

Strategy on screening fluorescent guanine analogues to specifically detect 5-carboxylcytosine adduct

Mingming Zhang(张明明), Zhizheng Cao(曹智正), Xixi Cui(崔习习), Xiaolin Chen(陈晓琳), Changzhe Zhang(张常哲), and Qingtian Meng(孟庆田)

Citation: Chin. Phys. B, 2026, 35 (**5**): 057109. DOI: 10.1088/1674-1056/ae12d3
http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

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

ARPES study of Y_2O_2Bi single crystals: Intrinsic electronic structure of Bi square nets

Yun-Bo Wu(吴云波), Tong-Rui Li(李彤瑞), Zhi-Peng Cao(曹志鹏), Zhan-Feng Liu(刘站锋), Yu-Liang Li(李昱良), Zheng-Ming Shang(尚政明), Xin Zheng(郑新), Hui Tian(田慧), Zong-Yi Wang(王宗一), Yu-Tong Bi(毕雨桐), Hao-Yang Zhou(周浩洋), Yi Liu(刘毅), Guo-Bin Zhang(张国斌), Zheng-Tai Liu(刘正太), Da-Wei Shen(沈大伟), Li-Dong Zhang(张李东), Sheng-Tao Cui(崔胜涛), and Zhe Sun(孙喆)

Chin. Phys. B, 2026, 35 (**4**): 047101. DOI: 10.1088/1674-1056/ae04d9

Topological phases in nitrogen-doped chevron graphene nanoribbons

Yixuan Gao(高艺璇), Xinxi Zeng(曾新喜), and Ruizi Zhang(张瑞梓)

Chin. Phys. B, 2026, 35 (**3**): 037103. DOI: 10.1088/1674-1056/ae27b2

Electronic structure study of the intermetallic compound $GdRhIn_5$

Yulu He(贺宇璐), Bo Wang(王博), Xiangfei Yang(杨向飞), Dengpeng Yuan(袁登鹏), Wei Feng(冯卫), Yaobo Huang(黄耀波), and Qiuyun Chen(陈秋云)

Chin. Phys. B, 2026, 35 (**3**): 037101. DOI: 10.1088/1674-1056/ae3123

First-principles insights into $NaMgPO_3S$ oxysulfide solid electrolyte

Jian Sun(孙健), Shaohui Ding(丁少辉), Daquan Yang(杨大全), Kan Zhang(张侃), and Huican Mao(毛慧灿)

Chin. Phys. B, 2026, 35 (**2**): 027105. DOI: 10.1088/1674-1056/adf17d

First-principles calculations on strain tunable hyperfine Stark shift of shallow donors in Si

Zi-Kai Zhou(周子凯) and Jun Kang(康俊)

Chin. Phys. B, 2025, 34 (**8**): 087102. DOI: 10.1088/1674-1056/addeba

---

Strategy on screening fluorescent guanine analogues to specifically detect 5-carboxylcytosine adduct

Mingming Zhang(张明明), Zhizheng Cao(曹智正), Xixi Cui(崔习习), Xiaolin Chen(陈晓琳), Changzhe Zhang(张常哲), and Qingtian Meng(孟庆田)†

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

(Received 24 July 2025; revised manuscript received 24 September 2025; accepted manuscript online 14 October 2025)

The adduct 5-carboxylcytosine (5caC), one of the products involved in demethylation pathway, has the potential to serve as a target for specifically deciphering proteins that regulate gene expression and chromatin organization. The fluorescence detection with its ultra-high sensitivity can break through the limitations of traditional detection methods and achieve single molecule detection of 5caC. In this work, the guanine analogue G^{r3} with large π-conjugation and large Stokes shift (89.25 nm) is selected as the fluorescence probe to specifically recognize and monitor 5caC. Because of its bright fluorescence when paired with C and a complete fluorescence quenching when paired with 5caC, this on-off switching feature of G^{r3} stems from excited-state intermolecular charge transfer. Furthermore, it is found that after linking with deoxyribose, this discriminative function remains unchanged. This strategy could contribute to detecting the presence of 5caC in the genetic material, and thus prevent diseases associated with this epigenetic modification.

Keywords: density functional theory, fluorescent probe, photochemistry, guanine analogues, 5-carboxylcytosine (5caC)

PACS: 71.15.Mb, 33.50.–j, 82.50.–m

DOI: 10.1088/1674-1056/ae12d3 **CSTR:** 32038.14.CPB.ae12d3

1. Introduction

Epigenetic modification is a regulatory mechanism with significant impact on various aspects in gene expression. It entails the covalent attachment of specific chemical groups to DNA, resulting in alleles with identical sequences being in distinct modified states. Normal DNA modifications contribute to maintaining cellular growth and metabolism, whereas aberrant conformational changes in natural bases can not only inhibit the transcription of tumor suppressor genes, but also result in genomic instability, which can lead to the formation of DNA adducts.[1–3] Accordingly, it is necessary to study DNA adducts for understanding biological growth, development, and disease treatment.

The methylation of the 5th carbon site of cytosine, 5-methylcytosine (5mC),[4] is an epigenetic modification type with diverse potential for variation. 5mC is a reversible modification of mammalian genomic DNA. A corresponding study has proposed a method to remove methylation and restore normal cytosine. The ten-eleven translocation (TET) family of proteins initiates the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC), followed by sequential oxidation to 5-formylcytosine (5fC) and ultimately to 5caC.[5] Both 5fC and 5caC are recognized and excised by thymine DNA glycosylase (TDG), and subsequently reverted to canonical cytosine via the base excision repair (BER) pathway.[6] Due to its position at the end of the oxidation process, the steady-state concentration of 5caC is significantly lower than that of the direct precursor 5fC.[7]

However, 5caC does be involved in the evolution of breast cancer, liver cancer, lung cancer, pancreatic cancer, prostate cancer and other malignant tumors by regulating the transcription process.[8] It not only exists and functions in brain tissue, but also has a certain impact on the activity of mammalian neurons.[9] To have a more comprehensive study on the potential functions and value of 5caC, it is particularly important to develop methods for its specific identification or detection. Electroanalysis, chromatography and other methods have been widely used in the detection of adducts and the monitoring analysis of pharmaceuticals and foods.[10] Several challenges remain, including the complexity of experimental conditions, quantitative accuracy issues, and challenges in sample preparation and chromatographic column maintenance.[11] By comparison, fluorescent base-based molecular probe technology can play different roles in the recognition of various complexes through precise chemical design, significantly reducing the risk of system errors caused by column efficiency decay or ionization efficiency fluctuations in traditional methods.[12]

