JOC*Note*

Expanding the Genetic Alphabet: Non-Epimerizing Nucleoside with the *py*DDA Hydrogen-Bonding Pattern

Daniel Hutter[†] and Steven A. Benner*[,†,‡]

Department of Chemistry and Department of Anatomy and Cell Biology, University of Florida, Gainesville, Florida 32611-7200

benner@chem.ufl.edu

Received June 25, 2003

Abstract: 6-Amino-3-(2′-deoxy-β-D-ribofuranosyl)-5-nitro-1*H*-pyridin-2-one (**4**), a *C*-glycoside exhibiting the nonstandard *py*DDA hydrogen-bonding pattern, was synthesized via Heck coupling. The nitro group greatly enhances the stability of the nucleoside toward acid-catalyzed epimerization without leading to significant deprotonation of the heterocycle at physiological pH. These results make nucleoside **4** a promising candidate for an expanded genetic alphabet.

The Watson−Crick nucleobase pair in DNA follows two rules of complementarity: size complementarity (large purines pair with small pyrimidines) and hydrogen-bonding complementarity (hydrogen bond donors from one base complement hydrogen bond acceptors from the other base).[1] A decade ago, we noted that 12 nucleobases forming six base pairs joined by mutually exclusive hydrogen-bonding patterns are possible within the Watson−Crick restraints and that these might be functionalized to enable a single biopolymer capable of both genetics and catalysis.[2] Expanded genetic alphabets have now been further explored in a variety of laboratories, and the possibility of a fully artificial system has been advanced.[3−7]

Various analyses of the interaction between polymerases and their substrates suggest that the polymerase seeks two unshared pairs of electrons in the minor groove, at position 3 of the purine (or analogue) and at position 2 of the pyrimidine (or analogue).[8] In addition, the base pairs that form three hydrogen bonds are expected to contribute more to duplex stability than pairs joined by just two hydrogen bonds.

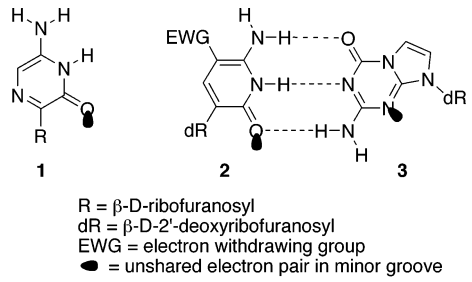


FIGURE 1. Nonstandard base pair with the *py*DDA·*pu*AAD[9] hydrogen-bonding pattern. Pyrimidine analogue **2** was believed to suffer epimerization less likely than known analogue **1**.[3f]

These conditions are fulfilled for the *py*DDA· *pu*AAD[9] base pair (Figure 1). While the purine base and the corresponding nucleoside **3** have been synthesized to study their antiviral activity as well as their capability to form purine·purine base pairs,[10] an unambiguous *py*DDA hydrogen-bonding pattern has so far only been reported for the RNA-pyrazine analogue **1**,[3f] while pseudoisocytidine is reported to exhibit this hydrogen-bonding pattern under certain conditons only.[5h] In addition to being a component of an expanded genetic information system, these analogues also exhibit at neutral pH the hydrogen-bonding pattern of protonated cytosine.

The *py*DDA hydrogen-bonding pattern can be achieved only with a *C*-glycoside, where the heterocycle is coupled to the sugar moiety by a C−C bond rather than a C−N bond, which is found in the standard nucleosides. A special feature of *C*-nucleosides is their susceptibility to epimerization if an electron-donating substituent is present in a suitable position on the heterocycle (Scheme 1). This phenomenon was described for pseudouridine

* To whom correspondence should be addressed. Tel: +352-392-7773. Fax: +352-392-7918.

† Department of Chemistry.

‡ Department of Anatomy and Cell Biology.

(1) (a) Watson, J. D.; Crick, F. H. C. *Nature* (London) **1953**, *171*, 737. (b) Watson, J. D.; Crick, F. H. C. *Nature* (London) **1953**, *171*, 964.

