Chemical Mutagenesis of an Emissive RNA Alphabet

Chemical mutagenesis of an emissive RNA alphabet

Alexander R. Rovira, Andrea Fin, and Yitzhak Tor*

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

Supporting Information Placeholder

ABSTRACT: An evolved fluorescent ribonucleoside alphabet comprising of isomorphic purine (**^{tz}A**, **^{tz}G**) and pyrimidine (**^{tz}U**, **^{tz}C**) analogs, all derived from isothiazolo[4,3-*d*]pyrimidine as a common heterocyclic core, is described. Structural and biochemical analyses illustrate that the nucleosides, particularly the *C*-nucleosidic purine analogs, are faithful isomorphic and isofunctional surrogates of their natural counterparts and show improved features when compared to an RNA alphabet derived from thieno[3,4-*d*]-pyrimidine. The restoration of the nitrogen in a position equivalent to the purines' N7 leads to "isofunctional" behavior, as illustrated by the ability of adenosine deaminase to deaminate **^{tz}A** as effectively as adenosine, the native substrate.

Fluorescent nucleoside analogs, which have been developed to address the non-emissive nature of the nucleobases found in DNA and RNA, have found great utility in biophysical analysis and discovery assays.[1] In addition to favorable photophysical features, a key element that dictates their utility and ultimate use is their structural similarity to the native counterparts. Characterized as their isomorphic nature, researchers have tried to maximize this trait by minimizing changes to the Watson-Crick (WC) pairing face while introducing fluorescence-enhancing electronic perturbations.[2] The impact of these alterations on the physical properties (e.g., tautomerization) and photophysics (excited state level and dynamics) are, however, frequently unpredictable and ultimately empirically assessed.

We have previously completed the first emissive RNA alphabet,[3] comprised of pyrimidine and purine analogs, all fundamentally derived from thieno[3,4-*d*]-pyrimidine as a common heterocyclic nucleus (Figure 1). While highly emissive and valuable as illustrated by several applications,[4] a limitation of the previously reported thieno analogs is the lack of the basic nitrogen, corresponding to N7 in the purine skeleton (Figure 1).[5] As many biomolecular interactions of purine nucleosides and nucleotides rely on the basicity and coordinating ability of this position, we have sought to reinstall this functionality into a new isomorphic RNA alphabet, with higher structural and electronic similarity to the native purines. Herein we report the synthesis, photophysical analysis, and performance of a next generation alphabet based on an isothiazolo[4,3-*d*]pyrimidine core. This atomic mutagenesis reinstates the basic moiety at the native position yielding analogs that better mimic a purine (Figure 1). This chemically evolved isomorphic alphabet was found to not only have unique photophysical features but also to be highly "isofunctional", illustrated by the ability of adenosine deaminase (ADA) to deaminate the adenosine analog as effectively as adenosine, the native substrate.

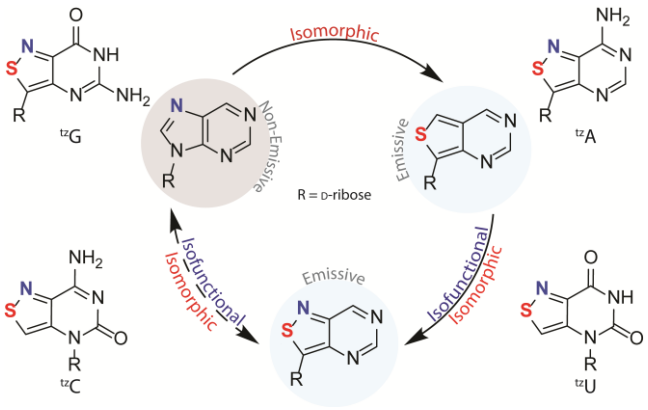
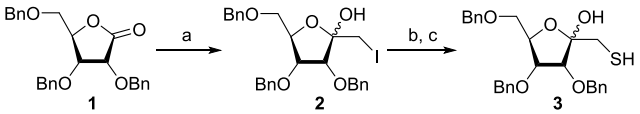


Figure 1. Evolution of an emissive RNA alphabet.