According to the differences in base modification strategies, fluorescent bases can be mainly divided into two categories: dye-modified nucleotides (e.g., nucleoside analogues labeled with anthocyanins)[13] and base modified fluorescent analogues (e.g., thioguanosine and extended pyrimidine derivatives).[14] Regarding the former, macromolecular nucleotides are widely used in experiments, including metal phthalocyanine class, rhodamine and other macromolecular flu-

†Corresponding author. E-mail: qtmeng@sdnu.edu.cn

© 2026 Chinese Physical Society and IOP Publishing Ltd. All rights, including for text and data mining, AI training, and similar technologies, are reserved.

orescent dye. Fluorescent dye labeled oligonucleotides have disadvantages such as significant influence of solvent polarity on fluorescence intensity, non-specific binding to various biomolecules (such as proteins and lipids), and background interference.[15] Macromolecular fluorophores are always used in vitro, and have the disadvantages of high cost, complicated operation, small detection flux, limited detection range, and contamination, which makes it difficult to monitor the presence of adducts such as 5caC.[16] Because of the inability to safely, harmlessly and endogenously detect adducts in real time, they are not the optimal choice for monitoring 5caC.

For the latter, our approach of using fluorescent base analogues to specifically recognize adducts provides a new one to accurately, sensitively, rapidly, and inexpensively detect adduct 5caC. It refers to the variation of photophysical properties due to the change of base charge distributions, i.e., by selecting suitable base analogues or modifying existing bases, it is possible to specifically regulate the charge transfer after binding to the corresponding base in DNA, thereby achieving the purpose of adjusting the relevant oscillator intensity.[17] The carboxylation at the 5th position of 5caC does not disrupt the linkage between 5caC and deoxyribose, thereby fulfilling the prerequisite for endogenous fluorescence detection of associated genes. Fluorescent bases have the characteristics of simplicity, rapidity, high selectivity and sensitivity, so they can be used as sensors to detect adducts in real time.

Because of the ability to avoid the problem of cytotoxicity and damage to normal cells caused by traditional drug treatment,[18] fluorescent probe, an analytical technique based on molecular recognition and photophysical signal conversion, is conducive to long-term research and treatment. It utilizes the fluorescence parameter changes (such as intensity, wavelength, lifetime, *etc*.) caused by the specific binding of fluorescent molecules to target substances to achieve high sensitivity and selectivity in target substance detection and imaging. In this regard, Saito and co-workers,[19,20] for instance, have developed several ring-expanded fluorescent nucleobase analogues, and they named them as base-discriminating fluoresecent nucleobases, which have been successfully implemented in the homogeneous detection of single nucleotide polymorphisms. Using the modification of fluorophore replacement based on the natural guanine G, Meng *et al*.[21] have designed a series of fluorescent nucleobase analogues, which were proposed as biosensors to ultra-sensitively detect the acetylated cytosine, a base that is highly correlated with active transcription and gene expression, in real time. In the field of single-molecule detection, a significant breakthrough has been made in the work such as the emergence of single pixel fluorescence labeling detection methods with high accuracy and low resource requirments,[22] and methods that can achieve small volume and high flux DNA sequencing through nanopores.[23] In recent years, our group has also published a series of articles on the specific recognition of adducts using fluorescent probes, such as the application of guanine analogue xG to specifically monitor the existence of O4 methylthymidine,[24] and utilizing thymine analogue T-R1 obtained by the fluorophore replacement to realize the specific recognition of Z, a nucleobase analogue of adenine.[25] These examples provide reference for the selection of fluorescent probes in this article to detect 5caC.

In the theoretical study of purine nucleoside derivatives, it was found that the configuration of G^{r3}, i.e., 8-(fur-2-yl) guanosine, adopts 8-H substituted guanosine derivatives with high quantum yield,[26] making G^{r3} a probe with research potential. Its specific recognition function for other bases is derived from the strict Watson–Crick (WC) base pairing rule in double stranded DNA. When there is normal or damaged cytosine, G^{r3} can form a relatively stable base pair with C or 5caC through hydrogen bonding. Accordingly, in this paper we investigated theoretically the photophysical properties of a series of probes, and the guanine analogue G^{r3} with large π conjugation and large Stokes shift is selected as the fluorescence base to specifically recognize and monitor the adduct 5caC.

The remaining part of the present work is arranged as follows: Section 2 systematically elaborates on the theoretical computational framework of this study, including the first principles computational method based on density functional theory (DFT), and its physical significance in system stability assessment and quantitative characterization of intermolecular interactions. In Section 3, we searched for 13 kinds of probes after a wide range of screening. Considering the planarity, wavelength difference, and oscillator intensity, as well as the fluorescence switching phenomenon we needed, the ideal probe G^{r3} was selected in the end. After the subsequent addition of deoxyribose and the use of other functional in verification processes, the qualitative results met expectations. At last, the summary and outlook are presented in Section 4.

2. Computational methods

The WC base pairs corresponding to the ground state (S_0) and the first excited state (S_1) of all considered deoxyribonucleosides were fully optimized using the density functional theory (DFT) method. Calculations were performed using the 6-311++G(d,p) basis set with the B3LYP-D3(BJ) functional. To reveal the bonding strength under the composite structure, we calculated the interaction energy as[27]

$$\Delta E = E_{\mathrm{AB}} - (E_{\mathrm{A}} + E_{\mathrm{B}}) + E_{\mathrm{BSSE}}. \tag{1}$$

Here the interaction energy (ΔE) is defined as the difference between the energy of the complex (E_{AB}) and the sum of the individual bases (E_{A}+E_{B}), and the addition of basis set superposition error (E_{BSSE}) energy helps with qualitative correctness. In order to analyze whether stable hydrogen bonds have been formed, we plotted reduced density gradient (RDG) scatter plots and contour maps, where RDG is a dimensionless parameter used in DFT to describe electronic non-uniformity,

and the corresponding expression can be read as[28,29]

$$S(r) = \frac{1}{2 \times (3\pi^2)^{1/3}} \frac{|\nabla \rho(r)|}{\rho(r)^{4/3}}, \tag{2}$$

in which ∇ is the gradient operator and $|\nabla \rho(r)|$ represents the modulus of the electron density gradient.

