

Article

Calculations of pK_a Values for a Series of Fluorescent Nucleobase Analogues

Sun Jeong Im , Alan J. Mlotkowski , H. Bernhard Schlegel  and Christine S. Chow *

Department of Chemistry, Wayne State University, Detroit, MI 48202, USA; sunim@wayne.edu (S.J.I.); mlotkowski@wayne.edu (A.J.M.); hbs@chem.wayne.edu (H.B.S.)

* Correspondence: cchow@wayne.edu

Abstract

Nucleobases play diverse structural and functional roles in biological systems. Understanding the fundamental properties of nucleobases is important for their applications as chemical probes of nucleic acid function. As the nucleobases are modified to tune their fluorescence or binding properties, their physical properties such as pK_a may also change. Unlike the canonical nucleobases, modified nucleobases are less well understood in terms of their acid-base properties. Previously, theoretical pK_a values of canonical, naturally modified, and aza-/deaza-modified nucleobases were determined. In this study, the theoretical pK_a values for 25 different fluorescent modified nucleobases (55 total pK_a values) were calculated by using an ab initio quantum mechanical method employing the B3LYP density functional with 6-31+G(d,p) basis set along with an implicit–explicit solvation model. The results of these computations are compared to known experimental pK_a values. The ability to estimate theoretical pK_a values will be beneficial for further development and applications of fluorescent nucleobases.

Keywords: pK_a ; fluorescent nucleoside; nucleobase modification



Academic Editor: Giacomo Dacarro

Received: 18 September 2025

Revised: 14 October 2025

Accepted: 17 October 2025

Published: 22 October 2025

Citation: Im, S.J.; Mlotkowski, A.J.; Schlegel, H.B.; Chow, C.S. Calculations of pK_a Values for a Series of Fluorescent Nucleobase Analogues. *Compounds* **2025**, *5*, 44.

<https://doi.org/10.3390/compounds5040044>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Nucleobases typically contain one or more proton donor and acceptor sites, which are involved in the formation of hydrogen bonds with neighboring nucleobases (Figure 1) [1]. Hydrogen bonding plays a significant role in DNA and RNA, stabilizing base pairs and ultimately impacting both local and global structure of the nucleic acid [2]. Understanding the acid-base properties of the varying functional groups of modified nucleobases is crucial for determining their ability to undergo hydrogen-bonding interactions and ultimately assessing their impact on folded, biologically relevant nucleic acid structures [3].

The canonical nucleobases have emissive properties, but the emission quantum yields are remarkably low compared to most fluorophores [4]. Decades of research have led to the development of nucleosides with modified physical and chemical properties, including nucleosides and nucleobases with enhanced emission capabilities [5–9]. The fluorescent analogues often have altered base-pairing and base-stacking interactions within the folded nucleic acid structure due to their modified structures [10–12]. One of the first reported fluorescent nucleosides is 2-aminopurine (2AP) [5], which has been used in numerous applications to study nucleic acid structure and/or small-molecule interactions [13–15]. This adenosine (A) analogue has the amino group at C2 instead of C6 (Figure 2) and demonstrates a significant increase in the emission quantum yield compared to A in aqueous conditions [16]. In addition, 2AP retains the ability to base pair with thymidine (T)

and uridine (U), or form altered wobble pairs with cytidine (C) [17,18], making it useful for various nucleic acid applications such as probing DNA or RNA conformational changes or interactions with other molecules [14,17,19]. Other examples of purine fluorescent nucleoside analogues are shown in Figures 2 and 3, with atom substitutions or scaffold extensions with functional group additions, conjugations, and ring fusions.

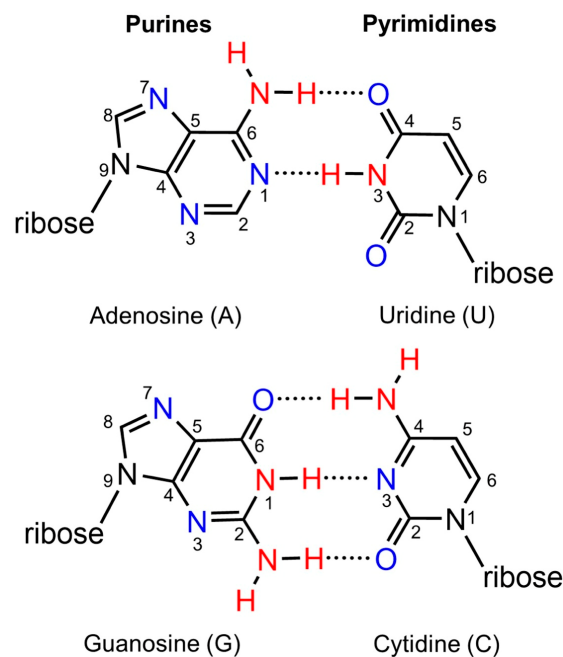


Figure 1. The four canonical nucleobase structures. The nucleobase structures of purines (A and G) and pyrimidines (U and C) are shown with the corresponding numbering system. Proton donors and acceptors are indicated in red and blue, respectively.

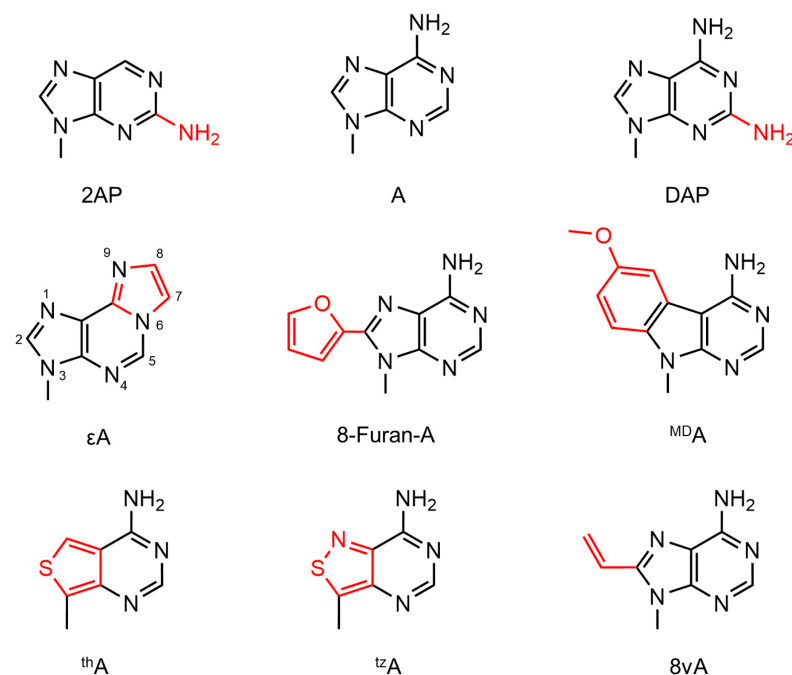


Figure 2. Structures of adenosine-derived fluorescent nucleobases. Differences from the canonical A structure are highlighted in red. The purine analogues employ the standard purine numbering system, except for ϵ A.

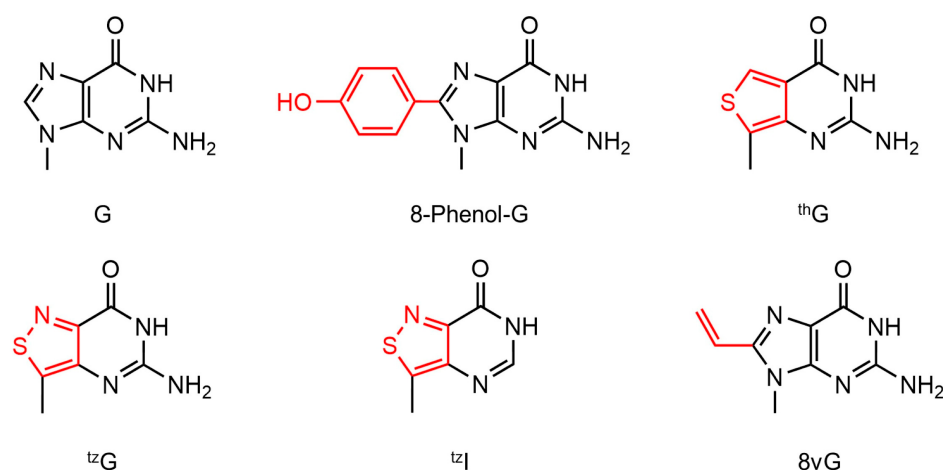


Figure 3. Structures of guanosine-derived fluorescent nucleobases. Differences from the canonical G structure are highlighted in red.

Another example of a fluorescent nucleobase is 5-methyl-2-pyrimidinone (m^5K), which is a pyrimidine analogue (Figure 4). Unlike canonical nucleobases, m^5K does not readily base pair and is useful for studying base-stacking interactions in single-stranded DNAs [7]. Thiophene-fused or thiophene-substituted U (^{th}U or 5- ^{th}U , respectively, Figure 5) with optimized emission quantum yields have been used in RNA oligonucleotides to monitor mismatched base pairs with C [20,21]. There are fewer examples of pyrimidine fluorescent nucleobase modifications compared to purines, but many involve ring expansions or conjugations (Figures 4 and 5).

