\pi^*)$ excited states of the luminescent purine analogues are mainly described by a single excited configuration, derived from the ground state by a HOMO → LUMO electronic transition, associated with a high oscillator strength in the gas and condensed phases. It is also interesting to mention that $^1(\pi\pi^*)$ corresponds to the S_1 state for ^{tz}G and

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

Table 3. ^{tz}A, ^{tz}G, and ^{tz}I Vertical Emission Energies (in eV) Computed at the SS(6)-CASPT2(18,13)/cc-pVDZ Level (from the MS-CASPT2/cc-pVDZ $^1(\pi\pi^*)_{min}$-Optimized Geometries (Vide Supra)) and the Corresponding Experimental (Gas Phase, 1,4-Dioxane, and Water) Values[29]

		1,4-dioxane		water	
molecule	gas	ASEC−FEG	exp.	ASEC−FEG	exp.
^{tz}A	3.23	3.35	3.03	3.04	3.02
^{tz}G	3.00	3.11	2.92	2.57	2.70
^{tz}I	3.63	3.50	3.33	3.48	3.29

^{tz}I, while for ^{tz}A, the S_1 excited state is of $^1(n\pi^*)$ nature in the gas phase, with $^1(\pi\pi^*)$ as the S_2 (more details in the next section). Nonetheless, in all cases, the emitting state was considered to be the $^1(\pi\pi^*)$ state, as described by us.[36]

The computed gas-phase emission energies (Table 3) can be compared with those obtained experimentally in 1,4-dioxane, a solvent with small dielectric constant. The gas-phase theoretical results differ from the experimental results obtained in 1,4-dioxane by 0.20 eV for ^{tz}A, 0.08 eV for ^{tz}G, and 0.30 eV for ^{tz}I, with the gas-phase emission energies being blue-shifted in comparison to the experimental values. By comparing the ^{tz}A (3.35 eV), ^{tz}G (3.11 eV), and ^{tz}I (3.50 eV) emission energies in 1,4-dioxane, computed with the ASEC−FEG method, with the reported experimental values, we notice that they are all blue-shifted (^{tz}A: 0.32 eV; ^{tz}G: 0.19 eV; ^{tz}I: 0.17 eV). Nevertheless, the theoretical emission energies are within the accuracy of the CASPT2 method, leading to good agreement with the experimental data.

The emission energies computed in water are also displayed in Table 3. The ASEC−FEG method presents a consistent red-shift from the gas phase to water for all molecules (^{tz}A: 0.19 eV; ^{tz}G: 0.43 eV; ^{tz}I: 0.15 eV). The same trend is observed experimentally; that is, experimental emission energies in the aqueous solution are red-shifted in relation to the computed gas-phase values (^{tz}A: 0.21 eV; ^{tz}G: 0.30 eV; ^{tz}I: 0.34 eV).

Table 4 contains Stokes shifts computed in this work in the gas phase and both solvents with the ASEC−FEG method. Our computed Stokes shifts in the gas phase for ^{tz}A and ^{tz}G, in comparison with the experimental values obtained in 1,4-dioxane, are overestimated by 0.17 and 0.13 eV, respectively, and underestimated by 0.21 eV for ^{tz}I. When using the ASEC−FEG for 1,4-dioxane, the differences decrease to 0.01 eV for ^{tz}A, 0.09 eV for ^{tz}G, and 0.12 eV for ^{tz}I, closer to the experimental value. In water, computed values obtained with the ASEC−FEG method agree with the reported experimental values for ^{tz}A and ^{tz}G, for which our values are overestimated by 0.08 and 0.04 eV, respectively; for ^{tz}I, the computed Stokes shift is underestimated by 0.17 eV. The Stokes shift computed by us for ^{tz}A in water (0.73 eV) compares with that of the CAM-B3LYP/PCM/6-31G(d) level of theory (0.52 eV) reported by Karsili et al.[60]

Table 4 also exhibits solvatochromic shifts from 1,4-dioxane to water (negative value indicates that the Stokes shift is larger in water). Experimental data reveal a moderate change for ^{tz}A (−0.05 eV) and ^{tz}I (−0.07 eV), and a larger value for ^{tz}G (−0.28 eV). We can observe that the ASEC−FEG method predicts good solvatochromic shift for all of the molecules, with the highest discrepancy of 0.13 eV, reported for ^{tz}G.

Photophysical Deactivation Pathways. The photophysical processes of ^{tz}A, ^{tz}G, and ^{tz}I were investigated by exploring the topology of their $^1(\pi\pi^*)$ bright state potential energy surfaces, using the minimum energy path (MEP) technique and the CASPT2//CASSCF protocol. From the Franck−Condon region, MEP calculations were performed at the CASSCF/cc-pVDZ level of theory and, at each point of the MEP, dynamical correlation energy was introduced via the SS(6)-CASPT2(18,13)/cc-pVDZ level of theory. The minimum energy crossing point (hereafter called conical intersection, CI) between the $^1(\pi\pi^*)$ and ground state ($(\pi\pi^*/GS)_{CI}$) was optimized at the MS(2)-CASPT2/cc-pVDZ level of theory (Supporting Information). The computed MEPs (gray background) are depicted in Figure 7 and optimized conical intersections in Figure 8.

Table 4. Theoretical Stokes Shift in the Gas Phase, 1,4-Dioxane, and Water computed with the ASEC−FEG Method[a]

		1,4-dioxane	
molecule	gas	ASEC−FEG	exp.
^{tz}A	0.77	0.61	0.60
^{tz}G	0.87	0.65	0.74
^{tz}I	0.40	0.49	0.61
		Water	
^{tz}A		0.73	0.65
^{tz}G		1.06	1.02
^{tz}I		0.51	0.68
		Solvatochromic Shift	
^{tz}A		−0.12	−0.05
^{tz}G		−0.41	−0.28
^{tz}I		−0.02	−0.07

[a]Experimental results are reported for comparison.[29] Theoretical absorption spectra were obtained from a previous work.[36] Solvatochromic shift in the Stokes shift from 1,4-dioxane to water was also computed (negative value indicates that the Stokes shift is larger in water). All of the values are reported in eV.

