

Photophysical property of fluorescent guanine analogs for selectively recognizing acetylated cytosine: A theoretical study

Xiaolin Chen(陈晓琳), Xixi Cui(崔习习), Yongkang Lyu(吕永康), Chenyang Zhang(张晨阳), Changzhe Zhang(张常哲), and Qingtian Meng(孟庆田)

Citation: Chin. Phys. B, 2025, 34 (**5**): 053102. DOI: 10.1088/1674-1056/adb8baJournal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Microscopic mechanism of plasmon-mediated photocatalytic H_2 splitting on Ag-Au alloy chain

Yuhui Song(宋玉慧), Yirui Lu(芦一瑞), Axin Guo(郭阿鑫), Yifei Cao(曹逸飞), Jinping Li(李金萍), Zhengkun Fu(付正坤), Lei Yan(严蕾), and Zhenglong Zhang(张正龙)

Chin. Phys. B, 2024, 33 (**3**): 033101. DOI: 10.1088/1674-1056/ad123d

Density functional theory investigation on lattice dynamics, elastic properties and origin of vanished magnetism in Heusler compounds CoMnV*Z* (*Z*= Al, Ga)

Guijiang Li(李贵江), Enke Liu(刘恩克), Guodong Liu(刘国栋), Wenhong Wang(王文洪), and Guangheng Wu(吴光恒)

Chin. Phys. B, 2021, 30 (**8**): 083103. DOI: 10.1088/1674-1056/ac0a6a

A strategy to improve the electrochemical performance of Ni-rich positive electrodes: Na/F-co-doped $LiNi_{0.6}Mn_{0.2}Co_{0.2}O_2$

Hui Wan(万惠), Zhixiao Liu(刘智骁), Guangdong Liu(刘广东), Shuaiyu Yi(易帅玉), Fei Gao(高飞), Huiqiu Deng(邓辉球), Dingwang Yuan(袁定旺), and Wangyu Hu(胡望宇)

Chin. Phys. B, 2021, 30 (**7**): 073101. DOI: 10.1088/1674-1056/abeeeb

CCSD(T) study on the structures and chemical bonds of AnO molecules (An=Bk-Lr)

Xiyuan Sun(孙希媛), Pengfei Yin(殷鹏飞), Kaiming Wang(王开明), and Gang Jiang(蒋刚)

Chin. Phys. B, 2021, 30 (**3**): 033101. DOI: 10.1088/1674-1056/abc3b1

Density functional theory study of formaldehyde adsorption and decomposition on Co-doped defective CeO_2 (110) surface

Yajing Zhang(张亚婧), Keke Song(宋可可), Shuo Cao(曹硕), Xiaodong Jian(建晓东), and Ping Qian(钱萍)

Chin. Phys. B, 2021, 30 (**10**): 103101. DOI: 10.1088/1674-1056/abff42

Photophysical property of fluorescent guanine analogs for selectively recognizing acetylated cytosine: A theoretical study

Xiaolin Chen(陈晓琳), Xixi Cui(崔习习), Yongkang Lyu(吕永康),
Chenyang Zhang(张晨阳), Changzhe Zhang(张常哲)†, and Qingtian Meng(孟庆田)‡

School of Physics and Electronics, Shandong Normal University, Jinan 250358, *China*

(Received 15 December 2024; revised manuscript received 15 January 2025; accepted manuscript online 21 February 2025)

The photophysical properties of fluorescent nucleobase analogs play a crucial role in nucleic acids detection and the investigation of their structural and functional characteristics. In this study, we computationally designed a series of quasi-intrinsic fluorescent probes according to natural guanine (G) for selectively identifying covalent N^4-acetylcytosine (4acC), a base that is highly correlated with active transcription and gene expression. This work aims to gain insight into the role of 4acC in biological regulation with minimal perturbation to the native DNA structure. The results indicate that these G-analogs possess extended π-conjugation in comparison with the natural guanine, which could yield efficient fluorescence emission and red-shifted absorption. Especially, the 8-thienyl-2'-deoxyguanosine (ThG) exhibits the highest fluorescence intensity and avoids self-absorption on account of the large Stokes shifts (> 67 nm). What is more, the fluorescence of ThG is unaffected to base pairing with natural cytosine, while the obvious fluorescence quenching is observed by virtue of the excited state intermolecular charge transfer after pairing with 4acC, so it is supposed as a promising biosensor for monitoring the fluorescence changes in the absence or presence of the 4acC. Additionally, the impact of binding deoxyribose on photophysical properties is explored to guarantee the biological applicability of the bright G-analogs in real environment.

Keywords: guanine analog, base modification, fluorescence probe, density functional theory

PACS: 31.15.E, 32.50.+d, 33.50.Dq, 34.70.+e

DOI: 10.1088/1674-1056/adb8ba **CSTR:** 32038.14.CPB.adb8ba

1. Introduction

Nucleic acid modifications of genomic DNA naturally occur within cells, and are often involved in critical biological functions, including regulating gene expression and stabilizing nucleic acid structure.[1,2] These modifications would not disrupt Watson–Crick (WC) pairing, but they can impact the interaction between DNA and proteins within the major groove of the double helix.[3] The N^4-acetylcytosine (ac^4C) is catalytically produced by the N-acetyltransferase 10 (NAT10), which is a highly conserved RNA modification.[4,5] The ac^4C is originally discovered in bacterial tRNA,[6] and subsequent studies have revealed that it can help to increase the fidelity of protein translation.[7] Although the modification is well-studied in RNA, the existence and epigenetic of N^4-acetyldeoxycytosine (4acC) role in cellular DNA are not satisfactorily understood. Inspired by the chemical modifications of RNA and DNA, such as m^6A/6mA,[8,9] m^5C/5mC,[10,11] and the RNA modification ac^4C,[12,13] Wang *et al.*[14] explored the existence and function of 4acC in DNA. It is found that the 4acC modification is mainly distributed in the euchromatin region of arabidopsis, with the highest abundance near the transcription start site of protein-coding genes among euchromatin regions, and its abundance is significantly positively correlated with active transcription and gene expression.[14–16] Consequently, the selective determination of 4acC in DNA is essential for exploring its biological regulation and interaction with other epigenetic information. It not only carries implications for understanding biological interactions and mechanism of 4acC formation at the molecular level, but also extends the DNA modified catalog of regulating gene expression.

