Conformational Analysis of 2′,3′-Didehydro-2′,3′-dideoxypyrimidine Nucleosides

Patrick Van Roey,*,† E. Will Taylor,‡ Chung K. Chu,‡ and Raymond F. Schinazi§

Contribution from the Medical Foundation of Buffalo, Inc., Buffalo, New York 14203, the Department of Medicinal Chemistry, College of Pharmacy, University of Georgia, Athens, Georgia 30602, and the Department of Pediatrics, Emory University School of Medicine and the Veterans Affairs Medical Center, Decatur, Georgia 30033

Received March 12, 1992

Abstract: The molecular conformations of 2′,3′-didehydro-2′,3′-dideoxypyrimidine nucleosides have been studied by X-ray crystallographic analysis and molecular mechanics calculations. In this paper, the crystal structures of the uridine, cytidine, and 5-ethyluridine analogues are reported. The structures of two crystal forms of the thymidine analogue and another crystal form of the cytidine analogue that were reported elsewhere (Gurskaya *et al. Int. J. Purine Pyrimidine Res.* **1991**, *2*, 55–60. Harte *et al. Biochem. Biophys. Res. Commun.* **1991**, *175*, 298–304. Birnbaum *et al. Nucleosides Nucleotides* **1989**, *8*, 1259–1269.) are included in the analysis. In total twelve independent observations of the molecular conformations of the unsaturated nucleosides are available because all six crystal forms have two molecules in the asymmetric unit. All molecules have the dihydrofuran ring in a nearly planar conformation with a slight tendency toward O4′-*endo*. The glycosylic link is observed in the unusual high-*anti* conformation in five of the twelve molecules. All but one of the 5′-hydroxyl groups are observed in the +*sc* conformation. The conformations are compared with those of their 2′,3′-dideoxy and 3′-substituted-2′,3′-dideoxypyrimidine analogues. Least-squares fitting of the central four atoms (N1, C1′, C2′, and O4′) of representative molecules of the three groups shows two distinct features of the furanose ring conformations of the saturated compounds relative to those of the unsaturated analogues: the exocyclic deviation at C5′ is smaller for the common C3′-*exo*/C2′-*endo* or C3′-*endo*/C2′-*exo* conformations than for the extreme C3′-*exo*/C4′-*endo* conformations observed frequently for the saturated and 3′-substituted analogues that have high anti-HIV activity. The opposite is observed for the endocyclic deviation at C3′. The molecular mechanics calculations, consisting of γ,χ-maps for the saturated and unsaturated thymidine analogues, confirm that the high-*anti* glycosylic link conformation is more accessible and that overall more area of conformational space is available for the unsaturated than for the saturated compound. This can be explained by the reduced ring puckering and the absence of out-of-plane hydrogen atom substituents at C2′ and C3′. This increased flexibility of the molecule may make it more amenable for phosphorylation by making the "active site" conformation more accessible.

Introduction

Nucleoside analogues that lack the 3′-hydroxyl group have been extensively studied as potential drugs for the treatment of AIDS. The 5′-triphosphate forms of these compounds inhibit HIV-1 reverse transcriptase.[1] Compounds with a wide variety of modifications of the deoxyribose moiety have been synthesized and tested for activity.[2,3] Attempts at correlating structure with activity have been performed for 2′,3′-dideoxynucleosides (D2N), for 3′-substituted analogues, including 3′-azido-3′-deoxythymidine (AZD2T, AZT), for 2′,3′-didehydro-2′,3′-dideoxynucleosides[3-7] (D4N), and for several other groups of modified nucleosides. Detailed analyses of the conformational properties aimed at identifying conformational features of the nucleosides that could be correlated with activity have been reported for the former two classes[8-13] but not for the unsaturated analogues. Many D4N nucleosides are potent inhibitors of HIV-1 reverse transcriptase.[14-17] The D4N compounds are consistently more active than their respective saturated D2N analogues. Of particular interest are the thymidine (D4T) and cytidine (D4C) analogues, which are nearly as potent as AZD2T. The goal of this study is to analyze the conformational properties of the D4N nucleosides and to determine if the observed conformations and associated molecular energy profiles can be useful in understanding the biological activities. This study parallels our earlier comparisons[9,10] of saturated and 3′-substituted anti-HIV nucleosides. These studies revealed that the very active saturated compounds, including AZD2T, D2C, and D2A, frequently adopt an unusual furanose

* To whom correspondence should be addressed. Present address: Wadsworth Center for Laboratories and Research, New York State Department of Health, Empire State Plaza, P.O. Box 509, Albany, NY 12201-0509.
† Medical Foundation of Buffalo.
‡ University of Georgia.
§ Emory University and Veterans Affairs Medical Center.

