CHEMICAL & PHARMACEUTICAL BULLETIN

Vol. 33, No.2 February 1985

Regular Articles

Chem. Pharm. Bull.
33(2) 455—463 (1985)

A Stacking Interaction between Thymine and Phenyl Rings: Crystal Structure of 3′-*O*-Benzoylthymidine Monohydrate

YURIKO YAMAGATA,*,[a] HIROTSUGU MIYASHIRO,[a] KEN-ICHI TOMITA,[a]
TOSHIMASA ISHIDA[b] and MASATOSHI INOUE[b]

Faculty of Pharmaceutical Sciences, Osaka University,[a] Yamadaoka, Suita, Osaka 565, Japan and Osaka College of Pharmacy,[b] 2-10-65 Kawai, Matsubara, Osaka 580, Japan

(Received June 25, 1984)

The crystal structure of 3′-*O*-benzoylthymidine has been determined by X-ray diffraction analysis. The crystals are monoclinic, space group $P2_1$ with the following cell dimensions: $a=12.528$ (6), $b=6.544$ (2), $c=10.770$ (5) Å, $\beta=102.80$ (5)°. The structure was solved by a combination of the Patterson and direct methods and refined to a crystallographic R factor of 0.053 by full-matrix least-squares refinement. The thymine and phenyl rings of neighboring molecules associate through stacking interactions. The glycosidic torsion angle ($\chi_{CN}=49.3$ (4)°) is in the *anti* region and the deoxyribose ring conformation is C(2′)-*endo*, while the conformation about the C(4′)–C(5′) bond is *gauche–gauche*.

Keywords——benzoyl-thymidine; stacking interaction; nucleoside conformation; X-ray analysis

Introduction

3′-*O*-Benzoylthymidine is utilized as a building block for the chemical synthesis of oligodeoxynucleosides in which the 3′-terminal nucleoside is a thymidine. Apart from its usefulness for chemical synthesis, this compound may serve as a model for investigations of the interactions between the nucleic acid base thymine and the aromatic side group of the amino acids phenylalanine and tyrosine. Nucleic acid–protein interactions are of considerable importance in molecular biology, but are not yet very well understood at the atomic level. It is well known that the mutual recognition of proteins and nucleic acids involves electrostatic, hydrogen bonding and stacking interactions. For example, the stacking interaction between the nucleic acid base guanine and the aromatic amino acid tyrosine is considered to play an important role in specific protein–nucleic acid recognition on the basis of the crystal structure of ribonuclease T_1–2′-guanylic acid complex.[1] Recently we reported that the prominent stacking detected in 7-methylguanine–tryptophan systems by means of X-ray crystallographic, spectroscopic and theoretical studies provides a possible mechanism for the role of the tryptophan residue in selective recognition at the 'cap' structure.[2] Some crystal structure analyses have indicated the existence of interesting stacking interactions between nucleic acid bases and aromatic amino acid components.[3-8] We carried out an X-ray analysis of

3′-*O*-benzoylthymidine in order to examine what kind of intra- and intermolecular interactions exist.

Results

The final positional and thermal parameters for non-hydrogen atoms and hydrogen atoms are given in Tables I and II, respectively.

Stacking Interactions and Crystal Packing

Figure 1 shows the crystal structure projected down the *b* axis. The most pronounced feature of the crystal structure is the alternating stacking of the thymidine base with the phenyl ring between adjacent molecules related by a two-fold screw axis, forming columns parallel to the *b* axis with a separation distance of about 3.3 Å, and a dihedral angle of 4.4 (2)°.

In this crystal, six unique hydrogen bondings were found, as tabulated in Table III, of which two weak ones are C–H---O type hydrogen bonds, *i.e.*, the intermolecular C(12′)–H---O (2) and the intramolecular C(6)–H---O(5′) hydrogen bonds. The water molecule is fixed by three hydrogen bonds with surrounding nucleoside molecules.

Molecular Conformation

The molecular conformation is shown by a stereographic drawing[9] in Fig. 2. The

TABLE I. Final Atomic Coordinates of Non-hydrogen Atoms (× 10^4) and Equivalent Isotropic Thermal Parameters with Their Estimated Standard Deviations in Parentheses

