ROYAL SOCIETY OF CHEMISTRY

PAPER

New insights into the structure and dynamics of the epigenetic modifications on DNA

Dineshbabu Takkella, [a] Javier Cerezo, [b] Lara Martinez-Fernandez [*c] and Krishna Gavvala [*a]

DNA methylation is a key epigenetic modification involved in genomic imprinting, X-chromosome inactivation, and repression of repetitive element transcription and transposition. Despite its biological significance, the impact of epigenetic modifications such as methylcytosine (mC) and hydroxymethylcytosine (hmC) on the structural and UV-induced dynamics of DNA remains poorly understood. Here, we employed the fluorescent nucleobase analogue 2-aminopurine (2Ap) in combination with steady-state and time-resolved spectroscopy, molecular dynamics, and quantum mechanical calculations to investigate these effects. Our findings reveal distinct differences in base stacking and helical stability between mC and hmC-modified DNA. mC-modified DNA predominantly adopts a stacked conformation, promoting efficient fluorescence quenching of 2Ap. In contrast, hmC-modified DNA displays both stacked and non-stacked conformations, leading to reduced base stacking and a more hydrophobic local environment, as indicated by blue-shifted emission spectra. Furthermore, although charge-transfer quenching occurs in all systems, hmC shows weaker charge-transfer character, resulting in higher fluorescence quantum yields and longer lifetimes. These results highlight the subtle but crucial role of hmC in modulating local DNA conformation and stability. Moreover, they demonstrate the effectiveness of 2Ap as a sensitive probe for detecting epigenetic modifications, offering deeper insights into the molecular mechanisms of DNA methylation and demethylation pathways.

1. Introduction

Genome is the major structural component of the cell that provides all the genetic information required for the growth and function of an organism. Modifications on the genome without altering the sequence are usually known as epigenetic modifications[1,2] and are essential for regulating cellular functions.[1–3] Therefore, the cell fate relies largely on the epigenetic Scheme 1. Chemical structures of cytosine and its epigenetic modifications on DNA induced by DNMT1 (PDB ID: 3PTA) and TET (PDB ID: 7NE3) enzymes mechanisms that control gene expression, which is critical for various biological and pathological processes.[1,2] DNA methylation, one of the key epigenetic modifications, involves the transfer of a methyl group from the cofactor S-adenosyl-L-methionine (SAM) to spec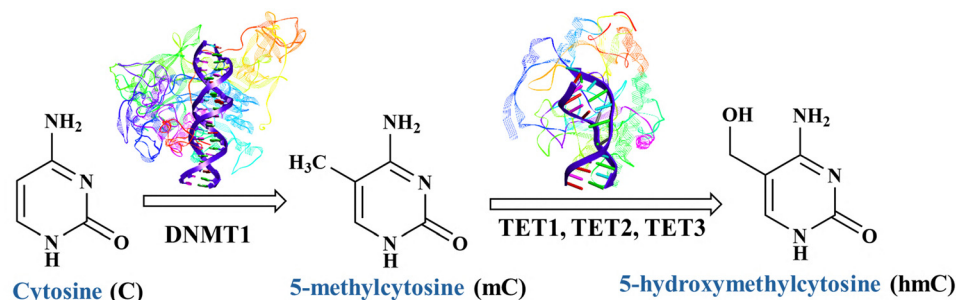ific sites in the DNA.[4–6] In eukaryotic cells, methylation takes place specifically on the C5 position of cytosine, predominantly within CpG islands of DNA.[7–9] Methylated cytosine (mC)[10,11] is often referred to as the fifth base of human DNA, accounting for roughly 1% of the bases in mammalian genomes,[12] and has a role in a plethora of biological processes, either functional or pathological.[6,13–17] This heritable DNA modification on cytosine base is catalysed by enzymes known as DNA methyltransferases (DNMTs). Subsequently, TET (ten-eleven-translocation) enzymes oxidizes 5-methylcytosine (mC) to 5-hydroxymethylcytosine (hmC) and further to 5-formyl cytosine (fC) and 5-carboxyl cytosine (caC) as shown in Scheme 1.[4,5] Though the oxidation mechanism of mC is well established, and their epigenetic role has been studied,[18–25] there is still room to understand their physical and conformational effects on DNA flexibility and its ability to wrap around histones to form nucleosomes.[26] Beyond this structural impact on DNA, the presence of these epigenetic bases, especially mC, has been connected with increase UV-induced damage,[13,14,17,27–29] pointing also to effects into DNA photoinduced dynamics. To fully unravel these sequence-dependent effects within the nearest-neighbour framework, we explored these issues using a variety of experimental and theoretical approaches.

In detail, we first examined the structure (using thermal melting, circular dichroism, and molecular dynamic (MD)

a Department of Chemistry, Indian Institute of Technology Hyderabad, Kandi, Sangareddy, Telangana-502284, India. E-mail: kgavvala@chy.iith.ac.in

b Departamento de Química, Facultad de Ciencias and Institute for Advanced Research in Chemistry (IADCHEM), Universidad Autónoma de Madrid, Campus de Excelencia UAM-CSIC, Cantoblanco, 28049 Madrid, Spain
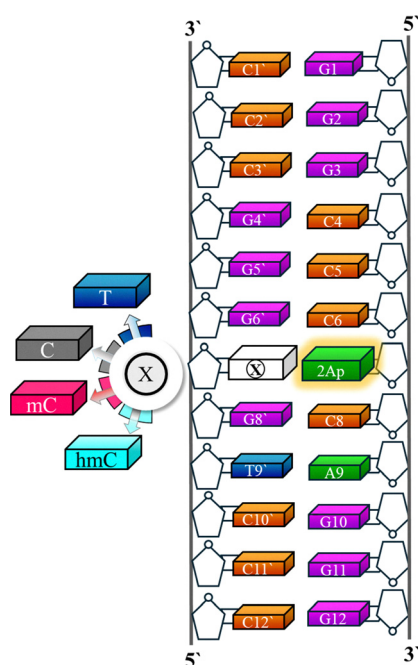
c Departamento de Química Física de Materiales, Instituto de Química Física Blas Cabrera, Consejo Superior de Investigaciones Científicas, IQF-CSIC, Calle Serrano 119, 28006 Madrid, Spain. E-mail: lmartinez@iqf.csic.es

© 2025 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2025, **6**, 1927–1940 | 1927

Scheme 1 Chemical structures of cytosine and its epigenetic modifications on DNA induced by DNMT1 (PDB ID: 3PTA) and TET (PDB ID: 7NE3) enzymes.

simulations) and dynamics of 12 bp (base pairs) duplexes containing epigenetic bases and labeled with 2-aminopurine (2Ap). As a model DNA duplex, we use a 12-bp duplex in which the A7 residue is replaced by 2Ap (Fig. 1). The ssDNA sequence is 5′-GGGCCC**2Ap**CAGGG-3′ (C**2Ap**C). In the corresponding dsDNA duplexes, the WC base pair of 2Ap is either C, T, mC, or hmC, denoted as C2ApC/GCG, C2ApC/GTG, C2ApC/GmCG, and C2ApC/GhmCG, respectively. In native CpG islands, cytosine (including mC and hmC) pairs with guanine; however, in this work, we deliberately positioned 2Ap opposite the modified cytosine. 2Ap's well-established ability to form wobble-like yet stable base pairs with cytosine.[30] Although structurally distinct from guanine, 2Ap preserves local duplex stability and exhibits high sensitivity to microenvironmental changes.[31–33] Embedding 2Ap within a GC-rich sequence that retains CpG-like structural features allowed us to probe the local structural and photophysical effects of mC/hmC modifications in a physiologically relevant context. Then, to enhance the

understanding of the photophysics of 2Ap-labeled DNA duplexes and their dependence on DNA sequence context, we investigate the UV absorption, emission, and time-resolved fluorescence, systematically varying the composition of bases opposite 2Ap. 2Ap has been selected as a fluorescent analogue since it can be incorporated into DNA, replacing the natural base without significantly disrupting DNA structure or its interactions with proteins.[34–36] Furthermore, 2Ap is highly responsive to its microenvironment, especially to stacking interactions, offering insights into its local environment,[31–33] which is of particular importance in our study. In this setup, we maintained a constant composition of the neighbouring bases (cytosine) and varied the opposite bases of 2Ap. The experimental data were complemented with MD simulations, which provide atomistic insight into the structural effects induced by 2Ap, and quantum mechanical (QM) calculations to elucidate the key photophysical processes (excited state nature and decays) in each duplex. Our integrated analysis offers an explanation for the observed spectroscopic properties and their dependence on 2Ap's neighbouring or opposite nucleobases, clarifying the various factors that influence the photophysical characteristics of 2Ap. These dynamic interaction properties could offer detailed information about the effectiveness and continuity of epigenetic replication and would be valuable for gaining further insights into conformational changes or protein binding events by analysing observed photophysical changes in relation to structural and electronic effects. Altogether, these discoveries uncovered a previously elusive mechanism of active DNA demethylation, ultimately leading to a wave of research in mammalian epigenetic regulation.

2. Materials and methods

2.1. DNA Sequences

All labeled and non-labeled DNA sequences were purchased from Integrated DNA Technologies, Inc. (U.S.A.). The 12-bp duplex oligonucleotide (ODN) sequence used here was 5′-GGGCCC**2Ap**CAGGG-3′/5′-CCCTG**X**GGGCCC3′ where X represents a modification site. In this study, the sequence has either matched (X = thymine, T) or mismatched base pairs (X= C, mC, and hmC) at the X position. The hemimethylated modifications included the incorporation of cytosine (C), methyl cytosine

Fig. 1 Schematic representation of sequences of the 2Ap-labeled DNA duplexes and base labeling used in this study.