(2) (a) Switzer, C. Y.; Moroney, S. E.; Benner, S. A. *J. Am. Chem. Soc.* **1989**, *111*, 8322. (b) Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* (London) **1990**, *343*, 33.

(3) (a) Piccirilli, J. A.; Krauch, T.; MacPherson, L. J.; Benner, S. A. *Helv. Chim. Acta* **1991**, *74*, 397. (b) Voegel, J. J.; Altorfer, M. M.; Benner, S. A. *Helv. Chim. Acta* **1993**, *76*, 2061. (c) Voegel, J. J.; von Krosigk, U.; Benner, S. A. *J. Org. Chem.* **1993**, *58*, 7542. (d) Heeb, N. V.; Benner, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 6929. (e) Voegel, J. J.; Benner, S. A. *J. Am. Chem. Soc.* **1994**, *116*, 6929. (f) von Krosigk, U.; Benner, S. A. *J. Am. Chem. Soc.* **1995**, *117*, 5361. (g) Voegel, J. J.; Benner, S. A. *Helv. Chim. Acta* **1996**, *79*, 1863. (h) Kodra, J. T.; Benner, S. A. *Synlett* **1997**, 939. (i) Jurczyk, S. C.; Battersby, T. R.; Kodra. J. T.; Park, J. H.; Benner, S. A. *Helv. Chim. Acta* **1999**, *82*, 1005. (j) Jurczyk, S. C.; Horlacher, J.; Devine, K. G.; Benner, S. A.; Battersby, T. R. *Helv. Chim. Acta* **2000**, *83*, 1517. (k) Rao, P.; Benner, S. A. *J. Org. Chem.* **2001**, *66*, 5012.

(4) (a) Service, R. F. *Science* **2000**, *289*, 232. (b) Wang, L.; Schultz, P. G. *Chem. Commun.* **2002**, 1.

(5) Modified H-bonding patterns: (a) Seela, F.; Wei, C. F. *Chem. Commun.* **1997**, 1869. (b) Seela, F.; He, Y. *J. Org. Chem.* **2003**, *68*, 367. (c) Hirao, I.; Ohtsuki, T.; Mitsui, T. Yokoyama, S. *J. Am. Chem. Soc.* **2000**, *122*, 6118. (d) Hirao, I.; Ohtsuki, T.; Fujiwara, T.; Mitsui, T.; Yokogawa, T.; Okuni, T.; Nakayama, H.; Takio, K.; Yabuki, T.; Kigawa, T.; Kodama, K.; Yokogawa, T.; Nishikawa, K.; Yokoyama, S. *Nature Biotechnol.* **2002**, *20*, 177. (e) Searls, T.; McLaughlin, L. W. *Tetrahedron* **1999**, *55*, 11985. (f) Chen, D. L.; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468. (g) Roberts, C.; Banduru, R.; Switzer, C. *J. Am. Chem. Soc.* **1997**, *119*, 4640. (h) Oyelere, A. K.; Strobel, S. A. *J. Am. Chem. Soc.* **2000**, *122*, 10259.

(6) Hydrophobic base pairs: (a) Kool, E. T. *Curr. Opin. Chem. Biol.* **2000**, *4*, 602. (b) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990. (c) Kool, E. T. *Acc. Chem. Res.* **2002**, *35*, 936. (d) Ogawa, A. K.; Wu, Y.; McMinn, D. L.; Liu, J.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2000**, *122*, 3274. (e) Berger, M.; Luzzi, S. D.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2002**, *124*, 1222.

(7) Metal chelation: (a) Weizman, H.; Tor, Y. *J. Am. Chem. Soc.* **2001**, *123*, 3375. (b) Zimmermann, N.; Meggers, E.; Schultz, P. G. *J. Am. Chem. Soc.* **2002**, *124*, 13684. (c) Tanaka, K.; Tengeiji, A.; Kato, T.; Toyama, N.; Shiro, M.; Shionoya, M. *J. Am. Chem. Soc.* **2002**, *124*, 12494.