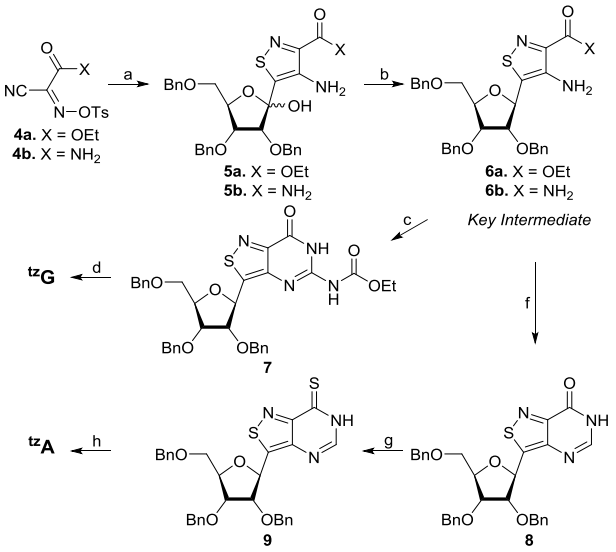
While standard conditions provided the isothiazole uridine and cytidine analogues **^{tz}U** and **^{tz}C**, respectively,[6] common approaches to forming the carbon–carbon glycosidic bonds of the adenosine and guanosine analogs failed under a wide variety of conditions.[7,8] A less common approach was required to form the desired connectivity.[9] A strategy starting from a ribofuranose-derived precursor as opposed to late-stage glycosylation with the nucleobase and desired sugar was conceived. To construct the *C*-glycosylated isothiazole ring, a ribofuranose derivative substituted with a primary thiol was synthesized (Scheme 1), starting from a known benzyl-protected precursor, which was subjected to Swern oxidation conditions.[10] Treatment with diiodomethane and methyllithium furnished the primary halide,[11] which was reacted with potassium thioacetate and reduced to give the primary thiol (Scheme 1).

Scheme 1. Synthesis of the sugar precursor[a]



[a] *Reagents and conditions*: (a) CH_2I_2, MeLi, toluene, –78 °C, 1 h, 67%. (b) Potassium thioacetate, DMF, rt, 6 h, 74%. (c) Et_2O, $LiAlH_4$, 0 °C to rt, 1 h, >90%.

Scheme 2. Synthesis of purine analogs[a]



[a] *Reagents and conditions*: (a) **3,** EtOH, morpholine, 5h, 0 °C to rt, 67% (**5a**), 79% (**5b**). (b) Et_3SiH, $BF_3.OEt_2$ DCM, –78 °C to rt, 4 h, 63% (**6a**), 79% (**6b**). (c) i. EtOC(O)NCS, CH_3CN, rt, 4 h; ii. EDCl, HMDS rt, 48 h, 63% (d) i. 1M NaOH, MeOH, 65 °C; ii. $HSCH_2CH_2SH$, BF_3-OEt_2, DCM, rt, 48 h, 59%. (f) $CH(OEt)_3$, Ac_2O, 120 °C, 16 h, 69%. (g) P_2S_5, Pyridine, 65 °C. 2 h, quant. (h) i. NH_3, MeOH, 70 °C, 16 h; ii. $HSCH_2CH_2SH$, BF_3-OEt_2, DCM, rt, 72 h, 32%.

Reactions of **3** with both amide and ester-substituted *N*-tosyl derivatives **4a** and **4b** furnished the cyclized **5a** and **5b**, respectively, in good yields (Scheme 2).[12] To set the stereocenter at the anomeric carbon, reduction of **5a** and **5b** with triethylsilane and BF_3-OEt_2 was found to be the most effective, yielding key precursors **6a** and **6b,** respectively (Scheme 2).[13] Only one diastereomer was isolated and found to be the desired one (see below). With this key substrate in hand, the synthesis of the protected guanosine analog was accomplished using a mild 2-step, one pot reaction with an isothiocyanate precursor.[14] After cleavage of the carbamate, **7** was deprotected using 1,2-ethanedithiol and BF_3-OEt_2 to yield the final nucleoside **^{tz}G**.[15,16] Synthesis of the adenosine analog was accomplished via initial construction of the inosine analog **8** using triethyl orthoformate. This was subsequently converted to the thioamide by treatment with P_2S_5 in pyridine followed by methanolic ammonia to give the protected final product **9**, which was then subjected to the same deprotection conditions to afford **^{tz}A** (Scheme 2).