(mC), or hydroxymethyl cytosine (hmC) at the modification site. 2Ap-labeled dsDNA duplexes were prepared by heating 30 μM of 2Ap-labeled ssDNA sequence and 30 μM complementary matched or mismatched ssDNA sequences at 90 °C for 5 minutes, followed by slow cooling to room temperature. All experiments were conducted in a buffer (pH 7.5) consisting of 25 mM TRIS-HCl with 150 mM NaCl.

2.2. Absorption spectroscopy

UV-Vis absorption spectra were recorded on the JASCO V-730 spectrophotometer. The concentration of single-stranded sequences was governed using extinction coefficients at 260 nm. The extinction coefficients for the 2Ap-labeled sequence 5′-GGGCCC**2Ap**CAGGG-3′ was 104 000 cm^{-1} M^{-1}. Extinction coefficients for the nonlabeled single stranded sequences 5′-CCCTG**T**GGGCCC3′, 5′-CCCTG**C**GGGCCC3′, 5′-CCCTG**mC**GGGCCC3′ and 5′-CCCTG**hmC**GGGCCC3′ were 99 400 cm^{-1} M^{-1}, 97 300 cm^{-1} M^{-1}, 95 600 cm^{-1} M^{-1}, and 97 300 cm^{-1} M^{-1}, respectively.

2.3. Fluorescence spectroscopy

The steady-state fluorescence spectra were recorded on the JASCO FP-8350 spectrofluorometer. For steady-state fluorescence studies, matched and mismatched 2Ap-labeled duplexes were excited at wavelengths of 290 nm and 315 nm. The excitation wavelength of 315 nm was selected to specifically excite the 2Ap. The measurements were supervised on 5 μM ODN duplexes. The apparent relative quantum yields (QYs) of sequence-dependent DNA in Tris-HCl buffer were determined using alone 2Ap as a standard ($\Phi = 0.68$).[37] Relative QYs were calculated from eqn (1).

$$\frac{\Phi_f}{\Phi_r} = \frac{I_f}{I_r} \times \frac{A_r}{A_f} \times \frac{n_f^2}{n_r^2} \qquad (1)$$

The subscripts f and r denote the sample and reference molecules, respectively. Here, Φ denotes the quantum yield, I represents the integrated area under the fluorescence emission spectrum, A denotes the absorbance at 315 nm, and n values correspond to the refractive index of the solvents used.

The melting curves of the duplexes (2 μM) in 25 mM TRIS-HCl buffer (pH = 7.5) with 150 mM NaCl were measured by monitoring the temperature-dependent fluorescence changes at 370 nm using a Horiba Fluoromax-4 spectrofluorometer equipped with a Peltier system controlled by a Thermo Scientific NESLAB RTE-7. The temperature was escalated from 15 to 95 °C at a heating rate of 0.5 °C min^{-1}. Melting temperatures were obtained from the melting curves following the method outlined in previous literature.[38]

2.4. Time-resolved fluorescence spectroscopy

Fluorescence lifetime was recorded using TCSPC (time-correlated single photon counting) setup by Horiba (Fluoro-Hub). Fluorescence decays were measured at 291 nm (NanoLED) excitation wavelength with an IRF (Instrumental response function) of <1 ns. Fluorescence decays were processed using Horiba's DAS6 package by convolving the IRF with

exponential decay functions. The fitting was performed following the mathematical model defined in eqn (2).[39]

$$I(t) = \sum \alpha_i e^{\left(\frac{-t}{\tau_i}\right)} \qquad (2)$$

where $I(t)$ signifies the exponential fluorescence decay, τ_i denotes fluorescence lifetime, and α_i represents the fractional contribution of the ith component. In TCSPC experiments, the average error was $\leq 5\%$.

A global analysis method was executed to analyze multi-wavelength decay data, enhancing the accuracy of decay component determination.[40,41] This approach assumes that the decay components τ_i are independent of the emission wavelength, while their amplitudes α_i change with wavelength. The DAS6 analysis tool incorporates an integrated global analysis fitting module. The DAS (Decay-associated spectra) were built based on eqn (3).

$$I_i(\lambda) = \frac{I(\lambda)\tau_i\alpha_i(\lambda)}{\sum\limits_i^n \tau_i\alpha_i(\lambda)} \qquad (3)$$

Here, $I(\lambda)$ represents the steady-state emission spectrum, while $\alpha_i(\lambda)$ denotes the wavelength-dependent amplitude. The DAS offers the spectra for each lifetime component, thereby revealing the intensity contribution of individual species to the total emission of the duplex. The real amplitudes corresponding to each lifetime species were calculated from $A_i = \frac{\int I_i(\lambda)\mathrm{d}\lambda}{\int I(\lambda)\mathrm{d}\lambda}$, here the numerator denotes the area under the DAS associated with species i, while the denominator represents the area under the steady-state emission spectrum.

2.5. Circular dichroism spectroscopy

The CD (circular dichroism) spectra were measured using a JASCO J-815 spectrophotometer. Each CD spectrum was the result of averaging three accumulated spectra of the sample and was recorded at a scanning speed of 100 nm min^{-1}. Each measurement's baseline was recorded using a 25 mM Tris-HCl buffer solution with 150 mM NaCl, at pH 7.5. The concentration of each dsDNA sequence used in the experiment was measured to be 5 μM.

2.6. Molecular dynamics simulations

A DNA double-strand with standard bases (GGGCCCGCAGGG) was generated using the DNA/RNA builder tool implemented in Avogadro.[42] The structure was subsequently modified with the same molecular editor to generate the non-standard versions, including 2Ap, mC, and hmC bases (Scheme S1). Interactions were treated with the AMBER-consistent parmbsc1 force field (FF),[43] which includes all standard nucleotides, as well as the solvent and counterions (Na$^+$). FF parameters for non-standard nucleotides (2Ap, mC, and hmC) were derived by selecting atom types within the FF by analogy and adding any missing bonded terms by similarity with those already implemented in the FF. Restricted electrostatic potential fitted (RESP) charges for non-standard residues were obtained using antechamber[44] at the

wB97XD[45]/cc-pVDZ level of theory, using Gaussian16.[46] The charges were adjusted to retain those of the sugar-phosphate backbone from standard AMBER while updating those in the nucleotide region, ensuring proper charge balance. This protocol was selected since its application to standard residues yields charges very similar to those included in the AMBER FF, validating the consistency of the derived parameters. The full parameter set is provided in Tables S1 and S2 of the SI. MD simulations were carried out using GROMACS 2021.4.[47] The DNA strands generated with Avogadro were solvated in a dodecahedral box, ensuring at least a 20 Å buffer between the strand and any box edge, followed by neutralization of the negative charges with sodium cations. The system was, then, energy-minimized to remove steric clashes and equilibrated under an NVT ensemble at 298.15 K for 500 ps using the Nosé–Hoover thermostat ($s = 1$ ps).[48,49] The production run was performed under an NPT ensemble at 298.15 K and 1 atm using the Nosé–Hoover thermostat ($\tau = 1$ ps) and the Parrinello–Rahman barostat ($\tau = 5$ ps)[50] for a total of 100 ns. Comparison with shorter trajectories (10 ns, shown in Fig. S1), confirm that the 100 ns trajectories are converged with respect to the structural parameters analysed. In all simulations, electrostatic interactions were evaluated using the PME algorithm,[51] with a real-part cutoff of 11 Å, while vdW interactions were treated with an 11 Å cutoff. A time step of 1 fs was employed, with constraints applied to all bonds involving hydrogen atoms with LINCS.[52] Additional simulations with a larger timestep of 2 fs (Fig. S1) provided similar results for one of the systems (T), and could be considered for follow-up investigations. The simulations were analysed by measuring the stacking angle between 2Ap and the adjacent bases within the same strand. This was quantified as the angle between the normal vectors to each ring plane, determined *via* the cross-product of vectors defined by selected atoms in each ring. The distribution of the stacking angles over the trajectories are obtained with an histogram with a bandwidth of 1°. The evolution of the stacking angle over the trajectories is shown in Fig. S2.

To assess statistical uncertainty in the stacking-angle distributions (Fig. S1 and S3), we applied a moving-block bootstrap.[53] For each trajectory, the time series of stacking angles was first analysed through the autocorrelation function to obtain the integrated autocorrelation time, τ_{int}, which ranged from ~10 ps up to several hundred ps depending on the system and analysed base. Block lengths of $5\tau_{int}$ were then used to resample the time series while preserving time correlations. For each system, 1000 bootstrap replicas were generated and histograms recomputed. The distributions were then plotted including the resulting 95% confidence bands for the probability density in each bin.