$B_{eq}=4/3\Sigma_i\Sigma_j\beta_{ij}\mathbf{a}_i\cdot\mathbf{a}_j$

Atom	x	y	z	B_{eq} (Å^2)
O(4)	1272 (2)	4480 (6)	3905 (2)	3.79 (9)
O(2)	4522 (2)	4238 (6)	2691 (2)	4.08 (10)
N(1)	2957 (2)	4512 (6)	1136 (2)	2.99 (9)
C(2)	3521 (3)	4371 (7)	2386 (3)	2.98 (10)
N(3)	2881 (2)	4373 (6)	3261 (2)	3.10 (9)
C(4)	1747 (3)	4480 (7)	3016 (3)	3.04 (10)
C(5)	1216 (3)	4589 (8)	1685 (3)	3.21 (11)
C(6)	1824 (3)	4595 (8)	820 (3)	3.35 (11)
C(7)	−21 (3)	4690 (12)	1328 (4)	5.22 (18)
C(1′)	3574 (3)	4651 (8)	126 (3)	3.30 (11)
C(2′)	3433 (3)	2825 (7)	−751 (4)	3.61 (13)
C(3′)	3514 (3)	3744 (7)	−2016 (3)	3.24 (12)
C(4′)	3107 (3)	5925 (7)	−1968 (3)	3.01 (11)
C(5′)	1938 (3)	6258 (10)	−2725 (4)	4.55 (16)
C(6′)	4853 (3)	4134 (7)	−3228 (3)	3.12 (12)
C(7′)	6022 (3)	4270 (7)	−3249 (3)	2.96 (10)
C(8′)	6837 (3)	4302 (7)	−2128 (3)	3.54 (11)
C(9′)	7922 (3)	4439 (8)	−2225 (4)	4.28 (13)
C(10′)	8198 (3)	4550 (8)	−3390 (4)	4.30 (14)
C(11′)	7387 (3)	4531 (8)	−4495 (4)	3.91 (13)
C(12′)	6299 (3)	4375 (7)	−4427 (3)	3.27 (11)
O(1′)	3160 (2)	6366 (5)	−645 (2)	3.58 (9)
O(3′)	4667 (2)	3785 (5)	−2069 (2)	3.38 (9)
O(5′)	1193 (2)	4897 (8)	−2327 (3)	5.46 (14)
O(6′)	4105 (2)	4347 (7)	−4156 (2)	4.37 (10)
O(W1)	640 (4)	1923 (8)	−4152 (3)	6.55 (17)

Table II. Atomic Coordinates of Hydrogen Atoms ($\times 10^3$)

Atom	x	y	z
H(3)	324 (4)	438 (11)	414 (4)
H(6)	149 (4)	472 (11)	−18 (5)
H(ME1)	−27 (4)	457 (12)	44 (4)
H(ME2)	−38 (4)	318 (10)	153 (5)
H(ME3)	−19 (4)	591 (11)	195 (5)
H(1′)	438 (4)	486 (11)	63 (4)
H(2″)	402 (4)	169 (10)	−42 (5)
H(2′)	268 (4)	213 (10)	−81 (4)
H(3′)	313 (4)	275 (10)	−270 (5)
H(4′)	361 (4)	682 (10)	−232 (4)
H(5′)	169 (4)	786 (10)	−260 (4)
H(5″)	187 (5)	588 (10)	−374 (5)
H(O5′)	96 (4)	376 (10)	−299 (5)
H(8′)	661 (4)	424 (11)	−130 (4)
H(9′)	853 (4)	457 (12)	−138 (4)
H(10′)	898 (4)	454 (11)	−350 (5)
H(11′)	761 (4)	455 (11)	−534 (4)
H(12′)	568 (4)	426 (11)	−528 (5)
H(W1)	13 (4)	76 (11)	−425 (4)
H(W2)	36 (4)	268 (10)	−497 (5)

All isotropic temperature factors are 4.1 Å^2.

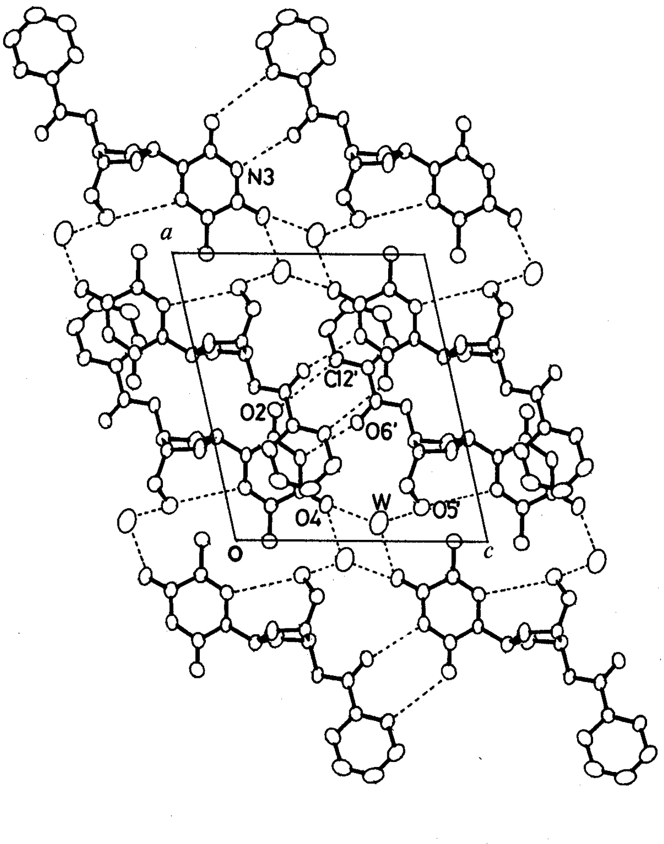

Fig. 1. The Crystal Structure Viewed Down the b Axis

Dotted lines indicate the hydrogen bonds.

relevant torsion angles are given in Table IV. The least-squares plane data for the two rings are given in Table V.