Currently, 4acC immunoprecipitation sequencing (4acC-IP-Seq) and ultrahigh performance liquid chromatography-electrospray ionization-tandem mass spectrometry (UPLC-ESI-MS) are available to detect the 4acC.[14] However, the enzymatic hydrolysis of nucleotides caused by these methods tends to disrupt the natural structure of DNA and affect biological interactions. Therefore, there is an urgent need to achieve selective response on microenvironments and position-specific conformations with minimal interference to the natural configuration. In response to the hereinbefore problems, the fluorescent nucleobase analogs (FNAs) have attracted much attention due to their high sensitivity and the easy accessibility.[17,18] An ideal FNA should satisfy at least two touchstones: Geometrically, it should be similar to the native nucleosides when incorporated in the target nucleic acid sequence without disturbing the native structure and interactions of the labelled nu-

†Corresponding author. E-mail: zhe852456@126.com
‡Corresponding author. E-mail: qtmeng@sdnu.edu.cn

© 2025 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

cleic acid. Electronically, the FNA should exhibit high fluorescence quantum yield and optimal spectroscopic characteristics when incorporated into the target nucleic acids.[19,20] Recently, many fluorescent analogs have been developed for molecular dynamics exploration, nucleic acid detection, as well as DNA structure and function determination. For example, the cyano quadracyclic adenine 2CNqA, due to the high and relatively sequence-independent quantum yield in DNA, was reported to promise as an interbase Förster resonance energy transfer (FRET) donor and a nucleic acid label.[21] Additionally, the trans-stilbene cytosine ^{ts}C was utilized to the real-time track of reversible conversions between i-motif, single-stranded and double-stranded structures derived from the human telomeric repeat sequence.[22] In recent years, our group also designed a series of the fluorescent C-analogs to monitor ^{AN}dG[23] as well as ^{ABP}dG adducts,[24] and the results show that the fluorescence probe can not only realize the identification of the covalent adduct, but also be expected to provide assistance in the cellular repair of DNA damage.

In this work, a theoretical investigation on the fluorescent characterizations of a series of bright G-analogs in aqueous solution was presented in order to achieve ultrasensitively identify the covalent 4acC. Herein, two modifications by fluorophore replacement (ThG, 8-NO_2G, ^{tz}G)[25–27] and ring-expansion (^{ttz}G, xxG, A1G)[19,28,29] were considered based on the complementary guanine due to the fact that the 4acC prefers to form a stable WC base pair with guanine. The calculations show that the G-analogs exhibits extended π-conjugation that could yield efficient fluorescence emission and red-shifted absorption. It is remarkable that the fluorescence of ThG is unaffected to base pairing with natural cytosine, while the obvious fluorescence quenching occurs after pairing with 4acC on account of the excited state intermolecular charge transfer (ESICT). In addition, the influence of binding deoxyribose on photophysical properties was examined to assess the biological applicability of the fluorescent G-analogs in real environment. This study theoretically proposes a series of quasi-intrinsic fluorescent probes for real-time detection of 4acC, which also shows a feasible way towards comprehending biological interactions and formation mechanism of 4acC.

2. Theoretical methods

Kasha's rule points out that upon excitation at any absorption wavelength the fluorescence emitted by a molecule originates from its lowest electronically excited state.[30] Therefore, we employed density functional theory (DFT) and time-dependent density functional theory (TDDFT) method to fully optimize the structures of base molecules at the stable ground state (S_0) and the first singlet excited state (S_1).[31–33] The functional B3LYP(D3)[34–36] with basis set 6-311++G(d,p)[37–39] was used in simulations. The solvent effect (water) on the spectroscopic properties was evaluated by employing the integral equation formalism of the polarized continuum model (IEFPCM).[40,41] Meanwhile, all optimized geometries were performed the vibrational frequency analysis to ensure the absence of imaginary frequencies for each obtained geometry.[42] Additionally, the electronic spectra with the three lowest singlet excited states were calculated in accordance with the fully optimized geometries at S_0 and S_1 states, and were broadened by employing the Gaussian function

$$G(w) = \frac{1}{c\sqrt{2\pi}} \exp\left(-\frac{(w-w_n)^2}{2c^2}\right), \tag{1}$$

in which $c = \mathrm{FWHM}/2\sqrt{2\ln 2}$ and the full width at half maximum (FWHM) is program-default 0.4 eV, w represents the spectral abscissa and the w_n denotes the excitation energy corresponding to transitions from the first to the n-th excited state.

The above-mentioned configurations and relative conclusions were also validated utilizing the PBE0(D3) and M062X(D3) functional.[43,44] Herein, the XYZ coordinates of molecules designed in this study were obtained in GaussianView, and all of the theoretical calculations were performed with the Gaussian16 A.03 software.[45] The related geometries, frontier molecular orbital (FMO) isosurfaces and hole-electron distribution[46] were plotted with the help of Visual Molecular Dynamics (VMD) software[47] according to the files generated by Multiwfn.[48] Furthermore, partial configurations were also plotted by the Chemcraft software.[49]

3. Results and discussion

In order to recognize the 4acC by using the proposed fluorescent probes, the basic photophysical properties of monomers are firstly calculated based on the stable configurations in Fig. S1 in supporting materials. Our results indicate that the fluorescence of 4acC possesses a weak oscillator intensity $f_{\mathrm{flu.}} = 0.07$ at around 354 nm, as shown in Fig. S2 and Table S1. Given the insufficient intrinsic optical emission of 4acC, two modifications by fluorophore replacement (ThG, 8-NO_2G, ^{tz}G) and ring-expansion (^{ttz}G, xxG, A1G) (Fig. 1) were made based on the natural guanine, which can form stable WC base pair with 4acC. On this basis, the spectroscopic characteristics of these analogs and base pairs were investigated to gain a comprehensive understanding of the fluorescence enhancement and the fluorescence quenching mechanism.

