The Crystal and Molecular Structure of 5-Fluorouracil

By LAWRENCE FALLON III

Department of Materials Science, University of Virginia, Charlottesville, Virginia 22901, U.S.A.

(Received 11 May 1973; accepted 14 June 1973)

5-Fluorouracil (C₄H₃FN₂O₂) is triclinic, space group $P\overline{1}$, with $a=9\cdot22$ (3), $b=12\cdot66$ (3), $c=12\cdot67$ (3) Å, $\alpha=89\cdot7$ (3), $\beta=43\cdot9$ (3), $\gamma=98\cdot6$ (3)°, Z=8. Intensities were measured with an automated diffractometer and Mo K α radiation. Patterson and symbolic addition methods failed to yield a complete solution, but gave a partial structure which was used to supply initial phases to the tangent formula. The complete structure was thus determined in space group P1. Least-squares refinement in space group P $\overline{1}$ gave a final R_w of 0.054 and R of 0.092. The unit cell contains two molecular layers, each with four molecules. Structural features of interest include several pseudo-centers of symmetry and several different stacking interactions.

Introduction

The properties of purine and pyrimidine bases are determined to a great extent by their hydrogen and π bonding systems. Several analog purine and pyrimidine bases have been developed which resemble structurally the natural bases. They alter the base hydrogen and π bonding characteristics, however, and greatly affect the properties of the nucleic acids containing them. These analog bases have been observed to have inhibitory and mutagenic effects upon the growth of certain tumors. Of the pyrimidine analogs, 5-fluorouracil (FU) is probably the most widely used, and is of great interest in clinical and experimental cancer chemotherapy. The exact mechanism of the inhibitory and mutagenic action of FU is not known, but it is thought to involve (a) an inhibition of DNA synthesis, (b) an inhibition of the formation of ribosomes, and (c) a faulty information transfer within the cell caused by the replacement of uracil by FU in mRNA (Mandel, 1969).

FU bears a close resemblance to uracil (see Fig. 1). Since structural differences, especially differences in stacking interactions, are thought to be important in the activity of halogenated pyrimidines, it is of interest to examine the molecular structure of FU for comparison with the structures of related systems.

Experimental measurements

A crystalline powder of FU (2,4-dioxo-5-fluoropyrimidine) was obtained from the Sigma Chemical Company. This powder quickly dissolves in hot water and, upon evaporation, forms large crystals of FU which are stable transparent prisms showing a pronounced cleavage plane. A crystal approximately $0.1 \times 0.2 \times 0.4$ mm was selected.

Precession, Weissenberg, and rotation photographs were taken at room temperature ($\sim 22 \,^{\circ}$ C) with Cu Ka radiation. The diffraction symmetry observed from the photographs is \overline{I} , and a measured unit cell is a=9.22 (3), b = 12.66 (3), c = 8.79 (3) Å, $\alpha = 80.2$ (3), $\beta = 90.8$ (3), $\gamma = 98.6$ (3)°, V = 998 Å³.

The density measured by pycnometry at room temperature (23.5 °C) was 1.728 g cm⁻³ which may be compared with a calculated value of 1.731 g cm⁻³ for Z = 8. A piezoelectric test was inconclusive. Because of the large Z, however, the space group was assumed to be $P\overline{1}$ with 4 molecules in the asymmetric unit. Intensities were measured for the equivalent of the asymmetric half of the copper sphere with a Picker four-circle card-controlled diffractometer and Mo Ka radiation. Symmetric 2θ scans were used with a rate of 1°/min and a scan width of 2°. Background intensities were counted at each scan limit for 1 min. The net intensity was set equal to the scan count minus the sum of the background counts. Monochromatic radiation was obtained by reflecting the primary beam from 002 of a graphite crystal. During the data collection, the intensities of three standard reflections were measured after each group of fifty data reflections. 4366 independent reflections were measured. A standard error σ_I was calculated for each intensity from

$$\sigma_I = \sqrt{(s+b_1+b_2)}$$

where s is the scan count, and b_1 and b_2 are the two background counts. If an intensity was observed to be less than $2\sigma_I/3$, the reflection was typed as unobserved and assigned the value of $2\sigma_I/3$. 1122 observed and 3244 unobserved reflections were obtained. The Lorentz-polarization correction, adapted for the geometry of the monochromator, was calculated with computer program *DATRDN* of the X-RAY System (1972) of crystallographic computer programs. No absorption correction was made.

Computer program NORNSF from the X-RAY System (1972) was used to calculate a set of normalized structure factors. The atomic scattering factors were taken from *International Tables for X-ray Crystallog-raphy* (1968). The Wilson plot gave B=3.779 Å² and K=1.1880. The statistical averages calculated from

the normalized structure factors are listed in Table 1, and yield additional evidence for the space group $P\overline{1}$. The magnitudes of the normalized structure factors show a hypercentric distribution corresponding to the presence of one or more pseudo-centers of symmetry.

