

Nonpolar Nucleoside Mimics as Active Substrates for Human Thymidine Kinases

Sarah K. Jarchow-Choy,[†] Elena Sjuvarsson,[‡] Herman O. Sintim,[†]
Staffan Eriksson,^{*,‡} and Eric T. Kool^{*,†}

Department of Chemistry, Stanford University, Stanford, California 94305, and Department of Anatomy, Physiology, and Biochemistry, Swedish University of Agricultural Sciences, Uppsala, Sweden

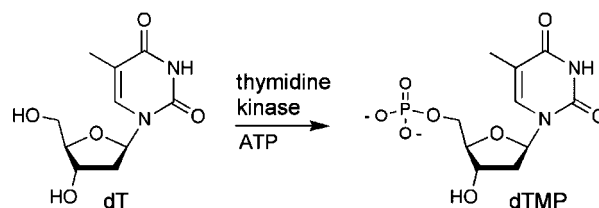
Received October 20, 2008; E-mail: staffan.eriksson@afb.slu.se; kool@stanford.edu

Abstract: We describe the use of nonpolar nucleoside analogues of systematically varied size and shape to probe the mechanisms by which the two human thymidine kinases (TK1 and TK2) recognize and phosphorylate their substrate, thymidine. Comparison of polar thymidine with a nonpolar isostere, 2,4-difluorotoluene deoxyriboside, as substrates for the two enzymes establishes that TK1 requires electrostatic complementarity to recognize the thymine base with high efficiency. Conversely, TK2 does not and phosphorylates the hydrophobic shape mimic with efficiency nearly the same as the natural substrate. To test the response to nucleobase size, thymidine-like analogues were systematically varied by replacing the 2,4 substituents on toluene with hydrogen and the halogen series (H, F, Cl, Br, I). Both enzymes showed a distinct preference for substrates having the natural size. To examine the shape preference, we prepared four mono- and difluorotoluene deoxyribosides with varying positions of substitutions. While TK1 did not accept these nonpolar analogues as substrates, TK2 did show varying levels of phosphorylation of the shape-varied set. This latter enzyme preferred toluene nucleoside analogues having steric projections at the 2 and 4 positions, as is found in thymine, and strongly disfavored substitution at the 3-position. Steady-state kinetics measurements showed that the 4-fluoro compound (**7**) had an apparent V_{max}/K_m value within 14-fold of the natural substrate, and the 2,4-difluoro compound (**1**), which is the closest isostere of thymidine, had a value within 2.5-fold. The results establish that nucleoside recognition mechanisms for the two classes of enzymes are very different. On the basis of these data, nonpolar nucleosides are likely to be active in the nucleotide salvage pathway in human cells, suggesting new designs for future bioactive molecules.

Introduction

As nucleosides enter the cell from the outside, they are recognized and phosphorylated by enzymes that are involved in the cellular nucleotide salvage pathway. Initial monophosphorylation by nucleoside kinase enzymes (Scheme 1) is an important rate-limiting step leading to their biological activity.¹ The monophosphates are further phosphorylated by the actions of a lower-selectivity nucleotide kinase followed by nucleotide diphosphate kinase, ultimately entering the nucleoside triphosphate pool from which DNA polymerases can select during replication and repair of the genome. There are two human cellular thymidine kinases: TK1 (the cytosolic kinase) and TK2 (the mitochondrial kinase). TK1 expression is upregulated in rapidly dividing cells, while TK2 is constitutively expressed in the mitochondria, where metabolism is always ongoing. The known substrate specificity of TK1 is restricted to thymidine and deoxyuridine, while that of TK2 also includes deoxycytidine. Although an X-ray crystal structure of human TK1 was

Scheme 1. Reaction Catalyzed by Thymidine Kinase



published recently, thymidine triphosphate was bound as a feedback inhibitor rather than as a phosphoryl accepting substrate.² No structure of TK2 is yet available. Thus, details of possible differential nucleoside recognition mechanisms by these enzymes are lacking.

Understanding the specificity of kinases is important for several reasons. First, these enzymes help achieve correct steady-state concentrations and balances of nucleotides for DNA replication; thus the enzymes are central to proper cellular function. Second, numerous clinically important nucleoside drugs are recognized and phosphorylated by the nucleoside kinases.³ For example, AZT and stavudine (d4T) are important antiviral compounds used in the treatment of AIDS, and

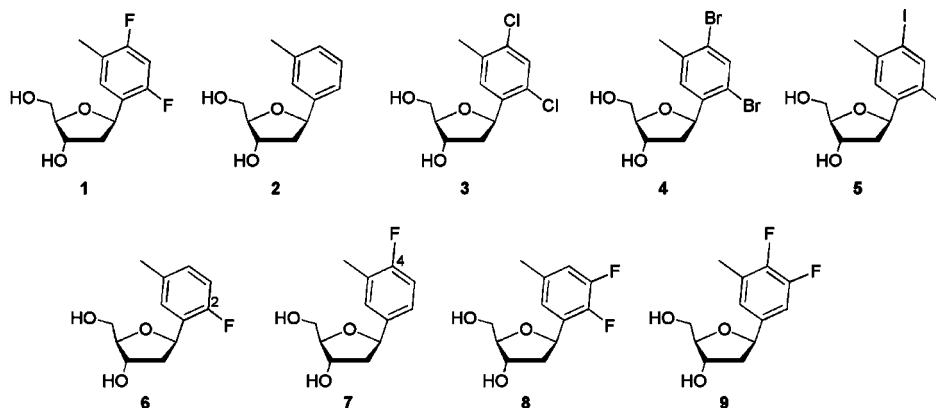
[†] Stanford University.

[‡] Swedish University of Agricultural Sciences.

(1) Al-Madhoun, A. S.; Tjarks, W.; Eriksson, S. *Mini Rev. Med. Chem.* **2004**, *4*, 341–50.

(2) Welin, M.; Kosinska, U.; Mikkelsen, N. E.; Carnrot, C.; Zhu, C.; Wang, L.; Eriksson, S.; Munch-Petersen, B.; Eklund, H. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 17970–17975.

(3) Van Rompay, A. R.; Johansson, M.; Karlsson, A. *Pharmacol. Ther.* **2003**, *100*, 119–39.



telbivudine is a recently approved therapeutic agent for HBV infection.⁴ Similarly, 5-fluorodeoxyuridine is an agent long used to treat breast cancer.⁵ All of these compounds are activated by thymidine kinases.

In addition to therapeutics, isotopically labeled thymidine derivatives are increasingly important as imaging agents; for example, 3'-¹⁸F-dT is now under wide study as a PET imaging agent,⁶ and ¹²³I-deoxyuridine is being investigated as a gamma-emitting tumor imaging agent.⁷ These are also dependent on thymidine kinase for their activities. A better understanding of how such compounds are recognized by TK1 and TK2 can give basic insight into their useful biological activities and may lead to new ideas for future bioactive nucleoside design.

