



## Tautomers of a Fluorescent G Surrogate and Their Distinct Photophysics Provide Additional Information Channels

This is the peer reviewed version of the following article:

*Original:*

Sholokh, M., Improta, R., Mori, M., Sharma, R., Kenfack, C., Shin, D., et al. (2016). Tautomers of a Fluorescent G Surrogate and Their Distinct Photophysics Provide Additional Information Channels. ANGEWANDTE CHEMIE. INTERNATIONAL EDITION, 55(28), 7974-7978 [10.1002/anie.201601688].

*Availability:*

This version is available <http://hdl.handle.net/11365/995751> since 2018-05-14T09:11:45Z

*Published:*

DOI:10.1002/anie.201601688

*Terms of use:*

Open Access

The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. Works made available under a Creative Commons license can be used according to the terms and conditions of said license.

For all terms of use and more information see the publisher's website.

(Article begins on next page)



# HHS Public Access

Author manuscript

*Angew Chem Int Ed Engl.* Author manuscript; available in PMC 2017 January 04.

Published in final edited form as:

*Angew Chem Int Ed Engl.* 2016 July 4; 55(28): 7974–7978. doi:10.1002/anie.201601688.

## Tautomers of a Fluorescent G Surrogate and Their Distinct Photophysics Provide Additional Information Channels

**Marianna Sholokh**<sup>#</sup>

Laboratoire de Biophotonique et Pharmacologie Faculté de Pharmacie, UMR 7213 CNRS, Université de Strasbourg 74 route du Rhin, 67401 Illkirch (France)

Department of Chemistry Kyiv National Taras Shevchenko University 60 Volodymyrska street, 01033 Kyiv (Ukraine)

**Dr. Roberto Improta**<sup>#</sup>

Consiglio Nazionale delle Ricerche Istituto di Biostrutture e Bioimmagini Via Mezzocannone 16, 80134 Napoli (Italy)

**Dr. Mattia Mori**<sup>#</sup>

Dipartimento di Biotecnologie, Chimica e Farmacia Università degli Studi di Siena Via Aldo Moro 2, 53100 Siena (Italy)

Center for Life Nano Science@Sapienza Istituto Italiano di Tecnologia viale Regina Elena 291, 00161 Roma (Italy)

**Rajhans Sharma**

Laboratoire de Biophotonique et Pharmacologie Faculté de Pharmacie, UMR 7213 CNRS, Université de Strasbourg 74 route du Rhin, 67401 Illkirch (France)

**Dr. Cyril Kenfack**

Laboratoire de Biophotonique et Pharmacologie Faculté de Pharmacie, UMR 7213 CNRS, Université de Strasbourg 74 route du Rhin, 67401 Illkirch (France)

Laboratoire O'Optique et Applications, Centre de Physique Atomique Moléculaire et Optique Quantique, Université de Douala BP 8580 Douala (Cameroon)

**Dr. Dongwon Shin**

Department of Chemistry and Biochemistry University of California, San Diego 9500 Gilman Dr, La Jolla, CA 92093-0358 (USA)

**Dr. Karine Voltz**

Department of Integrative Structural Biology, Institut de Génétique de Biologie Moléculaire et Cellulaire, INSERM U964 UMR 7104 CNRS, Université de Strasbourg, 67400 Illkirch (France)

**Dr. Roland H. Stote**

Department of Integrative Structural Biology, Institut de Génétique de Biologie Moléculaire et Cellulaire, INSERM U964 UMR 7104 CNRS, Université de Strasbourg, 67400 Illkirch (France)

<sup>#</sup> yves.mely@unistra.fr, ytor@ucsd.edu.

Supporting information for this article can be found under: <http://dx.doi.org/10.1002/anie.201601688>.