Table 1. Statistical properties of normalizedstructure factors

	Exper-	Theo	retical
	imental	Centric	Acentric
$\langle E \rangle$	0.685	0.798	0.886
$\langle E ^2 \rangle$	1.000	1.000	1.000
$\langle E^2 - 1 \rangle$	1.187	0.968	0.736
$\langle E^2 - 1 ^2 \rangle$	5.492	2.000	1.000

Structure determination and refinement

In the calculated Patterson map obvious planar features were present. These layers were inclined with respect to the planes of calculation. The unit cell was transformed for greater ease of analysis by the matrix [100/010/101]. The planar layers are then perpendicular to **a**^{*} and parallel to the *bc* plane. The transformed unit cell used from this point throughout the structure determination has the following parameters: a=9.22 (3), b=12.66 (3), c=12.67 (3) Å, $\alpha=89.7$ (3), $\beta=43.9$ (3), $\gamma=98.6$ (3)°. One of the major features in the Patterson map is a very large peak at (0.477, 0.216, 0.483), indicating that there are a large number of atoms in one layer related by this vector to atoms in the next. Several figures, roughly the expected shape of FU molecules, were observed in the map, but



Fig. 1. The chemical structure and numbering system of 5-fluorouracil.

Table 2. Final atomic coordinates ($\times 10^4$) and thermal parameters ($\times 10^6$)

Thermal parameters listed are coefficients in the temperature factor expression

 $\exp\left[-2\pi^2(U_{11}h^2a^{*2}+U_{22}k^2b^{*2}+U_{33}l^2c^{*2}+2U_{12}hka^*b^*\cos\gamma^*+2U_{13}hla^*c^*\cos\beta^*+2U_{23}klb^*c^*\cos\alpha^*\right].$

			•• ••		,	-			
	$x(\sigma)$	$y(\sigma)$	$z(\sigma)$	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
N(1,1)	2001 (18)	1047 (8)	4717 (13)	50418	45112	35163	7095	- 35201	- 8089
C(1.2)	2255 (23)	1482 (12)	3559 (16)	39527	37833	37835	1642	- 27595	484
N(1,3)	2306 (17)	2587 (8)	3450 (12)	45952	26448	43926	11720	- 35650	- 9775
C(1.4)	2152 (23)	3290 (11)	4349 (18)	26267	52310	30936	6067	- 17665	- 16020
C(1,5)	1972 (23)	2748 (12)	5473 (16)	41886	74185	24433	17180	- 25338	- 14247
C(1,6)	1859 (23)	1678 (11)	5675 (16)	31463	33463	45413	- 3772	- 23815	1971
O(1,7)	2369 (16)	918 (6)	2753 (10)	102834	31180	57977	22008	- 69201	- 14175
O(1,8)	2234 (16)	4268 (6)	4145 (11)	81545	25418	54516	11902	- 54447	- 3895
F(1,9)	1875 (12)	3428 (5)	6344 (8)	69885	44375	40358	14967	- 44138	- 5922
N(2,1)	2773 (19)	707 (8)	334 (13)	49206	23521	41030	3721	- 34905	- 4959
C(2,2)	2840 (24)	9687 (10)	9917 (18)	47517	61746	60526	5148	- 44625	- 14596
N(2,3)	2774 (18)	9569 (8)	8862 (13)	43070	37315	38681	3439	- 33843	4872
C(2,4)	2583 (23)	387 (10)	8259 (17)	35418	17405	41776	- 54	- 28918	- 7749
C(2,5)	2436 (22)	1433 (10)	8828 (16)	29153	43134	26899	- 6669	- 18433	5158
C(2,6)	2566 (22)	1563 (8)	9814 (16)	14327	26090	32752	- 2273	- 11909	9222
O(2,7)	2887 (16)	8919 (7)	476 (11)	91528	37303	61244	30130	- 66789	- 15456
O(2,8)	2428 (16)	225 (6)	7363 (11)	61421	26496	52527	3114	- 47955	- 1456
F(2,9)	2172 (12)	2213 (5)	8297 (8)	60992	36481	51400	10395	- 40802	- 1577
N(3,1)	2700 (18)	6737 (7)	763 (12)	38482	29589	32901	7792	- 29113	- 5893
C(3,2)	2720 (24)	6357 (12)	1772 (17)	20094	28977	44800	6108	-12507	- 4605
N(3,3)	2482 (17)	5225 (8)	2039 (12)	38232	25579	55898	12046	- 35093	- 19043
C(3,4)	2324 (22)	4515 (11)	1259 (15)	27466	51170	9759	- 4700	- 7598	10082
C(3,5)	2241 (23)	5021 (10)	262 (16)	49644	79118	19124	3793	-26334	-3100
C(3,6)	2435 (24)	6114 (11)	33 (16)	51991	34174	32032	1878	- 30479	- 5879
O(3,7)	2886 (15)	6935 (6)	2481 (10)	75207	24301	48568	7302	- 52808	-2872
O(3,8)	2220 (15)	3535 (6)	1463 (10)	61070	45722	43918	12991	- 45002	- 4828
F(3,9)	2041 (14)	4303 (5)	9573 (9)	123350	43818	74652	10719	- 83557	- 8533
N(4,1)	2780 (18)	7135 (7)	4704 (12)	35382	28728	20691	9950	-13882	- 7784
C(4,2)	2596 (24)	8110 (11)	5205 (17)	16898	88371	12642	4179	- 4295	-1487
N(4,3)	2654 (16)	8245 (8)	6283 (11)	36474	48489	15110	15531	- 19672	2376
C(4,4)	2794 (24)	7432 (11)	6956 (18)	30182	20580	64874	2747	-32195	- 11548
C(4,5)	3057 (23)	6438 (10)	6264 (18)	40803	43572	50883	23864	-30107	- 12403
C(4,6)	2998 (22)	6251 (9)	5257 (16)	34084	52978	42820	16066	- 28987	-4113
O(4,7)	2451 (15)	8912 (6)	4724 (10)	52997	40775	37220	5515	- 35778	-1541
O(4,8)	2842 (17)	7623 (6)	7847 (11)	67268	32038	38083	10057	- 41080	- 12063
F(4,9)	3271 (13)	5643 (5)	6829 (9)	90480	23721	70009	11268	- 67179	- 2093

Patterson methods failed to give a complete solution.

