Published in final edited form as:

J Am Chem Soc. 2011 September 28; 133(38): 14912–14915. doi:10.1021/ja206095a.

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

Abstract

A fluorescent ribonucleoside alphabet, comprised of highly emissive purine (**thA**, **thG**) and pyrimidine (**thU**, **thC**) analogs, all derived from thieno[3,4-*d*]pyrimidine as the heterocyclic nucleus, is described. Structural and biophysical analyses demonstrate that the emissive analogs are faithful isomorphic nucleoside surrogates. Photophysical analysis establishes 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 non-emissive nature of the nucleobases found in contemporary nucleic acids has triggered the development of emissive nucleoside analogs.¹ When carefully designed and employed, fluorescent nucleosides, nucleotides and oligonucleotides serve 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 analogs that display high structural resemblance to their native counterparts, a characteristic we commonly describe as isomorphism.^{1,3}

Although several emissive nucleoside analogs have been reported over the past several years,¹ a complete alphabet of isomorphic fluorescent ribonucleoside derived from a single heterocyclic core, has not been described.⁴ Here we disclose the design and synthesis, as well as structural, photophysical and preliminary biophysical characteristics of a complete ribonucleoside alphabet, comprised of highly emissive purine (**thA**, **thG**) and pyrimidine (**thU**, **thC**) analogs, 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 substituting 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 analogs, while C-glycosylation at the thiophene's C2 position is anticipated to provide a viable route to fluorescent purine C-nucleoside analogs. Appropriate elaboration of the H-bonding faces provides emissive analogs of U, C, A and G (Figure 1).⁶

The preparation of two analogs is briefly discussed.^{7,8} The synthesis of **thG** started with methyl 4-aminothiophene-3-carboxylate hydrochloride (Scheme 1a). Treatment with chloroformamidine hydrochloride in dimethylsulfone at 125 °C afforded 2-aminothieno[3,4-*d*]pyrimidin-4(3*H*)-one, which was protected as the dimethyl formamidine derivative.⁹ This

Corresponding Author ytor@ucsd.edu.

ASSOCIATED CONTENT

Supporting Information. Synthetic details, X-ray crystallography, photophysical data, thermal denaturation measurements, and MALDI-TOF MS spectra. This information is available free of charge via the Internet at <http://pubs.acs.org>.

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 **thG** was obtained by removing all protecting groups in methanolic ammonia. For the preparation of **thC**, the uridine analog **thU**⁷ was treated with 2,4,6-triisopropylbenzenesulfonyl chloride, activating position O4 for nucleophilic displacement (Scheme 1b). Treatment with saturated methanolic ammonia afforded **thC** in 37 % yield over 2 steps.¹³

Crystal structures of the modified ribonucleosides show they all display an *anti* orientation at their glycosidic linkage (Figure 2), similar to the preference seen with their natural counterparts (Table 1). This is particularly notable with **thU** and **thC**, which possess a fused thiophene ring at the pyrimidine's 5,6-position. This modification, however, does impact the observed sugar pucker, which appears to adopt an intermediate conformation partially resembling a C2'-endo conformation (Tables 1, S5, S6, and Figures S4, S6). In contrast, the purine analogs **thG** and **thA**, exhibit C2'-endo and C3'-endo ribose pucker in the solid state, respectively, which are the predominant conformations adopted by the natural ribonucleosides (Table 1).^{9,14} Indeed, overlaying the crystal structure of **thA** with that of adenosine (Figures 3, S3) shows minimal deviation of the sugar pucker (rms = 0.0521 Å). Similarly, minimal distortion of the ribose conformation is seen for **thG** (rms = 0.0525 Å; Figure S5).⁹

The fundamental spectroscopic properties of the modified nucleosides **thU**, **thC**, **thA** and **thG** are summarized in Table 1.⁹ The ground state absorption spectra in aqueous solutions show long wavelength maxima ranging from 304 nm (for **thU**) to 341 nm (for **thA**), all significantly red-shifted from the parent native nucleosides (Figure 4a). Excitation at the long wavelength absorption maximum gives rise to visible emission with maxima ranging from 409 nm (for **thU**) to 458 nm (for **thG**).⁹ Respectable to good quantum yields, ranging from 0.21 to 0.46 for **thA** and **thG**, respectively, are seen for all nucleosides, resulting in noteworthy brightness values (Table 1). Time resolved experiments show that the least emissive nucleoside analog (**thA**) displays a relatively short excited state life-time (3.9 ns), while the most emissive one (**thC**) has the longest (15.2 ns) (Figure S10). Spectra taken in dioxane typically show slightly bathochromically shifted absorption maxima and hypsochromically shifted emission maxima (Table 1), suggesting these chromophores possess a 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 nucleosides to be highly responsive, albeit to different extents (Figures 4b–d, S7, S8).^{17,18} For example, while the impact of solvent polarity on the Stokes shifts of **thC** is modest, its impact on the emission quantum yield is dramatic (Table 1). In contrast, polarity substantially impacts the Stokes shifts displayed by **thG**, while a minimal effect is seen on the emission quantum yield (Table 1).