To consider the aqueous environment, the integral equation formalism for polarizable continuum model (IEFPCM)[30] was adopted, and the vibrational frequency analysis ensures that the obtained stable configurations are free of imaginary frequencies. Based on the optimized structure of the S_0 and S_1 states, the present work broadens the electronic absorption and emission spectra using Gaussian function

$$G(\omega) = \frac{1}{c\sqrt{2\pi}} \exp\left(-\frac{(\omega - \omega_n)^2}{2c^2}\right), \tag{3}$$

where $c = \frac{\mathrm{FWHM}}{2\sqrt{2\ln 2}}$ belonging to scale parameters, and the full width at half maximum (FWHM) is set to 0.4 eV, ω_n is a parameter representing the wavelength at the center position of the spectrum, ω represents a random variable in a normal distribution. All calculations in this study were performed using the Gaussian 16 package,[31] and the frontier molecular orbital diagrams (FMOs)[32] of monomers and base pairs were visualized by the Multiwfn software[33] combined with the VMD program.[34]

3. Results and discussion

Initially, we explored the photochemical properties of 5caC based on its optimized configuration. By analyzing the emission spectra, we found that the fluorescence oscillator intensity of 5caC is low and lacks effective intrinsic fluorescence emission, so it is difficult to detect the presence of this adduct in the human body (Table 1). For effectively inhibiting the potential epigenetic toxicity damage of 5caC, we detected its presence through photochemical phenomena. To avoid greater harm to the human body caused by the involvement of 5caC in subsequent transcription processes, we will pair 5caC with natural bases and modified bases separately during the DNA replication process.

Here, we conducted a systematic and comprehensive search for the related probes, involving 13 guanine analogues (see Fig. S1 in the supporting information), which were selected based on changing the substituent group[35] and have better fluorescence properties, or by expanding the ring.[36] We also set some screening criteria according to different needs, and finally retained the fluorescent probes with greater research value. In addition, we also paired the initially selected probes with normal cytosine as well as adducts to see if their absorption and emission spectra match the characteristics of the fluorescent switch.[37] The final probe that meets the expectation was attached to deoxyribose to make the constructed compounds more close to the real DNA configuration in the human body, then the generation and analysis of spectra after pairing were repeated, so that they can further provide a stronger theoretical basis for its application as a fluorescent base to detect defect.

3.1. Photophysical properties of fluorescent guanine analogues

Since the natural base pairs are generally in a plane, the first factor we consider was planarity, which can help to maintain the stability of the DNA double helix structure. The probes we selected or designed should follow this principle of planarity so that they can be accommodated in the base sequence of DNA (Fig. S1). Secondly, we examine if the oscillator intensity of the fluorescence emission spectrum has reached the expectation. After a comparative analysis of the spectra peaks for the excited state, we found that the oscillator intensities of the selected fluorescent probes are higher than 0.10, so we did not exclude any probes. Meanwhile, some probes with oscillator strength values greater than 1.00 have appeared, and their light is very bright and easy for instruments to recognize. In the third step, we consider the Stokes shift of monomers, which refers to the red shift of the fluorescence spectrum compared to the corresponding absorption spectrum. When the Stokes shift is larger, the self-absorption is smaller, making it easier to emit stronger fluorescence. If the redshift of the fluorescence spectrum with respect to the absorption spectrum is less than 50 nm, it is difficult to be recognized by the instrument, thereby causing great distress in the application.

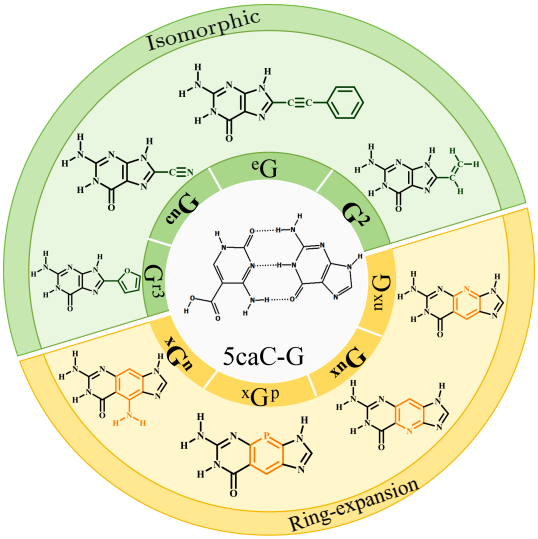


Fig. 1. Structural diagrams of the eight fluorescent probes that initially fit the screening scope.

Through using wavelength difference between the emission peak and the absorption peak, we found that the Stokes shifts of ^{6mi}G, G^{on}, G^{rn}, ^{3mi}G, and ^{s}G (Fig. S2)[38,39] do not satisfy the above requirement (see Table S1 in the supporting

information). Inevitably they were also excluded from the feasible research scope at this step. After the preliminary screening, we obtained 8 kinds of monomers which meet the expected requirements (Fig. 1). They show larger redshift and obvious photoluminescence compared with the natural guanine. In this study, the monomers used as fluorescent base analogues have longer wavelengths because longer wavelengths of light are less likely to scatter and can propagate deeper and more directly in tissues. At the same time, light with longer wavelengths can better avoid the absorption of major chromophores in biological tissues, thereby retaining more energy for deep penetration.

Then we paired this set of fluorescent analogues with normal cytosine and adduct 5caC, respectively, which reveals that they exhibit different photophysical properties (Fig. 2). The base analogues ^{xn}G, ^{x}G^{p}, ^{nx}G, G^{r3}, G^{2}, and ^{e}G[40–44] maintain bright luminescence when paired with C, whereas ^{x}G^{n}, and ^{cn}G[40–43] show quenching when paired with C because the oscillator intensity of the fluorescence spectra tends to be zero. Since what we are searching for is the base analogues whose pairing with C is fluorescent while with 5caC is fluorescence quenching, the latter two probes do not satisfy this "On to Off" prerequisite,[45] so they were eliminated from the selection range in this step.