Symbolic addition was next attempted. However, a correct structure was not found in any of the possible *E*-maps. The failure of symbolic addition will be the subject of a future paper.

A considerable amount of information had thus far been accumulated about the structure of FU, even though Patterson and symbolic addition methods did not give the complete structure. Drs Jerome and Isabella Karle suggested to the author that the tangent formula (Karle & Hauptman, 1956) is useful in such situations. Structure factors were calculated with one FU molecule in an orientation consistent with images in the Patterson map and previous *E*-maps. The position of this molecule was unknown with respect to the origin (that is, with respect to a center of symmetry). Space group P1 was then chosen as appropriate, and an arbitrary origin was assigned (Karle & Karle, 1971). The calculated phases from the partial structure were accepted if the corresponding |E| was greater than 2.5 and if the calculated structure factor $|F_c|$ was greater than the quantity $0.125F_o$. Of the 137 phases corresponding to reflections whose |E| calculated greater than 2.5, 77 were accepted. With the computer program TANGEN from the X-RAY System (1972) the tangent formula generated a total of 318 phases for the normalized structure factors greater than 1.8. If the initial 77 phases were allowed to refine during these calculations, the resultant E-map was unintelligible. If the initially accepted phases were held constant throughout, and only the others allowed to refine, the E-map then calculated showed the molecule corresponding to the initial partial structure. Also visible was a second molecule, identical in orientation with the first, but translated from it by a vector corresponding to the very large peak in the Patterson map. Structure factors were then calculated with these two molecules. Of the 200 |E|'s greater than 2.2, the $|F_c|$ of 105 reflections calculated greater than $0.250 F_o$, and thus were accepted. These were used to generate a total of 429 phases in four cycles of tangent refinement, to an |E| threshold of 1.6. If the 105 initially accepted phases were held constant for the entire four cycles, the Emaps calculated showed only the two molecules used to start the phasing. However, when these 105 phases

were held fixed for only two cycles and allowed to refine with all the phases for the final two cycles, the *E*-map then contained the whole structure. All the atoms were present as strong peaks in the map. There were very few false peaks and for the most part these were small. The centers of symmetry were casily found in the structure. The atomic coordinates were then translated so that space group PI could be used in least-squares refinement.

The coordinates of the 36 atoms in the asymmetric unit were read from the *E*-map and, after translation, were used as input to the least-squares program, CRYLSQ from the X-RAY System (1972). All atoms were initially assigned isotropic temperature factors of 3.0 Å^2 . The scale factor calculated by NORMSF, K =1.1880, was also used initially. The first few cycles of least-squares refinement used only the 876 largest reflections. No corrections were made for anomalous dispersion or extinction. Before the first cycle, R, calculated from $R = \sum ||F_o| - |F_c|| / \sum |F_o|$, was 0.348. Six cycles reduced it to 0.111. A difference synthesis indicated that anisotropic refinement was necessary. After two cycles R was reduced to 0.080. The full data set was used in further refinement cycles. Each structure factor was weighted by $w = 1/\sigma_F^2$ where

$$\sigma_F = \left(|F_o|^2 + \frac{\sigma(I)}{Lp} \right)^{1/2} - |F_o|.$$

 R_w was calculated with the observed reflections from $R_w = \sum w(|F_o| - |F_c|)^2 / \sum w |F_o|^2$.

After two more cycles of anisotropic refinement R_w dropped to 0.068. The 3244 unobserved reflections contributed to the parameter shifts only if they calculated greater than their observed value. It was decided that the structure had refined enough to attempt to find the hydrogen atoms. A difference synthesis showed no features with a peak height greater than 0.75 e Å⁻³. Peaks were located which corresponded to likely positions for all the hydrogen atoms except one. After 4 cycles of least squares, it became obvious that the hydrogen atoms were moving to geometrically unfavorable positions. These atoms were then removed from the refinement. Because $||F_o| - |F_c||$ was excessively large, 14 reflections were removed. This left 1109 observed and 3243 unobserved reflections. After two more



Fig.2. Stereoscopic view of unit cell. Atoms are represented by thermal ellipsoids drawn to 50% probability.

cycles with the 36 atoms of the asymmetric unit refining anisotropically, R_w fell to a final value of 0.054. The final value of R was 0.092. During the last cycle of least-squares calculations the average shift per error was 0.683 and the largest was 8.203. Final atomic parameters with their standard deviations are listed in Table 2. The final scale factor is K = 1.26 (6). Atoms are designated by symbols of the form A(m,n) where

Table 3. Observed and calculated structure factors Image: structure factors

Columns are h, $10F_o$, $10F_c$. Unobserved reflections are marked by *.

