NJC

PAPER

View Article Online
View Journal | View Issue

Cite this: *New J. Chem.*, 2015, 39, 9249

Watson–Crick base pairing, electronic and photophysical properties of triazole modified adenine analogues: a computational study†

Shubhajit Das,[a] Pralok K Samanta‡[b] and Swapan K Pati*[ab]

Received (in Montpellier, France) 19th June 2015,
Accepted 14th September 2015

DOI: 10.1039/c5nj01566a

www.rsc.org/njc

We employ first-principles Density Functional Theory (DFT) and time-dependent DFT (TDDFT) to elucidate structural, electronic and optical properties of a few recently reported triazole adenine nucleobase analogues. The results are compared against the findings obtained for both the natural adenine nucleobase and available experimental data. The optical absorption of these adenine analogues is calculated both in the gas-phase and in the solvent (methanol) using the Polarized Continuum Model (PCM). We find that all the analogues show a red-shifted absorption profile as compared to adenine. Our simulated emission spectra in the solvent compared fairly well with experimentally observed results. We investigate base paring ability of these adenine analogues with thymine. The calculations of the intrinsic stability of these base pairs ascertain that all the adenine analogues form a hydrogen bonded Watson–Crick base pair with similar H-bonding energy to that obtained for a natural adenine–thymine base pair. In our study, we provide a microscopic origin of the low-energy absorption and emission peaks, observed experimentally.

1 Introduction

In the quest to better understand the natural genetic system, nucleic acid modification[1] comprises an important approach. In principle, each of the three major components of nucleic acids, namely, the phosphodiester backbone, the deoxyribose or ribose sugar moieties and nucleobases can be modified separately.[2–14] However, nucleobase modifications continue to entice the nucleic acid research community as artificial nucleobases have already found potential applications in biology[15] and biotechnology.[16] These modified nucleobases exhibit certain attributes that are different from their natural counterparts. Purine and pyrimidine nucleobases, existing in deoxyribonucleic acid (DNA) and ribonucleic acid (RNA), are not fluorescence active due to their ultrafast excited state life time and very low quantum yields.[17] This virtually non-fluorescent nature of naturally occurring nucleobases has driven the exploration of their fluorescent analogues in the past three decades.[18–26] Now, the obvious question is why so much attention is paid to create fluorescence activity in a nucleobase? The answer lies in the uniqueness of fluorescence spectroscopy as a powerful tool for investigating complex biological systems.[27,28] Fluorescence not only allows capturing real-time dynamics of single molecules in living systems but also it can complement techniques, like, NMR spectroscopy and X-ray crystallography, which are routinely used for structural elucidations of nucleic acid containing systems.[29] Therefore, the use of modified nucleobases as fluorescence probes in nucleic acid chains enables a great deal of insight into the nucleic acid structure, dynamics and interactions with other biomolecules. Furthermore, the fluorescence signal from these nucleobase analogues can be monitored easily due to non-fluorescent natural counterparts offering no competing background signal. So, carefully designed fluorescent nucleic acid base analogues (FBA) are extremely valuable for biomedical applications.[25] However, modifications of a nucleic acid aiming at replacing a natural nucleobase within an oligonucleotide sequence by its fluorescent analogue imposes a great challenge because such a modification would most likely lead to serious structural and functional perturbations. The key challenge is to design nucleobase analogues which would show intrinsic fluorescence emission while causing minimal perturbations upon modifications. Evidently, fluorescent analogues possessing strong structural resemblance to natural nucleobases are important in this context. These analogues retain structural requirements for base pairing with one of the natural nucleobases through hydrogen bonds and therefore do not cause significant perturbations to the nucleic acid structures.

[a] *New Chemistry Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560064, India*
[b] *Theoretical Sciences Unit, Jawaharlal Nehru Centre for Advanced Scientific Research, Bangalore 560064, India. E-mail: pati@jncasr.ac.in; Fax: +91-80-22082767; Tel: +91-80-22082839*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c5nj01566a

‡ Present address: Solar and Photovoltaics Engineering Research Center, King Abdullah University of Science and Technology, Thuwal 23955-6900, Kingdom of Saudi Arabia.

Fluorescent nucleobase analogues have been studied extensively. For example, thymidine analogues synthesized by Tor *et al.*[18] show fluorescent properties. Pteridine analogues of guanosine and adenosine designed by Hawkins *et al.*[23] have been well studied. Highly responsive 2-aminopurine, most widely used adenine analogue, has also been characterized in detail.[24] Several theoretical studies on fluorescent nucleobase analogues are also reported in the literature in order to explore their structural, electronic, photophysical and base pairing properties.[30–33] Theoretical studies are important in the sense that they could be quite useful for the interpretation of experimental data and at the same time they may provide ways to design optimal fluorescent nucleobase analogues with a desired set of physiochemical properties. Thieno-modified[34–36] and size-expanded[36–39] RNA nucleosides have already been well studied in this respect.