With these goals in mind, here we use nonpolar nucleoside mimics to study the influences of nucleobase electrostatics and sterics on nucleobase recognition by human thymidine kinases. In human thymidine kinases, a range of nucleoside analogues has been previously studied *in vitro*,^{1,8} although the large majority of these analogues involve modifications to the sugar, and only a small number contain modifications to the base.⁸

Few studies have been carried out to systematically probe the hydrogen-bonding face of thymidine, despite the fact that it is this surface that is commonly believed to give thymidine its biochemical identity. Here, we show that one human thymidine kinase can accept nonpolar nucleosides with high efficiency, and that the two human thymidine kinases differ markedly from one another in their recognition mechanisms. Exploiting this mechanism of recognition could have important implications in future drug design, because the two kinases have distinct biological roles.¹ In addition, the results have bearing on the future use of unnatural nucleobases in synthetic genetic systems.⁹

Experimental Section

Synthesis of Nucleoside Analogues. The synthesis of nucleoside analogues **1–5** was carried out as described previously. Fluorinated analogues **6**, **7**, **8**, and **9** were prepared according to Scheme 2 (details below).

5-Bromo-3,4-difluorobenzaldehyde (8b) and 5-Bromo-2,3-difluorobenzaldehyde (9b). 3,4-Difluorobenzaldehyde (**8a**) (7.00 g, 49.23 mmol) was dissolved in concentrated H₂SO₄ (22.3 mL) and heated to 60 °C. To this was added *N*-bromosuccinimide (10.55 g,

59.10 mmol) in three portions over a period of 30 min. After being heated for ~3 h under N₂, the reaction mixture was poured into ice water. The product was extracted with hexanes, washed with water and brine, and then dried over Na₂SO₄.¹⁰ Purification by silica gel column chromatography (0–10% EtOAc in hexanes) yielded an orange liquid as product **8b** in 50% yield: ¹H NMR (CDCl₃, 500 MHz, ppm) δ 10.23 (d, 1H, *J*_{HF} = 2.5168), 7.77 (dd, 1H, *J*_{HF} = 9.2 Hz), 7.51 (dd, 1H, *J*_{HF} = 7.5 Hz); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -124.89 (*J*_{FF} = 15.2 Hz), -135.98 (*J*_{FF} = 11.3 Hz); ¹³C NMR (CDCl₃, 500 MHz, ppm) δ 189.45 (*J*_{CF} = 9.8 Hz), 155.00 (*J*_{CF} = 17.7 Hz), 152.92 (*J*_{CF} = 12.7 Hz), 151.30 (*J*_{CF} = 10.2 Hz), 149.30 (*J*_{CF} = 8.9 Hz). HRMS (EI⁺) calcd mass 219.9335 for [C₇H₃BrF₂O], found 219.9321 for [M]⁺.

The same procedure was done with 2,3-difluorobenzaldehyde (**9a**) to prepare **9b** in 45% yield: ¹H NMR (CDCl₃, 500 MHz) δ 10.29 (s, 1H), 7.77 (m, 1H), 7.59 (m, 1H); ¹⁹F NMR (CDCl₃, 400 MHz) δ -148.51, (m, *J*_{FF} = 15.4), -133.71 (m, *J*_{FF} = 15.4); ¹³C NMR (CDCl₃, 500 MHz) δ 184.55, 152.96, 151.69, 150.93, 149.62, 126.31. HRMS (EI⁺) calcd mass 219.9335 for [C₇H₃BrF₂O], found 219.9327 for [M]⁺.

5-Bromo-3,4-difluorobenzyl alcohol (8c) and 5-Bromo-2,3-difluorobenzyl alcohol (9c). **8b** (3.00 g, 13.57 mmol) was dissolved in 215 mL of ethanol, to which were added 1.30 g (39.60 mmol) of ammonium carbonate followed by 0.51 g (15.53 mmol) of sodium borohydride. Reaction monitored by TLC was complete after ~30 min stirring.¹¹ The reaction mixture was filtered through a pad of Celite, and the filtrate was concentrated to give the white solid **8c** in 80% yield: ¹H NMR (CDCl₃, 400 MHz, ppm) δ 7.37 (m, 2H), 4.66 (s, 2H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -137.79 (*J*_{FF} = 20.5 Hz), -138.60 (*J*_{FF} = 20.5 Hz); ¹³C NMR (CDCl₃, 500 MHz, ppm) δ 126.38 (*J*_{CF} = 5.0 Hz), 121.51, 121.38, 117.27, 117.00, 115.53, 115.40. HRMS (EI⁺) calcd mass 221.9492 for [C₇H₅BrF₂O], found 221.9423 for [M]⁺.

The same procedure was done with **9b** to give **9c** in 73% yield: ¹H NMR (CDCl₃, 400 MHz) δ 7.36 (m, 1H, *J*_{HH} = 0.9 Hz), 7.25 (m, 1H, *J*_{HF} = 2.9 Hz), 4.73 (d, 2H, *J*_{HF} = 6.5 Hz); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -135.77 (*J*_{FF} = 19.5, 10.6 Hz), -146.05 (*J*_{FF} = 19.5 Hz); ¹³C NMR (CDCl₃, 500 MHz, ppm) δ 151.12,

(4) Qi, X.; Xiong, S.; Yang, H.; Miller, M.; Delaney, W. E. *Antivir. Ther.* **2007**, *12*, 355–62.

(5) van Laar, J. A.; Rustum, Y. M.; Ackland, S. P.; van Groeningen, C. J.; Peters, G. J. *Eur. J. Cancer* **1998**, *34*, 296–306.

(6) Herrmann, K.; Wieder, H. A.; Buck, A. K.; Schoffel, M.; Krause, B. J.; Fend, F.; Schuster, T.; Meyer zum Buschenfelde, C.; Wester, H. J.; Duyster, J.; Peschel, C.; Schwaiger, M.; Dechow, T. *Clin. Cancer Res.* **2007**, *13*, 3552–8.

(7) Toyohara, J.; Hayashi, A.; Sato, M.; Gogami, A.; Tanaka, H.; Haraguchi, K.; Yoshimura, Y.; Kumamoto, H.; Yonekura, Y.; Fujibayashi, Y. *Nucl. Med. Biol.* **2003**, *30*, 687–96.

(8) (a) Hakimelahi, G. H.; Ly, T. W.; Moosavi-Movahedi, A. A.; Jain, M. L.; Zakerinia, M.; Davari, H.; Mei, H. C.; Sambaiah, T.; Moshfegh, A. A.; Hakimelahi, S. *J. Med. Chem.* **2001**, *44*, 3710–20. (b) Katona, C.; Timar, F.; Jeney, A.; Fischel, J. L.; Milano, G.; Pandi, E.; Kralovanszky, J. *Anticancer Drugs* **1999**, *10*, 561–7. (c) Al-Madhoun, A. S.; Eriksson, S.; Wang, Z. X.; Naimi, E.; Knaus, E. E.; Wiebe, L. I. *Nucleosides Nucleotides Nucl. Acids* **2004**, *23*, 1865–1874.

