Structural Chemistry, Vol. 16, No. 5, October 2005 (© 2005)
DOI: 10.1007/s11224-005-6058-9

X-ray Crystal and *Ab Initio* Structures of 3′,5′-di-*O*-Acetyl-*N*(4)-Hydroxy-2′-Deoxycytidine and Its 5-Fluoro Analogue: Models of the *N*(4)-OH-dCMP and *N*(4)-OH-FdCMP Molecules Interacting with Thymidylate Synthase

Adam Jarmuła,[1,4] Wojciech R. Rypniewski,[2] Krzysztof Felczak,[3] and Wojciech Rode[1]

The crystal and molecular structures of the 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-2′-deoxycytidine molecule and its 5-fluoro congener have been determined by X-ray single crystal diffraction. The 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-5-fluoro-2′-deoxycytidine molecule crystallizes in the space group *C*2 with the following unit cell parameters: $a = 21.72$ Å, $b = 8.72$ Å, $c = 8.61$ Å, and $\beta = 90.42^\circ$. 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-2′-deoxycytidine also belongs to the monoclinic space group *C*2 and the unit cell parameters are: $a = 39.54$ Å, $b = 8.72$ Å, $c = 22.89$ Å, and $\beta = 95.26^\circ$. The non-fluorine analogue demonstrates a rare example of crystal structure with five symmetry-independent molecules in the unit cell. All the molecules in both crystal structures have the sugar residue *anti* oriented with respect to the base, as well as have the N(4)-OH residue in *cis* conformation relatively to the N(3)-nitrogen atom. In addition to the molecular geometries from X-ray experiment, the optimized molecular geometries have been obtained with the use of theoretical *ab initio* calculations at the RHF/6-31G(d) level. The corresponding geometric parameters in the molecules of 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-2′-deoxycytidine and its 5-fluoro congener have been compared. The differences including the C(5)═C(6) bond shortening and C(4)—C(5)—C(6) angle widening in the fluorine analogue are discussed in this paper in relation to the molecular mechanism of enzyme, thymidylate synthase, inhibition by *N*(4)-hydroxy-2′-deoxycytidine monophosphate and its 5-fluoro congener.

KEY WORDS: Crystal structure; *ab initio* structure; nucleoside analogues of *N*(4)-OH-dCMP and *N*(4)-OH-FdCMP; thymidylate synthase.

INTRODUCTION

Thymidylate synthase (TS) (EC 2.1.1.45) is an important enzyme with a central role in DNA synthesis. It catalyzes the conversion of 2′-deoxyuridine-5′-monophosphate (dUMP) to 2′-deoxythymidine-5′-monophosphate (dTMP), involving a reductive methylation of N^5,N^{10}-methylenetetrahydrofolate (mTHF), functioning both as the one-carbon group donor and reductant

[1]Nencki Institute of Experimental Biology, Polish Academy of Sciences, Pasteura 3, 02-093 Warszawa, Poland.
[2]Institute of Bioorganic Chemistry, Polish Academy of Sciences, Noskowskiego 12-14, 61-704 Poznań, Poland.
[3]Institute of Biochemistry and Biophysics, Polish Academy of Sciences, Pawińskiego 5a, 02-106 Warszawa, Poland.
[4]To whom all correspondence should be addressed; e-mail: a.jarmula@nencki.gov.pl

Abbreviations: TS, thymidylate synthase; dUMP, 2′-deoxyuridine 5′-monophosphate; FdUMP, 5-fluoro-dUMP; *N*(4)-OH-dCMP, *N*(4)-hydroxy-2′-deoxycytidine 5′-monophosphate; *N*(4)-OH-FdCMP, 5-fluoro-*N*(4)-OH-dCMP; *N*(4)-OH-dCyd, *N*(4)-hydroxy-2′-deoxycytidine; *N*(4)-OH-FdCyd, *N*(4)-hydroxy-5-fluoro-2′-deoxycytidine; 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd, 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-2′-deoxycytidine; 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd, 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-5-fluoro-2′-deoxycytidine; mTHF, N^5,N^{10}-methylenetetrahydrofolate.

[1–4]. Thymidylate synthesis is the terminal step in the sole *de novo* biosynthetic pathway leading to dTMP. Consequently, TS inhibition blocks DNA synthesis and prevents cell proliferation. Therefore, the enzyme has long been recognized as a critical target for anticancer therapy [5]. Many compounds modeled after the substrate (dUMP) or cofactor (mTHF) have been tested as TS inhibitors and some are on clinical trial or on the market as drugs in anticancer, antiviral or antifungal chemotherapy [6–8].

Among dUMP analogues that have been studied to date, the most potent inhibitors are those with an electron-withdrawing substituent at the pyrimidine C(5) carbon [9–11]. In this group, 5-fluoro-dUMP (FdUMP) is the only compound, which has been approved for and is routinely used in chemotherapy (in the form of a pro-drug, 5-fluorouracil) [12, 13]. Inhibition of TS by FdUMP involves a time-dependent formation of the ternary covalent complex of TS with FdUMP and mTHF, resulting in the reaction being arrested, as the fluorine substituent fails to dissociate from the pyrimidine ring. This causes a slowly reversible inactivation of the enzyme.

A rare example of a dUMP analogue being a strong TS inhibitor in spite of lack of modification at C(5), is *N*(4)-hydroxy-2′-deoxycytidine 5′-monophosphate (*N*(4)-OH-dCMP), with the pyrimidine C(4) carbon substituted by the N(4)-OH residue. When used to inhibit the enzyme, *N*(4)-OH-dCMP demonstrates a similar behavior to FdUMP, causing in the presence of mTHF a time-dependent inactivation of the enzyme [14–16]. 5-Fluoro substitution in *N*(4)-OH-dCMP has been shown to potentiate the inhibition, with 5-fluoro-*N*(4)-OH-dCMP (*N*(4)-OH-FdCMP) probably retaining the same inhibitory mechanism as the parent compound [14].

