JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

COMMUNICATION

pubs.acs.org/JACS

Emissive RNA Alphabet

Dongwon Shin, Renatus W. Sinkeldam, and Yitzhak Tor*

Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California 92093-0358, United States

Supporting Information

ABSTRACT: A fluorescent ribonucleoside alphabet consisting of highly emissive purine (thA, thG) and pyrimidine (thU, thC) analogues, all derived from thieno[3,4-*d*]-pyrimidine as the heterocyclic nucleus, is described. Structural and biophysical analyses demonstrated that the emissive analogues are faithful isomorphic nucleoside surrogates. Photophysical analysis established that the nucleosides offer highly desirable qualities, including visible emission, high quantum yield, and responsiveness to environmental perturbations, traits entirely lacking in their native counterparts.

The nonemissive nature of the nucleobases found in contemporary nucleic acids has triggered the development of emissive nucleoside analogues.[1] When carefully designed and employed, fluorescent nucleosides, nucleotides, and oligonucleotides can play an unparalleled role in exploring fundamental biochemical transformations and facilitate the fabrication of biophysical and discovery assays.[1,2] A key criterion for the design and successful implementation of such probes is to minimize structural and functional perturbation, which is an inevitable consequence of replacing any native residue with a synthetic probe. This important constraint has led to the development of emissive nucleoside analogues that display strong structural resemblance to their native counterparts, a characteristic we commonly describe as isomorphicity.[1,3]

Although several emissive nucleoside analogues have been reported over the past several years,[1] a complete alphabet of isomorphic fluorescent ribonucleosides derived from a single heterocyclic core has not been described.[4] Here we disclose the design and synthesis as well as the structural, photophysical, and preliminary biophysical characteristics of a complete ribonucleoside alphabet consisting of highly emissive purine (thA, thG) and pyrimidine (thU, thC) analogues, all derived from thieno[3,4-*d*]-pyrimidine as the heterocyclic nucleus (Figure 1). This parent heterocycle can be viewed as a precursor to 5,6-modified emissive pyrimidines and, at the same time, as a purine mimic with thiophene substituted for the imidazole moiety.[5] As schematically illustrated in Figure 1, glycosylation with D-ribose at N1 of the heterocycle is expected to yield expanded and emissive pyrimidine nucleoside analogues, while C-glycosylation at the thiophene's C2 position is anticipated to provide a viable route to fluorescent purine C-nucleoside analogues. Appropriate elaboration of the H-bonding faces provides emissive analogues of U, C, A, and G (Figure 1).[6]

The preparation of two analogues is briefly discussed.[7,8] The synthesis of thG started with methyl 4-aminothiophene-3-carboxylate hydrochloride (Scheme 1a). Treatment with chloroformamidine

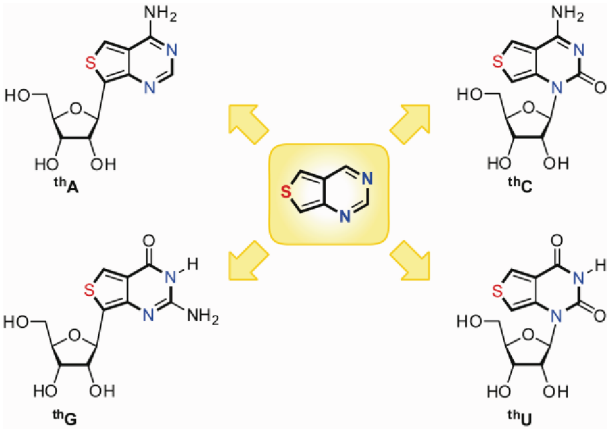
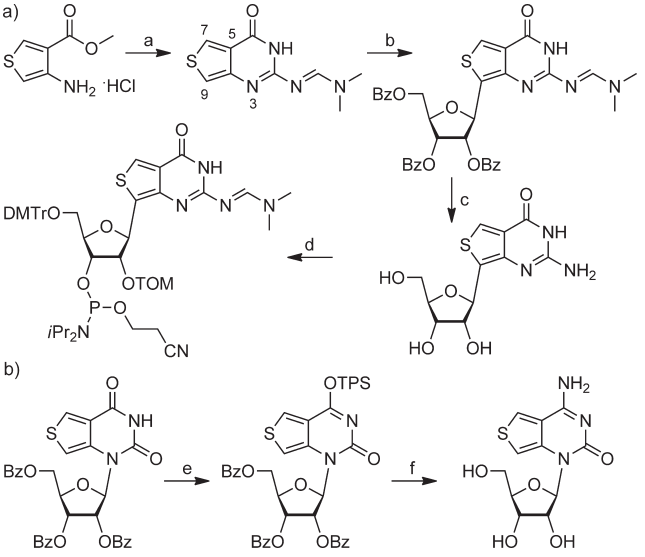


Figure 1. Emissive RNA Alphabet.