Modifications of a purine or a pyrimidine moiety by attaching additional conjugating linkers constitute a special class of modified nucleobases. This chromophoric extension invokes unique photophysical properties that are distinctly different from the parent nucleobase. Adenine shows a very short excited state lifetime and a low fluorescence quantum yield[40] and in this respect, non adiabatic dynamics for different ultrafast deactivation channels in adenine have been studied in detail.[41,42] Interestingly, adenine (A) analogues, modified in the purine 8 position, have been found to show fluorescent properties *e.g.* C8-modified triazole adenine analogues show fluorescence in the UV-Vis region with a high fluorescence quantum yield.[43] Another C8-modified adenine analogue has shown promising base mimicking and photophysical properties, both as a monomer and in DNA.[44] However, C8 modifications have been shown to impose destabilizing effects on the DNA duplex structure but modifications in the purine 7-position are well tolerated within the DNA helix.[45–47] Recently, Wilhelmsson and Grøtli *et al.* have synthesized a series of C7 substituted triazole adenine analogues which have been shown to have promising photophysical properties.[48] These analogues are obtained by introducing a triazole ring at the 7-position of the adenine ring.

Motivated by their work, we have performed detailed quantum chemical calculation in order to understand structural, energetic and photophysical properties of this new family of triazole adenine analogues (see Fig. 1 for structures). It is important that these analogues form H-bonded Watson–Crick base pairs with thymine, a natural complementary base partner of adenine, when they are incorporated into duplex DNA. In this respect, we calculate the intrinsic stability of each of the base pair, involving an adenine analogue and thymine (hereby denoted as ^{t}A–T pairs), and compare these results with naturally occurring adenine–thymine (A–T) base pairs. Our results show the stability of WC base pair formation between the modified adenine analogues and thymine. We also focus on the absorption and emission characteristics in these analogues and obtain reasonably good agreement with experimental findings. Furthermore, in this study, we provide a microscopic understanding of the low energy optical transitions in the absorption and emission spectra obtained experimentally.

2 Computational details

All electronic structure calculations are performed using the Gaussian 09[49] suit of programs. We have optimized the geometries of all analogue monomers using Becke's three-parameter hybrid

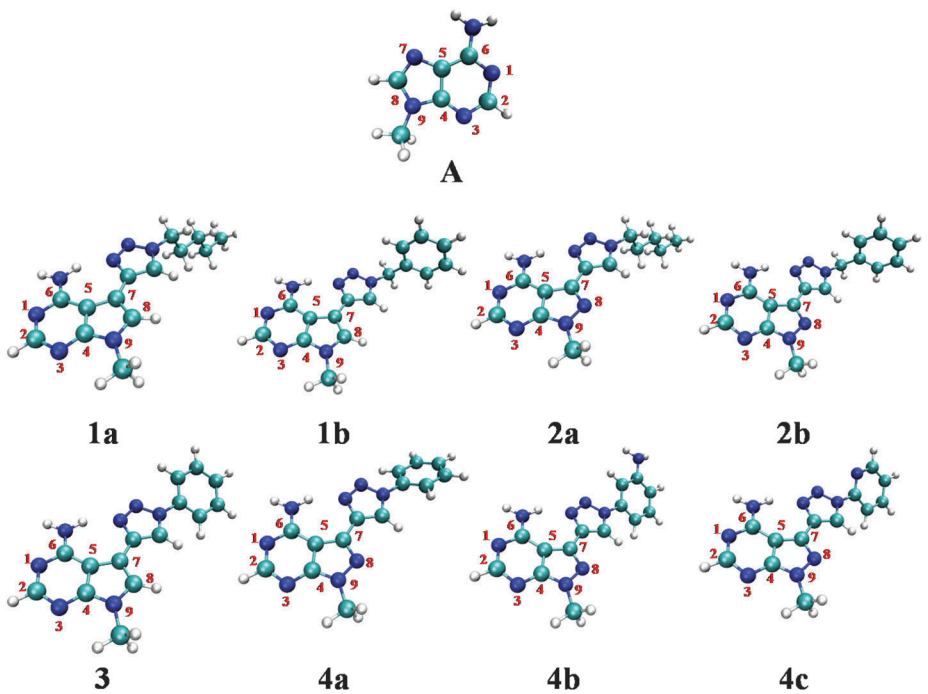

Fig. 1 B3LYP/6-31++G(d,p) level optimized structures of triazole adenine analogues. Atom colour code: blue (N), cyan (C), white (H).

exchange functional and Lee–Young–Parr correlation functional (B3LYP)[50–52] paired with the 6-31++G(d,p) basis set for all atoms. In addition to that, as non-covalent interactions are important for studying H-bonded WC base pairs, we have employed two different exchange and correlation functionals like M062X,[53] capable of capturing medium range electron correlation, and long range corrected ωB97XD[54,55] for the optimization of base pair geometries. Note that the structures and energetics of the H-bonded base pairs are strongly dependent on the level of computation. Electronic absorption spectra are calculated using time-dependent DFT (TDDFT) methods as implemented in the Gaussian 09 package. Emission spectra are calculated on optimized lowest optically excited state geometry (S_1) of each analogue using TDDFT methods. A frequency analysis is carried out to remove any vibrational unstable mode from the optimized structures and also to calculate zero-point energies. Implicit solvent (methanol in our case) effects have been taken into account by the polarizable continuum model (PCM).[56]