.#191-21 -9 529 -93 -1. 997 -8 #181-8	-9 129 121 -8 520 10 -7 171 -192	#1-12,-5 -6 56* 2	-3 86+ '62 8,-10,-4	H.25.04	1111 94• 19 -	11,81-2 11 579 -117	H,-2,-1 -8 879 -83	1 1 4	******	-: #: -:::	1 12	11.7.2 -1 13* -3	1 11. 17	
**************************************	N, 107	m,-18,-3	3 8 4					#1-11.0		11 22 23			1. 12. 11	
-11 59* 17 H,1,+# H,7,-11 -11 55* 15			*,-*,-*		3. 1				m.9.0 -6 569 100					3 23 4
-11 530 -24 -13 440 4 -7 910 14 -14 480 -14	N,11,-7	: :: ::		-5 5655 -5	11: 11	H.b2	*,-1,-1 -8 98* -97	H,-18,6 -5 999 -41	3 3 3	1 2. 31		2 414 17 +44,2		
-1. 464 -25 8.28		N,-8,-1	*****							3 .35 .32		3 8 3	42.1 -1 -10 -11	
	-1 16, 67		3 13. 19			1 110 -110	3 37 33	n,-1,a	******					18.3
**************************************	-7 138 -123 W,13,-7		1 21 3	*,-10,-3 -3		2 336 -322 4,7,+2	**************************************				 			
4,4,-1, -1, 5,2 1, -1,	-6 55* 66				35 32 - 			n,-4,4	-1 117 115	*.2.5		m, 9,2	1.00 -174	*,16,3
M,1,-1, M,4,-4	-7 56* 25	3 8 1	866 -7 9.1 -9	n13	#		*,1,*1				i (i i) 		3 22 .2	1 11 11
-14 182 711 55* 57 -16 51* -18 m,2,+10 -1 51* -18	×,•11,•4 •7 \$8* •⊾7	*****	3 3				384				: :: .:		3 5 4	1 11: 4
-10 547 65 Hullerd	*******	3 8 3	3 33 -33	**************************************	191 1911	·····		*****		1 24 - 24		a 196 197 mata.a		1
-14 50* -01 -14 51* 67	***	······			8: 4 E	· · · · · · · · · · · · · · · · · · ·	-1 141 172		· · · · · · · · · · · · · · · · · · ·	: :: .:	· ····	3 12 2	4,4,3	
**************************************	3 33 74	3 W 78	3 22 2		21: .: 21: .:		3 8 4		2		3 # 3			3 57+ -63 Ha-13a4
**************************************	······	P34-5						3 8 -8				1		
-18 54* 73 -8 44* -19 W164+12 H174+8	m76		****		*****		-2 /22 -/49			****	1 498 458 8,-9,2	3 12 19		3 944 -39 Ma-2444
	3 8 4			· · · · · · · · · · · · · · · · · · ·	···· ·		3 35 74						3 25 22	
-13 55 11 make-8	-1 10: -12		+3 183 86 m,-2,-4	··· 123 -122 ·	11. 21	121 14			-2 102 54 -1 337 -17 -1 110 -70	-1 757 -774 -1 757 -774				3 53* 51 *****
-14 55 51 -4 55 124 N(0,-10 -6 2/4 -253	3 (j. li		2 2 3	3 3 3					H.14.) -1 56- 54	n.s.,1	3 5 3	**************************************	1 11 17	
-11 56* -25 Hofferd -12 55* -1 -14 54* 14	3 33 11		-5 99 -73			******	****** -4 101 -41	*****					3 156 189 ****	1 51° 55 1 51° 5
-13 54* -4.	#,-*,-6		3 5 3			5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5 5			0 55° -17 #115.0				3 3 3	#1-1314
#4184-13 -1. 54* -3* -18 58* -5 -1 58* -4*	3 3 4	n,a,-5 -18 59+ 13	: : .i	3 35 6 3			-1 200 -121		-1 4449	1 12 - 13	******* ** *** **	-, 13, 2	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	
-8 6)* 43 4,-8,-8 -,-11,-8	-15 225 23		*,*,** -1 12 -41				****	1.32°	0 133 -129 m.10.0				1 /14	4 122 186 Martin
11 94 5% -9 5%* 3% #1+51+9 -8 55* 61		****				472	-7 52 -14	-3 231 272	3 🗱 33				4,3,3 -3 537 88	
-18 57* -23 H,22,+8 H,-3,-9 -9 56* 3	**************************************	3 # 4	H.1*	11-21-3 -9 575 -5 -7	••••••••	-2 -4• 53 #.137	-1 -11 -11	-1 327 -361 Ha-218	H15.1 -1 60* -3	1 17 44 1 19 44	*1-312 -6 58* 3	2 540 -125 11,14,2	-1 11 -141 -1 11 -141	1 22 - 21
-18 56+ 13 4,12,-4		-1 21 -21					-1 /42 /32 m.61		5 540 -27 M14.1				1 13	: :: :!
-11 54* 31 H-9,-7 -18 55* -52	******* *14 \$7* -\$3	·····		2									-,-,3	
-8 58* 22 Hu-14-8 -11 100 11 Hu-84-7	3 2 1		*.2	3 11 3 3		*. 147		-1 441 -388	····		/.7	2 575 22 15-2		1 424 A
	.06 -9 52 -12	*.35			191 -163		*.71			1 34- B *.*.1			115 -154	
	: :: ;			*,*,-3 -/		*******			*12,1 [] 201 - 21	: :: ::		2 660 -36 4, 16, 2	3 49. 24	1 118 -171 1 261 266
m,1,-9 -,-6,-7 -11 549 -12 -8 549 43	-19 33: 11	3 1 4	-18 58* -18			: :: n	-1 34 ² 161				****** ** \$7* \$7		3 22 .3	1 264 -232 Martin
-10 51- 53		-10 501 -12		-> 243 -223 Hiti-1 -0	1-5	*******	n.a1	- 199 - 29	•.•11,1 •• •1• •••			19.3		3 33 73