(9) (a) Sismour, A. M.; Benner, S. A. *Nucleic Acids Res.* **2005**, *33*, 5640–5646. (b) Henry, A. A.; Romesberg, F. E. *Curr. Opin. Chem. Biol.* **2003**, *7*, 727–733. (c) Hirao, I. *Curr. Opin. Chem. Biol.* **2006**, *10*, 622–627. (d) Krueger, A.; Lu, H.; Lee, A. H.-F.; Kool, E. T. *Acc. Chem. Rev.* **2007**, *40*, 141–150.

149.12, 148.31, 146.35, 131.88, 126.42, 119.69. HRMS (EI⁺) calcd mass 221.9492 for [C₇H₅BrF₂O], found 221.9478 for [M]⁺.

5-Bromo-3,4-difluoro-5-(iodomethyl)benzene (8d) and 5-Bromo-2,3-difluoro-5-(iodomethyl)benzene (9d). Cesium iodide (6.55 g, 25.23 mmol) was dissolved in 5 mL of anhydrous acetonitrile, to which were added **8c** (1.86 g, 8.41 mmol) dissolved in 5 mL of anhydrous acetonitrile and AlCl₃ (3.36 g, 25.2 mmol) dissolved in 5 mL of anhydrous acetonitrile. The orangish-yellow solution was allowed to stir 24 h and was then quenched with ice water and extracted with hexanes.¹² The organic layer was washed with water, brine, and dried over Na₂SO₄ to give 6.68 g of yellow product in 87% yield: ¹H NMR (CDCl₃, 400 MHz) δ 7.39 (t, 1H, J_{HF} = 9.5 Hz), 7.30 (t, 1H, J_{HF} = 8.5 Hz), 4.42 (s, 2H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -134.68, -137.17; ¹³C NMR (CDCl₃, 500 MHz, ppm) δ 150.61 (J_{CF} = 12.9 Hz), 148.62 (J_{CF} = 7.0 Hz), 135.35 (J_{CF} = 5.7 Hz), 122.43, 122.27, 118.85, 118.67. HRMS (EI⁺) calcd mass 331.8509 for [C₇H₄BrF₂I], found 331.8496 for [M]⁺.

The same procedure was done with **9c** to give **9d** in 80% yield: ¹H NMR (CDCl₃, 400 MHz) δ 7.25 (m, 2H, J_{HF} = 10.3 Hz), 4.35 (s, 2H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -134.28, -142.23; ¹³C NMR (CDCl₃, 400 MHz, ppm) δ 151.99, 139.24, 128.57, 120.77, 120.64, 115.55, 55.15. HRMS (EI⁺) calcd mass 331.8509 for [C₇H₄BrF₂I], found 331.8502 for [M]⁺.

5-Bromo-3,4-difluorotoluene (8e) and 5-Bromo-2,3-difluorotoluene (9e). **8d** (1.67 g, 5.02 mmol) was dissolved in 100 mL of THF, to which was added 16.70 g of a slurry of aqueous Raney Nickel. The mixture was stirred at room temperature for 1 h and filtered through a pad of Celite.¹³ The solvent was evaporated to yield the reduced compound **8e** in 65% yield: ¹H NMR (CDCl₃, 400 MHz, ppm) δ 7.34 (t, 1H, J_{HF} = 8.6), 7.04 (t, 1H, J_{HF} = 10.0 Hz), 2.33 (s, 3H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -124.41 (dd, J_{FF} = 22.4 Hz, 9.6 Hz), -136.63 (dd, J_{FF} = 22.4 Hz, 11.4 Hz); ¹³C NMR (CDCl₃, 500 MHz, ppm) δ 150.69, 148.21, 134.82, 125.70, 121.26 (J_{CF} = 15.7 Hz), 118.99, 104.98. HRMS (EI⁺) calcd mass 205.9543 for [C₇H₅BrF₂], found 205.9536 for [M]⁺.

The same procedure was done with **9d** to give **9e** in 46% yield: ¹H NMR (CDCl₃, 400 MHz, ppm) δ 7.14 (t, 1H, J_{HF} = 8.0), 7.09 (t, 1H, J_{HF} = 7.1), 2.28 (d, 3H, J_{HF} = 2.4); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -136.23 (dd, J_{FF} = 20.1 Hz, 9.0 Hz), -143.95 (d, J_{FF} = 20.1 Hz); ¹³C NMR (CDCl₃, 400 MHz, ppm) δ 151.95, 149.97, 149.45, 149.29, 147.57, 147.44, 30.55. HRMS (EI⁺) calcd mass 205.9543 for [C₇H₅BrF₂], found 205.9535 for [M]⁺.

General Procedure for C-Nucleoside Couplings. A previously described method of C-nucleoside coupling¹⁴ was used to yield new fluorinated aromatic nucleosides **6b**, **7b**, **8f**, and **9f** as the tetraisopropyl disiloxane intermediate. The synthesis involves the reaction of aryllithium derivatives of the fluorinated aromatic compounds with 3',5'-O-((1,1,3,3-tetraisopropyl)disiloxanediyl)-2'-deoxy-D-ribo-1,4-lactone in pentane at -78 °C to give a hemiketal, which is then reduced to give a disiloxane intermediate. This intermediate was then deprotected using tetrabutylammonium fluoride (TBAF) to give the free nucleoside compounds **6**, **7**, **8**, and **9**.

Procedure for Lactone Coupling Reaction To Yield Protected Nucleoside Analogues. **6a** (0.250 g, 1.32 mmol) was dissolved in anhydrous pentane (2 mL) and cooled to -78 °C. *t*-Butyllithium (1.5 M in pentane, 1 equiv) was added slowly to the mixture and stirred for 30 min under Ar. 3',5'-O-((1,1,3,3-Tetraisopropyl)disiloxanediyl)-2'-deoxy-D-ribo-1,4-lactone dissolved in anhydrous

THF (2 mL) was added slowly to the mixture and allowed to stir at -78 °C for 3 h. The reaction was quenched with saturated aqueous NH₄Cl at -78 °C and allowed to warm to room temperature. The solution was extracted with ether (2 × 10 mL) and washed with saturated aqueous NH₄Cl, water, and brine. The organic layer was dried over Na₂SO₄ and concentrated as a yellow oil.