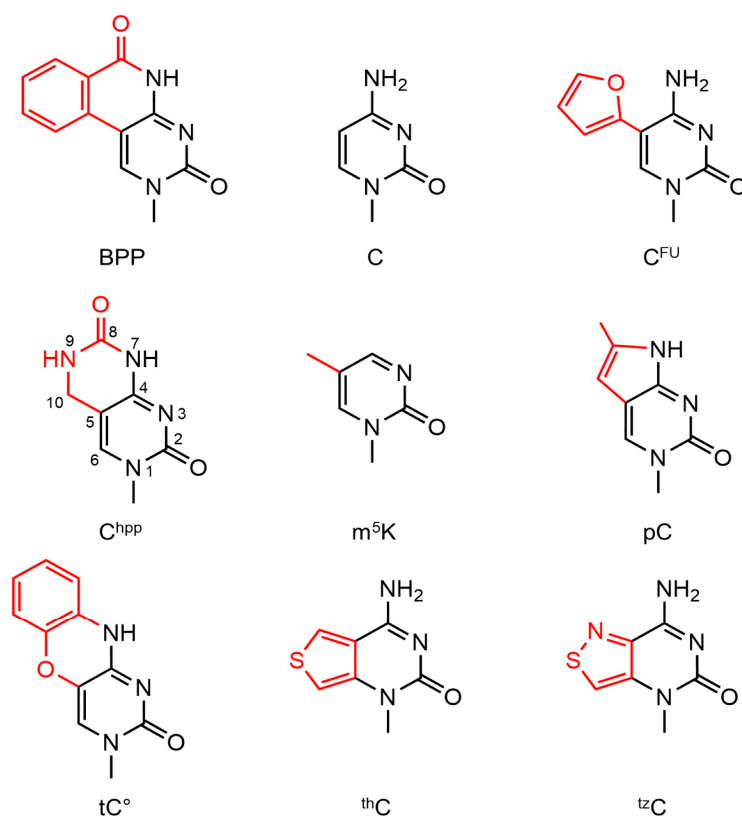


Figure 4. Structures of cytidine-derived fluorescent nucleobases. Substitutions to the canonical C structure are highlighted in red. The analogues shown employ the same numbering system as the canonical pyrimidine, except for the ring-fused analogues, pC, tC[°], and BPP, which follow the numbering system indicated for C^{hPP}.

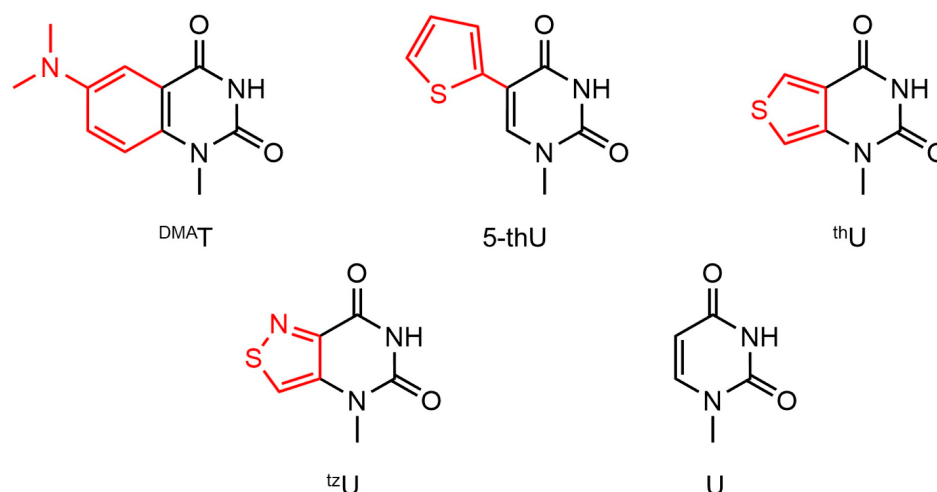


Figure 5. Structures of uridine-derived fluorescent nucleobases. Substitutions to the canonical U or T are indicated in red.

The acid-base properties of the fluorescent analogues generated for use in nucleic acid applications have not been fully delineated. Even in cases with reported experimental pK_a values, the values are often limited to atoms on the Watson–Crick face of the nucleobase. The other atoms of the nucleobases could be involved with interactions in noncanonical DNA or RNA structures or with proteins or other biomolecules. As such, knowing the pK_a values of all the various interaction sites of the nucleobase could be useful for a deeper understanding of the biological systems that employ these fluorescent analogues. Experimental determinations of the pK_a values are challenging, particularly when the nucleosides have low solubility or overlapping pK_a values for the individual protonation sites. Furthermore, extreme pH conditions needed to determine some pK_a values may damage the nucleoside. The use of computation overcomes some of these barriers and allows for determination of pK_a values in advance of time-consuming and costly syntheses and purification of the desired nucleosides.

Several methods for computing theoretical pK_a values have been employed [22–28]. In previous work on calculating the pK_a values of nucleobases, we used an implicit–explicit solvation model with the B3LYP density functional and the 6-31+G(d,p) basis set [25–27]. Small adjustments to the calculated pK_a values were made using a two-parameter linear regression fit to experimental pK_a values from the literature. As in our earlier studies, the ribose sugar was replaced by a methyl group to reduce the computational cost [25–27]. Previously, we examined natural and non-natural modified nucleobases, including aza-/deaza-modified nucleobases, and found the theoretical pK_a values closely matched the corresponding experimental pK_a values [25–27]. Using the same approach in this current study, we calculated 55 pK_a values for 25 different fluorescent nucleobase analogues.

2. Materials and Methods

The Gaussian 09 and 16 suite of programs [29,30] was used to determine the free energies of fluorescent analogue nucleobases (G16 uses a finer integration grid for density functional theory (DFT) than G09, but this does not result in any significant differences in the relative energies). The calculation method combined the B3LYP DFT [31,32] with the 6-31+G(d,p) basis set [33–35] and a hybrid implicit–explicit solvation model [25,28]. Explicit water molecules were placed at each hydrogen donor or acceptor site [25–28,36–38] of the fluorescent analogues and the SMD implicit solvation model [39] was used to mimic the bulk water environment surrounding the molecule. An overall workflow scheme is given in Figure S1. The water molecule positions for each nucleobase are provided in the

Supplementary Materials (Figures S2–S7). The ribose sugar was replaced by a methyl group to decrease the computational cost [25–27].

The lowest energy structure of the neutral nucleobase was calculated from several different explicit water molecule orientations. The selected neutral structure was either protonated or deprotonated, maintaining the same water molecule hydrogen-bonding relationship as the neutral species. The free energy difference between the neutral and protonated/deprotonated structures was used with the estimated experimental free energy of the proton in aqueous solution [40,41] (−270.297 kcal/mol) to determine the free energy for acid dissociation ($\Delta G_{dissoc(aq)}$) for the nucleobase structure. The calculated $\Delta G_{dissoc(aq)}$ was then used to determine the pK_a value of the given fluorescent analogue using Equation (1) (Tables S1–S4) (R is the ideal gas constant ($1.987 \text{ cal}\cdot\text{mol}^{-1}\cdot\text{K}^{-1}$), and T is the absolute temperature in Kelvin (298.15 K)).

$$pK_a = \frac{\Delta G_{dissoc(aq)}}{2.303RT} \quad (1)$$

The experimental pK_a values and the corresponding computationally determined values were plotted for purines and pyrimidines, and linear regression was used to obtain adjustments for the calculated values. The regression analysis included values from previous work using both naturally occurring (canonical) and non-natural modifications [26,27,42]. The linear adjustments for purines and pyrimidines were determined using Equations (2) and (3), respectively. The overall methodological scheme is shown in Figure S1.

$$\text{adjusted } pK_a = \frac{(\text{calculated } pK_a + 0.70)}{1.22} \quad (2)$$

$$\text{adjusted } pK_a = \frac{(\text{calculated } pK_a + 0.94)}{1.24} \quad (3)$$

3. Results

3.1. Determination of Nucleobase pK_a Values

Each fluorescent analogue has a distinct number of hydrogen-bond donors and acceptors and thus the number of explicitly placed water molecules around the nucleobase for hydrogen-bonding interactions varies. As with previous work, the theoretical pK_a values of the fluorescent analogues were adjusted using a linear correlation between the experimental and computational pK_a values of the nucleobases [26,27]. Different adjustments were applied for purines and pyrimidines due to their differing chemical structures (i.e., two versus one ring, respectively). A total of 55 pK_a values were determined for 13 purine and 12 pyrimidine fluorescent analogues. The purine analogue adjustment factor was updated from a previous study [27] by adding 2 new pK_a values of non-natural modifications (Table S5) [42] and 10 new pK_a values for the fluorescent analogues to the 24 data points of unmodified/modified nucleobases from the literature (Table S6) [26,27]. The pyrimidine analogue adjustment included 4 new pK_a values of non-natural modifications (Table S5) [42], 4 new experimental pK_a values for the fluorescent analogues, and 17 data points of previously reported nucleobases (Table S6) [26,27]. The mean absolute errors (MAEs) for purines and pyrimidines were 0.6 and 0.4, respectively. The pK_a values of the canonical nucleobases were calculated with Gaussian 16 in this study, resulting in improved theoretical pK_a values for the N7 of G and N3 of C compared to previous reports [26,27].

3.2. Purines

A total of 36 theoretical pK_a values of 13 purine fluorescent nucleobases were calculated. Most purine analogues have proton donor and acceptor sites at N1, N3, and N7 (Figures 2 and 3). All the purine nucleobases retained conventional numbering systems (Figure 1) except for 1, N^6 -ethenoadenosine (ϵ A) as shown in Figure 2. The linear regression between the computational and experimental pK_a values was used to obtain a two-parameter adjustment for these analogues. The linear fit was determined to be $1.22x - 0.70$ with an R^2 value of 0.95 (Figure 6). After adjustment, the theoretical pK_a values had a MAE of 0.6 pK_a units for the A and G analogues (Tables 1 and 2). If the calculated pK_a values are listed as “unfavorable”, the resulting pK_a values were determined to be less than -1.0 , for which protonation would not be energetically favorable.