2.7. Quantum mechanical calculations

From the extracted MD geometries, selected as representative from the lower and upper extremes of the stacking-angle distribution, one for each system, we have then optimized their ground state minima (S_{0min}) using density functional theory (DFT) and in particular the M052X[54,55] functional combined with the 6-31G(d) basis set. This method has been selected due to its good performance when treating this kind of system prone to populate charge transfer states[54] and has also already been tested for other fluorescent derivatives inside DNA.[56] Our QM model is based on the hexamer formed by C62ApC8/G6'XG8' (Fig. 1). Solvent (water) effects were also considered through an implicit polarizable continuum model (PCM).[57,58] Then, the absorption energies, intensities, and excited state character were obtained at this S_{0min} resorting to the time-dependent version of DFT (TD-DFT). Subsequently, the excited state potential energy surface was mapped by relaxation of the lowest-lying states at the same level of theory. At this point, the equilibrium version of PCM was selected as the most appropriate. The amount of charge transfer is estimated just through the difference between the Mulliken charges in the ground and excited state. All the calculations were done with Gaussian16 program.[46]

3. Results

3.1. Conformational/structural effect of DNA methylation markers

3.1.1. Thermal melting and circular dichroism analysis of epigenetic DNA duplexes.
The stability of duplexes was measured by their melting temperature (T_m), which was experimentally determined for the 2Ap-labeled DNA duplexes through temperature-dependent changes in fluorescence emission intensity at 370 nm (Fig. 2a and Table S3). Interestingly, the experimental melting temperature ($T_m = 59.5 \pm 0.14$ °C) for the C2ApC/GTG duplex closely matched the T_m predicted by DINAMelt[59] (58.5 °C) for the 5'-GGGCCC**A**CAGGG-3'/5'-CCCTG**T**GGGCCC-3'. This observation confirms that replacing adenine with 2Ap does not remarkably alter the melting temperature of DNA duplexes. As expected from the higher stability of 2Ap-T base pairs compared to 2Ap-C base pairs, the C2ApC/GTG duplexes exhibited significant thermal stability, with a T_m increase of ~9 ± 0.16 °C compared to C2ApC/GCG duplexes. The lower T_m values suggest that mismatched sequences are minimally stable compared to matched sequences. In the case of 2Ap-C base sequences, alterations from C to mC or hmC caused distinct changes in T_m. The C2ApC/GmCG and C2ApC/GhmCG duplexes exhibited only a very slight T_m increase of ~0.4 to 0.8 °C compared to C2ApC/GCG. The slight T_m increase may have originated from the marginal stabilization of mC and hmC compared to C. To further validate the effects of cytosine modifications, the T_m of duplexes containing a template guanine (G) paired with C, mC, and hmC on the complementary strand was measured (Fig. S4 and Table S4). The T_m of CGC/GCG (60.4 ± 0.4 °C) reflects higher stability than that of C2ApC/GCG (50.11 ± 0.17 °C). The CGC/GmCG reduces T_m by ~2.5 °C, and CGC/GhmCG by ~12 °C, compared to CGC/GCG, indicating that cytosine modifications destabilize G-based duplexes. Conversely, in 2Ap-C mismatches, mC and hmC slightly change T_m, indicating weak stabilization. These results highlight that, although 2Ap does not fully mimic guanine in base pairing, it
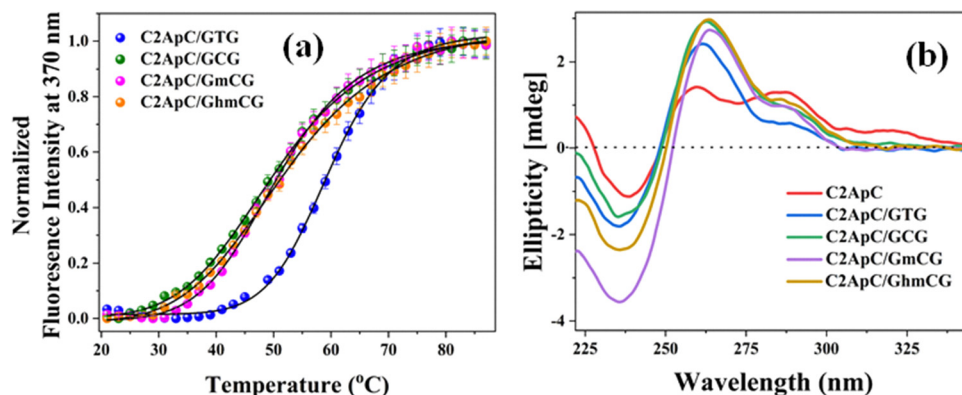
Fig. 2 Melting temperature curves (a) and CD spectra (b) of the 2Ap-labeled ssDNA matched, mismatched modified dsDNA sequences.

remains a reliable fluorescent probe for detecting subtle stability changes induced by cytosine modifications.

2Ap-DNA duplexes show positive CD bands (Fig. 2b) in the ~260 and ~290 nm along with a negative band near ~245 nm, which are typical spectral features of the B-form DNA structure.[32,60] The CD spectra of 2Ap-labeled ssDNA exhibit one extra positive peak at 325 nm, which corresponds to stacking interactions of adjacent 2Ap residues.[60] The 260 nm peak relates to base stacking interactions, while the 245 nm peak is associated with the helical structure of the DNA. There was a significant increase in ellipticity for both positive and negative bands in 2Ap-dsDNA compared to 2Ap-labeled ssDNA duplex. The lower ellipticity of the positive band in 2Ap-T base pair dsDNA compared to 2Ap-C base pair dsDNA is consistent with similar results examined in previous literature.[61] C2ApC/GCG, C2ApC/GmCG, and C2ApC/GhmCG duplexes exhibit equal ellipticity at the positive band, with a minor red shift in the peak position. C2ApC/GhmCG duplexes exhibited a ~2 nm red shift in the positive band when associated with the C2ApC/GTG duplexes, revealing subtle differences in their structural properties. A similar trend occurred at the 290 nm positive peak, where the C2ApC/GTG duplex exhibited lower ellipticity than the C2ApC/GCG and C2ApC/GmCG duplexes. The

C2ApC/GhmCG duplex shows a slight increase in ellipticity at 290 nm compared to C2ApC/GCG and C2ApC/GmCG, suggesting that hmC induces subtle structural changes, possibly affecting helical conformation and base stacking.

3.1.2. Molecular dynamics. In short, both thermal and CD experimental results point to moderate effects on the structure of 2Ap-labeled modified duplexes for the studied bases. To further assess this issue, classical MD simulations have been performed. After a 100 ns run, the stacking degree among adjacent bases was analyzed, being determined as the angle between the vectors perpendicular to each ring (Section 2.6), namely, the co-planarity of adjacent basis increases as the stacking angle approaches zero. The planarity of the consecutive 2Ap and C8 residues depends on X (recall Fig. 1 for labels), as shown in Fig. 3a, which reports the distribution of the corresponding stacking angle. The shaded regions denote the 95% confidence intervals obtained by block bootstrap; their narrow extent confirms that the observed differences and similarities between bases are statistically significant. Specifically, both hmC and C promote less-stacked configurations, with stacking angle distributions with broader tails toward non-planar geometries and reduced probability density around the maximum, whereas T and mC better preserve the co-planarity
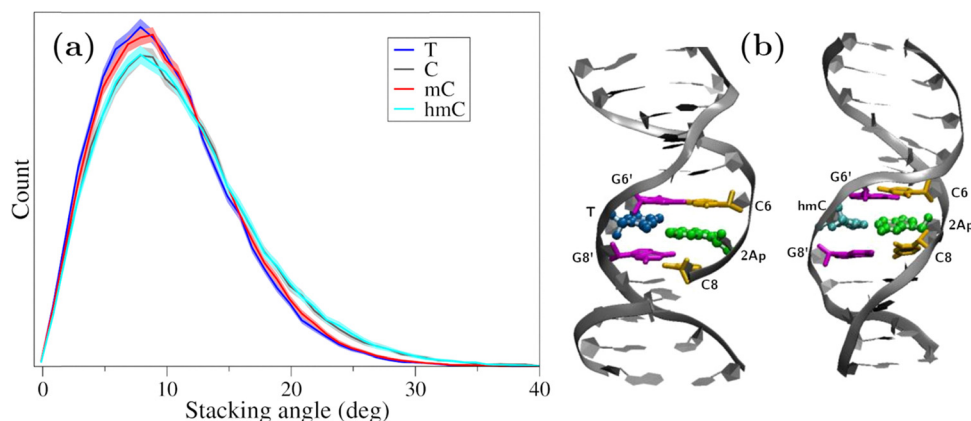

Fig. 3 (a) Stacking-angle distributions between 2Ap and C8 for duplexes containing the different X bases. Average angles are 10.0° (T), 11.0° (C), 10.2° (mC), and 11.1° (hmC). Shaded regions denote 95% confidence intervals obtained from block-bootstrap analysis. (b) Representative structures for T and hmC, including the labels for the nearby bases (color code: C6/C8 in orange, 2Ap in green; G6/G8 in magenta; T in blue, and hmC in cyan).

© 2025 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2025, **6**, 1927–1940 | **1931**

between 2Ap and C8. It should be noted that the tail of the distribution in Fig. 3a corresponds to transient spikes observed along the trajectory, which persist for only a few picoseconds. The more pronounced tails for hmC and C arise from more frequent distortions reaching larger stacking angles, as illustrated in Fig. S2. Despite this difference, completely solvent-exposed or unstacked conformations have not been located along the dynamics. Interestingly, the nature of the X base has a negligible effect on the stacking degree between 2Ap and C6, whereas the stacking of X with G6′ or G8′, which determines how the base accommodates within the strand, shows some dependence on the identity of X (see Fig. S3). Thus, although the global duplex structure is maintained, in agreement with the experimental results, the different nucleobases exhibit slightly different stacking patterns, with hmC showing the largest deviation from co-planarity.

To further confirm this hypothesis two "extreme" cases being X = T and hmC, were extracted from the MD (Fig. 3b) and reoptimized at the DFT level of theory, using as a model the hexamer formed by C6-2Ap-C8 + G6′-X-G8′. For each system representative cases of stacked (i.e., configurations with stacking angles below 1 degree) and less-stacked (i.e., configurations with stacking angles over 30 degrees) conformations were considered. Upon geometry optimization, it is confirmed that T favours stacked 2Ap-C8 arrangements (angles lower than 15 degrees), since even when starting from large angles, the optimization converged to smaller ones; whereas the opposite occurs for hmC, leading to less-stacked ones (angles between 25–35 degrees) even when starting from lower ones. Once it has been confirmed that, although small, some structural changes are induced by the modified bases (especially for hmC), their impact on the absorption/emission of UV light (Section 3.2) as well as in the excited state dynamics (Section 3.3) will be studied.