The crude product was allowed to dry overnight and was then dissolved in anhydrous CH₂Cl₂ (4 mL) and cooled to -78 °C under argon. Triethylsilane (3 equiv) and boron trifluoride diethyl etherate (3 equiv) were slowly added to the mixture, and the reaction proceeded for 4 h. The reaction was quenched with saturated aqueous NaHCO₃ and allowed to warm to room temperature. The solution was extracted with ether and washed with saturated aqueous NaHCO₃, water, and brine. The organic layer was dried over Na₂SO₄, concentrated as a yellow oil, and purified by flash silica gel chromatography (5:1 hexanes:ethyl acetate) to yield 3',5'-O-((1,1,3,3-tetraisopropyl)disiloxanediyl)-1'-2'-dideoxy-β-1'-(2-fluoro-5-methylphenyl)-D-ribofuranose (compound **6b**) in 26% yield: ¹H NMR (CDCl₃, 400 MHz) δ 7.28 (m, 1H), 7.00 (m, 1H, J_{HF} = 8.2 Hz), 6.88 (t, 1H, J_{HF} = 9.3 Hz), 5.30 (t, 1H, J_{HF} = 7.5 Hz, H-1'), 4.50 (dd, 1H, J_{HF} = 13.4, 6.5 Hz, H-3'), 4.12 (dd, 1H, H-5'), 3.98–3.92 (m, 1H, H-5'), 3.87–3.83 (m, 1H, H-4'), 2.46–2.40 (m, 1H, H-2'α), 2.30 (s, 3H, CH₃), 2.10–2.04 (m, 1H, H-2'β), 1.03 (m, 28H, 4 × -CH(CH₃)₂); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -124.88; ¹³C NMR (CDCl₃, 500 MHz) δ 159.09, 157.12, 133.73, 129.34, 127.52, 115.07, 114.92, 85.93, 73.22, 72.56, 63.22, 42.00, 29.94, 21.02, 17.84, 17.50, 17.34, 17.26, 13.73, 13.60, 13.16, 12.79. HRMS (EI⁺, Na⁺) calcd mass 491.2425 for [C₂₄H₄₀F₂O₄Si₂ + Na]⁺, found 491.2423 for [M + Na]⁺.

3',5'-O-((1,1,3,3-Tetraisopropyl)disiloxanediyl)-1'-2'-dideoxy-β-1'-(4-fluoro-5-methylphenyl)-D-ribofuranose (Compound 7b) in 21% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.15 (d, 1H, J_{HF} = 7.9 Hz), 7.11 (m, 1H), 6.94 t, 1H, J_{HF} = 8.9 Hz), 5.02 (t, 1H, J_{HF} = 7.0 Hz, H-1'), 4.52 (m, 1H, H-3'), 4.12–4.09 (m, 1H, H-5'), 3.92–3.86 (m, 1H, H-5'), 2.36–2.30 (m, 1H, H-2'α), 2.26 (d, 3H, CH₃, J_{HF} = 2.0 Hz), 2.07–2.01 (m, 1H, H-2'β), 1.06 (m, 28H, 4 × -CH(CH₃)₂); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -119.92; ¹³C NMR (CDCl₃, 400 MHz) δ 137.61, 129.16, 125.06, 115.11, 114.88, 89.89, 86.65, 78.80, 73.38, 63.78, 43.31, 29.98, 17.82, 17.70, 17.68, 17.48, 17.35, 14.79, 13.74, 13.63, 13.24, 12.74. HRMS (EI⁺, Na⁺) calcd mass 491.2425 for [C₂₄H₄₀F₂O₄Si₂ + Na]⁺, found 491.2424 for [M + Na]⁺.

3',5'-O-((1,1,3,3-Tetraisopropyl)disiloxanediyl)-1'-2'-dideoxy-β-1'-(2,3-fluoro-5-methylphenyl)-D-ribofuranose (Compound 8f) in 20% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.38 (dd, 1H, J_{HF} = 9.9 Hz), 6.94 (dd, 1H, J_{HF} = 10.4 Hz), 5.17 (t, 1H, J_{HF} = 6.0 Hz, H-1'), 4.51–4.47 (m, 1H, H-3'), 4.14 (dd, 1H, H-5'), 4.03–3.94 (m, 1H, H-5'), 3.87–3.84 (m, 1H, H-4'), 2.45–2.40 (m, 1H, H-2'α), 2.25 (s, 3H, CH₃), 2.2–2.06 (m, 1H, H-2'β), 1.12–1.02 (m, 28H, 4 × -CH(CH₃)₂); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -141.87 (J_{FF} = 12.2 Hz), -142.19 (J_{FF} = 12.2 Hz); ¹³C NMR (CDCl₃, 400 MHz) δ 149.52, 148.46, 138.35, 130.51, 127.20, 125.83, 123.37, 112.25, 105.62, 86.47, 72.82, 63.41, 43.29, 17.80, 17.68, 17.47, 17.32, 17.19, 14.75, 14.37, 13.73, 13.64, 13.20, 12.75. HRMS (EI⁺, Na⁺) calcd mass 509.2331 for [C₂₄H₄₀F₂O₄Si₂ + Na]⁺, found 509.2346 for [M + Na]⁺.

3',5'-O-((1,1,3,3-Tetraisopropyl)disiloxanediyl)-1'-2'-dideoxy-β-1'-(3,4-fluoro-5-methylphenyl)-D-ribofuranose (Compound 9f from 9e) in 18% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.00 (m, 1H, J_{HF} = 5.2 Hz), 6.90 (m, 1H, J_{HF} = 7.2 Hz), 4.99 (t, 1H, J_{HF} = 8.1 Hz), 4.49 (b, 1H, H-3'), 4.12 (m, 1H, H-5'), 4.09 (m, 1H, H-5'), 3.90–3.85 (m, 1H, H-4'), 2.37–2.31 (m, 1H, H-2'α), 2.28 (d, 3H, CH₃, J_{HF} = 2.1 Hz), 2.02–1.95 (m, 1H, H-2'β), 1.03 (m, 28H, 4 × -CH(CH₃)₂); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ -139.18 (J_{FF} = 21.2, 10.3 Hz), -144.58 (J_{FF} = 21.2 Hz); ¹³C NMR (CDCl₃, 500 MHz) δ 148.46, 138.34, 130.51, 127.20, 125.83, 123.37, 86.47, 72.82, 70.56, 63.41, 43.29, 17.8, 17.47, 17.32, 14.75, 13.73, 13.20,

(10) Rajesh, K.; Somasundaram, M.; Saiganesh, R.; Balasubramanian, K. K. *J. Org. Chem.* **2007**, *72*, 5867–5869.

(11) Mohanazadeh, F.; Hosini, M.; Tajbakhsh, M. *Monatsh. Chem.* **2005**, *136*, 2041–2043.

(12) Khan, K. M.; Zia-Ullah, S.; Perveen, S.; Maharvi, G. M.; Shah, S. T. A.; Ambreen, N.; Choudhary, M. I.; Atta-ur-Rahman; Voelter, W. *Lett. Org. Chem.* **2005**, *2*, 644–647.

(13) Barrero, A. F.; Alvarez-Manzaneda, E. J.; Chahboun, R.; Meneses, R.; Romera, J. L. *Synlett* **2001**, *4*, 485–488.

(14) Wichai, U.; Woski, S. A. *Org. Lett.* **1999**, *1*, 1173–1175.