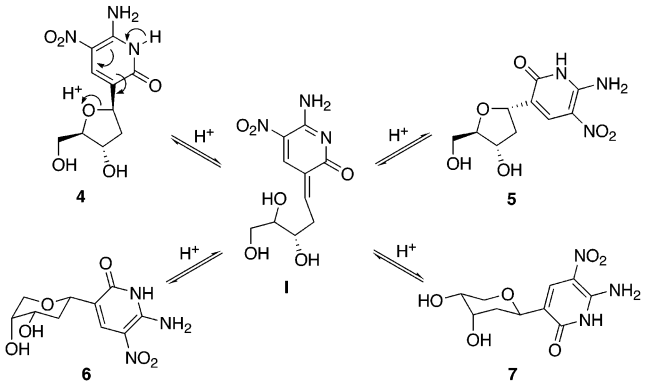
(8) Steitz, T. In *Biological Organization: Macromolecular Interactions at High Resolution*; Burnett, R. M., Vogel, H. J., Eds.; Academic Press: New York, 1987; pp 45−55.

(9) *py*DDA: Pyrimidine analogue carrying a H-bond donor−donor−acceptor motif proceeding from the major to the minor groove. *pu*AAD: purine analogue carrying an H-bond acceptor−acceptor−donor motif.

(10) (a) Kim, S. H.; Bartholomew, D. G.; Allen, L. B.; Robins, R. K.; Revankar, G. R.; Dea. P. *J. Med. Chem.* **1978**, *21*, 883. (b) Rosemeyer, H.; Seela, F. *J. Org. Chem.* **1987**, *52*, 5136. (c) Seela, F.; Amberg, S.; Melenewski, A.; Rosemeyer, H. *Helv. Chim. Acta* **2001**, *84*, 1996.

10.1021/jo034900k CCC: $25.00
Published on Web 11/13/2003

SCHEME 1

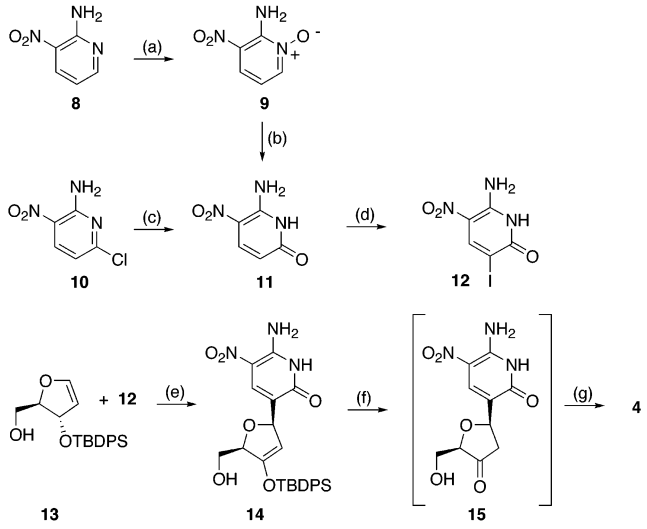


several decades ago[11] and for other *C*-nucleosides since.[12] In our laboratory, we have shown that *C*-nucleosides with 5′- and 3′-hydroxyls epimerize, usually in either acidic or basic medium, to a mixture of four isomers with the pyranose forms being the most abundant.[3e−g] With the hydroxyls protected or substituted, the epimerization limits itself to the α/*β*-forms of the furanoses. This epimerization is problematic, especially during the acidic deprotection step of the 5′-position during solid-phase oligonucleotide synthesis. But the nucleoside/oligonucleotide will also epimerize upon standing at pH 7 for a prolonged period of time. Considering the mechanism in Scheme 1, to prevent epimerization, a heterocycle was sought that prevents the electron "push" from the bond to this acidic proton. This could be achieved by putting an electron withdrawing group at the 5-position[13] (**2**) (Figure 1).