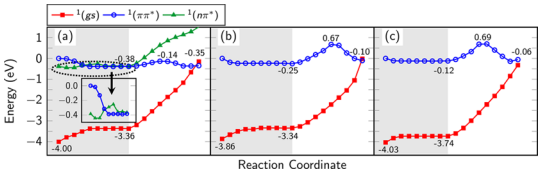


Figure 7. Energetic profiles of the ground and $^1(\pi\pi^*)$ bright states from the Franck−Condon (FC) region toward the $^1(\pi\pi^*)_{min}$ region, by means of the minimum energy path (MEP; gray background), computed at the SA(2)-CASSCF/cc-pVDZ level with the corresponding full-π active space, and from the $^1(\pi\pi^*)_{min}$ region toward the conical intersection with the ground state, by means of linear interpolation in internal coordinates (LIIC; white background), for (a) ^{tz}A, (b) ^{tz}G, and (c) ^{tz}I. At each point of the MEP, vertical excitation energies were computed at the SS(6)-CASPT2(18,13)/cc-pVDZ level of theory.

For the three isothiazole-derived nucleobases (Figure 7), the MEPs starting at the FC region show their $^1(\pi\pi^*)$ potential energy hypersurface evolving in a similar manner, with a barrierless pathway toward the $^1(\pi\pi^*)_{min}$ minimum region. For the ^{tz}A nucleobase, a lower $^1(n\pi^*)$ dark state was observed at the FC region, as it is known to be in 9*H*-adenine,[19,61] its canonical counterpart, for which the two lowest-lying excited states in the gas phase are S_1 $^1(n\pi^*)$ and S_2 $^1(\pi\pi^*)$, respectively.

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

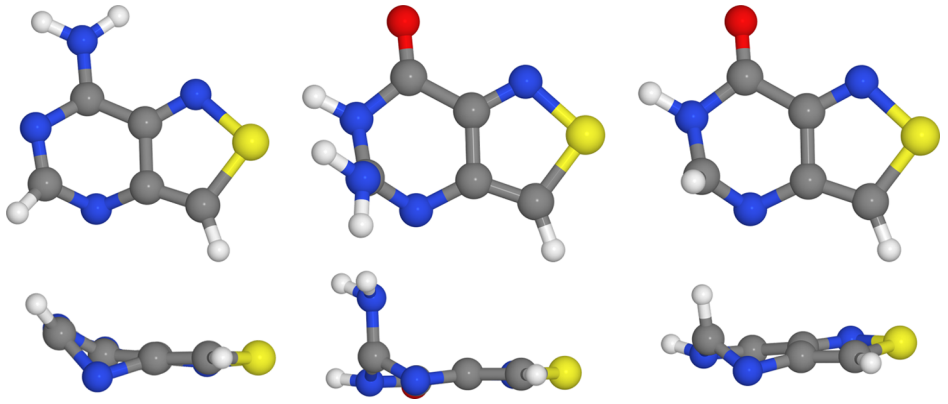


Figure 8. Frontal and side views of ^{tz}**A** (left), ^{tz}**G** (center), and ^{tz}**I** (right) $(\pi\pi^*/GS)_{CI}$ conical intersection structures optimized at the MS(2)-CASPT2/cc-pVDZ level of theory (see Methodology for details).

However, as the MEP evolves to the $^1(\pi\pi^*)_{min}$ region, the bright $^1(\pi\pi^*)$ excited state becomes the first excited state. Although a fraction of the population can be transferred from the $^1(\pi\pi^*)$ state to the $^1(n\pi^*)$ dark state via CI, partly quenching the ^{tz}**A** fluorescence as reported previously in ref 60, we believe that most of the electronic population remains in the bright state, as suggested experimentally; again, a similar behavior was also observed for 9*H*-adenine.[19,61] Furthermore, as mentioned previously, in condensed media, the $^1(\pi\pi^*)$ state is the lowest-lying excited state, reinforcing that the $(\pi\pi^*/n\pi^*)_{CI}$ conical intersection observed in the gas phase is not relevant for describing the photophysics of ^{tz}**A**. Taking the initial point of the MEP as a reference, the $^1(\pi\pi^*)_{min}$ is computed to be −0.38 eV lower in energy for ^{tz}**A**, −0.25 eV for ^{tz}**G**, and −0.12 eV for ^{tz}**I**. It is interesting to mention that the results displayed in Table 3 correspond to the vertical emission energies computed at the SS(6)-CASPT2(18,13)/cc-pVDZ level with the MS-CASPT2-optimized $^1(\pi\pi^*)_{min}$ structures; that is, ^{tz}**A**: 3.23 eV, ^{tz}**G**: 3.00 eV, and ^{tz}**I**: 3.63 eV. On the other hand, the emission energies displayed in Figure 7 were computed from the geometry of the last point of the MEP, obtained at the CASSCF/cc-pVDZ level employing the full π, π^* active space of the corresponding compound, at the SS(6)-CASPT2(18,13)/cc-pVDZ level; in short, ^{tz}**A**: 2.98 eV, ^{tz}**G**: 3.09 eV, and ^{tz}**I**: 3.62 eV. Therefore, by comparing the emission energies computed with the geometries optimized at the last point of the MEPs (obtained at the CASSCF level) to those computed with the MS-CASPT2-optimized geometries (Table 3), we conclude that they are in agreement, differing not more than 0.25 eV.

Once in the $^1(\pi\pi^*)_{min}$ minimum region, two deactivation mechanisms can be foreseen: (i) a nonradiative mechanism via a conical intersection with the ground state ($(\pi\pi^*/GS)_{CI}$) or (ii) relaxation to the ground state via fluorescence. The CI's accessibility was investigated connecting the $^1(\pi\pi^*)_{min}$ and $(\pi\pi^*/GS)_{CI}$ structures via linear interpolation in internal coordinates (LIIC), at the SS(6)-CASPT2(18,13)/cc-pVDZ level of theory. Furthermore, a better understanding of the energetic profile can be obtained by analyzing the initial and final structure geometries and their relative energies.