**Prof., Dr. Olga A. Zaporozhets**

Department of Chemistry Kyiv National Taras Shevchenko University 60 Volodymyrska street, 01033 Kyiv (Ukraine)

**Prof., Dr. Maurizio Botta**

Dipartimento di Biotecnologie, Chimica e Farmacia Università degli Studi di Siena Via Aldo Moro 2, 53100 Siena (Italy)

**Prof., Dr. Yitzhak Tor\***

Department of Chemistry and Biochemistry University of California, San Diego 9500 Gilman Dr, La Jolla, CA 92093-0358 (USA)

**Prof., Dr. Yves Mély\***

Laboratoire de Biophotonique et Pharmacologie Faculté de Pharmacie, UMR 7213 CNRS, Université de Strasbourg 74 route du Rhin, 67401 Illkirch (France)

# These authors contributed equally to this work.

**Abstract**

Thienoguanosine (<sup>th</sup>G) is an isomorphous nucleoside analogue acting as a faithful fluorescent substitute of G, with respectable quantum yield in oligonucleotides. Photophysical analysis of <sup>th</sup>G reveals the existence of two ground-state tautomers with significantly shifted absorption and emission wavelengths, and high quantum yield in buffer. Using (TD)-DFT calculations, the tautomers were identified as the H1 and H3 keto-amino tautomers. When incorporated into the loop of (-)PBS, the (-)DNA copy of the HIV-1 primer binding site, both tautomers are observed and show differential sensitivity to protein binding. The red-shifted H1 tautomer is strongly favored in matched (-)/(+)PBS duplexes, while the relative emission of the H3 tautomer can be used to detect single nucleotide polymorphisms. These tautomers and their distinct environmental sensitivity provide unprecedented information channels for analyzing G residues in oligonucleotides and their complexes.

**Keywords**

ab initio calculations; fluorescence; molecular modeling; nucleic acids; tautomerism

---

*The structure, acid-base features, and tautomeric equilibria of the canonical and non-canonical nucleobases found in nucleic acids have been the subject of intense investigation for decades.<sup>[1]</sup> While the role of minor tautomers in mutagenesis has been one of the primary foci,<sup>[2]</sup> recent observations suggest that such isomeric nucleobases also play key roles in regular nucleic acid structure and function.<sup>[3]</sup> As the population of distinct tautomeric forms is impacted by their micro-environment, this added level of complexity also provides opportunities to further advance our understanding of nucleic acid structure and dynamics.*

In this context, emissive nucleoside analogues, which have become powerful biophysical tools,<sup>[4]</sup> provide unique prospects. A tautomerizable nucleoside analogue, where the tautomers would have distinct absorption and emission spectra, could be instrumental for

investigating the micro-environment of a nucleobase with greater insight compared to tautomeric stable probes. Herein we analyze the photophysics of thienoguanosine, <sup>th</sup>G, a highly useful G surrogate,<sup>[5]</sup> and identify two environmentally sensitive ground-state tautomeric forms (Figure 1) which display distinct absorption and emission spectra. The equilibrium between the two tautomers is mainly governed by the hydrogen-bond donor properties of the solvent. Their observed sensitivity to the microenvironment was rationalized by ab initio calculations. By exploring single- and double-stranded <sup>th</sup>G-containing oligonucleotides, as well as DNA–protein complexes, we illustrate that this probe provides compelling biophysical information and greater insight compared to monochromatic or ratiometric fluorescent nucleosides.

The emission spectra of either <sup>th</sup>G or d<sup>th</sup>G in water and methanol are surprisingly complex.<sup>[5a,b]</sup> Excitation at both  $\lambda=360$  and  $380$  nm gives a similar emission spectrum (Figure 2a, orange) centered at  $\lambda=468$  nm. When the excitation energy is progressively increased, a blue-shifted emission with a maximum at  $\lambda=400$  nm appears and becomes dominant for excitation below  $\lambda=300$  nm (Figure 2a, magenta and blue). The simplest interpretation is that <sup>th</sup>G exhibits two ground-state species with shifted emission spectra. This hypothesis is supported by recording excitation spectra for various emission wavelengths (see Figure S1 in the Supporting Information). Since sample purity was rigorously maintained, the two ground-state species likely correspond to two tautomers, differing by their excitation and emission spectra. This conclusion is highly likely, since tautomers have also been observed for guanosine itself.<sup>[6]</sup>