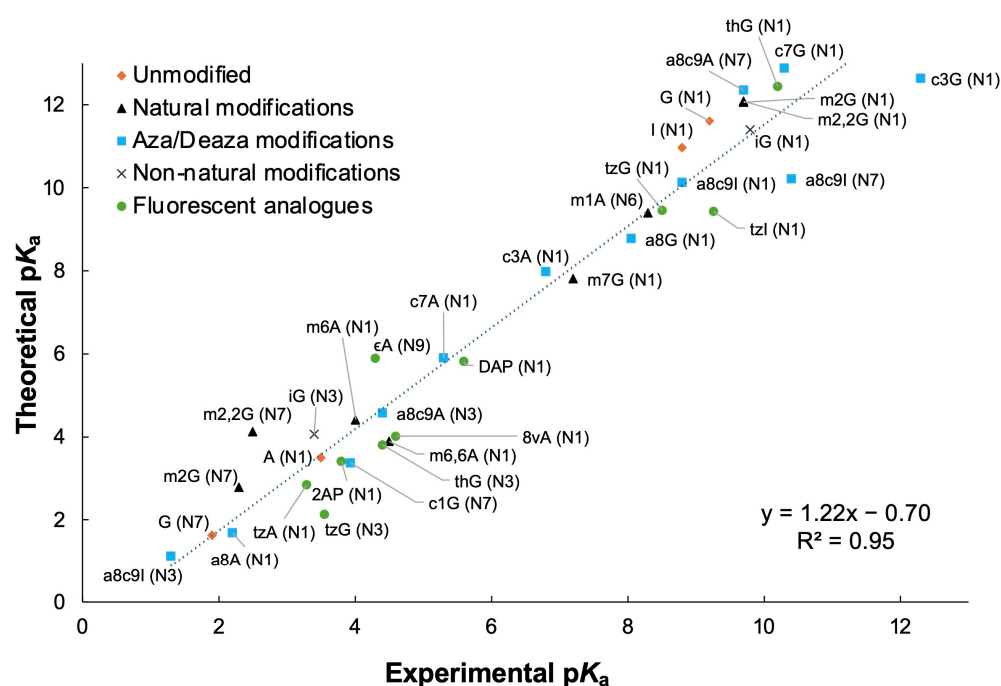


Figure 6. Plot and linear regression fit for the experimental and theoretical pK_a values of the fluorescent purine analogues. The data points for the linear correlation include unmodified (orange diamonds), naturally modified (black triangles), aza-/deaza-modified (blue squares), non-natural modifications (x marks), and fluorescent nucleobases (green circles).

The trend in the adjusted pK_a values for the N1 position of the fluorescent A analogues is as follows: $^{tz}A < 2AP \approx A < 8vA \approx 8\text{-Furan-A} < ^{th}A < DAP \ll ^{MD}A$ (with a range of 2.9–10.3) (Table 1). The pK_a value of ϵA at the N6 position (equivalent to the N1 position of A with respect to its location on the nucleobase) was not determined. The trend in pK_a values for the N3 sites of the A-type nucleobases is as follows: $2AP < 8vA \approx A < 8\text{-Furan-A} < ^{tz}A < DAP < ^{th}A \ll ^{MD}A$ (range of 0.6–9.2). The trend for the N7 site of the A analogues is: $A < 8vA \approx 2AP \approx \epsilon A (N1) < DAP$ (range of 1.3–3.5). The N1 position of ϵA is equivalent to the N7 position of A, and the pK_a values for the two species only differ by 1.3. The N4 position of ϵA is equivalent to the N3 position of A, but with unfavorable protonation compared to that of A (pK_a value of 1.3).

Table 1. Experimental and Adjusted pK_a Values for Fluorescent Analogues Derived from Adenosine.

Modification	Abbreviation	Position	Experimental pK_a Value ^a	Adjusted Theoretical pK_a Value ^b
2-aminopurine	2AP	N1	3.8 [43]	3.4
		N3		0.6
		N7		2.5
adenosine	A	N1	3.5 [44]	3.4
		N3		1.3
		N7		1.3
2,6-diaminopurine	DAP	N1	5.6 [45]	5.3
		N3		3.4
		N7		3.5
1, <i>N</i> ⁶ -ethenoadenosine	ϵ A	N1 (N7 ^c)	4.3 [46]	2.6
		N4 (N3 ^c)		unfavorable
		N9 (N6 ^c)		5.4
8-furan-adenosine	8-Furan-A	N1		3.9
		N3		2.0
		N7		unfavorable
methoxybenzodeazaadenosine	MD _A	N1		10.3
		N3		9.2
thieno[3,4, <i>d</i>]adenosine	th A	N1		4.8
		N3		5.0
isothiazolo[4,3, <i>d</i>]adenosine	^{tz} A	N1	3.3 [47]	2.9
		N3		3.0
		N7		unfavorable
8-vinyladenosine	8vA	N1	4.6 [48]	3.9
		N3		1.3
		N7		2.5

^a The experimental pK_a values were determined by UV titration [44,46], fluorescence spectroscopy [45,47], and NMR spectroscopy [43,48]. ^b The pK_a values lower than -1.0 are indicated as unfavorable. ^c The corresponding numbering system for adenosine.

Table 2. Experimental and Adjusted pK_a Values for Fluorescent Analogues Derived from Guanosine.

Modification	Abbreviation	Position	Experimental pK_a Value ^a	Adjusted Theoretical pK_a Value ^b
guanosine	G	N1	9.2 [49]	10.1
		N3		1.0
		N7	1.9 [49]	1.9
8-phenol-guanosine	8-Phenol-G	N1		9.4
		N3		0.2
		N7		unfavorable
thieno[3,4, <i>d</i>]guanosine	th G	N1	10.2 [50]	10.7
		N3		3.7
isothiazolo[4,3, <i>d</i>]guanosine	^{tz} G	N1	8.5 [47]	8.3
		N3		2.3
		N7		unfavorable
isothiazolo[4,3, <i>d</i>]inosine	^{tz} I	N1	9.3 [47]	8.3
		N3		-0.1
		N7		unfavorable
8-vinylguanosine	8vG	N1		9.2
		N3		0.03
		N7		3.2

^a The experimental pK_a values were determined by fluorescence spectroscopy [47,50] and NMR spectroscopy [49].

^b The pK_a values lower than -1.0 are indicated as unfavorable. ^c The experimental pK_a value in the literature was assigned to N7 [47], but the theoretical calculations suggest that assignment to N3 is more appropriate.

The G analogues have the following pK_a value trend for the N1 site: ${}^{\text{tz}}\text{G} \approx {}^{\text{tz}}\text{I} < 8\text{vG} \approx 8\text{-Phenol-G} < \text{G} < {}^{\text{th}}\text{G}$ (range of 8.3–10.7) (Table 2). The trend for the N3 position of the G nucleobases is as follows: ${}^{\text{tz}}\text{I} < 8\text{vG} < 8\text{-Phenol-G} < \text{G} < {}^{\text{tz}}\text{G} < {}^{\text{th}}\text{G}$ (range of -0.1 – 3.7). For all G analogues examined, protonation at the N7 site is unfavorable, except for 8vG (pK_a value of 3.2).

3.3. Adenosine Fluorescent Analogues

The A nucleobase modifications have a wide range of pK_a values, and several values differ significantly from those of the canonical, unmodified nucleobase. The N1 pK_a value of 2AP is equivalent to that of A, but the N3 and N7 sites have comparatively lower ($\Delta pK_a - 0.7$) or higher ($\Delta pK_a + 1.2$) pK_a values, respectively. The 2AP analogue has been reported to base pair with U or T in a standard geometry or with C in a wobble configuration [17,18,51,52]. Another widely used fluorescent analogue diaminopurine (DAP) has higher pK_a values compared to those of A at all three protonation sites ($\Delta pK_a + 1.9$, $+2.1$, and $+2.2$ for N1, N3, and N7, respectively). The DAP nucleobase has the potential for an additional hydrogen bond with U or T in comparison to canonical A-U/T base pairs because of the extra amine group, which increases base-pair stability [53,54] and decreases flexibility of nucleic acid structures [55]. The altered pK_a values of DAP could further impact the stability of U/T pairing.

The 8vA analogue has a vinyl group at position 8 that provides an extended π -system and improves the fluorescence quantum yield [56]. This analogue has only slightly higher N1 and N7 pK_a values ($\Delta pK_a + 0.5$ and $\Delta pK_a + 1.2$, respectively) and an equivalent N3 pK_a value compared to those of A. As such, 8vA is an ideal fluorescent replacement of A residues in canonical DNA or RNA duplexes. Indeed, when 8vA was incorporated into DNA duplexes, the resulting DNAs displayed greater thermodynamic stability than duplexes with 2AP and the oligonucleotides were detected with greater sensitivity [57]. In contrast, the N7 pK_a value of 8vA is impacted by the presence of the vinyl group, which could alter the thermodynamic stability of noncanonical DNA structures that involve the N7 group [58,59]. The 8-Furan-A analogue could also serve as a good fluorescent replacement of A because the N1 and N3 sites have only slight perturbations to the pK_a values (ΔpK_a of $+0.5$ and $+0.7$, respectively).