12.75. HRMS (EI⁺, NH₃) calcd mass 504.2777 for [C₂₄H₄₀F₂O₄Si₂ + NH₄]⁺, found 504.2775 for [M + NH₄]⁺.

Procedure for Deprotection of Nucleoside Analogues. 6b (0.06 g, 0.12 mmol) was dissolved in dry THF (3 mL). Tetrabutylammonium fluoride (1.0 M, THF, 0.36 mL) was added dropwise, and the solution was allowed to stir for 2 h at room temperature. Upon completion, the reaction was quenched with 5% potassium bicarbonate (1 mL) and extracted twice with ether. The organic layer was washed with 5% potassium bicarbonate, water, and brine and then concentrated to give a yellow oil. Silica gel chromatography (0–10% CH₃OH in CH₂Cl₂) gave a white, crystalline solid, 1'-2'-dideoxy-β-1'-(2-fluoro-5-methylphenyl)-D-ribofuranose (**6**) from **6b** in 67% yield: ¹H NMR (CDCl₃, 400 MHz) δ 7.21 (d, 1H, J_{HF} = 7.5 Hz), 7.04 (m, 1H), 6.92 (dd, 1H, J_{HF} = 9.8 Hz), 5.34 (dd, 1H, J_{HF} = 16.5, 5.5 Hz), 4.48 (br, 1H), 4.04–4.01 (m, 1H), 3.89–3.75 (m, 2H), 2.33–2.29 (m, 1H), 2.32 (s, 3H), 2.12–2.04 (m, 1H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ –123.88; ¹³C NMR (CDCl₃, 400 MHz) δ 129.71, 127.87, 124.60, 115.80, 107.72, 87.29, 74.04, 63.56, 40.56, 29.96, 24.02, 21.03. HRMS (EI⁺, NH₃) calcd mass 244.1349 for [C₁₂H₁₅FO₃ + NH₄]⁺, found 244.1347 for [M + NH₄]⁺.

1'-2'-Dideoxy-β-1'-(4-fluoro-5-methylphenyl)-D-ribofuranose (7) from **7b** in 70% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.16 (d, 1H, J_{HF} = 8.5 Hz), 7.12 (m, 1H, J_{HF} = 7.5 Hz), 6.97 (t, 1H, J_{HF} = 6.5 Hz), 5.11 (dd, 1H, J_{HF} = 9.8, 6.5 Hz), 4.45 (m, 1H), 4.02–3.99 (m, 1H), 3.87–3.83 (m, 1H), 3.78–3.73 (m, 1H), 2.27 (s, 3H, J_{HF} = 2.0 Hz), 2.26–2.21 (m, 1H), 2.07–1.98 (m, 1H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ –119.33 Hz; ¹³C NMR (CDCl₃, 500 MHz) δ 129.44, 125.27, 115.20 (J_{CF} = 23.7 Hz), 87.47, 79.88, 74.04, 63.63, 59.29, 44.34, 29.92, 24.36, 20.02. HRMS (EI⁺, NH₃) calcd mass 244.1349 for [C₁₂H₁₅FO₃ + NH₄]⁺, found 244.1348 for [M + NH₄]⁺.

1'-2'-Dideoxy-β-1'-(2,3-fluoro-5-methylphenyl)-D-ribofuranose (8) from **8f** in 67% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 7.26 (1H), 6.94 (dd, 1H, J_{HF} = 8.8 Hz), 5.25 (dd, 1H, J_{HF} = 15.1, 4.8 Hz), 4.47 (br, 1H), 4.00 (m, 1H), 3.86–3.82 (m, 2H), 2.28–2.21 (m, 1H), 2.28 (s, 3H), 1.89–1.81 (m, 1H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ –141.11, –144.53; ¹³C NMR (CDCl₃, 400 MHz) δ 138.72, 136.54, 128.07, 118.36, 114.05, 104.53, 86.94, 63.25, 55.91, 42.93, 29.79, 18.56. HRMS (EI⁺, NH₃) calcd mass 262.1255 for [C₁₂H₁₄F₂O₃ + NH₄]⁺, found 262.1253 for [M + NH₄]⁺.

1'-2'-Dideoxy-β-1'-(3,4-fluoro-5-methylphenyl)-D-ribofuranose (9) from **9f** in 60% Yield. ¹H NMR (CDCl₃, 400 MHz) δ 6.98 (dd, 1H, J_{HF} = 10.7, 7.7 Hz), 6.90 (d, 1H, J_{HF} = 6.3 Hz), 5.07 (m, 1H, J_{HF} = 9.8, 5.8 Hz), 4.43 (m, 1H), 4.02–3.99 (m, 1H), 3.85–3.73 (m, 2H), 2.38–2.33 (m, 1H), 2.29 (s, 3H), 1.99, 1.93 (m, 1H); ¹⁹F NMR (CDCl₃, 400 MHz, ppm) δ –138.36 (J_{FF} = 11.6 Hz), –143.97 (J_{FF} = 2.14 Hz); ¹³C NMR (CDCl₃, 500 MHz) δ 137.09, 127.32, 123.60, 112.28, 111.28, 84.50, 79.10, 73.67, 63.37, 44.29, 29.79, 19.8. HRMS (EI⁺, NH₃) calcd mass 262.1255 for [C₁₂H₁₄F₂O₃ + NH₄]⁺, found 262.1254 for [M + NH₄]⁺.

Methods for Preparation of Kinases TK1 and TK2. Recombinant human TK1 and TK2 were expressed and purified using a bacterial vector system *E. coli* BL21 (DE3) pLysS.^{15,16} This vector allows efficient purification of target proteins with an affinity tag of six histidines giving a rapid and complete purification in one step. The vector also contains a thrombin cleavage site fused to the N-terminus of the coding sequence. The His-tag could be effectively removed using Thrombin (Pharmacia Biotech). Addition of 10 mM DTT and 20% glycerol to samples was done directly after elution. The purity of the enzymes and the extent of digestion with Thrombin were determined using SDS-PAGE.

Kinetics Methods. The adenosine 5'-triphosphate transfer assay was performed with 0.05 μM [^γ-³²P]-ATP (10 μCi/μL), 100 μM

ATP, 50 mM Tris-HCl (pH 7.6), 5 mM MgCl₂, 100 mM KCl, 0.5 mg/mL BSA, 10 mM DTT, and different concentrations of nucleoside analogues. The reaction was initiated by adding enzyme followed by incubation at 37 °C and terminated by boiling after 20 min. Two microliters of the supernatant was applied to PEI cellulose FTLC plates. Chromatography was performed for 8–12 h with isobutyric acid:NH₄OH:H₂O (66:1:33) (v/v) as the mobile phase. The products of the kinase reaction were detected by autoradiography or phosphoimager analysis. The spots were excised and eluted with 0.5 mL of 0.2 M KCl/0.1 M HCl (1:1, v/v) and quantified by liquid scintillation counting or by phosphorimager (Fuji BAS 25000/LAS 1000, FUJIFILM, 1&1-Imaging & Information) with Image Reader V 1.7E. The apparent K_m and V_{max} values were determined using the Michaelis–Menten equation and nonlinear regression analysis with the SigmaPlot 2001 program (Enzyme Kinetics from SPSS Science). Kinetics values were obtained using a single enzyme preparation and were repeated at least twice with similar values.