-41 457 23 -4 537 34	#121-6 -18 59* 27					3 6 6		-1 30 -11	-2 216 219	4.1 .7 .7. 23	-1 -1	N,-14,3		
		-5 393 383 9,5,-5	-18 807 -18	-1 102 104 -3 102 104	186 -181 	: :: ::	3 3 4	-1 52 L2 H.1.8	#1-1+11	3 3 1	n.a.z	1 35 -22	i iir is 1.1.1	1
-11 117 105 -4 51 7	*,3,**	3 # 3		2,-3	# j	*******	·····		· · · · · · · · · · · · · · · · · · ·			1 12 32	3 3 3	3 23 25
H, 4, -4 -1, 54* -59 -11 54* -1 52* -29		- 25 - 22	3 3 3		199 -212 191 -245		······································	-5 50 - 195 -5 -2573 -3 -34 - 85	41-9,1 -5 59* 4	*****				
-13 13* 137 131 144	- 12 12	H. 6, +5	-11 12 13	3 # 3	•.•3,-2 ·	*		-1 25 7 -1 25 7 -1 35 7			3 .33 .33	*******	1 104 271	
-11 547 55 -8 467 -81 -10 547 51	1 8 1			n.31 -1 -1 5446 -5		······································	*,18,-1 •8 *** -12*	4.2.0 -8 644 51	1 3 4		: :: ::	1 106 -64	****3 ***	H,-4,4
-1 51 -74 -9454-7 Nobert -14 -941 -7		3 2 3	2 13. 11		1+3 1+1 +,-2,-2	*****			: : :	1	-1 101 101			
	*.5.~5 *12 53* 17	1 1 1	14 12 12	3 # 3 3			-1 257 211 -2 116 116 -1 127 -246	-1 111 -107	1 10	: :: :	4 611 -629 6 650 667			
				: 1: .:	125 - 125 103 - 346	4141					: +us -706 +.2.2	2 96 -1-6 H16,3	, a, -46 	3 Ju28 M,-J
-12 31 -1 32 14 5-1-7	*.66 -9 51* -27		2 12 10					1 4 3			3 3 2			
			3 8 3		114 .125		-2 349 343		• • \$2 · • 2•	11 22 A				
- 14 36 17 3 6 	•	3 14 24		3 35 .3	*****		3 23 4		· · · · · · · · · · · · · · · · · · ·			*,**,3 ** *** ***		
-H 35 7 -H 65 -H		3 # 3	-,+,-,-,-	-4 129 -138 -1 -5 198 183 -1 -6 182 182 -1		*****	-1 101 44	:1 23 -22			··· 3.2			n,-2,4
Hallard	*.*.*b	÷.: .:		4,6,-1					#5.1	2 3 3		1		-1 19 -19
······································		-1 54° 57	-• ••• -••	-4 51° 72 -7 122 137 -6 166 -193 -9	****** *** -{*	471	3 55 75	-3 155 139 -2 196 245 -1 156 -165	······································		-1 114 115	: :: :		1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
-11 57 -11 -1 11 -1			-1 1 -1				-1 11 - 17	#,5,0	1 100 100 1 200 100	1 3. 1				3 1175 -1167 *****
	-11 12 13		3 5 -9	н,7,+3 -4 -1 -2 -8 -8 -2	21 22. 21 22.		-, 14, -1 -6 58* 14		-6 36* 74		 -1 -1111-	2 410 -24 11,-743	4,14,3 -4 53+ -34	-1 161 -246
4,-4,-4 -11 33 - 1 4,-4,-4 -11 33 - 1 -1 -1 - 1 -1 -1 -1 -1	-7 1.12 -1.24		**************************************									3 33 33	-1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1 -	
n,-7,-3 n,5,-7	4.286 -8 138 93	H.125	3 4 3	: # # :: ·	11	3 3 2	+.151	· · · · · · · · · ·	-1 141 124 6 324 -101	******	-2 573 616 -1 517 -567 6 989 -1083	-1 47* 1 154 341 119 -177		
-1 64* -41 -11 64* 1 -1, 52* 4 -1, 64* -1 -1, 64* -1 -1, 64* -1	-7 124 -94					-, 19+ 11 -,+5,-1		·····		1 33 .2	1 12 13	*.**.1	 	
-1 36* 15 -7 252 4 #,-5,-6 #,6,-7	: :: ::	····	3 5 4					3 8 3		: :: :	3 22 3	3 # 3		
-1, 33, -17, -11, 50, -d -1, 35, -6, -10, 52, -d	: :: :: ::		2 14. 13						1,11,11	· ····································				3 244 255 3 254 255
4,-4,-8 -7 -44 - 5 -7 -74 - 11	3 1: 3	8,14,-3 -7 5613	3 8 3	3 8 3 3		: 3: **	4,-15,4 -1 115 -113	*.*.*	12 12 13	1 50		i 131 - 111 	*****	8,2,5 -5 529 -16
-1 53+ 20 -1,7,-7 h-3,-7 -1,9 32 -10	-6 53° 21 H,13,-6	3 8: 3	: :: ::		-,-,-?		N,-14.2			•	2 225 234 Mahari	3 2 3		
-1 12* -41 -1 13* -5 -7 12* -41 -1 13* -5 -7 17* 1		3 8 3	3 15 .4	· · · · · · · · · · · · · · · · · · ·		•••••• •• ••• ••	-1 11 -11	3 8 3	33. 1					
-12 544 412 H141-7 -1 514 8 -3 524 -21	-,1-,-4 -: 11: 11	H13	: :: ::	2 11 4 3			1 11 11	1 8 3	H1-1.1			4 119 -117 m,-1,1	1 175 -145 1 134 - 15	· · · · · ·
-8 155 154 Hy+1,-8 -,4,-7 -11 58* 34	-6 564 -26 H,15,-6	** 60* -74	3 23 -11			3 2 3	n, 330 - 14 N, 1218	: :: ::		· ····		3 3 3	3 13 3	2 83 2
				-			-4 46* 47	-1 17 11	-2 31 61	*1 218 1*1		1 14 - 14	-2 52 -5	-1 -1 -15