The overall strategy for our synthesis of the nucleoside (Scheme 2) focused on the Heck coupling of the iodinated heterocycle with the correct hydrogen bonding pattern (**12**) to the known[14] glycal **13**.

The initial attempt to synthesize the pyridinone **11** from 2-amino-3-nitropyridine (**8**) via Katada rearrangement of *N*-oxide **9** was unsatisfactory. The *N*-oxide was obtained easily and in good yield by oxidation with *m*-CPBA.[15] The subsequent rearrangement gave mostly tar, however, and only about 10−15% of the pyridinone, despite several variations in reaction conditions and workup.[16] We therefore switched to the less elegant route starting with commercial 2,6-dichloro-3-nitropyridine. Selective aminolysis led to the known 6-chloro derivative **10**,[17] which was hydrolyzed to **11** with aqueous sodium hydroxide. Iodination at the 5-position with *N*-iodosuccinimide in DMF gave **12**.

SCHEME 2[a]



[a] Key: (a) *m*-CPBA, acetone, rt; (b) (1) KOAc, Ac_2O, Δ, (2) MeOH−NH_3, rt; (c) NaOH, EtOH/H_2O, Δ; (d) *N*-iodosuccinimide, DMF, rt; (e) cat. $Pd(OAc)_2$·$2Ph_3As$, Et_3N, DMF, 60 °C; (f) TBAF, THF, 0 °C; (g) $NaBH(OAc)_3$, MeCN/MeCOOH, 0 °C.

Heck couplings of iodinated heterocycles to the glycal **13** and subsequent deprotection and reduction have been reported previously by several groups.[18−21] The bulky TBDPS group at the 3′-position is assumed to direct addition of the heterocycle to the *β*-face, and the free 5′-hydroxyl leads to stereospecific reduction of the ketone by complexation with the borate. According to these literature procedures, palladium acetate with triphenylarsine was used as the catalyst system for the Heck coupling, with anhydrous DMF as the solvent. Triethylamine was substituted for tributylamine as the base, since it was easier to remove during purification, and 1.2 equiv of glycal proved to be sufficient. After several days at 60 °C, *β*-nucleoside **14** was obtained in good yield. Quick removal of the 5′-protective group with TBAF at 0 °C gave ketone **15** without any significant epimerization. The crude ketone was reduced with $NaBH(OAc)_3$ to the final nucleoside **4**, part of which was purified by reversed-phase HPLC for the physicochemical studies.

The pK_a of **4** was measured by titrating an aqueous solution of the nucleoside with dilute aqueous solutions of HCl and NaOH and recording a UV spectrum at

(11) (a) Cohn, W. E. *J. Biol. Chem.* **1960**, *235*, 1488. (b) Chambers, R. W.; Kurkov, V.; Shapiro, R. *Biochemistry* **1963**, *2*, 1192.

(12) (a) Ohrui, H.; Jones, G. H.; Moffatt, J. G.; Maddox, M. L.; Christensen, A. T.; Byram, S. K. *J. Am. Chem. Soc.* **1975**, *97*, 4602. (b) De Bernardo, S.; Weigele, M. *J. Org. Chem.* **1976**, *41*, 287. (c) Pankiewicz, K.; Matsuda, A.; Watanabe, K. A. *J. Org. Chem.* **1982**, *47*, 485.

(13) The different numbering systems of different heterocycles create confusion when attempting to compare substituents placed upon them. For this discussion, we use the pyrimidine numbering system by analogy to the pyridine rings that we describe.