(1) Mitsuya, H.; Yarchoan, R.; Kageyama, R.; Broder, S. *FASEB J.* **1991**, *5*, 2369–2381.
(2) Schinazi, R. F.; Mead, J. R.; Feorino, P. M. *AIDS Res. Hum. Retroviruses* **1992**, *8*, 963–990.
(3) Nasr, M.; Litterst, C.; McGowan, J. *J. Antiviral Res.* **1990**, *14*, 125–148.
(4) Mitsuya, H.; Yarchoan, R.; Broder, S. *Science* **1990**, *249*, 1533–1544.
(5) Chu, C. K.; Schinazi, R. F.; Arnold, B. H.; Cannon, D. L.; Doboszewski, B.; Bhadti, V. B.; Gu, Z. P. *Biochem. Pharmacol.* **1988**, *39*, 35–43.
(6) Chu, C. K.; Schinazi, R. F.; Ahn, M. K.; Ullas, G. V.; Gu, Z. P. *J. Med. Chem.* **1989**, *32*, 612–617.
(7) Balzarini, J.; Kang, G.-J.; Dalal, M.; Herdewijn, P.; De Clercq, E.; Broder, S.; Johns, D. G. *Mol. Pharmacol.* **1987**, *32*, 162–167.
(8) Van Roey, P.; Salerno, J. M.; Duax, W. L.; Chu, C. K.; Ahn, M. K.; Schinazi, R. F. *J. Am. Chem. Soc.* **1988**, *110*, 2277–2282.
(9) Van Roey, P.; Salerno, J. M.; Chu, C. K.; Schinazi, R. F. *Proc. Natl. Acad. Sci. U.S.A.* **1989**, *86*, 3929–3933.
(10) Taylor, E. W.; Van Roey, P.; Schinazi, R. F.; Chu, C. K. *Antiviral Chem. Chemother.* **1990**, *1*, 163–173.
(11) Van Roey, P.; Taylor, E. W.; Chu, C. K.; Schinazi, R. F. *Ann. N. Y. Acad. Sci.* **1990**, *616*, 29–40.
(12) Birnbaum, G. I.; Lin, T.-S.; Prusoff, W. H. *Biochem. Biophys. Res. Commun.* **1988**, *151*, 608–614.
(13) Camerman, A.; Mastropaolo, D.; Camerman, N. *Proc. Natl. Acad. Sci. U.S.A.* **1987**, *84*, 8239–8242.
(14) Lin, T.-S.; Schinazi, R. F.; Prusoff, W. H. *Biochem. Pharmacol.* **1987**, *36*, 2713–2718.
(15) Mansuri, M. M.; Starrett, J. E., Jr.; Ghazzouli, I.; Hitchcock, M. J. M; Sterzycki, R. Z.; Brankovan, V.; Lin, T.-S.; August, E. M.; Prusoff, W. H.; Sommadossi, J.-P.; Martin, J. C. *J. Med. Chem.* **1989**, *32*, 461–466.
(16) Hamamoto, Y.; Nakashima, H.; Matsui, T.; Matsuda, A.; Ueda, T.; Yamamoto, N. *Antimicrob. Agents Chemother.* **1987**, *31*, 907–910.
(17) Balzarini, J.; Pauwels, R.; Herdewijn, P.; De Clercq, E.; Cooney, D. A.; Kang, G. J.; Dalal, M.; Johns, D. G.; Broder, S. *Biochem. Biophys. Res. Commun.* **1986**, *140*, 735–742.

0002-7863/93/1515-5365$04.00/0

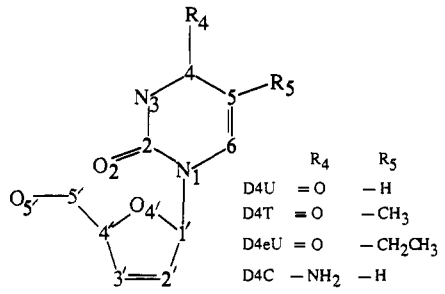

Figure 1. Schematic diagram of the structure of D4N pyrimidine nucleosides with the numbering scheme used in the text and tables. This numbering scheme and all definitions of conformational parameters are consistent with the recommendations of the IUPAC-IUB.[21,22]

ring conformation, C3′-*exo*/C4′-*endo* twist. *Endo* and *exo* refer to displacements of the atoms above and below the plane of the ring, respectively. This conformation forces C5′ into an extreme axial position. We have hypothesized that this conformation may be important in facilitating the phosphorylation of the 5′-hydroxyl, which is the first step in conversion of the nucleoside to the active 5′-triphosphate nucleotide.

In this paper we report the crystal structures of D4C, 2′,3′-didehydro-2′,3′-dideoxyuridine (D4U), and 2′,3′-didehydro-2′,3′-dideoxy-5-ethyluridine (D4eU). The crystal structures of two crystal forms of D4T[18,19] and of another form of D4C[20] were previously reported. The molecular conformations and geometries observed in all six structures are compared with those of related saturated and 3′-substituted nucleosides including D2T, D2U, AZD2T, and 3′-(propyl-2-ene)-2′,3′-dideoxyuridine (PED2U), an inactive 3′-substituted analogue. Figure 1 shows the molecular formulas and the standard nucleoside numbering scheme[21,22] which is used in the text and tables.

The crystallographic analyses are complemented by a complete analysis of the conformational energetics and rotational barriers of D4T and D2T. Energy-isocontoured γ,χ-maps based on data derived by molecular mechanics calculations are reported for both compounds to explore the differences in conformational flexibility and energy of the saturated and unsaturated analogues.

Experimental Section

X-ray Crystallography. The compounds were prepared as previously described.[23,24] The crystal data, structure determination, and refinement results are summarized in Table I. Crystals of all three compounds were obtained by slow evaporation methods. As is not unusual for nucleoside crystals, the crystals tend to be very thin needles and are not suitable for highly accurate geometric analysis but are adequate for conformational analysis based on the comparison of several compounds. All data with $F < 3.0\sigma(F)$ were considered unobserved, where $\sigma^2(F) = (k/(\mathrm{Lp}\ I))[\sigma^2(I) + (0.02I)^2]$. The structures were determined by direct methods, using the programs MULTAN[25] or SHELXS.[26] The structures were refined by full-matrix least-squares, minimizing the function $\sum w(F_o - F_c)^2$, where $w = 1/\sigma^2(F)$ for the observed data and $w = 0$ for the unobserved data. All hydrogen atoms, except for the O5′A hydrogen in the structure of D4U, were located in difference maps and refined after the anisotropic refinement of the non-hydrogen atoms had converged. Other programs used include Blessing's data reduction package,[27] the least-squares fitting program FITMOL,[28] and the plotting program ORTEP-II.[29]

Table I. Crystallographic Data for the 2′,3′-Didehydro-2′,3′-dideoxypyrimidine Structures D4U, D4eU, and D4C

	D4U	D4eU	D4C
chemical formula	$C_9H_{10}N_2O_4$	$C_{11}H_{14}N_2O_4 \cdot {}^1/_2H_2O$	$C_9H_{11}N_3O_3$
solvent	acetonitrile	ethanol	methanol
space group	$P\bar{1}$	$P\bar{1}$	$P2_12_12$
crystal size (mm)	0.1 × 0.15 × 0.4	0.1 × 0.2 × 0.25	0.08 × 0.12 × 0.80
a (Å)	5.484(1)	5.934(1)	7.047(3)
b (Å)	9.438(2)	9.897(2)	38.56(1)
c (Å)	9.733(2)	10.919(2)	7.214(3)
α (deg)	72.24(2)	71.53(1)	90
β (deg)	78.73(2)	78.27(1)	90
γ (deg)	88.03(2)	75.80(1)	90
V (Å^3)	470.3(4)	584.1(2)	1960.6(7)
Z	2	2	8
D_c (g cm^{-3})	1.482	1.404	1.417
radiation/filter	Mo/Nb	Mo/Nb	Cu/Ni
crystal temp	ambient	160 K	ambient
resolution	$2 < 2\theta < 55$	$2 < 2\theta < 55$	$3 < 2\theta < 154$
total data	4027	3407	3060
unique data	2763	2703	2386
structure det	SHELXS	MULTAN	SHELXS
data > $3\sigma(F)$	2292	2094	2168
R	0.064	0.069	0.050
R_w	0.041	0.046	0.064
R_{all}	0.078	0.091	0.053
S	2.213	2.180	1.583
final diff map			
minimum (e Å^{-3})	−0.375	−0.620	−0.516
maximum	0.430	0.638	0.376

