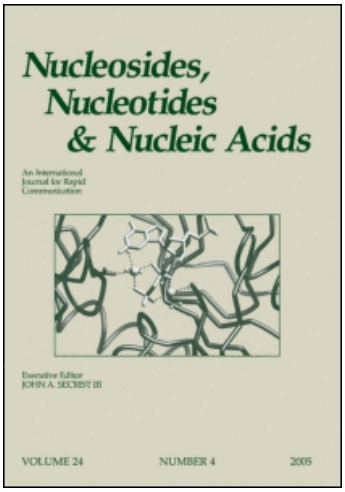


Nucleosides, Nucleotides and Nucleic Acids
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597286

Definitive Solution Structures for the 6-Formylated Versions of 1-(βD-Ribofuranosyl)-, 1-(2'-Deoxy-β-D-Ribofuranosyl)-, and 1-β-D-Arabinofuranosyluracil, and of Thymidine

Michael P. Groziak[a]; Ronghui Lin[a]; William C. Stevens[b]; Linda L. Wotring[c]; Leroy B. Townsend[c]; J. Balzarini[d]; M. Witvrouw[d]; E. De Clercq[d]
[a] Department of Chemistry and Biochemistry, Southern Illinois University, Carbondale, IL, USA [b] Nuclear Magnetic Resonance Facility, Southern Illinois University, Carbondale, IL, USA [c] College of Pharmacy, The University of Michigan, Ann Arbor, USA [d] Rega Instituut, Katholieke Universiteit Leuven, Leuven, Belgium

To cite this Article Groziak, Michael P. , Lin, Ronghui , Stevens, William C. , Wotring, Linda L. , Townsend, Leroy B. , Balzarini, J. , Witvrouw, M. and De Clercq, E.(1996) 'Definitive Solution Structures for the 6-Formylated Versions of 1-(βD-Ribofuranosyl)-, 1-(2'-Deoxy-β-D-Ribofuranosyl)-, and 1-β-D-Arabinofuranosyluracil, and of Thymidine', Nucleosides, Nucleotides and Nucleic Acids, 15: 5, 1041 — 1057
To link to this Article: DOI: 10.1080/07328319608002033
URL: http://dx.doi.org/10.1080/07328319608002033

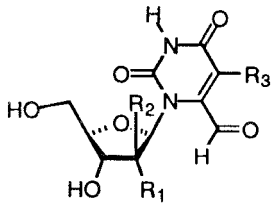
NUCLEOSIDES & NUCLEOTIDES, 15(5), 1041-1057 (1996)

DEFINITIVE SOLUTION STRUCTURES FOR THE 6-FORMYLATED VERSIONS OF 1-(β-D-RIBOFURANOSYL)-, 1-(2'-DEOXY-β-D-RIBOFURANOSYL)-, AND 1-β-D-ARABINOFURANOSYLURACIL, AND OF THYMIDINE

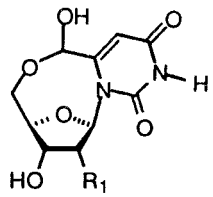
Michael P. Groziak,*[a] Ronghui Lin,[a] William C. Stevens,[b]
Linda L. Wotring,[c] Leroy B. Townsend,[c] J. Balzarini,[d] M. Witvrouw,[d] and E. De Clercq[d]
[a]Department of Chemistry and Biochemistry and [b]Nuclear Magnetic Resonance Facility,
Southern Illinois University, Carbondale, IL 62901-4409 USA
[c]College of Pharmacy, The University of Michigan, Ann Arbor, 48109-1065 USA
[d]Rega Instituut, Katholieke Universiteit Leuven, B-3000 Leuven, Belgium

Abstract: ROESY and NOESY NMR spectroscopic analyses of the ribofuranosyl (**1a**), 2'-deoxyribofuranosyl (**1b**), and arabinofuranosyl (**1c**) derivatives of 6-formyluracil in $(CD_3)_2SO$ and D_2O solutions have established that each exclusive 7,O5'-cyclic hemiacetal diastereomer of **1a,b** and the major 7,O2'-cyclic hemiacetal diastereomer of **1c** possess the 7*R* configuration. In addition, (7*R*)-**1c** has been shown to be thermodynamically more stable than (7*S*)-**1c**, contrary to our previous indication. A new, higher yielding synthetic route to **1a** has been developed, **1b** has been obtained for the first time in crystalline form, the route to **1c** has been modified to better accommodate large scale preparations, and a new, fourth member of this class, 6-formylthymidine (**1d**), has been synthesized and its solution structures in $(CD_3)_2SO$, D_2O, and CD_3OD have been determined. Antitumor and antiviral evaluations of **1a-c** have revealed no significant levels of activity.

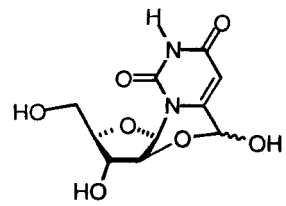
We recently reported the synthesis of the first three fully unprotected members of a new class of pyrimidine nucleosides possessing a formyl group at the pyrimidine C6 position.[1] The carboxaldehyde functionality in the 1-β-D-ribofuranosyl, 1-(2'-deoxy-β-D-ribofuranosyl), and 1-(β-D-arabinofuranosyl) derivatives of 6-formyluracil (**1a-c**, respectively) is so highly electrophilic that each one of these nucleosides exhibits a strong tendency towards existing in rare cyclic hemiacetal form, both in protic and aprotic solution as well as in the solid state. As the search for new antiviral and antitumor agents has been expanding recently to include spiro- and cyclonucleosides,[2-4] we felt it important to evaluate the potential for bioactivity in these materials and at the same time, try to improve upon their



1a, R_1 = OH, R_2 = R_3 = H
b, R_1 = R_2 = R_3 = H
c, R_1 = R_3 = H, R_2 = O H
d, R_1 = R_2 = H, R_3 = Me



One 7,O5'-cyclized diastereomer for **1a** (91% in $(CD_3)_2SO$, 33% in D_2O) **1b** (98% in $(CD_3)_2SO$, 75% in D_2O)



7,O2'-cyclized diastereomers for **1c** (100% in $(CD_3)_2SO$ and D_2O)

syntheses, and use high-field NMR techniques to better define certain configurational and conformational aspects of the unusual transglycosidic hemiacetal-bridged solution structures[4] exhibited by **1a-c**. In addition, we wished to prepare a new member of this new class of nucleosides, 6-formylthymidine (**1d**), in order to examine its array of solution structures to assess the effect of the presence of a substituent at the pyrimidine C5 position.

EXPERIMENTAL SECTION

Materials and Methods. Melting points were determined on a Thomas-Hoover UniMelt capillary apparatus and are uncorrected. Radial preparative-layer chromatography was performed on a Chromatotron instrument using Merck silica gel-60 PF254 as adsorbent. Flash column chromatography was performed with 230-400 mesh ASTM Merck silica gel-60. TLC analyses were performed on Analtech 250 μ silica gel GF Uniplates. TMEDA, *n*-BuLi in hexanes, *t*-BuLi in pentane, 1.0 M TBAF in THF, iPr_2NH, HCO_2Et, $CH_2(OMe)_2$, and TfOH were purchased from the Aldrich. The *n*-BuLi was titrated by the modified Watson-Eastham method. THF was dried by distillation from Na-benzophenone ketyl under Ar. The HCO_2Et was dried over and distilled from P_2O_5 under Ar. Mass spectral and elemental microanalyses were obtained from the University of Illinois.