3.2. Effect on 2Ap-DNA photophysical properties

3.2.1. Steady-state absorption and emission properties of 2Ap-labeled DNA.
The UV absorption spectra for ssDNA, as well as matched and mismatched dsDNA duplexes, were measured (Fig. 4a). In the 2Ap-labeled DNA duplexes, UV absorption in the 240–280 nm range is governed by the normal bases. This dominance occurs from 2Ap's weak absorbance (0.2%) at

260 nm and its 11:1 ratio to normal bases.[62] A small 2Ap absorption shoulder band appears at 307–315 nm, enabling accurate estimation of base contributions,[36,63] since the main absorption band of 2Ap-labeled DNA is red-shifted relative to the absorption band of free 2Ap. Unlike ssDNA, dsDNA contains stacked bases that increase its stability and decrease its absorbance (hypochromicity).[64] Fig. 4a clearly illustrates a decrease in absorbance for dsDNA duplexes compared to ssDNA (C2ApC). The monitored hypochromic effect suggests that Watson–Crick (WC) pairing with a complementary C2ApC/GTG residue promotes optimal stacking of 2Ap with adjacent nucleobases, as already discussed in Section 3.1. In contrast, mismatched duplexes (C2ApC/GCG, C2ApC/GmCG, and C2ApC/GhmCG) show a small increase in absorbance compared to matched duplexes, which may be due to the structural deviation of the flanking base pairs. Nonetheless, all the duplex sequences exhibit same absorption maximum with slightly different amplitudes. The fluorescence emission spectra were measured upon excitation at 290 nm (Fig. S5) and 315 nm (Fig. 4b). As shown in Fig. 4b, ssDNA has a huge emission intensity compared to dsDNA duplexes. The complementary sequence of C2ApC/GTG duplex showed very low emission intensity compared to ssDNA. Mismatched DNA duplexes display higher fluorescence intensity compared to matched sequences. This observation could be related to reduced base-pair stability or conformational changes in the DNA duplexes, potentially induced by cytosine methylation status. Despite the intensity changes, spectral shifts were observed (Fig. S6a). A slight blue shift (~6 nm) was monitored in the dsDNA duplex emission spectra compared to the ssDNA. The emission maxima were ~376 nm, ~374 nm, ~374 nm, ~372 nm, and ~370 nm for C2ApC, C2ApC/GTG, C2ApC/GCG, C2ApC/GmCG, and C2ApC/GhmCG, respectively. C2ApC/GhmCG exhibits distinct spectral changes, displaying a 2 nm blue shift compared to C2ApC/GmCG and a 4 nm blue shift compared to C2ApC/GCG.

The fluorescence excitation spectrum of the 2Ap-labeled ssDNA exhibits two excitation bands (Fig. 4c). The initial band of the excitation spectrum lies between 245 and 285 nm, representing the energy transfer from adjacent DNA bases to 2Ap. The fluorescence intensity in the initial band region correlates with energy transfer efficiency, which increases with
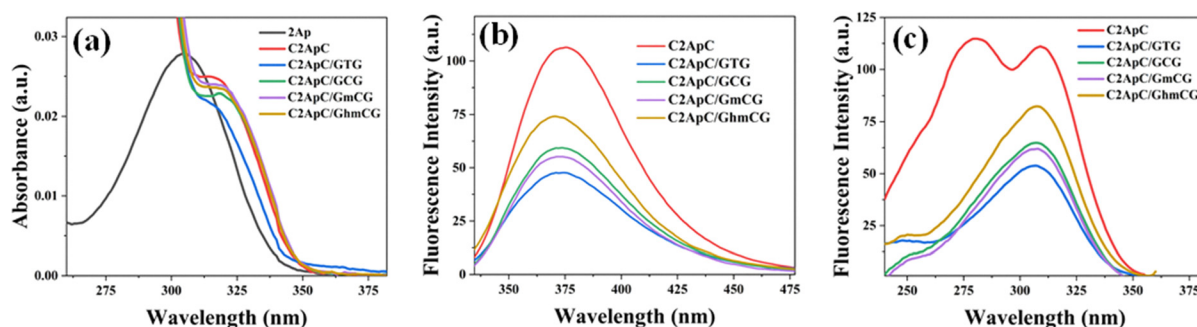

Fig. 4 Absorption (a) and fluorescence emission ((b) collected at 315 nm excitation wavelength) and excitation ((c) collected at 370 nm emission wavelength) spectra of 2Ap, labeled ssDNA, matched, mismatched, methylated, and hydroxymethylated dsDNA sequences.

lower temperatures.[62] A second excitation band represents the direct excitation peak of 2Ap ($\sim$309 nm, which is consistent with the shoulder observed in the absorption spectrum, also located in this region of the spectrum and associated with 2Ap absorption). In the case of 2Ap-labeled dsDNA duplexes, a single excitation band is detected at $\sim$306 nm. At temperatures above 15 °C, it is not feasible to significantly visualize a second excitation band, typically between 255 and 280 nm.[65] Furthermore, dsDNA duplexes show $\sim$50% lower energy transfer efficiency related to ssDNA, which is why the initial excitation band is not observable in dsDNA at room temperature.[62] The slight blue shift ($\sim$3 nm) in the excitation band of 2Ap-labeled dsDNA compared to ssDNA is attributed to enhanced base stacking interactions in the duplex (Fig. S6b). This shift reflects the structural rigidity and reduced solvent exposure of the 2Ap bases. The excitation spectra also reflected the trends observed in the emission intensity changes for matched and mismatched DNA duplexes.

The QYs of 2Ap-labeled DNA duplexes were significantly lower than those observed for free 2Ap.[66] The QY of ssDNA (C2ApC) was 0.0173, while that of C2ApC/GTG was 0.0049 (Section 3.3). This decrease was attributed to significant stacking interactions between T and 2Ap in the dsDNA duplex. The relative QY of C2ApC/GTG is relatively close to those of C2ApC/GCG (0.0063) and C2ApC/GmCG (0.0059) (Section 3.3), although below them. In the case of C2ApC/GhmCG, the QY (0.01) is approximately double that of C2ApC/GCG and C2ApC/GTG.

3.2.2. Vertical absorption and emission energies. The absorption spectra of the labeled duplexes were then simulated by computing the vertical absorption energies and intensities on top of the optimized ground state minima of the four systems (T-stacked and less-stacked and hmC-stacked and less stacked). The absorption maxima and the global shape of the spectra are almost unaffected either by the nature of the X base or by the stacking degree (Fig. S7), in agreement with the experimental trends (Fig. 4a). A more detailed analysis of the lowest lying excited states is collected in Table 1 and Table S5). For the four systems, the first two excited states are located in 2Ap, being the S_1 bright and S_2 dark. Then, the third state S_3 is located on the X nucleobase, and it is dark for T and partially "bright" for hmC. The following states arise from mixtures between the bright $\pi\pi^*$ states of G's and C's and excited state with different amounts of charge transfer (CT). Among the four studied geometries, hmC-stacked stands in terms of CT stability with the S_6 showing a strong CT character involving 2Ap and C8 nucleobases (Fig. S8). Some conclusions can be extracted from the close inspection of Table 1 (Table S5): (1) stacked systems have slightly more CT compared to less-stacked systems, and (2) the stacking degree also affects the nature of the bases involved in the first clear CT state, being 2Ap and C8 in stacked ones and G8 and X in the less stacked systems.

To further understand the fluorescence emission energies and yields, the S_1 bright excited state and the first two states with significant CT character ($>$0.3 a.u.) were optimized for the

Table 1 Excited state character, vertical absorption energies (in eV), oscillator strength, and charge transfer character (in a.u.) computed at the PCM/TD-M052X/6–31G(d) level of theory

Character	ΔE	f	CT	Character	ΔE	f	CT
T-less-stacked				T-stacked			
S_1 $\pi\pi^*$ 2aP	4.71	0.2187	0.0	$\pi\pi^*$ 2aP	4.62	0.1870	0.0
S_2 $n\pi^*$ 2aP	5.07	0.0052	0.0	$n\pi^*$ 2aP	5.06	0.0048	0.0
S_3 $n\pi^*$ T	5.17	0.0096	0.0	$n\pi^*$ T	5.17	0.0070	0.0
S_4 $\pi\pi^*$ G6′	5.18	0.0689	0.1	$\pi\pi^*$ G6′$_{+C6+C8}$	5.21	0.0968	0.1
S_5 $\pi\pi^*$ G8′	5.21	0.0613	0.1	$\pi\pi^*$ G8′$_{+C6+C8}$	5.21	0.1018	0.1
S_6 $\pi\pi^*$ C6	5.29	0.1663	0.1	$\pi\pi^*$ G8′ + CT	5.24	0.0881	0.3
S_7 $\pi\pi^*$ C8$_{+C6}$	5.29	0.3268	0.1	$\pi\pi^*$ C6$_{+G6′}$	5.27	0.1788	0.1
S_8 CT G8-> T	5.38	0.0538	0.7	CT 2aP-> C8	5.32	0.1576	0.5

Character	ΔE	f	CT	Character	ΔE	f	CT
hmC-less-stacked				hmC-stacked			
S_1 $\pi\pi^*$ 2aP	4.59	0.2119	0.0	$\pi\pi^*$ 2aP	4.50	0.1507	0.1
S_2 $n\pi^*$ 2aP	4.96	0.0035	0.0	$n\pi^*$ 2aP	4.95	0.0078	0.0
S_3 $\pi\pi^*$ hmC	5.06	0.1152	0.0	$\pi\pi^*$ hmC	5.06	0.0798	0.2
S_4 $\pi\pi^*$ G6′$_{+C6}$	5.23	0.0995	0.1	$\pi\pi^*$ G6′$_{+C6}$	5.21	0.1498	0.1
S_5 $\pi\pi^*$ G8′$_{+C8}$	5.26	0.0408	0.1	$\pi\pi^*$ G8′$_{+C8}$	5.25	0.0796	0.0
S_6 $\pi\pi^*$ C6 $_{+G6′}$	5.30	0.1478	0.0	CT 2aP-> C8	5.27	0.0439	0.7
S_7 $\pi\pi^*$ C8 $_{+G8′}$	5.31	0.3352	0.2	$\pi\pi^*$ C6 $_{+G6′}$	5.28	0.1548	0.0
S_8 CT G8-> hmC	5.37	0.0533	0.5	$\pi\pi^*$ C8 $_{+CT}$	5.33	0.1913	0.2

four systems. The S_1 state evolves towards a minimum in the first excited Potential Energy Surface (PES) for all of them (Fig. 5 and Table S6). The minimum, that is located on the 2Ap base presents emission energies in the 3.70–3.80 eV energy range, that correspond with Stokes shifts of $\sim$0.8–0.9 eV, except for the hmC less-stacked system whose emission energy is almost 4 eV (blue shifted, Stokes shift $\sim$0.54 eV), in line with experimental trends (Stokes shift 0.56 eV, for C2ApC/GhmCG Fig. 4b). This latter minimum resembles the minimum for the isolated 2Ap base in solution that we have also optimized for comparison, and has an emission energy of 4.04 eV.