Molecular Mechanics Calculations. The nucleoside analogues were modeled using the SYBYL program (Version 5.4; Tripos Associates, St. Louis, MO). The conformations were first adjusted to fully minimize the molecular energies without taking electrostatic terms into account. Partial atomic charges for the Coulombic potential term in molecular mechanics were then determined by the AM1 semiempirical quantum mechanics method. The TRIPOS force field[30] was used for the molecular mechanics calculations, unmodified except for the addition of a *gauche* potential energy parameter (in this case uniquely required for the O4′–C4′–C5′–O5′ torsion).[31] This *gauche* potential has been shown to be necessary for accurate modeling of the conformation energetics of the furanose ring in nucleosides[32,33] and to contribute significantly to the conformational trend for saturated anti-HIV nucleosides described previously.[8,9]

For both D4T and D2T, a total of 1296 (36^2) conformations were generated and fully energy minimized while constraining γ and χ to the desired combinations of values, systematically sampling γ,χ torsional space. This was accomplished using GRID search, with γ and χ both being driven through a full 360° range, with 10-deg increments. The criterion for termination of each minimization was an energy change of less than 0.0001 kcal/mol. A distance-dependent dielectric constant equal to $4r$ was used, where r is the distance in Å between the two centers in question. This choice has been shown to yield results comparable to

(18) Gurskaya, G. B. Bochkarev, A. V.; Zhdanov, A. S.; Dyatkina, N. B.; Krayevsky, A. A. *Int. J. Purine Pyrimidine Res.* **1991**, *2*, 55–60.
(19) Harte, W. E., Jr.; Starrett, J. E., Jr.; Martin, J. C.; Mansuri, M. M. *Biochem. Biophys. Res. Commun.* **1991**, *175*, 298–304.
(20) Birnbaum, G. I.; Giziewicz, J.; Lin, T.-S.; Prusoff, W. H. *Nucleosides Nucleotides* **1989**, *8*, 1259–1269.
(21) IUPAC-IUB Commission on Biochemical Nomenclature (CBN). *Eur. J. Biochem.* **1970**, *15*, 203–208.
(22) IUPAC-IUB Joint Commission on Biochemical Nomenclature (CBN). *Eur. J. Biochem.* **1983**, *131*, 9–15.
(23) Chu, C. K.; Bhadti, V. S.; Doboszewski, B.; Gu, Z. P.; Kosugi, Y.; Pullaiah, K. C.; Van Roey, P. *J. Org. Chem.* **1989**, *54*, 2217–2225.
(24) Cosford, N. D. P.; Schinazi, R. F. *J. Org. Chem.* **1991**, *56*, 2161–2165.
(25) Germain, G.; Main, P.; Woolfson, M. M. *Acta Crystallogr., Sect. A: Cryst. Phys., Diffr., Theor. Gen. Crystallogr.* **1971**, *A27*, 368–376.
(26) Sheldrick, G. M. SHELX84; University of Gottingen: Gottingen, F.R.G., 1984.
(27) Blessing, R. H. *Crystallogr. Rev.* **1987**, *1*, 3–57.
(28) Smith, G. D. Personal communication.
(29) Johnson, C. K. ORTEP-II. Report ORNL-5138; Oak Ridge National Laboratory: Oak Ridge, TN, 1976.
(30) Clark, N. M.; Cramer, R. D., III; Van Opdenbosch, N. *J. Comput. Chem.* **1989**, *10*, 982–1012.
(31) We did *not* use AM1 to optimize the geometry of the five-membered ring. Like other investigators,[36] we have found that this produces an unrealistic, virtually flat furanose ring. This can lead to absurd conclusions, such as a recently published AM1 study of AZT (Sabio, M.; Topiol, S. *J. Comput. Chem.* **1992**, *40*, 478–491) for which most of the low-energy conformers were in the $\chi \sim -270°$ quadrant of pseudorotational space, which is well-known to be precisely the least favored region due to the axial positioning of the base and the 5′-carbinol. With a flat ring, no substituent can be axial. As a result one of the most important determinants of the molecular conformations of nucleosides with puckered rings was ignored in this study.
(32) Olson, W. K. *J. Am. Chem. Soc.* **1982**, *104*, 278–286.
(33) Pearlman, D.; Kim, S.-H. *J. Biomol. Struct. Dyn.* **1985**, *3*, 99–125.