Spectral deconvolution yields well separated emission and absorption spectra of the individual tautomers in buffer (Figure 2b,c). Thus, by judiciously selecting the excitation and emission wavelengths, each tautomer can be individually excited and observed. The fluorescence quantum yield (QY) is found to be constant ( $0.49\pm 0.03$ ) over a large range of excitation wavelengths ( $\lambda=290$ – $375$  nm) and close to earlier reported values,<sup>[5a,c]</sup> thus indicating that the two forms possess similar QY values. The individual absorption spectrum of the red-shifted form (Figure 2c) is very similar to the spectra of <sup>th</sup>G and d<sup>th</sup>G in 1,4-dioxane.<sup>[5a,b]</sup> By using the molar extinction coefficient of <sup>th</sup>G in 1,4-dioxane ( $\epsilon^{333}=4530\text{M}^{-1}\text{cm}^{-1}$ ),<sup>[5a]</sup> it is possible to calculate the concentration of the red-shifted tautomer in buffer and deduce the proportion (44%) and the molar extinction coefficient ( $\epsilon^{313}\approx 4600\text{M}^{-1}\text{cm}^{-1}$ ) of the blue-shifted tautomer. Importantly, the red-shifted tautomer, excited at  $\lambda=350$ – $380$  nm, is highly photostable (see Figure S2). In contrast, extended illumination at higher energies (e.g.,  $\lambda=325$  nm), where both tautomers absorb, show continuously diminished emission, thus suggesting that the blue-shifted tautomer is less photostable and that the two tautomers are equilibrating.

The spectroscopic properties of <sup>th</sup>G were comparatively characterized in various solvents (see Table S1). In methanol, ethanol, and *n*-butanol (see Figures S3 and S4) spectra comparable to those in buffer are obtained, thus indicating that <sup>th</sup>G also exists in the two tautomeric forms in these solvents. In all other tested solvents, the emission spectra are independent on the excitation wavelength and indicate that only the red-shifted tautomer is present (see Figure S3c). Its emission maximum correlates well with the empirical polarity index  $E_T(30)$  of the solvent (see Figure S5). Interestingly, although *N,N*-

dimethylformamide, acetonitrile, and methanol are all rather polar [ $E_T(30) > 43$  kcal mol<sup>-1</sup>,  $\epsilon > 32$ ], the blue-shifted tautomer is seen only in methanol, thus strongly suggesting hydrogen-bonding stabilization. This proposal is further substantiated by the deconvoluted absorption spectra in polar protic solvents (see Figure S6), as they show that the relative concentration of the blue-shifted isomer linearly increases with solvent proticity (see Figure S6d and Table S1). Thus, the equilibrium between the two <sup>th</sup>G tautomers is dependent on the hydrogen-bond donor ability of the solvent.

To assist in identifying the two emissive isomers, the ground-state energy minima of five hypothetical <sup>th</sup>G tautomers were optimized (Figure 3) in the gas phase, 1,4-dioxane, and water at the DFT level, by using PBE0 and M052X functionals and including solvent effects with the Polarizable Continuum Model (PCM; see Tables S2 and S3). The keto-amino <sup>th</sup>G–H1 tautomer (Figure 3a) appears largely dominant over the other tautomers (Figures 3b–e), with the exception of water where it is only 0.11 eV more stable than <sup>th</sup>G–H3, when including only bulk solvent effects. Therefore, the two <sup>th</sup>G keto-amino tautomers, analogous to the most stable tautomer of guanine in solution,<sup>[7]</sup> are likely populated in water.

Independent of the inclusion of one solvent molecule in the computational model (see Figures S7 and S8, and Table S2), the <sup>th</sup>G–H1 tautomer appears as the main contributor to the observed spectroscopic properties of <sup>th</sup>G in 1,4-dioxane and is therefore assigned to the red-shifted isomer.<sup>[8]</sup> The <sup>th</sup>G–H1 and <sup>th</sup>G–H3 tautomers are found to be almost isoenergetic in water when solute–solvent hydrogen bonds are considered (see Figure S7 and Table S2). These data suggest that both tautomers likely contribute to the spectroscopic properties of <sup>th</sup>G in water and that the blue-shifted isomer corresponds to the <sup>th</sup>G–H3 tautomer. The computed energy difference between the two tautomers (<0.05 eV, that is, <400 cm<sup>-1</sup>) is beyond the expected accuracy of our method, thus explaining why the molar fraction of the <sup>th</sup>G–H3 tautomer (see Table S2) does not perfectly match with the experimental value (0.44).<sup>[9]</sup> The computed vertical absorption and emission energies (see Table S4)<sup>[10]</sup> indicate that, independent of the functional, the lowest-energy transition in water for both tautomers corresponds to a bright  $\pi\pi^*$  S<sub>0</sub>→S<sub>1</sub> transition attributed to a HOMO→LUMO excitation.<sup>[11]</sup>