Scheme 1. Syntheses of thG and thC[a]



[a] Reagents and conditions: (a) (i) Chloroformamidine hydrochloride, $DMSO_2$, 125 °C, 77%; (ii) dimethylformamide dimethyl acetal, DMF, 98%. (b) β-D-Ribofuranose 1-acetate 2,4,5-tribenzoate, $SnCl_4$, $MeNO_2$, 65 °C, 64%. (c) NH_3/MeOH, 65 °C, 84%. (d) (i) Dimethylformamide dimethyl acetal, DMF, 77%; (ii) DMTrCl, Py, 70%; (iii) TOMCl, tBu_2SnCl_2, iPr_2NEt, DCE, 13%. (iv) 2-Cyanoethyl *N*,*N*-diisopropylchlorophosphoramidite, iPr_2NEt, DCM, 0 °C—RT, 67%. (e) 2,4,6-Triisopropylbenzenesulfonyl chloride, TEA, DMAP, CH_2Cl_2. (f) NH_3/MeOH, 70 °C, 37% over two steps.[9,13]

Received: June 30, 2011
Published: August 25, 2011

Table 1. Structural[a] and Photophysical[b] Data for Thieno[3,4-*d*]pyridmidine Nucleoside Analogues

		rmsd (Å)[a]								
	sugar pucker	ribose	base	solvent	λ_{abs} (ε)	λ_{em} (Φ)	$\Phi\varepsilon$	τ	Stokes shift	polarity sensitivity[c]
th**A**	C3′-endo	0.0521	0.157	water	341 (7.44)	420 (0.21)	1562	3.9	5950	68.9
				dioxane	345 (7.83)	411 (0.14)	1096	3.2	5080	
th**C**	C2′-endo	0.294	0.045	water	320 (4.53)	429 (0.41)	1857	15.2	8300	27.3
				dioxane	326 (4.21)	422 (0.01)	42	5.0	7550	
th**G**	C2′-endo	0.0525	0.158	water	321 (4.15)	453 (0.46)	1909	14.8	9580	107.2
				dioxane	333 (4.53)	424 (0.50)	2265	13.0	6890	
th**U**	C1′-exo	0.240	0.047	water	304 (3.16)	409 (0.41)	1296	11.5	8860	80.8
				dioxane	304 (3.50)	378 (0.04)	140	1.0	6690	

[a] Rmsd values were calculated by overlaying the X-ray crystal structures of the natural nucleosides (adenosine-ADENOS10, cytidine-CYTIDI10, guanosine-GUANSH10 and uridine-BEURID10 from the CCDC) with the structures of the synthetic analogues reported here.[9] For th**C** and th**U**, the corresponding pyrimidine regions were overlaid. th**G**^{dmf} was crystallized and used[15] [b] λ_{abs}, λ_{em}, ε, τ, and Stokes shift are reported in nm, nm, 10^3 M^{-1} cm^{-1}, ns, and cm^{-1}, respectively. All photophysical values reflect the average of two independent measurements, and the standard errors of the mean values are smaller than 1, 0.26, and 0.5 for Φ, Stokes shift, and τ, respectively. [c] Sensitivity to polarity, expressed in $cm^{-1}/(kcal\ mol^{-1})$, is equal to the slope of the linearization depicted in Figure 4c,d.

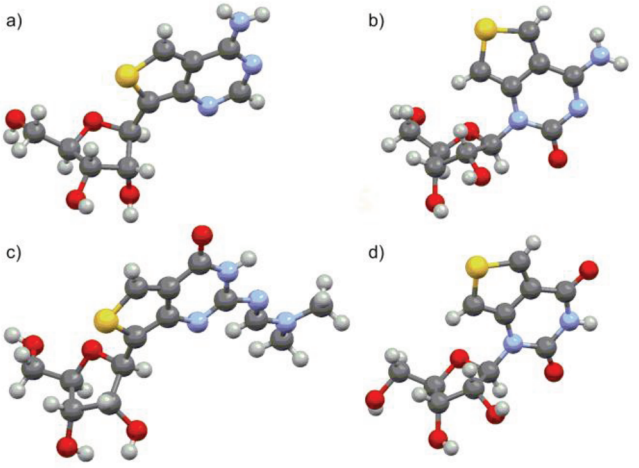


Figure 2. X-ray crystal structures of thieno[3,4-*d*]pyrimidine analogues: (a) th**A**; (b) th**C**; (c) th**G**^{dmf}; (d) th**U**.[15]