Crystal structure determination confirmed the proposed structure and anomeric configuration of the modified ribonucleosides (Figure 2 and Tables S1–S4). In the solid state, the purine analogs **^{tz}A** and **^{tz}G** showed an anti orientation at the glycosidic linkages, while the pyrimidines **^{tz}U** and **^{tz}C** were found to be in the syn orientation (Figure 2). Rewardingly, analysis of **^{tz}G**'s crystal packing pattern shows pairing through both the WC and Hoogsteen faces, which is identical to the pattern seen for guanosine in the solid state (Figure 3).[17] This intermolecular hydrogen bonding arrangement, illustrating the restoration of a "functional" Hoogsteen face and "N7", suggests that **^{tz}G** is likely to share **G**'s H-bonding and tautomeric preferences.

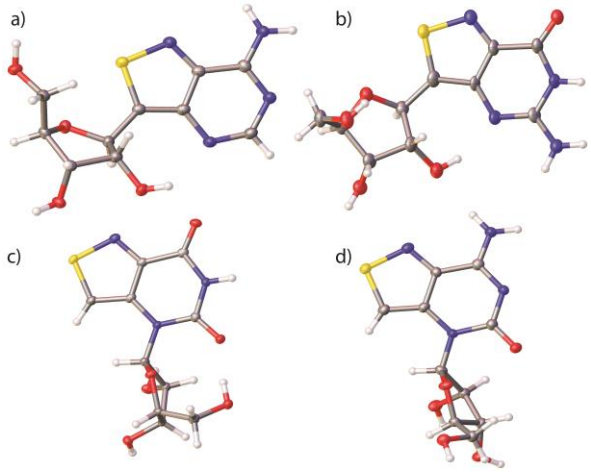


Figure 2. X-ray crystal structures of isothiazolo[4,3-*d*]pyrimidine analogs:(a) **^{tz}A** (b) **^{tz}G** (c) **^{tz}U** (d) **^{tz}C**.

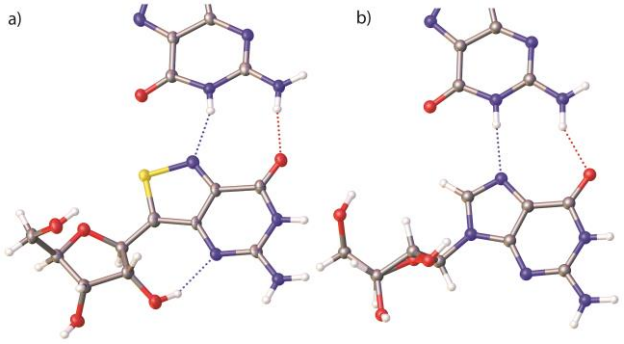


Figure 3. Comparison of intermolecular H-bonding seen in the crystal structures of (a) **^{tz}G** and (b) **G**.

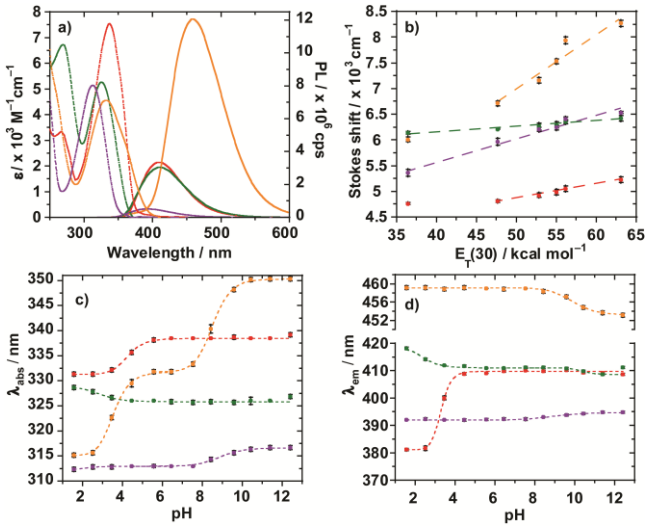


Figure 4. (a) Absorption (dashed lines) and emission (solid lines) spectra of **^{tz}A** (red), **^{tz}C** (green), **^{tz}G** (orange) and **^{tz}U** (purple) in water. (b) Stokes shift correlation versus solvent polarity ($E_T(30)$) of water/dioxane mixtures for **^{tz}A** (red), **^{tz}C** (green), **^{tz}G** (orange) and **^{tz}U** (purple). (c) Absorption maxima and (d) emission maxima variation versus pH for **^{tz}A** (red), **^{tz}C** (green), **^{tz}G** (orange) and **^{tz}U** (purple).