3.1. Photophysical properties of isolated G-analogs

The optimized configuration of the G-analogs (see Fig. 1) retains the purine infrastructures and WC hydrogen bond surfaces, which could ensure the native structures and biological functions of nucleic acid with minimal disturbance.

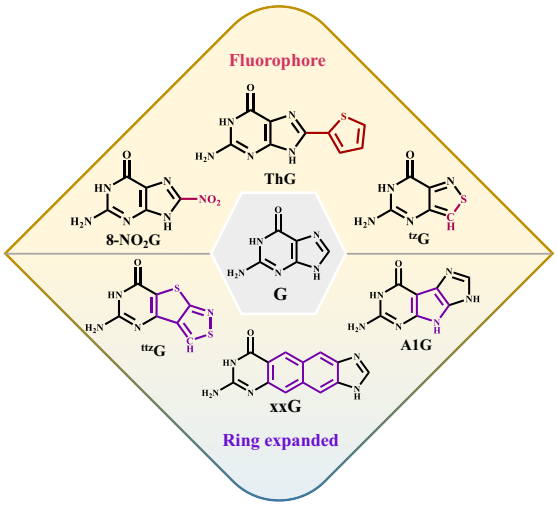


Fig. 1. Chemical structure formula of guanine and modified G-analogs.

Spectroscopic characteristics (Table S2 and Figs. S3(a) and S3(b)) of these modified monomers indicate that G-analogs possess bright photoluminescence and distinct red-shifted absorption (23 nm–163 nm) arisen from extended π-conjugation. Among these fluorescent analogs, the ThG exhibits the highest fluorescence intensity ($f_{\text{flu.}} = 0.73$) and can avoid self-absorption on account of large Stokes shifts (> 67 nm), so it will be explored in detail as an example to elucidate the underlying mechanism of the fluorescence enhancement.

Firstly, the potential energy curves (PECs) of ThG in solution are constructed, where the dihedral angle δ(N1-C2-C3-S4) is varied in steps of 20° to explore the stable configuration of ThG base (see Fig. S4). It can be seen that one minimum is identified at $\delta = 180°$ along the minimal energy pathways for rotation of the the thiophene group within the ThG in both S_0 and S_1 states. As illustrated in Fig. S5, the geometrical structure of the ThG monomer in solution is close to natural guanine in shape, thereby minimizing potential interference with the helical structure and function of nucleic acids. Moreover, the preservation of the fundamental purine structure and WC hydrogen bonding ensures the ThG base can combine with complementary target bases to form a stable DNA double-strand with the assistance of DNA polymerase.[50] Based on this, the electronic spectra of ThG were obtained attending to the optimized geometry in Fig. 2(a) to uncover the function of the modified groups in spectroscopic characteristics. As demonstrated in Fig. 2(b), ThG has a sharp absorption peak at 341 nm ($S_0 \rightarrow S_1$, $f_{\text{abs.}} = 0.67$) and a weaker peak at 275 nm ($S_0 \rightarrow S_2$, $f_{\text{abs.}} = 0.12$), which is red-shifted compared with the native G base (243 nm), and this characteristic allows its selective photo-excitation in the presence of the natural bases. The two peaks correspond to the HOMO→LUMO (HOMO: highest occupied molecular orbital, LUMO: lowest unoccupied molecular orbital) (98.8%) and the HOMO→LUMO+1 (95.4%), respectively (Table 1), and the obvious red-shift can be attributed to the decreased