The first interpretations of the effect of stronger affinity of *N*(4)-OH-FdCMP to TS as compared to *N*(4)-OH-dCMP pointed to a hypothetical intramolecular hydrogen bond between the fluorine atom and the *N*(4)-OH residue as the cause of the proposed better stabilization of the rotameric *imino-anti* (*trans*) against the *imino-syn* (*cis*) form in *N*(4)-OH-FdCMP [14]. That interpretation is based on the observation that only the *trans* form of *N*(4)-OH-dCMP actively inhibits thymidylate synthase. However, the subsequent *ab initio* studies proved that the intramolecular hydrogen bond would be too weak to considerably improve the internal stability of *N*(4)-hydroxy-5-fluorocytosine [17]. Furthermore, recent results of the molecular mechanics calculations indicated very similar populations of the *trans* conformer in 1-methyl-*N*(4)-hydroxycytosine and its C(5)-fluoro derivative [18]. Thus, although some interplay between the C(4)- and C(5) substituents may exist, the true reasons for a stronger inhibitory potency of the 5-fluoro congener of *N*(4)-OH-dCMP against TS may likely lie elsewhere.

In order to learn about those reasons, crystallographic and theoretical comparative studies of 3′,5′-di-*O*-acetyl derivatives of *N*(4)-hydroxy-2′-deoxycytidine (*N*(4)-OH-dCyd) and *N*(4)-hydroxy-5-fluoro-2′-deoxycytidine (*N*(4)-OH-FdCyd) were made. The results indicate that a stronger affinity of *N*(4)-OH-FdCMP for TS may originate from the structural changes caused locally by the C(5)-fluoro substituent, rather than interactions between the C(4)–N–OH and C(5)–F groups. These changes involve the C(5)═C(6) double bond shortening and C(4)—C(5)—C(6) angle widening, producing a local strain in the pyrimidine ring and rendering the C(5)═C(6) bond more susceptible to the nucleophilic addition of thymidylate synthase active site cysteine residue, an initial step in the enzyme catalysis.

METHODS

X-Ray Diffraction

3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-2′-deoxycytidine (3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd) and 3′,5′-di-*O*-acetyl-*N*(4)-hydroxy-5-fluoro-2′-deoxycytidine (3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd) were synthesized according to the protocol described in Felczak *et al*. [18]. Crystals of the both compounds were grown at room temperature by slow evaporation of the solvent containing saturated methanol and ethanol mixed at the molar ratio close to 1:2. Colorless crystals of X-ray diffraction quality were formed after 2 days.

Single crystal measurements were carried out at room temperature on the synchrotron X31 EMBL beam line at DESY/Hamburg, using a MAR IP detector and a radiation wavelength of 0.702 Å. The data were collected using a $\omega/2\theta$ scan technique and corrected for Lorentz and polarization effects but no absorption correction was applied. A total of 108201 and 17349 reflections were measured, and 12211 and 2583 reflections were found unique, for 3′,5′-di-*O*-acetyl derivatives of *N*(4)-OH-dCyd and *N*(4)-OH-FdCyd, respectively.

Both structures were solved by direct methods using SHELXS [19] and refined by the full-matrix least-squares optimization on F^2 over all unique reflections with SHELXL-97 [20]. All the non-hydrogen atoms were refined anisotropically, whereas the hydrogen atoms were placed in calculated positions and their thermal parameters were refined isotropically. Atomic scattering factors were taken from International Tables for Crystallography (1992, Volume C) [21]. Anomalous dispersion

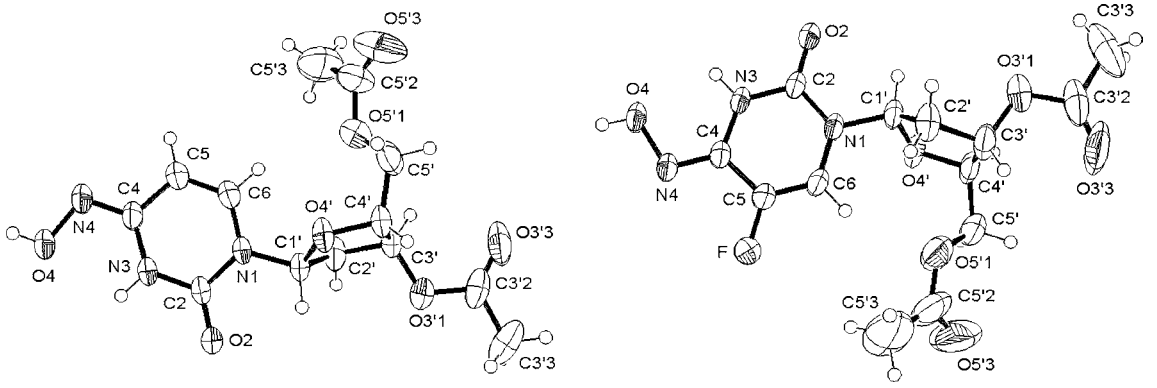

Fig. 1. ORTEP diagram at the 50% probability level and atom numbering for 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd (left) and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd. One of the five symmetry-independent molecules of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd is shown.

factors were calculated with the program FPRIME [22, 23]. In the course of refinement, the heavy atoms of some of the acetyl ($-COCH_3$) groups displayed large displacement parameters (see Figs. 1 and 4), indicating a possibility of disorder. Attempts were made to refine those groups as disordered but these models did not improve the fit of diffraction data. In the final refinement cycles R_1 converged to 0.0571 and 0.0497 for 3′,5′-di-*O*-acetyl derivatives of *N*(4)-OH-dCyd and *N*(4)-OH-FdCyd, respectively.