Table 1. Photophysical properties of isothiazolo[4,3-*d*]pyrimidine nucleoside analogs

	solvent	λ_{abs} (ε)[a]	λ_{em} (Φ)[a]	$\Phi\varepsilon$	Stokes shift[a]	polarity sensitivity[b]	pK_a[c] abs	pK_a[c] em
^{tz}A	water	338 (7.79)	410 (0.05)	413	5.23	27.7	4.25	3.29
	dioxane	342 (7.42)	409 (0.03)	193	4.76			
^{tz}C	water	325 (5.45)	411 (0.05)	289	6.42	10.5	2.94	2.46, 10.38
	dioxane	333 (5.03)	419 (0.04)	181	6.14			
^{tz}G	water	333 (4.87)	459 (0.25)	1203	8.27	102.0	3.55, 8.51	9.88
	dioxane	339 (4.65)	425 (0.17)	539	6.01			
^{tz}U	water	312 (5.17)	392 (0.01)	41	6.53	45.4	2.25, 8.88	8.94
	dioxane	314 (5.20)	377 (< 0.01)	21	5.36			
^{tz}I	water	316 (7.63)	377 (0.01)	46	5.13	12.7	9.26	7.83
	dioxane	315 (6.62)	372 (< 0.01)	26	4.79			

[a] λ_{abs}, ε, λ_{em} and Stokes shift are reported in nm, 10^3 M^{-1} cm^{-1}, nm and 10^3 cm^{-1} respectively. All photophysical values reflect the average of at least three independent measurements. [b] Sensitivity to solvent polarity reported in cm^{-1}/ (kcal mol^{-1}) is equal to the slope of the linear fit in Figure 4b. [c] pK_a values reflect the average over three independent measurements and are equal to the inflection point determined by the fitting curves in Figure 4c and 4d. See supporting table S5 for expanded data including experimental errors.

The fundamental spectroscopic properties, the sensitivity towards environmental polarity, and the spectroscopically-derived pK_a values of the modified nucleoside analogs are listed in Table 1. The ground-state absorption spectra in aqueous solution displayed bathochromic-shifted maxima compared to the corresponding native nucleosides ranging from 312 to 338 nm for **^{tz}U** to **^{tz}A**, respectively (Figure 4a). Visible emission maxima ranging from 392 nm (for **^{tz}U**) to 459 nm (for **^{tz}G**) were observed for all modified nucleosides upon excitation at their maxima. The emission quantum yield of the purine analogs, 0.25 for **^{tz}G** and 0.05 for **^{tz}A**, were higher than the pyrimidine analogs, with 0.01 for **^{tz}U** and 0.05 for **^{tz}C** in water.

The absorption spectra taken in dioxane showed batho and hypochromic shifts for **^{tz}A**, **^{tz}C** and **^{tz}G** in comparison to the aqueous solutions, while no significant variations were observed for **^{tz}U**. The emission intensity is sensibly lower in dioxane. The fluorescence maxima displayed a remarkable hypsochromic shift for **^{tz}G** and **^{tz}U**, a slight blue-shift for **^{tz}A,** and a bathochromic shift for **^{tz}C**. This suggests a charge-transfer character of their excited states (Figure S5). This is manifested for all nucleosides, albeit to different extents, in their responsiveness toward polarity changes, as seen by the linear correlations between the measured Stokes shifts and microscopic solvent polarity parameters (Figure 4b, Table 1).

All new nucleosides display sensitivity towards pH variations, thus facilitating the extraction of pK_a values (Figure 4c, 4d and Table 1). The deprotonation of **^{tz}U** between pH 8 and 11 is characterized by a red shift of its absorption maximum yielding a pK_a value of 8.9, which is comparable to the values reported for N3 deprotonation in uridine (pK_a 9.20–9.25).[18] **^{tz}C** is the least responsive to pH changes, displaying minor blue shifts in both the absorption and emission spectra (Figure 4c and 4d), and yielding two pK_a values (Table 1). **^{tz}A** was characterized by a bathochromic shift both in its absorption and emission maxima upon deprotonation of N1 (pK_a= 4.25 and 3.29, respectively) in close proximity to the reported values for adenosine (pK_a 3.6–4.2).[19] pH titration of **^{tz}G** showed two distinct red-shifted transitions of the absorption maximum and two different isosbestic points at 325 and 333 nm (Figure S2), assigned to deprotonation of N7 and N1 (pK_a = 3.55 and 8.51, respectively). These values correlate well with the reported values for guanosine (pK_a = 3.2–3.3 and 9.2–9.6, respectively).[20] Taken together, these observations not only illustrate the responsiveness of these nucleosides, but further indicate the ability of the isothiazole ring to mimic the basic imidazole moiety in the native purines.