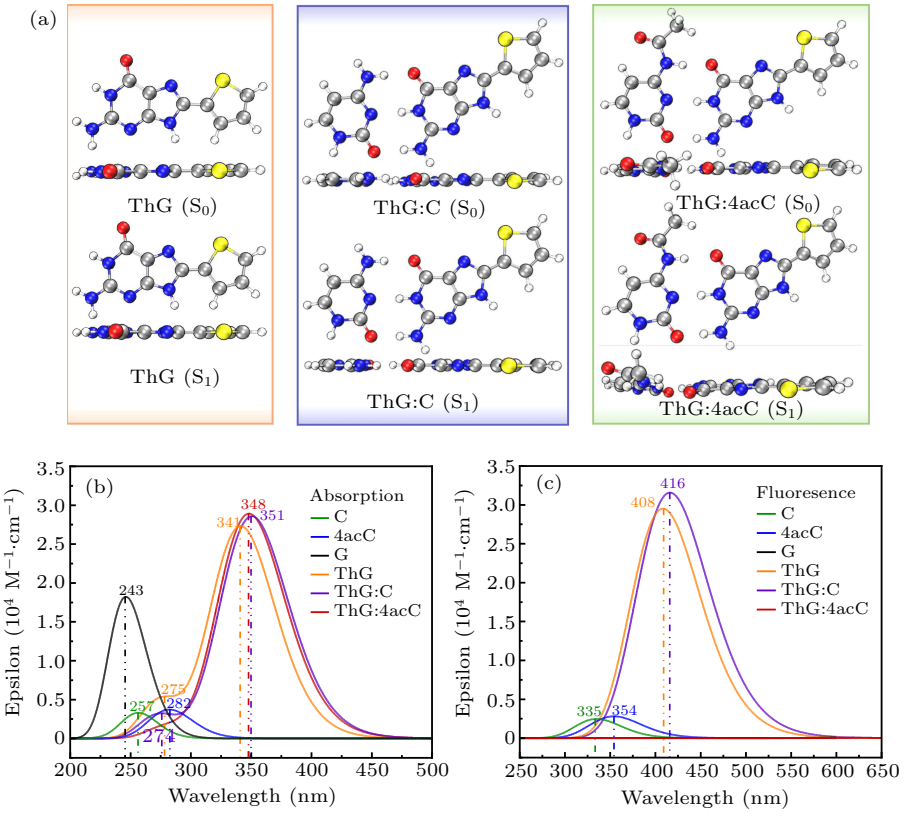


Fig. 2. Diagram of the ThG, ThG:C, ThG:4acC structure (a), absorption spectra of ThG monomer and corresponding base pairs (b), fluorescence spectra of ThG monomer and corresponding base pairs (c). Atomic colors: white (H), gray (C), blue (N), red (O), yellow (S).

energy difference between the FMOs involved in the transitions (Fig. 3). Subsequent to vertical excitation, the emission process from the relaxed structure of the S_1 state is simulated (Fig. 2(c)). The ThG monomer possesses a strong oscillator intensity ($f_{\text{flu.}} = 0.73$) with fluorescence emission observed at 408 nm. As shown in Fig. 3 and Table 1, the fluorescence is dominated by LUMO→HOMO (99.6%), which allows bright emission process and presents $\pi^* \rightarrow \pi$ characteristic. The large Stokes shift of 67 nm exhibited by the ThG monomer favors its application in fluorescence detection. This difference between excitation and emission maxima minimizes self-absorption, thereby improving detection sensitivity.

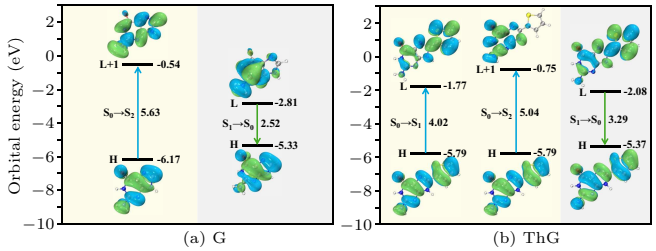


Fig. 3. View of the FMOs (isovalue = 0.02 a.u.) involved in the excitation and emission processes for the natural G (a) and the modified ThG (b). Green and blue signify the positive isosurface and negative isosurface, respectively. H: HOMO; L: LUMO.

Table 1. Absorption/fluorescence wavelengths (λ), oscillator strengths ($f_{\text{abs.}}$ and $f_{\text{flu.}}$) and state assignments for G and ThG. CI represents the composition index of the orbitals involved in the electronic transitions. Type and Gap represent the characteristic and energy gap of the orbital transitions, respectively.

Optical spectral	Base	Transition	λ (nm)	$f_{\text{abs.}}/f_{\text{flu.}}$	Assignment (CI/Type/Gap (eV))
Absorption	G	$S_0 \rightarrow S_2$	243	0.34	HOMO→LUMO+1 (85.2%/$\pi\pi^*$/5.63)
	ThG	$S_0 \rightarrow S_1$	341	0.67	HOMO→LUMO (98.8%/$\pi\pi^*$/4.02)
		$S_0 \rightarrow S_2$	275	0.12	HOMO→LUMO+1 (95.4%/$\pi\pi^*$/5.04)
Fluorescence	G	$S_1 \rightarrow S_0$	–	0.01	LUMO→HOMO (99.6%/πn/2.52)
	ThG	$S_1 \rightarrow S_0$	408	0.73	LUMO→HOMO (99.6%/$\pi\pi^*$/3.29)

3.2. The effects of base pairing on the photophysical properties

As previously mentioned, the G-analogs could pair with the targeted 4acC by hydrogen bonds, thus the impact of base pairing on the photophysical characterizations is studied in order to evaluate the application of optical probe to recognize the 4acC. The results reveal that these analogs exhibit different optical behaviors with respect to the base pairing (Table S2 and Figs. S3(c) and S3(d)). The analogs (8-NO_2G, ^{tz}G, ^{ttz}G, xxG) are observed to have a bright emission upon base pairing with both C and 4acC. Furthermore, the fluorescence of A1G is almost completely quenched upon base pairing with C and 4acC. In particular, the photoluminescence of ThG remains unaffected when base pairing with C, whereas a quench is observed in the presence of 4acC.

To assess the usefulness of an indicator in the detection of 4acC, the interaction energy (ΔE) of the ThG after pairing with the 4acC base was calculated utilizing the following formula:[51]

$$\Delta E = E_{\text{AB}} - (E_{\text{A}} + E_{\text{B}}) + E_{\text{BSSE}}, \tag{2}$$

where ΔE stands for the energy differences between the base pairs (E_{AB}) and the base monomers (E_{A} and E_{B}), E_{BSSE} represents the superposition error of the basis set. The results with correction for BSSE show that the interaction energy of ThG:4acC in S_0 state is −17.48 kcal/mol, which is comparable to the natural base pair G:C (−17.19 kcal/mol). Accordingly, ThG can form stable WC base pair with 4acC in the double helix structure of DNA. Subsequently, the relevant spectra based on the optimized stable configuration are obtained in Fig. 2, and the pertinent information of the transition is presented in Table 2. The absorption peak of ThG:4acC is situated at 348 nm ($f_{\text{abs.}} = 0.71$) (Fig. 2(b)), which is mainly due to the HOMO→LUMO+1 (98.9%) transition corresponding to $S_0 \rightarrow S_2$. The base pairing results in a slight red-shift of the absorption peak by 7 nm compared with the ThG monomer (341 nm). For the emission process, it can be illustrated in Fig. 2(c) that the fluorescence intensity of ThG:4acC is nearly zero, which indicates that the fluorescence of ThG is totally quenched after binding with 4acC. The deactivation phenomenon is linked to the FMOs involved in the $S_1 \rightarrow S_0$ transition (Fig. 4(b)), where the LUMO→HOMO (99.8%) contribution is predominant. As demonstrated in Fig. 4(b), the decay process is accompanied by significant ESICT characteristic, namely the electron transfers from the ThG part (S_2) to the 4acC moiety (S_1), which produces significant optical changes.