The $(\pi\pi^*/GS)_{CI}$ structures for ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** structures are located slightly (0.03, 0.10, 0.06 eV, respectively) above the last point of the corresponding MEPs (Figure 7). The final structures obtained along the MEP, as mentioned before, resemble their $^1(\pi\pi^*)_{min}$-optimized geometries, obtained at the MS-CASPT2/cc-pVDZ level, indicating that the CASPT2//CASSCF protocol is appropriate. The $(\pi\pi^*/GS)_{CI}$ and $^1(\pi\pi^*)_{min}$ geometries, on the other hand, differ much from each other. The $^1(\pi\pi^*)_{min}$ structures are mainly planar, while $(\pi\pi^*/GS)_{CI}$ exhibits large distortions on the six-membered rings, while the five-membered rings are planar (Figure 8). Significant changes are observed at the C_2 position for the three molecules, with the bonded moiety placed perpendicular to the six-membered ring for ^{tz}**G** and ^{tz}**I**; for ^{tz}**A**, the N_1 and C_2 positions are distorted. In comparison to the $^1(\pi\pi^*)_{min}$-optimized structure, the C_2N_3 bonds are elongated in all cases. On the other hand, the N_1C_2 is shortened for ^{tz}**A** and stretched for ^{tz}**G** and ^{tz}**I**.

The ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** $(\pi\pi^*/GS)_{CI}$ geometries can be compared to those CIs reported previously for adenine,[61,62] guanine,[63] and inosine (hypoxanthine).[64] For a more detailed comparison, we will make use of the Cremer−Pople[65] and Boeyens[66] parameters (Q, in Å; Φ and Θ, both in deg) for analyzing the six-membered ring conformation, as displayed in Table 5; it is worth mentioning that the $(\pi\pi^*/GS)_{CI}$ structures for adenine,[61] guanine,[63] and inosine[64] were optimized at the CASSCF level. We also compared our structures to that reported by Karsili et al. computed at the SA(2)-CASSCF-(6,6)/6-31G(d) level for ^{tz}**A** (in their work, coined as iTPA),[60] neglecting the dynamical correlation effects during the optimization.

Table 5. Cremer−Pople and Boeyens Parameters for the Six-Membered Ring Conformation Analysis for ^{tz}A, ^{tz}G, ^{tz}I, and Their Natural Nucleobase Analogues Adenine, Guanine, and Inosine (Hypoxanthine)[a]

	Q/Å	Φ/deg	Θ/deg	Boeyens[66]
^{tz}**A** (this work)	0.63	125	74	^{3,6}B
^{tz}**A** (Karsili et al.[60])	0.55	92	65	$^3S_2$
adenine (Serrano-Andrés et al.[61])	0.54	76	66	$^3S_2$
adenine (Barbatti and Lischka[62])	0.72	118	82	^{3,6}B
^{tz}**G**	0.47	68	57	E_2
guanine (Serrano-Andrés et al.[63])	0.49	67	60	E_2
^{tz}**I**	0.51	78	60	$^3S_2$
inosine (Guo et al.[64])	0.51	65	62	E_2

[a]For ^{tz}**A**, we also report the structure obtained by Karsili et al.[60]

The ring conformation analysis reveals that the optimized ^{tz}**A** $(\pi\pi^*/GS)_{CI}$ differs from its corresponding counterpart reported for adenine.[61] The ^{tz}**A** $(\pi\pi^*/GS)_{CI}$ has a ^{3,6}B boat

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

Table 6. Bond Length Alternation Coordinates for Bonds around the C_2 Position (N_1C_2 and C_2N_3 Chemical Bonds) for ^{tz}A, ^{tz}G, and ^{tz}I at the Franck−Condon ($S_{0,min}$), $^1(\pi\pi^*)_{min}$, Transition State (TS) and $(\pi\pi^*/GS)_{CI}$ Regions[a]

	β_X			
	$S_{0,min}$	$S_1\ ^1(\pi\pi^*)_{min}$	TS	$(\pi\pi^*/GS)_{CI}$
^{tz}A	0.042 [0.091]	−0.052 [−0.107]	[−0.051]	−0.154 [−0.140]
^{tz}G	0.079 [0.112]	−0.026 [−0.027]	[−0.020]	−0.024 [0.040]
^{tz}I	0.075 [0.106]	−0.028 [−0.032]	−[0.094]	−0.032 [0.036]

[a]The values of β_X are in Å and it is defined as $\beta_X = r_{N_1C_2} - r_{C_2N_3}$, where the subscript X represents the potential energy hypersurface region. The values computed with the CASSCF-optimized geometries employing the same active spaces and basis sets are displayed in square brackets.

conformation (Table 5), with the positions N_3 and C_6 out of the plane, whereas the adenine is classified as a skew-boat $^3S_2$. More recently, Karsili and co-workers[60] reported $(\pi\pi^*/GS)_{CI}$ for ^{tz}A in close resemblance to that reported for adenine, belonging to the same Boeyens's group $^3S_2$. We have to keep in mind that both structures were optimized at the same level of theory, which might explain why both structures are indexed at the same Boyens group. It is also interesting to mention that Barbatti and Lischka[62] reported, for the first time, a ^{3,6}B conical intersection for adenine, computed at the MR-CIS(6,5)/SA-3-CAS-(12,10)/6-31G* level of theory.

For ^{tz}G and 9*H*-guanine,[63] its canonical nucleobase counterpart, the ring conformation analysis indicates similar $(\pi\pi^*/GS)_{CI}$ conical intersections, characterized by a E_2 envelope conformation at the C_2 position (Table 5). It is also interesting to recall that for 9*H*-guanine in the gas phase,[63] $^1(\pi\pi^*)$ is always the lowest-lying excited state, evolving in a barrierless way from FC to the conical intersection with the ground state.