Crystal data and structure refinement are summarized in Table I. The absolute configurations could not be experimentally determined with the used wavelength due to the absence of heavier atoms (only the first-row atoms were present). The natural D-dCyd enantiomers were thus assigned and are reported. Table II lists selected bond lengths and angles for the non-hydrogen atoms. The software used to prepare the material for publication and the CIF file (see supplementary material) were the WinGX program package [24] and the enCIFer program [25]. Drawings were prepared with the ORTEP-3 for Windows program [26] (Figs. 1, 3 and 4), or with the ISIS/Draw program [27] (Figs. 2 and 5).

Ab Initio Modeling

The starting point for theoretical calculations were the crystallographic coordinates (converted to Cartesian format) of the molecules of 3′,5′-di-*O*-acetyl derivatives of *N*(4)-OH-dCyd and *N*(4)-OH-FdCyd. The geometry optimization was done *ab initio* at the RHF level, involving initial optimization using the standard 3-21G basis set, followed by re-optimization using the 6-31G(d) basis set, which includes polarization functions on the non-hydrogen atoms. The calculations were performed with the GAMESS-US version May 19, 2004 (R3) program [28] using the default convergence criteria. The RHF/6-31G(d) harmonic vibrational frequencies were also calculated to ascertain that the geometries converged to the stable ones (no negative/imaginary frequencies were found; results not shown). The selected geometric parameters resulting from the optimizations are listed in Table III. A single set of parameters is reported for the non-fluorine analogue, since its five crystallographic molecules (see section "Crystal Structures") converged to almost identical geometries.

RESULTS AND DISCUSSION

Crystal Structures

An ORTEP view of the 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd molecules with thermal ellipsoids and labeling scheme is shown in Fig. 1. All of the sugar moieties are envelopes, with puckering mode being C(2′)-endo in the 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd molecules *A*, *B*, *C*, *D*, and *E*, and in the molecule of 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd. The conformation with respect to the glycosidic bond between the sugar and the base is *anti* in all the molecules of the both analogues. The corresponding O(4′)–C(1′)–N(1)–C(2) torsion angles are listed in Table IV.

In the structures of both compounds the *N*(4)-hydroxycytosine moieties exist in the *imino* *N*(3)-H tautomeric form, which is in agreement with the findings from earlier studies in the gas phase ([17], and references therein), in solution ([29], and references therein), and in the crystalline state [30]. All the base rings are planar, with the *N*(4)-OH groups lying in the ring planes. The conformations of the *N*(4)-OH group are

Table I. Crystal Data and Structure Refinement for 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd

Crystal data	3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd	3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd
CCDC deposit no.	CCDC 258521	CCDC 258522
Empirical formula	$(C_{13}H_{17}N_3O_7)_5$	$C_{13}H_{16}FN_3O_7$
Formula weight	1636.48	345.29
Crystal color, habit	Colorless, needle	Colorless, needle
Crystal size (mm)	0.5 × 0.2 × 0.1	0.5 × 0.2 × 0.1
Crystal system	Monoclinic	Monoclinic
Space group	*C*2	*C*2
Unit cell dimensions		
a (Å)	39.54(5)	21.72(3)
b (Å)	8.720(10)	8.720(10)
c (Å)	22.89(3)	8.610(10)
β (°)	95.26(10)	90.42(10)
Volume (Å^3)	7859 (17)	1631 (3)
Z	4	4
Density calculated (g/cm^3)	1.383	1.406
Temperature (K)	293 (2)	293 (2)
Wavelength (Å)	0.702	0.702
Absorption coefficient (mm^{-1})	0.075	0.081
F(000)	3440	720
Index ranges	$0 \le h \le 55$	$0 \le h \le 30$
	$0 \le k \le 12$	$0 \le k \le 12$
	$-32 \le l \le 32$	$-12 \le l \le 12$
θ range (°)	0.88–29.75	1.85–29.92
Reflections collected	108201	17349
Independent reflections	12211	2583
Refinement method		Full-matrix least-squares on F^2
Data/restraints/parameters	12211/1/1130	2583/1/236
Goodness-of-fit on F^2	1.039	1.037
Final *R* indices [$I > 2\sigma(I)$]	$R_1 = 0.0571$, $wR_2 = 0.1704$	$R_1 = 0.0497$, $wR_2 = 0.1377$
Final *R* indices (all data)	$R_1 = 0.0599$, $wR_2 = 0.1763$	$R_1 = 0.0508$, $wR_2 = 0.1398$
Extinction coefficient	0.00045(17)	0.006(2)
Largest diff. peak (e/Å^3)	0.761	0.284
Largest diff. hole (e/Å^3)	−0.290	−0.229

cis (Fig. 2), with the hydroxyl pointed toward the N(3) atom of the base in all the molecules of 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd (for values of the torsion angles N(3)—C(4)—N(4)—O(4) and C(5)—C(4)—N(4)—O(4) see Table IV). This preference for the *cis* form corresponds with the available *ab initio*, spectroscopic, molecular mechanics, and crystallographic data. The *ab initio* RHF 6/31G** calculations showed the *cis* form to have significantly lower energy compared to the *trans* form (Fig. 2) for the molecules of *N*(4)-hydroxycytosine and its 5-fluoro congener [17]. In accordance with the latter, the molecule of 1-methyl-*N*(4)-hydroxycytosine, isolated in low-temperature inert gas matrixes, was found from IR spectra to be in the *cis* form [31]. In addition, a series of various 5-substituted 1-methyl-*N*(4)-hydroxycytosines, investigated using molecular mechanics, exhibited invariantly very low populations of the *trans* form, with the hydroxyl pointed toward C(5), relatively to much higher populations of the *cis* conformer [18]. The population of the *trans* form at 298 K was estimated to be 2×10^{-2} for both 1-methyl-*N*(4)-hydroxycytosine and 1-methyl-*N*(4)-hydroxy-5-fluorocytosine. Finally, the only entry in the Cambridge Structural Database (CSD) corresponding to the class of 5-substituted [*N*(4)-hydroxy]-cytosines/-2′-deoxycytidines, 1,5-dimethyl-*N*(4)-hydroxycytosine [30], is also the *cis* conformer, similar to the structures reported here.