The glycosidic torsion angle, χ_{CN} 49.3 (2) °, is in the normal *anti* region. The deoxyribose ring has the usual C(2′)-*endo* pucker with $P = 160.6°$ and $\tau_m = 30.2°$ as the pseudorotation

Table III. Distances (Å) and Angles (°) for Hydrogen Bonds

D–H	A		Distance D---A	Distance H---A	Angle D–H---A
N(3)–H	O(6′)	$(x, y, z+1)$	2.864 (4)	1.91 (6)	174 (5)
C(12′)–H	O(2)	$(x, y, z-1)$	3.397 (5)	2.34 (6)	171 (5)
C(6)–H	O(5′)	(x, y, z)	3.310 (5)	2.26 (6)	167 (5)
O(5′)–H	O(W)	(x, y, z)	2.744 (6)	1.72 (6)	176 (5)
O(W)–H	O(4)	$(-x, y-1/2, -z)$	2.941 (6)	2.05 (6)	149 (5)
O(W)–H	O(4)	$(x, y, z-1)$	2.922 (6)	2.18 (6)	129 (5)

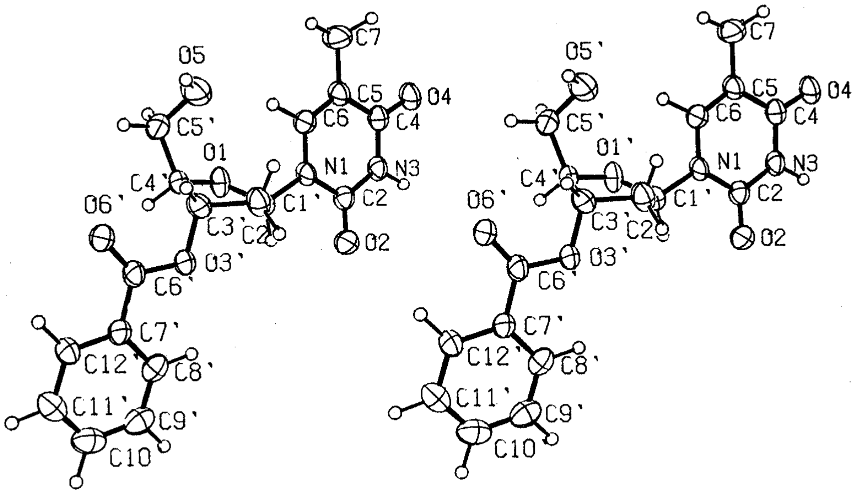

Fig. 2. Stereoview of the Molecular Conformation with the Atomic Numbering

Table IV. Selected Torsion Angles (°)

χ_{CN}	C(6)–N(1)–C(1′)–O(1′)	49.3 (4)
τ_0	C(4′)–O(1′)–C(1′)–C(2′)	−19.0 (4)
τ_1	O(1′)–C(1′)–C(2′)–C(3′)	29.4 (4)
τ_2	C(1′)–C(2′)–C(3′)–C(4′)	−28.5 (4)
τ_3	C(2′)–C(3′)–C(4′)–O(1′)	18.0 (4)
τ_4	C(3′)–C(4′)–O(1′)–C(1′)	0.5 (4)
ψ	C(3′)–C(4′)–C(5′)–O(5′)	55.9 (5)

parameters;[10] but the τ_m (maximum amplitude) is lower than average (38.7 °)[11] indicating less puckering of the deoxyribose ring.

The orientation of the C(5′)–O(5′) bond is *gauche–gauche*. The *anti* conformation about the glycosidic bond and the *gauche–gauche* orientation of the C(5′)–O(5′) bond are thought to be essential for the formation of the C(6)–H---O(5′) hydrogen bond, as pointed out in a number of nucleosides and nucleotides.[12–15] We therefore computed the intramolecular distances between H(6) and O(5′) in 64 crystal structures of pyrimidine nucleosides and nucleotides with the *anti* and *gauche–gauche* conformations (Table VI), and we found that 40 (63%) of the 64 crystal structures have intramolecular H(6)---O(5′) distances of less than 2.51 Å. This frequency of short C(6)–H---O(5′) contacts suggests that in nucleosides and nucleotides with *anti* form the *gauche–gauche* conformer is in part stabilized by the C(6)–H---O(5′) hydrogen bond.[16]