In the same way, by comparing ^{tz}I and inosine $(\pi\pi^*/GS)_{CI}$ structures, we observe similar *Q* parameters (0.51 Å for both molecules), Φ (78° for ^{tz}I and 65° for inosine), and Θ (60° for ^{tz}I and 62° for inosine). Nevertheless, they are included in distinct groups, with ^{tz}I belonging to the $^3S_2$ skew-boat conformation group and inosine to an envelope E_2 group.

After analyzing the conformation of the $^1(\pi\pi^*)_{min}$ and $(\pi\pi^*/GS)_{CI}$ structures, we will focus our attention on the LIIC path connecting them. As displayed in Figure 7, the $(\pi\pi^*/GS)_{CI}$ region could be reached from the $^1(\pi\pi^*)_{min}$ structure only after surmounting the energetic barriers amounting to 0.25, 0.92, and 0.81 eV for ^{tz}A, ^{tz}G, and ^{tz}I, respectively, hindering the nonradiative population transfer mechanism taking place at the $(\pi\pi^*/GS)_{CI}$ region. Therefore, the extra energy of the $^1(\pi\pi^*)$ state will be released by fluorescence from the $^1(\pi\pi^*)_{min}$ minimum region. Comparing to the work of Karsili et al. about the ^{tz}A molecule, they found that the $^1(\pi\pi^*)_{min}$ structure is lower in energy than $(\pi\pi^*/GS)_{CI}$, even though no barrier was found between the structures.[60] As mentioned above, it is also worth keeping in mind that the last point of the MEPs along the $^1(\pi\pi^*)$ hypersurfaces from the Franck−Condon region, which corresponds to the $^1(\pi\pi^*)_{min}$ minima region, are slightly more stable than the optimized $(\pi\pi^*/GS)_{CI}$ structures (Figure 7), reinforcing that most of the electronic population will remain on the corresponding $^1(\pi\pi^*)_{min}$ minima region.

Aware about the LIIC limitations and that the computed LIIC pathway represents an upper limit to the real path, leading to overestimated energetic barriers, we decided to estimate the energetic barrier by optimizing the transition states (TS) between the critical points for all of the molecules at the SA(2)-CASSCF/cc-pVDZ level, because it is not computationally feasible to carry it out at the MS-CASPT2 level. In comparison to the barriers computed with the LIIC pathways, those obtained with optimized TS structures are systematically lower for all molecules (0.18 eV for ^{tz}A, 0.37 eV for ^{tz}G, and 0.68 eV for ^{tz}I), but still high to be overcome. We conclude that these molecules fluoresce in both the gas phase and condensed media, despite no experimental evidence in the gas phase to date. A recent investigation[67] about thienoguanosine (thG), a highly fluorescent analogue of guanosine and ^{tz}G, carried out at DFT and CASSCF levels, showed a deactivation mechanism similar to that reported by us for ^{tz}G, with a corresponding energetic barrier ≥0.4 eV.

A closer look at the geometric changes observed from the Franck−Condon to the $(\pi\pi^*/GS)_{CI}$ region gives us some clues about the energetic barrier nature. As the main geometric changes are related to the 6-membered rings, specifically around the C_2 position, we decided to use a new coordinate ($\beta_X = r_{N_1C_2} - r_{C_2N_3}$, the subscript X represents the potential energy hypersurface region, namely, FC, $^1(\pi\pi^*)_{min}$, TS, or CI region) that coined the bond length alternation (BLA) to quantify the alteration of chemical bonds around the C_2 atom.

As shown in Table 6, for the three compounds studied in this work, the N_1C_2 bonds are longer than the C_2N_3 bonds (by 0.04 Å for ^{tz}A and 0.08 Å for ^{tz}G and ^{tz}I) at the Franck−Condon region; the same is true for the CASSCF geometries, but with larger β values. Around the $^1(\pi\pi^*)_{min}$ energy regions, the bonds lengths are alternated, that is, the C_2N_3 bonds are longer than the N_1C_2 bonds (by about 0.05 Å for ^{tz}A and 0.03 Å for ^{tz}G and ^{tz}I); the CASSCF-optimized geometries exhibit the same trends. The transition states on the potential energy surface of the $S_1\ ^1(\pi\pi^*)$ excited state of all molecules show different results. For instance, β_{TS} for ^{tz}A and ^{tz}G exhibits the same trend as that observed at the $^1(\pi\pi^*)_{min}$ region. On the other hand, the ^{tz}I molecule presents a significant change from −0.03 to 0.09 Å in this coordinate, which can provide an explanation for the higher energy barrier observed for this molecule, at least at the CASSCF level. In the CI region, a distinct trend is observed. For ^{tz}A, β_X is negative, while for ^{tz}G and ^{tz}I the values of β_{CI} are positive and similar, around 0.02 and 0.03 Å, respectively, with the same trend noticed for the CASSCF structures. The ^{tz}G and ^{tz}I are similar around the C_2 position throughout the S_1 deactivation pathway, with the most striking difference at the TS region, computed at another level of theory. Both molecules differ drastically from ^{tz}A at the $(\pi\pi^*/GS)_{CI}$ region, which is reflected in the Boeyens classification (Table 5).