To preliminarily explore the impact of such isomorphous nucleotides, and particularly the incorporation of a modified purine, on the stability of RNA duplexes, the 17-mer oligonucleotide **1a**, placing a **thG** at a central position was synthesized using the **thG** phosphoramidite (Scheme 1a), and compared to the unmodified strand **1b** (Figure 5).⁹ Hybridization to the complementary oligonucleotide **2a** and all mismatched strands **2b–d**, was followed by thermal denaturation experiments. Replacing a G residue with **thG** yields a modified duplex **1a/2a**, which is slightly more stable than the native unmodified one **1b/2a** ($\Delta T_m = +1.0$ °C) (Figure 5). Importantly, an entirely analogous stability trend is seen for all mismatched **thG**-containing duplexes **1a/2b–d** when compared to the corresponding native hybrids **1b/2b–d** (Figure 5b). This suggests minimal perturbation of the resulting duplexes, and corroborates a highly isomorphous character for such an emissive purine analog.

One drawback of many emissive nucleoside analogs, including the classical 2-aminopurine,^{1,19,20} is their significant quenching upon incorporation into oligonucleotides, a result of diverse quenching pathways facilitated by neighboring nucleobases. Of significance is therefore the observation that oligonucleotide **1a**, having the emissive thG “sandwiched” between two potentially quenching G residues, still displays strong visible emission, with a relative quantum yield of 0.10 (Figure S13).^{9,21}

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

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

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 RW, Greco NJ, Tor Y. *Chem. Rev.* 2010; 110:2579–2619. [PubMed: 20205430]
2. (a) Okamoto A, Saito Y, Saito I. *J. Photochem. Photobiol. C: Photochem. Rev.* 2005; 6:108–122. (b) Wilson JN, Kool ET. *Org. Biomol. Chem.* 2006; 4:4265–4274. [PubMed: 17102869] (c) Dodd DW, Hudson RHE. *Mini-Reviews in Organic Chemistry.* 2009; 6:378–391. (d) Wilhelmsson LMQ. *Rev. Biophys.* 2010; 43:159–183. (e) Kimoto M, Cox RS III, Hirao I. *Exp. Rev. Molec. Diagnos.* 2011; 11:321–331.
3. (a) N. Greco NJ, Tor Y. *J. Am. Chem. Soc.* 2005; 127:10784–10785. [PubMed: 16076156] (b) Tor Y, Del Valle S, Jaramillo D, Srivatsan SG, Rios A, Weizman H. *Tetrahedron.* 2007; 63:3608–3614. (c) Tor Y. *Pure & Appl. Chem.* 2009; 81:263–272.
4. Kool has reported a complete set of emissive expanded RNA nucleoside analogs Hernández AR, Kool ET. *Org. Lett.* 2011; 13:676–679. [PubMed: 21214222]
5. Fused heteroaromatics are known to be emissive. See: Wettack FS, Klapthor R, Shedd A, Koeppe M, Janda K, Dwyer P, Stratton K. *The photophysics of several condensed ring heteroaromatic compounds.* NBS Special Publication (United States). 1978; 526:60–62. Zander M, Kirsch GZ. *Naturforsch.* 1989; 44a:205–209.
6. Note, related nucleosides have previously been explored for their antiviral and chemotherapeutic potential, but never as fluorescent nucleoside analogs See: (a) Rao SP, Rao KVB, Otter BA, Klein RS. *Tetrahedron Lett.* 1988; 29:3537–3540. (b) Patil SA, Otter BA, Klein RS. *Nucleos. Nucleot.* 1990; 9:937–956. (c) Patil SA, Otter BA, Klein RS. *J Heterocyclic. Chem.* 1993; 30:509–515.
7. We have previously reported the synthesis of the thU analog (a) Srivatsan SG, Weizman H, Tor Y. *Org. Biomol. Chem.* 2008; 6:1334–1338. [PubMed: 18385838] (b) Srivatsan SG, Greco NJ, Tor Y. *Angew. Chem. Int. Ed.* 2008; 47:6661–6665.
8. The synthesis of the thA analog closely follows Klein’s report.^{6c}
9. See SI for additional data.