hydrochloride in dimethylsulfone at 125 °C afforded 2-aminothieno[3,4-*d*]pyrimidin-4(3*H*)-one, which was protected as the dimethylformamidine derivative.[9] This protecting group facilitated the regioselective Friedel−Crafts C-glycosylation affording the β-isomer only, an outcome which is similar to a Vorbrüggen glycosylation reaction.[10,11] Crystal structure analysis confirmed the assigned stereochemistry (see below).[12] The corresponding nucleoside th**G** was obtained by removing all of the protecting groups in methanolic ammonia. For the preparation of th**C**, the uridine analogue th**U**[7] was treated with 2,4,6-triisopropylbenzenesulfonyl chloride, activating position O4 for nucleophilic displacement (Scheme 1b). Treatment with saturated methanolic ammonia afforded th**C** in 37% yield over two steps.[13]

Crystal structures of the modified ribonucleosides showed that they all display an anti orientation at their glycosidic linkages (Figure 2), similar to the preference seen with their natural counterparts (Table 1). This is particularly notable with th**U** and th**C**, which possess a fused thiophene ring at the 5,6-position of the pyrimidine. This modification, however, does impact the observed sugar pucker, which appears to adopt an intermediate

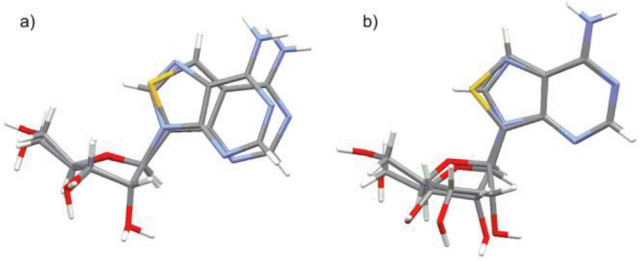


Figure 3. Comparison of the X-ray crystal structures of th**A** and adenosine (ADENOS10 from the CCDC): (a) overlay of the ribose rings (rmsd = 0.0521 Å); (b) overlay of the nucleobases (rmsd = 0.157 Å).

conformation partially resembling a C2′-endo conformation [Table 1; also see Tables S5 and S6 and Figures S4 and S6 in the Supporting Information (SI)]. In contrast, the purine analogues th**G** and th**A** exhibit C2′-endo and C3′-endo ribose pucker, respectively, in the solid state, which are the predominant conformations adopted by the natural ribonucleosides (Table 1).[9,14] Indeed, overlaying the crystal structure of th**A** with that of adenosine (Figure 3 and Figure S3) shows minimal root-mean-square deviation (rmsd) of the sugar pucker (0.0521 Å). Similarly, minimal distortion of the ribose conformation is seen for th**G** (rmsd = 0.0525 Å; Figure S5).[9,15]

The fundamental spectroscopic properties of the modified nucleosides th**U**, th**C**, th**A**, and th**G** are summarized in Table 1.[9] The ground-state absorption spectra in aqueous solutions showed long-wavelength maxima ranging from 304 nm (for th**U**) to 341 nm (for th**A**), all significantly red-shifted from those of the parent native nucleosides (Figure 4a). Excitation at the long-wavelength absorption maximum gave rise to visible emission with maxima ranging from 409 nm (for th**U**) to 453 nm (for th**G**).[9] Respectable to good quantum yields, ranging from 0.21 to 0.46 for th**A** and th**G**, respectively, were seen for all of the nucleosides, resulting in noteworthy brightness values (Table 1). Time-resolved experiments showed that the least-emissive nucleoside analogue (th**A**) displays a relatively short excited-state lifetime (3.9 ns), while the most emissive one (th**C**) has the longest (15.2 ns) (Figure S10). Spectra taken in dioxane typically