We have also investigated the solvation effects of 1,4-dioxane and water on the photochemical pathways described above, using the ASEC−FEG method to include solvation on selected points along the photochemical paths, as minima, barriers, and conical intersection structures. It is worth pointing out that the geometries were not optimized in the solvent, but in the gas

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

Table 7. Energetic Barriers (Computed as the Energy Difference between the $^1(\pi\pi^*)_{min}$-Optimized Geometry (E_{min}) and the Energy of the Structure Corresponding to the Maximum of the LIIC Path from the $^1(\pi\pi^*)_{min}$ to the $(\pi\pi^*/GS)_{CI}$ Structure (Barrier (LIIC)) and as the Energy Difference between the Optimized Transition State Structure and the $^1(\pi\pi^*)_{min}$-Optimized Geometry (Barrier (TS))) and Energy Difference between the Minimum and the Conical Intersection ($E_{min} - E_{CI}$) for Gas and Condensed Phases (1,4-Dioxane and Water)[a]

	barrier (LIIC)			barrier (TS)			$E_{min} - E_{CI}$		
molecule	gas	dioxane	water	gas	dioxane	water	gas	dioxane	water
^{tz}**A**	0.25	0.24	0.38	0.18	0.13	0.24	−0.16	0.30	0.52
^{tz}**G**	0.92	1.11	0.99	0.37	0.28	0.29	−0.30	0.87	0.06
^{tz}**I**	0.81	0.94	0.98	0.68	0.60	0.76	−0.26	0.22	0.06

[a]Vertical excitation energies were computed at the SS(6)-CASPT2(18,13)/cc-pVDZ level; all of the values in eV.

phase only, and solvation effects were simulated frozen conical intersection geometries as optimized in the gas phase. As the conical intersection geometries were not reoptimized in the presence of the solvents, we do not expect the energies computed in the solution to be degenerate. Indeed, the energy gaps computed in solution (smaller than 0.4 eV) are larger than in the gas phase (where they are not larger than 0.1 eV), suggesting that solvation effects affect the geometry, energy, and location of the CI. Nonetheless, we do believe that the most relevant fact is the height of the barriers that must be surmounted to achieve the conical intersection regions. As the barriers are high, the conical intersections are inaccessible (in the gas phase or solution) and, consequently, the compounds will fluoresce. In all cases, the position of the CIs is computed using the energy of the $^1(\pi\pi^*)_{min}$ state, at the corresponding conical intersection-optimized geometry, as the reference value. Computed values in either the gas phase or the solvent are displayed in Table 7.

For the fluorescent nucleoside analogues, solvation effects (1,4-dioxane and water) increase the energetic barrier provided by the LIIC scan along the pathway from $^1(\pi\pi^*)_{min}$ to $(\pi\pi^*/GS)_{CI}$ structures, with the largest variation in dioxane being 0.19 eV for ^{tz}**G** followed by ^{tz}**I** (0.13 eV). As for the aqueous medium, an increase of 0.17 eV in the energetic barrier is observed for ^{tz}**I** (with increases of 0.07 eV for ^{tz}**G** and of 0.13 eV for ^{tz}**A**). We also studied solvation effects on the energetic barrier, using the optimized transition state structure instead of the corresponding maximum of the linear interpolation path. As it can be noticed in Table 7, it is not possible to observe any trend about solvation effects.

The relative position of $^1(\pi\pi^*)_{min}$ and $(\pi\pi^*/GS)_{CI}$ (Table 7) observed for ^{tz}**A** increases to 0.30 eV in 1,4-dioxane and 0.52 eV in water. In the opposite way, for ^{tz}**G**, the energy difference between the structures increases to 0.87 eV in 1,4-dioxane, but in water they are placed even closer in energy, separated by only 0.06 eV. Solvent effects change the energetic difference observed for ^{tz}**I** in the gas phase, to about 0.22 eV in 1,4-dioxane and 0.06 eV in water. Therefore, solvent effects enhance fluorescence via vibrational energy transfer to the solvents, as mentioned before.[60]

The above discussion can be summarized as follows: the photochemical deactivation pathways of fluorescent nucleoside analogues ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** are similar; after absorption of energy, the $^1(\pi\pi^*)$ state is populated and follows downhill toward its $^1(\pi\pi^*)_{min}$ region, from which the $(\pi\pi^*/GS)_{CI}$ region cannot be reached due to the presence of a high energetic barrier. Therefore, the excess of energy is released from the $^1(\pi\pi^*)_{min}$ region as fluorescence; solvent effects do not alter the mechanism, inducing only small changes on the energetic barrier.

A similar pattern was described by us for the pyrimidine fluorescent nucleosides.[37] The difference relies on the energetic separation between minima and optimized $(\pi\pi^*/GS)_{CI}$ structures, which are due to the blockage of the C_4C_5 rotation, that would be the main geometrical change on the deactivation path of pyrimidines. For the ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** purine fluorescent nucleosides, the main geometrical changes observed on the CI structures are not at the C_4C_5 bond, but at the N_1C_2 and C_2N_3 bonds, as observed in the canonical bases. Therefore, the fluorescence is not due to the blockage of the C_4C_5 rotation, but probably due to the red-shift of the absorption energy which would not allow that the molecules have the necessary energy to surpass the energy barrier and reach the CI region. Therefore, although purine and pyrimidine fluorescent nucleoside analogues share common deactivation pathways, the molecular mechanisms behind the emission differ slightly.

CONCLUSIONS

The photophysical deactivation pathways of three modified purine nucleobases ^{tz}**A**, ^{tz}**G**, and ^{tz}**I** were investigated at the CASPT2/cc-pVDZ level, using Sequential QM/MM (Monte Carlo simulations) to unravel the structural changes they present on the respective excited states and the solvent shell structures around them. The mechanisms that result in their fluorescence in gas phase, 1,4-dioxane, and aqueous media were also explored. A global analysis of the solvation shells in the ground and excited states revealed that the profile of the minimum pairwise distribution function is quite similar, with two well-defined solvation shells. Nevertheless, looking into the solvent molecules that make hydrogen bonds with the solute, noticeable changes between the ground and excited states as well as the radial distribution function evolving these atomic sites were identified. As for the photophysical deactivation mechanisms, once excited to their respective lowest $^1(\pi\pi^*)$ state, they can reach the state minima in a barrierless path. The pathway to the $(\pi\pi^*/GS)_{CI}$ region is blocked by significant energy barriers, which hinders a possible internal conversion mechanism. Therefore, the main deactivation path is the emission from the $^1(\pi\pi^*)_{min}$ region.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jpca.0c03398.