10. Other protecting groups, such as pivaloyl, at N^2 hampered this Friedel-Crafts reaction leading to undesired products.
11. For a general review on C-nucleosides, see: Stambasky J, Hocek M, Kocovsky P. *Chem. Rev.* 2009; 109:6729–6764. [PubMed: 19761208]
12. See SI for an HMBC experiment.
13. Unmodified ^{13}C U (18% over two steps) was also recovered; see SI.
14. Altona C, Sundaralingam M. *J. Am. Chem. Soc.* 1972; 94:8205–8212. [PubMed: 5079964]
15. Attempts to obtain suitable crystals of the unprotected ^{13}C G have so far been unsuccessful.
16. Solvent induced tautomerization can also contribute to solvent sensitivity.
17. (a) Sinkeldam RW, Tor Y. *Org. Biomol. Chem.* 2007; 5:2523–2528. [PubMed: 18019524] (b) Sinkeldam RW, Greco NJ, Tor Y. *Chembiochem.* 2008; 9:706–709. [PubMed: 18286575]
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 effects of pure solvents and their mixtures on the photophysical features of fluorophores.
19. (a) Ward DC, Reich E, Stryer L. *J Biol. Chem.* 1969; 244:1228–1237. [PubMed: 5767305] (b) Jean JM, Hall KB. *Proc. Natl. Acad. Sci. USA.* 2001; 98:37–41. [PubMed: 11120885] (c) Jean JM, Hall KB. *Biochemistry.* 2002; 41:13152–13161. [PubMed: 12403616]
20. (a) Kirk SR, Luedtke NW, Tor Y. *Bioorg Med. Chem.* 2001; 9:2295–2301. [PubMed: 11553468] (b) Tam V, Kwong D, Tor Y. *J. Am. Chem. Soc.* 2007; 129:3257–3266. [PubMed: 17319662]
21. Relative quantum yield, determined by comparing the emission of **1a** to that of ^{13}C G under analogous conditions (See SI).
22. (a) Greco NJ, Sinkeldam RW, Tor Y. *Org. Lett.* 2009; 11:1115–1118. [PubMed: 19196162] (b) Xie Y, Dix AV, Tor Y. *J. Am. Chem. Soc.* 2009; 131:17605–17614. [PubMed: 19908830] (c) Xie Y, Dix AV, Tor Y. *Chem. Commun.* 2010; 46:5542–5544. (d) Xie Y, Maxson T, Tor Y. *J. Am. Chem. Soc.* 2010; 132:11896–11897. [PubMed: 20690779] (e) Sinkeldam RW, Wheat AJ, Boyaci H, Tor Y. *ChemPhysChem.* 2010; 12:567–570. [PubMed: 21344595] (f) Sinkeldam RW, Marcus P, Uchenik D, Tor Y. *ChemPhysChem.* 2011; 12:2260–2265. [PubMed: 21698743]
23. (a) Robins MJ, Barr PJ. *Tetrahedron Lett.* 1981; 22:421–424. (b) Robins MJ, Barr PJ. *J. Org. Chem.* 1983; 48:1854–1862. (c) Liu CH, Martin CT. *J. Mol. Biol.* 2001; 308:465–475. [PubMed: 11327781] (d) Hudson RHE, Choghamarani AG. *Nucleos. Nucleot. Nucl.* 2007; 26:533–537.
24. (a) Okamoto A, Tainaka K, Saito I. *J. Am. Chem. Soc.* 2003; 125:4972–4973. [PubMed: 12708835] (b) Okamoto A, Tainaka K, Saito I. *Tetrahedron Lett.* 2003; 44:6871–6874.
25. Wilhelmsson LM, Holmén A, Lincoln P, Nielsen PE, Nordén B. *J Am. Chem. Soc.* 2001; 123:2434–2435. [PubMed: 11456897]
26. A few emissive purine analogs have been reported but no biophysical applications have been described, see: (a) Greco NJ, Tor Y. *Tetrahedron.* 2007; 63:3515–3527. [PubMed: 18431439] (b) Dyrager C, Borjesson K, Diner P, Elf A, Albinsson B, Wilhelmsson LM, Grotli M. *Eur. J. Org. Chem.* 2009:1515–1521. (c) Dierckx A, Diner P, El-Sagheer AH, Kumar JD, Brown T, Grotli M, Wilhelmsson LM. *Nucleic Acids Res.* 2011; 39:4513–4524. [PubMed: 21278417]
27. The relatively high quantum yield of the emissive nucleosides and their sensitivity to environmental polarity suggest utility in diverse biophysical and discovery assays, which include, for example, monitoring RNA conformational changes as well as RNA–protein and RNA–small molecule interactions.

Figure 1.
Emissive RNA Alphabet.

Figure 2.
X-ray crystal structures of thieno[3,4-*d*]pyrimidine analogs: a) **thA**, b) **thC**, c) **thG^dmf** and d) **thU**.¹⁵

Figure 3. Comparison of the X-ray crystal structure of **thA** with adenosine (ADENOS10 from CCDC): a) overlay of the ribose rings (rms 0.0521 Å), b) overlay of the nucleobases (rms 0.157 Å).