To demonstrate the utility of the new analogs and the impact of restoring the basic N7 in evolving the thieno alphabet into the isothiazolo one, we selected to probe the enzymatic deamination of adenosine to inosine. This important metabolic transformation is catalyzed by adenosine deaminase (ADA).[21] We have previously reported the ability of ADA to recognize and deaminate **thA**, the thieno[3,4-*d*]-pyrimidine-based analog of the naturally occurring adenosine, to **thI** (Figure 5a), the corresponding inosine derivative. While of significance in and of itself, the enzymatic conversion of **thA** to **thI** was, however, approximately 20-times slower compared to that of adenosine.[4b] Since ADA has been crystallographically shown to form a H-bond to the N7 of its substrate,[21a,b] we hypothesized that restoration of this functionality in **^{tz}A**, the new adenosine surrogate, should facilitate its deamination compared to **thA**.

Steady state absorption and emission spectra, taken at defined time-intervals upon addition of ADA to **^{tz}A**, showed a fast and efficient conversion of **^{tz}A** to **^{tz}I** (Figure S6). Real-time continuous measurements, relying on the photophysical differences between **^{tz}A** and **^{tz}I**, **thA** and **thI** and **A** and **I**,[4b,22] show that while the deamination reaction of **thA** is indeed sluggish, ADA deaminates **^{tz}A** to **^{tz}I** at the same rate as it deaminates **A** to **I** (Figure 5b). The reaction half-life for **^{tz}A** deamination, calculated assuming a pseudo-first order reaction, was comparable to the one observed for adenosine, the native substrate ($t_{½}$ = 39 and 57 s, respectively) and substantially shorter than the one obtained for **thA** ($t_{½}$ = 818 s). This remarkable initial deamination rate of **^{tz}A** by ADA, substantiating our hypothesis, was confirmed by HPLC analyses (Figure 5c inset, Figure S8). Taken together, these observations highlight the improved functionality of the isothiazolopyrimidine over the thienopyrimidine core.

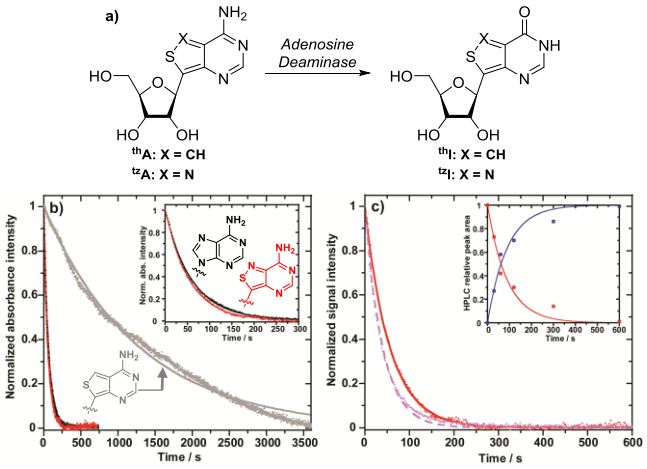


Figure 5. (a) Deamination of th**A** and ^{tz}**A** (b) Enzymatic deamination of **A** to **I** (black), ^{tz}**A** to ^{tz}**I** (red) and th**A** to th**I** (grey) with ADA monitored by real-time absorption at 260, 340 and 340 nm respectively. Inset: zoomed in region between 0 and 300 seconds. (c) Enzymatic deamination of ^{tz}**A** to ^{tz}**I** by ADA monitored by real-time absorption (red) at 340 nm and real time emission (light purple) at 410 nm (excitation at 322 nm). Inset: HPLC relative peak area variation at different time-points for ^{tz}**A** (red) and ^{tz}**I** (blue) monitored at 340 nm

In summary, we introduce a second-generation family of emissive nucleoside analogs, based on an isothiazolopyrimidine scaffold. This atomic mutation results in higher isomorphicity and significantly improved functionality when compared to the thienopyrimidine-based RNA alphabet. The presence of the isothiazole core with its nitrogen in the equivalent position to N7 of the native purines, restores the Hoogsteen face, as well as the basicity and native H bonding ability of these surrogates, as illustrated structurally (for ^{tz}**G**) and biochemically (for ^{tz}**A** deamination by ADA). These observations, indicating that the newly introduced purine surrogates display improved structural and functional characteristics, are of significance, since very few emissive, isomorphic and non-perturbing purine analogs have so far been made and biophysically exploited.