TABLE V. Least-Squares Planes

(a) Thymine ring	Deviation (Å)	(b) Phenyl ring	Deviation (Å)
N(1)	0.007 (4)[a]	C(7′)	−0.002 (4)[a]
C(2)	−0.007 (4)[a]	C(8′)	0.000 (5)[a]
N(3)	−0.000 (4)[a]	C(9′)	−0.001 (5)[a]
C(4)	0.005 (5)[a]	C(10′)	0.000 (5)[a]
C(5)	−0.003 (5)[a]	C(11′)	0.007 (5)[a]
C(6)	−0.005 (5)[a]	C(12′)	−0.006 (5)[a]
O(2)	−0.034 (5)	C(6′)	0.003 (5)[a]
O(4)	0.019 (5)	O(6′)	0.163 (5)
C(7)	−0.011 (7)	O(3′)	−0.161 (5)
C(1′)	0.086 (6)		

Equations of the planes

(a) $0.0381x+0.9981y+0.0476z-3.1276=0.0$

(b) $0.0639x+0.9976y-0.0255z-6.2362=0.0$

a) Atoms used to calculate the least-squares planes.

x, y, and z refer to the orthogonal coordinates system (Å) with x along $\boldsymbol{a}$, y along $\boldsymbol{b}$, and z along $\boldsymbol{c}^*$.

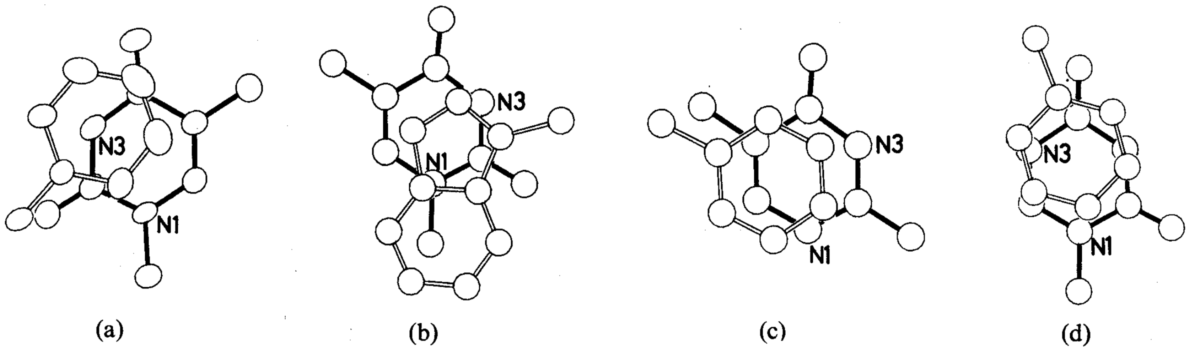

Fig. 3. Stacking Modes between Pyrimidine and Aromatic Ring

Each pair is projected down normal to the pyrimidine plane. a) This work. b) 1-[3-(1*H*-Indol-3-yl)propyl]thymine.[5] c) DL-2-(α-Hydroxybenzyl)oxythiamine.[6] d) 5-[*N*-(L-Phenylalanyl)amino]uridine.[7]

Bond Distances and Angles

The bond distances and angles are listed in Tables VII and VIII, respectively. The bond distances and angles in the thymine moiety are in good agreement with the average values (Table IX) for 7 independent thymidine derivatives. The bond distances of the sugar moiety are normal as compared with those of the furanose ring.[17] The endocyclic bond angles are somewhat different from values of the *S*-type (C(2′)-*endo*) furanose ring, *i.e.*, C(1′)–C(2′)–C(3′) and C(2′)–C(3′)–C(4′) are significantly larger than those of standard *S*-type conformer[16] by 1.8 and 2.2°, respectively, probably because of the flattening of the deoxyribose ring and of the correlation between the endocyclic bond angles and the pseudorotation parameters.[11] The mean bond distances and angles of the phenyl ring, 1.387 (7) Å and 120.0 (9)° are in very good agreement with the standard values.