Figure 4.

a) Absorption (dashed lines) and emission (solid lines) spectra of **thA** (green), **thC** (purple), **thG** (orange), and **thU** (blue) in water. b) Absorption (dashed lines) and emission (solid lines) spectra of **thU** in water (blue) and dioxane (red) and mixtures thereof (10, 30, and 70 v/v % water in dioxane, black lines). Emission intensities are corrected to reflect an optical density of 0.1 at λ_{ex} . c) Stokes shift vs. $E_T(30)$ correlations for purines **thA** (solid squares) and **thG** (open squares), and d) Stokes shift vs. $E_T(30)$ correlations for pyrimidines **thC** (solid circles) and **thU** (open circles). The orange lines represent linear fits. Note: the data point for **thA** in dioxane [$E_T(30) = 36.0 \text{ kcal mol}^{-1}$] has been excluded from linearization.^{9,18}

Figure 5.

Thermal melts of thG-containing RNA. All samples are containing 1.0 uM of each oligo, 20 mM Na cacodylate, 100 mM NaCl, 0.5 mM EDTA and pH 7.0. a) Absorbance vs. temperature denaturation profiles, b) Comparison of T_m values for duplexes of containing thG (**1a**) and G (**1b**) with their complementary and mismatched bases. Errors of two independent T_m measurements were less than 0.6 °C.

Scheme 1. Syntheses of **thG** and **thC.a**

^a *Reagents and conditions:* (a) (i) chloroformamidine hydrochloride, DMSO₂, 125 °C, 77 %; (ii) dimethylformamide dimethyl acetal, DMF, 98 %; (b) β-D-ribofuranose 1-acetate 2,4,5-tribenzoate, SnCl₄, MeNO₂, 65 °C, 64 %; (c) NH₃/MeOH, 65 °C, 84 %; (d) (i) dimethylformamide dimethyl acetal, DMF, 77 %; (ii) DMTrCl, Py, 70 %; (iii) TOMCl, *t*Bu₂SnCl₂, *i*Pr₂NEt, DCE, 13 %; (iv) 2-cyanoethyl *N,N*-diisopropylchloro phosphoramidite, *i*Pr₂NEt, DCM, 0 °C–RT, 67%.; (e) 2,4,6-triisopropylbenzenesulfonyl chloride, TEA, DMAP, CH₂Cl₂; (f) NH₃/MeOH, 70 °C, 37 % over 2 steps.^{9–13}

Table 1

Structural^a and Photophysical^b data of thieno[3,4-*d*]pyrimidine nucleoside analogs

	Sugar pucker	rms ($\text{\AA}/^\circ$)		Solvent	$\lambda_{\text{abs}}(\text{e})$	$\lambda_{\text{em}}(\Phi)$	$\Phi \times \epsilon$	τ	Stokes Shift	Polarity sensitivity ^c
		Ribose	Base							
thA	C3'-endo	0.0521	0.157	water	341 (7.44)	420 (0.21)	1562	3.9	5,950	73.6
				dioxane	345 (7.83)	411 (0.14)	1096	3.2	5,080	
thC	C2'-endo	0.294	0.045	water	320 (4.53)	429 (0.41)	1857	15.2	8,300	27.3
				dioxane	326 (4.21)	422 (0.01)	42	5.0	7,550	
thG	C2'-endo	0.0525	0.158	water	321 (4.15)	458 (0.46)	1909	14.8	9,580	107.2
				dioxane	333 (4.53)	424 (0.50)	2265	13.0	6,890	
thU	C1'-exo	0.240	0.047	water	304 (3.16)	409 (0.41)	1296	11.5	8,860	80.8
				dioxane	304 (3.50)	378 (0.04)	140	1.0	6,690	

^a rms values were calculated by overlaying the X-ray crystal structure of the natural nucleosides (adenosine-ADENOS10, cytidine-CYTID10, guanosine-GUANSH10 and uridine-BEURID10 from CCDC), with the structures of the synthetic analogs reported here.⁹ For **thC** and **thU** the corresponding pyrimidine region was overlaid. **thG** was crystalized and used.¹⁵

^b λ_{abs} , λ_{em} , ϵ , τ and Stokes shift are reported in nm, $\times 10^3 \text{ M}^{-1} \text{ cm}^{-1}$, ns, and cm^{-1} , respectively. All photophysical values reflect the average of two independent measurements and standard errors of mean values are smaller than 1, 0.26, and 0.5 for Φ , Stokes shift, and τ , respectively.

^c Sensitivity to polarity, expressed in $\text{cm}^{-1}/\text{Kcal mol}^{-1}$ is equal to the slope of the linearization depicted in Figures 4c and 4d