The hydrogen bonding patterns are similar in the 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd structures. There are no intramolecular H-bonds, and the only intermolecular H-bonds are those stabilizing the conformation of bases. Each molecule in the both structures appears to participate in four such contacts with two neighboring molecules, forming two types of hydrogen bonds, O(4)—H(4)···O(2) (with the H(4)···O(2) distance of 1.87–1.99 Å, and the O(4)—H(4)···O(2) angle of 152.1–162.0°) and

Table II. Selected Bond Distances and Angles in the Crystal Structures of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd

	FdCyd	dCyd-A	dCyd-B	dCyd-C	dCyd-D	dCyd-E	Mean of *A*–*E*
Bond distances (Å)							
N(1)—C(2)	1.360(3)	1.367(3)	1.371(3)	1.362(4)	1.363(3)	1.358(4)	1.364
C(2)—N(3)	1.368(3)	1.365(3)	1.360(3)	1.370(3)	1.367(3)	1.363(3)	1.365
N(3)—C(4)	1.380(3)	1.375(4)	1.380(4)	1.389(4)	1.370(4)	1.383(4)	1.379
C(4)—C(5)	1.435(3)	1.428(4)	1.433(4)	1.425(4)	1.419(4)	1.427(4)	1.426
C(5)—C(6)	1.320(3)	1.332(4)	1.329(4)	1.314(4)	1.338(4)	1.326(4)	1.328
C(6)—N(1)	1.383(3)	1.384(4)	1.375(4)	1.394(4)	1.375(4)	1.374(5)	1.380
C(2)—O(2)	1.226(3)	1.216(4)	1.226(4)	1.221(4)	1.222(4)	1.219(4)	1.221
C(4)—N(4)	1.287(3)	1.277(4)	1.283(3)	1.279(3)	1.286(4)	1.288(3)	1.283
N(4)—O(4)	1.400(3)	1.413(4)	1.406(3)	1.411(3)	1.406(4)	1.412(4)	1.410
C(5)—F	1.334(3)	—	—	—	—	—	—
Bond angles (°)							
C(2)—N(1)—C(6)	121.6(2)	121.7(2)	121.1(2)	121.3(2)	121.2(2)	121.4(2)	121.3
C(2)—N(1)—C(1′)	119.5(2)	119.9(2)	119.2(2)	118.7(3)	118.8(2)	117.0(3)	118.7
C(6)—N(1)—C(1′)	118.9(2)	118.4(2)	119.5(2)	120.0(2)	120.0(2)	120.5(2)	119.7
O(2)—C(2)—N(1)	122.2(2)	122.2(2)	121.7(2)	122.0(2)	122.4(2)	121.7(2)	122.0
O(2)—C(2)—N(3)	121.8(2)	121.8(2)	122.2(2)	121.9(3)	121.3(2)	121.7(3)	121.8
N(1)—C(2)—N(3)	116.0(2)	116.1(2)	116.1(2)	116.1(3)	116.3(3)	116.5(3)	116.2
C(2)—N(3)—C(4)	126.0(2)	124.8(2)	124.9(2)	124.8(2)	124.9(2)	124.5(2)	124.8
N(3)—C(4)—N(4)	123.6(2)	123.9(3)	123.7(2)	124.0(2)	123.3(3)	123.6(3)	123.7
C(5)—C(4)—N(4)	122.3(2)	119.4(3)	119.7(3)	120.2(3)	120.2(3)	120.4(3)	120.0
N(3)—C(4)—C(5)	114.1(2)	116.8(3)	116.7(2)	115.9(2)	116.4(2)	116.1(2)	116.4
C(4)—N(4)—O(4)	108.2(2)	111.2(3)	110.5(2)	110.5(2)	109.9(3)	109.8(3)	110.4
C(4)—C(5)—C(6)	121.1(2)	119.0(3)	118.1(3)	119.9(3)	119.2(3)	119.2(3)	119.1
C(5)—C(6)—N(1)	121.1(2)	121.6(3)	123.1(2)	121.9(3)	121.9(3)	122.2(3)	122.1
C(4)—C(5)—F	117.6(2)	—	—	—	—	—	—
C(6)—C(5)—F	121.3(2)	—	—	—	—	—	—

N(3)—H(3)···N(4) (with the H(3)···N(4) distance of 2.15–2.31 Å, and the N(3)—H(3)···N(4) angle of 151.3–153.3°), each molecule participating in both H-bonding pairs as both donor and acceptor. The fluorine atom in 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd appears unlikely to be involved in hydrogen bond as its closest potential intermolecular contact, which is to the O(4)—H(4) proton, has an unfavorable distance (d(H(4)···F) = 2.71 Å) and angular orientation (∠(O(4)—H(4)···F) = 90.4°) for hydrogen bond formation. On the other hand, the positions of the hydrogen atoms could not be reliably determined from difference Fourier maps and were generated geometrically, and thus, in fact, can be to some extent different, meaning that the existence of the O(4)—H(4)···F hydrogen bond cannot be definitely excluded.

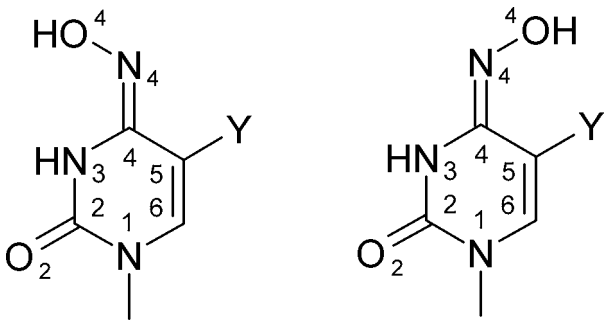

Fig. 2. The *cis* (left) and *trans* orientations of the N(4)-OH group with respect to the pyrimidine ring.