3 Results and discussion

3.1 Molecular structures of the adenine nucleobase analogues

All adenine analogues considered in this study are obtained from purine 7-modification of natural adenine by a triazole moiety. These analogues differ from each other primarily by the substituent attached with the triazole ring. These substitutions are of two kinds; aliphatic (*n*-pentyl group in **1a** and **2a**, benzyl group in **1b** and **2b**) and aromatic (phenyl group for **3** and **4a**, 3-aminophenyl for **4b** and pyridine for **4c**). Furthermore, in some of the analogues (**2a**, **2b**, **4a**, **4b**, **4c**) a nitrogen atom has been incorporated at the purine 8-position in place of a carbon atom. That makes the analogue-pair **1a** and **2a** to be of exactly the same structure except at the purine 8-position where the first analogue of the pair contains a carbon atom and the other one a nitrogen atom. Similar is the case for **1b** and **2b** and also **3** and **4a**. The ground state geometry of all eight adenine analogues is optimized at the B3LYP/6-31++G(d,p) level of theory in the gas phase and their fully optimized structures are shown in Fig. 1. Methanol is used as a solvent in accordance with the experiments. The optimized structure of the all the analogues shows a weak intramolecular hydrogen bonding interaction present between a proton of the exocyclic amine of the purine ring and one of the triazole nitrogens (see Table S1 in the ESI†). As a result, the adenine (purine) and triazole rings are found to be nearly coplanar in all the analogues.

3.2 Watson–Crick base pairing

In duplex DNA, adenine forms a Watson–Crick (WC) base pair with thymine through two H-bonds (one N–H···N and one N–H···O). To compare the formation and stability of the WC base pair formed by these analogue nucleobases, we first perform molecular electrostatic potential (MEP) calculation (see Fig. S1, ESI†) on geometry optimized structures of each analogue to assess qualitatively how the electronic distribution in the adenine ring is affected by the triazole modification; then, we determine intrinsic stability of the ^{t}A–T base pairs. The MEP on individual analogues is visualized through mapping their values onto a surface (of constant electron density) which eventually reflects the molecular boundaries. Comparison of calculated MEP on each analogue with its natural counterpart adenine ascertains the fact that triazole substitution at the C7-position of adenine has a minimal effect on the electronic distribution on the purine moiety. Thus, on a purely electronic point of view, it seems to be reasonable to predict that these analogues are well-suited to form WC base pairs with thymine.

Unlike other unnatural base pairs (UBPs), these analogues indeed form H-bondings similar to the WC base pairing as found in natural nucleobase pair A–T.[57–59] The optimized structures for all ^{t}A–T base pairs are shown in Fig. 2. At the same level of theory, the H-bond donor–acceptor distances for N–H···O and N–H···N are found to be 2.935 Å and 2.878 Å, respectively, in a A–T pair. The N–H···O distance is found to be

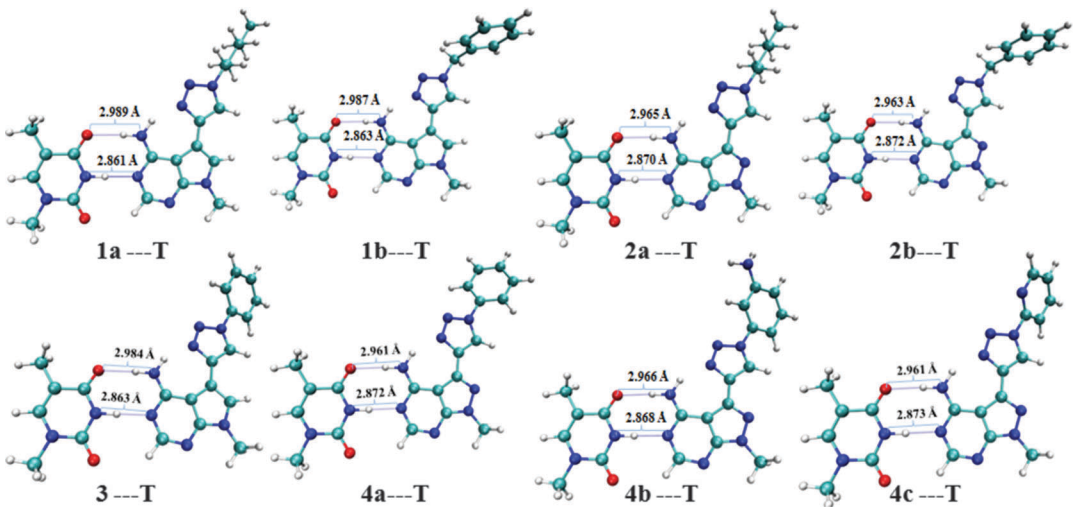

Fig. 2 Structures of the WC base-pairs between triazole analogues and thymine optimized at the B3LYP/6-31++G(d,p) level. Atom color code: blue (N), cyan (C), white (H).

Table 1 Calculated base pairing energies (in kcal mol^{-1}) and their BSSE corrected values for adenine analogue–thymine base pairs with various exchange and correlation energy functionals

	ΔE_{Pair}			ΔE_{Pair} + BSSE			ΔE_{Pair} + BSSE + ZPE		
Systems	B3LYP	M062X	ωB97XD	B3LYP	M062X	ωB97XD	B3LYP	M062X	ωB97XD
A–T	−12.71	−15.03	−16.27	−11.99	−14.22	−15.54	−10.74	−13.20	−14.40
1a–T	−11.77	−14.30	−15.61	−11.03	−13.50	−14.85	−10.06	−13.34	−13.71
1b–T	−11.76	−14.31	−15.51	−11.05	−13.51	−14.77	−10.09	−12.98	−13.73
2a–T	−12.07	−14.53	−15.89	−11.33	−13.69	−14.89	−10.31	−13.62	−14.64
2b–T	−12.10	−14.60	−15.86	−11.35	−13.75	−15.09	−10.39	−13.33	−14.04
3–T	−11.82	−14.37	−15.60	−11.09	−13.54	−14.84	−10.11	−13.33	−14.00
4a–T	−12.14	−14.66	−15.84	−11.40	−13.81	−15.06	−10.43	−13.54	−14.18
4b–T	−12.14	−14.63	−15.84	−11.39	−13.78	−15.06	−10.36	−13.76	−14.21
4c–T	−12.03	−14.40	−15.68	−11.28	−13.55	−14.90	−10.31	−13.18	−14.06

slightly increased while the N–H···N distance remains more or less the same in all ^{t}A–T pairs. This small change in N–H···O distances can be qualitatively understood to a great extent by the fact that C7-triazole modifications in these analogues would most likely be accommodated in the major groove.