Discussion

Although many model compounds for both nucleic acid bases and protein components have been investigated by means of X-ray analysis, stacking interactions have rarely been observed,[3-8] partly because of the weakness of this interaction force and partly because of the

TABLE VI. The C(6)–H---O(5′) Interactions of Pyrimidine Nucleosides and Nucleotides

	R-Factor	H(6)–O(5′) (Å)	C(6)–O(5′) (Å)	C(6)–H(6) (Å)	C(6)–H(6)- O(5′) (°)	CHI (°)	PSI (°)	TAU3 (°)	PSI′ (°)
ACURID	0.034	3.332	4.266	1.015	153.68	55.87	49.61	-32.34	87.93
ADBURM	0.138	2.405	3.435	1.039	170.60	20.03	36.21	-41.11	83.35
ADURPO10	0.060	2.327	3.176	1.004	141.64	5.02	56.87	-35.54	83.78
ADYPUR10	0.057	2.494	3.463	0.995	164.41	28.98	57.53	-41.11	73.55
ADYPUR10	0.057	2.436	3.390	0.987	162.30	29.05	58.57	-40.88	77.31
ARACYP	0.034	2.892	3.471	0.831	128.51	25.93	64.92	30.18	157.76
ASTHYM10	0.054	2.347	3.348	1.039	161.14	54.08	51.33	23.76	147.20
BAGXIQ	0.060	2.654	3.632	1.051	154.56	21.49	46.33	-43.16	72.22
BELZEX	0.056	2.494	3.337	0.859	167.53	26.48	49.08	-38.38	87.04
BEURID10	0.033	2.222	3.112	1.005	146.95	18.28	45.94	-34.53	80.79
BEURID10	0.033	2.621	3.548	0.937	170.14	24.26	39.63	-39.52	72.81
BEZGES	0.038	2.542	3.398	1.010	142.30	9.81	47.63	-37.64	81.18
BIVVIL	0.028	2.801	3.647	1.014	141.24	39.97	50.24	23.45	146.91
BIVVIL	0.028	2.432	3.469	1.075	161.73	55.67	57.80	25.50	146.75
BOFWIC	0.041	3.328	4.263	1.096	143.93	54.44	53.10	-30.64	87.15
BOFWOI	0.035	2.378	3.480	1.103	176.43	28.94	51.24	-37.79	76.08
BOLHEP	0.111	2.638	3.626	1.067	153.89	26.54	56.30	-45.97	76.81
BRURID10	0.070	2.268	3.231	1.011	158.76	52.52	52.02	21.70	146.73
BURDMS	0.067	2.347	3.363	1.086	155.17	62.12	52.61	29.00	151.17
CAGUCP10	0.082	2.336	3.295	0.982	165.36	21.21	57.31	-40.05	79.63
CDURID	0.036	2.692	3.639	1.056	149.16	61.66	49.97	20.52	142.13
CLURID10	0.032	2.347	3.218	0.894	164.79	51.44	52.29	24.58	146.38
CXMURD	0.047	2.592	3.552	1.020	156.59	49.44	60.02	23.61	148.75
CYCYPH10	0.064	2.044	3.151	1.138	163.32	59.15	48.49	22.05	151.47
CYTIAC	0.045	2.783	3.583	0.922	145.84	41.82	43.83	28.11	152.05
CYTIAC01	0.073	2.673	3.430	0.932	138.77	39.28	45.51	27.87	150.18
CYTIDI10	0.056	2.328	3.233	0.982	152.78	18.37	47.13	-34.74	83.00
CYTPCD10	0.038	2.484	3.373	0.995	148.62	16.14	58.73	-33.53	78.66
DAURID01	0.039	2.553	3.406	0.915	155.28	60.42	63.58	26.90	155.63
DOCYPO	0.023	2.912	3.029	0.892	88.87	-6.18	56.42	32.58	153.40
DOCYPO01	0.070	2.874	3.002	1.042	86.83	-6.08	57.25	31.11	153.20
DOCYTC	0.035	2.512	3.250	1.067	125.53	-0.09	46.24	-31.85	81.41
DTURID	0.039	2.475	3.419	1.036	151.18	19.57	40.71	-35.79	79.25
DXCYTD	0.054	2.242	3.384	1.208	156.41	44.14	62.39	27.15	148.20
DXCYTD	0.054	2.336	3.267	0.979	158.56	17.79	56.67	-34.16	82.33
ETCYTC	0.045	2.273	3.331	1.073	168.11	42.65	58.35	28.69	152.33
GUPCYT20	0.054	2.341	3.289	0.961	168.50	31.71	50.36	-40.55	77.24
HMDOUR	0.050	2.886	3.936	1.062	169.65	56.44	46.11	4.19	123.27
HXHTYM	0.055	3.525	4.478	1.012	157.91	68.73	49.32	-18.15	105.28
HXURID	0.026	3.086	3.979	0.983	151.84	42.07	52.81	16.82	140.42
IPDXUR	0.061	2.362	3.357	1.081	152.37	64.41	52.96	22.53	145.24
KURDMP	0.035	2.716	3.823	1.128	166.79	71.23	46.73	22.31	142.33
KURDPI	0.059	2.416	3.366	1.012	155.94	45.80	51.95	17.65	140.32
KURDPI	0.059	2.397	3.334	1.028	151.00	40.86	53.92	20.52	142.67
LEUAUR10	0.080	2.629	3.549	1.014	150.81	50.54	41.25	28.60	153.07
MCYTMS10	0.076	2.502	3.296	0.837	158.63	45.45	61.84	27.04	148.82
MEURID	0.075	2.513	3.429	0.925	170.27	29.43	49.41	-38.20	82.06
MXEURD	0.035	2.339	3.256	0.947	162.87	56.56	52.90	14.26	143.41
MXURID	0.085	2.297	3.242	0.999	157.37	25.41	54.89	-35.91	77.37
OMCYTD20	0.069	2.549	3.313	0.779	167.26	47.42	55.86	29.87	151.53
OMCYTD20	0.069	2.488	3.359	0.931	155.75	43.61	51.63	27.65	151.06
PRYURD10	0.027	2.485	3.360	0.915	160.12	61.84	52.15	23.62	148.24
RFURPD	0.049	2.348	3.298	0.968	166.91	53.52	50.55	23.28	146.75
SURIDP	0.043	2.894	3.769	0.918	159.70	45.10	42.89	17.97	144.18
SURIPH10	0.145	2.821	3.833	1.166	144.74	44.44	49.63	18.52	146.88
TCYTDH	0.039	2.219	3.189	0.989	166.43	20.30	54.60	-35.35	82.43
THPYUR	0.039	2.378	3.339	1.011	158.56	49.61	55.92	27.57	153.43
TMAMUR10	0.063	2.355	3.272	0.954	160.95	15.07	50.71	-39.10	82.56
TMCMUR	0.119	2.425	3.410	1.085	150.35	16.10	47.84	-28.72	76.50
TYMCXA	0.039	2.276	3.258	0.994	169.64	43.95	97.19	28.79	150.15
TYMCXA	0.039	2.342	3.309	0.990	165.23	52.74	85.94	23.21	145.51
URIDMP10	0.051	2.392	3.382	0.998	171.53	55.80	58.56	25.82	149.88
UROAME	0.040	2.172	3.128	1.010	157.40	34.28	58.85	-36.02	79.97
ZZZAPA10	0.034	2.727	3.127	1.044	102.71	3.41	51.40	21.90	148.94