NMR Methods. ^1H and ^{13}C NMR spectra were recorded at 300 and 75 MHz, respectively, on a Varian VXR-300, or at 500 and 125 MHz, respectively, on a VXR-500 instrument. 2D ^1H-^1H Homonuclear shift correlation (COSY) and 2D ^1H-^{13}C heteronuclear shift correlation (HETCOR) NMR spectra were obtained on the VXR-300 instrument. These spectra were recorded with tetramethylsilane or 2,2-dimethyl-2-silapentane-5-sulfonic acid, sodium salt (DSS) (both δ = 0.0 for ^1H), and $CDCl_3$ (δ = 77.0 for ^{13}C) or $(CD_3)_2SO$ (δ = 39.5 for ^{13}C) or CD_3OD (δ = 49.0 for ^{13}C) or 1,4-dioxane (δ = 66.5 for ^{13}C) as internal reference. ROESY NMR spectra of **1a-c** in $(CD_3)_2SO$ and D_2O solution were recorded at 30 °C and were obtained in phase-sensitive mode by the method of Kessler et al.[5] That for **1c** in $(CD_3)_2SO$ solution is shown in Fig. 3. All nOe

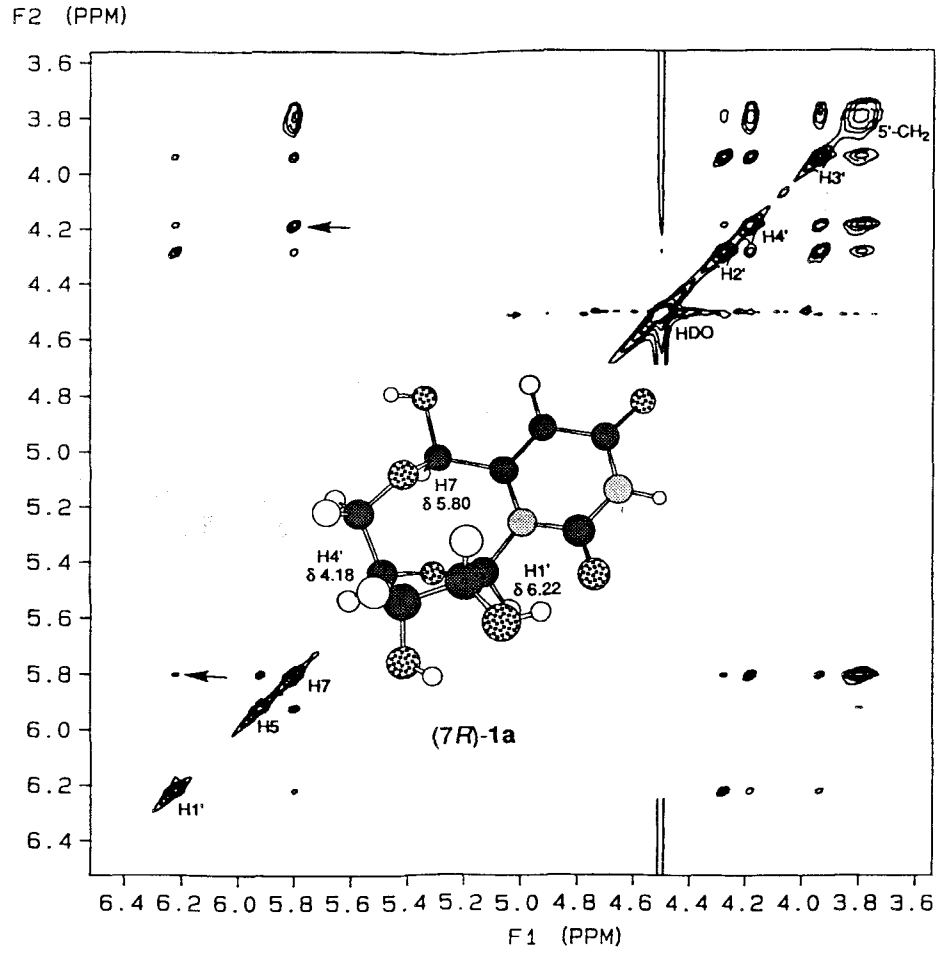


FIG. 1 Selected region of a 500 MHz NOESY NMR spectral plot of **1a** in 70:30 $(CD_3)_2SO/D_2O$ at -15 °C showing diagnostic crosspeaks at F1/F2 = δ 5.80/6.22 and 5.80/4.18 ppm due to nOe interactions of H7/H1' and H7/H4', respectively.

crosspeaks were negative. Phase-sensitive NOESY spectra of **1a** and **1b** shown in Figs. 1 and 2, respectively, were obtained by the method of States et al.,[6] at -15 °C in 70:30 $(CD_3)_2SO/D_2O$ solution. All ROESY and NOESY analyses were conducted at 500 MHz and employed a mixing time of 500 ms.

Antiproliferative Evaluation in Human Tumor Cell Lines in Vitro. Compounds **1a-c** were evaluated in the MCF-7 breast adenocarcinoma (American Type Culture Collection HTB 22) and HT-29 colon adenocarcinoma (American Type Culture Collection

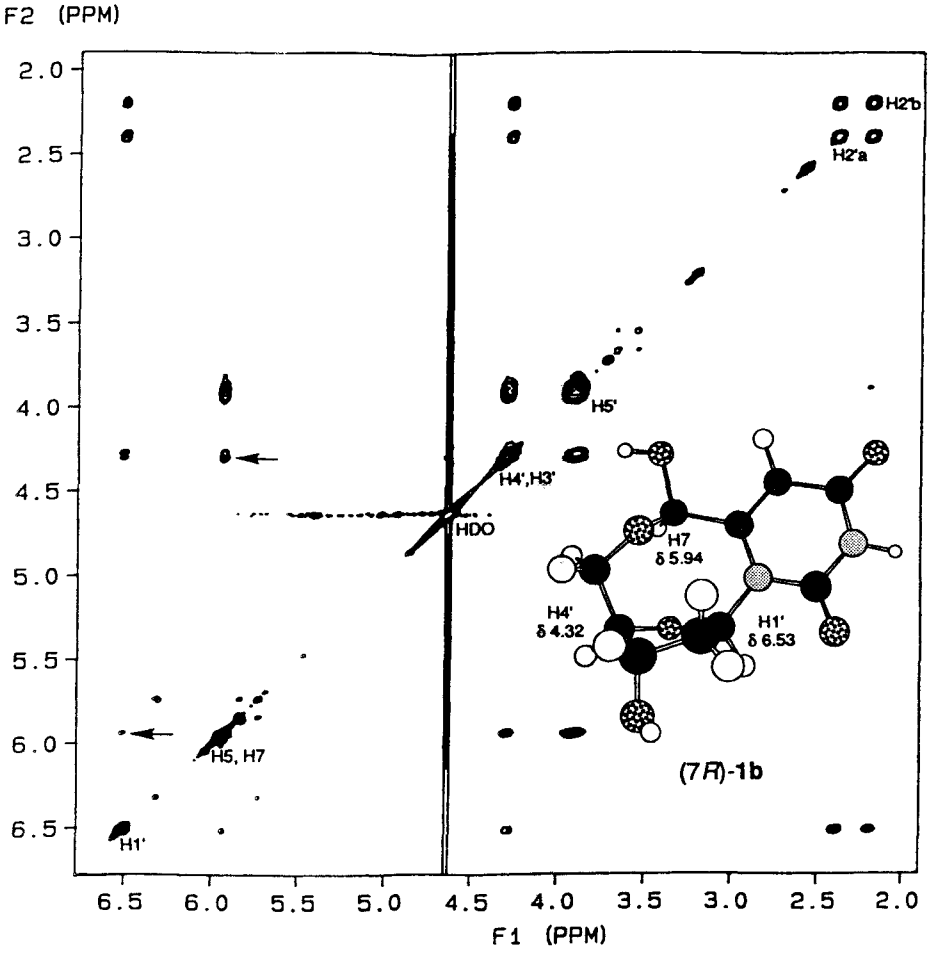


FIG. 2 Selected region of a 500 MHz NOESY NMR spectral plot of **1b** in 70:30 $(CD_3)_2SO/D_2O$ at -15 °C showing diagnostic crosspeaks at F1/F2 = δ 5.94/6.53 and 5.94/4.32 ppm due to nOe interactions of H7/H1' and H7/H4', respectively.