Ring-substituted or ring-fused A analogues also exhibit a range of pK_a differences in comparison to the parent A. Thiophene-attached A (${}^{\text{th}}\text{A}$) has higher pK_a values for both the N1 and N3 positions ($\Delta pK_a + 1.4$ and $+3.7$, respectively). The pK_a values for isothiazole-substituted A (${}^{\text{tz}}\text{A}$) also display small differences from A (ΔpK_a of -0.5 for N1 and ΔpK_a of $+1.7$ for N3). The ${}^{\text{tz}}\text{A}$ and ${}^{\text{th}}\text{A}$ analogues retain the ability to pair with the complementary bases (T or U) in a canonical manner [50].

The standard purine numbering system is not applied for ϵA (Figure 2). The pK_a value for the N1 position of ϵA is 1.3 units higher than the corresponding N7 of A. The N9 position of ϵA , which is unique to this analogue, has a pK_a value of 5.4. The corresponding amine (N6 position) of A has a pK_a value of 16.9 (Table S1). The pK_a value at the N6 position of ϵA was not determined because the explicit water molecule was observed to drift away from the expected position during the geometry optimization. The attached ring impedes DNA base pairing [60] but maintains similar binding properties as A to contractile proteins such as actin [61] and myosin [62,63]. As such, the fluorescent ϵA analogue is useful for probing nucleotide binding sites on proteins [64–66].

The ${}^{\text{MD}}\text{A}$ analogue with a methoxybenzene moiety, an electron-rich functional group, displays the largest differences in pK_a values compared to A ($\Delta pK_a + 6.9$ and $+7.9$ for N1 and N3, respectively). The ${}^{\text{MD}}\text{A}$ nucleoside can form stable base pairs with both T and C. No difference in base-pair stability between A-T and ${}^{\text{MD}}\text{A}$ -T pairs was observed;

whereas the melting temperature of an ^{MD}A-C wobble base pair was shown to be 11 °C higher than that of an A-C pair at pH 7 [67]. This difference in stability is supported by the pK_a values determined for ^{MD}A in this study. A typical A-C mismatch has one hydrogen bond between the N6 amine of A and N3 of C. The standard A N1 would only be protonated at low pH (pK_a 3.4), which would facilitate hydrogen bonding with the O2 of C and stabilize the A-C pair. The calculated pK_a value of the N1 site of ^{MD}A is significantly higher (pK_a 10.3), indicating that the N1 site would be protonated at pH 7. Therefore, our results are consistent with thermodynamic stabilization of the ^{MD}A-C pair relative to the A-C pair, due to additional hydrogen-bond formation at neutral pH.

3.4. Guanosine Fluorescent Analogues

Modification at the C8 position of G with vinyl or phenol moieties leads to altered pK_a values compared to the unmodified counterpart. The 8vG analogue, which has been used to study nucleic acid quadruplex structures [68], has slightly lower N1 and N3 pK_a values (ΔpK_a of −0.9 and −1.0, respectively) and a higher N7 pK_a value (ΔpK_a of +1.3). The 8vG has similar ΔpK_a values as 8vA, in which presence of the vinyl group decreases acidity of the N7 site compared to that of the corresponding unmodified nucleobase. The phenol moiety at the C8 position of 8-Phenol-G alters resonance stabilization of the purine structure, leading to changes in the pK_a values. Overall, the 8-Phenol-G analogue displays a similar pK_a trend as 8vG (ΔpK_a of −0.7 for N1 and ΔpK_a of −0.8 for N3).

The ring-substituted G analogues also have altered pK_a values compared to the parent G. The isothiazole-modified G (^{tz}G) and inosine (^{tz}I) analogues have the same pK_a value (8.3) for the N1 site, which differs from that of G (ΔpK_a of −1.8). In contrast, the pK_a values for N3 show opposite trends when compared to G (ΔpK_a of +1.3 for ^{tz}G and ΔpK_a of −1.1 for ^{tz}I). Of note, the experimental pK_a value of ^{tz}G (3.6) was reported for the N7 position [47]. This experimental value was obtained using fluorescence spectroscopy, from which it is not possible to distinguish the specific protonation sites, N3 and N7. In the present study, the pK_a value for N3 of ^{tz}G is 2.3 and protonation of the N7 is unfavorable. Therefore, the experimental value for N7 was reassigned to N3 (Table 2).

The thG nucleoside can exist in two ground-state tautomeric forms with protonation at N1 or N3 (referred to as thG-H1 and thG-H3, respectively) [69]. The thG-H1 tautomer is favored over the thG-H3 tautomer in duplexes because the H1 form preserves canonical base pairing with C [50]. In the present study, the pK_a values for N1 and N3 of thG-H1 were determined and observed to be higher than the corresponding values of G (ΔpK_a of +0.6 and +2.7, respectively). The slightly higher stability of a thG-C base pair compared to canonical G-C pair in a duplex DNA model system is consistent with the altered pK_a value for the N1 position of thG [20]. Perhaps somewhat surprisingly, only a small alteration in the N1 pK_a value is observed for G with the thiophene modification, supporting its usefulness as a fluorescent analogue for duplex DNA or RNA systems.

3.5. Pyrimidines

A total of 19 theoretical pK_a values of 12 pyrimidine-based fluorescent analogues were determined. The major protonation or deprotonation site of pyrimidines is N3, but additional sites of protonation/deprotonation are available for the modified nucleobases shown in Figures 4 and 5. A two-parameter adjustment was employed for the pyrimidine analogues with a regression of 1.24x − 0.94 with an R² value of 0.97 (Figure 7). The MAE for the adjusted pK_a values was 0.4 pK_a units (Tables 3 and 4).

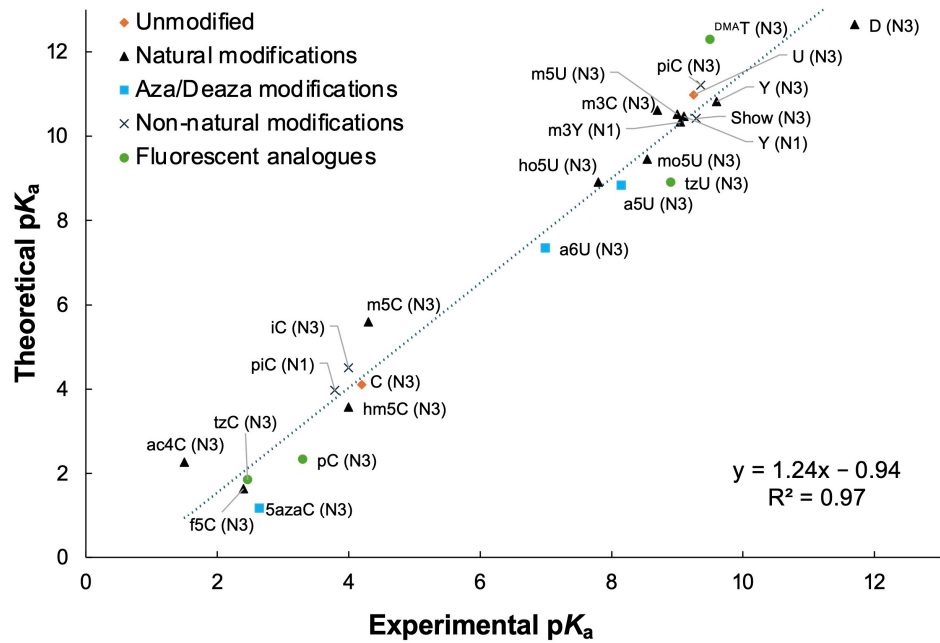


Figure 7. Plot and linear regression fit for the experimental and theoretical pK_a values of the fluorescent pyrimidine analogues. The data points for the linear correlation include unmodified (orange diamonds), naturally modified (black triangles), aza-/deaza-modified (blue squares), non-natural modifications (x marks), and fluorescent nucleobases (green circles).

Table 3. Experimental and Adjusted pK_a Values for Fluorescent Analogues Derived from Cytidine.

Modification	Abbreviation	Position	Experimental pK_a Value ^a	Adjusted Theoretical pK_a Value ^b
benzopyridopyrimidine	BPP	N3		0.9
		N7		9.4
cytidine	C	N3	4.2 [70]	4.0
5-furan-cytidine	C ^{FU}	N3		2.1
locked bicyclic 4- <i>N</i> -carbamoylcytidine	C ^{hpp}	N3		0.6
		N7		5.7
		N9		11.4
5-methyl-2-pyrimidinone	m ⁵ K	N3		3.5
pyrrolocytidine	pC	N3	3.3 [71]	2.6
		N7		13.2
1,3-diaza-2-oxophenoxazine	tC ^o	N3		2.1
		N7		11.6
thieno[3,4- <i>d</i>]cytidine	thC	N3		4.2
isothiazolo[4,3- <i>d</i>]cytidine	tzC	N3	2.5 [47]	2.2
		N7		unfavorable

^a The experimental pK_a values were determined by UV titration [70] and fluorescence spectroscopy [47,71]. ^b The pK_a values lower than -1.0 are indicated as unfavorable.

Table 4. Experimental and Adjusted pK_a Values for Fluorescent Analogues Derived from Uridine.

Modification	Abbreviation	Position	Experimental pK_a Value ^a	Adjusted Theoretical pK_a Value ^b
<i>N,N</i> -dimethylaniline-2-deoxythymidine	DMA _T	N3	9.5 [72]	10.6
5-(thien-2-yl)uridine	5-thU	N3		9.3
thieno[3,4- <i>d</i>]uridine	th U	N3		9.8
isothiazolo[4,3- <i>d</i>]uridine	^{tz} U	N3 N7	8.9 [47]	7.9 unfavorable
uridine	U	N3	9.2 [73]	9.6

^a The experimental pK_a values were determined by UV titration [73] and fluorescence spectroscopy [47,72]. ^b The pK_a values lower than -1.0 are indicated as unfavorable.