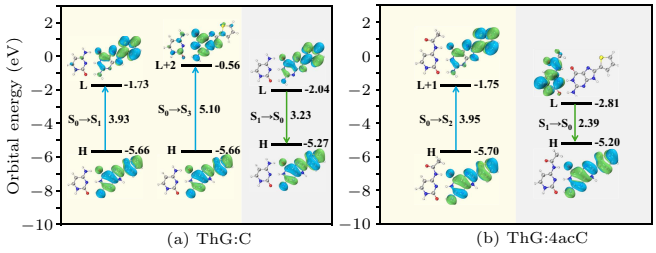


Fig. 4. View of the FMOs (isovalue = 0.02 a.u.) involved in the excitation and emission processes for ThG:C (a) and ThG:4acC (b). Green and blue signify the positive isosurface and negative isosurface, respectively. H: HOMO; L: LUMO.

For comparison, the effect of pairing with cytosine on the spectra characteristics is further investigated. As exhibited in

Fig. 2(b), ThG:C exhibits a sharp peak at 351 nm ($S_0 \rightarrow S_1$) with an oscillator intensity of $f_{abs.} = 0.70$, which is mainly involved in the HOMO→LUMO transition (99.0%). In S_1 state, the fluorescence emission spectra calculation of ThG:C shows that the peak is at 416 nm with an oscillator intensity $f_{flu.} = 0.78$ (Fig. 2(c)), which is consistent with the value of single ThG (408 nm, $f_{flu.} = 0.73$). As shown in Fig. 4(a), the $S_1 \rightarrow S_0$ transition is mainly associated with LUMO→HOMO (99.6%) and belongs to the $\pi^* \rightarrow \pi$ transition (bright state). The above discussion indicates that the pairing with cytosine has no visible effect on the fluorescence, whereas the fluorescence quenching induced by ESICT is observed after pairing with 4acC, so it can be used as an ultra-sensitive fluorescence probe for selective detection of 4acC. The relevant conclusions were further validated with PBE0-D3 and M062X-D3 functional to confirm the feasibility of the results, which has been successfully employed in similar systems (see Tables S3 and S4).[52,53]

Table 2. Absorption/fluorescence wavelengths (λ), oscillator strengths ($f_{abs.}$ and $f_{flu.}$), and state assignments for ThG:C and ThG:4acC. CI represents the composition index of the orbitals involved in the electronic transitions. Type and Gap represent the characteristic and energy gap of the orbital transitions, respectively.

Optical spectral	Base	Transition	λ (nm)	$f_{abs.}/f_{flu.}$	Assignment (CI/Type/Gap (eV))
Absorption	ThG:C	$S_0 \rightarrow S_1$	351	0.70	HOMO→LUMO (99.0%/$\pi\pi^*$/3.93)
		$S_0 \rightarrow S_3$	274	0.08	HOMO→LUMO+2 (94.0%/$\pi\pi^*$/5.10)
	ThG:4acC	$S_0 \rightarrow S_2$	348	0.71	HOMO→LUMO+1 (98.9%/$\pi\pi^*$/3.95)
Fluorescence	ThG:C	$S_1 \rightarrow S_0$	416	0.78	LUMO→HOMO (99.6%/$\pi\pi^*$/3.23)
	ThG:4acC	$S_1 \rightarrow S_0$	–	0.00	LUMO→HOMO (99.8%/$\pi\pi^*$/2.39)

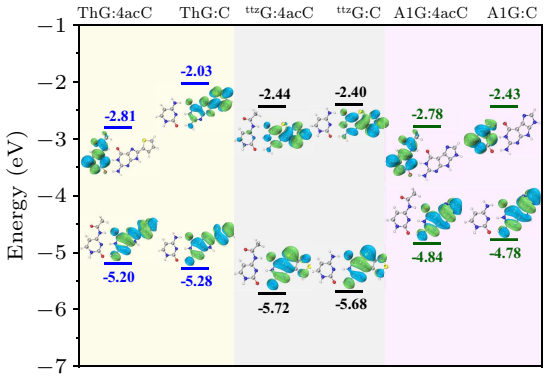


Fig. 5. Electron distribution and orbital energies involved in the emission processes of base pairs. ThG:4acC, ThG:C: blue solid line; ^{ttz}G:4acC, ^{ttz}G:C: black solid line; A1G:4acC, A1G:C: green solid line.

To illustrate the optical mechanism of these G-analogs pairing with 4acC and C, the electron distribution and orbital energies in S_1 state of two representative analogs (^{ttz}G and A1G) were investigated. As depicted by the blue solid lines in Fig. 5, the fluorescence of ThG:4acC and ThG:C has unique LUMO→HOMO contribution, and the phenomenon can be attributed to deactivation processes that the electron transfers from the 4acC part (LUMO) to the ThG moiety (HOMO). Furthermore, the HOMO/LUMO of ThG:C is predominantly located in the ThG region and is similar to those involved in S_1 state of single ThG, resulting in insensitivity to fluorescence after pairing with the C-base. However, the base pairs ^{ttz}G:4acC and ^{ttz}G:C remain efficiently fluorescent, whose decay process is mainly contributed by LUMO→HOMO, and the fluorescence is generated by the ^{ttz}G(π^*)→ ^{ttz}G(π) (bright states), which does not generate significant fluorescence changes to recognize 4acC in DNA (black solid line in Fig. 5). For the base pairs A1G:4acC and A1G:C, the fluorescence quenching is observed and the $S_1 \rightarrow S_0$ transition in them is primarily connected with the LUMO→HOMO, both of which are characterized by the emission from the LUMO orbital of pyrimidine (4acC/C) to the HOMO orbital of purine (A1G) (green solid lines in Fig. 5). The above discussions of fluorescence reaction fully suggest that ThG can serve as a quasi-intrinsic probe for detecting the presence or absence of 4acC.