The 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd structure represents a rarely encountered example of a crystallographic structure with five molecules in the asymmetric part of unit cell. The molecular packing in the unit cell is shown in Fig. 3. The arrangement in the asymmetric unit (depicted in Fig. 4) is such that the five molecules spread along three planes approximately parallel to the *a*-axis as follows: two molecules per plane along the right (molecules *A* and *C*) and middle (molecules *D* and *E*) planes and the molecule *B* along the left plane. This arrangement is secured *via* the O(4)—H(4)···O(2) and N(3)—H(3)···N(4) hydrogen bonding, linking the molecules: *A* with *C*, and *D* with *E*. The differences in respective geometric parameters among the molecules *A*–*E* are not large, with smaller differences between the corresponding angles than bond lengths (*cf.* Table II). The largest variations concern the C(5′2)—C(5′3) and C(5′2)—O(5′3) bond lengths in the —O-Ac substituent on C(5′) (*cf.* Table V). This effect

Table III. Selected Bond Distances and Angles in the *Ab Initio* Structures of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd

	FdCyd	dCyd
Bond distances (Å)		
N(1)—C(2)	1.368	1.372
C(2)—N(3)	1.370	1.366
N(3)—C(4)	1.382	1.383
C(4)—C(5)	1.455	1.454
C(5)—C(6)	1.318	1.326
C(6)—N(1)	1.391	1.386
C(2)—O(2)	1.200	1.201
C(4)—N(4)	1.259	1.262
N(4)—O(4)	1.384	1.389
C(5)—F	1.323	—
Bond angles (°)		
C(2)—N(1)—C(6)	121.4	121.1
C(2)—N(1)—C(1′)	116.6	116.3
C(6)—N(1)—C(1′)	122.0	122.6
O(2)—C(2)—N(1)	122.7	122.4
O(2)—C(2)—N(3)	121.6	121.9
N(1)—C(2)—N(3)	115.6	115.7
C(2)—N(3)—C(4)	127.0	126.5
N(3)—C(4)—N(4)	124.8	123.7
C(5)—C(4)—N(4)	121.8	121.5
N(3)—C(4)—C(5)	113.4	114.8
C(4)—N(4)—O(4)	110.7	110.8
C(4)—C(5)—C(6)	120.9	119.2
C(5)—C(6)—N(1)	121.6	122.7
C(4)—C(5)—F	117.4	—
C(6)—C(5)—F	121.7	—

results from considerable thermal motions of some of the acetyl groups, reflected in high thermal displacement parameters for the atoms of those groups (see Figs. 1 and 4).

Structural Consequences of Fluorine Substitution in View of TS-Catalyzed Reaction Mechanism

In addition to the molecular geometries obtained in the refinement of the crystal structures of the 3′,5′-di-*O*-acetyl derivatives of *N*(4)-OH-dCyd and *N*(4)-OH-FdCyd, one set of the optimized geometries of isolated molecules has been derived from the theoretical *ab initio* RHF/6-31G(d) calculations. Both sets, the crystallographic and *ab initio* molecular geometries have been analyzed and compared in the search for structural differences between the fluorine and non-fluorine analogues, as that could help to explain the origin of more potent inhibition of thymidylate synthase caused by 5-fluoro-*N*(4)-OH-dCMP in comparison to *N*(4)-OH-dCMP. However, it would be a highly impractical challenge to compare each of the five values of a given parameter in 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd with that parameter single value in 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd. Therefore, for the sake of compromise and simplicity, each bond length and angle in 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd was assigned the mean of the five measured values (see Table II) and only then put into the comparison.

The geometry analysis in this study has been done assuming that structural dependencies observed in cytidine-based *anti* nucleosides can model the activity functions in the corresponding *anti* nucleotides in the TS reaction. The assumption appears reasonable, as the reaction of TS, in which covalent bonds with nucleotide are formed, is limited to the nucleotide base ring, whose geometry may be expected to be negligibly influenced by the presence (or lack of) of the phosphate moiety or the -O-Ac substituents of the 2′-deoxyribose ring. For further discussion in this regard, see Jarmuła *et al*. [32].

The experimental solid-state (crystallographic) and the theoretical isolated-molecule (*ab initio*) geometries are, in general, in good agreement. They show no substantial differences between respective bond lengths or angles (*cf*. Table II vs. Table III), although some parameters differ relatively more than the rest (*cf*. the C(2)═O(2), C(4)═N(4), C(4)—C(5), and N(4)—O(4) bond lengths). In addition, the *ab initio* parameters corresponding to the pyrimidine ring and the N(4)-OH residue have been compared and are in very good agreement with the SCF/6-31G** molecular geometries (not shown here) calculated for *N*(4)-hydroxycytosine and its C(5)-fluoro derivative [17].

In view of both the experimental (Table II) and theoretically optimized (Table III) molecular geometries of

Table IV. Selected Torsion Angles (°) in the Crystal Structures of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd

	FdCyd	dCyd-A	dCyd-B	dCyd-C	dCyd-D	dCyd-E
O(4′)—C(1′)—N(1)—C(2)	−129.7(2)	−121.9(3)	−123.2(2)	−133.6(3)	−123.6(2)	−153.5(3)
O(4′)—C(1′)—N(1)—C(6)	49.2(3)	55.6(4)	51.6(3)	47.7(4)	53.9(3)	38.5(4)
C(2′)—C(1′)—N(1)—C(2)	112.9(3)	121.1(3)	120.0(3)	109.2(3)	118.8(3)	89.7(3)
C(2′)—C(1′)—N(1)—C(6)	−68.2(3)	−61.4(4)	−65.2(3)	−69.5(4)	−63.6(3)	−78.3(4)
N(3)—C(4)—N(4)—O(4)	−0.4(3)	1.0(5)	0.9(4)	0.0(4)	0.5(4)	−0.3(4)
C(5)—C(4)—N(4)—O(4)	178.9(2)	−178.9(3)	179.8(2)	−179.5(2)	178.5(3)	−179.7(2)