The adjusted theoretical pK_a values at the N3 site for the fluorescent C analogues have the following trend: $C^{hpp} \approx BPP < C^{FU} \approx tC^\circ \approx {}^{tz}C < pC < m^5K < C \approx {}^{th}C$ (range of 0.6–4.2). For the ring-fused analogues, the trend in pK_a values for the N7 position are as follows: $C^{hpp} \ll BPP < tC^\circ < pC$ (range of 5.7–13.2). The N3 site of the modified U fluorescent analogues has the following pK_a trend: ${}^{tz}U < 5\text{-thU} < U \approx {}^{th}U < \text{DMA}_T$ (range of 7.9–10.6).

3.6. Cytidine Fluorescent Analogues

Many fluorescent C analogues that have been reported in the literature have ring-fused modifications, although some have functional group additions or deletions (e.g., C^{FU} and m^5K) (Figure 4). Analogues such as m^5K have been used to study protein-DNA complex interactions [74]. The pK_a value for the N3 position of m^5K is only slightly lower than that of C (ΔpK_a of -0.5), so interactions of N3 in the absence of the amine group of C can be examined. Another modification with addition of a furan moiety at position 5 of C (C^{FU}) increases the acidity of N3 compared to C (ΔpK_a of -1.9). Thiophene- and isothiazole-modified C nucleobases show a similar trend as the purine analogues in which N3 has a negligible difference in pK_a for ${}^{th}C$ and C (ΔpK_a of $+0.2$) but is more acidic for ${}^{tz}C$ (ΔpK_a of -1.8).

Four ring-fused C analogues that were examined in this study display decreased pK_a values at the N3 position compared to C (ΔpK_a of -3.4 , -3.1 , -1.9 , and -1.4 for C^{hpp} , BPP, tC° , and pC, respectively). The N7 position of these ring-fused analogues is equivalent to the N4 amine of C. The N7 pK_a value of C^{hpp} is considerably more acidic than that of the other ring-fused modified analogues. In contrast, the N9 pK_a value of C^{hpp} is more basic compared to the N7 position, with more similarity to the N7 of tC° . Derivatives of C^{hpp} have been reported in the literature [75], which could be the focus of future pK_a calculations. For example, it may be of interest to seek ring-fused analogues with N3 pK_a values closer to the parent C.

The N7 pK_a value of pC is highly basic compared to other sites in the fluorescent C analogues examined in this study. The pC hydrogen-bond donor and acceptor pattern matches that of C with favorable pairing to G residues. The presence of pC has been shown to have a modest stabilizing effect on a DNA duplex compared to C [76]. The steady-state fluorescence of pC is only quenched when it pairs with G, making pC a useful probe for examining distortions in DNA/RNA hybrids such as those in the human immunodeficiency virus type-1 [77].

The adjusted pK_a value at the N7 position of tC° is 11.6. This tC° analogue can base pair with G [78] but has also been shown to pair with protonated C to form an intercalated

motif (i-motif) as shown in Figure 8 [79,80]. Due to a lower N3 pK_a value, tC° is not protonated at pH 5, which is necessary for i-motif formation with C, and the structure is better controlled [80]. Furthermore, tC° can be utilized to monitor the i-motif formation because of differing fluorescence emission of the folded and unfolded states [79,80]. The unique pH dependence of the stability and quenching of tC° in the i-motif allows it to be a useful probe for monitoring i-motif formation inside cells [80].

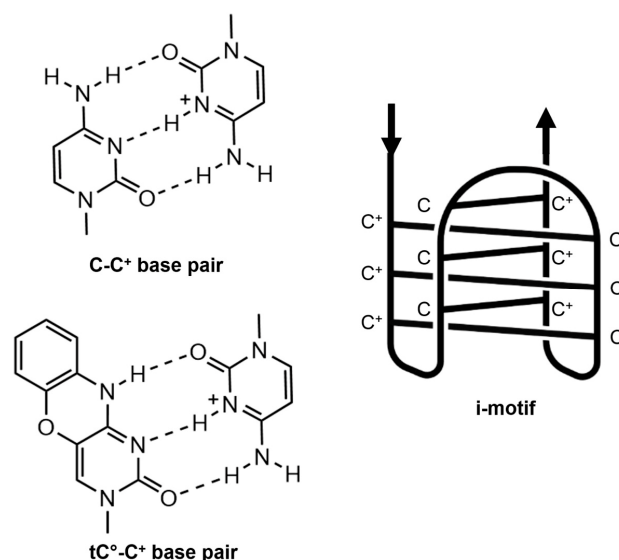


Figure 8. Hemiprotonated C-C or tC° pairing that contributes to i-motif formation by C-rich strands. The i-motif structure consists of C-C⁺ base pairs in an anti-parallel orientation. The canonical C nucleobase can be replaced with tC° for fluorescent monitoring of i-motif structure formation. The i-motif structure was created with Biorender.com.

3.7. Uridine Fluorescent Analogues

Addition of a thiophene functional group at position 5 of U (5-thU) leads to a slightly lower pK_a value (N3 position) compared to that of the canonical nucleobase (ΔpK_a of -0.3), and a ring-conjugated thiophene had little impact on the pK_a (^{th}U , ΔpK_a of $+0.2$). As such, 5-thU and ^{th}U serve as ideal fluorescent analogues of U. As with other fluorescent analogues, 5-thU exhibits altered fluorescence depending on its position and neighboring sequences in DNA and can be used to monitor DNA hybridization [81]. Similarly, fluorescence is significantly decreased when ^{th}U base pairs with A compared to mismatched pairs [21]. Due to these characteristics, ^{th}U is utilized as a fluorescent probe for single nucleotide polymorphism (SNP) detection.

The ^{tz}U analogue has a similar ΔpK_a trend as the isothiazole-fused nucleobases mentioned earlier (purines and C) (at N3, the ΔpK_a is -1.7 compared to U). Finally, the fluorescent nucleobase ^{DMA}T has the highest pK_a value at the N3 position among the U analogues examined (ΔpK_a of $+1.0$), but this change in pK_a is not likely to have much impact on its ability to pair with A. In studies with DNA model systems, only a slight thermal destabilization of duplexes containing A- ^{DMA}T base pairs was observed compared to A-T pairs, supporting the use of ^{DMA}T as a fluorescent probe [72].

4. Discussion

The theoretical pK_a values of 25 fluorescent analogues were determined by using a previously established ab initio quantum chemical protocol with an implicit–explicit solvation model, B3LYP density functional, and 6-31+G(d,p) basis set [25–28,31–39]. An adjustment factor for the theoretical pK_a values of purine and pyrimidine analogues was applied by comparing calculated and known experimental values. The differences

between the theoretical and experimental pK_a values of the fluorescent analogues are within 1.0 pK_a unit, except for the N9 position of ϵA and N3 of tG and ${}^{DMA}T$. The adjusted theoretical pK_a values for the fluorescent analogues show good agreement, with MAEs of 0.6 and 0.4 pK_a units for the purines and pyrimidines, respectively. This study expands the nucleoside/nucleobase database for natural [26] and non-natural [27] (e.g., aza/deaza) modifications with 55 values for the fluorescent analogues. The combined databases now include over 75 different nucleobases and more than 150 pK_a values.

The theoretical pK_a values complement existing experimental data by providing unknown or ambiguous pK_a values for fluorescent nucleotide analogues. Many studies employ pH titrations of nucleosides and monitor changes in the UV or fluorescence spectra [44–47,70,71,73]. However, those experiments are sometimes unable to identify the specific protonation/deprotonation sites if multiple proton donor/acceptors are present in a molecule, particularly when the pK_a values are overlapping. For example, experimental pK_a values of tG were determined by fluorescence spectroscopy with two titration curves [47]. The experimental and adjusted theoretical pK_a values for N1 of tG are comparable (8.5 and 8.3, respectively). In contrast, the second experimental pK_a value of tG was assigned to N7, but the N7 and N3 sites are both potential proton acceptors. Our study reveals that assignment of the second experimental pK_a value to N3 is more appropriate because protonation of the N7 position of tG was observed to be energetically unfavorable.

The present study highlights how nucleobases can be tuned for both fluorescent properties and pK_a values, particularly at key positions that are involved with canonical and noncanonical pairing interactions without significantly perturbing the overall structures of the nucleic acids. The wide range of pK_a values for the purine and pyrimidine fluorescent nucleobases (varying by 2.4 to 7.3 pK_a units for each site) is consistent with the range for the natural modifications. For comparison, the N1 of A displayed a range of 3.3 pK_a units for methylation at different sites [26], 5.5 pK_a units for aza/deaza substitutions [27], and 7.3 pK_a units for the fluorescent analogues examined (Table 1).

The broad range of acid-base properties of the fluorescent analogues is valuable for their applications in nucleic acid systems under differing pH conditions in cellular environments or other biological milieu. For example, a stable i-motif structure requires hemiprotonated C-C⁺ base pairs, which can exist in the pH range of 4.5 to 6.5 [82]. Replacing C with tC° within the i-motif structure allows monitoring of the conformational states in cells through fluorescence changes, while also contributing to preservation of the i-motif structure in a broader range of pH environments [80]. The wide range of pK_a values of the fluorescent analogues could also be utilized for developing pH-dependent DNA sensors that display different conformational states in acidic or basic conditions and can be monitored in live cells [83].