3.3. The effects of linking to deoxyribose

With the intention of further demonstrating the biological applicability of the modified G-analogs in the double helix structure of DNA, the impact of binding with deoxyribose on photophysical properties is assessed. In accordance with the fully optimized geometries with linking to deoxyribose (Fig. 6(a)), the absorption and fluorescence spectra of base monomers and the relevant base pairs are simulated. As displayed in Fig. 6(b), the isolated ThdG displays a maximum absorption peak at 318 nm ($S_0 \rightarrow S_1$, $f_{abs.} = 0.56$), which is mainly contributed by HOMO→LUMO (98.6%) (Fig. S6) and possesses distinct blue shifts (∼ 23 nm) compared with the free ThG base. In the experiment of Rankin *et al.*,[54] the maximum absorption wavelength of ThdG is determined to be 284 nm in an aqueous solution, which is close to our calculated value of 318 nm. For emission, the fluorescence peak is red-shifted from 408 nm (ThG) to 409 nm (ThdG), corresponding to a negligible change in fluorescence strength ($\Delta f_{flu.} = 0.01$) (Fig. 6(c)), which is in good accordance with the experimental emission value of 414 nm.[54]

In respect to the base pairs (ThdG:d4acC and ThdG:dC), the electronic spectra are simulated at the optimized S_0 and S_1 structures, respectively (Fig. 6). As expected, the fluorescence of ThdG is still unaffected to base pairing with dC, whereas the obvious fluorescence quenching is observed to be

effective upon pairing with the d4acC. In detailed, the fluorescent emission peak of ThdG:dC is located at around 418 nm ($f_{\text{flu.}} = 0.79$), which is characterized by LUMO→HOMO (99.6%), corresponding to electronic transition of the $\pi^* \rightarrow \pi$ type (Fig. S6 and Table S5). In contrast, the fluorescence of ThdG:d4acC is totally quenched with the negligible oscillator intensity. As exhibited in Fig. S6, the $S_1 \rightarrow S_0$ transition of ThdG:d4acC is related with LUMO→HOMO (99.8%) and exhibits the distinct ESICT characteristic with the electron transition from the ThdG moiety (S_2) to d4acC part (S_1), which leads to the fluorescence quenching of base pairs. In addition, the hole–electron distribution graph and C_{hole}–C_{ele} graph defined by smooth description of the electron distribution are also utilized to evaluate the electronic excitation characteristics. This results intuitively suggest that the ThdG:d4acC undergoes the apparent intermolecular charge transfer, that is, the electron migrates from the ThdG base to the d4acC base during the $S_0 \rightarrow S_1$ transition process (Fig. 7), which provides further evidence for the charge transfer process by the ThdG as the electron donor and the d4acC base as the electron acceptor. Additionally, we also take into account factors such as acidity and lipophilicity in living organisms to ensure that the fluorescent ThdG molecule with an extended π-conjugation can efficiently enter the nucleus (SI-I). It follows that all the results obtained after deoxyribose connection are consistent with the above conclusions and thus the ThG base can function as a valid optical indicator to detect the 4acC bases in real biological environment.

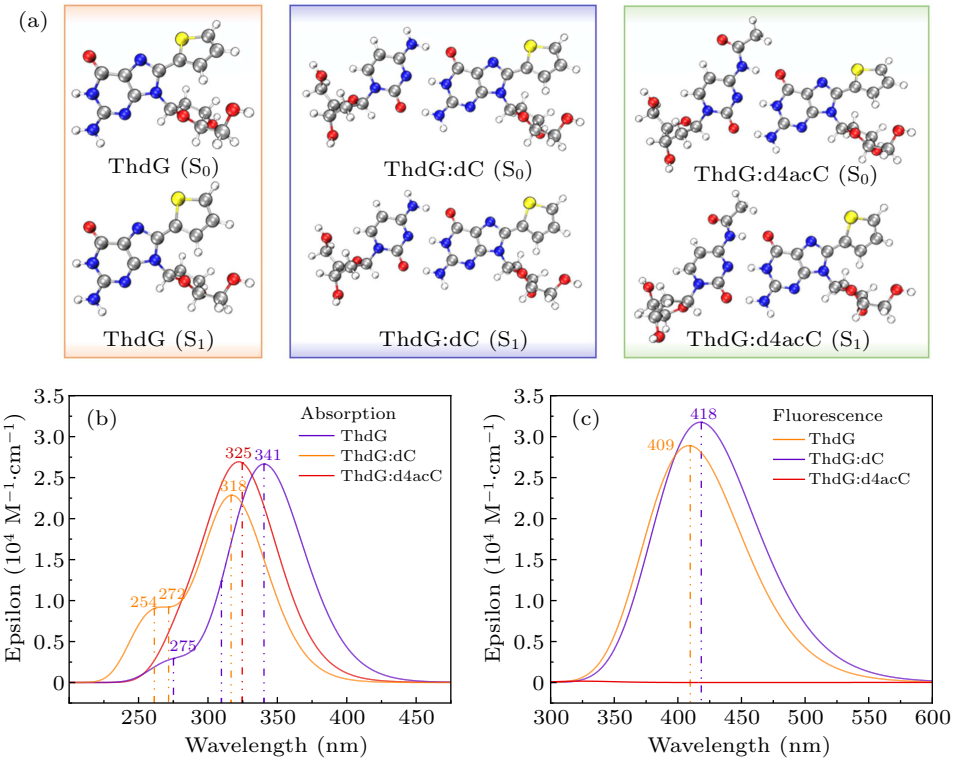


Fig. 6. Diagram of the ThdG, ThdG:dC, ThdG:d4acC structures (a), absorption spectra of ThdG monomer and corresponding base pairs (b), fluorescence spectra of ThdG monomer and corresponding base pairs (c). Atomic colors: white (H), gray (C), blue (N), red (O), yellow (S).