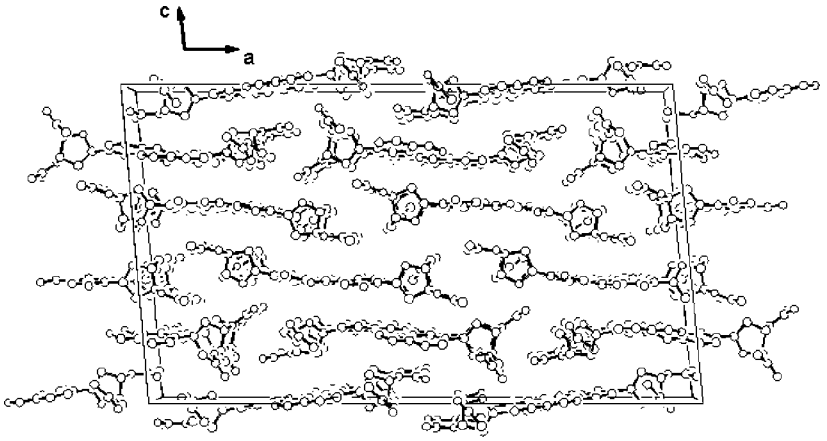

Fig. 3. Crystal packing of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd viewed normal to the *ac*-plane. Hydrogen atoms are omitted for clarity.

3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd, the corresponding bond distances and angles are quite similar in both compounds. This includes the substituted sugar ring (with the exception of the acetyl groups in the crystal structures for reasons mentioned earlier in the paper; sugar parameters not shown), as well as the base ring and the N(4)-OH residue (with some exceptions discussed later). However, in spite of the overall good agreement between the two pyrimidine ring geometries, the vicinity of C(5) shows certain specific structural differences, clearly resulting from the presence of alternative C(5) substituents: hydrogen or fluorine.

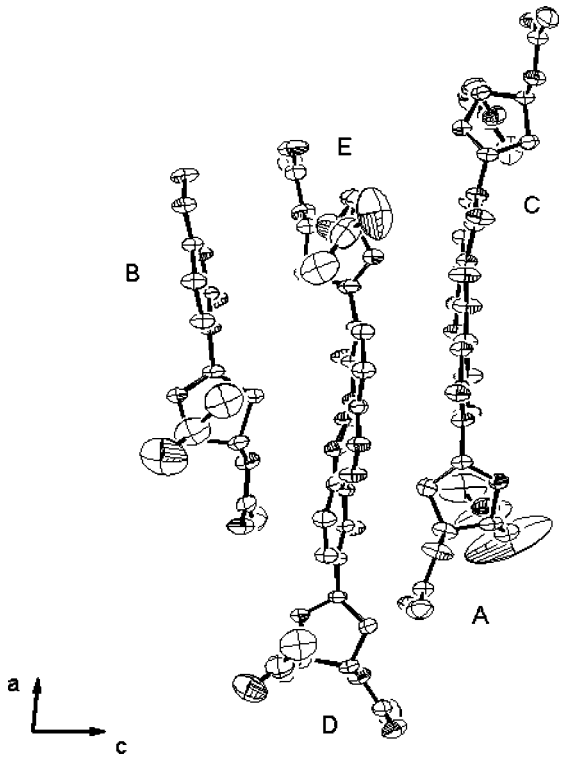

Fig. 4. Asymmetric unit of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd viewed down the *b*-axis. Five symmetry-independent molecules *A*–*E* are shown. Thermal ellipsoids are drawn at the 40% probability level. Hydrogen atoms are omitted for clarity.

The effect of fluorine atom substitution on the pyrimidine ring structure can most sensibly be assessed by comparing the *ipso*-angle C(4)—C(5)—C(6) in the molecules of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd. The corresponding values are 119.1° and 121.1° (experimental) or 119.2° and 120.9° (theoretical), respectively, showing the C(4)—C(5)—C(6) angle to be considerably widened in response to the fluorine substitution. The widening of the C(4)—C(5)—C(6) angle affects the nearby angles C(5)—C(4)—N(4), C(5)—C(4)—N(3), C(4)—N(4)—O(4), C(4)—N(3)—C(2) and C(5)—C(6)—N(1), causing relaxation of their geometries and resulting in clearly different angle values from one analogue to the other (observed in crystal and also, but to a smaller extent, in *ab initio* structures; *cf.* Table II vs. Table III, respectively). Such changes in angle values may be expected to be coupled with changes in neighboring bond lengths. In a classic example of the benzene ring, an electronegative substituent such as the fluorine atom would affect the ring by causing cooperative structural changes involving increase of the endocyclic angle at the substituted carbon and equal shortening of the two equivalent bonds formed by this carbon with neighbor atoms in the ring (Fig. 5). This effect has long been known and several explanations with various theoretical backgrounds, including the Walsh–Bent rule [33–35] and the valence-shell electron-pair repulsion theory [36–38],