(14) (a) Ireland, R. E.; Thaisrivongs, S.; Vanier, N.; and Wilcox, C. S. *J. Org. Chem.* **1980**, *45*, 48. (b) Larsen, E.; Jorgensen, P. T.; Sofan, M. A.; Pedersen, E. B. *Synthesis*, **1994**, 1037. (c) Walker, J. A., II; Chen, J. J.; Wise, D. S.; Townsend, L. B. *J. Org. Chem.* **1996**, *61*, 2219. (d) Cameron, M. A.; Cush, S. B.; Hammer, R. P. *J. Org. Chem.* **1997**, *62*, 9065.

(15) Deady, L. W. *Synth. Commun.* **1977**, 509.

(16) (a) Daeniker, H. U.; Druey, J. *Helv. Chim. Acta* **1958**, *41*, 2148. (b) Taylor, E. C.; Driscoll, J. S. *J. Org. Chem.* **1960**, *25*, 1716. (c) Markgraf, J. H.; Brown, Jr., H. B.; Mohr, S. C.; Peterson, R. G. *J. Am. Chem. Soc.* **1962**, *85*, 958. (d) McKillop, A.; Bhagrath, M. K. *Heterocycles* **1985**, *23*, 1697. (e) Sato, N.; Miwa, N.; Suzuki, H.; Sakakibara, T. *J. Heterocycl. Chem.* **1994**, *31*, 1229.

(17) Radl, S.; Hradil, P. *Collect. Czech. Chem. Commun.* **1991**, *56*, 2420.

(18) (a) Farr, R. N.; Outten, R. A.; Cheng, J. C.-Y.; Daves, G. D., Jr. *Organometallics* **1990**, *9*, 3151. (b) Zhang, H.-C.; Daves, G. D., Jr. *J. Org. Chem.* **1992**, *57*, 4690. (c) Zhang, H.-C.; Daves, G. D., Jr. *Organometallics* **1993**, *12*, 1499.

(19) (a) Hsieh, H.-P.; McLaughlin, L. W. *J. Org. Chem.* **1995**, *60*, 5356. (b) Chen, D. L.; McLaughlin, L. W. *J. Org. Chem.* **2000**, *65*, 7468. (c) Searls, T.; Chen, D. L.; Lan, T.; McLaughlin, L. W. *Biochemistry* **2000**, *39*, 4375. (d) Lan, T.; McLaughlin, L. W. *Bioorg. Chem.* **2001**, *29*, 198.

(20) Coleman, R. S.; Madaras, M. L. *J. Org. Chem.* **1998**, *63*, 5700.

(21) With 5′-OH protected: (a) Chen, J. J.; Walker, J. A., II; Liu, W.; Wise, D. S.; Townsend, L. B. *Tetrahedron Lett.* **1995**, *36*, 8363. (b) Walker, J. A., II; Chen, J. J.; Hinkley, J. M.; Wise, D. S.; Townsend, L. B. *Nucleosides Nucleotides* **1997**, *16*, 1999.

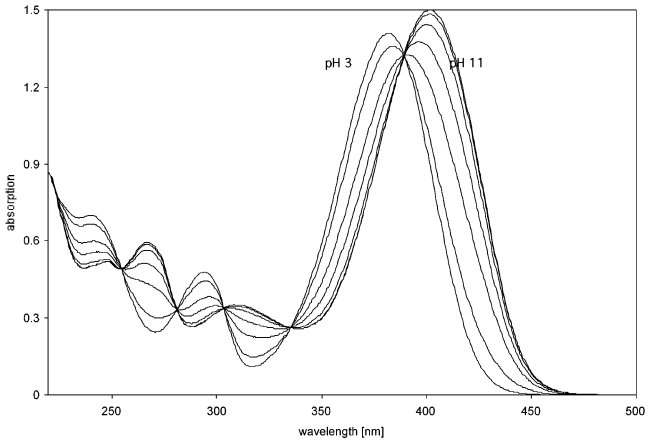


FIGURE 2. Tritration of nucleoside analogue **4** (UV absorbance versus wavelength) showing a pK_a of 7.8 ± 0.1.