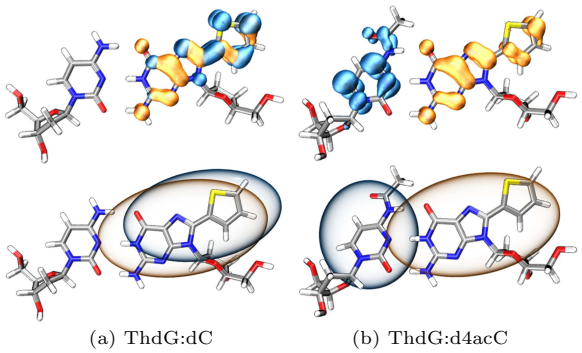


Fig. 7. Hole–electron distribution graph (isovalue = 0.002 a.u.) and C_{hole}–C_{ele} graph (isovalue = 0.0002 a.u.) of ThdG:dC (a) and ThdG:d4acC (b). Blue and orange zones signify the electron and hole distributions, respectively.

4. Conclusions

In this work, the photophysical properties of two types of modified fluorescent G-analogs, proposed as biosensor to ultra-sensitively detect the 4acC in real time, are theoretically investigated. These G-analogs retain the purine underlying architectures and WC hydrogen bond surfaces, which could guarantee the native structures and biological functions of nucleic acid with minimal disturbance. Compared with the native Guanine, these G-analogs possess extended π-conjugation that could yield efficient fluorescence emission and red-shifted absorption. Especially, the ThG exhibits the highest fluorescence intensity and avoids self-absorption on account of the large Stokes shifts (> 67 nm). Crucially, the fluorescence of

ThG remains unaffected by base pairing with natural cytosine, while the obvious fluorescence quenching is noticed by virtue of the ESICT after pairing with 4acC, which can be presented as a promising biosensor for monitoring fluorescence changes in the absence or presence of the 4acC. Furthermore, the impact of binding deoxyribose on photophysical properties is explored to ensure the biological applicability of these bright G-analogs in real environment. In summary, the proposed fluorescent analogs obtained through fluorophore modification and ring expansion benefit the DNA research field, which carries implications on comprehending biological interactions and mechanism of 4acC formation at the molecular level, and offering valuable insights into essential biological processes.

Acknowledgements

Project supported by the National Natural Science Foundation of China (Grant Nos. 12274265 and 11804195) and the Natural Science Foundation of Shandong Province, China (Grant No. ZR2022MA006).