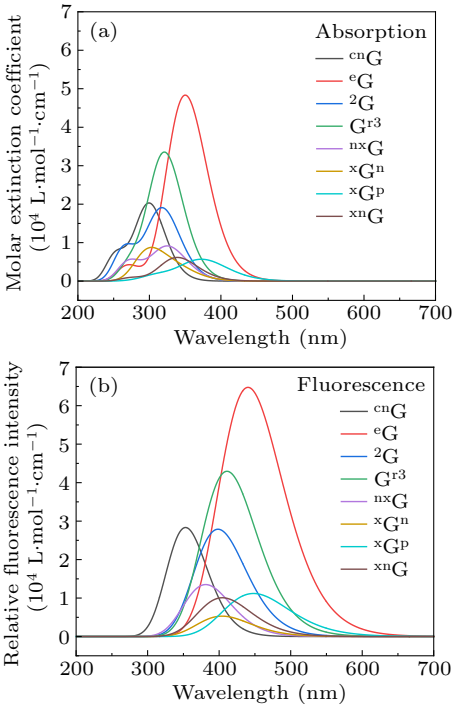


Fig. 2. Absorption spectrum (a) and emission spectrum (b) of eight fluorescent bases.

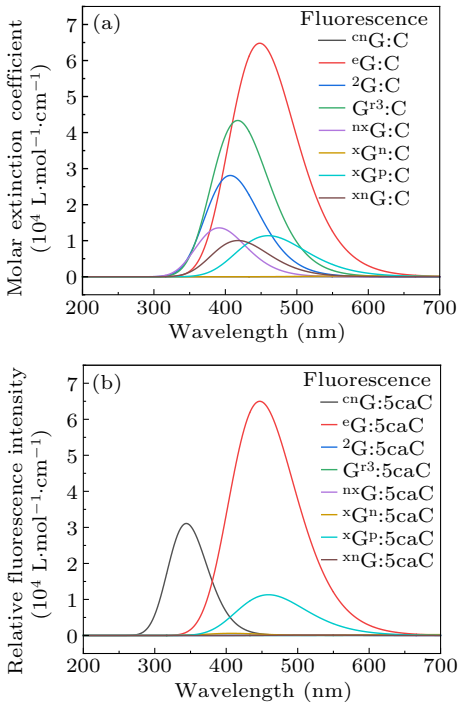


Fig. 3. Fluorescence spectra of eight fluorescent probes paired with cytosine C (a) and 5caC (b).

Table 1. Major transition contributions of monomers and base pairs using the B3LYP functional.

Optical spectral	Structure	Transition	λ	$f_{abs.}/f_{flu.}$	Assignment (CI/Gap (eV)/Type)
Absorption	G^{r3}	$S_0 \rightarrow S_1$	321.74	0.82	$H \rightarrow L$ (98.9%/4.22/$\pi\pi^*$)
	C	$S_0 \rightarrow S_1$	256.76	0.08	$H \rightarrow L$ (92.7%/5.49/$\pi\pi^*$)
	5caC	$S_0 \rightarrow S_1$	265.38	0.04	$H \rightarrow L$ (67.8%/5.27/$\pi\pi^*$)
	G^{r3}:C	$S_0 \rightarrow S_1$	334.27	0.41	$H \rightarrow L$ (98.4%/4.10/$\pi\pi^*$)
	G^{r3}:5caC	$S_0 \rightarrow S_2$	330.07	0.84	$H \rightarrow L{+}1$ (98.7%/4.13/$\pi\pi^*$)
Fluorescence	G^{r3}	$S_1 \rightarrow S_0$	410.99	1.06	$L \rightarrow H$ (99.6%/3.46/$\pi^*\pi$)
	C	$S_1 \rightarrow S_0$	334.96	0.06	$L \rightarrow H$ (97.3%/4.50/$\pi^*\pi$)
	5caC	$S_1 \rightarrow S_0$	382.05	0.01	$L \rightarrow H$ (97.2%/4.02/$\pi^*\pi$)
	G^{r3}:C	$S_1 \rightarrow S_0$	417.27	1.07	$L \rightarrow H$ (99.6%/2.97/$\pi^*\pi$)
	G^{r3}:5caC	$S_1 \rightarrow S_0$	632.51	0.00	$L \rightarrow H$ (99.8%/1.96/$\pi^*\pi$)

Figure 3 shows the photophysical phenomenon when pairing these eight fluorescent probes with adduct 5caC. Observing the values of the fluorescence spectra, it is found that the three probes (^{e}G, ^{cn}G, ^{x}G^{p}) do not quench because of their large fluorescence oscillator intensity, therefore they do not meet the rules for selection of fluorescent probes. Among the remaining viable probes, we found that G^{r3} has a Stokes shift of about 90 nm, and shows a large fluorescence oscillator in-

tensity ($f_{\text{flu.}} = 1.06$) in the fluorescence spectrum of the S_1 state (Table 1). Obviously, it follows the mechanism of real-time monitoring of compounds, and the detection limit is at the single-molecule level, making it the preferred choice.

3.2. Optical response for pairing with adduct 5caC and natural cytosine C

Based on the optimized configuration, we analyzed the spectral information about the transition from the ground state to the excited state of these ligands. From Table 1, we found that the oscillator intensity of the absorption and emission spectra of adduct 5caC decays to zero ($f_{\text{abs.}} = 0.04$, $f_{\text{flu.}} = 0.01$), which confirms that the adduct lacks the natural fluorescence, so it is difficult to be identified in the human body. Also found in this table is that the absorption and emission spectra of the natural cytosine have a low oscillator intensity ($f_{\text{abs.}} = 0.08$, $f_{\text{flu.}} = 0.06$), so we used the base analogues to explore their possibility to specifically recognize pathological base. In addition, it can be found that whether when monomerized or when paired with normal cytosine or adducts, the maximum absorption wavelength difference between G^{r3} monomer and ligands is small, and the absorption peaks are also large ($f_{\text{abs.}} \geq 0.41$).