Thermodymanic stability of a DNA oligonucleotide sequence depends on various factors and among those, intrinsic stability of base pairing interaction certainly plays an important role. In this respect, we compute pairing energies (ΔE_{Pair}) of all adenine analogues with natural thymine and compare the strength of pairing interaction against natural A–T base pairs. Pairing energies are also calculated at the B3LYP/6-31++G(d,p) level of theory in the gas phase using the following equation:

$$\Delta E_{Pair} = \Delta E_{BP} - \Delta E_{t_{A/A}} - \Delta E_{T} \quad (1)$$

where, ΔE_{BP}, $\Delta E_{t_{A/A}}$ and ΔE_{T} are optimized total energy of the corresponding WC base pair, individual analogue (adenine) and natural thymine, respectively. The pairing energy for A–T base pairs is calculated to be −12.71 kcal mol^{-1} while that for **1a–T**, **1b–T**, **2a–T**, **2b–T**, **3–T**, **4a–T**, **4b–T** and **4c–T** are found to be −11.77 kcal mol^{-1}, −11.76 kcal mol^{-1}, −12.07 kcal mol^{-1}, −12.10 kcal mol^{-1}, −11.82 kcal mol^{-1}, −12.14 kcal mol^{-1}, −12.14 kcal mol^{-1} and −12.03 kcal mol^{-1}, respectively. These results clearly suggest that the stability of a WC base pair formed by these analogues and thymine is very close to that of a natural A–T base pair. However, a small decrease in pairing energy in ^{t}A–T base pairs is consistent with the changes in the separation distances in these base pairs compared to A–T. At this point, we use two different energy functionals such as ωB97XD and M062X, with the same basis set as before, for computing the pairing energies of ^{t}A–T pairs mainly to take into account weak non-covalent interactions in these base-pairs. We have found that the pairing energies increase (3–4 kcal mol^{-1}) with both functionals suggesting their ability to include medium and long range correlations with respect to B3LYP. The pairing energies are further corrected for basis set superposition error (BSSE)[60] and ZPE at the same level of theory and we see that these corrections do not significantly alter the numerical values. In short, the overall findings remain unchanged. Table 1 contains all the BSSE and ZPE corrected ΔE_{Pair} values calculated with different functionals for all ^{t}A–T pairs.

3.3 Photophysical properties

In this section, we focus on the photophysical properties of all adenine analogues. Vertical excitation energies and oscillator strengths are calculated from time-dependent DFT (TDDFT) calculations using B3LYP exchange and energy correlation functional. Optical absorption spectra are calculated in the presence of a methanol solvent using PCM. The states corresponding to the lowest energy transitions in these analogues are assigned after detailed MO analyses. Emission spectra are

Table 2 Three lowest vertical excitation energies (E), oscillator strengths (f) and state assignments (with % contribution) of adenine and triazole adenine analogues in methanol

Analogues	E (eV)	f	Assignment
A	4.90	0.267	H → L (88%)
	5.10	0.002	H−1 → L (95%)
	5.16	0.053	H → L+1 (70%)
1a	4.10	0.079	H → L (90%)
	4.46	0.190	H → L+1 (87%)
	4.76	0.149	H → L+2 (87%)
1b	4.05	0.049	H → L (92%)
	4.32	0.107	H → L+1 (95%)
	4.47	0.066	H → L+2 (92%)
2a	4.26	0.219	H → L (92%)
	4.79	0.067	H−1 → L (77%)
	4.82	0.082	H → L+1 (79%)
2b	4.25	0.213	H → L (92%)
	4.72	0.065	H → L+1 (86%)
	4.79	0.043	H−1 → L (65%)
3	3.61	0.055	H → L (98%)
	4.30	0.169	H → L+1 (92%)
	4.35	0.044	H → L+2 (92%)
4a	4.05	0.174	H → L (98%)
	4.43	0.166	H → L+1 (87%)
	4.68	0.234	H−1 → L (87%)
4b	3.92	0.092	H → L (90%)
	4.10	0.155	H−1 → L (92%)
	4.32	0.009	H → L+1 (85%)
4c	3.92	0.105	H → L (98%)
	4.35	0.143	H → L+1 (92%)
	4.46	0.086	H → L+2 (85%)

calculated with optimized first excited geometry (S_1) of each nucleobase analogues in methanol.

We have calculated the absorption spectra of adenine in the same level of theory to make a direct comparison between the analogues and the parent nucleobase. Adenine shows three lowest energy transitions as π–π* (4.90 eV), n–π* (5.10 eV) and π–π* (5.16 eV) in methanol while for the analogues, all the three lowest energy transitions are found to be of π–π* character. Vertical excitation energies, oscillator strengths and state assignments for three lowest singlet transitions of adenine and triazole analogues are shown in Table 2. Relevant molecular orbitals dominating these excitations are depicted in Fig. 3 (for adenine see Fig. S2, ESI†).