ASSOCIATED CONTENT

Supporting Information

Synthetic details, X-ray crystallographic data (CIF), photophysical data, enzymatic protocols and HPLC traces. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

ytor@ucsd.edu

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

We thank the National Institutes of Health for generous support (GM 069773), the Chemistry & Biochemistry MS Facility, and the UCSD X-ray crystallography Facility.

REFERENCES

1. (a) Sinkeldam, R. W.; Greco, N. J.; Tor, Y. *Chem. Rev.* **2010**, *110*, 2579–2619; (b) Wilhelmsson, M. *Q. Rev. Biophys.* **2010**, *43*, 159–183; (c) Hawkins, M. E. *Cell Biochem. Biophys.* **2001**, *34*, 257–281; (c) Wilson, J. N.; Kool, E. T. *Org. Biomol. Chem.* **2006**, *4*, 4265–4274; (d) Okamoto, A.; Saito, Y.; Saito, I. *J. Photochem. Photobiol. C* **2005**, *6*, 108–122; (e) Dodd, D. W.; Hudson, R. H. E. *Mini-Rev. Org. Chem.* **2009**, *6*, 378–391; (f) Kimoto, M.; Cox, R. S., III; Hirao, I. *Expert Rev. Mol. Diagn.* **2011**, *11*, 321–331; (g) Rist, M. J.; Marino, J. P. *Curr. Org. Chem.* **2002**, *6*, 775–793; (h) Wierzchowski, J. ; Antosiewiczb, J. M.; Shugar, D. *Mol. BioSyst.*, **2014**, *10*, 2756–2774.

2. (a) Gaied, N. B.; Glasser, N.; Ramalanjaona, N.; Beltz, H.; Wolff, P.; Marquet, R.; Burger, A.; Mely, Y. *Nucleic Acids Res.* **2005**, *33*, 1031–1039; (b) Nadler, A.; Strohmeier, J.; Diederichsen, U. *Angew. Chem. Int. Ed.* **2011**, *50*, 5392–5396; (c) Dierckx, A.; Miannay, F.-A.; Gaied, N. B.; Preus, S.; Björck, M.; Brown, T.; Wilhelmsson, L. M. *Chem. Eur. J.*, **2012**, *18*, 5987–5997. (d) Dumat, B.; Bood, M.; Wranne, M. S.; Lawson, C. P.; Foller Larsen, A.; Preus, S.; Streling, J.; Gradén, H.; Wellner, E.; Grøtli, M.; Wilhelmsson, L. M. *Chem. Eur. J.*, **2015**, *21*, 4039–4048.

3. Shin, D.; Sinkeldam, R. W.; Tor, Y. *J. Am. Chem. Soc.* **2011**, *133*, 14912–14915.

4. (a) Liu, W.; Shin, D.; Tor, Y.; Cooperman, B. S. *ACS Chem. Biol.* **2013,** *8*, 2017–2023; (b) Sinkeldam, R. W.; McCoy, L. S.; Shin, D.; Tor, Y. *Angew. Chem., Int. Ed.* **2013**, *52*, 14026–14030; (c) McCoy, L. S.; Shin, D.; Tor, Y. *J. Am. Chem. Soc.* **2014**, *136*, 15176–15184; (d) Otomo, H.; Park, S.; Yamamoto, S.; Sugiyama, H. *RSC Adv.* **2014**, *4*, 31341–31344; (e) Park, S.; Otomo, H.; Zheng, L.; Sugiyama, H. *Chem. Commun.* **2014**, *50*, 1573–1575; (f) Sholokh, M.; Sharma, R.; Shin, D.; Das, R.; Zaporozhets, O. A.; Tor, Y.; Mely, Y. *J. Am. Chem. Soc.* **2015**, *137*, 3185–3188.