Table 3 (cont.)

a arreada arrea A manta arreada Arreada A arreada Arread	ander versten services versten försten försten försten verste verste förete förete av som verste verste verste 1944 förstenen förstende försteten försten försten försten företen förste setrete besten förete setter beste v	and and and and and and a set of and and and and and and and a set of a set The set of a set of a set of the set of the set of the set of a set	arten erreten anderen jandenen jandenen jandenen jandenen jandenen jandenen jandenen jandenen janden janden ja 1944 janden jandenen jandenen jandenen jandenen jandenen jandenen jandenen jandenen jandenen janden janden jande		anda anamada anamada 1980 (1980-1984 (1980-1984 (1980-1984 (1980-1984 (1980-1984 (1980-1984 (1980-1984 (1980-1984 (1980-19 1944 - Adada (1980-1984 - Adada Adada Maradas Maradas Marada Marida Adada Adada Adada Adada (1980-1984 (1980-			erennend erennende er 1994deren § merendete § erennende § mitteren § ferennende § merenne årennende erennende	terre freite freit is eine state betreit betreit state state betreit b	interest int	unerente intereste in Antereste intereste in Atomierte intereste in	a control and the second state of the second s	a farant farant farant farant farant farant farant faran faran faran fara fara (ara ara ara ara ara farant far 1) arant (arant (arant (arant) arant (arant) aran (ara (ara (ara (ara (ara (ara (ara	and and the first
aradit aradi aradi aradi aradi aradit ar 1999 - Santa Sa	u version version sector to the sector of the sector secto	 ************************************	14 contactor à avenue contactor à accordant à formant à avenue contactor à de la contactor à avenue de la contactor à ave	 Marrier Lands Marrier Mar	and the second s	serve strandster standard sources and server standard sources and server sources and serv	H summer meren m	errerent errerente errerente errerente errerente errerente errerente errerent (erreretent (errerente)) errerent (errerente) errerente errerente erreretete errererente erreretete erreretetete erreretetete		image	arren farren f 1976 - Arren farren f		and a set of the set o	

A is the atom type, m refers to the number of the molecule in the asymmetric unit, and n refers to the atom position within a molecule as shown in Fig. 1. Observed and calculated structure factors after the last refinement cycle are given in Table 3.

Discussion of the structure

A stereoscopic view of the contents of the unit cell, drawn by computer program *ORTEP* (Johnson, 1965), is shown in Fig. 2. Atoms are represented by thermal ellipsoids drawn to 50% probability. Bond lengths and angles, listed in Table 4, are similar to FU molecular dimensions determined earlier (Kim & Rich, 1967; Voet & Rich, 1969).

The structural information earlier deduced from the Patterson map and the reflection statistics is indeed present: the two layers in the unit cell are identical and are translated with respect to one another by a nonlattice translation vector corresponding to the very large peak in the Patterson map; and each layer contains several pseudo-centers of symmetry (Fig. 3). The molecules within each layer are arranged in a way that suggests a fourfold rotation axis in projection. This symmetry is destroyed in three dimensions when the lavers stack upon each other.

Hydrogen bonds are drawn as dashed lines in Fig. 3. The molecules in each layer are hydrogen bonded to one another using all probable bonding sites. The lengths of these bonds (Table 5) are similar to those observed in other structures of this type.

The molecules are hydrogen bonded in a ring around the apparent fourfold axis within each layer. This kind of pattern has been observed in several other structures involving purines and pyrimidines. Hydrogen bonding in uracil (Parry, 1954) is different from this scheme,



Fig. 3. View of one molecular layer. This drawing contains the equivalent of nearly four unit cells.