It has been shown from Fig. 2 that the molar extinction coefficients[46] of the absorption spectra of G^{r3} monomer and ligands formed with cytosine are greater than 40000, corresponding to the electron cloud distribution of the $\pi \rightarrow \pi^*$ transition from S_0 state to S_1 state. In addition, the spectra of the bases we chose as our research subjects only have one peak, and the more special one is the absorption spectrum of G^{r3}:5caC, which corresponds to the transition from the ground state to the second excited state. Moreover, when we combine the information in Fig. 3 for analysis, it is found that the spectra have only one absorption peak, with the configuration index (CI)[47] being equal to 98.7%. When observing the emission spectra of the involved base pairs, it is found that both G^{r3} and G^{r3}:C have larger fluorescence oscillator intensity ($f_{\text{abs.}} \geq 1.06$), while the oscillator intensity of the fluorescence of G^{r3}:5caC is 0.0005. Obviously, the latter one has the fluorescence tending to quench, which demonstrates that there are very different optical phenomena for normal and diseased base during the formation of base pairs, and that G^{r3} could be used as quasi-instinctive fluorescent probe to selectively detect DNA adduct.

From Fig. 4, we found that the natural base guanine has a large orbital energy gap ($\Delta E_{\text{gap}} = 5.34$ eV), so it is more difficult for electron from the highest occupied molecular orbital (HOMO) to be excited into the lowest unoccupied molecular orbital (LUMO). On the other hand, the G^{r3} monomer has a small energy gap ($\Delta E_{\text{gap}} = 4.22$ eV) and CI is equal to 98.9%, so the transition is easy to occur from the ground state to the first excited state. By examining Table 1, we found that the energy gap of the electron excitation process of cytosine and 5caC is very large ($\Delta E_{\text{gap}} \geq 5.27$ eV), and the energy gap of the electron inactivation process is not small ($\Delta E_{\text{gap}} \geq 4.02$ eV). Since the transition probability is not only influenced by the transition selection rule, but also affected by the orbital gap value, the synergistic effect of this dual high band gap leads to an extremely low probability of electron transitions in the Franck–Condon region (oscillator strength $f \approx 0$), which significantly reduces its UV induced electron excitation efficiency.

FMOs can visualize the changes in the electron cloud distribution of the HOMO and the LUMO during the electron transition. Due to the Kasha rule,[48] the electrons in the excited state will relax from S_n to S_1. Therefore, we mainly analyzed the phenomenon from S_1 to S_0. From Fig. 4, we found that the electron clouds between S_1 and S_0 states of the modified guanine analogue G^{r3} have a larger overlap. The transition that takes place when electrons are inactivated from the excited state involves orbital transformation $\pi^* \rightarrow \pi$, in which the fluorescence emission occurs. While the electron cloud overlap between the ground state and excited state of guanine is relatively small, so the difference between them illustrates the superiority of the fluorescent base of G^{r3} over natural guanine. We compared the values of certain bases involved in our work with the values of bases of the same configuration from other literature, and found that the differences were small, which gives our work a high level of credibility. (S.I. D1)

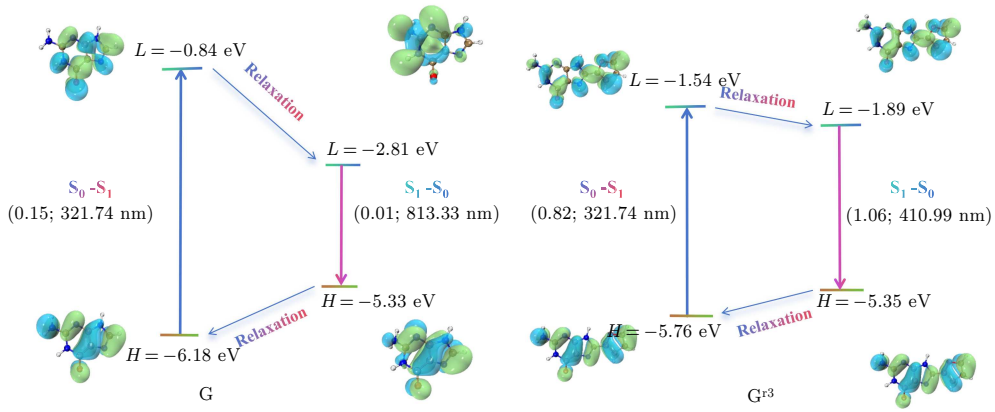


Fig. 4. FMOs (isovalue = 0.02) in the ground and excited states of guanine and G^{r3}.

From the photoexcitation process in Fig. 5, we can see that when G^{r3} is paired with cytosine, it has a much smaller orbital energy gap ($\Delta E_{gap} = 4.10$ eV). Subsequently, we analyzed the distribution and transfer of electron clouds after G^{r3} paired with these two different bases. When G^{r3} is paired with C, the area of the electron cloud distribution becomes larger from the HOMO to the LUMO during absorption of light ($f_{abs.} = 0.41$), then the inactivation of electrons from the excited state will lead to the electron clouds distributed in similar positions from LUMO to HOMO, resulting in a higher fluorescence oscillator intensity ($f_{flu.} = 1.07$). And when forming G^{r3}:5caC ligand, the energy gap is also low ($\Delta E_{gap} = 4.13$ eV), G^{r3} and 5caC form a conjugate enhancement system, which can significantly improve the degree of orbital hybridization. Through hybridization, the spatial distribution, energy level arrangement and symmetry matching of the electron cloud can be optimized to enhance the overlapping integration of ground state and excited state wave functions. As a result, the transition dipole moment is increased, and the transition is prone to occur. After photoexcitation ($S_0 \rightarrow S_2$), the excited state system (S_2) should dissipate energy to S_1 and then convert from S_1 to S_0 following the Kasha rule. During inactivation, electrons of G^{r3}:5caC return from the LUMO orbital to the HOMO orbital, in which the carboxyl group of 5caC (–COOH) acts as a strong electron attracting group that captures electrons on the G^{r3} expansion ring in the excited state, forming a charge transfer state of $G^{r3} \rightarrow$5caC, resulting in a fluorescence intensity close to zero.