Cartesian coordinates of the optimized structures (PDF)

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844

AUTHOR INFORMATION

Corresponding Author

Antonio Carlos Borin − *Department of Fundamental Chemistry, Institute of Chemistry, University of São Paulo, 05508-000 São Paulo, SP, Brazil;* orcid.org/0000-0003-3047-2044; Email: ancborin@iq.usp.br

Authors

Adalberto Vasconcelos Sanches de Araújo − *Department of Fundamental Chemistry, Institute of Chemistry, University of São Paulo, 05508-000 São Paulo, SP, Brazil*

Danillo Valverde − *Institute of Physics, University of São Paulo, 05508-090 São Paulo, SP, Brazil*

Sylvio Canuto − *Institute of Physics, University of São Paulo, 05508-090 São Paulo, SP, Brazil;* orcid.org/0000-0002-9942-8714

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.jpca.0c03398

Author Contributions

[§]A.V.S.d.A. and D.V. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

A.C.B. and A.V.S.d.A. thank the CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico) for research fellowships (project numbers 302318/2017-8 and 153104/2015-5, respectively). A.C.B. thanks FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) for the research grant 2018/19454-5. D.V. acknowledges financial support from FAPESP, under grant 2017/02612-4. S.C. thanks CAPES for the BioMol project 23038.004630/2014-35 and the National Institute of Science and Technology Complex Fluids (INCT-FCx) with the CNPq grant 141260/2017-3 and FAPESP grant 2014/50983-3. Computations were performed on resources provided by HPC through Superintendência de Tecnologia da Informação da Universidade de São Paulo. The authors acknowledge the National Laboratory for Scientific Computing (LNCC/MCTI, Brazil) for providing HPC resources of the SDumont supercomputer (http://sdumont.lncc.br), which have contributed to the research results reported within this paper.

REFERENCES

(1) Daniels, M.; Hauswirth, W. W. Fluorescence of the Purine and Pyrimidine Bases of the Nucleic Acids in Neutral Aqueous Solution at 300 K. *Science* **1971**, *171*, 675−677.

(2) Morgan, J. P.; Daniels, M. Polarised Fluorescence from the Thymine Family of DNA Bases in Neutral Aqueous Solution at ∼300 K. *Chem. Phys. Lett.* **1979**, *67*, 533−537.

(3) Improta, R.; Santoro, F.; Blancafort, L. Quantum Mechanical Studies on the Photophysics and the Photochemistry of Nucleic Acids and Nucleobases. *Chem. Rev.* **2016**, *116*, 3540−3593.

(4) Mai, S.; Richter, M.; Marquetand, P.; González, L. Excitation of Nucleobases from a Computational Perspective II: Dynamics. In *Photoinduced Phenomena in Nucleic Acids I—Topics in Current Chemistry*; Barbatti, M.; Borin, A. C.; Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355, pp 99−153.

(5) Giussani, A.; Segarra-Martín, J.; Roca-Sanjuán, D.; Merchán, M. Excitation of Nucleobases from a Computational Perspective I: Reaction Paths. In *Photoinduced Phenomena in Nucleic Acids I—Topics in Current Chemistry*; Barbatti, M.; Borin, A. C.; Ullrich, S., Eds.; Springer International Publishing: Cham, Switzerland, 2015; Vol. 355, pp 57−97.

(6) *Photoinduced Phenomena in Nucleic Acids I—Nucleobases in the Gas Phase and in Solvents*; Barbatti, M.; Borin, A. C.; Ullrich, S., Eds.; Springer, 2015; Vol. 355.

(7) *Photoinduced Phenomena in Nucleic Acids II—DNA Fragments and Phenomenological Aspects*; Barbatti, M.; Borin, A. C.; Ullrich, S., Eds.; Springer, 2015; Vol. 356.

(8) Gustavsson, T.; Improta, R.; Markovitsi, D. DNA/RNA: Building Blocks of Life Under UV Irradiation. *J. Phys. Chem. Lett.* **2010**, *1*, 2025−2030.

(9) Barbatti, M.; Borin, A. C.; Ullrich, S. *Photoinduced Phenomena in Nucleic Acids I—Nucleobases in the Gas Phase and in Solvents*; Barbatti, M.; Borin, A. C.; Ullrich, S., Eds.; Springer: 2015; Vol. 355, pp 1−32.

(10) Lakowicz, J. R. *Principles of Fluorescence Spectroscopy*; Springer, New York, 2006.

(11) Hell, S. W. Toward Fluorescence Nanoscopy. *Nat. Biotechnol.* **2003**, *21*, 1347−1355.

(12) *Fluorescent Analogs of Biomolecular Building Blocks, Design and Applications*; Wilhelmsson, M.; Tor, Y., Eds.; Wiley: New Jersey, 2016.

(13) Hirao, I. Unnatural Base Pair Systems for DNA/RNA-based Biotechnology. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622−627.

(14) Krueger, A. T.; Kool, E. T. Redesigning the Architecture of the Base Pair: Toward Biochemical and Biological Function of New Genetic Sets. *Chem. Biol.* **2009**, *16*, 242−248.

(15) Krueger, A. T.; Kool, E. T. Model Systems for Understanding DNA Base Pairing. *Curr. Opin. Chem. Biol.* **2007**, *11*, 588−594.

(16) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, *110*, 2579−2619.

(17) Kimoto, M.; Cox, R. S.; Hirao, I. Unnatural Base Pair Systems for Sensing and Diagnostic Applications. *Expert Rev. Mol. Diag.* **2011**, *11*, 321−331.

(18) Nordlund, T. M.; Andersson, S.; Nilsson, L.; Rigler, R.; Gräslund, A.; McLaughlin, L. W. Structure and Dynamics of a Fluorescent DNA Oligomer Containing the EcoRI Recognition Sequence: Fluorescence, Molecular Dynamics, and NMR Studies. *Biochemistry* **1989**, *28*, 9095−9103.

(19) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. Adenine and 2-Aminopurine: Paradigms of Modern Theoretical Photochemistry. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8691−8696.