	Atoms are labeled as shown in Fig. 1.								
Central atom N(1)	Atom 1 C(2)	Atom 3	Molecule 1 1·40 (3)	Molecule 2 1·39 (2)	Molecule 3 1·38 (3)	Molecule 4 1·37 (2)			
	C(6) C(2)	C(6)	1·39 (3) 122 (1)	1·37 (3) 124 (2)	1·35 (3) 125 (1)	1·43 (3) 122 (2)			
C(2)	N(1) N(3) O(7) N(1) N(1) N(3)	N(3) O(7) O(7)	$ \begin{array}{r} 1.40 (3) \\ 1.40 (2) \\ 1.20 (3) \\ 116 (2) \\ 121 (2) \\ 123 (2) \end{array} $	1·39 (2) 1·39 (4) 1·23 (3) 116 (2) 123 (2) 121 (2)	1·38 (3) 1·42 (2) 1·24 (3) 117 (2) 124 (1) 118 (2)	1.37 (2) 1.42 (3) 1.26 (3) 119 (1) 124 (2)			
N(3)	C(2) C(4)	0(1)	$ \begin{array}{r} 1.40 (2) \\ 1.39 (3) \\ 1.37 (2) \end{array} $	$ \begin{array}{r} 1 \cdot 39 (4) \\ 1 \cdot 39 (3) \end{array} $	1.42(2) 1.41(3)	1·42 (3) 1·41 (3)			
C(4)	C(2) N(3) C(5) O(8) N(3) N(3) C(5)	C(4) C(5) O(8) O(8)	$127 (2) 1 \cdot 39 (3) 1 \cdot 46 (4) 1 \cdot 24 (2) 112 (1) 122 (2) 126 (2)$	$124 (1) 1 \cdot 39 (3) 1 \cdot 49 (2) 1 \cdot 26 (3) 115 (2) 122 (1) 123 (2)$	$122 (2) 1 \cdot 41 (3) 1 \cdot 47 (3) 1 \cdot 24 (2) 115 (1) 120 (2) 125 (2)$	125 (1) 1·41 (3) 1·49 (2) 1·19 (3) 110 (2) 120 (1) 129 (2)			
C(5)	C(4) C(6) F(9) C(4) C(4) C(6)	C(6) F(9) F(9)	1·46 (4) 1·35 (2) 1·36 (2) 125 (2) 113 (1) 122 (2)	1.49 (2) 1.35 (4) 1.34 (2) 121 (2) 116 (2) 123 (1)	1.47 (3) 1.37 (2) 1.34 (2) 122 (2) 113 (1) 125 (2)	1·49 (2) 1·34 (4) 1·35 (3) 128 (2) 112 (2) 120 (1)			
C(6)	N(1) C(5) N(1)	C(5)	1·39 (3) 1·35 (2) 118 (2)	1·37 (3) 1·35 (4) 119 (1)	1·35 (3) 1·37 (2) 118 (2)	1·43 (3) 1·34 (4) 116 (1)			
O(7) O(8) F(9)	C(2) C(4) C(5)		1·20 (3) 1·24 (2) 1·36 (2)	1.23 (3) 1.26 (3) 1.34 (2)	1.24(3) 1.24(2) 1.34(2)	1·26 (3) 1·19 (3) 1·35 (3)			

Table 4. Bond lengths and angles

Table 5. Hydrogen-bond lengths

2·80 (2) Å
2.84 (3)
2.83 (3)
2.77 (2)
2.75 (1)
2.77 (3)
2.76 (3)
2.82 (2)

however. In uracil, one oxygen atom forms two hydrogen bonds to nitrogen atoms in neighboring molecules while the other forms no hydrogen bonds at all.

Unfortunately, hydrogen-bonding data on the other halogenated uracils are not yet available.

Another interesting feature of the FU structure, also shown in Fig. 3, is the localization of several fluorine atoms in the same area. Such a highly electronegative concentration is somewhat unusual. These fluorine-tofluorine distances are given in Table 6.

Table 6. Fluorine-to-fluorine distances for atomswithin the same layer

$F(1,9) \cdots F(2,9)$	3·11 (2) Å
$F(2,9) \cdots F(3,9)$	3·08 (1)
$F(3,9) \cdots F(4,9)$	3.15 (2)
$F(4,9) \cdots F(1,9)$	3.16(1)



onto (100). Black circles represent atoms in layer at $x \sim \frac{1}{4}$.

The overlap of molecular layers is shown in two views. In Fig. 4 the molecular layer at $x \sim \frac{1}{4}$ is illustrated with the layer at $x \sim -\frac{1}{4}$. Fig. 5 shows the layer at $x \sim \frac{3}{4}$ with the layer at $x \sim \frac{1}{4}$. Several types of stacking interaction are visible. In general, there is very little actual base overlap. This is typical of halogenated analogs (Bugg, Thomas, Sundaralingam & Rao, 1971). The fluorine atoms in molecules 1, 2 and 4 interact with ring atoms of type N(1) or C(5) in bases above or below them. This is also typical of halogenated analog structures. The fluorine atom in molecule 3, however, does not appear to interact with any atoms in layers above or below it, and it has no approaches less than 3.5 Å in these layers. In many cases, there are polar groups (C=O) positioned over parallel rings, but in no case does the uracil-type stacking interaction occur (Parry, 1954). This interaction involves the positioning of an N-H group over the ring of an adjacent base.

Several of the closest interlayer distances are listed in Table 7. Of special interest here are the fluorinenitrogen distances. The distance between atom F(1,9) in the layer at $x \sim \frac{1}{4}$ and atom N(3, 1) in the layer at $x \sim -\frac{1}{4}$ is especially small (2.96 Å), and suggests a strong interaction. It is possible that these fluorine-nitrogen interactions cause a disordering of several of the hydrogenatom positions resulting in difficulty in the refinement of some of the hydrogen positional parameters.