Results

First, we investigated the effects of nucleobase electrostatics on TK1 and TK2 substrate recognition by comparing the natural substrate thymidine with the isosteric difluorotoluene deoxy-nucleoside (Figure 1).¹⁷

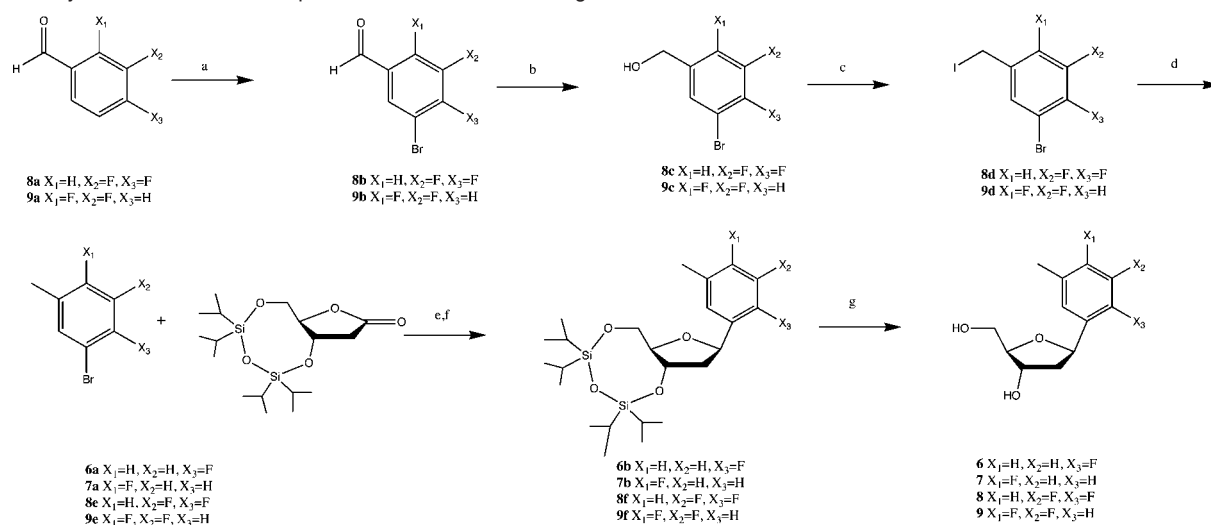
The latter has size and shape nearly identical to that of the natural compound, but with considerably lower local atomic charge densities, and a greatly attenuated ability to undergo hydrogen bonding.^{17a,e} We evaluated phosphorylation of the nucleosides in vitro (using purified kinases) by use of ^γ-³²P-ATP and thin layer chromatographic (TLC) analysis of products, with quantitation by autoradiography. Table 1 shows data from one time point at varied nucleoside concentrations. The results show (see first and second entries in the table) that TK1 phosphorylated the nonpolar isostere **1** at 3.8% of the yield of the natural compound (at 100 μM substrate, 20 min), which suggests a substantial dependence of this enzyme on electrostatics of the nucleobase. In contrast to this, TK2 phosphorylated the analogue **1** at nearly as high a yield as the natural compound (in agreement with a preliminary study that included **1**),^{8c} indicating that this mitochondrial enzyme requires little electrostatic charge or hydrogen-bonding ability to efficiently recognize and phosphorylate a substrate.

A polar substrate such as thymidine might be recognized by enzymes either by use of complementary electrostatics, or by providing a sterically complementary active site pocket, or both. To probe steric selectivity of the two kinases, we employed a series of nonpolar thymidine mimics with systematically varied size.¹⁸ In these nucleoside analogues, the projecting groups at the 2 and 4 positions (which are carbonyl groups in thymine) are varied by substituting toluene with H, F, Cl, Br, and I substituents (Figure 2). Thus, the range of compounds varies by 1 Å in size across the series. In electrostatic terms, the new compounds (**2–5**) are even less polar than the difluoro compound (**1**), which is measurably quite hydrophobic as a result of its hydrofluorocarbon aromatic “nucleobase” ring. Thus, electrostatic effects across the series are expected to be quite small.

- (17) (a) Kool, E. T.; Morales, J. C.; Guckian, K. M. *Angew. Chem., Int. Ed.* **2000**, *39*, 990–1009. (b) Kool, E. T.; Sintim, H. O. *Chem. Commun.* **2006**, 3665–3675. (c) Schweitzer, B. A.; Kool, E. T. *J. Org. Chem.* **1994**, *59*, 7238–7242. (d) Guckian, K. M.; Kool, E. T. *Angew. Chem., Int. Ed.* **1998**, *36*, 2825–2828. (e) Wang, X.; Houk, K. N. *Chem. Commun.* **1998**, 2631–2632.
- (18) Kim, T. W.; Delaney, J. C.; Essigmann, J. M.; Kool, E. T. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 15803–15808.

(15) Usova, E. V.; Eriksson, S. *Eur. J. Biochem.* **1997**, *248*, 762–766.

(16) Chottiner, E. G.; Shewach, D. S.; Datta, N. S.; Ashcraft, E.; Gribbin, D.; Ginsburg, D.; Fox, I. H.; Mitchell, B. S. *Proc. Natl. Acad. Sci. U.S.A.* **1991**, *88*, 1531–1535.

Scheme 2. Synthetic Scheme for Preparation of Nucleoside Analogues 6–9^a

^a Reagents and conditions: (a) NBS (1.2 equiv), H₂SO₄, 60 °C; (b) NaBH₄, (NH₄)₂CO₃, EtOH; (c) CsI, AlCl₃, CH₃CN; (d) Raney Ni, THF, room temperature; (e) *t*-BuLi, pentane, THF, -78 °C; (f) Et₃SiH, BF₃-OEt₂, CH₂Cl₂, -78 °C; (g) TBAF, THF.

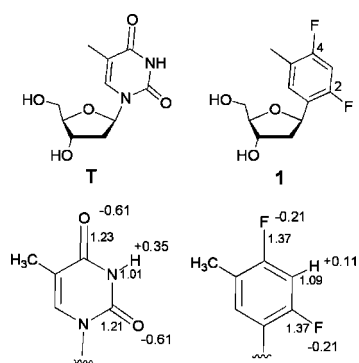


Figure 1. Comparison of structures and electrostatics for thymidine (T) and low-polarity nucleoside isostere **1**. Bond lengths (Å) and calculated atomic charges are taken from ref 17a.

The results of phosphorylation with these variably sized compounds are presented graphically in Figure 2B, and yields are tabulated in Supporting Information Table S1. The data establish that both TK1 and TK2 show substantial steric preferences for the difluoro analogue, which is the analogue closest in size to natural thymidine. Both smaller (H) and larger substituents (Cl, Br, I) lead to lower phosphorylation yields. This size preference may be attributed to both enzymes having relatively rigid active sites that best accommodate the size of the natural substrate.