Table V. Bond Lengths (Å) with Largest Differences in Value Across Five Independent Molecules in the Crystal Structure of 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd

	dCyd-A	dCyd-B	dCyd-C	dCyd-D	dCyd-E	Mean
C(5′2)—C(5′3)	1.466(12)	1.504(10)	1.497(8)	1.484(9)	1.498(11)	1.490
C(5′2)—O(5′3)	1.241(18)	1.185(8)	1.195(5)	1.181(6)	1.157(9)	1.192

have been offered. However, in the pyrimidine ring, unlike in the benzene ring, a uniform electron distribution around the ring is not preserved. The factor analysis of a sample consisting of C(5)–halogen–uracil derivatives extracted from CSD showed that the C(4)—C(5) bond length was influenced from different factors and weaker than the C(5)═C(6) bond length (see Fig. 5) [32, 39]. Comparison of the molecules of 3′,5′-di-*O*-acetyl derivatives of *N*(4)-OH-dCyd and *N*(4)-OH-FdCyd shows a similar influence of the fluorine atom on the bonds adjacent to C(5). The C(5)═C(6) bond distances in 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd are 1.328 and 1.320 Å (X-ray diffraction), or 1.326 and 1.318 Å (*ab initio* study), respectively, showing a clear trend of shortening of the C(5)═C(6) bond in response to the C(5)-fluoro substitution. The distances of the second adjacent bond, C(4)—C(5), are 1.426 and 1.435 Å (X-ray diffraction), or 1.454 and 1.455 Å (*ab initio* study) in 3′,5′-di-*O*-Ac-*N*(4)-OH-dCyd and 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd, respectively. They show that the C(4)—C(5) bond becomes longer (crystal structure) or remains almost unaltered (*ab initio* structure) in response to the C(5)-fluoro substitution, thus indicating a different influence of the fluorine atom on the C(4)—C(5) than C(5)═C(6) bond. Other most differing parameters include the N(4)—O(4) bond length, which according to the crystal structures is by 0.01 Å longer in the non-fluorine analogue (*cf.* Table II). The latter is weaker reflected in the theoretical structures, where this difference amounts to only 0.005 Å (*cf.* Table III). The shortening of the N(4)—O(4) bond in the crystal structure of 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd may originate from repulsive interactions in the unit cell between the F and O(4) atoms, which are located within a distance of 2.84 Å from each other, i.e. closer than their van der Waals distance.

The opening of the C(4)—C(5)—C(6) angle together with the shortening of the C(5)═C(6) bond, observed in response to the C(5)-fluoro substitution, results in a local strain in the molecule of 3′,5′-di-*O*-Ac-*N*(4)-OH-FdCyd. That strain strengthens a potential reactivity of the C(5)═C(6) bond. Assuming the same effect takes place in the pyrimidine ring of the molecule of the fluorine-substituted *N*(4)—OH—FdCMP nucleotide, it may explain a more potent inhibitory activity against TS demonstrated by *N*(4)—OH—FdCMP compared to the non-fluorine parent analogue, *N*(4)—OH—dCMP. This explanation refers to the known mechanism of TS reaction, where the nucleophilic attack on the pyrimidine C(6) atom is a first step in enzyme catalysis [40, 1–4]. It is worth noting that a similar effect of the C(5)-fluoro substituent has been previously observed and proposed by us as an explanation for enhanced affinities to TS demonstrated by the 5-fluoro-dUMP and 2,4-dithio-5-fluoro-dUMP inhibitors, as compared to the respective analogues, dUMP and 2,4-dithio-dUMP [32, 39].

In conclusion, a higher inhibitory potency of *N*(4)—OH—FdCMP than *N*(4)—OH—dCMP toward TS-catalyzed reaction may originate from the structural changes caused locally by the C(5)-fluoro substituent, rather than interactions between the C(4)—N—OH and C(5)—F groups.

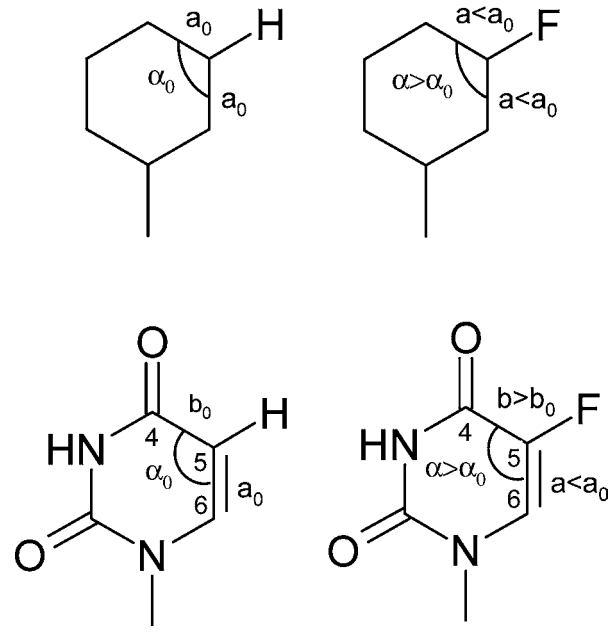

Fig. 5. The effect of the fluorine substituent on local geometry in the benzene ring (top) and in the uracil ring (bottom). Note that in the uracil ring the effect on *a* (the C(5)═C(6) bond length) is opposite to that on *b* (the C(4)—C(5) bond length).

SUPPLEMENTARY MATERIAL

CCDC-258521 and CCDC-258522 contain the supplementary crystallographic data for this paper. These data can be obtained free of charge at www.ccdc.cam.ac.uk/conts/retrieving.html [or from the Cambridge Crystallographic Data Centre (CCDC), 12 Union Road, Cambridge CB2 1EZ, UK; fax: +44(0)1223–336033; e-mail: deposit@ccdc.cam.ac.uk].

ACKNOWLEDGMENTS

We would like to thank Prof. Krzysztof Woźniak for his advice in projecting the crystallographic experiment. Dr. Michał K. Cyrański and Mrs. Paulina Rogowska helped us in searching the CSD and retrieving its entries.