It is unfortunate that quantitative data on the stacking of 5-chloro-, 5-bromo-, and 5-iodo-uracil are not yet available, since the comparisons would be interesting. From preliminary results by Freeman, Chastain & Bugg (1970), we know that chloro- and bromo-uracil are nearly isostructural and stack so that the halogen has close contact with an atom of the ring of a parallel base. They have also reported that iodouracil stacking is like non-halogenated pyrimidine derivatives; that is, with a polar group overlapping an adjacent base. A comparison of the various stacking patterns in the FU structure seems to indicate that FU stacking patterns, themselves, are not as specific as we might have believed earlier, but that certain characteristic interactions do occur.

Conclusion

It is believed that halogenated pyrimidine derivatives effect base-stacking interactions in ways that may cause changes in possible hydrogen-bonding geometry. For example, in some cases normal Watson–Crick base pairing could be destroyed. These altered interactions

Table 7. Interlayer distances

Atom 1 at $x \sim \frac{1}{4}$	Atom 2 at $x \sim \frac{3}{4}$	Distance (Å)	Atom 1 at $x \sim \frac{1}{4}$	Atom 2 at $x \sim -\frac{1}{4}$	Distance (Å)
F(1.9)	C(4,5)	3.21(2)	F(1,9)	N(3,1)	2.96 (1)
F(1.9)	F(4.9)	3.16(1)	F(1,9)	C(3,2)	3.09 (2)
F(2.9)	N(4,1)	3.18 (1)	F(2,9)	N(2,3)	3.46 (1)
F(4,9)	C(1,4)	3.27(2)	F(4,9)	C(3,5)	3.44 (2)

Distances less than 3.3 Å and which involve fluorine atoms are listed. Two distances greater than 3.3 Å are listed since they are the only approaches that the fluorine atoms make in that direction. Atom F(3,9) makes no interlayer approaches less than 3.5 Å.

could be connected with activity of FU in some situations, for example when FU is substituted into RNA. The variety of stacking interactions observed in the FU structure demonstrates that a particular stacking geometry seems to be affected by a number of external constraints. The matter is therefore not simple, and it is difficult to generalize about the stacking geometry of FU in other environments. Interactions as strong as those observed between fluorine atoms and nitrogen atoms of adjacent pyrimidine rings have no counterpart in naturally occuring purine or pyrimidine bases. These



Fig. 5. Projection of molecular layer at $x \sim \frac{1}{4}$ and layer at $x \sim \frac{1}{4}$ onto (100). Black circles represent atoms in layer at $x \sim \frac{1}{4}$.

kinds of interaction should be investigated further, since they are liable to recur in various environments and could be connected with the mutagenic action of FU.

The author is indebted to Dr Fredrik A. Muller and to Dr Roger Hanscom for their suggestions and assistance. Additional advice which proved critical in yielding a solution to this problem was given by Drs Jerome and Isabella Karle. Support of the research program by USPHS Research Grant DE 02111 from the National Institute of Dental Research, National Institutes of Health, Bethesda, Maryland is also gratefully acknowledged.

References

- BUGG, C. E., THOMAS, J. M., SUNDARALINGAM, M. & RAO, S. T. (1971). *Biopolymers*, **10**, 175–219.
- FREEMAN, G. R., CHASTAIN, B. & BUGG, C. E. (1970). Amer. Cryst. Assoc. Meeting Program and Abstracts. Paper B7, p. 26.
- International Tables for X-ray Crystallography. (1968). Vol. III, pp. 201–209. Birmingham: Kynoch Press.
- JOHNSON, C. K. (1965). ORTEP. Report ORNL-3794, Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KARLE, J. & HAUPTMAN, H. (1956). Acta Cryst. 9, 635-651.
- KARLE, I. L. & KARLE, J. (1971). Acta Cryst. B27, 1891-1898.
- KIM, SUNG-HOU & RICH, A. (1967). Science, 158, 1046– 1048.
- MANDEL, H. G. (1969). Prog. Mol. Subcell. Biol. 1, 82-135. PARRY, G. S. (1954). Acta Cryst. 7, 313-320.
- VOET, D. & RICH, A. (1969). J. Amer. Chem. Soc. 91, 3069-3075.
- X-ray System (1972). Technical Report TR-192 of the Computer Scince Center, Univ. of Maryland. Version of June, 1972.

Acta Cryst. (1973). B29, 2556

The Crystal Structure of KIBr₂.H₂O

BY STUART SOLED AND GENE B. CARPENTER

Metcalf Chemical Laboratories, Brown University, Providence, Rhode Island 02912, U.S.A.

(Received 9 June 1973; accepted 11 June 1973)

Crystals of potassium dibromoiodide monohydrate, KIBr₂. H₂O, are orthorhombic, space group *Pnnm*, with $a = 12 \cdot 183$ (5), $b = 13 \cdot 046$ (5), $c = 4 \cdot 390$ (3) Å, Z = 4. All atoms lie on mirror planes. The structure was refined from diffractometer data by least-squares methods to $R = 5 \cdot 8$ %. The structure consists of nearly linear chains of halogen atoms extending infinitely along b. The two independent IBr₂⁻ anions are linear and symmetrical, both with I-Br bond lengths of $2 \cdot 71$ Å. The closest O \cdots Br distance is $3 \cdot 60$ Å, indicative of weak hydrogen bonding.

Introduction

Wells, Wheeler & Penfield (1892) were the first to report preparing 'potassium dibromoiodide'. Later, Cremer & Duncan (1932) proved by measurements of dissociation pressures that both hydrated and anhydrous versions of the compound exist. Although singlecrystal X-ray studies of compounds containing triiodide