Interestingly, small differences in the shape of the frontier orbitals involved in the electronic transition (see Figure S9) result in fairly large differences in the computed vertical excitation energy, so that the absorption maximum of the <sup>th</sup>G–H1 tautomer in water is red-shifted by 30–40 nm, with respect to that of <sup>th</sup>G–H3. The absorption maxima predicted for the two forms, namely  $\lambda = 330 \sim 350$  nm (depending on the solvation model) for <sup>th</sup>G–H1 and  $\lambda = 300 \sim 310$  nm for <sup>th</sup>G–H3, are very close to the experimental ones (see Figures S10a, 2c and Table S1).<sup>[12]</sup> TD-DFT excited-state geometry optimizations (employing either PBE0 or M052X) predict a stable S<sub>1</sub> minimum for both tautomers in all examined solvents.<sup>[13]</sup> This S<sub>1</sub> minimum is characterized by a fairly large oscillator strength, of about 80% of the value computed for absorption. This minimum contrasts guanosine, for which the same functionals predict a barrierless decay to S<sub>0</sub>, through an effective conical intersection.<sup>[14]</sup> Both <sup>th</sup>G tautomers therefore show promising electronic features with potentially robust emissive states. The computed emission wavelengths of  $\lambda = 448$  and 383 nm for the <sup>th</sup>G–H1 and <sup>th</sup>G–H3 tautomers, respectively,<sup>[15]</sup> (see Figure S10b), are in good agreement with the

spectroscopic data in buffer (Figure 2b). Moreover, by weighting the contribution of the different tautomers with a simple Boltzmann equation, the computed fluorescence spectra (see Figure S10b) are consistent with the experimental ones. Taken together, PCM/(TD)DFT calculations indicate that the <sup>th</sup>G–H1 and <sup>th</sup>G–H3 tautomers are responsible of the observed photophysical features of <sup>th</sup>G.

To examine the ground-state equilibrium between the tautomers of <sup>th</sup>G in oligonucleotides, the DNA equivalent of the 18-mer primer binding site of HIV-1 was selected as a biologically relevant model (see Figure S11). It forms a stem-loop of known three-dimensional structure<sup>[16]</sup> and is involved in the second strand transfer of HIV-1 reverse transcription.<sup>[17]</sup> Deoxy-<sup>th</sup>G (d<sup>th</sup>G), which exhibits spectroscopic properties very similar to <sup>th</sup>G,<sup>[5b,c]</sup> substitutes the dG7 residue in the loop [labeled d<sup>th</sup>G7(–)PBS; Figure 4a inset]. Comparing the emission spectra at various excitation wavelengths clearly shows that both d<sup>th</sup>G tautomers are present in the (–)PBS loop (Figure 4a). In contrast, when d<sup>th</sup>G7(–)PBS is annealed to its complementary (+)PBS strand (see Figure S11), forming the d<sup>th</sup>G7(–)/(+)PBS duplex (Figure 4c, inset), the normalized emission spectra obtained at different excitation wavelengths all overlap, thus indicating that the d<sup>th</sup>G–H1 tautomer is predominant in the double-stranded form (Figure 4c). Although not attributed to the two tautomers disclosed here, a similar switch from a two-band to a single-band emission was previously observed upon transition from single- to double-stranded structures in model <sup>th</sup>G- and d<sup>th</sup>G-labeled sequences,<sup>[5a,b]</sup> thus indicating that the tautomeric shift reported here is not unique for (–)PBS.