The above data showed that while TK1 is sensitive to electrostatics, TK2 is not, thus allowing it to accept a nonpolar nucleoside analogue readily. Our data show that both enzymes have a distinct size preference for a difluoro-substituted benzene relative to smaller or larger variants. However, because all analogues have substituents at the 2,4 positions, the above data did not address whether the active site is most sensitive simply to the size/volume of the nucleobase, or rather its shape¹⁹ (i.e., the steric projections in specific locations) during substrate recognition. To test this explicitly for TK2, which accepts nonpolar analogues readily, we prepared a third set of nonpolar nucleoside analogues (**6–9**) as possible substrates. In the first

Table 1. Yields for 5'-Phosphorylation of Nucleoside Analogues by Human Thymidine Kinases 1 and 2 (TK1 and TK2)^a

nucleoside analogue	concentration (μM)	yields (relative, %) ^b	
		TK1	TK2
thymidine	100	100	100
2,4-diF (1)	100	3.8	82
	200	10	116
	400	21	173
2,4-diH (2)	200	<0.01	0.7
	400	0.03	nd
2-F (6)	50	nd	3.3
	100	nd	6.5
	200	<0.01	9.9
	400	<0.01	20.4
4-F (7)	50	nd	16.1
	100	nd	25.8
	200	<0.01	42.7
	400	<0.01	60.2
2,3-diF (8)	50	nd	2.7
	100	nd	5.4
	200	<0.01	9.1
	400	<0.01	14.4
3,4-diF (9)	50	nd	0.6
	100	nd	2.5
	200	<0.01	3.8
	400	<0.01	6.1

^a Conditions: 20 min incubation in a buffer containing 100 μM ATP, 50 mM Tris·zmd/HCl (pH 7.6), 10 mM MgCl₂, 50 μg/mL BSA, and 10 mM dithiothreitol, incubated at 37 °C. ^b Yields are normalized to that obtained with thymidine at 100 μM.

two of these (**6, 7**), there is a single fluorine substituent at the 2 or 4 position to test individually the importance of the two steric projections associated with the natural substrate.

The remaining two compounds (**8, 9**) are the same size as the high-activity substrate **1**, but with the 2- or 4-substituent moved to the 3-position. These latter two test variations in shape without substantially altering size or electrostatic charge on the substituents. Details of the synthetic preparation and characterization of these compounds are given in the Experimental Section. The data for phosphorylation yields with these substrates are shown in Table 1; as expected, they were not measurable substrates for TK1, but were accepted with varying efficiency by TK2.

(19) Sintim, H. O.; Kool, E. T. *Angew. Chem., Int. Ed.* **2006**, *45*, 1974–1979.

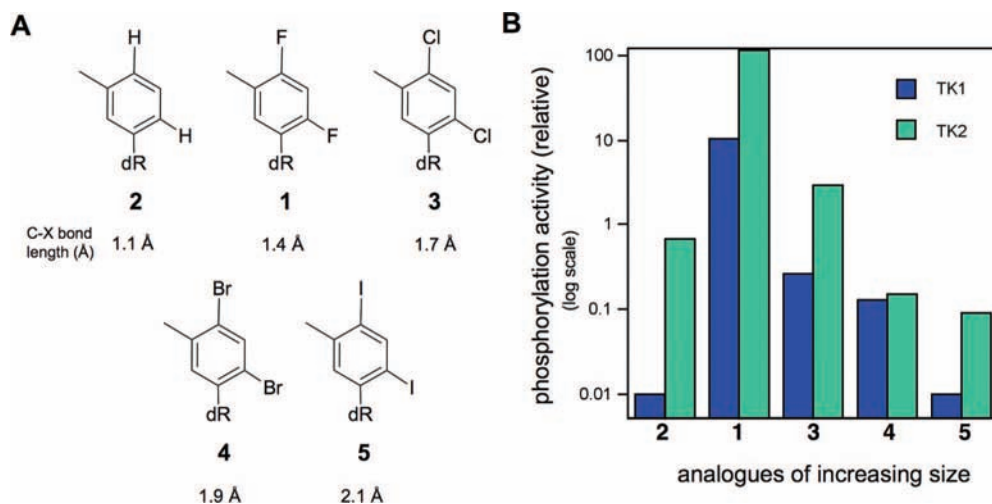


Figure 2. Variably sized thymidine mimics as substrates for human thymidine kinases. (A) Analogue sizes as measured by C–X bond length (R = 2-deoxyribose). (B) Phosphorylation activities with TK1 and TK2 as a function of analogue size; data are taken from Table S1, 200 μM concentration, and are normalized to thymidine. Note log scale.

The experimental results for these analogues with TK2 show that a single substituent at the 2 or 4 position is beneficial to activity, and that the 4-substituent is more important for activity than is the 2-substituent. Indeed, the 4-fluorotoluene analogue **7** is a relatively efficient substrate for this enzyme, although not as good as compound **1**, which has both of these substitutions. Importantly, compounds **8** and **9**, which have shapes altered from the original **1**, are relatively poor substrates. Thus, we conclude that the human TK2 enzyme clearly prefers benzene (pyrimidine-like) substrates in which there is a steric projection at both the 2- and 4-positions, and in which there is no projection (even a small fluorine) at the 3-position, which, of course, is the steric definition of thymine itself. What is surprising is how sensitive the enzyme is to small steric changes: altering a hydrogen substituent to fluorine results in only a ~ 0.25 Å increase in steric size, and yet the enzyme clearly senses this difference at all three positions on the hydrogen-bonding edge of the pyrimidine.

To gather more information on the effects of sterics in recognition by TK2, we carried out more in-depth studies of nucleoside phosphorylation at varied nucleotide concentration for the series **1**, **6**, **7**, **8**, **9**, and fit the data to obtain apparent Michaelis–Menten parameters (see Figure 3 and Experimental Section). The results are shown in Table 2, and detailed data are given in Table S1. Limited solubility of some of the nonpolar substrates prevented us from using high enough concentrations to achieve true saturation for most of the analogues except **1**, thus increasing uncertainty in the kinetic parameters. Given that limitation, the data for **1**, **7**, **8**, **9** show that, in general, the differences between the substrates are in most cases reflected in altered K_m values rather than in changes in V_{max} . The better substrates (e.g., **1**, **7**) show apparent V_{max} values close to or higher than natural thymidine; however, apparent K_m values are less favorable by a factor of 3–27.

Interestingly, the altered-shape analogues **8** and **9** do show a lowered apparent V_{max} value, suggesting that substitution at the 3-position may result in some misalignment in the active site, hindering the reaching of a rate-limiting competent complex. Remarkably, the two best analogues are quite efficient substrates, with the 4-fluoro analogue **7** showing apparent efficiency ($V_{max}/$

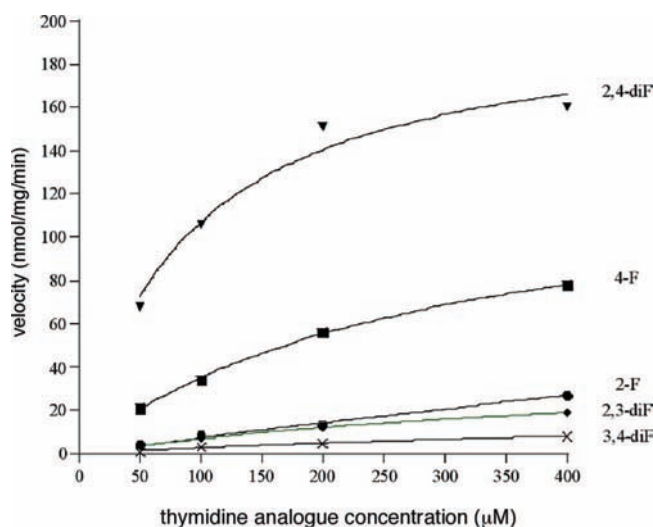


Figure 3. Concentration curves for phosphorylation of 2,4-diF (\blacktriangledown), 4-F (\blacksquare), 2-F (\bullet), 2,3-diF (\blacklozenge), and 3,4-diF (\times) analogues by TK2, using 0.05 μM [γ - ^{32}P]-ATP (10 $\mu\text{Ci}/\mu\text{L}$), 100 μM ATP, 50 mM Tris-HCl (pH 7.6), 5 mM MgCl_2 , 100 mM KCl, 0.5 mg/mL BSA, 10 mM DTT. The reaction was initiated by adding enzyme followed by incubation at 37 $^\circ\text{C}$ for 20 min.