At this point, for the purpose of clear representation, eight analogues are divided into two groups: (A) analogues containing aliphatic substituents (**1a**, **1b**, **2a**, **2b**) and (B) analogues containing aromatic substituents (**3**, **4a**, **4b**, **4c**). First, we concentrate on the optical properties of analogues containing aliphatic substituents. **1a** and **1b** show similar absorptive properties to their sister analogues, **2a** and **2b**. Basically in both of these pairs, effective chromophoric length remains more or less the same since the π-conjugation cannot extend beyond the triazole moiety. So, qualitatively, their optical properties are expected to be little affected by the non-conjugated *n*-pentyl or benzyl groups. Our calculations predict that in **1a** (**1b**), three lowest energy transitions are located near 4.10 (4.06), 4.46 (4.32) and 4.57 (4.47) eV with the strongest oscillator strength for the second one, which is mainly dominated by the configurational transition of the HOMO to LUMO+1. The first π–π* transition is dominated by the HOMO to LUMO transition, while the third transition is dominated by the configurations HOMO and LUMO+2. Therefore, the experimentally observed peak for **1a** (**1b**) at 282 (281) nm corresponds to the second π–π* transition (in methanol). We find that for **2a** (**2b**) three lowest energy transitions are located near 4.27 (4.26), 4.79 (4.72), 4.82 (4.79) eV. The lowest energy π–π* transition, dominated by HOMO to LUMO transition, shows the highest oscillator strength and this peak can be related to the experimentally observed peak at 290 nm for both analogues **2a** and **2b**. Other two π–π* states in **2a** and **2b** are dominated by HOMO−1 to LUMO and HOMO to LUMO+1 configurational transitions.

Adenine analogue **3** shows peaks at 3.61, 4.30 and 4.34 eV with the strongest intensity for the second one, which is mainly dominated by configurations HOMO and LUMO+1. The first π–π* transition is dominated by the configuration HOMO to LUMO while the third transition is dominated by the HOMO and LUMO+2 configurations. Therefore, the experimentally observed peak for **3** at 280 nm corresponds to the second π–π* transition. For **4a** (**4b**, **4c**) three lowest energy transitions are located near 4.05 (3.92, 3.92), 4.43 (4.10, 4.35) and 4.68 (4.31, 4.46) eV. In all three analogues the experimentally observed lowest energy peaks can be related to the first π–π*

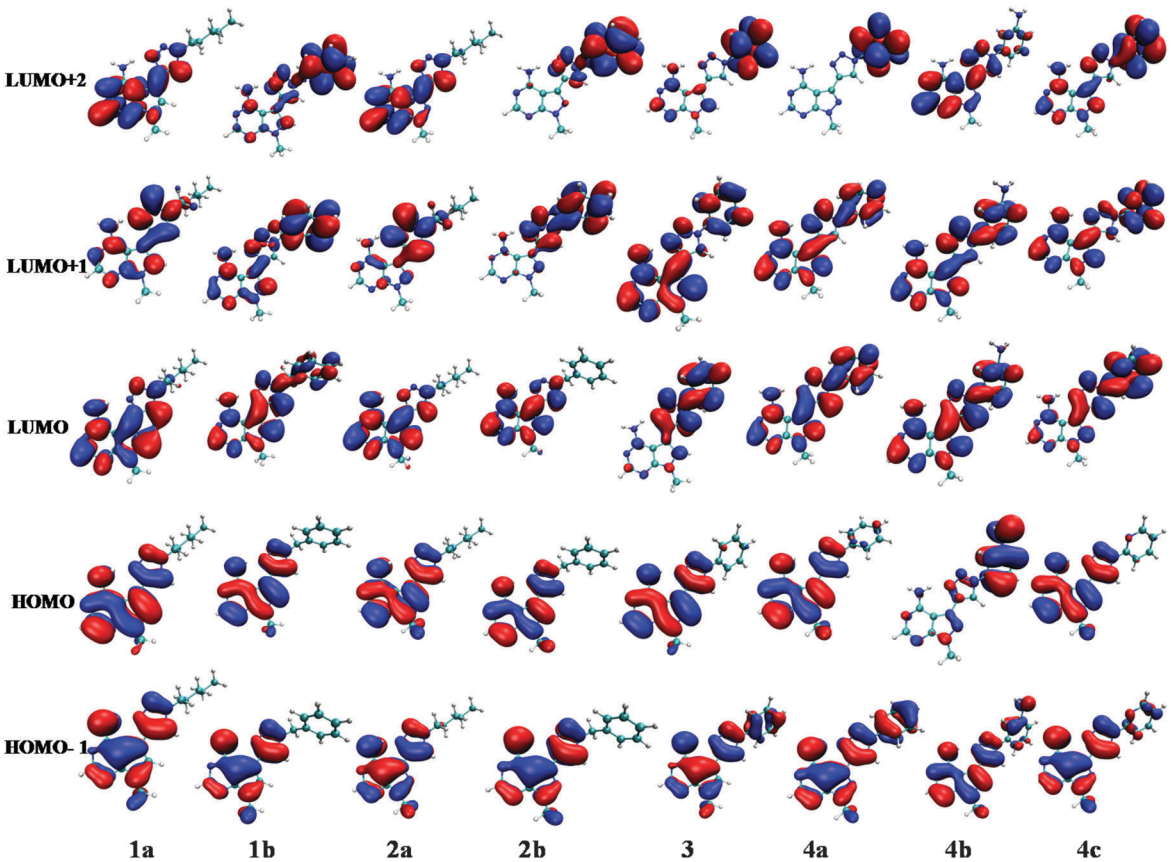

Fig. 3 Calculated frontier molecular orbitals of the triazole adenine analogues.