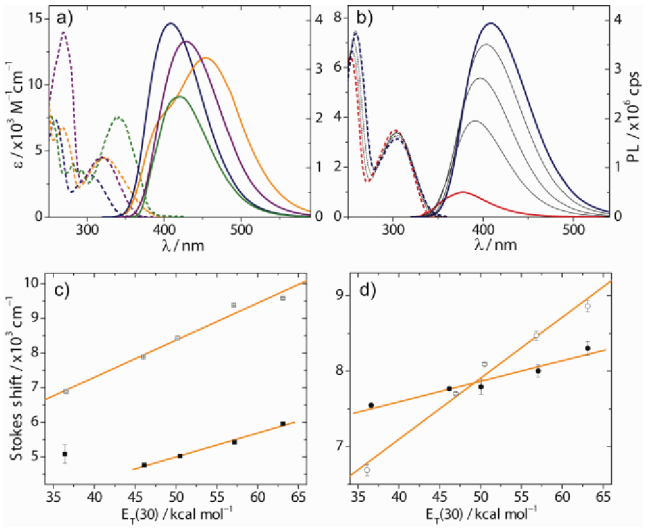


Figure 4. (a) Absorption (dashed lines) and emission (solid lines) spectra of th**A** (green), th**C** (purple), th**G** (orange), and th**U** (blue) in water. (b) Absorption (dashed lines) and emission (solid lines) spectra of th**U** in water (blue) and dioxane (red) and mixtures thereof (10, 30, and 70% v/v water in dioxane, black lines). Emission intensities were corrected to reflect an optical density of 0.1 at λ_{ex}. (c, d) Stokes shifts vs $E_T(30)$ correlations for (c) purines th**A** (■) and th**G** (□) and (d) pyrimidines th**C** (●) and th**U** (○). The orange lines represent linear fits. Note: the data point for th**A** in dioxane [$E_T(30)$ = 36.0 kcal mol^{-1}] was excluded from the linear fit.[9,18]

showed slightly bathochromically shifted absorption maxima and hypsochromically shifted emission maxima (Table 1), suggesting that these chromophores possess charge-transfer character in their excited states.[16] Sensitivity to environmental polarity, assessed by determining the Stokes shifts and emission intensities in water/dioxane mixtures, showed all of the nucleosides to be highly responsive, albeit to different extents (Figure 4b−d and Figures S7 and S8).[17,18] For example, while the impact of solvent polarity on the Stokes shifts of th**C** was modest, its impact on the emission quantum yield was dramatic (Table 1). In contrast, polarity substantially impacted the Stokes shifts displayed by th**G**, while a minimal effect on the emission quantum yield was seen (Table 1).

In a preliminary exploration of the impact of such isomorphic nucleotides (particularly the incorporation of a modified purine) on the stability of RNA duplexes, the 17-mer oligonucleotide **1a** containing th**G** at a central position was synthesized using the th**G** phosphoramidite (Scheme 1a) and compared to the unmodified strand **1b** (Figure 5).[9] Hybridization to the complementary oligonucleotide **2a** and all of the mismatched strands **2b**−**d** was followed by thermal denaturation experiments. Replacing a G residue with th**G** yielded a modified duplex **1a**/**2a** that was slightly more stable than the native unmodified one **1b**/**2a** (ΔT_m = +1.0 °C) (Figure 5). Importantly, an entirely analogous stability trend was seen for all of the mismatched th**G**-containing duplexes **1a**/**2b**−**d** relative to the corresponding native hybrids **1b**/**2b**−**d** (Figure 5b). This suggests minimal perturbation of the resulting duplexes, corroborating the highly isomorphic character of the emissive purine analogue.

One drawback of many emissive nucleoside analogues, including the classical 2-aminopurine,[1,19,20] is their significant

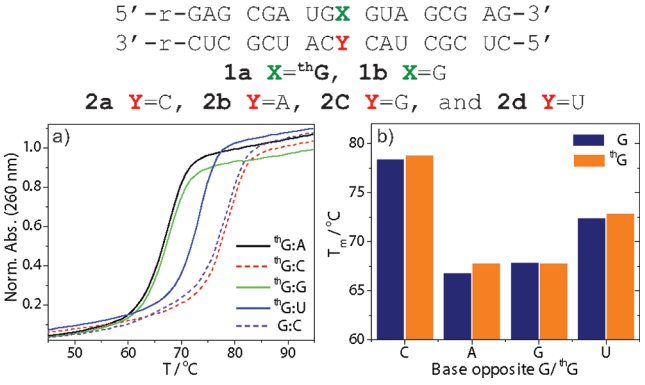


Figure 5. Thermal melts of th**G**-containing RNA. All samples contained each oligo at 1.0 μM, 20 mM sodium cacodylate, 100 mM NaCl, and 0.5 mM EDTA (pH 7.0): (a) absorbance vs temperature denaturation profiles; (b) comparison of T_m values for duplexes of th**G** (**1a**) and **G** (**1b**) with their complementary and mismatched bases. Errors based on two independent T_m measurements were less than 0.6 °C.