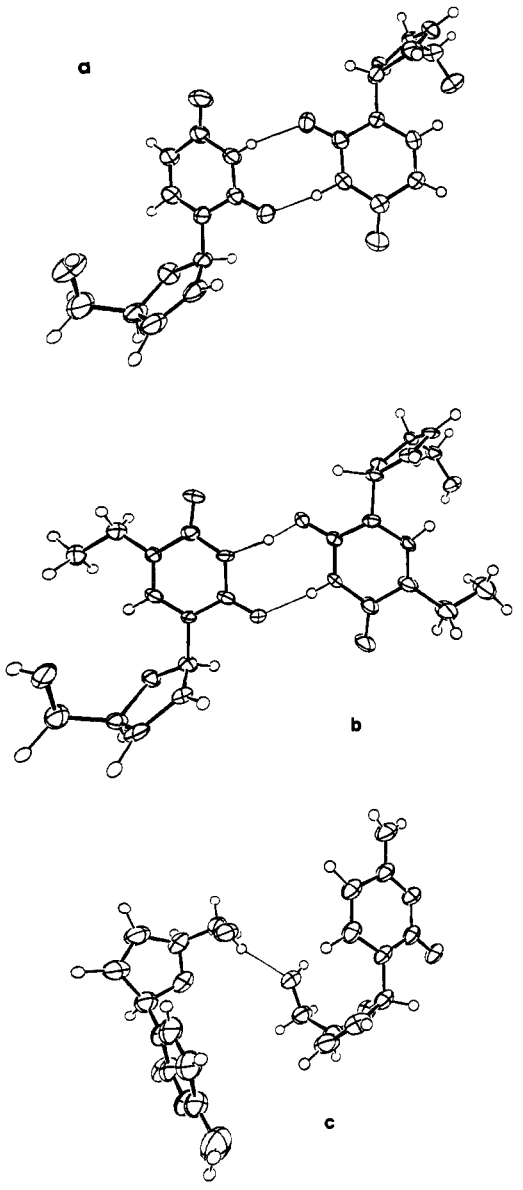

Figure 2. Molecular conformations of the pairs of unique molecules in the structures of (a) D4U, (b) D4eU, and (c) D4C. The base pairing hydrogen bonding between independent molecules observed in the structures of D4U and D4eU is shown. Thermal ellipsoids are shown at 50% probability.

much more computationally intensive explicit solutions of the Poisson equation for molecules in a hydrated crystal environment.[34] This is desirable in this case, since we are comparing the theoretical results to experimental crystallographic data. In addition, the environment of the solvated or enzyme-bound drug is certainly closer to that of a hydrated crystal than to a vacuum. The furanose ring conformation was allowed to vary dynamically as each conformational search proceeded. The searches took an average of 5 h of CPU time on a Silicon Graphics IRIS 4D/240.

The resultant databases of energy-minimized nucleoside conformations were analyzed using the TABLE and GRAPH options of SYBYL to generate energy-contoured γ,χ-maps and graphs of various torsional barriers for the two analogues.

Results

The molecular conformations of nucleosides have been extensively reviewed, primarily by Saenger.[35] Three parameters describe the most important features of the conformation of a nucleoside molecule: the geometry of the glycosylic link (torsion angle χ, C2–N1–C1'–O4'), the furanose ring puckering (pseudorotation angle P), and the orientation of the 5'-hydroxyl group (torsion angle γ, C3'–C4'–C5'–O5').

As part of ongoing crystallographic studies of potential anti-HIV nucleosides, the structures of D4U, D4T, D4eU, and D4C were determined and analyzed. While this paper was in preparation, we realized that this structure of D4T was published. Gurskaya *et al.*[18] were able to obtain better crystals, and as a result their structure determination was more accurate. Therefore, the results of their study are used in this analysis although all the structural features observed in both structure determinations are identical to within experimental error. The structures of D4T[19] and D4C[20] in different crystal forms were also previously reported and are included in this analysis.

All six structures have two nucleoside molecules in the asymmetric unit. Therefore, twelve independent observations of the molecular conformations of D4N pyrimidine nucleosides are available for analysis. Figure 2 shows the molecular conformations of both molecules of the three not previously reported structures. Table II lists the most important conformational parameters of all six structures. Independent molecules are distinguished by capital letters A, B, C, and D, where C and D refer to the molecules in the alternate crystal forms of D4T and of D4C. Three distinct conformations of the glycosylic link (torsion angle χ, C2–N1–C1'–O4') are observed: $\chi \sim -170°$, $\chi \sim -120°$, and $\chi \sim -90°$. The latter conformation, called high-*anti*, is very unusual for pyrimidine nucleosides. The position of the 5'-hydroxyl group is *+sc* in all but one molecule. The furanose rings are nearly planar with a slight trend to a O4'-*endo* or O4'-*exo* conformation. Table III lists the root mean square deviations from planarity for the five-membered rings and the deviations of O4' from the least-squares plane through C1', C2', C3', and C4' for the three new structures. D4U has one molecule with a planar ring and one that is O4'-*endo*. In D4eU and D4C, one molecule is slightly O4'-*endo* while the other is slightly O4'-*exo*.

Table II. Important Geometric and Conformational Parameters of the Twelve Conformations of 2',3'-Didehydro-2',3'-dideoxypyrimidines[a]

compound	conformation	N1–C1' (Å)	C1'–O4' (Å)	χ (deg)	γ (deg)
D4U	A	1.472(7)	1.415(6)	−96.1(5)	53.2(6)
	B	1.499(6)	1.392(6)	−178.0(4)	54.1(6)
D4T[18]	A	1.505(3)	1.386(4)	−172.6	54.1
	B	1.487(4)	1.423(4)	−85.1	55.6
D4T[19]	C	1.477	1.400	−118.0	60.6
	D	1.502	1.401	−174.0	53.8
D4eU	A	1.48(1)	1.45(1)	−98.6(8)	48.(1)
	B	1.50(1)	1.41(1)	−168.4(6)	56.(1)
D4C	A	1.476(6)	1.427(4)	−82.7(3)	55.1(5)
	B	1.448(5)	1.435(5)	−87.8(4)	47.3(5)
D4C[20]	C	1.468(4)	1.404(4)	−118.7(4)	49.8(4)
	D	1.499(4)	1.401(4)	−160.2(4)	165.0(4)
AZD2T[8]	A	1.460(4)	1.432(5)	−124.4(3)	50.9(4)
	B	1.502(4)	1.395(4)	−173.6(3)	173.4(3)

[a] Parameters for AZD2T are shown for comparison. Standard deviations, where available, are included in parenthesis.

Table III. Deviations (Å) from Least-Squares Planes through the Sugar Rings of the Three New Structures[a]

compound	conformation	rms deviation 5-atom plane	dist O4' from plane C1',C2',C3',C4'
D4U	A	0.027	−0.094
	B	0.004	0.012
D4eU	A	0.011	−0.039
	B	0.022	0.074
D4C	A	0.017	−0.057
	B	0.023	0.073

(34) Pickersgill, R. W. *Protein Eng.* **1988**, *2*, 247–248.
(35) Saenger, W. *Principles of Nucleic Acid Structure*; Springer-Verlag: New York, 1984.