Distinct behavior was seen for mismatched duplexes between d<sup>th</sup>G7(–)PBS and complementary (+)PBS oligonucleotides, where d<sup>th</sup>G was placed opposite A, T, or G (Figure 4d). In contrast to the fully complementary duplex, where the d<sup>th</sup>G–H3 tautomer is nearly absent, its relative contribution as estimated by the ratio of the fluorescence intensities at  $\lambda=375$  and 550 nm,  $I_{375}/I_{550}$ , increases by a factor of three and five in the mismatched duplexes with opposite dG and dA, respectively (see Table S5). For the mismatched duplex with opposite dT, the difference with the matched duplex is marginal, but the two duplexes can be easily discriminated by the twofold difference in their quantum yield (see Table S5). This difference likely results both from a change in polarity (as suggested by the changes in the positions of the d<sup>th</sup>G–H1 emission maximum; see Table S5) and in the quenching by the flanking nucleobases, as a result of the different geometries adopted by d<sup>th</sup>G and its neighbors in the two duplexes. The relative emission of the two d<sup>th</sup>G tautomers and the d<sup>th</sup>G quantum yield are therefore highly sensitive to the nature of the opposite base and can thus be used in combination to detect single nucleotide polymorphisms.

To further illustrate the potential applications of the two spectrally distinct d<sup>th</sup>G tautomers when in oligonucleotides, we investigated their response to binding of the HIV-1 nucleocapsid NC(11-55) peptide to the (–)PBS loop.<sup>[16b]</sup> Titration with NC(11-55) protein resulted in a strong increase of the d<sup>th</sup>G–H1 peak of d<sup>th</sup>G7(–)PBS, but only a marginal increase in the d<sup>th</sup>G–H3 peak (shift in the H3/H1 emission ratio from 1.1 to 0.8), thus indicating that the relative emission of the two tautomers is sensitive to protein binding (Figure 4b). As NC(11-55) was reported to direct the G7 base toward the exterior of the

loop<sup>[16b]</sup> and restrict its collisions with the neighboring bases,<sup>[18]</sup> it appears that the d<sup>th</sup>G–H1 tautomer is more sensitive than d<sup>th</sup>G–H3 to these changes.

To shed light on the biophysical observations, molecular dynamics (MD) simulations using the ff12SB AMBER force field were performed on the NMR structure of P(–)PBS DNA,<sup>[16b]</sup> a truncated form of (–)PBS lacking the 3′ overhang (see Figure S11). Analysis of MD trajectories (0.2 μs of unbiased MD trajectory production) and thermodynamic parameters unequivocally shows that there are no differences in the behavior of either dG or the two d<sup>th</sup>G tautomers in two representative P(–)PBS structures (see Figures S12, S13a,b, and S14). In contrast, analysis of local motion within the Watson–Crick base pair established by either dG or d<sup>th</sup>G at position 7 in the (–)/(+)PBS DNA duplex clearly shows that d<sup>th</sup>G–H1 has the same behavior as dG, whereas d<sup>th</sup>G–H3 pairs with the counterpart dC with lower stability (see Figure S13d). A local structural analysis of MD trajectories further confirms that d<sup>th</sup>G–H1 forms the three canonical hydrogen-bonds with dC as observed for guanine (Figure S15a,b), while d<sup>th</sup>G–H3 contacts dC in multiple non-canonical complexes (see Figures S15c and S16). Overall, and consistently with experimental observations, the replacement of dG with d<sup>th</sup>G–H3 in the (–)/(+)PBS DNA duplex is noticeably less favorable than the replacement with d<sup>th</sup>G–H1, from a thermodynamic and conformational viewpoint.<sup>[19]</sup> Finally, MD simulations reveal that the two tautomers are mainly in the *anti*-conformation in both the stem-loop and the duplex (see Table S6).