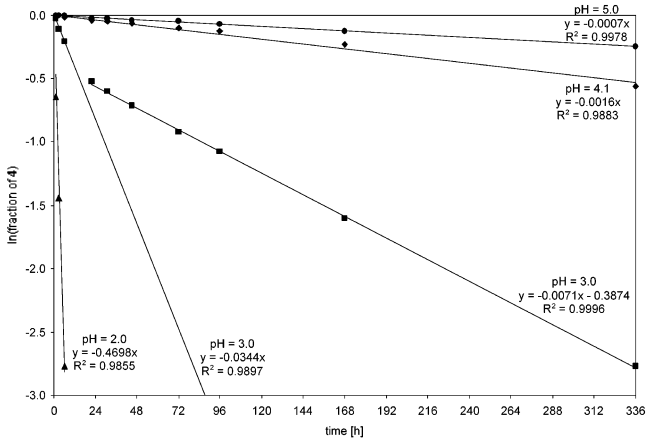


FIGURE 3. Acid-catalyzed epimerization of nucleoside **4**. Natural log of the fraction of **4** remaining after time *t* versus *t* at pH 2.0, 3.0, 4.1, and 5.0 (37 °C, aqueous).

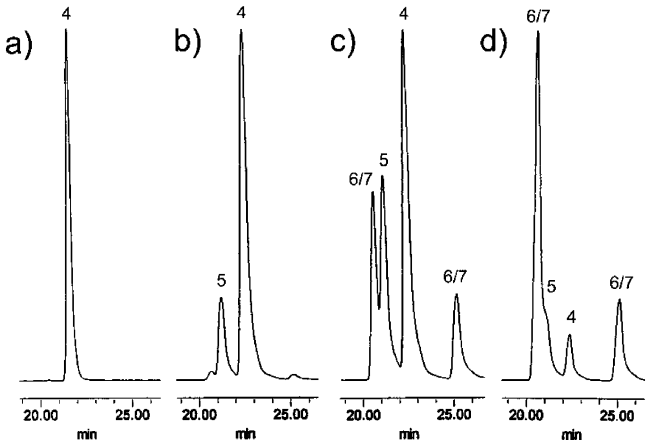


FIGURE 4. HPLC traces at different stages of the epimerization process at pH 3: (a) $t = 0$ h (pure **4**); (b) $t = 6$ h (preequilibrium including **4** and **5**); (c) $t = 72$ h (appearance of **6** and **7**); (d) $t = 336$ h (global equilibrium).

different pH values. The resulting UV spectra, displaying several isosbestic points, are collected in Figure 2. Analysis of the curves gave a pK_a value of 7.8 ± 0.1, which is ca. 2 pK_a units lower than those for the natural nucleosides. Nevertheless, nucleoside **4** should be largely (>80%) protonated at physiological pH and still more protonated when embedded into a polyanionic DNA or RNA chain. Such embedding generally elevates the pK_a of the nucleotide by about 0.5 units.[22]

To measure the rate of epimerization of the nucleoside at neutral pH, a sample was dissolved in D_2O and the ^{1}H NMR was measured after several time intervals. The spectrum did not change even after 1 week. In particular, the NMR signal arising from H-6 of the heterocycle remained unchanged. This suggested that the compound was stable toward epimerization at neutral pH and allowed for analysis by HPLC to measure the epimerization rate at other pH values.

For those measurements, the pH of an aqueous solution of the nucleoside (ca. 5 mM) was adjusted with dilute aqueous NaOH or HCl to the desired values. These solutions were incubated at constant temperature. After given time intervals, aliquots were removed, neutralized and subjected to analysis by reversed-phase HPLC.

At high pH (pH 11.0), no epimerization was detected at 23 °C even after 14 days. The HPLC trace showed still only one peak, corresponding to the starting material. This is consistent with the specific acid catalysis mechanism proposed by Voegel and Benner[3g] for their *C*-glycoside and suggests that *C*-glycosides employ a specific acid mechanism for epimerization somewhat more generally.