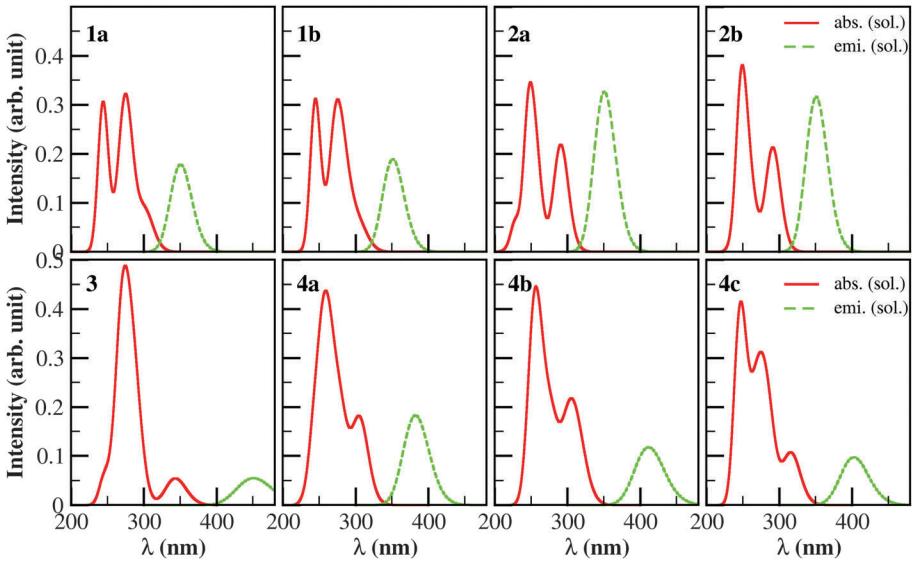

Fig. 4 Absorption and emission spectra of triazole adenine analogues.

transition which is dominated by transition from HOMO to LUMO configurations. Other two π–π* transitions in **4a**, **4b** are dominated by configurations HOMO−1 to LUMO and HOMO to LUMO+1.

Overall, triazole adenine analogues absorb light at lower energy corresponding to natural adenine and therefore allows for selective excitation. Basically, the analogues differ from adenine in three aspects; imidazole ring modifications, introduction of the triazole moiety and further substitutions. In order to separate out the effect of these three contributions, we have performed TDDFT calculations in a series of model structures (see Fig. S3, ESI†). Our calculations show that the introduction of a triazole ring causes a red shift in the absorption spectrum compared to parent backbones (Fig. S4, ESI†) due to increased conjugation. Attachment of non-conjugating aliphatic substituents (analogues **1a**, **1b**, **2a** and **2b**) results in no further red shift in optical absorption, whereas the same for aromatic substituents (conjugating with triazole ring, analogues **3**, **4a–c**) results in a further red shift in their absorption spectra (Fig. 4).

As we have already mentioned, emission spectra of all adenine analogues are calculated in methanol and therefore, these results are directly comparable to experimentally observed spectra. Emission peaks for analogues in methanol are found to be at 350 nm, 351 nm, 354 nm, 382 nm, 411 nm and 402 nm for **1a**, **2a**, **2b**, **4a**, **4b**, **4b** and **4c**, respectively. These peaks are in good agreement with experimentally observed emission peaks at 369 nm, 343 nm, 340 nm, 345 nm, 468 nm, 359 nm for **1a**, **2a**, **2b**, **4a**, **4b** and **4c**, respectively, in the same solvent. Oscillator strength of emission maxima in **3** is found to be very small which agrees well with experiment. Analogue **4b** shows a Stokes shift of 96 nm which is maximum among the analogues considered. Emission peak energies and corresponding oscillator strength for all the analogue systems are listed in Table 3.

Table 3 Emission peaks of the adenine analogues

Analogues	E_{flu} (eV)	Osc. str.	τ (ns)
1a	3.54	0.178	10.31
1b	3.53	0.189	9.78[61]
2a	3.54	0.327	5.64
2b	3.50	0.316	5.96
3	2.75	0.055	55.41
4a	3.24	0.183	11.98
4b	3.01	0.117	21.60
4c	3.08	0.097	24.99

The fluorescence efficiency of these nucleobase analogues can be described in terms of their radiative lifetime (τ). We use emission (fluorescent) energy (E_{flu}) and oscillator strength (f) to calculate τ for spontaneous emission by using Einstein transition probabilities according to the formula (in a.u.).[62,63]

$$\tau = \frac{c^3}{2E_{flu}^2 f} \quad (2)$$

where, c is the velocity of light. Longer lifetime refers to poor light-emitting efficiency while shorter lifetime implies enhanced fluorescence quantum yield. The computed lifetime τ in this family of analogues follows the trend **2a** < **2b** < **1b** < **1a** < **4a** < **4b** < **4c** < **3** which is also found to be the sequence of decreasing oscillator strength, f. Hence, oscillator strength surely plays a crucial role in determining the radiative lifetime.