HTB 38) cell lines. Both cell lines were grown in RPMI 1640 medium with 5% fetal bovine serum, 37 °C, 5% CO_2 in air. Control population doubling times were 24 hours for MCF-7 and 17 hours for HT-29 cells. Cells were harvested and distributed in 96-well plates: 2000 cells/well for MCF-7 and 1000 cells/well for HT-29. The next day, the test compounds were added at the concentrations indicated, and after 5 days incubation with the compounds the extent of cell growth was assayed using the sulforhodamine B assay.[7] These protocols were optimized for each cell line to provide exponential cell proliferation,

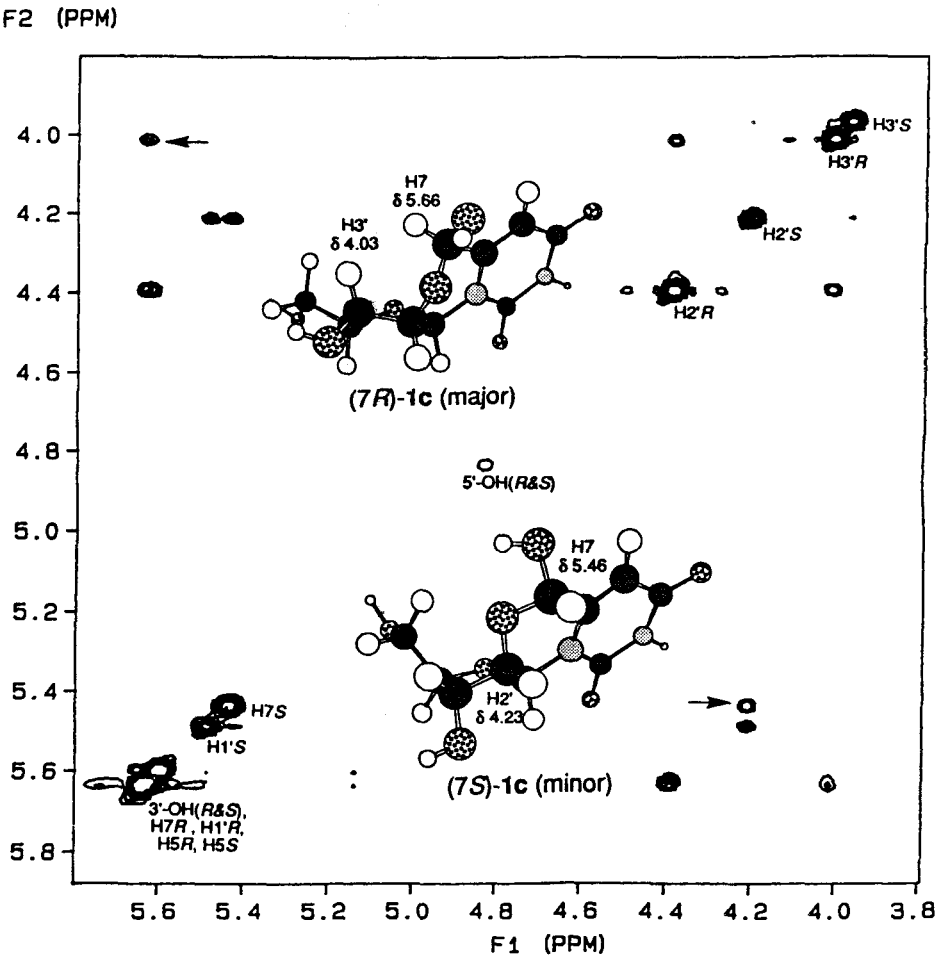


FIG. 3 Selected region of a 500 MHz ROESY NMR spectral plot of **1c** in $(CD_3)_2SO$ solution showing diagnostic crosspeaks at F1/F2 = δ 5.66/4.03 and 5.46/4.23 ppm due to strong nOe interactions of H7/H3' for (7*R*)-**1c** and H7/H2' for (7*S*)-**1c**, respectively.

and exponential increase of the optical density in the sulforhodamine B dye assay, for control cells throughout the incubation period. The average optical density obtained for treated wells (16 wells in each experiment), as % of the average optical density for control wells (16 wells in each experiment), is reported as the % of control in Table I. The IC_{50} was defined as the concentration required to decrease the value to 50% of control, and it was derived graphically from a plot of % of control against log of concentration.

Compound	Concentration (μM)	HT-29 cells (% of control)[a]	MCF-7 cells (% of control)
1a	100	99 (n=2)[b]	99 (n=3)
1b	100	102 (n=4)	58 (n=5)
	200	92 (n=2)	32 (n=2)
1c	100	98 (n=2)	98 (n=3)

[a]% of control optical density in the sulforhodamine B dye assay after 5 days incubation with the test compounds.
[b]n = number of independent experiments the results of which were averaged to obtain the value reported.

TABLE II. Anti-HIV-1 and -HIV-2 Activity and Cytotoxic Properties of **1a-c**

Compound	HIV-1[a] EC_{50} (μg/mL)	HIV-2[b] EC_{50} (μg/mL)	CC_{50} (μg/mL)[c]
1a	> 100	> 100	> 100
1b	> 199.6	> 199	78.0 ± 31.1
1c	> 100	> 100	> 100

1-(5-*O*-*t*-Butyldimethylsilyl-2,3-*O*-isopropylidene-β-D-ribofuranosyl)-6-formyluracil (6). A solution of 5'-*O*-(*t*-butyldimethylsilyl)-2',3'-*O*-isopropylideneuridine (**5**)[8] (2.00 g, 5.0 mmol) in 7 mL of dry THF was added over 15 min to a freshly prepared, -78 °C solution of LDA (13.7 mmol of iPr_2NH treated with 13.0 mmol of BuLi) in dry THF (10 mL). The solution was stirred at -78 °C for 1.25 h, then 2 mL of HCO_2Et was added, and the reaction mixture was stirred for an additional 2 h at -78 °C. The mixture was quenched while cold by the addition of HOAc (0.75 mL, 13.0 mmol), was allowed to warm slowly to r.t., and was then rotary evaporated under reduced pressure at ca. 50 °C. The residue was partitioned between CH_2Cl_2 and sat. aq. $NaHCO_3$, and the combined organic extracts were dried ($MgSO_4$) and rotary evaporated to give a residue that was separated by sequential radial chromatography using 3% and then 5% MeOH/CH_2Cl_2 as eluent in portionwise fashion on 4 mm thick plates to afford 1.55 g (72%) of pure **6**: mp 84-85 °C. ^1H NMR ($CDCl_3$) δ 9.69 (s, 1H, CHO), 8.15 (bs, 1H, NH), 6.38 (d, 1H, H1'), 6.19 (s, 1H, H5), 5.05 (d of d, 1H, H2'), 4.79 (m, 1H, H3'), 4.12 (d of d, 1H, H4'), 3.78 (m, 2H, H5'), 1.55 and 1.43 (each s, each 3H, Me_2C), 0.88 (s, 9H, Me_3C),

0.50 (s, 6H, Me_2Si). ^{13}C NMR ($CDCl_3$) δ 184.8 (CHO), 162.2 (C4'), 150.0 and 147.4 (C2/C6), 114.6 and 111.6 (C5, Me_2C), 90.9, 88.1, 84.3, and 80.9 (C1', C4', C2', C3'), 63.4 (C5'), 27.2 and 25.3 (Me_2C), 25.9 (Me_3C), 22.0 and 18.5 (Me_2Si). LR-EIMS, *m/e* 411.2 (M^+-Me), 396.2 (M^+-Me_3C). LR-CIMS, *m/e* 427.3 (MH^+), 396.2 (M^+-Me_3C). Anal. Calcd for $C_{19}H_{30}N_2O_7Si$: C, 53.50; H, 7.09; N, 6.57. Found: C, 53.62; H, 7.08; N, 6.58.