While many different types of fluorescent nucleobase analogues have been designed and synthesized [60], isothiazole- and thiophene-substituted variants of all four canonical nucleosides have been generated [20,47], allowing for comparison of the pK_a trends. The fused-ring thiophene substitutions (${}^{th}U$ and ${}^{th}C$) do not significantly impact the N3 pK_a values of the pyrimidine analogues. Similarly, the N1 pK_a value of ${}^{th}G$ is shifted by only 0.6 units, and the ${}^{th}A$ analogue exhibits a slightly increased pK_a value (ΔpK_a of +1.4) compared to A. In contrast, the fused-ring isothiazole substitutions on the pyrimidine rings display larger changes in the N3 pK_a values (ΔpK_a values of -1.9 for ${}^{th}U$ vs. tU and -2.0 for ${}^{th}C$ vs. tC) as well as for the N1 of purine analogues (ΔpK_a values of -2.4 for ${}^{th}G$ vs. tG and -1.9 for ${}^{th}A$ vs. tA). These examples illustrate how the nucleobases can be tuned with multiple or single atom changes for optimizing the fluorescent properties and pK_a values, including positions that are involved with canonical and noncanonical pairing interactions.

As with our prior studies [26,27], the model systems using the first explicit solvation shell in combination with the B3LYP/6-31+G(d,p) level of theory provided reliable pK_a values that matched well with the reported experimental values, regardless of the ring size or type of modification. One of the more challenging modifications was ϵA , which is a three-ring system with multiple sites of protonation. The pK_a value of the N6 (N1*) position of ϵA could not be determined because water placement at that position was not energetically favorable for geometry optimization. However, the larger size of the nucleobase did not impact our ability to determine pK_a values for the other positions of ϵA . Similarly, the expanded ring systems for the pyrimidines have additional (de)protonation sites but behaved like the purine systems.

The present study only focused on 25 analogues among hundreds of fluorescent nucleobases that have been reported in the literature [4,60]. Determining the pK_a values for those additional fluorescent analogues could reveal wider ranges of values that could facilitate future design and biological or biotechnology applications. Theoretical pK_a computations in advance of carrying out synthetic procedures to generate the fluorescent analogues could also be advantageous.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/compounds5040044/s1>, Figure S1: Methodological Workflow Diagram of the pK_a Calculations of the Fluorescent Analogues; Figure S2: Images of Optimized Structures of Adenosine Fluorescent Analogues; Figure S3: Images of Optimized Structures of Guanosine Fluorescent Analogues; Figure S4: Images of Optimized Structures of Cytidine Fluorescent Analogues; Figure S5: Images of Optimized Structures of Uridine Fluorescent Analogues; Figure S6: Images of Optimized Structures of Non-natural Modifications; Figure S7: Cartesian Coordinates of Optimized Structures; Table S1: pK_a Calculations of the Adenosine Analogues; Table S2: pK_a Calculations of the Guanosine Analogues; Table S3: pK_a Calculations of the Cytidine Analogues; Table S4: pK_a Calculations of the Uridine Analogues; Table S5: pK_a Calculations of the Non-natural Modifications [42]; Table S6: The Experimental and Computational pK_a Values used for Linear Regression Adjustments of Purines and Pyrimidines.

Author Contributions: Conceptualization, S.J.I., A.J.M. and C.S.C.; Methodology (Computation), S.J.I., A.J.M. and H.B.S.; Software, H.B.S.; Validation and Analysis, S.J.I., A.J.M. and H.B.S.; Data Curation, S.J.I. and A.J.M.; Writing—Original Draft Preparation, S.J.I., A.J.M. and C.S.C.; Writing—Review and Editing, S.J.I., A.J.M., H.B.S. and C.S.C.; Visualization, S.J.I.; Supervision, H.B.S. and C.S.C.; Project Administration, C.S.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported in part by a grant from the National Science Foundation (CHE1856437).

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding author.

Acknowledgments: The Wayne State University High Performance Computing Services Teams provided computational support.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

A	Adenosine
C	Cytidine
DFT	Density functional theory

DNA	Deoxyribonucleic acid
G	Guanosine
i-motif	Intercalated motif
MAE	Mean absolute error
pK _a	Negative logarithm of the acid dissociation constant
ΔpK _a	Difference in pK _a values
RNA	Ribonucleic acid
SNP	Single nucleotide polymorphism
T	Thymidine
th	Thiophene
tz	Isothiazole
U	Uridine

References

1. Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, NY, USA, 1984.
2. Šponer, J.; Leszczynski, J.; Hobza, P. Electronic Properties, Hydrogen Bonding, Stacking, and Cation Binding of DNA and RNA Bases. *Biopolymers* **2001**, *61*, 3–31. [\[CrossRef\]](#)
3. Shan, S.O.; Loh, S.; Herschlag, D. The Energetics of Hydrogen Bonds in Model Systems: Implications for Enzymatic Catalysis. *Science* **1996**, *272*, 97–101. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Tor, Y. Isomorphous Fluorescent Nucleosides. *Acc. Chem. Res.* **2024**, *57*, 1325–1335. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Ward, D.C.; Reich, E.; Stryer, L. Fluorescence Studies of Nucleotides and Polynucleotides: I. Formycin, 2-Aminopurine Riboside, 2,6-Diaminopurine Riboside, and Their Derivatives. *J. Biol. Chem.* **1969**, *244*, 1228–1237. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Secrist, J.A.; Barrio, J.R.; Leonard, N.J. A Fluorescent Modification of Adenosine Triphosphate with Activity in Enzyme Systems: 1,N6-Ethenoadenosine Triphosphate. *Science* **1972**, *175*, 646–647. [\[CrossRef\]](#)
7. Wu, P.; Nordlund, T.M.; Gildea, B.; McLaughlin, L.W. Base Stacking and Unstacking as Determined from a DNA Decamer Containing a Fluorescent Base. *Biochemistry* **1990**, *29*, 6508–6514. [\[CrossRef\]](#)
8. Hurley, D.J.; Seaman, S.E.; Mazura, J.C.; Tor, Y. Fluorescent 1,10-Phenanthroline-Containing Oligonucleotides Distinguish between Perfect and Mismatched Base Pairing. *Org. Lett.* **2002**, *4*, 2305–2308. [\[CrossRef\]](#)
9. Wojciechowski, F.; Hudson, R.H.E. Fluorescence and Hybridization Properties of Peptide Nucleic Acid Containing a Substituted Phenylpyrrolocytosine Designed to Engage Guanine with an Additional H-Bond. *J. Am. Chem. Soc.* **2008**, *130*, 12574–12575. [\[CrossRef\]](#)
10. Stolar, T.; Pearce, B.K.D.; Etter, M.; Truong, K.N.; Ostojić, T.; Krajnc, A.; Mali, G.; Rossi, B.; Molčanov, K.; Lončarić, I.; et al. Base-Pairing of Uracil and 2,6-Diaminopurine: From Cocrystals to Photoreactivity. *iScience* **2024**, *27*, 109894. [\[CrossRef\]](#)
11. Mandal, S.; Hebenbrock, M.; Müller, J. A Dinuclear Silver(I)-Mediated Base Pair in DNA Formed from 1,N6-Ethenoadenine and Thymine. *Inorg. Chim. Acta* **2018**, *472*, 229–233. [\[CrossRef\]](#)
12. Sinkeldam, R.W.; Greco, N.J.; Tor, Y. Fluorescent Analogs of Biomolecular Building Blocks: Design, Properties, and Applications. *Chem. Rev.* **2010**, *110*, 2579–2619. [\[CrossRef\]](#) [\[PubMed\]](#)
13. Soulière, M.F.; Haller, A.; Rieder, R.; Micura, R. A Powerful Approach for the Selection of 2-Aminopurine Substitution Sites to Investigate RNA Folding. *J. Am. Chem. Soc.* **2011**, *133*, 16161–16167. [\[CrossRef\]](#)
14. Datta, M.; Liu, J. A DNA Aptamer for 2-Aminopurine: Binding-Induced Fluorescence Quenching. *Chem. Asian J.* **2024**, *19*, e202400817. [\[CrossRef\]](#)
15. Matarazzo, A.; Hudson, R.H.E. Fluorescent Adenosine Analogs: A Comprehensive Survey. *Tetrahedron* **2015**, *71*, 1627–1657. [\[CrossRef\]](#)
16. Serrano-Andrés, L.; Merchán, M.; Borin, A.C. Adenine and 2-Aminopurine: Paradigms of Modern Theoretical Photochemistry. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 8691–8696. [\[CrossRef\]](#)
17. Nordlund, T.M.; Andersson, S.; Nilsson, L.; Rigler, R.; Graeslund, A.; McLaughlin, L.W. Structure and Dynamics of a Fluorescent DNA Oligomer Containing the EcoRI Recognition Sequence: Fluorescence, Molecular Dynamics, and NMR Studies. *Biochemistry* **1989**, *28*, 9095–9103. [\[CrossRef\]](#)
18. Sowers, L.C.; Fazakerley, G.V.; Eritja, R.; Kaplan, B.E.; Goodman, M.F. Base Pairing and Mutagenesis: Observation of a Protonated Base Pair between 2-Aminopurine and Cytosine in an Oligonucleotide by Proton NMR. *Proc. Nat. Acad. Sci. USA* **1986**, *83*, 5434–5438. [\[CrossRef\]](#)
19. Guest, C.R.; Hochstrasser, R.A.; Sowers, L.C.; Millar, D.P. Dynamics of Mismatched Base Pairs in DNA. *Biochemistry* **1991**, *30*, 3271–3279. [\[CrossRef\]](#)
20. Shin, D.; Sinkeldam, R.W.; Tor, Y. Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2011**, *133*, 14912–14915. [\[CrossRef\]](#)