Experimentally, **2a** and **2b** show a ten-fold increment in quantum yields compared to their sister analogues, **1a** and **1b**. We also find that **2a** and **2b** have a relatively short lifetime and comparatively a larger oscillator strength than other analogues. Although, the definition of quantum yield involves the presence or the absence of various deactivation channels which are non-radiative in nature, our preliminary results point towards that. In fact, our fluorescence data for these two analogues suggest that the excitons are highly localized (in some parts of the systems in

all the MOs involved in optical processes, see Fig. 3) and so they would recombine faster and efficiently.

4 Conclusions

Base paring properties of a newly designed series of 7-substituted triazole adenine analogues have been investigated by detailed quantum mechanical *ab initio* calculations. All the adenine analogues retain Watson–Crick base pairing ability with thymine and the stability of these base pairs is comparable to the natural Watson–Crick base pair, A–T. We have calculated the optical properties of these analogues and found that all the analogues absorb light at lower energy corresponding to π–π* transitions in comparison to their natural counterpart, adenine. Contrary to adenine, these triazole analogues are found to fluoresce in the UV-Visible region. Our calculated absorption and emission maxima follow the experimental trends fairly well, measured in a solvent, methanol. Moreover, we have found that two of these analogues, **2a** and **2b**, are the most promising members in this family due to their high fluorescence efficiency. It would be interesting to monitor these two analogues inside the DNA duplex environment. The main point is that our findings on their interesting photophysical properties clearly suggest that they would be potential candidates for probing the nucleic acid structure, dynamics and recognition.

Acknowledgements

SD thanks Council of Scientific and Industrial Research (CSIR), New Delhi for Junior Research Fellowship (JRF). SKP acknowledges research support from DST, Government of India.

References

1 A. J. Cobb, *Org. Biomol. Chem.*, 2007, **5**, 3260–3275.
2 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500.
3 E. Lescrinier, R. Esnouf, J. Schraml, R. Busson, H. Heus, C. Hilbers and P. Herdewijn, *Chem. Biol.*, 2000, **7**, 719–731.
4 K.-U. Schöning, P. Scholz, S. Guntha, X. Wu, R. Krishnamurthy and A. Eschenmoser, *Science*, 2000, **290**, 1347–1351.
5 J. C. Chaput and J. W. Szostak, *J. Am. Chem. Soc.*, 2003, **125**, 9274–9275.
6 S. Sinha, P. H. Kim and C. Switzer, *J. Am. Chem. Soc.*, 2004, **126**, 40–41.
7 K.-H. Jung and A. Marx, *Cell. Mol. Life Sci.*, 2005, **62**, 2080–2091.
8 A. K. Jissy and A. Datta, *J. Phys. Chem. Lett.*, 2013, **5**, 154–166.
9 J. A. Piccirilli, T. Krauch, S. E. Moroney and S. A. Benner, *Nature*, 1990, **343**, 33–37.
10 T. J. Matray and E. T. Kool, *Nature*, 1999, **399**, 704–708.
11 H. Weizman and Y. Tor, *J. Am. Chem. Soc.*, 2001, **123**, 3375–3376.
12 I. Hirao, Y. Harada, M. Kimoto, T. Mitsui, T. Fujiwara and S. Yokoyama, *J. Am. Chem. Soc.*, 2004, **126**, 13298–13305.
13 I. Hirao, T. Mitsui, M. Kimoto and S. Yokoyama, *J. Am. Chem. Soc.*, 2007, **129**, 15549–15555.
14 Y. Doi, J. Chiba, T. Morikawa and M. Inouye, *J. Am. Chem. Soc.*, 2008, **130**, 8762–8768.
15 E. T. Kool, J. C. Morales and K. M. Guckian, *Angew. Chem., Int. Ed.*, 2000, **39**, 990–1009.
16 S. A. Benner, S. Hoshika, M. Sukeda, D. Hutter, N. Leal, Z. Yang and F. Chen, *Nucleic Acids Symp. Ser.*, 2008, 243–244.
17 D. Onidas, D. Markovitsi, S. Marguet, A. Sharonov and T. Gustavsson, *J. Phys. Chem. B*, 2002, **106**, 11367–11374.
18 N. J. Greco and Y. Tor, *J. Am. Chem. Soc.*, 2005, **127**, 10784–10785.
19 P. Sandin, L. M. Wilhelmsson, P. Lincoln, V. E. Powers, T. Brown and B. Albinsson, *Nucleic Acids Res.*, 2005, **33**, 5019–5025.
20 P. Sandin, K. Börjesson, H. Li, J. Mårtensson, T. Brown, L. M. Wilhelmsson and B. Albinsson, *Nucleic Acids Res.*, 2008, **36**, 157–167.
21 A. Okamoto, Y. Saito and I. Saito, *J. Photochem. Photobiol., C*, 2005, **6**, 108–122.
22 D. A. Berry, K.-Y. Jung, D. S. Wise, A. D. Sercel, W. H. Pearson, H. Mackie, J. B. Randolph and R. L. Somers, *Tetrahedron Lett.*, 2004, **45**, 2457–2461.
23 M. E. Hawkins, *Cell Biochem. Biophys.*, 2001, **34**, 257–281.
24 D. Ward, E. Reich and L. Stryer, *J. Biol. Chem.*, 1969, **244**, 1228–1237.
25 R. W. Sinkeldam, N. J. Greco and Y. Tor, *Chem. Rev.*, 2010, **110**, 2579–2619.
26 L. M. Wilhelmsson, *Q. Rev. Biophys.*, 2010, **43**, 159–183.
27 S. K. Yang, X. Shi, S. Park, T. Ha and S. C. Zimmerman, *Nat. Chem.*, 2013, **5**, 692–697.
28 S. Preus and L. M. Wilhelmsson, *ChemBioChem*, 2012, **13**, 1990–2001.
29 A. Vallée-Bélisle and S. W. Michnick, *Nat. Struct. Mol. Biol.*, 2012, **19**, 731–736.
30 H. Liu, Q. Song, Y. Yang, Y. Li and H. Wang, *J. Mol. Model.*, 2014, **20**, 2100.
31 Y. P. Yurenko, R. O. Zhurakivsky, M. Ghomi, S. P. Samijlenko and D. M. Hovorun, *J. Phys. Chem. B*, 2007, **111**, 6263–6271.
32 G. Brancolini and R. Di Felice, *J. Chem. Phys.*, 2011, **134**, 205102.
33 P. K. Samanta and S. K. Pati, *Phys. Chem. Chem. Phys.*, 2015, **17**, 10053–10058.
34 P. K. Samanta, A. K. Manna and S. K. Pati, *J. Phys. Chem. B*, 2012, **116**, 7618–7626.
35 M. Gedik and A. Brown, *J. Photochem. Photobiol., A*, 2013, **259**, 25–32.
36 P. K. Samanta and S. K. Pati, *New J. Chem.*, 2013, **37**, 3640–3646.
37 D. Varsano, A. Garbesi and R. Di Felice, *J. Phys. Chem. B*, 2007, **111**, 14012–14021.
38 L. Zhang, H. Li, X. Chen, R. I. Cukier and Y. Bu, *J. Phys. Chem. B*, 2009, **113**, 1173–1181.
39 S. Rai, S. Ranjan, H. Singh and U. D. Priyakumar, *RSC Adv.*, 2014, **4**, 29642–29651.