6-Formyl-1-(β-D-ribofuranosyl)uracil (1a). A solution of **6** (0.50 g, 1.2 mmol) in 5 mL of dry THF was treated dropwise with 10 mL of 50% aq. TFA. The resulting mixture was stirred at r.t. for 24 h and then was rotary evaporated to dryness in vacuo at ca. 50 °C. Residual TFA was then removed by azeotropic rotary coevaporation with EtOH and H_2O. The resultant pale yellow solid was then recrystallized from MeOH to afford 270.5 mg (85%) of pure **1a**, identical to that obtained previously[1a] by ^1H NMR.

1-(2-Deoxy-β-D-ribofuranosyl)-6-formyluracil (1b). This nucleoside was prepared by our reported procedure,[1b] but was isolated for the first time as a microcrystalline solid by purification by repetitive radial chromatography (10% MeOH/CH_2Cl_2 as eluent) followed by recrystallization from absolute EtOH: mp 150-155 °C (dec). By ^1H NMR, **1b** in CD_3OD solution exists as a 20:35:45 ternary mixture composed of the (7*R*)-7,O5'-cyclic hemiacetal and the C7 epimeric methyl hemiacetal diastereomers. For the 7,O5'-cyclic hemiacetal: δ 6.66 (1H, pseudo-t, H1'), 6.08 (1H, s, H5), 5.98 (1H, s, hemiacetal CH), 4.31-4.29 (2H, m, H3' and H4'), 4.00 (1H, d of d, H5'a), 3.92 (1H, d, H5'b), 2.51 (1H, m, H2'a), 2.31 (1H, m, H2'b); $^3J_{1'\text{-}2'}$ = 6.7, $^3J_{4'\text{-}5'a}$ = 4.5, $^2J_{5'a\text{-}5'b}$ = 12.9 Hz. For the methyl hemiacetals: δ 6.34 (1H, d of d, H1' minor), 6.22 (1H, d of d, H1' major), 5.91 (1H, s, H5 major), 5.87 (1H, s, H5 minor), 5.45 (1H, s, H7 major), 5.40 (1H, s, H7 minor), 5.52-5.48 (1H, m, H3'), 3.86-3.82 (1H, m, H4'), 3.79-3.68 (2H, m, 5'-CH_2), 2.89-2.83 (1H, m, H2'a), 2.15-2.08 (1H, m, H2'b); $^3J_{1'\text{-}2'}$ = 6.2 (major), 6.1 Hz (minor).

1-(β-D-Arabinofuranosyl)-6-formyluracil (1c). A suspension of TLC-pure 1-β-D-arabinofuranosyluracil[9] (20 g, 81.9 mmol) in 800 mL of $CH_2(OMe)_2$ containing 5 mL of TfOH was stirred at room temperature for 18 h. By then, the reaction mixture had become nearly homogeneous and TLC analysis (5% MeOH/CH_2Cl_2) indicated that the expected tri-*O*-methoxymethylated derivative was present together with small amounts of di-*O*-methoxymethylated and other byproducts. The distribution of products (by TLC) was not affected by employing a larger quantity of TfOH or by extending the reaction time. After suction filtration, the filtrate was neutralized by adding 30% aq. NH_4OH and stirring vigorously. The layers were separated, and the aqueous layer was extracted with Et_2O (3 x 100 mL). The combined organic solutions were dried ($MgSO_4$) and rotary evaporated to an oil that was separated by column chromatography (5% MeOH/CH_2Cl_2 as eluent) to

afford 14.62 g (45%) of 1-(2,3,5-tris-*O*-methoxymethyl-β-D-arabinofuranosyl)uracil as an oil. Formylation and deprotection performed as previously reported[1b] then gave **1c**. ^1H NMR analysis of this compound revealed that it exists as a 3:1 mixture of 2'-cyclonucleoside diastereomers in $(CD_3)_2SO$ solution as before. A closer inspection of the ^1H, ^{13}C, 2D COSY and HETCOR NMR spectra necessitates our amending the previous[1b] ^1H and ^{13}C spectral assignments for the C_5H and hemiacetal-CH moieties of the major isomer (*7R*) of **1c** in $(CD_3)_2SO$: ^1H NMR δ 5.66 (d, J = 5.7 Hz, collapses to a singlet upon addition of D_2O, 1H, hemiacetal-CH), 5.61 (s, 1H, H5); ^{13}C NMR δ 99.0 (C5), 86.9 (hemiacetal-CH).

[3,5-*Bis*-*O*-(*t*-butyldimethylsilyl)-2-deoxy-β-D-ribofuranosyl]-6-formylthymine (8). A solution of **7**[10] (1.977 g, 4.2 mmol) in 20 mL of dry THF at -78 °C was treated with TMEDA (3.0 mL, 20.0 mmol) and then with 16.1 mL of a 1.24 M solution (20.0 mmol) of *n*-BuLi in hexanes. The solution was stirred at -78 °C for 2 h, then 2.0 mL (excess) of HCO_2Et were added and the reaction mixture was stirred for 30 min at -78 °C. The mixture was quenched while cold by adding 1.1 mL of HOAc, and the volatiles were then removed by rotary evaporation under reduced pressure at r.t. The residue was partitioned between CH_2Cl_2 and sat. aq. $NaHCO_3$, the layers were separated, and the aqueous layer was washed several times with CH_2Cl_2. The organic solutions were combined, dried ($MgSO_4$), and then were rotary evaporated to a mixture that was separated by repetitive radial and ascending preparative chromatography using 5% CH_3OH/CH_2Cl_2 and 2:1 EtOAc/hexanes as eluents to afford 251 mg (12%) of **8**, characterized on the basis of its ^1H and ^{13}C NMR and LR mass spectral properties: ^1H NMR ($CDCl_3$) δ 10.04 (1H, s, CHO), 8.8 (1H, br, NH), 6.01 (1H, pseudo-t, H1'), 4.51 (1H, m, H3'), 3.92 (1H, m, H4'), 3.88-3.68 (2H, m, 5'-CH_2), 2.60 and 2.30 (each 1H, each s, 2'-CH_2), 2.06 (3H, s, Me), 0.89 (18H, br, two Me_3C), 0.09 and 0.08 (each 6H, each s, each Me_2Si); $^3J_{1'-2'}$ = 7.2 Hz. ^{13}C NMR ($CDCl_3$) δ 186.5 (CHO), 163.1 (C4), 149.3 (C2), 143.7 (C6), 114.3 (C5), 88.6 and 88.3 (C4'and C1'), 71.2 (C3'), 61.8 (C5'), 40.6 (C2'), 25.7 (Me_3C), 17.9 (Me_2Si), 9.5 (Me). LR-CIMS, *m/e* 385.3 (M^+-Me_3C+2). LR-FDMS, *m/e* 498 (M^+), 385 (M^+-2 Me_3C+1), 231 (M^+-2 Me_3C-base+2).