In summary, through a careful analysis of its spectroscopic properties as a free nucleoside and when incorporated into oligonucleotides, thienoguanosine <sup>th</sup>G was observed to exhibit two ground-state tautomers with significantly shifted absorption and emission spectra. Quantum mechanical calculations unambiguously identified the two tautomers as being the keto-amino tautomers, <sup>th</sup>G–H1 and <sup>th</sup>G–H3. MD studies further suggested <sup>th</sup>G–H1 behaves similarly to its native counterpart in both the single- and double-stranded structures studied here, whereas the <sup>th</sup>G–H3 tautomer behaves comparably to G only in the loop of a stem-loop DNA. When incorporated into double-stranded sequences, <sup>th</sup>G–H3 tautomerizes to the favorable and benign <sup>th</sup>G–H1 tautomer, which forms a stable Watson–Crick base pair. The ratio of the two tautomers and their relative emission were found to be highly sensitive to the nucleic acid strandedness, to the nature of the opposite base in DNA duplexes, as well as to protein binding. The tautomerism of the isomorphous <sup>th</sup>G, which is associated with distinct and highly emissive states, thus constitutes a highly useful additional channel of information that provides an unprecedented window into features of substituted G residues in oligonucleotides.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The work was supported by a fellowship from the Ministère de la Recherche (M.S.), the European Project THINPAD “Targeting the HIV-1 Nucleocapsid Protein to fight Antiretroviral Drug Resistance” (FP7-Grant Agreement 601969), Agence Nationale de la Recherche (ANR blanc Fluometadn and FEMTOSTACK), Agence Nationale de Recherche sur le SIDA, French-Ukrainian Dnipro program, the Université de Strasbourg, the Centre National de la Recherche Scientifique (CNRS), the Institut de la Santé et de la Recherche Médicale (INSERM),

Progetto Bilaterale CNR/CNRS, and the US National Institutes of Health (GM 069773). Computing time was provided at the French national computing centers by GENCI (Grand Equipement National de Calcul Intensif) and the Meso-center for High Performance Computing at the Université de Strasbourg and supported by the project EQUIP@MESO.

## References

- [1]. a) Miles HT. Proc. Natl. Acad. Sci. USA. 1961; 58:791. [PubMed: 13770642] b) Wong YP, Wong KL, Kearns DR. Biochem. Biophys. Res. Commun. 1972; 49:1580. [PubMed: 4344814] c) Lee GC, Chan SI. J. Am. Chem. Soc. 1972; 94:3218. [PubMed: 5032498] d) Robinson H, Gao YG, Bauer C, Roberts C, Switzer C, Wang AHJ. Biochemistry. 1998; 37:10897. [PubMed: 9692982] e) Blas JR, Luque FJ, Orozco M. J. Am. Chem. Soc. 2004; 126:154. [PubMed: 14709079]
- [2]. a) Drake JW, Baltz RH. Annu. Rev. Biochem. 1976; 45:11. [PubMed: 786147] b) Topal MD, Fresco JR. Nature. 1976; 263:285. [PubMed: 958482] c) Shugar D, Kierdaszuk B. Proc. Int. Symp. Biomol. Struct. Interactions, Suppl. J. Biosci. 1985; 8:657. d) Wang WN, Hellenga HW, Beese LS. Proc. Natl. Acad. Sci. USA. 2011; 108:17644. [PubMed: 22006298]
- [3]. a) Westhof E. FEBS Lett. 2014; 588:2464. [PubMed: 24950426] b) Singh V, Fedeles BI, Essigmann JM. RNA. 2015; 21:1. [PubMed: 25516996] c) Kimsey I, Al-Hashimi HM. Curr. Opin. Struct. Biol. 2014; 24:72. [PubMed: 24721455]
- [4]. a) Sinkeldam RW, Greco NJ, Tor Y. Chem. Rev. 2010; 110:2579. [PubMed: 20205430] b) Tanpure AA, Pawar MG, Srivatsan SG. Isr. J. Chem. 2013; 53:366. c) Phelps K, Morris A, Beal PA. ACS Chem. Biol. 2012; 7:100. [PubMed: 22148740] d) Wilhelmsson LM, Rev Q. Biophys. 2010; 43:159. e) Wilson JN, Kool ET. Org. Biomol. Chem. 2006; 4:4265. [PubMed: 17102869] f) Hawkins ME. Methods Enzymol. 2008; 450:201. [PubMed: 19152862]
- [5]. a) Shin D, Sinkeldam RW, Tor Y. J. Am. Chem. Soc. 2011; 133:14912. [PubMed: 21866967] b) Park S, Otomo H, Zheng L, Sugiyama H. Chem. Commun. 2014; 50:1573. c) Sholokh M, Sharma R, Shin D, Das R, Zaporozhets OA, Tor Y, Mely Y. J. Am. Chem. Soc. 2015; 137:3185. [PubMed: 25714036]
- [6]. a) Colominas C, Luque FJ, Orozco M. J. Am. Chem. Soc. 1996; 118:6811. b) Nir E, Janzen C, Imhof P, Kleinermanns K, de Vries MS. J. Chem. Phys. 2001; 115:4604. c) Marian CM. J. Phys. Chem. A. 2007; 111:1545. [PubMed: 17274608]
- [7]. Lee YJ, Jang YH, Kim Y, Hwang S. Bull. Korean Chem. Soc. 2012; 33:4255.
- [8]. Interestingly, the stability of the <sup>th</sup>G–H3 tautomer increases with polarity and particularly with the hydrogen-bonding ability of the solvent.
- [9]. For the examined solvents, PBE0 and M052X provide similar indications, thus suggesting that our conclusions are robust with respect to the choice of the functional. Inclusion of vibrational effects does not substantially affect the conformational equilibrium between <sup>th</sup>G–H1 and <sup>th</sup>G–H3 tautomers.
- [10]. Computed at the PCM/TD-PBE0 and PCM/TD-M052X level.
- [11]. a) Samanta PK, Manna AK, Pati SK. J. Phys. Chem. B. 2012; 116:7618. [PubMed: 22671305] b) Samanta PK, Pati SK. New J. Chem. 2013; 37:3640.
- [12]. In addition to the relative position of the lowest energy peak, the general shape of the absorption spectra of the <sup>th</sup>G–H1 and <sup>th</sup>G–H3 tautomers was found to be very close to experimental ones, thus supporting their assignment.
- [13]. a) Gedik M, Brown A, Photochem J. Photobiol. A. 2013; 259:25. b) Samanta PK, Pati SK. Phys. Chem. Chem. Phys. 2015; 17:10053. [PubMed: 25785569]
- [14]. a) Karunakaran V, Kleinermanns K, Improta R, Kovalenko SA. J. Am. Chem. Soc. 2009; 131:5839. [PubMed: 19341282] b) Improta R. Chem. Eur. J. 2014; 20:8106. [PubMed: 24828154]
- [15]. According to PCM/6-311+G(2d,2p) calculations on a <sup>th</sup>G-6H<sub>2</sub>O model.
- [16]. a) Johnson PE, Turner RB, Wu ZR, Hairston L, Guo J, Levin JG, Summers MF. Biochemistry. 2000; 39:9084. [PubMed: 10924101] b) Bourbigot S, Ramalanjaona N, Boudier C, Salgado GFJ, Roques BP, Mély Y, Bouaziz S, Morellet N. J. Mol. Biol. 2008; 383:1112. [PubMed: 18773912]
- [17]. Darlix JL, Godet J, Ivanyi-Nagy R, Fossé P, Mauffret O, Mély Y. J. Mol. Biol. 2011; 410:565. [PubMed: 21762801]