5. Mizrahi, R. A.; Shin, D.; Sinkeldam, R. W.; Phelps, K. J.; Fin, A.; Tantillo, D. J.; Tor, Y.; Beal, P. A. *Angew. Chem., Int. Ed.* **2015**, *54*, 8713–8716.

6. The synthesis of the pyrimidine analogs is described in the supporting information (Scheme S5).

7. (a) Coté, G. L.; Flitsch, S.; Ito, Y.; Kondo, H.; Nishimura, S.; Yu, B.; Fraser-Reid, B. O.; Tatsuta, K.; Thiem, J. *Glycoscience: Chemistry and Chemical Biology*. 2 ed.; Springer-Verlag Berlin Heidelberg: 2008; (b) Iddon, B. *Heterocycles* **1995**, *41*, 533–593; (c) Stambasky, J.; Hocek, M.; Kocovsky, P. *Chem. Rev.* **2009**, *109*, 6729–6764.

8. Attempts to glycosylate the nucleobase using palladium coupling chemistry, grignard formation, alkyllithium chemistry, and standard Vorgbruggen glycosylation conditions were all unsucessful.

9. Wamhoff, H.; Berressem, R.; Nieger, M. *J. Org. Chem.* **1993**, *58*, 5181–5185.

10. See Supplementary Information (Scheme S1).

11. Bessieres, B.; Morin, C. *Synlett.* **2000**, 1691–1693.

12. Gewald, K.; Bellmann, P. *Liebigs Ann. Chem.* **1979**, 1534–1546.

13. van Rijssel, E. R.; van Delft, P.; Lodder, G.; Overkleeft, H. S.; van der Marel, G. A.; Filippov, D. V.; Codee, J. D. *Angew. Chem. Int. Ed.* **2014**, *53*, 10381–10385.

14. Lecoutey, C.; Fossey, C.; Rault, S.; Fabis, F. *Eur. J. Org. Chem* **2011**, 2785–2788.

15. Fuji, K.; Ichikawa, K.; Node, M.; Fujita, E. *J. Org. Chem.* **1979**, *44*, 1661–1664.

16. Seley, K. L.; Zhang, L.; Hagos, A.; Quirk, S. *J. Org. Chem.* **2002**, *67*, 3365–3373.

17. (a) Additionally, an overlay of the purine analogs and natural nucleosides may be found on Figure S1; (b) Marsh, R. E.; Bugg, C. E.; Thewalt, U. *Acta Cryst.* **1970**, *B26*, 1089–1101.

18. (a) Simpson, R. B. *J. Am. Chem. Soc.* **1964**, *86*, 2059–2065; (b) Luyten, I.; Pankiewicz, K. W.; Watanabe, K. A.; Chattopadhyaya, J. *J. Org. Chem.* **1998**, *63*, 1033–1040.

19. (a) Christensen, J. J.; Rytting, J. H.; Izatt, R. M. *Biochemistry* **1970**, *9*, 4907–4913; (b) Kapinos, L. E.; Operschall, B. P.; Larsen, E.; Sigel, H. *Chem. Eur. J.* **2011**, *17*, 8156–8164.

20. (a) Bundari, S. *The Merck Index*, 12th ed.; Merck and Co., Inc.: Whitehouse Station, NJ, 1996; (b) Sigel, H.; Massoud, S. S.; Corfù, N. A. *J. Am. Chem. Soc.* ***1994***, *116*, 2958–2971; (c) Thapa, B.; Schlegel, H. B. *J. Phys. Chem. A* **2015**, *119*, 5134-5144; (d) Kampf, G; Kapinos, L. E.; Griesser, R.; Lippert, B.; Sigel, H. *J. Chem. Soc., Perkin Trans. 2* **2002**, 1320–1327.

21. (a) Kinoshita, T.; Nishio, N.; Nakanishi, I.; Sato, A.; Fujii, T. *Acta Crystallogr. Sect. D* **2003**, *59*, 299–303; (b) Wilson, D. K.; Rudolph, F. B.; Quiocho, F. A. *Science* **1991**, *252*, 1278–1284; (c) Kinoshita, T.; Nakanishi, I.; Terasaka, T.; Kuno, M.; Seki, N.; Warizaya, M.; Matsumura, H.; Inoue, T.; Takano, K.; Adachi, H.; Mori, Y.; Fujii, T. *Biochemistry* **2005**, *44*, 10562–10569.

22. See Supplementary Information (Figure S6).

Insert Table of Contents artwork here