21. Srivatsan, S.G.; Weizman, H.; Tor, Y. A Highly Fluorescent Nucleoside Analog Based on Thieno[3,4-d]Pyrimidine Senses Mismatched Pairing. *Org. Biomol. Chem.* **2008**, *6*, 1334–1338. [[CrossRef](#)] [[PubMed](#)]
22. Shields, G.C.; Seybold, P.G. *Computational Approaches for the Prediction of pKa Values*; CRC Press: Boca Raton, FL, USA; London, UK; New York, NY, USA, 2013.
23. Ho, J.; Coote, M.L. A Universal Approach for Continuum Solvent pKa Calculations: Are We There Yet? *Theor. Chem. Acc.* **2010**, *125*, 3–21. [[CrossRef](#)]
24. Thapa, B.; Raghavachari, K. Accurate pKa Evaluations for Complex Bio-Organic Molecules in Aqueous Media. *J. Chem. Theory Comput.* **2019**, *15*, 6025–6035. [[CrossRef](#)] [[PubMed](#)]
25. Thapa, B.; Schlegel, H.B. Calculations of pKa's and Redox Potentials of Nucleobases with Explicit Waters and Polarizable Continuum Solvation. *J. Phys. Chem. A* **2015**, *119*, 5134–5144. [[CrossRef](#)] [[PubMed](#)]
26. Jones, E.L.; Mlotkowski, A.J.; Hebert, S.P.; Schlegel, H.B.; Chow, C.S. Calculations of pKa Values for a Series of Naturally Occurring Modified Nucleobases. *J. Phys. Chem. A* **2022**, *126*, 1518–1529. [[CrossRef](#)] [[PubMed](#)]
27. Mlotkowski, A.J.; Schlegel, H.B.; Chow, C.S. Calculated pKa Values for a Series of Aza- and Deaza-Modified Nucleobases. *J. Phys. Chem. A* **2023**, *127*, 3526–3534. [[CrossRef](#)]
28. Thapa, B.; Schlegel, H.B. Improved pKa Prediction of Substituted Alcohols, Phenols, and Hydroperoxides in Aqueous Medium Using Density Functional Theory and a Cluster-Continuum Solvation Model. *J. Phys. Chem. A* **2017**, *121*, 4698–4706. [[CrossRef](#)]
29. 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.; Nakatsuji, H.; et al. *Gaussian 09, Revision D.01*; Gaussian Inc.: Wallingford, CT, USA, 2014.
30. 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.; Nakatsuji, H.; et al. *Gaussian 16 Rev. B.01*; Gaussian Inc.: Wallingford, CT, USA, 2016.
31. Lee, C.; Yang, W.; Parr, R.G. Development of the Colle-Salvetti Correlation-Energy Formula into a Functional of the Electron Density. *Phys. Rev. B* **1988**, *37*, 785–789. [[CrossRef](#)]
32. Becke, A.D. Density-Functional Thermochemistry. III. The Role of Exact Exchange. *J. Chem. Phys.* **1993**, *98*, 5648–5652. [[CrossRef](#)]
33. Hehre, W.J.; Ditchfield, R.; Pople, J.A. Self-Consistent Molecular Orbital Methods. XII. Further Extensions of Gaussian-Type Basis Sets for Use in Molecular Orbital Studies of Organic Molecules. *J. Chem. Phys.* **1972**, *56*, 2257–2261. [[CrossRef](#)]
34. Hariharan, P.C.; Pople, J.A. The Influence of Polarization Functions on Molecular Orbital Hydrogenation Energies. *Theor. Chim. Acta* **1973**, *28*, 213–222. [[CrossRef](#)]
35. Francel, M.M.; Pietro, W.J.; Hehre, W.J.; Binkley, J.S.; Gordon, M.S.; DeFrees, D.J.; Pople, J.A. Self-Consistent Molecular Orbital Methods. XXIII. A Polarization-Type Basis Set for Second-Row Elements. *J. Chem. Phys.* **1982**, *77*, 3654–3665. [[CrossRef](#)]
36. Bryantsev, V.S.; Diallo, M.S.; Goddard, W.A. Calculation of Solvation Free Energies of Charged Solutes Using Mixed Cluster/Continuum Models. *J. Phys. Chem. B* **2008**, *112*, 9709–9719. [[CrossRef](#)] [[PubMed](#)]
37. Zhang, S. A Reliable and Efficient First Principles-Based Method for Predicting pKa Values. III. Adding Explicit Water Molecules: Can the Theoretical Slope be Reproduced and pKa Values Predicted More Accurately? *J. Comput. Chem.* **2012**, *33*, 517–526. [[CrossRef](#)]
38. da Silva, E.F.; Svendsen, H.F.; Merz, K.M. Explicitly Representing the Solvation Shell in Continuum Solvent Calculations. *J. Phys. Chem. A* **2009**, *113*, 6404–6409. [[CrossRef](#)]
39. Marenich, A.V.; Cramer, C.J.; Truhlar, D.G. Universal Solvation Model Based on Solute Electron Density and on a Continuum Model of the Solvent Defined by the Bulk Dielectric Constant and Atomic Surface Tensions. *J. Phys. Chem. B* **2009**, *113*, 6378–6396. [[CrossRef](#)] [[PubMed](#)]
40. Camaioni, D.M.; Schwerdtfeger, C.A. Comment on “Accurate Experimental Values for the Free Energies of Hydration of H⁺, OH[−], and H₃O⁺”. *J. Phys. Chem. A* **2005**, *109*, 10795–10797. [[CrossRef](#)] [[PubMed](#)]
41. Kelly, C.P.; Cramer, C.J.; Truhlar, D.G. Aqueous Solvation Free Energies of Ions and Ion-Water Clusters Based on an Accurate Value for the Absolute Aqueous Solvation Free Energy of the Proton. *J. Phys. Chem. B* **2006**, *110*, 16066–16081. [[CrossRef](#)]
42. Mlotkowski, A.J. The Impacts of Modification on Acid Dissociation Constants of Nucleobases and Structure of Helix 69 RNA. Ph.D. Dissertation, Wayne State University, Detroit, MI, USA, 2023.
43. Sowers, L.C.; Mhaskar, D.N.; Khwaja, T.A.; Goodman, M.F. Preparation of Imino and Amino N-15 Enriched 2-Aminopurine Deoxynucleoside. *Nucleosides Nucleotides* **1989**, *8*, 23–34. [[CrossRef](#)]
44. Harkins, T.R.; Freiser, H. Adenine-Metal Complexes_{1,2}. *J. Am. Chem. Soc.* **1958**, *80*, 1132–1135. [[CrossRef](#)]
45. Ma, L.; Kartik, S.; Liu, B.; Liu, J. From General Base to General Acid Catalysis in a Sodium-Specific DNAzyme by a Guanine-to-Adenine Mutation. *Nucleic Acids Res.* **2019**, *47*, 8154–8162. [[CrossRef](#)]
46. Penzer, G.R. The Solution Conformation and Some Spectroscopic Properties of 1, N6-Ethenoadenosine Monophosphate, a Fluorescent Analogue of AMP. *Eur. J. Biochem.* **1973**, *34*, 297–305. [[CrossRef](#)] [[PubMed](#)]
47. Rovira, A.R.; Fin, A.; Tor, Y. Chemical Mutagenesis of an Emissive RNA Alphabet. *J. Am. Chem. Soc.* **2015**, *137*, 14602–14605. [[CrossRef](#)]