1-(2-Deoxy-β-D-ribofuranosyl)-6-formylthymine (1d). A solution of **8** (150 mg, 0.30 mmol) in 0.5 mL of dry THF was treated dropwise with 1.0 mL of a 1.0 M solution of TBAF in THF, and the reaction mixture was stirred for 1 h at r.t. The residue remaining after a rotary evaporative removal of volatiles was purified by repetitive column chromatography using 10% CH_3OH/CH_2Cl_2 as eluent, and an aqueous solution of the chromatographically-pure **1d** thus obtained was subjected to lyophilization for >24 h to provide 67.4 mg (83%) of product as a highly hygroscopic white foam. By ^1H and ^{13}C NMR, **1d** in $(CD_3)_2SO$ solution exists exclusively as the aldehyde: ^1H NMR ($(CD_3)_2SO$)

δ 11.35 (1H, br, NH), 9.95 (1H, s, CHO), 6.04 (1H, pseudo-t, H1'), 5.26 (1H, exchanges with D_2O, d, 3'-OH), 4.88 (1H, exchanges with D_2O, t, 5'-OH), 4.13 (1H, m, H3'), 3.73 (1H, m, H4'), 3.49 (2H, m, 5'-CH_2), 2.24 (1H, d of d, H2'a), 2.15 (1H, d of d of d, H2'b), 1.82 (3H, s, Me); $^3J_{1'-2'}$ = 6.4, $^3J_{3'-3'OH}$ = 4.7, $^3J_{5'-5'OH}$ = 5.3 Hz. ^{13}C NMR ($(CD_3)_2SO$) δ 188.0 (CHO), 163.6 (C4), 149.4 (C2), 143.6 (C6), 110.5 (C5), 87.9 and 86.1 (C4'/C1'), 69.9 (C3'), 60.6 (C5'), 40.0 (C2'), 9.6 (Me). LR-CIMS, *m/e* 155.1 (100%, B^+). LR-EIMS, *m/e* 154.1 (40%, B^+-1). LR-ESIMS (Electrospray ionization) (MeOH), *m/e* 303.1 (M^++1: hemiacetal), 271.2 (M^++1: aldehyde). LR-FABMS (3-NBA), *m/e* 307.1 (35%, M^++H_2O+1), 289.1 (65%, M^++2H_2O+1).

By ^1H and ^{13}C NMR, **1d** in CD_3OD solution exists as a 47:47:6 ternary mixture composed of the C7 epimeric methyl hemiacetal diastereomers and the aldehyde. For the methyl hemiacetals: ^1H NMR δ 6.73 (1H, pseudo-t, both H1'), 5.85 and 5.84 (each 1H, each s, each H7), 4.51 (1H, m, both H3'), 3.82-3.70 (3H, m, both H4' and both 5'-CH_2), 2.80 (1H, m, both H2'a), 2.18 (1H, m, both H2'b), 1.99 and 1.98 (each 3H, each s, each Me); ^{13}C NMR δ 166.3 (C4), 152.0 (C2), 149.7 and 149.5 (each C6), 109.6 (C5), 95.2 and 95.1 (each C7), 89.7 and 89.5 (each C4'/C1'), 88.8 and 88.7 (each C4'/C1'), 72.5 and 72.4 (each C3'), 63.8 (C5'), 39.5 and 39.4 (each C2'), 10.1 and 10.0 (each Me). For the aldehyde: ^1H NMR δ 10.02 (1H, s, CHO), 6.15 (1H, pseudo-t, H1').

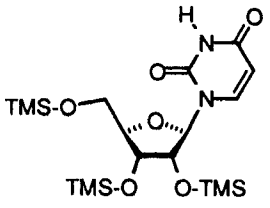
By ^1H and ^{13}C NMR, **1d** in either 2:1 $(CD_3)_2SO/D_2O$ or in D_2O solution exists as an equimolar mixture of the hydrate and the aldehyde. For the hydrate: ^1H NMR ($(CD_3)_2SO/D_2O$) δ 6.69 (1H, d of d, H1'), 4.33 (1H, m, H3'), 3.72-3.58 (3H, m, H4' and 5'-CH_2), 2.65 (1H, m, H2'a), 2.15 (1H, m, H2'b), 1.88 (3H, s, Me); $^3J_{1'-2'}$ = 5.8 Hz; ^1H NMR (D_2O) δ 6.81 (1H, d of d, H1'), 4.53 (1H, m, H3'), 3.91(1H, m, H4'), 3.85-3.3.79 (2H, m, 5'-CH_2), 2.87 (1H, m, H2'a), 2.25 (1H, m, H2'b), 1.98 (3H, s, Me); $^3J_{1'-2'}$ = 5.4 Hz. For the aldehyde: ^1H NMR ($(CD_3)_2SO/D_2O$) δ 9.92 (CHO), 6.01(1H, pseudo-t, H1'), 4.15 (1H, m, H3'), 3.77 (1H, m, H4'), 3.60-3.52 (2H, m, 5'-CH_2), 2.32 (1H, m, H2'a), 2.25 (1H, m, H2'b), 1.82 (3H, s, Me); $^3J_{1'-2'}$ = 7.0 Hz; ^1H NMR (D_2O) δ 10.33 (1H, s, CHO), 6.13 (1H, pseudo-t, H1'), 4.37 (1H, m, H3'), 3.96 (1H, m, H4'), 3.79-3.72 (2H, m, 5'-CH_2), 2.59 (1H, m, H2'a), 2.37 (1H, m, H2'b), 1.96 (3H, s, Me); $^3J_{1'-2'}$ = 7.2 Hz. For both: ^{13}C NMR ($(CD_3)_2SO/D_2O$) δ 189.3 (CHO), 166.0 and 165.6 (each C4), 151.3 and 150.7 (each C2), 150.8 (C6 hydrate), 145.0 (C6 aldehyde), 112.8 (C5 aldehyde), 108.1 (C5 hydrate), 88.5, 88.3, 87.8, and 87.5 (each C4'/C1'), 86.7 (C7 hydrate), 71.9 (C3' hydrate), 71.0 (C3' aldehyde), 63.1 (C5' hydrate), 61.7 (C5' aldehyde), 40.7 (C2' aldehyde), 38.8 (C2' hydrate), 10.5 and 10.4 (each Me).

A ^1H NMR monitoring of a D_2O solution of **1d** gave evidence of slow hydrolysis at r.t. ($t_{1/2}$ ca. 24 h). After complete hydrolysis, 6-formylthymine (**9**)[11] was isolated in

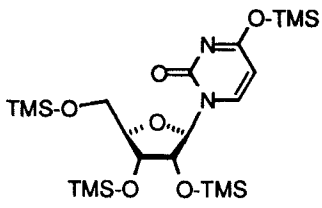
quantitative yield. Similar to that of **1d**, ^{1}H NMR analysis of **9** revealed that it exists as a mixture (4:1) of the hydrate and the aldehyde in D_2O. For the hydrate: ^{1}H NMR δ 6.07 (1H, s, $CH(OD)_2$), 1.90 (1H, s, CH_3). For the aldehyde: ^{1}H NMR δ 10.09 (1H, s, CHO), 2.24 (1H, s, CH_3). ^{1}H NMR analysis of **9** in CD_3OD solution revealed it to exist as the methyl hemiacetal: ^{1}H NMR δ 5.54 (1H, s, hemiacetal-CH), 1.89 (1H, s, CH_3); ^{13}C NMR δ 92.6 (hemiacetal-CH), 9.0 (CH_3).