Table IV. Intermolecular Hydrogen-Bonding Geometry of 2′,3′-Didehydro-2′,3′-dideoxypyrimidines

donor···acceptor	D···A (Å)	D–H (Å)	H···A (Å)	D–H···A (Å)	(acceptor symmetry)
			D4U		
N3A–HN3A···O2B	2.877(6)	1.01(4)	1.87(4)	172(3)	(x, y, z)
O5′A–HO5′A···O5′B	3.119(8)				$(-2.0 + x, -1.0 + y, z)$
N3B–HN3B···O2A	2.895(5)	0.82(3)	2.09(4)	170(3)	(x, y, z)
O5′B–HO5′B···O5′A	2.964(8)	0.74(6)	2.24(6)	168(5)	$(1.0 + x, 1.0 + y, -1.0 + z)$
			D4eU		
N3A–HN3A···O2B	2.86(1)	1.14(8)	1.77(8)	165(6)	(x, y, z)
O5′A–HO5′A···O1C	2.64(1)	0.81(8)	1.85(8)	164(5)	$(x, 1.0 + y, -1.0 + z)$
N3B–HN3B···O2A	2.80(1)	0.88(8)	1.92(5)	173(4)	(x, y, z)
O5′B–HO5′B···O5′A	2.75(1)	0.96(8)	1.80(8)	167(6)	$(1.0 + x, -1.0 + y, 1.0 + z)$
O1C–H1CA···O5′B	2.71(1)	0.98(7)	1.76(6)	162(4)	$(-2.0 + x, y, z)$
O1C–H1CB···O4′B	2.75(1)	1.15(7)	1.69(7)	152(5)	$(-1.0 + x, y, z)$
			D4C		
N4A–H4AA···O2A	2.837(4)	0.78(3)	2.09(3)	166(3)	$(x, y, 1.0 + z)$
N4A–H4AB···O5′B	3.015(4)	0.83(4)	2.21(4)	164(3)	$(-0.5 + x, 1.5 - y, 2.0 - z)$
O5′A–H5′A···N3A	2.830(4)	0.78(4)	2.06(4)	168(3)	$(1.0 - x, 2.0 - y, z)$
N4B–H4BA···O2B	2.967(4)	1.18(4)	1.83(4)	160(3)	$(1.0 + x, y, z)$
N4B–H4BB···N3B	3.191(5)	1.19(5)	2.44(5)	119(3)	$(0.5 + x, 1.5 - y, 3.0 - z)$
O5′B–H5′B···O5′A	2.742(5)	0.74(5)	1.99(5)	165(4)	$(1.5 + x, -0.5 + y, 2.0 - z)$

Table IV lists the intermolecular hydrogen bonds of the new structures. D4U is isostructural with one form of D4T,[18] having similar cell dimensions and molecular geometries and identical intermolecular hydrogen-bonding patterns. The crystal structure of D4eU is also related except for the insertion of a water molecule which forms hydrogen bonds with two O5′-hydroxyl groups and one O4′-ether. These structures have symmetric N3···O2 base–pairing hydrogen bonding between the independent molecules in the asymmetric unit and cross-links between O5′-hydroxyl groups, interrupted by the water molecule in D4eU. The intermolecular interactions in the new structure of D4C do not involve base pairing or stacking. All contacts between the independent molecules in this structure consist of hydrogen bonds involving the O5′-hydroxyl groups and one amino group. The remaining amino hydrogens are involved in hydrogen bonds to carbonyl groups of symmetry-related molecules.

The calculated energy-contoured γ,χ-maps for D4T and D2T are shown in Figure 3. The D4T map clearly shows that the four conformations[18,19] observed in the crystal structures are all located within the same energy well, clustered around the computed global minimum at *+sc,anti*. All are inside the 1.5 kcal/mol contour. As expected, both maps show three minima for γ, with *ap* and *+sc* essentially isoenergetic and *–sc* about 0.6 kcal/mol higher, due to the *gauche* effect. This is also clearly seen in Figure 4, which compares the torsional barriers for interconversion of the three O5′ rotamers in D4T and D2T.

Several striking features are observed when comparing the γ,χ-maps for D4T and D2T. A much greater proportion of γ,χ-space is available to D4T, as measured by the percentage of the map lying within the 2 kcal/mol contour: 20% as compared to 12% for D2T. The increased flexibility of the D4 compound is also evidenced by the fact that the entire γ,χ-map for D4T is "flatter" than that for D2T; the peaks are generally several kcal/mol lower for the former. A clear result of this increased flexibility is the greater accessibility of the high-*anti* regions of the map for the D4N analogue. This is in complete agreement with the trend toward the unusual high-*anti* in the solid state for the set of D4N pyrimidine nucleosides reported here.

Regarding the pseudorotational state of the furanose ring, the set of D4T conformations used to calculate Figure 3a consisted almost entirely of variants of the O4′-*endo* conformation, with only moderate ring puckering amplitude. This is also highly consistent with the crystallographically observed conformations. The γ,χ-map for D2T is considerably more convoluted than the D4T map. This is in part a reflection of the fact that an entire spectrum of ring puckering conformations involving C3′-*endo*/C2′*exo*, C3′-*exo*/C2′-*endo* twist and envelope conformations was observed in the set of D2T conformers used to calculate the γ,χ-map.

The lower barrier to rotation around the C4′–C5′ bond in D4T relative to D2T is illustrated in Figure 4. This torsional barrier graph shows that interconversion between the three O5′ orientations *ap*, *+sc*, and *–sc* is significantly enhanced in the D4N compounds, particularly between *+sc* and *–sc*, which is along the preferred pathway for interconversion between the two lowest energy orientations, *+sc* and *ap*.[33] The barrier to rotation between *ap* and *–sc* is about 0.5 kcal/mol lower for D4T, and that between *–sc* and *+sc* is about 1 kcal/mol lower. This difference is primarily caused by the reduced van der Waals interactions that result from the lack of 2′- and 3′-"up" substituents in the D4N analogue. In contrast, the barrier to direct interconversion between *+sc* and *ap* (without passing through *–sc*) is about the same in both compounds, which is not surprising, since this highest barrier involves interactions between O5′ and the pyrimidine base, rather than between O5′ and C3′ or C2′ hydrogens.

Similarly, rotations of the base in D4T are facilitated relative to those in D2T primarily by the lack of a C2′-"up" hydrogen atom. Examination of the contour levels between equivalent *syn*- and *anti*-states in Figure 3a and b shows that the barrier to their interconversion is generally about 1 kcal/mol lower in D4T as compared to D2T.