quenching upon incorporation into oligonucleotides, which is a result of diverse quenching pathways facilitated by neighboring nucleobases. Of significance therefore is the observation that oligonucleotide **1a**, having the emissive th**G** "sandwiched" between two potentially quenching G residues, still displayed strong visible emission with a relative quantum yield of 0.10 (Figure S13).[9,21]

The structural, biophysical, and spectroscopic characteristics of this set of emissive RNA nucleosides illustrate highly desirable traits, including native Watson−Crick faces, unparalleled structural isomorphicity with respect to native nucleosides, minimal perturbation upon incorporation into duplexes, and intense visible emission. In particular, we note that while numerous emissive pyrimidine analogues have been reported in recent years,[1,22−25] highly emissive and responsive isomorphic purines are rare, with the classical 2-aminopurine, introduced in 1969,[19a] still being the most prevalently used.[26] Expanding the repertoire of isomorphic purine analogues, as demonstrated here for the thieno[3,4-*d*]-pyrimidine-based th**A** and th**G** nucleosides, is therefore significant and is expected to advance the biophysical probing of RNA-related processes.[27]

■ ASSOCIATED CONTENT

Supporting Information. Synthetic details, X-ray crystallographic data (CIF), photophysical data, thermal denaturation measurements, and MALDI−TOF MS spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author
ytor@ucsd.edu

■ ACKNOWLEDGMENT

We thank the National Institutes of Health for generous support (GM 069773) and the National Science Foundation for an instrumentation grant (CHE-0741968).

REFERENCES

(1) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579.

(2) (a) Okamoto, A.; Saito, Y.; Saito, I. *J. Photochem. Photobiol., C* **2005**, *6*, 108. (b) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265. (c) Dodd, D. W.; Hudson, R. H. E. *Mini-Rev. Org. Chem.* **2009**, *6*, 378. (d) Wilhelmsson, L. M. *Q. Rev. Biophys.* **2010**, *43*, 159. (e) Kimoto, M.; Cox, R. S., III; Hirao, I. *Expert Rev. Mol. Diagn.* **2011**, *11*, 321.

(3) (a) Greco, N. J.; Tor, Y. *J. Am. Chem. Soc.* **2005**, *127*, 10784. (b) Tor, Y.; Del Valle, S.; Jaramillo, D.; Srivatsan, S. G.; Rios, A.; Weizman, H. *Tetrahedron* **2007**, *63*, 3608. (c) Tor, Y. *Pure Appl. Chem.* **2009**, *81*, 263.

(4) Kool has reported a complete set of emissive expanded RNA nucleoside analogues. See: Hernández, A. R.; Kool, E. T. *Org. Lett.* **2011**, *13*, 676.

(5) Fused heteroaromatics are known to be emissive. See: (a) Wettack, F. S.; Klapthor, R.; Shedd, A.; Koeppe, M.; Janda, K.; Dwyer, P.; Stratton, K. *NBS Spec. Publ. (U. S.)* **1978**, *526*, 60. (b) Zander, M.; Kirsch, G. *Z. Naturforsch.* **1989**, *44a*, 205.

(6) It should be noted that related nucleosides have previously been explored for their antiviral and chemotherapeutic potential but never as fluorescent nucleoside analogues. See: (a) Rao, S. P.; Rao, K. V. B.; Otter, B. A.; Klein, R. S.; Ren, W.-Y. *Tetrahedron Lett.* **1988**, *29*, 3537. (b) Patil, S. A.; Otter, B. A.; Klein, R. S. *Nucleosides Nucleotides* **1990**, *9*, 937. (c) Patil, S. A.; Otter, B. A.; Klein, R. S. *J. Heterocycl. Chem.* **1993**, *30*, 509.

(7) We previously reported the synthesis of the th**U** analogue. See: (a) Srivatsan, S. G.; Weizman, H.; Tor, Y. *Org. Biomol. Chem.* **2008**, *6*, 1334. (b) Srivatsan, S. G.; Greco, N. J.; Tor, Y. *Angew. Chem., Int. Ed.* **2008**, *47*, 6661.

(8) The synthesis of the th**A** analogue closely followed Klein's report.[6c]

(9) See the SI for additional data.

(10) Other protecting groups (e.g., pivaloyl) at N^2 hampered this Friedel–Crafts reaction, leading to undesired products.