disturbance by unfavorable crystal packing force. As an exceptional case, it has been reported that 7-methylguanine and 1-methyladenine residues having a positive charge introduced by *N*-methylation exhibit strong stacking interactions with tryptophan residue,[2,3] while only thymine as a normal nucleic acid base has a stacking interaction with the aromatic ring, as found in the present crystal and 1-[3-(1*H*-indol-3-yl)-propyl]thymine.[5] These results support

TABLE VII. Bond Distances (Å) for the Non-hydrogen Atoms

N(1)–C(2)	1.378 (5)	O(1′)–C(1′)	1.424 (5)
C(2)–N(3)	1.366 (5)	C(4′)–C(5′)	1.525 (6)
N(3)–C(4)	1.388 (5)	C(5′)–O(5′)	1.423 (7)
C(4)–C(5)	1.442 (5)	C(3′)–O(3′)	1.458 (5)
C(5)–C(6)	1.328 (5)	O(3′)–C(6′)	1.340 (4)
C(6)–N(1)	1.385 (5)	C(6′)–O(6′)	1.217 (5)
C(2)–O(2)	1.227 (5)	C(6′)–C(7′)	1.473 (5)
C(4)–O(4)	1.233 (5)	C(7′)–C(8′)	1.398 (5)
C(5)–C(7)	1.514 (6)	C(8′)–C(9′)	1.389 (6)
N(1)–C(1′)	1.470 (5)	C(9′)–C(10′)	1.376 (6)
C(1′)–C(2′)	1.509 (6)	C(10′)–C(11′)	1.383 (6)
C(2′)–C(3′)	1.513 (6)	C(11′)–C(12′)	1.385 (6)
C(3′)–C(4′)	1.521 (5)	C(12′)–C(7′)	1.388 (5)
C(4′)–O(1′)	1.441 (5)		