Discussion

The twelve independent observations of the molecular conformations of the D4N pyrimidine nucleosides included in this study reveal the following trends. The high-*anti* glycosylic link geometry ($\chi \sim -90°$) is observed in five of the twelve molecules. This conformation is unusual for pyrimidines but is not uncommon for purines. This conformation places the C1′–C2′ in the plane of the base with C2′ eclipsed with C6 of pyrimidine and C8 of purine nucleosides. The steric interaction between C2′ and C6 in the pyrimidine (six-membered ring) is much greater than that between C2′ and C8 in the purine (five-membered ring) nucleoside. The hydrogen atom on the pyrimidine C6 points down to the C2′ hydrogen atom of the furanose ring while the purine base does not have a substituent that points down. In the D4N pyrimidine nucleosides, the high-*anti* conformation is allowed because there is no hydrogen atom at C2′ that points up toward the base and steric hindrance with C6 is greatly reduced. One can conclude that reduction in the steric interactions, whether this is a result of the less bulky base or less bulky furanose, stabilizes the high-*anti* conformation of nucleosides.

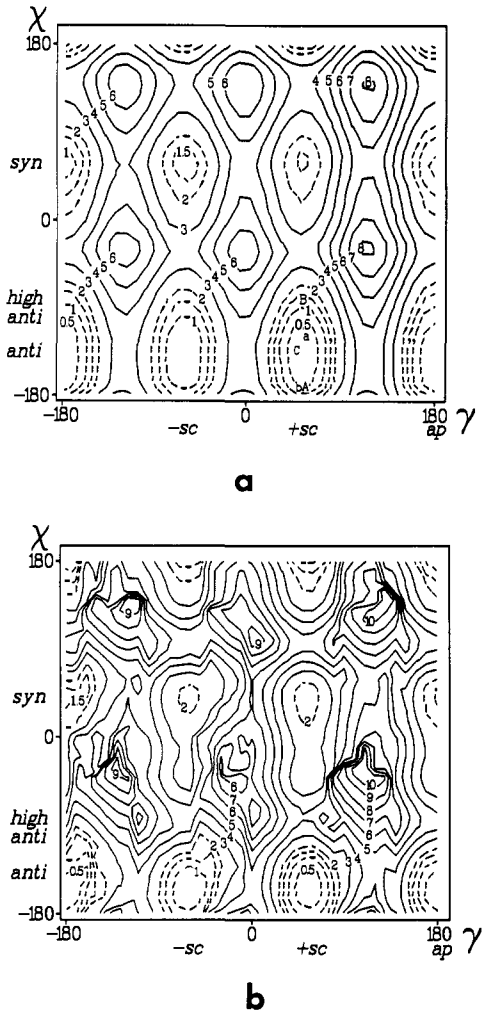

Figure 3. Energy-isocontoured γ,χ-maps for (a) D4T and (b) D2T, showing the dependence of relative enthalpy (the contoured variable) upon variations in γ and χ. Energy minima are highlighted by dotted contours up to 2 kcal/mol, at 0.5 kcal/mol intervals. Solid lines represent higher energy states from 3 to 10 kcal/mol above the global minimum, contoured at 1 kcal/mol intervals. Note that, as has been described previously for conventional nucleosides,[33] equivalent *syn* conformations are about 1.0 kcal/mol higher than *anti* and that *−sc* is about 0.6 kcal/mol higher than *+sc* and *ap*. The most striking feature of the map for D4T (a) when compared to that of D2T (b) is its "flatness", indicated by lower peaks and saddles and by a greater percentage of the map lying within the 2 kcal/mol contour. The high-*anti* regions of D4T are easily accessible within 1 or 1.5 kcal/mol from the minimum while they are at least 2 kcal/mol above the minimum in D2T. These results suggest an unusual degree of molecular flexibility for a nucleoside analogue. Solid-state D4T conformations are indicated. A and B refer to the two conformations in the structure of Gurskaya *et al.*,[18] and a and b are the conformations reported by Harte *et al.*[19] These are all in the same energy well, clustered around the global minimum, indicated by C, at *+sc,anti*. The solid-state D2T conformers reported previously by Van Roey *et al.*[9] are not indicated but lie in the two energy wells at *+sc,anti* and *ap,anti*.

Comparison of the conformations of the furanose rings of all twelve molecules with those of the saturated and 3'-substituted analogues has been performed by least-squares fitting of the central four atoms, N1, C1', C2', and O4', of these compounds with the corresponding atoms in AZD2T, D2T, D2U, and PED2U. The results of representative examples of these fits are listed in Table V and illustrated in Figure 5. For this comparison, the saturated pyrimidine nucleosides can be classified in three groups. Two groups have "standard" furanose ring conformations in which C2' and C3' are displaced out of the plane of the ring: C3'-*exo*/C2'-*endo* ($P \sim 170°$) and C3'-*endo*/C2'-*exo* ($P \sim 10°$). The first conformation has been observed in molecule A of AZD2T

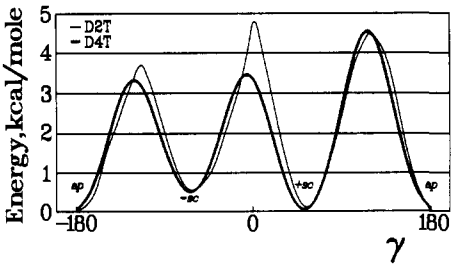

Figure 4. Torsional barrier to rotation about γ (C3'–C4'–C5'–O5') in D4T as compared to D2T. Interconversion between the three rotamers is relatively facilitated in D4T, primarily due to the substantially lower barrier between *+sc* and *−sc*. This difference originates from the lack of 2'- and 3'-"up" substituents in the D4 compound.