(11) For a general review of C-nucleosides, see: Stambasky, J.; Hocek, M.; Kocovsky, P. *Chem. Rev.* **2009**, *109*, 6729.

(12) See the SI for an HMBC experiment.

(13) Unmodified th**U** (18% over two steps) was also recovered (see the SI).

(14) Altona, C.; Sundaralingam, M. *J. Am. Chem. Soc.* **1972**, *94*, 8205.

(15) Attempts to obtain suitable crystals of the unprotected th**G** were unsuccessful.

(16) Solvent-induced tautomerization can also contribute to solvent sensitivity.

(17) (a) Sinkeldam, R. W.; Tor, Y. *Org. Biomol. Chem.* **2007**, *5*, 2523. (b) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *ChemBioChem* **2008**, *9*, 706.

(18) It is not uncommon to observe the greatest deviation from linearity in pure solvents, since solvent mixtures tend to suppress specific solvent–solute interactions present only in the former. See ref 17a for a discussion related to the effects of pure solvents and their mixtures on the photophysical features of fluorophores.

(19) (a) Ward, D. C.; Reich, E.; Stryer, L. *J. Biol. Chem.* **1969**, *244*, 1228. (b) Jean, J. M.; Hall, K. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 37. (c) Jean, J. M.; Hall, K. B. *Biochemistry* **2002**, *41*, 13152.

(20) (a) Kirk, S. R.; Luedtke, N. W.; Tor, Y. *Bioorg. Med. Chem.* **2001**, *9*, 2295. (b) Tam, V. K.; Kwong, D.; Tor, Y. *J. Am. Chem. Soc.* **2007**, *129*, 3257.

(21) The relative quantum yield was determined by comparing the emission of **1a** to that of th**G** under analogous conditions (see the SI).

(22) (a) Greco, N. J.; Sinkeldam, R. W.; Tor, Y. *Org. Lett.* **2009**, *11*, 1115. (b) Xie, Y.; Dix, A. V.; Tor, Y. *J. Am. Chem. Soc.* **2009**, *131*, 17605. (c) Xie, Y.; Dix, A. V.; Tor, Y. *Chem. Commun.* **2010**, *46*, 5542. (d) Xie, Y.; Maxson, T.; Tor, Y. *J. Am. Chem. Soc.* **2010**, *132*, 11896. (e) Sinkeldam, R. W.; Wheat, A. J.; Boyaci, H.; Tor, Y. *ChemPhysChem* **2010**, *12*, 567. (f) Sinkeldam, R. W.; Marcus, P.; Uchenik, D.; Tor, Y. *ChemPhysChem* **2011**, *12*, 2260.

(23) (a) Robins, M. J.; Barr, P. J. *Tetrahedron Lett.* **1981**, *22*, 421. (b) Robins, M. J.; Barr, P. J. *J. Org. Chem.* **1983**, *48*, 1854. (c) Liu, C. H.; Martin, C. T. *J. Mol. Biol.* **2001**, *308*, 465. (d) Hudson, R. H. E.; Choghamarani, A. G. *Nucleosides, Nucleotides Nucleic Acids* **2007**, *26*, 533.

(24) (a) Okamoto, A.; Tainaka, K.; Saito, I. *J. Am. Chem. Soc.* **2003**, *125*, 4972. (b) Okamoto, A.; Tainaka, K.; Saito, I. *Tetrahedron Lett.* **2003**, *44*, 6871.

(25) Wilhelmsson, L. M.; Holmén, A.; Lincoln, P.; Nielsen, P. E.; Nordén, B. *J. Am. Chem. Soc.* **2001**, *123*, 2434.

(26) A few emissive purine analogues have been reported, but no biophysical applications have been described. See: (a) Greco, N. J.; Tor, Y. *Tetrahedron* **2007**, *63*, 3515. (b) Dyrager, C.; Borjesson, K.; Diner, P.; Elf, A.; Albinsson, B.; Wilhelmsson, L. M.; Grotli, M. *Eur. J. Org. Chem.* **2009**, 1515. (c) Dierckx, A.; Diner, P.; El-Sagheer, A. H.; Kumar, J. D.; Brown, T.; Grotli, M.; Wilhelmsson, L. M. *Nucleic Acids Res.* **2011**, *39*, 4513.

(27) The relatively high quantum yields of the emissive nucleosides and their sensitivity to environmental polarity suggest utility in diverse biophysical and discovery assays, which include, for example, monitoring of RNA conformational changes as well as RNA–protein and RNA–small-molecule interactions.