RESULTS AND DISCUSSION

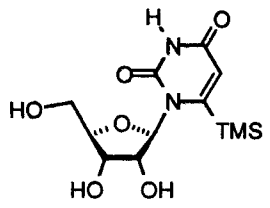
Syntheses. We first sought new syntheses of **1a-c** that were based on the transient TMS protection methodology developed by Jones' group,[12] even though we were aware of work from Pichat's group[13] which indicated that a dianion generated by the BuLi-mediated dideprotonation of 2',3',5'-tris-*O*-(trimethylsilyl)uridine (**2**) led to a mixture of C-5- and C-6-functionalized products. We prepared **2** and then used LDA and HCO_2Et in attempts to prepare **1a** according to a Miyasaka uridine-3,6-diyllithium-based approach,[14] but failed. Still, we considered that 4,2',3',5'-tetrakis-*O*-(trimethylsilyl)uridine (**3**)[13b] might lead to **1a** if a uridin-6-yllithium derivative of it could be generated and formylated, so we sought this pertrimethylsilylated nucleoside. When equilibration of uridine in HMDS solution containing a catalytic amount of $(NH_4)_2SO_4$[15] gave only a 1:2 mixture of **2** and **3** (isolated and characterized by ^{1}H NMR), we used more forcing conditions (TMS-Cl, Et_3N, C_6H_6) to complete the conversion to the desired **3**, obtained in near-quantitative yield. After



2



3

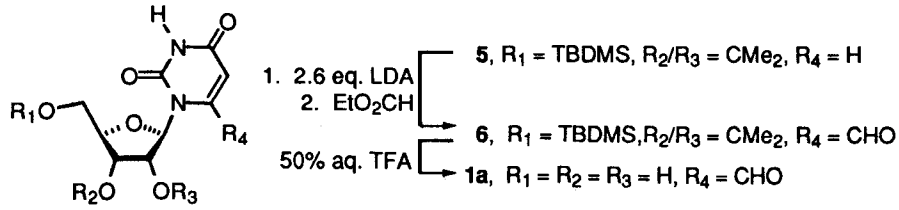


4

combining LDA and **3** in ratios of 1:1, 2:1, or 5:1 for monodeprotonation in THF solution at -78 °C, however, attempted formylation by quench with HCO_2Et followed by hydrolysis consistently led to the isolation of uridine and 6-(trimethylsilyl)uridine (**4**).[16] Based on this outcome, we concluded that most (or all) of any uridin-6-yllithium species we had managed to generate had undergone intra- or intermolecular TMS group transfer faster than the desired formylation, and so abandoned the transient TMS protection approaches.

We then proceeded to develop a more traditional, four step route to **1a** from uridine in which each intermediate and the final target could be prepared in high yield and with a simple workup procedure. The known 5'-*O*-TBDMS-2',3'-*O*-isopropylideneuridine (**5**)[8]

was readily prepared in > 85% overall yield in two steps from uridine, and was converted via its 3,6-diyl dianion into the 6-formyluridine **6** in higher yield (72%) than that (50%)[1a] obtained by us previously for the similar formylation of the 2',3',5'-tri-*O*-MOM-protected uridine. That the 72% yield, which coincidentally matches that reported recently by Megati et al.[17] for a similar formylation of the 5'-*O*-(methoxymethyl) analog of **5**, was higher than

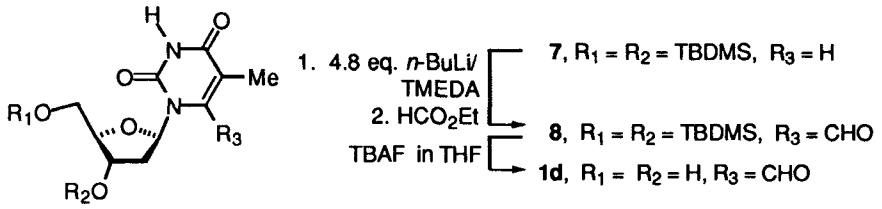


before is likely a consequence of a reduction in steric hindrance to the production of a *syn* conformer uridin-3,6-diyl dianion effected by the presence of a 2',3'-*O*-isopropylidene protecting group, as pointed out by Miyasaka in studies with dilithiated **5**.[18] Target **1a** was obtained in 85% yield by deprotection of **6** in aqueous TFA.

The 6-formylated 2'-deoxyribonucleoside **1b** was prepared as previously reported,[1b] but the purification protocol was improved. In the past, it had proven difficult yet possible to remove residual TBAF from **1b** by repetitive silica gel chromatography alone. This difficulty has now been circumvented by adding an EtOH recrystallization step after the chromatographic separations, a procedure that has afforded **1b** for the first time as a microcrystalline solid rather than as an amorphous semisolid. This result is particularly important as **1b** is the only nucleoside of the three uridines to show any hint of bioactivity (*vide infra*).

In the synthesis of **1c**, we have achieved a dramatic reduction in the reaction mixture volume in the tri-*O*-MOM protection step. By increasing the amount of TfOH catalyst, we have found it possible to use 40 mL of pure $CH_2(OMe)_2$ as reagent/solvent per gram of 1-β-D-arabinofuranosyluracil instead of the 500 mL of 1:1 $CH_2(OMe)_2/CH_2Cl_2$ used previously. This new protocol lowers the yield somewhat (45% instead of 60%), but is much preferred for large scale preparations. Uridine itself gives the 2',3'-*O*-methylidene derivative as the major product under these new conditions (by ^1H NMR), but the use of the four-step synthesis of **1a** described above obviates the need for 2',3',5'-tris-*O*-(methoxymethyl)uridine.

Synthesis of 6-formylthymidine (**1d**) began with the preparation of 3',5'-bis-*O*-TBDMS-thymidine (**7**),[10] which was then converted to aldehyde **8** in 12% yield by treatment with 4.8 equiv. of *n*-BuLi/TMEDA in THF solution for 2 h at -78 °C followed by equilibration with HCO_2Et for 30 min, also at -78 °C. Despite many efforts, we were unable to effect an improvement in the yield by adjusting the time allowed for dianion formation or its formylation, or by using *t*-BuLi/TMEDA, LTMP, or LDA in place of the *n*-BuLi/TMEDA. Deprotection of **8** by treatment with TBAF gave **1d** in 83 % yield.

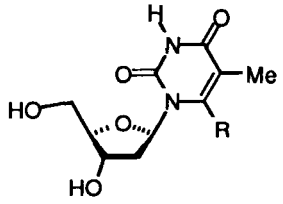


NMR Analyses. ROESY NMR spectroscopy[5] was used to conduct a detailed examination the various solution species of **1a-c** in $(CD_3)_2SO$ and in D_2O, and NOESY NMR spectroscopy[6] was used to evaluate **1a,b** in a 70:30 $(CD_3)_2SO/D_2O$ binary solvent solution at low temperature (-15 °C). ROESY NMR is the preferred method for small molecules because of the inherent low intensity of their nOe's, but it often generates artifacts due to HOHAHA contributions.[19] On the other hand, NOESY NMR does not work well at room temperature for compounds in this molecular mass range, and thus mixed solvent systems are often a necessity for achieving requisite low temperature experimental conditions. According to either method, the 7,O5'-hemiacetal of the ribonucleoside **1a** and of the 2'-deoxyribonucleoside **1b** in $(CD_3)_2SO$ and in D_2O solution possess the same 7*R* stereogenic center as that reported earlier in the X-ray crystal structure determination of **1a** deposited from water.[1b] Figs. 1 and 2 display regions of the NOESY contour plots for $(CD_3)_2SO/D_2O$ solutions of **1a** and **1b**, respectively, that show the structurally diagnostic ^1H-^1H nOe crosspeaks due to H7/H1' and H7/H4' through-space interactions. Fig. 3 shows a region of a ROESY contour plot for the binary C7 epimeric diastereomer mixture of **1c** 7,O2'-cyclic hemiacetal structures in $(CD_3)_2SO$ solution in which the structurally diagnostic ^1H-^1H nOe crosspeaks of H7/H3' for the 7*R*, major, and the H7/H2' for the 7*S*, minor constituent are readily apparent. Corroborative ROESY NMR results were obtained from $(CD_3)_2SO$ solutions of **1a-b** and of D_2O solutions of **1a-c**. The structures accompanying the Figs. are molecular mechanics energy minima consistent with the 2D NMR data, and were generated by MM2 force-field calculations implemented by Chem3D Pro, version 3.1.3, of Cambridge Scientific Computing, Inc. They are provided as suggested conformations only, and are not necessarily global minima in every case. We note that those of **1a,b** are quite similar to that determined rigorously by Megati et al. for $O^{5',6}$-methanocytidine.[20]