Table 2. Apparent Michaelis–Menten Kinetics Parameters for 5'-Phosphorylation of Nucleoside Analogues by Human Thymidine Kinase 2 (hTK2)

nucleoside analogue	$K_m(\text{app})^a$ (μM)	$V_{max(\text{app})}^a$ (nmol/min/mg)	V_{max}/K_m
thymidine (dThd)	23 ± 4	140 ± 10	6.0
2,4-diF (1)	64 ± 10	150 ± 10	2.4
2-F (6)	nd	nd	nd
4-F (7)	410 ± 40	180 ± 10	0.44
2,3-diF (8)	570 ± 40	46 ± 2	0.08
3,4-diF (9)	630 ± 420	21 ± 9	0.03

^a K_m and V_{max} values were determined by curve fitting with the SigmaPlot program using a range of concentrations as follows: dThd and compound **1** (3–200 μM); compounds **7**, **8**, **9** (12–400 μM).

K_m) within 14-fold of the natural substrate, and the 2,4-difluoro analogue **1** within 2.5-fold of natural thymidine.

Discussion

The current experiments establish clearly that the two human thymidine kinases recognize their substrates by very different

chemical criteria. Examination of the nucleoside binding site of human TK1² (with analogy to the related *U. urealyticum* enzyme structure²⁰) suggests that polar protein backbone groups form hydrogen bonds to all three polar groups (the two carbonyls and imino proton) of thymine (see Supporting Information Figures S1–S3). The substrate nucleobase is well buried, with no access to solvent. The current results with TK1 suggest that loss of these hydrogen bonds with concomitant desolvation of these side chain groups is costly, making nonpolar structures poor substrates. Interestingly, combined analysis of the sequence of TK2 along with the homologous dNK structure²¹ suggests that TK2 also makes hydrogen bonds to the base, in this case, from a glutamine to the O4 and NH groups of thymine. However, the base binding pocket of this latter enzyme appears to be more solvent accessible than that of TK1, and direct solvent access near O2 appears possible in TK2 (see Figures S1–S3). Thus, for TK2 it seems reasonable that a nonpolar substrate can occupy the active site without invoking full desolvation costs. Moreover, greater solvent accessibility would be expected to lower the energetic contribution of H-bonds by presenting a higher local dielectric.

A previous study of the phosphorylation of unnatural nucleosides (primarily with modified indole nucleobases) by a fruit fly kinase, *Dm-dNK*, demonstrated the flexibility of that enzyme's active site and the small influence of electronic effects.²² However, that study also concluded that hydrogen bonding is necessary for effective enzymatic activity. This is in contrast to the current results with human TK2, which show that hydrogen bonding is not essential for effective phosphorylation. This difference is somewhat surprising, given the homology of dNK with TK2, but it is possible that small sequence differences are the cause or that steric constraints in *Dm-dNK* were too restrictive for the analogues tested.

The current results have potentially useful implications for future nucleoside design. For example, there is the possibility

of selectively targeting one of the two classes of thymidine kinases over the other.²³ Knowing the mechanisms of TK1 and TK2 substrate selectivity might enable design of antiviral therapeutics, where mitochondrial toxicity resulting from unintended phosphorylation by TK2 and subsequent incorporation of the analogue nucleotide into mitochondrial DNA can be avoided. In addition, ways to prevent inhibition of TK2 phosphorylation of the natural mitochondrial thymidine substrate by the presence of a competing antiviral analogue are desirable, because such a mechanism of mitochondrial toxicity has been demonstrated in case of AZT.²⁴

Interestingly, some related difluorobenzene nucleoside analogues have been shown to be relatively nontoxic in human cell lines.²⁵ The current data suggest that they may well have been phosphorylated by kinases in these cells. Their fate after this potential phosphorylation remains unknown, but the possibility of their activity in additional cellular nucleotide-processing pathways is intriguing and merits further study.

The current data also have implications in the design of bases and base pairs for augmenting the natural genetic system. Future nucleobase designs, if meant to function in the nucleotide salvage pathway, will need to address what types of cells are to be the host, and to determine whether TK1 or TK2 will be the intended target. Polar designs may be best suited to TK1 in rapidly dividing cells, and nonpolar designs should be targeted to TK2. Alternatively, one may need to engineer kinases with new selectivities.²⁶ The current data suggest some molecular strategies that may alter this selectivity, and the current nucleoside analogues may be useful as probes of new enzymes that use thymidine as substrate.

Acknowledgment. This work was supported by NIH grant GM-072705 and by grants to S.E. from the Swedish Research Council.

Supporting Information Available: Data for kinetics analysis of kinases, NMR spectra of nucleoside analogues, and active site models. This material is available free of charge via the Internet at <http://pubs.acs.org>.

JA808244T

- (20) Kosinska, U.; Carnrot, C.; Eriksson, S.; Wang, L.; Eklund, H. *FEBS J.* **2005**, *272*, 6365–6372.
- (21) Johansson, K.; Ramaswamy, S.; Ljungcrantz, C.; Knecht, W.; Piskur, J.; Munch-Petersen, B.; Eriksson, S.; Eklund, H. *Nat. Struct. Biol.* **2001**, *8*, 616–620.
- (22) Wu, Y.; Fa, M.; Tae, E. L.; Schultz, P. G.; Romesberg, F. E. *J. Am. Chem. Soc.* **2002**, *124*, 14626–14630.
- (23) Hannigan, B. M.; Barnett, Y. A.; Armstrong, D. B.; McKelvey-Martin, V. J.; McKenna, P. G. *Cancer Biother.* **1993**, *8*, 189–97.

- (24) Lynx, M. D.; McKee, E. E. *Biochem. Pharmacol.* **2006**, *72*, 239–243.
- (25) Wang, Z. X.; Duan, W.; Wiebe, L. I.; Balzarini, J.; De Clercq, E.; Knaus, E. E. *Nucleosides Nucleotides Nucl. Acids* **2001**, *20*, 41–58.
- (26) Gerth, M. L.; Lutz, S. *J. Mol. Biol.* **2007**, *370*, 742–751.