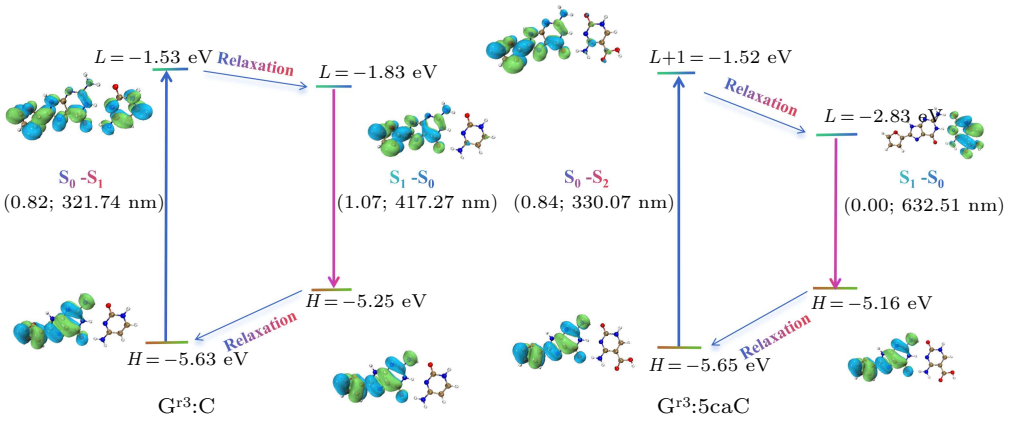


Fig. 5. FMOs (isovalue = 0.02) of G^{r3} pairing bases in the ground and excited states with C and 5caC.

When G^{r3} is paired with 5caC, corresponding to the absorption spectrum, the electron excited transition is from the ground state to the second excited state. So, there is an overlap of the electron clouds between the S_0 state and the S_2 state, creating a high absorption peak. During the excited state inactivation process, there is a large shift of the electron cloud from 5caC to G^{r3}, so there is a very low overlap between them, and thus the fluorescence quench phenomenon based on the excited state intermolecular charge transfer (ESICT)[49] is produced. We found that the quenching phenomenon corresponding to the case when the oscillator strength is 0 is due to a large shift of the electron cloud with almost no overlapping.

3.3. The role of deoxyribose linkage

To explore even further the feasibility of applying the fluorescent probe G^{r3}, we adopted the strategy of linking the relevant bases to deoxyribose fragments to form the nucleoside, which is closer to the DNA conformation of the human body. We performed a transition process analysis via Multiwfn (Fig. 6). After obtaining the major transition contribution by collating the information (Table 2), we plotted the FMOs for the excitation and inactivation processes (Fig. 7) and compared the similarities and differences with the previous correlated diagram and table.

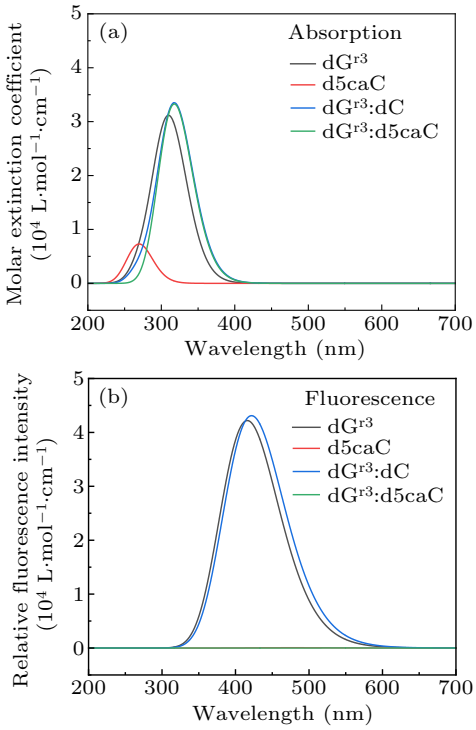


Fig. 6. Absorption spectrum (a) and emission spectrum (b) of nucleotides linked to deoxyribose.

Table 2. Major transition contributions after linking of deoxyribose under B3LYP generalization.

Optical spectra	Structure	Transition	λ	$f_{abs.}/f_{flu.}$	Assignment (CI/Gap (eV)/Type)
Absorption	dG^{r3}	$S_0 \rightarrow S_1$	311.39	0.74	$H \rightarrow L$ (98.6%/4.40/$\pi\pi^*$)
	dC	$S_0 \rightarrow S_1$	263.44	0.16	$H \rightarrow L$ (93.1%/5.38/$\pi\pi^*$)
	d5caC	$S_0 \rightarrow S_1$	271.78	0.16	$H \rightarrow L$ (79.2%/5.19/$\pi\pi^*$)
	dG^{r3}:dC	$S_0 \rightarrow S_1$	320.85	0.53	$H \rightarrow L$ (99.0%/4.28/$\pi\pi^*$)
	dG^{r3}:d5caC	$S_0 \rightarrow S_2$	318.08	0.81	$H \rightarrow L$+1 (98.8%/4.32/$\pi\pi^*$)
Fluorescence	dG^{r3}	$S_1 \rightarrow S_0$	416.10	1.04	$L \rightarrow H$ (99.6%/3.43/$\pi\pi^*$)
	dC	$S_1 \rightarrow S_0$	335.76	0.10	$L \rightarrow H$ (96.9%/4.49/$\pi\pi^*$)
	d5caC	$S_1 \rightarrow S_0$	476.84	0.00	$L \rightarrow H$ (99.3%/3.39/$\pi\pi^*$)
	dG^{r3}:dC	$S_1 \rightarrow S_0$	421.84	1.06	$L \rightarrow H$ (99.7%/3.39/$\pi\pi^*$)
	dG^{r3}:d5caC	$S_1 \rightarrow S_0$	624.86	0.00	$L \rightarrow H$ (99.8%/2.34/$\pi\pi^*$)