(20) Ludwig, V.; do Amaral, M. S.; da Costa, Z. M.; Borin, A. C.; Canuto, S.; Serrano-Andrés, L. 2-Aminopurine Non-Radiative Decay and Emission in Aqueous Solution: A Theoretical Study. *Chem. Phys. Lett.* **2008**, *463*, 201−205.

(21) Gaied, N. B.; Glasser, N.; Ramalanjaona, N.; Marquet, R.; Burger, A.; Mely, Y. 8-Vinyl-deoxyadenosine, an Alternative Fluorescent Nucleoside Analog to 2'-deoxyribosyl-2-aminopurine with Improved Properties. *Nucleic Acids Res.* **2005**, *33*, 1031−1039.

(22) Nadler, A.; Strohmeier, J.; Diederichsen, U. 8-Vinyl-2'-deoxyguanosine as a Fluorescent 2'-deoxyguanosine Mimic for Investigating DNA Hybridization and Topology. *Angew. Chem., Int. Ed.* **2011**, *50*, 5392−5396.

(23) Passow, K. T.; Harki, D. A. 4-Cyanoindole-2'-deoxyribonucleoside (4CIN): A Universal Fluorescent Nucleoside Analogue. *Org. Lett.* **2018**, *20*, 4310−4313.

(24) Burns, D. D.; Teppang, K. L.; Lee, R. W.; Lokensgard, M. E.; Purse, B. W. Fluorescence Turn-On Sensing of DNA Duplex Formation by a Tricyclic Cytidine Analogue. *J. Am. Chem. Soc.* **2017**, *139*, 1372−1375.

(25) Specht, E. A.; Braselmann, E.; Palmer, A. E. A Critical and Comparative Review of Fluorescent Tools for Live-Cell Imaging. *Annu. Rev. Physiol.* **2017**, *79*, 93−117.

(26) Xu, W.; Chan, K. M.; Kool, E. T. Fluorescent nucleobases as tools for studying DNA and RNA. *Nat. Chem.* **2017**, *9*, 1043−1055.

(27) Saito, Y.; Hudson, R. H. E. Base-modified fluorescent purine nucleosides and nucleotides for use in oligonucleotide probes. *J. Photochem. Photobiol., C* **2018**, *36*, 48−73.