Table V. Least-Squares Fit of D4N Pyrimidine Nucleosides with AZD2T, D2T, and PED2U[a]

compound	molecule conformation	rms dist (refined)	max dist (refined)	C3'	C4'	C5'	O5'
	AZD2T[8] Molecule A ($\chi = -124.4°$, $P = 174.9°$)						
D4C	A	0.036	0.043	0.77	0.38	0.50	1.01
	B	0.026	0.035	0.55	0.15	0.21	0.56
D4C[20]	C	0.023	0.032	0.58	0.14	0.30	0.72
	D	0.029	0.034	0.76	0.44	0.65	2.36
D4T[18]	A	0.037	0.048	0.66	0.22	0.28	0.81
	B	0.026	0.035	0.59	0.16	0.28	0.78
D4T[19]	C	0.079	0.115	0.53	0.14	0.68	1.25
	D	0.031	0.036	0.70	0.27	0.32	0.81
D4U	A	0.026	0.031	0.58	0.14	0.28	0.74
	B	0.032	0.034	0.72	0.31	0.37	0.84
D4eU	A	0.029	0.038	0.67	0.26	0.31	0.85
	B	0.040	0.054	0.58	0.20	0.23	0.59
	AZD2T[8] Molecule B ($\chi = -173°$, $P = 215.3°$)						
D4C	A	0.035	0.049	0.42	0.18	0.80	1.19
	B	0.037	0.048	0.24	0.43	1.12	1.45
D4C	C	0.028	0.042	0.26	0.45	1.24	1.30
	D	0.024	0.043	0.41	0.13	0.54	0.82
D4T	A	0.028	0.056	0.30	0.36	1.12	1.18
	B	0.026	0.043	0.27	0.42	1.18	1.20
D4T	C	0.080	0.120	0.20	0.63	1.56	0.85
	D	0.026	0.045	0.36	0.30	1.01	1.21
D4U	A	0.027	0.037	0.26	0.44	1.24	1.27
	B	0.028	0.042	0.38	0.26	0.91	1.24
D4eU	A	0.026	0.042	0.33	0.33	1.07	1.15
	B	0.046	0.062	0.26	0.42	1.15	1.42
	PEDDU[9] ($\chi = -161.6°$, $P = 16°$)						
D4C	A	0.027	0.034	0.51	0.05	0.53	3.04
	B	0.027	0.043	0.71	0.22	0.41	2.89
D4C	C	0.020	0.028	0.69	0.24	0.35	2.71
	D	0.018	0.028	0.52	0.10	0.79	1.38
D4T	A	0.022	0.040	0.61	0.15	0.35	2.76
	B	0.020	0.028	0.67	0.21	0.34	2.73
D4T	C	0.075	0.107	0.73	0.44	0.38	2.18
	D	0.020	0.030	0.57	0.10	0.39	2.88

[a] Atoms fitted are N1, C1', O4', and C2'. Distances shown are the root mean square and maximum deviations for the fitted atoms and the deviations for remaining atoms of the ribose moiety (in Å).

and molecule A of D2T. The second conformation has been observed for the second molecule of D2T, D2U, and PED2U, which are all compounds that happen to have weak anti-HIV activity. The third conformation, C3'-*exo*/C4'-*endo* ($P \sim 210°$), is much less commonly observed in nucleosides in general but occurs frequently in the structures of the very potent anti-HIV compounds, namely, the second molecule in AZD2T, AZD2U, D2C, and D2A. Two distinct features can be identified in these comparisons: endocyclic differences, specifically the position of C3', and exocyclic differences, the position of C5'. At C3', the first and second class of saturated compounds differ somewhat more (~0.6 Å) from the unsaturated compounds than the third class (~0.3 Å). The distortions from the planar ring of classes one and two are essentially pure envelope puckering displacements