- [18]. Godet J, Ramalanjaona N, Sharma KK, Richert L, De Rocquigny H, Darlix JL, Duportail G, Mély Y. *Nucleic Acids Res.* 2011; 39:6633. [PubMed: 21543454]
- [19]. To rule out any possible bias by the selected force field used for the MD simulations (AMBER ff12SB force field) similar computations were performed using the CHARMM all-atom force field for nucleic acids. The results of these calculations are highly comparable and are detailed in the Supporting Information.

Author Manuscript

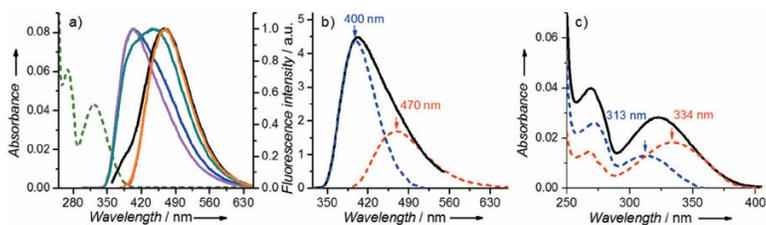
Author Manuscript

Author Manuscript

Author Manuscript



**Figure 1.** Structures of guanosine (G) and the two emissive tautomers of thienoguanosine (<sup>th</sup>G-H1 and <sup>th</sup>G-H3). (d)Rib=D-ribose or 2'-deoxy-D-ribose.