The reason for this difference is that for the other three systems, hmC-stacked and T-stacked or T-less-stacked, the minima, even if localized on 2Ap, have a small CT character with the adjacent C6 base (0.1–0.2 a.u., Table S6). This partial CT character is also reflected in the lower oscillator strength (f) of those minima, being at a maximum of 0.2, compared to hmC-less-stacked and isolated 2Ap, where it amounts to 0.3. If only this minimum would be present, and if we assume that the non-radiative decays are the same, hmC-less stacked would be more emissive compared to the rest of the systems. Then, for all the systems, at least one minimum with strong CT character was located. This is in agreement with the quenching of 2Ap fluorescence within DNA observed experimentally (Fig. 4 and Table 2). In hmC-stacked the CT is so strong as observed in absorption that when the S_{11} is optimized the system evolves to a degeneracy region with the ground state, in which a bond is created between hmC and G. In the less-stacked system, the CT is not so strong and, although slightly, the CT minima are less stable than the bright ones. In T-stacked and less-stacked the CT minimum involving the G8-T is as stable as the bright one.
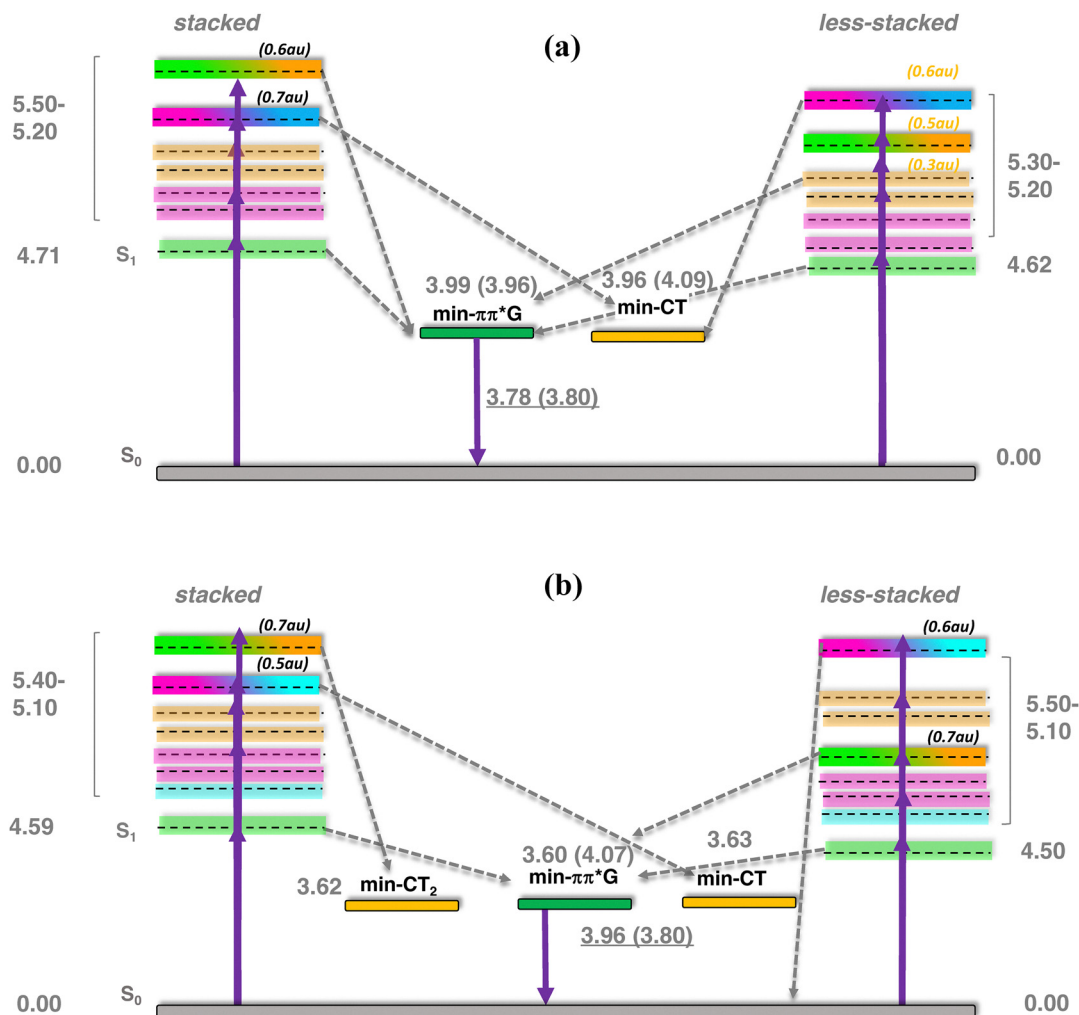
© 2025 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2025, **6**, 1927–1940 | **1933**

Fig. 5 Excited state potential energy surface for C2ApC/GTG (a) and C2ApC/GhmCG (b). The Franck–Condon region is shown for both stacked (left) and less-stacked (right) cases, whereas the evolution towards the minima is shown in the middle. Adiabatic energies (respect to the ground state at Franck–Condon) and emission energies underlined (in eV) for both stacked and less-stacked (parenthesis) systems.

Table 2 Decay-associated spectral properties, rate constants and QY of 2Ap-labeled ssDNA, matched, mismatched and modified dsDNA sequences

Sample	τ_1 (ns)	A_1	f_1	τ_2 (ns)	A_2	f_2	τ_3 (ns)	A_3	f_3	τ_{av} (ns)	k_r ($10^7 \times s^{-1}$)	k_{nr} ($10^7 \times s^{-1}$)	QY
C2ApC	0.61	0.53	0.09	2.35	0.21	0.14	10.00	0.27	0.76	3.47	0.49	28.34	0.0173
C2ApC/GTG	0.61	0.21	0.03	2.35	0.32	0.15	10.00	0.42	0.83	5.12	0.09	19.45	0.0049
C2ApC/GCG	0.61	0.25	0.03	2.35	0.32	0.14	10.00	0.44	0.83	5.27	0.11	18.87	0.0063
C2ApC/GmCG	0.61	0.24	0.03	2.35	0.33	0.15	10.00	0.43	0.82	5.18	0.11	19.19	0.0059
C2ApC/GhmCG	0.61	0.24	0.03	2.35	0.31	0.14	10.00	0.45	0.83	5.35	0.18	18.50	0.0100

3.3. Excited state dynamics impacted by epigenetic bases in DNA

3.3.1. Time-resolved fluorescence and decay-associated spectral analysis. The fluorescence decay of both matched and mismatched DNA duplexes is shown at a maximum emission wavelength of 370 nm (Fig. S9). Rather than focusing on a particular emission maximum wavelength, we investigated the full emission range utilizing the DAS spectra, which allows for the segregation of lifetime components and provides insights

into molecular conformational alteration.[56,67] Fluorescence decay measurements were recorded for the 2Ap-labeled matched and mismatched DNA duplexes across the 320–430 nm emission range with 10 nm intervals to generate the decay-associated spectra, DAS (Table 2). The 2Ap fluorescence decay was well-fitted with three lifetime components (Table S7) for all matched and mismatched duplexes, each correlated with a specific conformational state. ssDNA has a lower average lifetime compared to dsDNA, observed because the fluorescence lifetime of 2Ap varies based on its position within the

1934 | *RSC Chem. Biol.*, 2025, **6**, 1927–1940

© 2025 The Author(s). Published by the Royal Society of Chemistry

ODN, with time constants ranging from $\sim$50 ps to 10 ns.[36,37,62,68,69] The high amplitude populations (α_1) are associated with the lifetime (τ_1), which corresponds to the most strongly stacked conformations of 2Ap. Based on these high percentages, it is clear that 2Ap contributes to stacking interactions with neighbouring bases in the duplex, also suggesting that 2Ap has minimal impact on the structure of the duplex. As a result, the 2Ap residue is stacked efficiently within the duplex. The long lifetimes (τ_2 and τ_3) are likely a result of unstacked or extrahelical conformations.

The amplitude value of A_i was determined by integrating areas under the DAS profile of each lifetime component, normalized to the total steady-state fluorescence intensity. The fractional intensity values (f_i) were calculated by $f_i = \dfrac{A_i \tau_i}{\langle \tau \rangle}$. The non-radiative and radiative rate constants were determined by $k_{nr} = \dfrac{(1 - QY)}{\langle \tau \rangle}$ and $k_r = \dfrac{QY}{\langle \tau \rangle}$ respectively. In DAS experiments, the average error was $\leq$ 5%.

As represented in Fig. S10 and Table 2, the DAS spectra of ssDNA exhibited a high contribution from the τ_1 lifetime component relative to the steady-state emission profile. Furthermore, significant contributions from three lifetime components are witnessed across the entire emission profile.