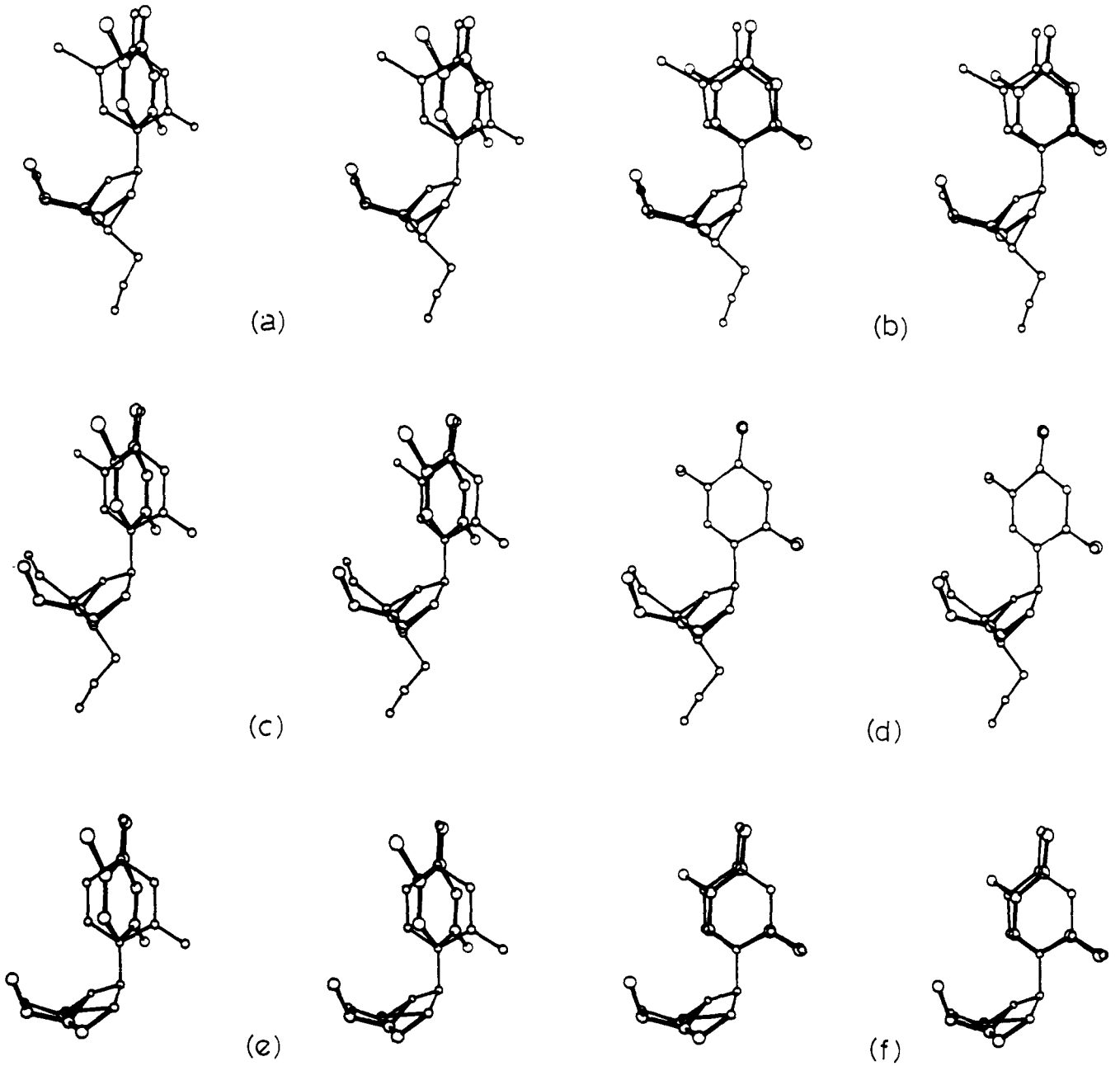

Figure 5. Stereo diagrams showing the superpositions of the $\chi = -90°$ and $\chi = -170°$ molecules of D4T (bold) on representative saturated 2′,3′-dideoxypyrimidines (fine line), produced by least-squares fitting of the central four atoms (N1, C1′, C2′, and O4′): (a) D4T B/AZD2T A; (b) D4T A/AZD2T A; (c) D4T B/AZD2T B; (d) D4T A/AZD2T B; (e) D4T B/D2U; (f) D4T A/D2U. The main differences in the conformations of the furanose rings are at C3′ in parts a, b, e, and f and at C5′ in parts c and d. The high-*anti* glycosylic link conformation of D4T B (a, c, e) is not observed for the saturated analogues. Note that the orientation of the base of D4T B (left column) does not match with those of any other compounds (high-*anti* conformation).

of C3′ out of the plane of the ring, with the difference between the two classes being limited to the puckering direction. In these conformations the deviations in the position of C5′ are relatively small (~0.3 Å). The third class shows a smaller deviation at C3′ (~0.3 Å) but a much larger deviation at C5′ (~1.0 Å). The twist involving C3′ and C4′ yields reduced deviation from the unsaturated conformation but moves C5′ to an extreme axial position above the ring and closer to the base. This greatly affects the location of the O5′-hydroxyl group relative to the position of the base.[11] Overall the furanose conformation of the unsaturated compounds does not match any one of the three classes of saturated compounds. The near planar ring does not allow C5′ to be axial. This constraint on the position of C5′ in the unsaturated compounds and the resulting effect on the position of O5′ are not consistent with our previous hypothesis of the need for a C5′ position close to the base as a requirement for activity.[9,11]

The results of the molecular mechanics calculations are consistent with the crystallographic results on the accessibility of the high-*anti* conformation for the glycosylic link. The conformers observed in the solid state for D4T, which is typical for this set of compounds, are clustered around the computed global energy minimum (Figure 3a). The calculations also rationalize the observed tendency toward the unusual high-*anti* conformations in terms of reduced torsional barriers in D4T and increased molecular flexibility relative to the saturated analogue.

The O5′ rotamers of all crystallographically observed conformations of the D4 pyrimidines are in the +*sc*,*anti* state, with the exception of one molecule of the previously reported crystal form of D4C.[20] The molecular mechanics calculations suggest that, for an isolated molecule, the energy well for the *ap*,*anti* state is of approximately equal depth and size. This apparent inconsistency can be explained by the fact that the orientation

of the O5' rotamer in the crystal structure is largely determined by the orientation in which O5' can form an energetically favorable intermolecular hydrogen bond. The *ap,anti* state is observed in other nucleoside analogues, including D4A.[23]

There is some degree of conformational similarity between the D4 pyrimidine nucleosides and their analogues with an *exo*-oriented three-membered ring fused to the 2',3'-ribose bond. Conformational analysis of these compounds[36] shows that they also have a fairly flat sugar ring and a preferred glycosylic angle range ($\chi \sim -105°$) that is near to the high-*anti* region. This structural modification reduces the steric interaction between the base and the 2'-"up" hydrogen by fixing it in an equatorial orientation, very similar to the effect of the 2',3'-double bond.

It is of considerable interest that another potent class of anti-HIV nucleoside analogues, those in which the 3'-methylene unit has been replaced by a heteroatom (either sulfur or oxygen), should also be expected to have reduced barriers to interconversion of the O5' rotamers, which again, as for the D4 compounds, is due to the lack of 3'-"up" substituents. Conformational analysis of these compounds is presently in progress.[37-39]

(36) Koole, L. H.; Neidle, S.; Crawford, M. D.; Krayevski, A. A.; Gurskaya, G. V.; Sandstroem, A.; Wu, J.-C.; Tong, W.; Chattopadhyaya, J. *J. Org. Chem.* **1991**, *56*, 6884–6892.

(37) Norbeck, D. W.; Spanton, S.; Broder, S.; Mitsuya, H. *Tetrahedron Lett.* **1989**, *30*, 6263–6269.

(38) Chu, C. K.; Ahn, S. K.; Kim, H. O.; Beach, J. W.; Alves, A. J.; Jeong, L. S.; Islam, Q.; Van Roey, P.; Schinazi, R. F. *Tetrahedron Lett.* **1991**, *32*, 3791–3794.

Conclusion

The conformations of the furanose rings of pyrimidine nucleosides can be divided into four distinct classes, three for saturated nucleosides and one for the 2',3'-unsaturated compounds. None of these conformations can be designated as being the "active site" conformation for phosphorylation at O5'. However, the greater flexibility of the unsaturated compounds with corresponding higher activity when compared to the saturated analogues and the greater tendency of the highly active compounds to adopt unusual conformations, C3'-*exo*/C4'-*endo* furanose ring puckering and high-*anti* glycosylic link geometry for compounds with saturated and unsaturated furanose rings, respectively, may indicate the need to be able to adopt an unusual conformation to facilitate enzymatic phosphorylation.

Acknowledgment. This research was supported in part by the Father Vincent Grant for AIDS Research (P.V.R.) and USPHS Grant AI-30392001 (E.W.T.) and by the Department of Veterans Affairs (R.F.S.).

Supplementary Material Available: Tables of atomic coordinates and thermal parameters (6 pages); tables of observed and calculated structure factors for D4U, D4eU, and D4C (53 pages). Ordering information is given on any current masthead page. A SYBYL molecular database containing the files obtained by the GRID search may be requested from E.W.T.

(39) Kim, H. O.; Ahn, S. K.; Alves, A. J.; Beach, J. W.; Jeong, L. S.; Islam, Q.; Van Roey, P.; Schinazi, R. F. *J. Med. Chem.* **1992**, *35*, 1987–1995.