Earlier,[1b] we had reported that the relative proportion of **1c** diastereomers appeared to change slowly over time to favor the initially less abundant one, namely (7*S*)-**1c**, which therefore could be identified as the thermodynamically more stable isomer. We have now found this not to be the case. Recrystallization of freshly prepared 9:5 binary diastereomeric mixture of (7*R*)- and (7*S*)-**1c** does indeed afford a microcrystalline solid mixture enriched in the latter (3:5 7*R*/7*S* by integration of the H2' resonances in the ^1H NMR

solution spectrum), but a ^1H NMR monitoring of this new mixture held at 75-80 °C in either $(CD_3)_2SO$ or D_2O solution revealed a slow return to the initial ratio of ca. 2:1 7*R*/7*S*. Thus, (7*S*)-**1c** is actually the thermodynamically *less* stable of these two diastereomers. In addition, a ^1H NMR spectral analysis of the mother liquor remaining after recrystallization revealed it to be enriched in the 7*R* isomer (5:1 7*R*/7*S*), permitting us to conclude that the change in component ratio after recrystallization is a result of a simple solubility difference in which (7*S*)-**1c** is less soluble and crystallizes out of solution to a greater extent than does its diastereomeric counterpart.

In great contrast to the uridines **1a-c**, thymidine **1d** has exhibited no cyclic hemiacetal structural form in any solution environment examined to date. Compound **1d**, characterized on the basis of its ^1H, ^{13}C, ^1H-^1H COSY, ^1H-^{13}C HETCOR NMR, and LR mass spectral properties, was found to exist exclusively in aldehyde form when in anhydrous $(CD_3)_2SO$ solution, even upon warming (50-60 °C). Gradient addition of D_2O to this solution up to a 33% v/v extent produced an equimolar mixture of hydrate and aldehyde forms identical to that exhibited by **1d** in pure D_2O solution. Deoxynucleoside **1d** shows a susceptibility toward undergoing slow glycosidic bond hydrolysis in this latter solution. Assignment of the hydrate structure in the $(CD_3)_2SO/D_2O$ medium was made based upon the diagnostic ^{13}C NMR chemical shift (δ 86.7 ppm) of C7. The only difference between thymidine **1d** and deoxyuridine **1b** is the presence of the 5-Me group in the former, and the contrast between the structural form array exhibited by them in CD_3OD solution is especially striking. An equimolar mixture of C7 epimeric methyl hemiacetal diastereomers accounts for 94% of **1d** in this solvent, the remainder being the aldehyde. Deoxyuridine **1b** in the same solvent exists as a 20:35:45 ternary mixture of (7*R*)-7,O5'-cyclic hemiacetal and two C7 epimeric methyl hemiacetal diastereomers. We have not yet determined the



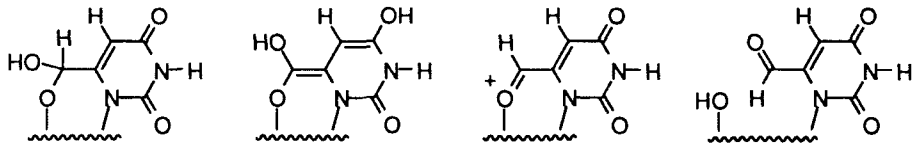
1d, R = CHO (100% in $(CD_3)_2SO$, 50% in D_2O and 2:1 $(CD_3)_2SO/D_2O$, 6% in CD_3OD)
R = $CH(OH)_2$ (50% in D_2O and 2:1 $(CD_3)_2SO/D_2O$)
R = $CH(OH)OCH_3$ (94% in CD_3OD)

absolute configuration of these **1b** methyl hemiacetals, and have no ready explanation for the slight apparent stereofacial selectivity associated with their formation.

Antitumor and Antiviral Activity Studies. Compounds **1a-c** were evaluated for growth inhibitory activity in two human tumor cell lines: HT-29 colon carcinoma and MCF-7 breast carcinoma. As shown in Table I, the 2'-deoxynucleoside **1b** exhibited a slight antiproliferative activity in MCF-7 cells, but none of the compounds had detectable activity in HT-29 cells. The IC_{50} for **1b** in MCF-7 cells was 130 μM. In anti-HIV-1 and -2 activity screens (Table II), none of the 6-formyluridine derivatives **1a-c** was signifi-

cantly active at 100 μg/mL concentration levels. The 2'-deoxynucleoside **1b** showed a toxicity CC_{50} level of 78 μg/mL, while neither **1a** nor **1c** was significantly cytotoxic. While none of these compounds exhibited significant antiviral or antiproliferative activity, the slight cytotoxicity activity observed for the 2'-deoxynucleoside **1b** in MCF-7 and CEM cells suggested that the thymidine analog **1d** might be of interest. Its synthesis has now been accomplished, but biological evaluations will be difficult due to the aforementioned hydrolytic lability in water.

Structural Equilibria. How do the C7 epimeric 7,O2'-cyclic hemiacetals of **1c** (and perhaps the 7,O5' ones of **1a,b**) interconvert? There are at least three mechanisms to consider. The first of these stems from the expectation that proton lability at C7, a position vinylogously α to the pyrimidine 4-keto group, should be facile when the C7-H7 bond vector is oriented in near perpendicular fashion to the plane of the uracil ring, as it is in the cyclic hemiacetal solid-state structure of **1a**[1b] and in the solution cyclic hemiacetal forms of **1a-c**, by NMR. Lack of a stereofacially-specific proton transfer in enol **A** (below) back to C7 would indeed be expected to lead to diastereomer interconversion, but this would neces-



1a-c cyclic hemiacetals **A** **B** **C**

sarily proceed with C7 deuterium acquisition upon equilibration of **1a-c** in D_2O solution. We have found that this does not occur. The second mechanism involves an anionization of the hydroxyl group at C7 and a stereofacially-indiscriminate hydroxide group return in the resultant species **B**, leading to diastereomer interconversion. If such a high energy, cross-conjugated oxonium ion as **B** were to be generated from **1a-c**, however, alkyl acetals would have been obtained upon recrystallization of these compounds. They were not. Indeed, **1a-c** were found resistant to acetal formation under conventional conditions (e.g, heating in lower primary alcohols containing TfOH as catalyst). The third and most likely mechanistic possibility focuses on the "open" hydroxy-aldehyde species **C** inarguably present and in equilibrium with the hemiacetals. Diastereomer interconversion in **1c** would be a consequence of formyl group (or glycosidic bond) rotation in **C** prior to reformation of the hemiacetal linkage. The direct NMR observation of small amounts of aldehyde or hydrate species in solutions of **1a,b** is consistent with this hypothesis.