40 C. E. Crespo-Hernández, B. Cohen, P. M. Hare and B. Kohler, *Chem. Rev.*, 2004, **104**, 1977–2020.
41 M. Barbatti, Z. Lan, R. Crespo-Otero, J. J. Szymczak, H. Lischka and W. Thiel, *J. Chem. Phys.*, 2012, **137**, 22A503.
42 F. Plasser, R. Crespo-Otero, M. Pederzoli, J. Pittner, H. Lischka and M. Barbatti, *J. Chem. Theory Comput.*, 2014, **10**, 1395–1405.
43 C. Dyrager, K. Börjesson, P. Dinér, A. Elf, B. Albinsson, L. M. Wilhelmsson and M. Grøtli, *Eur. J. Org. Chem.*, 2009, 1515–1521.
44 A. Dierckx, P. Dinér, A. H. El-Sagheer, J. D. Kumar, T. Brown, M. Grøtli and L. M. Wilhelmsson, *Nucleic Acids Res.*, 2011, 1–12.
45 F. Seela, M. Zulauf, M. Sauer and M. Deimel, *Helv. Chim. Acta*, 2000, **83**, 910–927.
46 F. Seela and M. Zulauf, *J. Chem. Soc., Perkin Trans. 1*, 1999, 479–488.
47 F. Seela and H. Thomas, *Helv. Chim. Acta*, 1995, **78**, 94–108.
48 C. P. Lawson, A. Dierckx, F.-A. Miannay, E. Wellner, L. M. Wilhelmsson and M. Grøtli, *Org. Biomol. Chem.*, 2014, **12**, 5158–5167.
49 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. 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, Ö. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *Gaussian 09 Revision A.01*, Gaussian Inc., Wallingford CT, 2009.
50 A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
51 C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1988, **37**, 785–789.
52 B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Lett.*, 1989, **157**, 200–206.
53 Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215–241.
54 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.
55 J.-D. Chai and M. Head-Gordon, *J. Chem. Phys.*, 2008, **128**, 084106.
56 G. Scalmani and M. J. Frisch, *J. Chem. Phys.*, 2010, **132**, 114110.
57 D. A. Malyshev, K. Dhami, T. Lavergne, T. Chen, N. Dai, J. M. Foster, I. R. Corrêa and F. E. Romesberg, *Nature*, 2014, **509**, 385–388.
58 K. Dhami, D. A. Malyshev, P. Ordoukhanian, T. Kubelka, M. Hocek and F. E. Romesberg, *Nucleic Acids Res.*, 2014, **42**, 10235–10244.
59 S. Jahiruddin and A. Datta, *J. Phys. Chem. B*, 2015, **119**, 5839–5845.
60 S. F. Boys and F. Bernardi, *Mol. Phys.*, 2002, **100**, 65–73.
61 The analogue **1b** shows emission at 425 nm with a very poor oscillator strength (0.0025) in methanol which is quite surprising because we expect a similar emissive behaviour as structurally close analogue **1a**. To check this, we have reoptimized the first excited state geometry of **1b** in CIS/6-31++g(d,p) level of theory and calculated the vertical transition energy using TDB3LYP method. The emission energy is calculated to be 350 nm with a oscillator strength of 0.189. This result of **1b** is very much comparable with **1a**.
62 R. C. Hilborn, 2002, arXiv preprint physics/0202029.
63 V. Lukeš, A. Aquino and H. Lischka, *J. Phys. Chem. A*, 2005, **109**, 10232–10238.