(28) Hirashima, S.; Han, J. H.; Tsuno, H.; Tanigaki, Y.; Park, S.; Sugiyama, H. New Size-Expanded Fluorescent Thymine Analogue: Synthesis, Characterization, and Application. *Chem. - Eur. J* **2019**, *25*, 9913−9919.
(29) Rovira, A. R.; Fin, A.; Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015**, *137*, 14602−14605.
(30) Rovira, A. R.; Fin, A.; Tor, Y. Expanding a Fluorescent RNA Alphabet: Synthesis, Photophysics and Utility of Isothiazole-derived Purine Nucleoside Surrogates. *Chem. Sci.* **2017**, *8*, 2983−2993.
(31) Samanta, P. K.; Manna, A. K.; Pati, S. K. Thieno Analogues of RNA Nucleosides: A Detailed Theoretical Study. *J. Phys. Chem. B* **2012**, *116*, 7618−7626.
(32) Gedik, M.; Brown, A. Computational Study of the Excited State Properties of Modified RNA Nucleobases. *J. Photochem. Photobiol., A* **2013**, *259*, 25−32.
(33) Chawla, M.; Poater, A.; Oliva, R.; Cavallo, L. Structural and Energetic Characterization of the Emissive RNA Alphabet Based on the Isothiazolo[4,3-d]pyrimidine Heterocycle Core. *Phys. Chem. Chem. Phys.* **2016**, *18*, 18045−18053.
(34) Zhang, L.; Kong, X.; Wang, M.; Zheng, M. Photophysical Properties of the Isomorphic Emissive RNA Nucleobase Analogues and Effect of Water Solution, Ribose, and Base Pairing: A theoretical Study. *Int. J. Quantum Chem.* **2017**, *117*, No. e25377.
(35) Miertuš, S.; Scrocco, E.; Tomasi, J. Electrostatic interaction of a solute with a continuum. A direct utilizaion of AB initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **1981**, *55*, 117−129.
(36) Valverde, D.; de Araujo, A. V. S.; Borin, A. C.; Canuto, S. Electronic Structure and Absorption Spectra of Fluorescent Nucleoside Analogues. *Phys. Chem. Chem. Phys.* **2017**, *19*, 29354−29363.
(37) Valverde, D.; de Araujo, A. V. S.; Canuto, S.; Borin, A. C. Photophysics of Emissive ^{tz}C[Isothiazolo-Cytidine] and ^{tz}U-[Isothiazolo-Uridine] Pyrimidine Analogues. *ChemPhotoChem* **2019**, *3*, 916−924.
(38) Coutinho, K.; Canuto, S. Solvent Effects in Emission Spectroscopy: A Monte Carlo Quantum Mechanics Study of the $n \leftarrow \pi^*$ Shift of Formaldehyde in Water. *J. Chem. Phys.* **2000**, *113*, 9132−9139.
(39) Orozco-Gonzalez, Y.; Coutinho, K.; Peon, J.; Canuto, S. Theoretical Study of the Absorption and Nonradiative Deactivation of 1-Nitronaphthalene in the Low-Lying Singlet and Triplet Excited States Including Methanol and Ethanol Solvent Effects. *J. Chem. Phys.* **2012**, *137*, No. 054307.
(40) Finley, J.; Malmqvist, P.-Å.; Roos, B. O.; Serrano-Andrés, L. The Multi-State CASPT2 Method. *Chem. Phys. Lett.* **1998**, *288*, 299−306.
(41) Roos, B. O. *Advances in Chemical Physics: Ab Initio Methods in Quantum Chemistry Part 2*; Lawley, K. P., Ed.; John Wiley & Sons, Inc.: Chichester, U.K., 1987; Vol. *69*, pp 399−445.
(42) Dunning, T. H., Jr. Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron Through Neon and Hydrogen. *J. Chem. Phys.* **1989**, *90*, 1007−1023.
(43) De Vico, L.; Olivucci, M.; Lindh, R. New General Tools for Constrained Geometry Optimizations. *J. Chem. Theory Comput.* **2005**, *1*, 1029−1037.
(44) Andersson, K.; Malmqvist, PÅ.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. Second-Order Perturbation Theory With a CASSCF Reference Function. *J. Phys. Chem. A.* **1990**, *94*, 5483−5488.
(45) Andersson, K.; Malmqvist, PÅ.; Roos, B. O. Second-Order Perturbation Theory with a Complete Active Space Self-Consistent Field Reference Function. *J. Chem. Phys.* **1992**, *96*, 1218−1226.
(46) Andersson, K.; Roos, B. O. *Advanced Series in Physical Chemistry*; Yarkony, D. R., Ed.; World Scientific, 1995; Vol. *2*, pp 55−109.
(47) Forsberg, N.; Malmqvist, PÅ. Multiconfiguration Perturbation Theory with Imaginary Level Shift. *Chem. Phys. Lett.* **1997**, *274*, 196−204.
(48) Ghigo, G.; Roos, B. O.; Malmqvist, PÅ. A Modified Definition of the Zeroth-Order Hamiltonian in Multiconfigurational Perturbation Theory (CASPT2). *Chem. Phys. Lett.* **2004**, *396*, 142−149.
(49) Zobel, J. P.; Nogueira, J. J.; González, L. The IPEA Dilemma in CASPT2. *Chem. Sci.* **2017**, *8*, 1482−1499.
(50) Aquilante, F.; Autschbach, J.; Carlson, R. K.; Chibotaru, L. F.; Delcey, M. G.; De Vico, L.; Fdez. Galván, I.; Ferré, N.; Frutos, L. M.; Gagliardi, L.; et al. Molcas 8: New Capabilities for Multiconfigurational Quantum Chemical Calculations Across the Periodic Table. *J. Comput. Chem.* **2016**, *37*, 506−541.
(51) Georg, H. C.; Canuto, S. Electronic Properties of Water in Liquid Environment. A Sequential QM/MM Study Using the Free Energy Gradient Method. *J. Phys. Chem. B* **2012**, *116*, 11247−11254.
(52) Coutinho, K.; Georg, H.; Fonseca, T.; Ludwig, V.; Canuto, S. An Efficient Statistically Converged Average Configuration for Solvent Effects. *Chem. Phys. Lett.* **2007**, *437*, 148−152.
(53) Okuyama-Yoshida, N.; Nagaoka, M.; Yamabe, T. Transition-State Optimization on Free Energy Surface: Toward Solution Chemical Reaction Ergodography. *Int. J. Quantum Chem.* **1998**, *70*, 95−103.
(54) Nagaoka, M. Structure Optimization of Solute Molecules Via Free Energy Gradient Method. *Bull. Korean Chem. Soc.* **2003**, *24*, 805−808.
(55) Cezar, H. M.; Canuto, S.; Coutinho, K. DICE: AMonte Carlo Code for Molecular Simulation Including the Configurational Bias Monte Carlo Method. *J. Chem. Inf. Model.* **2020**, *7*, 3472−3488.
(56) Georg, H. C.; Coutinho, K.; Canuto, S. Solvent Effects on the UV-Visible Absorption Spectrum of Benzophenone in Water: A Combined Monte Carlo Quantum Mechanics Study Including Solute Polarization. *J. Chem. Phys.* **2007**, *126*, No. 034507.
(57) Coutinho, K.; Canuto, S. The Sequential Monte Carlo-Quantum Mechanics Methodology. Application to the Solvent Effects in the Stokes Shift of Acetone in Water. *J. Mol. Struct.: THEOCHEM* **2003**, *632*, 235−246.
(58) Swiatla-Wojcik, D. Evaluation of the Criteria of Hydrogen Bonding in Highly Associated Liquids. *Chem. Phys.* **2007**, *342*, 260−266.
(59) Mezei, M.; Beveridge, D. L. Theoretical Studies of Hydrogen Bonding in Liquid Water and Dilute Aqueous Solutions. *J. Chem. Phys.* **1981**, *74*, 622−632.
(60) Karsili, T. N. V.; Thodika; Nguyen, M. L.; Matsika, S. The Origin of Fluorescence in DNA Thio-Analogues. *Chem. Phys.* **2018**, *515*, 434−440.
(61) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. A Three-State Model for the Photophysics of Adenine. *Chem. - Eur. J.* **2006**, *12*, 6559−6571.
(62) Barbatti, M.; Lischka, H. Nonadiabatic Deactivation of 9H-Adenine: A Comprehensive Picture Based on Mixed Quantum-Classical Dynamics. *J. Am. Chem. Soc.* **2008**, *130*, 6831−6839.
(63) Serrano-Andrés, L.; Merchán, M.; Borin, A. C. Three-State Model for the Photophysics of Guanine. *J. Am. Chem. Soc.* **2009**, *130*, 2473−2484.
(64) Guo, X.; Lan, Z.; Cao, Z. Ab Initio Insight Into Ultrafast Nonadiabatic Decay of Hypoxanthine: Keto-N7H and Keto-N9H Tautomers. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10777−10782.
(65) Cremer, D. T.; Pople, J. A. General Definition of Ring Puckering Coordinates. *J. Am. Chem. Soc.* **1975**, *97*, 1354−1358.
(66) Boeyens, J. C. A. The Conformation of Six-Membered Rings. *J. Cryst. Mol. Struct.* **1978**, *8*, 317−320.
(67) Martinez-Fernandez, L.; Gavvala, K.; Sharma, R.; Didier, P.; Richert, L.; Segarra Martì, J.; Mori, M.; Mely, Y.; Improta, R. Excited-State Dynamics of Thienoguanosine, an Isomorphic Highly Fluorescent Analogue of Guanosine. *Chem. - Eur. J.* **2019**, *25*, 7375−7386.

https://dx.doi.org/10.1021/acs.jpca.0c03398
J. Phys. Chem. A 2020, 124, 6834−6844