48. Roday, S.; Saen-oon, S.; Schramm, V.L. Vinyldeoxyadenosine in a Sarcin-Ricin RNA Loop and Its Binding to Ricin Toxin A-Chain. *Biochemistry* **2007**, *46*, 6169–6182. [[CrossRef](#)]
49. Levene, P.A.; Bass, L.W. *Nucleic Acids*; Chemical Catalog Company: New York, NY, USA, 1931.
50. Dziuba, D.; Didier, P.; Ciacco, S.; Barth, A.; Seidel, C.A.M.; Mély, Y. Fundamental Photophysics of Isomorphous and Expanded Fluorescent Nucleoside Analogues. *Chem. Soc. Rev.* **2021**, *50*, 7062–7107. [[CrossRef](#)]
51. Fagan, P.A.; Fàbrega, C.; Eritja, R.; Goodman, M.F.; Wemmer, D.E. NMR Study of the Conformation of the 2-Aminopurine:Cytosine Mismatch in DNA. *Biochemistry* **1996**, *35*, 4026–4033. [[CrossRef](#)]
52. Law, S.M.; Eritja, R.; Goodman, M.F.; Breslauer, K.J. Spectroscopic and Calorimetric Characterizations of DNA Duplexes Containing 2-Aminopurine. *Biochemistry* **1996**, *35*, 12329–12337. [[CrossRef](#)]
53. Lamm, G.M.; Blencowe, B.J.; Sparoat, B.S.; Iribarren, A.M.; Ryder, U.; Lamond, A.I. Antisense Probes Containing 2-Aminoadenosine Allow Efficient Depletion of U5 snRNP from HeLa Splicing Extracts. *Nucleic Acids Res.* **1991**, *19*, 3193–3198. [[CrossRef](#)]
54. Hoheisel, J.D.; Lehrach, H. Quantitative Measurements on the Duplex Stability of 2,6-Diaminopurine and 5-Chloro-Uracil Nucleotides Using Enzymatically Synthesized Oligomers. *FEBS Lett.* **1990**, *274*, 103–106.
55. Bailly, C.; Waring, M.J. The Use of Diaminopurine to Investigate Structural Properties of Nucleic Acids and Molecular Recognition between Ligands and DNA. *Nucleic Acids Res.* **1998**, *26*, 4309–4314. [[CrossRef](#)] [[PubMed](#)]
56. Kodali, G.; Kistler, K.A.; Narayanan, M.; Matsika, S.; Stanley, R.J. Change in Electronic Structure upon Optical Excitation of 8-Vinyladenosine: An Experimental and Theoretical Study. *J. Phys. Chem. A* **2010**, *114*, 256–267. [[CrossRef](#)] [[PubMed](#)]
57. Gaied, N.B.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mély, Y. 8-Vinyl-Deoxyadenosine, an Alternative Fluorescent Nucleoside Analog to 2'-Deoxyribosyl-2-Aminopurine with Improved Properties. *Nucleic Acids Res.* **2005**, *33*, 1031–1039. [[CrossRef](#)]
58. Seela, F.; Zulauf, M. 7-Deazaadenine-DNA: Bulky 7-Iodo Substituents or Hydrophobic 7-Hexynyl Chains Are Well Accommodated in the Major Groove of Oligonucleotide Duplexes. *Chem. Eur. J.* **1998**, *4*, 1781–1790. [[CrossRef](#)]
59. Booth, J.; Cummins, W.J.; Brown, T. An Analogue of Adenine that Forms an “A:T” Base Pair of Comparable Stability to G:C. *Chem. Commun.* **2004**, *19*, 2208–2209. [[CrossRef](#)]
60. Xu, W.; Chan, K.M.; Kool, E.T. Fluorescent Nucleobases as Tools for Studying DNA and RNA. *Nat. Chem.* **2017**, *9*, 1043–1055. [[CrossRef](#)]
61. Thames, K.E.; Cheung, H.C.; Harvey, S.C. Binding of 1,N6-Ethanoadenosine Triphosphate to Actin. *Biochem. Biophys. Res. Commun.* **1974**, *60*, 1252–1261. [[CrossRef](#)]
62. McCubbin, W.D.; Willick, G.E.; Kay, C.M. Enzymatic Studies on the Interaction of Myosin and Heavy Meromyosin with 1,N6-Ethanoadenosine Triphosphate (ϵ ATP), a Fluorescent Analog of ATP. *Biochem. Biophys. Res. Commun.* **1973**, *50*, 926–933. [[CrossRef](#)]
63. Willick, G.E.; Oikawa, K.; McCubbin, W.D.; Kay, C.M. Equilibrium Dialysis Binding Studies of 1,N6-Ethanoadenosine Diphosphate (ϵ ADP) to Myosin, Heavy Meromyosin, and Subfragment One. *Biochem. Biophys. Res. Commun.* **1973**, *53*, 923–928. [[CrossRef](#)]
64. Miki, M.; Mihashi, K. Fluorescence and Flow Dichroism of F-Actin- ϵ -ADP; the Orientation of the Adenine Plane Relative to the Long Axis of F-Actin. *Biophys. Chem.* **1976**, *6*, 101–106. [[CrossRef](#)] [[PubMed](#)]
65. Harvey, S.C.; Cheung, H.C. Fluorescence Studies of 1,N6-Ethanoadenosine Triphosphate Bound to G-Actin: The Nucleotide Base is Inaccessible to Water. *Biochem. Biophys. Res. Commun.* **1976**, *73*, 865–868. [[CrossRef](#)] [[PubMed](#)]
66. Perkins, W.J.; Wells, J.A.; Yount, R.G. Characterization of the Properties of Ethanoadenosine Nucleotides Bound or Trapped at the Active Site of Myosin Subfragment 1. *Biochemistry* **1984**, *23*, 3994–4002. [[CrossRef](#)] [[PubMed](#)]
67. Okamoto, A.; Tanaka, K.; Saito, I. DNA Logic Gates. *J. Am. Chem. Soc.* **2004**, *126*, 9458–9463. [[CrossRef](#)]
68. Nadler, A.; Strohmeier, J.; Diederichsen, U. 8-Vinyl-2'-Deoxyguanosine as a Fluorescent 2'-Deoxyguanosine Mimic for Investigating DNA Hybridization and Topology. *Angew. Chem. Int. Ed.* **2011**, *50*, 5392–5396. [[CrossRef](#)] [[PubMed](#)]
69. Sholokh, M.; Improta, R.; Mori, M.; Sharma, R.; Kenfack, C.; Shin, D.; Voltz, K.; Stote, R.H.; Zaporozhets, O.A.; Botta, M.; et al. Tautomers of a Fluorescent G Surrogate and Their Distinct Photophysics Provide Additional Information Channels. *Angew. Chem. Int. Ed.* **2016**, *55*, 7974–7978. [[CrossRef](#)] [[PubMed](#)]
70. Shugar, D.; Fox, J.J. Spectrophotometric Studies of Nucleic Acid Derivatives and Related Compounds as a Function of pH. I. Pyrimidines. *Biochim. Biophys. Acta* **1952**, *9*, 199–218. [[CrossRef](#)]
71. Tinsley, R.A.; Walter, N.G. Pyrrolo-C as a Fluorescent Probe for Monitoring RNA Secondary Structure Formation. *RNA* **2006**, *12*, 522–529. [[CrossRef](#)] [[PubMed](#)]
72. Mata, G.; Schmidt, O.P.; Luedtke, N.W. A Fluorescent Surrogate of Thymidine in Duplex DNA. *Chem. Commun.* **2016**, *52*, 4718–4721. [[CrossRef](#)]
73. Hall, R. *The Modified Nucleosides in Nucleic Acids*; Columbia University Press: New York, NY, USA, 1971.
74. Singleton, S.F.; Shan, F.; Kanan, M.W.; McIntosh, C.M.; Stearman, C.J.; Helm, J.S.; Webb, K.J. Facile Synthesis of a Fluorescent Deoxycytidine Analogue Suitable for Probing the RecA Nucleoprotein Filament. *Org. Lett.* **2001**, *3*, 3919–3922. [[CrossRef](#)]

75. Wada, T.; Kobori, A.; Kawahara, S.I.; Sekine, M. Synthesis and Properties of Oligodeoxyribonucleotides Containing 4-N-Acetylcytosine Bases. *Tetrahedron Lett.* **1998**, *39*, 6907–6910. [[CrossRef](#)]
76. Datta, K.; Johnson, N.P.; Villani, G.; Marcus, A.H.; von Hippel, P.H. Characterization of the 6-Methyl Isoxanthopterin (6-MI) Base Analog Dimer, a Spectroscopic Probe for Monitoring Guanine Base Conformations at Specific Sites in Nucleic Acids. *Nucleic Acids Res.* **2012**, *40*, 1191–1202. [[CrossRef](#)]
77. Dash, C.; Rausch, J.W.; Le Grice, S.F. Using Pyrrolo-Deoxycytosine to Probe RNA/DNA Hybrids Containing the Human Immunodeficiency Virus Type-1 3' Polypurine Tract. *Nucleic Acids Res.* **2004**, *32*, 1539–1547. [[CrossRef](#)]
78. Wilhelmsson, L.M.; Holmén, A.; Lincoln, P.; Nielsen, P.E.; Nordén, B. A Highly Fluorescent DNA Base Analogue that Forms Watson–Crick Base Pairs with Guanine. *J. Am. Chem. Soc.* **2001**, *123*, 2434–2435. [[CrossRef](#)]
79. Reilly, S.M.; Lyons, D.F.; Wingate, S.E.; Wright, R.T.; Correia, J.J.; Jameson, D.M.; Wadkins, R.M. Folding and Hydrodynamics of a DNA i-Motif from the c-MYC Promoter Determined by Fluorescent Cytidine Analogs. *Biophys. J.* **2014**, *107*, 1703–1711. [[CrossRef](#)] [[PubMed](#)]
80. Mir, B.; Serrano-Chacón, I.; Medina, P.; Macaluso, V.; Terrazas, M.; Gandioso, A.; Garavís, M.; Orozco, M.; Escaja, N.; González, C. Site-Specific Incorporation of a Fluorescent Nucleobase Analog Enhances i-Motif Stability and Allows Monitoring of i-Motif Folding Inside Cells. *Nucleic Acids Res.* **2024**, *52*, 3375–3389. [[CrossRef](#)] [[PubMed](#)]
81. Noé, M.S.; Sinkeldam, R.W.; Tor, Y. Oligodeoxynucleotides Containing Multiple Thiophene-Modified Isomorphous Fluorescent Nucleosides. *J. Org. Chem.* **2013**, *78*, 8123–8128. [[CrossRef](#)] [[PubMed](#)]
82. Abou Assi, H.; Garavís, M.; González, C.; Damha, M.J. i-Motif DNA: Structural Features and Significance to Cell Biology. *Nucleic Acids Res.* **2018**, *46*, 8038–8056. [[CrossRef](#)]
83. Modi, S.; Nizak, C.; Surana, S.; Halder, S.; Krishnan, Y. Two DNA Nanomachines Map pH Changes Along Intersecting Endocytic Pathways Inside the Same Cell. *Nat. Nanotechnol.* **2013**, *8*, 459–467. [[CrossRef](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.