REFERENCES

1. Santi, D. V.; Danenberg, P. V. In *Folates and Pterines*; Blakely, R. L.; Benkovic, S. J., Eds.; Wiley: New York, 1984; Vol. 1, *Chemistry and Biochemistry of Folates*; pp. 345–398.
2. Carreras, C. W.; Santi, D. V. *Annu. Rev. Biochem.* **1995,** *64*, 721.
3. Stroud, R. M.; Finer-Moore, J. S. *Biochemistry* **2003,** *42*, 239.
4. Finer-Moore, J. S.; Santi, D. V.; Stroud, R. M. *Biochemistry* **2003,** *42*, 248.
5. Danenberg, P. V. *Biochim. Biophys. Acta* **1977,** *473*, 73.
6. Heidelberger, C.; Danenberg, P. V.; Moran, R. G. *Adv. Enzymol.* **1983,** *54*, 57.
7. Jackman, A. L.; Jones, T. R.; Calvert, A. H.; In *Experimental and Clinical Progress in Cancer Chemotherapy*; Muggia, F. M., Ed.; Martinus Nijhoff Publishers: Boston, 1985; pp. 155–210.
8. Ealick, S. E.; Armstrong, S. R. *Curr. Opin. Struct. Biol.* **1993,** *3*, 861.
9. Lewis, C. A., Jr.; Dunlap, R. B. In *Topics in Molecular Pharmacology*; Burgen, A. S. V.; Roberts, G. C. K., Eds.; Elsevier/North-Holland Biomedical: New York, 1981; pp. 170–219.
10. Santi, D. V. *J. Med. Chem.* **1980,** *23*, 103.
11. De Clercq, E.; Balzarini, J.; Torrence, P. F.; Mertes, M. P.; Schmidt, C. L.; Shugar, D.; Barr, P. J.; Jones, A. S.; Verhelst, G.; Walker, R. T. *Mol. Pharmacol.* **1981,** *19*, 321.
12. Hartman, K. R.; Heidelberger, C. J. *Biol. Chem.* **1958,** *236*, 3006.
13. Heidelberger, C. *Prog. Nucl. Acid Res. Mol. Biol.* **1965,** *4*, 1.
14. Rode, W.; Zieliński, Z.; Dzik, J. M.; Kulikowski, T.; Bretner, M.; Kierdaszuk, B.; Cieśla, J.; Shugar, D. *Biochemistry* **1990,** *29*, 10835.
15. Lorenson, M. Y.; Maley, G. F.; Maley, F. *J. Biol. Chem.* **1967,** *242*, 3332.
16. Goldstein, S.; Pogolotti, A. L.; Garvey, E. P., Jr.; Santi, D. V. *J. Med. Chem.* **1984,** *27*, 1259.
17. Leś, A.; Adamowicz, L.; Rode, W. *Biochim. Biophys. Acta* **1993,** *1173*, 39.
18. Felczak, K.; Miazga, A.; Poznański, J.; Bretner, M.; Kulikowski, T.; Dzik, J. M.; Gołos, B.; Zieliński, Z.; Cieśla, J.; Rode, W. *J. Med. Chem.* **2000,** *43*, 4647.
19. Sheldrick, G. M. *Acta Crystallogr.* **1990,** *A46*, 467.
20. Sheldrick, G. M. SHELXL-97, *Program for Crystal Structure Refinement*; University of Gottingen: Germany, 1997; Release 97-2.
21. Wilson, A. J. C., Ed. *International Tables for Crystallography*; Kluwer Academic Publishers: Dordrecht, 1992; Volume C.
22. Kissel, L. P.; Pratt, R. H. *Acta Crystallogr.* **1990,** *A46*, 170.
23. Cromer, D. T. *J. Appl. Crystallogr.* **1983,** *16*, 437.
24. Farrugia, L. J. *J. Appl. Crystallogr.* **1999,** *32*, 837.
25. Allen, F. H.; Johnson, O.; Shields, G. P.; Smith, B. R.; Towler, M. *J. Appl. Crystallogr.* **2004,** *37*, 335.
26. Farrugia, L. J. *J. Appl. Crystallogr.* **1997,** *30*, 565.
27. ISIS/Draw; MDL Information Systems, Inc., 1990–2001; Ver. 2.4.
28. Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993,** *14*, 1347.
29. Niedźwiecka-Hornaś, A.; Kierdaszuk, B.; Stolarski, R.; Shugar, D. *Biophys. Chem.* **1998,** *71*, 87.
30. Shugar, D.; Huber, C. P.; Birnbaum, G. I. *Biochim. Biophys. Acta* **1976,** *447*, 274.
31. Stepanenko, T.; Lapinski, L.; Sobolewski, A. L.; Nowak, M. J.; Kierdaszuk, B. J. *Phys. Chem. A* **2000,** *104*, 9459.
32. Jarmuła, A.; Anulewicz, R.; Leś, A.; Cyrański, M. K.; Adamowicz, L.; Bretner, M.; Felczak, K.; Kulikowski, T.; Krygowski, T. M.; Rode, W. *Biochim. Biophys. Acta* **1998,** *1382*, 277.
33. Walsh, A. D. *Discuss. Faraday Soc.* **1947,** *2*, 18.
34. Bent, H. A. *Chem. Rev.* **1961,** *61*, 275.
35. Bent, H. A. *J. Inorg. Nucl. Chem.* **1961,** *19*, 43.
36. Gillespie, R. J.; Nyholm, R. S. *Quart. Rev.* **1957,** *11*, 339.
37. Gillespie, R. J. *Molecular Geometry*; Van-Nostrand-Reinhold: London, 1972.
38. Gillespie, R. J.; Hargittai, I. In *The VSEPR Model of Molecular Geometry*; Allyn & Bacon: Needham Heights, MA, 1991.
39. Jarmuła, A.; Cyrański, M. K.; Leś, A.; Krygowski, T. M.; Rode, W. *Pol. J. Chem.* **1998,** *72*, 1958.
40. Matthews, D. A.; Villafranca, J. E.; Janson, C. A.; Smith, W. W.; Welsh, K.; Freer, S. *J. Mol. Biol.* **1990,** *214*, 937.