Acta Cryst. (1958). **11**, 453

The Structures of the Pyrimidines and Purines. VII. The Crystal Structure of Caffeine

By D. June Sutor

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(*Received* 7 *October* 1957 *and in revised form* 4 *December* 1957)

Caffeine crystallizes in the monoclinic space group $P2_1/a$ with $a = 14{\cdot}8$, $b = 16{\cdot}7$, $c = 3{\cdot}97$ Å, $\beta = 97{\cdot}0°$. The crystal structure was solved by an application of the isomorphous-replacement method and a consideration of the possible hydrogen bond system in the crystal. A comparison is made of the intramolecular distances with those in other purines and a pyrimidine, and indicates that steric hindrance must be allowed for in a theoretical calculation of bond lengths in relatively complicated molecules of this type. Evidence for and against the existence of a short hydrogen bond between water molecules is given.

Introduction

In pursuance of the investigation of structures of pyrimidines and purines and related compounds, the crystal structure of caffeine, 1:3:7 trimethyl 2:6 dihydroxy purine (see Fig. 1) has been determined.

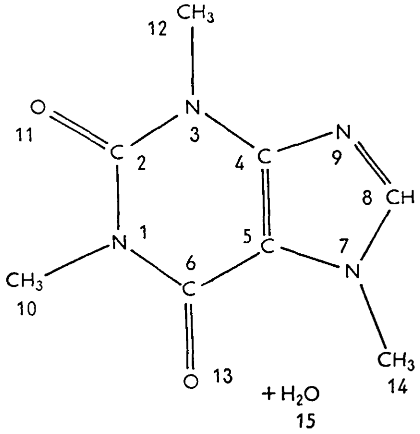


Fig. 1. Caffeine, showing the numbering system used.

Crystallographically, it is nearly isomorphous with theophylline (Fig. 4) (Sutor, 1958), and a comparison of bond lengths in the two should yield information concerning the effects of a substituent in the iminazole ring on the dimensions of the purine ring.

Experimental

Commercially obtained caffeine was recrystallized by dissolving in hot water, and on cooling feathery needles of thin cross-section were obtained. Generally, this method did not yield crystals suitable for X-ray work, but if the hot caffeine solution was allowed to cool slowly in a Dewar flask, a crystal of not too thin cross-section was sometimes obtained. Caffeine effloresces readily and within a few days gives only powder photographs; however, crystals can be preserved for longer by coating with silicone grease.

The unit-cell dimensions and space group were obtained from rotation and Weissenberg photographs:

$$a = 14{\cdot}8 \pm 0{\cdot}1,\ b = 16{\cdot}7 \pm 0{\cdot}1,\ c = 3{\cdot}97 \pm 0{\cdot}03\ \text{Å}\ ,\quad \beta = 97{\cdot}0 \pm 0{\cdot}5°\ .$$

Systematic absences are $h0l$ with $h = 2n+1$, $0k0$ with $k = 2n+1$; space group $P2_1/a\text{–}C_{2h}^5$.

The water content of crystalline caffeine has been the subject of several investigations. Waters & Beal (1946) have shown that the stable form under normal conditions is the anhydrous salt and that the water content of hydrated caffeine stable at high humidities (97–100%) corresponds to 5/6 H_2O per mole of caffeine. The existence or non-existence of the monohydrate has not been proved. The densities of three different samples of freshly prepared crystals have been determined by the method of flotation, using carbon tetrachloride and ether. The values so obtained, 1·444, 1·451 and 1·450 g.cm.$^{-3}$, correspond to molecular weights of 211·8, 212·7 and 212·7 for four molecules per unit cell (the molecular weight of the monohydrate is 212). The error in the molecular-weight determination has been estimated as 1·3% or 0·15 H_2O, by using the errors in the cell dimensions given above and by taking the error in the density determination as 0·006 g.cm.$^{-3}$; the error in weighing is negligible because of the large volume (25 ml.) of solution weighed.

The $hk0$ and $hk1$ reflexions obtainable with Cu$K\alpha$ radiation were recorded on Weissenberg photographs, using the multiple-film technique, and the equi-inclination method for the $hk1$ layer line. Intensities were measured visually and scaling factors and corrections for spot shape were applied as for theophylline.