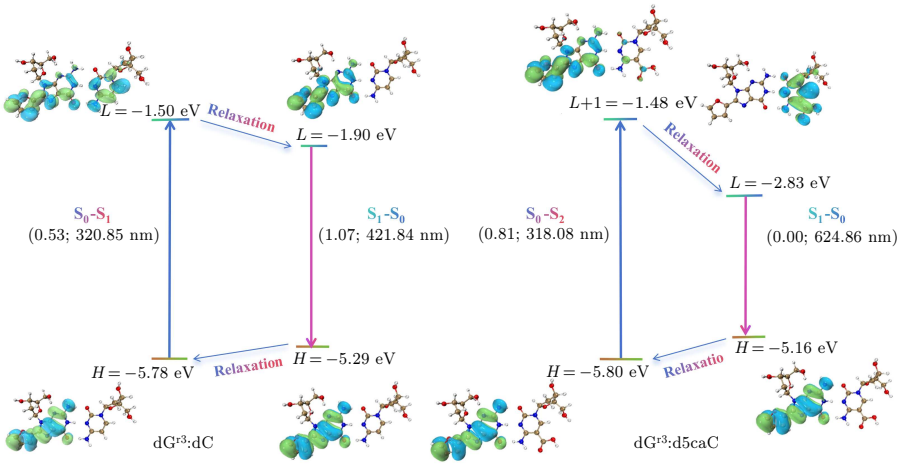


Fig. 7. FMOs of paired base under the B3LYP functional after addition of deoxyribose.

Subsequently, we investigated relative photophysical properties of these molecules based on the most stable configuration. From Table 2, we can find that after linking with deoxyribose, the d5caC monomer absorbs part of the light ($f_{abs.} = 0.16$), but during the inactivation process from the first excited state (S_1) to the ground state (S_0), there is almost no fluorescence ($f_{flu.} = 0.00$). The fluorescence intensity of the dC monomer is elevated compared to its state prior to the linking of deoxyribose ($\Delta f_{flu.} = 0.10$). Meanwhile, the maximum excitation wavelength of dG^{r3} is relatively large, so the emitted light is more likely to penetrate tissue cells.

The fluorescent base dG^{r3} has similar wavelength for the peak of the absorption ($\Delta\lambda_{abs.} = -10.35$ nm) and fluorescence spectra ($\Delta\lambda = 5.11$ nm) and less variation in the intensity of the absorption ($f_{abs.} = -0.08$) and emission spectra ($f_{flu.} = -0.02$) compared to G^{r3}. It shows that dG^{r3} and dG^{r3}:dC have an increase of the fluorescence emission spectra in comparison with the absorption spectra in oscillator intensity aspect. In the context of absorption spectra, dG^{r3}:dC exhibits a smaller value of maximum absorption wavelength, characterized by a slight blue shift ($\Delta\lambda = -13.42$ nm), and demonstrates a marginally higher oscillator strength ($\Delta f_{flu.} = 0.12$) relative to G^{r3}:C. Consistent with the phenomenon when no deoxyribose was added, dG^{r3} possesses brighter fluorescence when paired with dC, with a distinct peak near 421.84 nm ($S_1 \rightarrow S_0$, $f_{flu.} = 1.06$), and electrons mainly transfer from LUMO to HUMO (CI = 99.7%). Meanwhile, electrons migrate from the d5caC (LUMO) fragment to the dG^{r3} (HOMO) fragment, so dG^{r3}:d5caC exhibits ESICT phenomenon due to extensive charge transfer between excited states (Fig. 7, Table 2). The quenching phenomenon where the oscillator strength of dG^{r3} and d5caC pairs is close to zero strongly confirms that dG^{r3} can be used as a biosensor to identify the presence of 5caC.

Subsequently, to further validate the conclusions obtained, we repeated the operation of the above process using another functional PBE0 (D3).[50] We first compiled the absorption and emission spectra data of monomers and ligands (see Table S2). Secondly, we plotted the absorption and emission spectra after linking the deoxyribose (Fig. S6). Then we counted the main transition contributions of various base monomers and ligands connected to deoxyribose (Table S3). Finally, we obtained the FMO maps of G^{r3} paired with normal cytosine and adduct 5caC after adding deoxyribose (Fig. S7). By examining the spectral information of different functionals closely related to the above conclusion, we found that the qualitative conclusion is consistent with the description using B3LYP functionals. Although there may be slight fluctuations in the values, the overall trend of these two functionals is generally consistent, thus providing stronger evidence to support our above conclusion after this verification.

4. Conclusions

This work introduces novel fluorescent guanine analogue probes designed to exhibit superior optical characteristics for the detection of 5caC. Our calculations reveal that G^{r3} demonstrates pronounced fluorescence quenching behavior upon interaction with the carcinogenic adduct 5caC, facilitating the realization of endogenous monitoring within biological systems. Due to its reduced frontier orbital energy gap, G^{r3} exhibits a substantial redshift and a high fluorescence oscillator strength. This property allows to emit intense fluorescence when paired with native cytosine, thereby providing distinct experimental signatures that enable clear differentiation between the two nucleobases after forming ligands. By plotting the FMOs, we can intuitively find that the excited fluorescence oscillator strength of G^{r3}:5caC is based on the degree of overlap of electron cloud distribution and thus conforms to the ESICT feature of transition from the π^* LUMO to the π HOMO. So G^{r3} is a good probe suitable for 5caC monitoring due to its desirable photochemical properties. Compared with traditional methods such as chromatography, the G^{r3} fluorescent probe method not only has the single molecule level sensitivity, short detection time, but also can achieve early diagnosis and treatment with high safety. To more closely mimic the human DNA environment, we conducted additional investigations on the results of the linking of deoxyribose. We discovered that the qualitative relationship remains unchanged, and the maximum absorption wavelengths are very similar to the data of the unconnected deoxyribose original configuration. This work not only is expected to provide theoretical guidance for the design of quasi-intrinsic optical biosensors and contribute to understanding the mechanism of the covalent adducts formation, but also can enable better dynamic real-time studies of DNA structure, function and interactions through specific pairing, which can be useful for the design of self-assembled DNA materials.