Nonetheless, a clear differentiation is observed in the τ_2 and τ_3 components, with the longer lifetime component (τ_3) predominantly seeming on the red side of the emission spectrum. As represented in Table 2 and Fig. 6, the amplitudes of individual lifetime components show minimal differences between the matched and mismatched duplexes. The τ_3 lifetime component exhibits a significantly superior amplitude contribution across the entire steady-state emission spectrum associated with the τ_2 and τ_1 components in the 2Ap-labeled dsDNA duplex. However, the average lifetime values from the DAS spectra show a slight difference, with C2ApC/GhmCG exhibiting an τ_{av} of 5.35 ns compared to 5.12 ns for C2ApC/GTG. It is interesting to note that the QY increase in the hmC clearly correlated with the slight increase in the (α_3) population.

This suggests a significant contribution of segmental motions, indicating the high flexibility of the duplexes. The non-radiative (k_{nr}) and radiative (k_r) rate constants were derived through an evaluation of the QY and the average excited-state lifetimes (τ_{av}). Since k_r values are relative to the extinction coefficients of the chromophores,[39] the reduced k_r values align with the hypochromism witnessed in the absorption spectra. Furthermore, our data show that the slightly higher QY values of hmC duplexes, compared to thymine duplexes, are primarily due to decreased k_{nr}. This confirms the decrease in additional
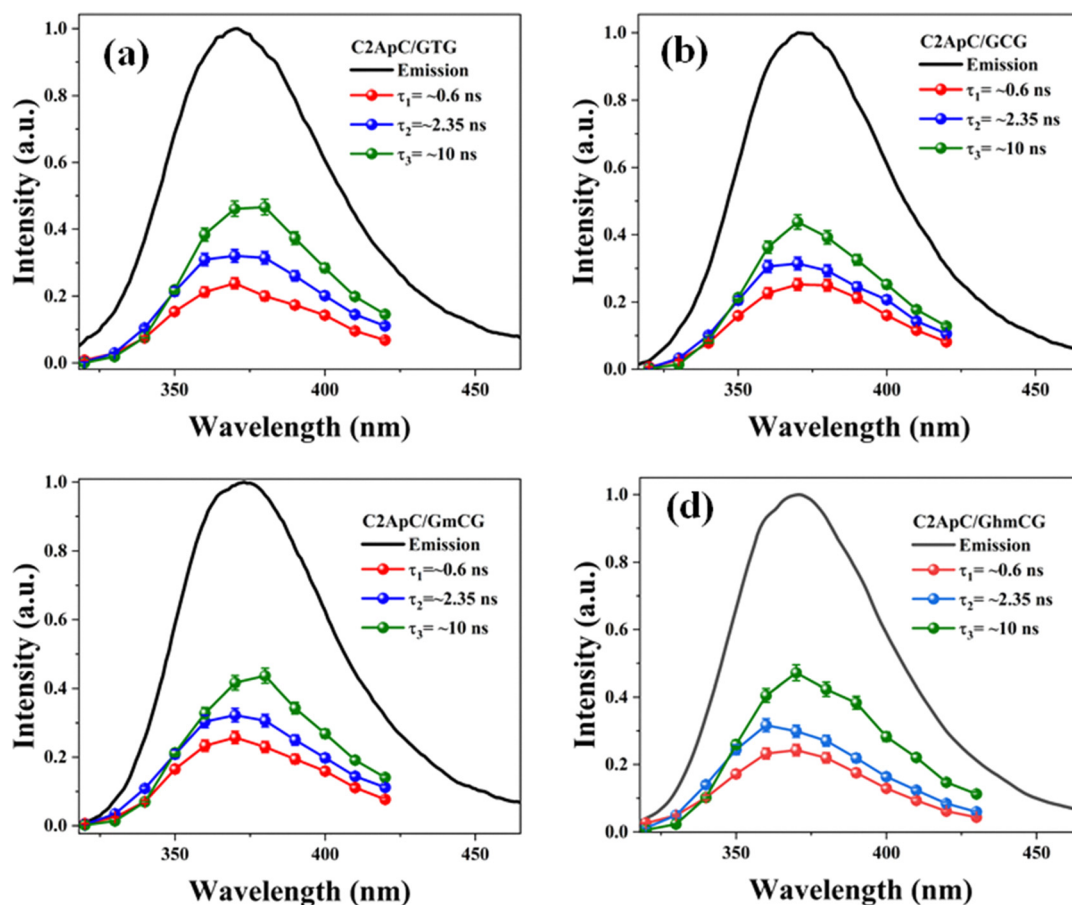

Fig. 6 DAS spectra of matched, mismatched, and modified dsDNA duplexes. The black lines represent the normalized steady-state fluorescence spectra.

non-radiative pathways in hmC modified duplexes compared to matched duplexes. The k_{nr} values (Table 2) exhibit a strong correlation with the stacking behaviour observed in MD simulations (Fig. 3). Precisely, the two systems with reduced stacking interactions (hmC and C) correspond to those with the lowest k_{nr} values.
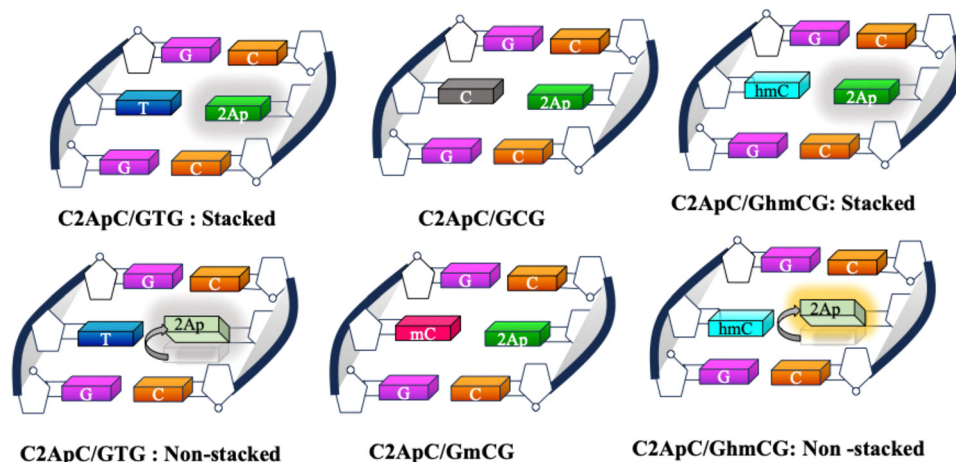
Discussion

Epigenetic modifications, such as methylation and demethylation, play a crucial role in regulating DNA structure and function. Accurately probing the structural dynamics of modified nucleobases within these pathways is essential for understanding their biochemical impact. 2Ap, a fluorescent adenine analogue with high environmental sensitivity, serves as a valuable probe for detecting local conformational changes in DNA. However, a deeper understanding of its photophysical behaviour within nucleic acids is needed to expand its applications in studying modified DNA. This study integrates experimental and theoretical approaches to examine the structural and spectral effects of methylation, hydroxymethylation, and demethylation in DNA duplexes, with 2Ap strategically positioned to report on local changes.

The fluorescence properties of 2Ap in DNA provide a deep insight into the structural dynamics, conformational fluctuations, and interaction mechanisms of DNA and RNA.[69–72] Its fluorescence behaviour encompassing intensity, lifetime, and decay kinetics is intricately modulated by its immediate nucleic acid environment. A major observation is the significant quenching of 2Ap fluorescence in dsDNA compared to ssDNA, which occurs primarily due to base stacking interactions and the restriction of the excited-state dynamics of 2Ap within the duplex.[31,34,37,69,71] This quenching mechanism is driven by both dynamic (collisional) and static (ground-state complex) interactions between 2Ap and neighbouring nucleotides, with varying contributions depending on the sequence context and structural arrangement of the DNA.[31,34,37,69,71] This quenching is further influenced by charge-transfer interactions, as demonstrated in studies of 2Ap paired with a mismatched base pair (cytosine), and by theoretical studies[73,74] (see below), where an equilibrium exists between neutral wobble and protonated Watson–Crick configurations.[75] 2Ap fluorescence QYs are sensitive to base stacking, with mismatched duplex DNA exhibiting higher yields than well-matched duplexes, and ssDNA showing the highest yields due to the absence of base stacking.[36] Furthermore, its fluorescence lifetime also varies depending on its position within the ODN, with time constants ranging from ∼50 ps to 10 ns.[36,37,62,68] In short, the photophysical properties of 2Ap fluorescence are influenced by DNA sequence context, making 2Ap an effective probe for distinguishing variations in DNA stability and structural integrity. Methylation of cytosines at specific positions in 2Ap-labeled dsDNA can destabilize neighbouring residues to the same extent, leading to increased fluorescence QYs and longer fluorescence lifetimes.[36] This sequence-dependent modulation underscores the sensitivity of 2Ap fluorescence to local DNA conformations and interactions.