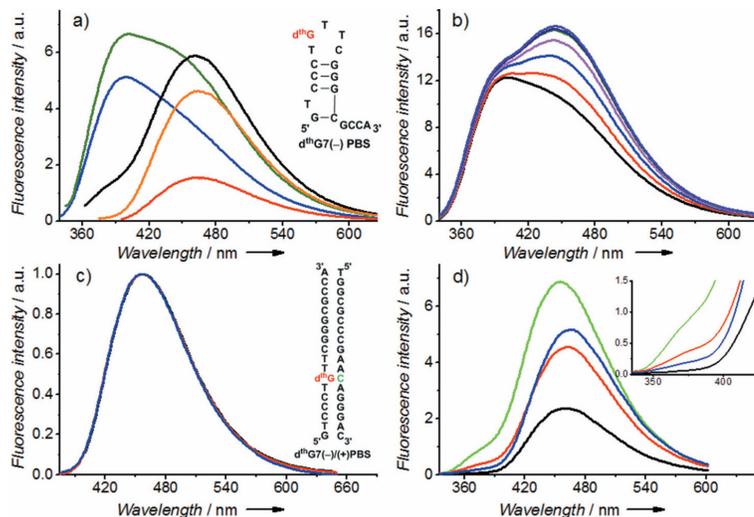


**Figure 2.**

Absorption and emission spectra of <sup>th</sup>G in TRIS-HCl buffer 25 mM, pH 7.5, 30 mM NaCl, 0.2 mM MgCl<sub>2</sub>. a) Absorption (green dashed line) and emission spectra of <sup>th</sup>G at different excitation wavelengths:  $\lambda=283$  nm (magenta line);  $\lambda=298$  nm (blue);  $\lambda=320$  nm (green);  $\lambda=345$  nm (black); and  $\lambda=360$  nm (orange). The emission spectra were normalized at their maxima. The normalized emission spectrum at  $\lambda_{exc}=380$  nm fully overlaps that at  $\lambda_{exc}=360$  nm and is not shown. b) Deconvoluted emission spectrum of <sup>th</sup>G, obtained at  $\lambda_{exc}=283$  nm. c) Deconvolution of the absorption spectrum of <sup>th</sup>G (black line) in its two ground-state forms (colors as in b).



**Figure 3.** Schematic drawing of the <sup>th</sup>G tautomers which have been calculated: a) keto-amino <sup>th</sup>G-H1, b) keto-amino <sup>th</sup>G-H3, c) enol-amino, d) keto-imino, and e) enol-imino.



**Figure 4.**

Emission spectra of  $d^{\text{th}}\text{G7}(-)\text{PBS}$  (a, b) and  $d^{\text{th}}\text{G7}(-)/(+)\text{PBS}$  duplexes (c, d). a) Emission spectra of  $d^{\text{th}}\text{G7}(-)\text{PBS}$  recorded at different excitation wavelengths:  $\lambda=298$  nm (blue),  $\lambda=320$  nm (green),  $\lambda=345$  nm (black),  $\lambda=360$  nm (orange), and 380 nm (red). Inset: structure of  $d^{\text{th}}\text{G7}(-)\text{PBS}$ , the G7 residue (red) is replaced by  $d^{\text{th}}\text{G}$ . b) Emission spectra of  $d^{\text{th}}\text{G7}(-)\text{PBS}$  in the absence (black) and in the presence of 1 to 6 equivalents of NC(11-55) protein (red to violet) at  $\lambda=320$  nm excitation wavelength. c) Emission spectra of the matched  $d^{\text{th}}\text{G7}(-)/(+)\text{PBS}$  duplex at the same excitation wavelengths as in (a). Inset: structure of  $d^{\text{th}}\text{G7}(-)/(+)\text{PBS}$  duplex. In mismatched duplexes, the C residue in green is replaced by a A, G or T. d) Emission of the matched and mismatched  $d^{\text{th}}\text{G7}$ -labeled  $(-)/(+)\text{PBS}$  duplexes at  $\lambda=310$  nm excitation wavelength. The base opposite to  $d^{\text{th}}\text{G}$  is C (black, native duplex), T (blue), G (red), or A (green). Inset: zoom of the blue part of the spectra. The buffer was as in Figure 2.