At low pH (2.0−5.0) and 37 °C (used to speed the reaction), HPLC analysis showed the formation of four different compounds, as expected from the mechanism shown in Scheme 1. The rate of epimerization was strongly dependent upon pH (Figure 3).

Plots of the log of the progress of the reaction versus time were curved, however. We interpreted this as evidence for a two-step process that forms **I** from **4** prior to the formation of **5**−**7**. While several alternatives are possible, we prefer a model that involves the achievement of a preequilibrium including **4**, **I**, and **5**, prior to the achievement of a global equilibrium that includes **6** and **7**.

This is because no evidence was obtained for the presence of substantial amounts of **I** at any point in the reaction cycle (see also Figure 2). This requires that the rate for conversion of **I** to **4** must be smaller than the rate for conversion of **4** to **I** under these conditions. The curvature in the log progress plots can thus not be explained by a preequilibrium between **I** and **4** alone but rather suggests the achievement of a preequilibrium including **4**, **I** and **5**.

The notion that the formation of a five membered furanose ring is faster than the formation of a pyranose ring, even though the pyranose product is more stable, is not inconsistent with general concepts of reactivity in similar systems.[23]

Further evidence for this two step model can be found by comparing the HPLC traces at different stages of the epimerization process, as shown for pH 3 in Figure 4.

(22) Saenger, W. *Principles of Nucleic Acid Structure*; Springer: New York, 1984; p 107.

(23) Baldwin, J. E.; Thomas, R. C.; Kruse, L. I.; Silberman, L. *J. Org. Chem.* **1977**, *42*, 3846.

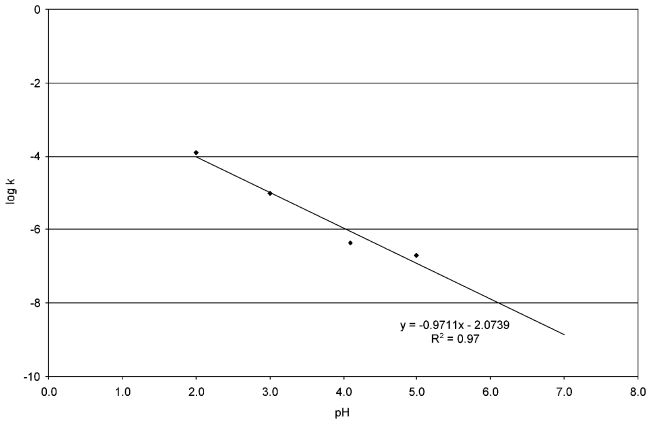


FIGURE 5. Replot of log $k_{apparent}$ versus pH, showing a first-order dependence of the epimerization reaction with hydrogen activity.

The HPLC traces at other pH values showed the same pattern.

The initial rate of epimerization displays pseudo first-order rate constants $k_{pH2} \approx 10^{-4}$ s^{-1} at pH 2.0 and $k_{pH5} \approx 2 \times 10^{-7}$ s^{-1} at pH 5.0 (Figure 3). A replot of log k versus pH is linear within experimental error with a slope of −0.971 (Figure 5), suggesting that the rate process is first order in hydrogen ion concentration consistent with specific acid catalysis.

For the design of nonepimerizable nucleosides belonging to the *C*-glycoside class, a tradeoff evidently must be made between having a heterocycle that remains protonated at physiological pH (that is, it has a high pK_a) and the stability of the nucleoside analogue with respect to epimerization. A stronger electron withdrawing substituent at C-5 evidently lowers the rate of epimerization. But it evidently lowers the pK_a as well. The nucleoside analog **4** described here seems to strike that balance quite well.

Acknowledgment. Support of the National Institutes of Health (GM-54048) is gratefully acknowledged.

Supporting Information Available: Experimental procedures, characterization data, and ^{1}H and ^{13}C NMR spectra for all new compounds. Experimental details for the measurement of pK_a and epimerization rate of **4**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO034900K