References

[1] Chen K, Zhao B X S M and He C A 2016 *N Cell Chem. Biol.* **23** 74
[2] Lyu Y K, Chen S, Zhao Y, Yuan H X, Zhang C Y, Zhang C Z and Meng Q T 2024 *Phys. Chem. Chem. Phys.* **26** 12552
[3] Zhao L Y, Song J H, Liu Y B, Song C X and Yi C Q 2020 *Protein Cell* **11** 792
[4] Tsai K, Vasudevan A A J, Campos C M, Emery A, Swanstrom R and Cullen B R 2020 *Cell Host Microbe* **28** 306
[5] Hao H J, Liu W C, Miao Y J, Ma L, Yu B C, Liu L S, Yang C J, Zhang K, Chen Z, Yang J W, Zheng Z H, Zhang B, Deng F, Gong P, Yuan J H, Hu Z L and Guan W X 2022 *Nucleic Acids Res.* **50** 9339
[6] Stern L and Schulman L H 1987 *J. Biol. Chem.* **253** 6132
[7] Kumbhar B V, Kamble A D and Sonawane K D 2013 *Cell Biochem. Biophys.* **66** 797
[8] Yao F Y, Zhong F M, Jiang J Y, Cheng Y, Xu S, Liu J, Lin J, Zhang J, Li S Q, Li M Y, Xu Y M, Huang B and Wang X Z 2024 *Genes Dis.* **11** 993
[9] Fu Y, Luo G Z, Chen K, Deng X, Yu M, Han D L, Hao Z Y, Liu J Z, Lu X Y, Doré L C, Weng X C, Ji Q J, Mets L and He C 2015 *Cell* **161** 879
[10] Yang X, Yang Y, Sun B F, Chen Y S, Xu J W, Lai W Y, Li A, Wang X, Bhattarai D P, Xiao W, Sun H Y, Zhu Q, Ma H L, Adhikari S, Sun M, Hao Y J, Zhang B, Huang C M, Huang N, Jiang G B, Zhao Y L, Wang H L, Sun Y P and Yang Y G 2017 *Cell Res.* **27** 606
[11] Greenberg M V C and Bourc'his D 2019 *Nat. Rev. Mol. Cell Biol.* **20** 590
[12] Sas-Chen A, Thomas J M, Matzov D, Taoka M, Nance K D, Nir R, Bryson K M, Shachar R, Liman G L S, Burkhart B W, Gamage S T, Nobe Y, Briney C A, Levy M J, Fuchs R T, Robb G B, Hartmann J, Sharma S, Lin Q, Florens L, Washburn M P, Isobe T, Santangelo T J, Shalev-Benami M, Meier J L and Schwartz S 2022 *Nature* **583** 638
[13] Arango D, Sturgill D, Alhusaini N, Dillman A A, Sweet T J, Hanson G, Hosogane M, Sinclair W R, Nanan K K, Mandler M D, Fox S D, Zengeya T T, Andresson T, Meier J L, Coller J and Oberdoerffer S 2018 *Cell* **175** 1872
[14] Wang S, Xie H R, Mao F, Wang H Y, Wang S, Chen Z L, Zhang Y X, Xu Z H, Xing J M, Cui Z K, Gao X Q, Jin H M, Hua J, Xiong B and Wu Y F 2022 *Genome Biol.* **23** 5
[15] Zhou J X, Wang X, Wei Z, Meng J and Huang D Y 2022 *Mol. Ther. Nucl. Acids* **30** 337
[16] Chillar K, Yin Y P, Apostle A and Fang S Y 2023 *Org. Biomol. Chem.* **21** 9005
[17] Zhang J R, Jia Y, Xing X Y, Qin M M, Wu Z B, Zhong Y Q, Liu L L, Sun S Q, Li P, Wang H Y and Zhao G J 2021 *SmartMat* **2** 554
[18] Fang Y, Wang Q, Xiang C L, Liu G J and Li J J 2023 *Biosensors* **13** 610
[19] Zhang L B and Bu Y X 2008 *J. Phys. Chem. B* **112** 10723
[20] Dziuba D, Didier P, Ciaco S, Barth A, Seidel C A M and Mély Y 2021 *Chem. Soc. Rev.* **50** 7062
[21] del Nogal A W, Füchtbauer A F, Bood M, Nilsson J R, Wranne M S, Sarangamath S, Pfeiffer P, Rajan V S, El-Sagheer A H, Dahlén A, Brown T, Grotli M and Wilhelmsson L M 2020 *Nucleic Acids Res.* **48** 7640
[22] Karimi A, Wang K X, Basran K, Copp W and Luedtke N W 2023 *Bioconjugate Chem.* **34** 972
[23] Zhao Y, Cui X X, Song Y Z, Zhang C Z and Meng Q T 2021 *Spectrochim. Acta A* **260** 119926
[24] Duan L J, Zhang X, Zhao Y, Meng Q T and Zhang C Z 2023 *Phys. Chem. Chem. Phys.* **25** 3859
[25] Fadock K L and Manderville R A 2017 *ACS Omega* **2** 4955
[26] Saito Y, Taguchi H, Fujii S, Sawa T, Kida E, Kabuto C, Akaike T and Arimoto H 2008 *Chem. Commun.* **45** 5984
[27] Valverde D, de Araujo A V S, Borin A C and Canuto S 2017 *Phys. Chem. Chem. Phys.* **19** 29354
[28] Zhang L B, Wang M, Zheng M M and Kong X M 2019 *Int. J. Quantum Chem.* **119** e25870
[29] Zhang L B, Ren T Q, Tian J X, Yang X Q, Zhou L Z and Li X M 2013 *J. Phys. Chem. B* **117** 3983
[30] Kasha M 1950 *Discussions of the Faraday Society.* **9** 14
[31] Kumar A and Sevilla M D 2008 *J. Am. Chem. Soc.* **130** 2130
[32] Qi Y T, Wang Y, Tang Z, Liu J Y, Hou Y M, Gao Z Q, Tian J and Fei X 2020 *J. Mol. Liq.* **314** 113614
[33] Zhao G J, Yang Y F, Zhang C Y, Song Y Z and Li Y Q 2021 *J. Lumin.* **230** 117741
[34] Becke A D 1993 *J. Chem. Phys.* **98** 5648
[35] Goerigk L, Hansen A, Bauer C, Ehrlich S, Najibi A and Grimme S 2017 *Phys. Chem. Chem. Phys.* **19** 32184
[36] Zhan H B, Zhang H W, Jiang J J, Wang Y, Fei X and Tian J 2022 *Chin. Phys. B* **31** 038201
[37] Zhong Y Q, Chen Y, Feng X, Sun Y, Cui S, Li X Z, Jin X N and Zhao G J 2020 *J. Mol. Liq.* **302** 112562
[38] Zhao J F, Jin B and Tang Z 2022 *Phys. Chem. Chem. Phys.* **24** 27660
[39] Su S Y, Liang X N and Fang H 2022 *Chin. Phys. B* **31** 038202
[40] Mennucci B, Cances E and Tomasi J 1997 *J. Phys. Chem. B* **101** 10506
[41] Tomasi J, Mennucci B and Cammi R 2005 *Chem. Rev.* **105** 2999
[42] Yang Y F, Shi W, Chen Y P, Ma F C and Li Y Q 2021 *J. Lumin.* **229** 117698
[43] Yang Y F, Yang L J, Ma F C, Li Y Q and Qiu Y 2023 *Chin. Phys. B* **32** 057801
[44] Adamo C and Barone V 1999 *J. Chem. Phys.* **110** 6158
[45] Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Scalmani G, Barone V, Petersson G A and Nakatsuji H 2016 Gaussian 16 Revision A. 03 (Gaussian Inc., Wallingford CT)
[46] Liu Z Y, Lu T and Chen Q X 2020 *Carbon* **165** 461
[47] Humphrey W, Dalke A and Schulten K 1996 *J. Mol. Graph.* **14** 33
[48] Lu T and Chen F W 2012 *J. Comput. Chem.* **33** 580
[49] Andrienko G A 2016 Chemcraft 1.8. http://www.chemcraftprog.com (accessed 08/10/2016)
[50] Eid J, Fehr A, Gray J, *et al.* 2009 *Science* **323** 133
[51] van Duijneveldt F B, van Duijne-veldt-vande Rijdt J G C M and van Lenthe J H 1994 *Chem. Rev.* **94** 1873
[52] Seidu Y S, Roy H A and Rodgers M T 2021 *J. Phys. Chem. A* **125** 5939
[53] Green J A, Jouybari M Y, Aranda D, Improta R and Santoro F 2021 *Molecules* **26** 1743
[54] Rankin K M, Sproviero M, Rankin K, Sharma P, Wetmore S D and Manderville R A 2012 *J. Org. Chem.* **77** 10498