Base stacking interactions stabilize our here studied 2Ap-labeled matched DNA duplex (C2ApC/GTG) by promoting a highly ordered conformation, in view of the results of the thermal, CD spectra, and MD simulations (Fig. 2 and 3). This is manifested as reduced absorbance compared to ssDNA (Fig. 4a). However, both MD and absorption measurements (larger intensity compared to matched duplex) demonstrate that mismatched duplexes disrupt to some extent this stacking, but keeping the global dsDNA structure, with special attention to C2ApC/GCG and C2ApC/GhmCG (Fig. 2). Notably, the C2ApC/GhmCG duplex exhibits a slight increase in ellipticity at 290 nm compared to C2ApC/GCG and C2ApC/GmCG, suggesting that hmC induces subtle structural changes (Fig. 2b). This could reflect an alteration in helical conformation or base stacking, likely influenced by the hydroxymethyl group at the 5-position of cytosine. Such modification could impact base stacking and hydrogen bonding interactions, leading to a structurally distinct DNA duplex. These findings suggest that hmC may alter the DNA's local secondary structure, potentially influencing its flexibility and base-pairing properties, providing further insight into the structural dynamics of modified cytosine-containing duplexes. This intensity dependence of the low energy region of the spectra on the nature of the X bases is also captured by the oscillator strength (Table 1) for the S_1 $\pi\pi^\star$ excited state. Indeed, the C2ApC/GTG stacked system is 10% smaller than the C2ApC/GhmCG less stacked one. Despite these differences in intensities and the energies of the lowest lying state, the global maximum is not very much affected by the inclusion of the epigenetic bases (Fig. S7). On the other hand, the fluorescence intensities of ssDNA and dsDNA variants differ significantly (Fig. 4), with ssDNA showing higher intensity due to less base stacking. Mismatched duplexes exhibit increased fluorescence compared to the matched duplexes. A blue shift in emission spectra (Fig. S6) further suggests that cytosine modifications, especially hmC, influence duplex conformation and stability. Such spectral shifts indicate changes in base stacking interactions and DNA conformation, resulting in a more hydrophobic environment with a lower dielectric constant around 2Ap.[76]

In the same line, 2Ap-labeled dsDNA shows markedly reduced QYs compared to free 2Ap. For instance, ssDNA (C2ApC) exhibits a QY of 0.0173, while the matched duplex (C2ApC/GTG) is at 0.0049 due to enhanced base stacking (Table 2). Comparable QYs in C2ApC/GCG and C2ApC/GmCG contrast with the higher QY (0.01) of C2ApC/GhmCG. It should be noted that C2ApC/GhmCG is the system that also displays the larger radiative rate (k_r), which leads to a significant increase in QY. Focussing on the non-radiative rate (k_{nr}), which is more directly related to the stacking, C2ApC/GhmCG also displays the lower rate, again consistent with the lesser stacking. In this case, it is also worth mentioning that the trend of k_{nr} values correlates quite well with the stacking behaviour observed in the MD simulations (Fig. 3), with hmC and C, which display larger k_{nr}, being the system with larger deviations from perfect stacking. QM results support these findings, revealing that all the studied systems present two kinds of

1936 | *RSC Chem. Biol.*, 2025, **6**, 1927–1940

© 2025 The Author(s). Published by the Royal Society of Chemistry

Scheme 2 Schematic representation of stacked and non-stacked confirmations for matched, mismatched, methylated, and hydroxymethylated duplexes.

excited minima: a bright $\pi\pi^*$ minimum localized on 2Ap responsible for the fluorescence, plus a CT minimum responsible for the fluorescence quenching compared to the free 2Ap. This fluorescence quenching due to the presence of CT states between 2Ap and adjacent bases has already been reported in the literature for reduced DNA models, such as dinucleotides or trimers.[73,74] Other potential quenching mechanisms occurring through CI or exciplexes have also been characterized in the literature.[74,77,78] Our increased ds hexamer model allows not only to see how efficient these CT minima are in the duplex but to estimate how different CT pathways compete. Indeed, we observe two coupled effects that may affect the fluorescence quenching. On one hand, the bright $\pi\pi^*$ 2Ap minimum has some degree of 2amP- >C6 CT, except for the C2ApC/GhmCG less-stacked system (Table S6 and Scheme 2). Interestingly, this duplex is the one showing the larger fluorescence yield (Table 2). On the other hand, the main quenching effect should be due to the loss of excited state population towards G8- >T/hmC CT minima, which a priori are not directly related to 2Ap. Although both intra and interstrand CT states are present in the FC region, we have not located signs for proton transfer reactions as in canonical DNA.[79–81] It must be noted that we have not attempted to locate $\pi\pi^*$ minima localized on hmC or mC since those excited states have been characterized to decay in the ps timescale, that is shorter than our ns lifetimes,[82,83] due to the accessible S_1/S_0 CI.[29,83,84] Other epigenetic bases, such as 5-formyl-cytosine, do survive longer but are prone to populate triplet states,[85,86] which is out of the scope of the present paper. Lastly, TCSPC analysis revealed that the matched and mismatched duplexes show similar amplitude distributions; however, the C2ApC/GhmCG duplex exhibits a slightly longer average lifetime and an enhanced long-lifetime (α_3) population. This suggests that hmC may destabilize neighbouring residues, reducing non-radiative decay pathways. Consequently, the lower k_{nr} value in C2ApC/GhmCG duplexes correlates with their modestly increased QY, highlighting the sensitivity of 2Ap fluorescence to subtle conformational and dynamic variations in DNA.

Conclusions

Herein, we unveiled the structural and dynamic impact of the epigenetic markers on DNA with the help of an intrinsic fluorescent nucleobase analogue, 2-aminopurine (2Ap), combined with molecular dynamics and quantum mechanical calculations. To the best of our knowledge, this study provides the first comprehensive characterization of the photophysical properties of 2Ap within dsDNA through a combined multi-technique experimental and multi-level computational approach. The results demonstrate that while all 2Ap-labeled duplexes exhibit significant base stacking interactions, keeping the overall structure of dsDNA, subtle variations in stacking degrees occur as a function of the epigenetic base. Indeed, our findings reveal that hmC modifications modulate local DNA conformation by diminishing base stacking and helical stability without compromising overall duplex integrity. This structural modulation results in a more hydrophobic environment around 2Ap, as indicated by blue-shifted emission spectra. Although CT-induced quenching is present in all systems, hmC shows reduced CT character compared to thymine, leading to enhanced fluorescence QYs, although still small, and longer lifetimes. These insights underscore the sensitivity of 2Ap fluorescence to subtle DNA conformational changes and suggest that hmC could play a regulatory role in biological processes by influencing DNA dynamics and local stability. This work not only expands the utility of 2Ap as an environmental probe but also highlights the potential epigenetic implications of hmC in modulating DNA structural and photophysical properties, opening the route for further studies with even more sensitive probes. Such modulation of DNA dynamics offers important insights into the role of epigenetic modifications in cellular function and provides a potential avenue for understanding the molecular underpinnings of epigenetic regulation.

Conflicts of interest

There are no conflicts to declare.

© 2025 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2025, **6**, 1927–1940 | **1937**

Data availability

The data supporting this article have been included as part of the supplementary information (SI) and the results of the molecular dynamics simulations have been uploaded to DOI: **10.5281/zenodo.17312412**. Supplementary information contains MD details and results as well as more excited state analysis from the QM calculations and further experimental results. See DOI: **https://doi.org/10.1039/d5cb00207a**.

Acknowledgements

References

1 B. Carter and K. Zhao, *Nat. Rev. Genet.*, 2021, **22**, 235–250.

2 C. D. Allis and T. Jenuwein, *Nat. Rev. Genet.*, 2016, **17**, 487–500.

3 E. Kriukienė, M. Tomkuvienė and S. Klimašauskas, *Chem. Soc. Rev.*, 2024, **53**, 2264–2283.

4 M. Tahiliani, K. P. Koh, Y. Shen, W. A. Pastor, H. Bandukwala, Y. Brudno, S. Agarwal, L. M. Iyer, D. R. Liu and L. Aravind, *Science*, 2009, **324**, 930–935.

5 S. Ito, L. Shen, Q. Dai, S. C. Wu, L. B. Collins, J. A. Swenberg, C. He and Y. Zhang, *Science*, 2011, **333**, 1300–1303.

6 Z. D. Smith and A. Meissner, *Nat. Rev. Genet.*, 2013, **14**, 204–220.

7 T. H. Bestor and V. M. Ingram, *Proc. Natl. Acad. Sci. U. S. A.*, 1983, **80**, 5559–5563.

8 D. N. Cooper and H. Youssoufian, *Hum. Genet.*, 1988, **78**, 151–155.

9 P. A. Jones, *Nat. Rev. Genet.*, 2012, **13**, 484–492.

10 T. B. Johnson and R. D. Coghill, *J. Am. Chem. Soc.*, 1925, **47**, 2838–2844.

11 R. Holliday and J. E. Pugh, *Science*, 1975, **187**, 226–232.

12 C. Dahl and P. Guldberg, *Biogerontology*, 2003, **4**, 233–250.

13 V. J. Cannistraro and J.-S. Taylor, *J. Mol. Biol.*, 2009, **392**, 1145–1157.

14 S.-G. Jin, W. Xiong, X. Wu, L. Yang and G. P. Pfeifer, *Genomics*, 2015, **106**, 322–330.

15 M. F. Denissenko, J. X. Chen, M.-S. Tang and G. P. Pfeifer, *Proc. Natl. Acad. Sci. U. S. A.*, 1997, **94**, 3893–3898.

16 W. M. Rideout, G. A. Coetzee, A. F. Olumi and P. A. Jones, *Science*, 1990, **249**, 1288–1290.

17 R. Drouin and J.-P. Therrien, *Photochem. Photobiol.*, 1997, **66**, 719–726.

18 H. Stroud, S. Feng, S. Morey Kinney, S. Pradhan and S. E. Jacobsen, *Genome Biol.*, 2011, **12**, 1–8.

19 F. Battistini, P. D. Dans, M. Terrazas, C. L. Castellazzi, G. Portella, M. Labrador, N. Villegas, I. Brun-Heath, C. González and M. Orozco, *PLoS Comput. Biol.*, 2021, **17**, e1009547.

20 Y. Wang, X. Zhang, F. Wu, Z. Chen and X. Zhou, *Chem. Sci.*, 2019, **10**, 447–452.

21 T. Zeng, L. Liu, T. Li, Y. Li, J. Gao, Y. Zhao and H.-C. Wu, *Chem. Sci.*, 2015, **6**, 5628–5634.

22 C. G. Lian, Y. Xu, C. Ceol, F. Wu, A. Larson, K. Dresser, W. Xu, L. Tan, Y. Hu and Q. Zhan, *Cell*, 2012, **150**, 1135–1146.