TABLE VIII. Bond Angles (°) for the Non-hydrogen Atoms

C(2)–N(1)–C(6)	121.2 (3)	C(3′)–C(4′)–O(1′)	106.2 (3)
N(1)–C(2)–N(3)	114.9 (3)	C(4′)–O(1′)–C(1′)	110.5 (3)
C(2)–N(3)–C(4)	127.0 (3)	C(2′)–C(3′)–O(3′)	107.9 (3)
N(3)–C(4)–C(5)	114.7 (3)	C(3′)–C(4′)–C(5′)	113.9 (3)
C(5)–C(6)–N(1)	122.9 (4)	C(5′)–C(4′)–O(1′)	109.2 (3)
C(2)–N(1)–C(1′)	119.1 (3)	C(4′)–C(5′)–O(5′)	111.4 (4)
C(6)–N(1)–C(1′)	119.6 (3)	C(3′)–O(3′)–C(6′)	114.7 (3)
O(2)–C(2)–N(1)	122.6 (3)	C(7′)–C(6′)–O(3′)	113.8 (3)
O(2)–C(2)–N(3)	122.5 (3)	C(7′)–C(6′)–O(6′)	124.6 (3)
O(4)–C(4)–N(3)	120.1 (3)	O(3′)–C(6′)–O(6′)	121.6 (3)
O(4)–C(4)–C(5)	125.2 (3)	C(6′)–C(7′)–C(8′)	121.7 (3)
C(4)–C(5)–C(7)	118.4 (4)	C(6′)–C(7′)–C(12′)	117.9 (3)
C(6)–C(5)–C(7)	122.4 (4)	C(8′)–C(7′)–C(12′)	120.4 (3)
N(1)–C(1′)–C(2′)	114.0 (3)	C(7′)–C(8′)–C(9′)	118.4 (4)
N(1)–C(1′)–O(1′)	107.2 (3)	C(8′)–C(9′)–C(10′)	121.3 (4)
O(1′)–C(1′)–C(2′)	106.4 (3)	C(9′)–C(10′)–C(11′)	120.0 (4)
C(1′)–C(2′)–C(3′)	103.1 (3)	C(10′)–C(11′)–C(12′)	119.9 (4)
C(2′)–C(3′)–C(4′)	104.6 (3)	C(7′)–C(12′)–C(11′)	120.0 (4)
C(4)–C(5)–C(6)	119.2 (4)	C(4′)–C(3′)–O(3′)	109.1 (3)

TABLE IX. The Average Values of Bond Distances (Å) and Angles (°) for the Seven Known Thymine Bases

N(1)–C(2)	1.379 (11)	N(1)–C(6)	1.388 (15)
C(2)–N(3)	1.380 (8)	C(2)–O(2)	1.217 (11)
N(3)–C(4)	1.381 (9)	C(4)–O(4)	1.233 (18)
C(4)–C(5)	1.444 (9)	C(5)–C(7)	1.502 (8)
C(5)–C(6)	1.340 (10)	N(1)–C(1′)	1.467 (12)
C(6)–N(1)–C(2)	121.5 (4)	N(1)–C(2)–O(2)	124.1 (6)
N(1)–C(2)–N(3)	114.2 (5)	N(3)–C(2)–O(2)	121.7 (10)
C(2)–N(3)–C(4)	127.2 (5)	N(3)–C(4)–O(4)	119.9 (7)
N(3)–C(4)–C(5)	115.6 (6)	C(5)–C(4)–O(4)	124.6 (9)
C(4)–C(5)–C(6)	118.0 (5)	C(4)–C(5)–C(7)	118.9 (8)
C(5)–C(6)–N(1)	123.4 (5)	C(6)–C(5)–C(7)	123.0 (9)
C(2)–N(1)–C(1′)	118.2 (18)	C(6)–N(1)–C(1′)	120.2 (15)

the spectroscopic evidence that the aromatic rings of tyrosine and tryptophan interact with nucleosides and nucleotides in general and with thymidine in particular.[18,19)]

As shown in Fig. 3, the most characteristic feature of the known stacking modes between pyrimidine and aromatic rings is the considerable ring–ring overlapping, compared with the stacking patterns between pyrimidine–pyrimidine bases.[20)] The more extensive overlapping of pyrimidine–aromatic rings as compared with pyrimidine–pyrimidine bases could be essential for the specific recognition of nucleic acid by protein.

Experimental

3′-*O*-Benzoylthymidine was synthesized as described in a previous paper[21)] and crystallized from dioxane–water (50 : 50 v/v) mixture. The crystal data are as follows: Chemical formula $C_{17}H_{18}N_2O_6 \cdot H_2O$; M_r 364.4; space group $P2_1$; $a = 12.528$ (6), $b = 6.544$ (2), $c = 10.770$ (5) Å, $\beta = 102.80$ (5) °; $V = 861.0$ (7) Å^3; $Z = 2$; $D_m = 1.393$ (3) (determined by floatation in benzene–carbon tetrachloride) and $D_x = 1.405$ Mg m^{-3}; $F(000) = 384$; μ (CuK_α) = 0.89 mm^{-1}.

The intensity data were collected on a Rigaku automatic four-circle diffractometer with CuK_α radiation within $\sin\theta/\lambda = 0.58$ Å^{-1} using the $\omega - 2\theta$ scan technique. The intensities of 1524 independent reflections were corrected for Lorentz and polarization factors, but not for absorption.

The structure was solved by a combination of the Patterson function and direct methods; the peaks around the origin were used for assignment of the six-membered ring orientation and the minimum function based on this orientation revealed the relative position of the other intramolecular six-membered ring. These data were used for the calculation of *E*-values by using the program MULTAN 78.[22)] An *E* map computed with the best set of phases (combined figure of merit = 2.95) revealed the position of 19 atoms in the thymine and benzoyl rings. Subsequent Fourier syntheses showed the peaks for all remaining non-hydrogen atoms.