Acknowledgements

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

References

[1] Lacagnina S 2019 *Am. J. Lifestyle Med.* **13** 165
[2] Yuan B F 2019 *Chem. Res. Toxicol.* **33** 695
[3] Tomkova M and Schuster-Böckler B 2018 *Trends Genet.* **34** 627
[4] Ito S, D'Alessio A C, Taranova O V, Hong K, Sowers LC and ZhangY 2010 *Nature* **466** 1129
[5] Ponnaluri V K C, Maciejewski J P and Mukherji M 2013 *Biochem. Biophys. Res. Commun.* **436** 115
[6] Palit D and Manna D 2024 *Inorg. Chem. Front.* **11** 7930
[7] Nanan K K, Sturgill D M, Prigge M F, Thenoz M, Dillman A A, Mandler M D and Oberdoerffer S 2019 *Iscience* **19** 326
[8] Chowdhury B, Cho I H, Hahn N and Irudayaraj J 2014 *Anal. Chim. Acta* **852** 212
[9] Zhang L J, Li Z J, Chen G Z, Huang Q L, Zhang J J, Wen J, Ye X X and Cai C 2016 *Anal. Methods* **8** 5812
[10] Abdel-Rehim M, Bielenstein M and Askemark Y 2003 *Anal. Chim. Acta* **492** 253
[11] Tjabadi E and Mketo N 2019 *TrAC - Trends Anal. Chem.* **118** 207
[12] Papp L A, Imre S, Bálint I, Lungu A, Mărcutiu P E, Papp J and Ion V 2024 *J. Chromatogr. Sci.* **62** 191
[13] Koshel B, Birdsall R and Chen W B 2020 *J. Chromatogr.* **1137** 121906
[14] Carlsson B, Karlsson L, Ärlemalm A, Sund S and Appell M L 2024 *Anal. Bioanal. Chem.* **416** 6711
[15] Wycisk V, Achazi K, Hirsch O, Kuehne C, Dernedde J, Hagg R and Licha K 2017 *ChemistryOpen* **6** 437
[16] Zhao Y S, Song Q Y, Lin Y B, Chu F, Wei Y C, Liu S, Pan C J, Quan L and Wang Y Q 2022 *J. Appl. Polym. Sci.* **139** 51625
[17] Chudinov M V 2020 *Chem. Heterocycl. Compd.* **56** 636
[18] Johns A C and Campbell M T 2022 *Cancer J.* **28** 469
[19] Okamoto A, Tainaka K and Saito I 2003 *J. Am. Chem. Soc.* **125** 4972
[20] Okamoto A, Tainaka K, Fukuta T and Saito I 2003 *J. Am. Chem. Soc.* **125** 9296
[21] Chen X L, Cui X X, Lyu Y K, Zhang C Y, Zhang C Z and Meng Q T 2025 *Chin. Phys. B* **34** 053102
[22] Chen H H and Lu C C 2021 *J. Comput. Biol.* **28** 195
[23] Myong S 2016 *Methods* **105** 1
[24] Zhang C Y, Zhao Y, Cui M L, Zhang C Z and Meng Q T 2023 *Spectrochim. Acta, Part A* **296** 122691
[25] Zhang X, Duan L J, Zhang Y N, Meng Q T and Zhang C Z 2023 *J. Mol. Liq.* **389** 122864
[26] Saito Y and Hudson R H E 2018 *J. Photochem. Photobiol., C* **36** 48
[27] Racioppi S, Sironi A and Macchi P 2020 *Phys. Chem. Chem. Phys.* **22** 24291
[28] Sharma R, Zhang J and Ohlin C A 2016 *Phys. Chem. Chem. Phys.* **18** 8235
[29] Demirpolat A, Akman F and Kazachenko A S 2022 *Mol.* **27** 6129
[30] Cossi M, Scalmani G, Rega N and Barone V 2002 *J .Chem. Phys.* **117** 43
[31] Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Scalmani G, Barone V, Mennucci B and Petersson G A in: Wallingford C T (ed.), Gaussian 16 Revision A.03, Gaussian Inc, 2013
[32] Fukuzawa K and Tanaka S 2022 *Curr. Opin. Struc. Biol.* **72** 127
[33] Lu T and Chen F W 2012 *J. Comput. Chem.* **33** 580
[34] Humphrey W, Dalke A and Schulten K 1996 *J. Mol. Graph.* **14** 33
[35] Benitah A, Qualley D F and Woski S A 2022 *Nucles. Nucleot. Nucl.* **41** 1187
[36] Fei N, Wang Z K, He P, Kang Q W, Wang Y, Wang C Y and Li Y Z 2024 *Org. Chem. Front.* **11** 6797
[37] Guille-Collignon M, Delacotte J, Lemaître F, Labbé E and Buriez O 2021 *Chem. Rec.* **21** 2193
[38] Moriya S, Shibasaki H, Kohara M, Kuwata K, Imamura Y, Demizu Y, Kurihara M, Kittaka A and Sugiyama T 2021 *Bioorg. Med. Chem. Lett.* **39** 127850
[39] Duan L J, Zhang X, Zhao Y, Meng Q T and Zhang C Z 2023 *Phys. Chem. Chem. Phys.* **25** 3859
[40] Bucardo M S and Wu Y 2021 *ACS Chem. Biol.* **16** 1208
[41] Cong D, Li Y, Ludford III P T and Tor Y 2022 *Chem. Eur. J.* **28** e202200994
[42] Riensch N A, Swoboda L, Lik A, Krummenacher I, Brauschweig H and Helten H 2021 *Z. Anorg. Allg. Chem.* **647** 421
[43] Qi P, Yi M J, Song A X and Hao J C 2022 *Langmuir* **38** 7099
[44] Steinbuch K B, Bucardo M and Tor Y 2024 *ACS Omega* **9** 36778
[45] Sivadas D K, Ravi S, Nagarasu P, Johnson D, Thiruvenkatam V, Anthony S P and Madh V 2023 *Mol. Syst. Des. Eng.* **8** 1462
[46] Alrahili M 2024 *Russ. J. Phys. Chem. B* **18** 343
[47] Hoyer C E, Lu L, Hu H, Shumilov K D, Sun S C, Knecht S and Li X S 2023 *J. Cleaner Prod.* 158
[48] Xie M C, Cai J H, Zhang G Y, Jung S Y, Dong D F, Zhang Z Y, Zhou D Y Liao L S and Li T 2023 *Mater. Chem. Front.* **7** 4092
[49] Bhakta V, Chatterjee S, Bhattacharyya A and Guchhait N 2024 *J. Phys. Chem. A* **128** 8914
[50] Blaško M, Mach P, Antušek A and Urban M 2018 *J. Phys. Chem. A* **122** 4591