23 M. Iurlaro, G. Ficz, D. Oxley, E.-A. Raiber, M. Bachman, M. J. Booth, S. Andrews, S. Balasubramanian and W. Reik, *Genome Biol.*, 2013, **14**, R119.

24 M. W. Kellinger, C.-X. Song, J. Chong, X.-Y. Lu, X. He and D. Wang, *Nat. Struct. Mol. Biol.*, 2012, **19**, 831–833.

25 S. Nakano, T. Suzuki, L. Kawarada, H. Iwata, K. Asano and T. Suzuki, *Nat. Chem. Biol.*, 2016, **12**, 546–551.

26 T. T. Ngo, J. Yoo, Q. Dai, Q. Zhang, C. He, A. Aksimentiev and T. Ha, *Nat. Commun.*, 2016, **7**, 10813.

27 Y.-H. You and G. P. Pfeifer, *J. Mol. Biol.*, 2001, **305**, 389–399.

28 H. Ikehata and T. Ono, *J. Radiat. Res.*, 2011, **52**, 115–125.

29 L. Martinez-Fernandez, A. Banyasz, L. Esposito, D. Markovitsi and R. Improta, *Sig. Transduct. Target. Ther.*, 2017, **2**, 17021.

30 D. Dziuba, P. Didier, S. Ciaco, A. Barth, C. A. Seidel and Y. Mély, *Chem. Soc. Rev.*, 2021, **50**, 7062–7107.

31 J. M. Jean and K. B. Hall, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 37–41.

32 M. A. O'Neil and J. K. Barton, *J. Am. Chem. Soc.*, 2002, **124**, 13053–13066.

33 C. Wan, T. Fiebig, O. Schiemann, J. K. Barton and A. H. Zewail, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 14052–14055.

34 A. Dallmann, L. Dehmel, T. Peters, C. Mügge, C. Griesinger, J. Tuma and N. P. Ernsting, *Angew. Chem., Int. Ed.*, 2010, **34**, 5989–5992.

35 R. Eritja, B. E. Kaplan, D. Mhaskar, L. C. Sowers, J. Petruska and M. F. Goodman, *Nucleic Acids Res.*, 1986, **14**, 5869–5884.

36 V. J. Greiner, L. Kovalenko, N. Humbert, L. Richert, C. Birck, M. Ruff, O. A. Zaporozhets, S. Dhe-Paganon, C. Bronner and Y. Mely, *Biochemistry*, 2015, **54**, 6012–6020.

37 A. C. Jones and R. K. Neely, *Q. Rev. Biophys.*, 2015, **48**, 244–279.

38 V. Kilin, K. Gavvala, N. P. Barthes, B. Y. Michel, D. Shin, C. Boudier, O. Mauffret, V. Yashchuk, M. Mousli and M. Ruff, *J. Am. Chem. Soc.*, 2017, **139**, 2520–2528.

39 J. R. Lakowicz, *Principles of fluorescence spectroscopy*, Springer, 2006.

40 J. R. Knutson, J. M. Beechem and L. Brand, *Chem. Phys. Lett.*, 1983, **102**, 501–507.

41 J. M. Beechem, M. Ameloot and L. Brand, *Chem. Phys. Lett.*, 1985, **120**, 466–472.

42 M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek and G. R. Hutchison, *J. Cheminf.*, 2012, **4**, 17.

43 I. Ivani, P. D. Dans, A. Noy, A. Pérez, I. Faustino, A. Hospital, J. Walther, P. Andrio, R. Goñi, A. Balaceanu, G. Portella, F. Battistini, J. L. Gelpí, C. González, M. Vendruscolo, C. A. Laughton, S. A. Harris, D. A. Case and M. Orozco, *Nat. Methods*, 2016, **13**, 55–58.

44 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.

45 J.-D. Chai and M. Head-Gordon, *Phys. Chem. Chem. Phys.*, 2008, **10**, 6615–6620.

46 M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, Williams, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith, R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman and D. J. Fox, *Gaussian 16*, Wallingford, CT, 2016.

47 S. Pronk, S. Páll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess and E. Lindahl, *Bioinformatics*, 2013, **29**, 845–854.

48 S. Nosé, *Mol. Phys.*, 1984, **52**, 255–268.

49 W. G. Hoover, *Phys. Rev. A.*, 1985, **31**, 1695–1697.

50 M. Parrinello and A. Rahman, *J. Appl. Phys.*, 1981, **52**, 7182–7190.

51 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.

52 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

53 S. Mignani and R. Rosa, *Comput. Phys. Commun.*, 1995, **92**, 203–213.

54 Y. Zhao and D. G. Truhlar, *Acc. Chem. Res.*, 2008, **41**, 157–167.

55 Y. Zhao, N. E. Schultz and D. G. Truhlar, *J. Chem. Theory Comput.*, 2006, **2**, 364–382.

56 J. Kuchlyan, L. Martinez-Fernandez, M. Mori, K. Gavvala, S. Ciaco, C. Boudier, L. Richert, P. Didier, Y. Tor and R. Improta, *J. Am. Chem. Soc.*, 2020, **142**, 16999–17014.

57 J. Tomasi, B. Mennucci and R. Cammi, *Chem. Rev.*, 2005, **105**, 2999–3093.

58 S. Miertuš, E. Scrocco and J. Tomasi, *Chem. Phys.*, 1981, **55**, 117–129.

59 N. R. Markham and M. Zuker, *Nucleic Acids Res.*, 2005, **33**, W577–W581.

60 N. P. Johnson, W. A. Baase and P. H. Von Hippel, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 3426–3431.

61 S. M. Law, R. Eritja, M. F. Goodman and K. J. Breslauer, *Biochemistry*, 1996, **35**, 12329–12337.

62 D.-G. Xu and T. M. Nordlund, *Biophys. J.*, 2000, **78**, 1042–1058.

63 D. Xu, K. O. Evans and T. M. Nordlund, *Biochemistry*, 1994, **33**, 9592–9599.

64 J. M. Berg, J. L. Tymoczko and L. Stryer, *Biochemistry (loose-leaf)*, Macmillan, 2007.

65 T. M. Nordlund, D. Xu and K. O. Evans, *Biochemistry*, 1993, **32**, 12090–12095.

66 D. Ward, E. Reich and L. Stryer, *J. Biol. Chem.*, 1969, **244**, 1228–1237.

67 D. Takkella, L. Martinez-Fernandez and K. Gavvala, *J. Photochem. Photobiol. A: Chem.*, 2024, **457**, 115928.

68 T. Goel, T. Mukherjee, B. J. Rao and G. Krishnamoorthy, *J. Phys. Chem. B*, 2010, **114**, 8986–8993.

69 T. Ramreddy, B. J. Rao and G. Krishnamoorthy, *J. Phys. Chem. B*, 2007, **111**, 5757–5766.

70 J. D. Ballin, S. Bharill, E. J. Fialcowitz-White, I. Gryczynski, Z. Gryczynski and G. M. Wilson, *Biochemistry*, 2007, **46**, 13948–13960.

71 P. Manoj, C.-K. Min, C. Aravindakumar and T. Joo, *Chem. Phys.*, 2008, **352**, 333–338.

72 M. Sholokh, R. Sharma, D. Shin, R. Das, O. A. Zaporozhets, Y. Tor and Y. Mély, *J. Am. Chem. Soc.*, 2015, **137**, 3185–3188.

73 J. M. Jean and K. B. Hall, *Biochemistry*, 2002, **41**, 13152–13161.

74 J. Liang, Q. L. Nguyen and S. Matsika, *Photochem. Photobiol. Sci.*, 2013, **12**, 1387–1400.

75 L. C. Sowers, Y. Boulard and G. V. Fazakerley, *Biochemistry*, 2000, **39**, 7613–7620.

76 T. J. Su, B. A. Connolly, C. Darlington, R. Mallin and D. T. Dryden, *Nucleic Acids Res.*, 2004, **32**, 2223–2230.

77 J. Liang and S. Matsika, *J. Am. Chem. Soc.*, 2011, **133**, 6799–6808.

78 J. Liang and S. Matsika, *J. Am. Chem. Soc.*, 2012, **134**, 10713–10714.

79 L. Martinez Fernandez, F. Santoro and R. Improta, *Acc. Chem. Res.*, 2022, **55**, 2077–2087.

80 L. Martinez-Fernandez and R. Improta, *Faraday Discuss.*, 2018, **207**, 199–216.

81 L. M. Fernández and R. Improta, *Nucleic Acid Photophysics and Photochemistry*, Springer, 2024, pp. 29–50.

82 X. Wang, Y. Yu, Z. Zhou, Y. Liu, Y. Yang, J. Xu and J. Chen, *J. Phys. Chem. B*, 2019, **123**, 5782–5790.

© 2025 The Author(s). Published by the Royal Society of Chemistry

RSC Chem. Biol., 2025, **6**, 1927–1940 | 1939

83 L. Martínez-Fernández, A. Pepino, J. Segarra-Martí, J. Jovaisaite, I. Vaya, A. Nenov, D. Markovitsi, T. Gustavsson, A. Banyasz and M. Garavelli, *J. Am. Chem. Soc.*, 2017, **139**, 7780–7791.

84 P. Kabacinski, M. Romanelli, E. Ponkkonen, V. K. Jaiswal, T. Carell, M. Garavelli, G. Cerullo and I. Conti, *J. Phys. Chem. Lett.*, 2021, **12**, 11070–11077.

85 X. Wang, L. Martínez-Fernández, Y. Zhang, P. Wu, B. Kohler, R. Improta and J. Chen, *J. Am. Chem. Soc.*, 2024, **146**, 1839–1848.

86 X. Wang, L. Martínez-Fernández, Y. Zhang, K. Zhang, R. Improta, B. Kohler, J. Xu and J. Chen, *Chem. – Eur. J.*, 2021, **27**, 10932–10940.