Why is it that both C7 epimeric 7,O2'-cyclic **1c** hemiacetal diastereomers are ever present in good relative amounts in solutions, but only the (7*R*)-7,O5' one has ever been observed for **1a,b**? This difference is related to the fact that it is only in **1a,b** that the

hemiacetal linkage is formed directly over the β face of the furanose ring. The detection of only the 7*R* epimer of **1a,b** cyclic hemiacetals could be a consequence its greater stability compared to the 7*S* counterparts (a thermodynamic argument), or of a lower energy pathway towards their formation (a kinetic one), or of both. The thermodynamic argument is supported by the observation that only the 7*R* epimers can experience dipole stabilization via through-space interaction between the furanose ring O4' lone pair electrons and the antibonding molecular orbital antipode of the C7-O7 σ-bond in somewhat of a "virtual" anomeric effect. From the solid-state structure determination of **1a**,[1b] the O4'-C7-O7 "angle" is nearly optimal at 7.8°, and the O4'-C7 distance is somewhat long at 2.77 Å. The kinetic argument is more complex, and arises from the observation that hemiacetal linkage formation in **1a,b** occurs in a highly asymmetric environment by way of a conformer of the "open" hydroxy-aldehyde form that is *anti*-delineating with respect to O4'-C1'-N1-C2 glycosidic torsion angle and thereby highly restrictive of the formyl group's rotational mobility. A careful analysis of kinetic pathways available for the formation of the C7 epimers reveals that they are directionally-dependent upon glycosidic bond rotation and have different energy requirements.

Why does 6-formylthymidine (**1d**) exhibit no evidence of a tendency toward existing in 7,O5'-cyclic hemiacetal form? The explanation depends on whether the kinetic or thermodynamic argument above is adopted. The formyl group in **1d** (and in **9**) is certainly less susceptible towards facile 1,2-addition than is that in **1b**, as evidenced by the presence of the aldehyde form in $(CD_3)_2SO$ and even D_2O solution, but this reduction in electrophilicity is likely not the only reason for lack of detectable amounts of 7,O5'-cyclic hemiacetal(s). The 5-Me substituent in **1d**, by impeding the C6-formyl group's rotational mobility, could be disallowing the adoption of that C6-C7 bond rotamer requisite of a bonding proximity between the C5' hydroxyl group and the *re*-face of the C6 formyl group directly over the β face of the furanose ring. Alternatively or additionally, an unfavorable interaction between the 5-Me and 7-OH groups in the NMR-absent 7*R* cyclic hemiacetal form of **1d** may have elevated its steric energy.

Acknowledgments. This work was supported by a grant from the Office of Research Development and Administration of Southern Illinois University at Carbondale. Work at the Rega Institute is supported in part by the Biomedical Research Programme of the European Commission. WCS thanks Dr. David Vander Velde for suggesting the low-temperature NOESY method. The authors also thank Mr. Paul W. Chrenka for expert technical assistance in the antiproliferative evaluations.

REFERENCES

1. (a) Groziak, M. P.; Koohang, A. *J. Org. Chem.* **1992**, *57*, 940; (b) Groziak, M. P.; Koohang, A.; Stevens, W. C.; Robinson, P. D. *J. Org. Chem.* **1993**, *58*, 4054.
2. (a) Kittaka, A.; Tanaka, H.; Odanaka, Y.; Ohnuki, K.; Yamaguchi, K.; Miyasaka, T. *J. Org. Chem.* **1994**, *59*, 3636; (b) Garg, N.; Hossain, N.; Chattopadhyaya, J. *Tetrahedron* **1994**, *50*, 5273; (c) Camarasa, M.-J.; Pérez-Pérez, M.-J.; San-Félix, A.; Balzarini, J.; De Clercq, E. *J. Med. Chem.* **1992**, *35*, 2721.
3. (a) Hsu, L.-Y.; Wise, D. S.; Drach, J. C.; Townsend, L. B. *Chin. Pharm. J.* **1991**, *43*, 275; (b) Hsu, L.-Y.; Wise, D. S.; Kucera, L. S.; Drach, J. C.; Townsend, L. B. *J. Org. Chem.* **1992**, *57*, 3354.
4. Otter, B. A.; Patil, S. A.; Spada, M. R.; Jelicks, L. A.; Yoshimura, Y.; Matsuda, A.; Klein, R. S. *Nucleosides Nucleotides* **1992**, *11*, 615. These workers prepared the 2',3'-*O*-isopropylidenated **1a** and characterized its 7,O5'-hemiacetal solution structure.
5. Kessler, H.; Griesinger, C.; Kerssebaum, R.; Wagner, K.; Ernst, R. R. *J. Am. Chem. Soc.* **1987**, *109*, 607.
6. States, D. J.; Haberkorn, R. A.; Ruben, D. J. *J. Magn. Res.* **1982**, *48*, 286.
7. Skehan, P. J.; Storeng, R.; Scudiero, D. A.; Monks, A.; McMahon, J. B.; Vistica, D. T.; Warren, J. T.; Bokesch, H.; Kenney, S.; Boyd, M. R. *J. Natl. Cancer Inst.* **1990**, *82*, 1107.
8. Tanaka, H.; Hayakawa, H.; Shibata, S.; Haraguchi, K.; Miyasaka, T. *Nucleosides Nucleotides* **1992**, *11*, 319.
9. Schinazi, R. F.; Chen, M. S.; Prusoff, W. H. *J. Med. Chem.* **1979**, *22*, 1273.
10. Ogilvie, K. K. *Can. J. Chem.* **1973**, *51*, 3799.
11. Zee-Cheng, K.-Y.; Cheng, C. C. *J. Heterocycl. Chem.* **1967**, *4*, 163.
12. Ti, G. S.; Gaffney, B. L.; Jones, R. A. *J. Am. Chem. Soc.* **1982**, *104*, 1316.
13. (a) Pichat, L.; Godbillon, J.; Herbert, M. *Bull. Soc. Chim. Fr.* **1973**, 2715; (b) Pichat, L.; Godbillon, J.; Herbert, M. *Bull. Soc. Chim. Fr.* **1973**, 2709.
14. (a) Tanaka, H.; Nasu, I.; Miyasaka, T. *Tetrahedron Lett.* **1979**, 4755; (b) Tanaka, H.; Hayakawa, H.; Miyasaka, T. *Chem. Pharm. Bull.* **1981**, *29*, 3565; (c) Tanaka, H.; Hayakawa, H.; Miyasaka, T. *Tetrahedron* **1982**, *38*, 2635; (d) Tanaka, H.; Hayakawa, H.; Iijima, S.; Haraguchi, K.; Miyasaka, T. *Tetrahedron* **1985**, *41*, 861.
15. Fessenden, R.; Crowe, D. F. *J. Org. Chem.* **1961**, *26*, 4638.
16. Ikehira, H.; Matsuura, T.; Saito, I. *Tetrahedron Lett.* **1984**, *25*, 3325.
17. Megati, S.; Sodum, R.; Otter, G. M.; Klein, R. S.; Otter, B. A. *Bioorg. Med. Chem. Lett.* **1994**, *4*, 469.
18. Hayakawa, H.; Tanaka, H.; Maruyama, Y.; Miyasaka, T. *Chem. Lett.* **1985**, 1401.

19. Neuhaus, D.; Williamson, M. "The Nuclear Overhauser Effect in Structural and Conformational Analysis", VCH Publishers: 1989, New York, p. 318.
20. Megati, S.; Ealick, S. E.; Naguib, R. N. M.; el Kouni, M. H.; Klein, R. S.; Otter, B. A. *Nucleosides Nucleotides* **1994**, *13*, 2151.

Received December 13, 1994
Accepted January 5, 1996