The structure was refined by the full-matrix least-squares method with anisotropic temperature factors. The positions of all hydrogen atoms were located on a difference Fourier map. The final refinement, including hydrogen atoms with isotropic temperature factors, reduced the *R* value to 0.053 ($R_w = 0.070$) for 1476 non-zero reflections. The quantity minimized was $\Sigma w (|F_o| - k|F_c|)^2$, where $w = 1/(\sigma^2|F_o| + 0.1139|F_o| + 0.0027|F_o|^2)$ and k is a scale factor. The highest peak height in the final difference Fourier map is 0.18 eÅ^{-3}. The average and maximum ratios of parameter shift to e.s.d. after refinement were 0.04 and 0.19, respectively, for non-hydrogen atoms. All numerical calculations were carried out on an ACOS 900 computer at the Crystallographic Research Center, Institute for Protein Research, Osaka University, by using the programs of UNICS (1979).[23)] The atomic scattering factors used were those cited in "International Tables for X-ray Crystallography," (1974).[24)]

The crystallographic data for related nucleosides and nucleotides were obtained from the Cambridge Crystallographic Database, September, 1983, using the programs BIBSER, RETRIVE and GEOM.

References and Notes

1) U. Heinemann and W. Saenger, *Nature* (London), **299**, 27 (1982).
2) T. Ishida, M. Katsuta, M. Inoue, Y. Yamagata and K. Tomita, *Biochem. Biophys. Res. Commun.*, **115**, 849 (1983).
3) T. Ishida, H. Usami, M. Inoue, Y. Yamagata and K. Tomita, *Biochem. Biophys. Res. Commun.*, **107**, 746 (1982).
4) M. Sundaralingam and S. K. Arora, *J. Mol. Biol.*, **71**, 49 (1972).
5) D. Voet, *J. Am. Chem. Soc.*, **102**, 2071 (1980).
6) W. Shin, J. Pletcher and M. Sax, *J. Am. Chem. Soc.*, **101**, 4365 (1979).
7) H. M. Berman, W. C. Hamilton and R. J. Roussean, *Biochemistry*, **12**, 1089 (1973).
8) M. Sax, P. Pulsinelli and J. Pletcher, *J. Am. Chem. Soc.*, **99**, 1396 (1977).
9) C. K. Johnson, ORTEP II, Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, 1976.
10) C. Altona and M. Sundaralingam, *J. Am. Chem. Soc.*, **94**, 8205 (1972).
11) E. Westhof and M. Sundaralingam, *J. Am. Chem. Soc.*, **102**, 1493 (1980).
12) E. Shefter and K. N. Trueblood, *Acta Crystallogr.*, **18**, 1067 (1965).
13) T. Srikrishnan and R. Parthasarathy, *Acta Crystallogr., Sect. B*, **34**, 1363 (1978).
14) T. Srikrishnan, S. M. Fridey and R. Parthasarathy, *J. Am. Chem. Soc.*, **101**, 3739 (1979).
15) B. Hingerty, P. J. Bond, R. Langridge and F. Rottman, *Acta Crystallogr., Sect. B*, **33**, 1349 (1977).
16) W. Saenger, "Principles of Nucleic Acid Structure," Springer-Verlag, New York, 1984, pp. 80—82.
17) M. Sundaralingam, *Jerusalem Symp. Quantum Chem. Biochem.*, **5**, 417 (1973).
18) T. Montenay-Garestier and C. Helene, *Biochemistry*, **10**, 300 (1971).
19) C. Helene, T. Montenay-Garestier and J-L. Dimicoli, *Biochim. Biophys. Acta*, **254**, 349 (1971).

20) C. E. Bugg, J. M. Thomas, M. Sundaralingam and S. T. Rao, *Biopolymers*, **10**, 175 (1971).
21) S. Uesugi, T. Shida, H. Miyashiro, K. Tomita, M. Ikehara, Y. Kobayashi and Y. Kyogoku, *Nucleic Acids Res.*, Symposium Series No. 11, 237 (1982).
22) P. Main, S. E. Hull, L. Lessinger, J. P. Declercq and M. M. Woolfson, MULTAN 78, A System of Computer Programs for the Automatic Solution of Crystal Structure from X-Ray Diffraction Data. Univs. of York, England and Louvain, Belgium, 1978.
23) T. Ashida (ed.) "The Universal Crystallographic Computing System-Osaka," Library of Programs, Computing Center, Osaka Univ., 1979.
24) J. A. Ibers and W. C. Hamilton (eds.) "International Tables for X-Ray Crystallography," Vol. IV, Kynoch Press, Birmingham, 1974, pp. 72—73.