Determination and refinement of the structure

All attempts to completely solve the crystal structure from the sharpened Patterson function failed. It was possible to obtain the two orientations of a symmetrical hexa-substituted six-membered ring, but the point of attachment of the iminazole ring and the location of the molecules in the unit cell could not be determined. Calculation of the Patterson superposition function for the hexa-substituted six-membered ring, using one of the above orientations, should pick out the molecular centre of this symmetrical group, but all possible structures based on the indicated centre would not refine, and this was later shown to be the incorrect centre. Finally, the x and y coordinates were obtained by the isomorphous-replacement method (Sutor, 1956), and refined by Fourier and then difference syntheses (Cochran, 1951), to give an R factor of 12·5% for the observed reflexions. The final Fourier projection is shown in Fig. 2. The individual scattering factors for carbon, nitrogen and oxygen of Hoerni & Ibers (1954) were employed in the calculation of structure factors, and both structure-factor and Fourier-series computations were carried out on the EDSAC. The hydrogen atoms, except those of the water molecule, were clearly visible in the difference syntheses, and their contribution to the structure factors was included in the final stages of refinement. Some of the carbon, nitrogen and oxygen atoms were given additional isotropic and anisotropic temperature factors, and values of the constants α, β and ψ (see Hughes, 1941, also Sutor, 1958) are listed in Table 1. These factors have not completely corrected for anisotropic vibration, and the R factor could probably be reduced further by changing α and β. The scaling curve is of the form $F_o = F_c \exp\,[-1{\cdot}6 \sin^2 \theta]$. The contribution of the water molecule to this projection was estimated as 0·8 from its peak height in the difference projections.

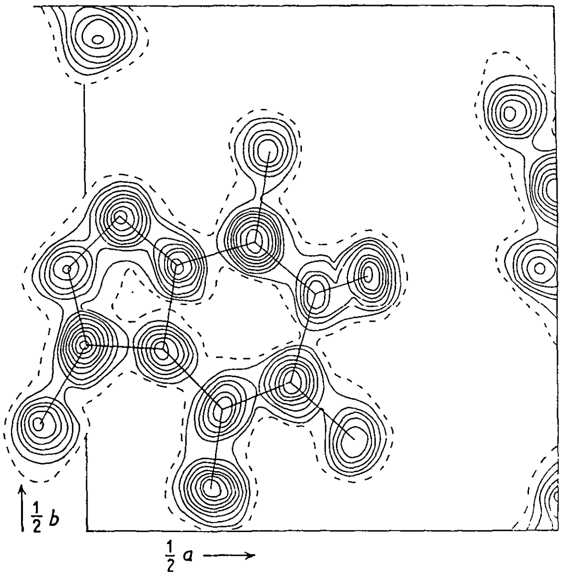


Fig. 2. Final $hk0$ Fourier projection of caffeine. Contours are at arbitrary but equal intervals; zero contour broken.

Table 1. *Individual isotropic and anisotropic temperature factors for the $hk0$ and $hk1$ layer lines*

	$hk0$			$hk1$		
	α	β	ψ (°)	α	β	ψ (°)
C_2	0·8	0·8	100	−0·4	—	—
C_4	0·8	0·8	165	−0·2	—	—
C_5	−0·3	—	—	−0·4	—	—
C_6	−0·2	—	—	−0·4	—	—
C_8	0·5	0·5	95	−0·2	—	—
C_{10}	0·5	0·5	89	—	—	—
C_{14}	0·5	0·5	79	−0·4	—	—
N_1	—	—	—	−0·4	—	—
N_3	0·5	0·5	90	−0·4	—	—
N_7	0·8	0·8	99	−0·4	—	—
N_9	−0·3	—	—	−0·2	—	—
O_{11}	0·8	0·8	89	−0·2	—	—
O_{13}	0·8	0·8	161	—	—	—
O_{15}	0·8	0·8	0	0·8	0·8	0

The z coordinates were calculated from a consideration of the possible hydrogen bond system in the crystal, and refined by hkl generalized projections (Cochran & Dyer, 1952) and hkl difference generalized projections (Sutor, 1958), the x and y coordinates from the final $hk0$ projection being employed throughout. The last difference maps showed maxima and minima which could be ascribed to the hydrogen atoms, but z coordinates were calculated for those attached to C_{10}, C_{12} and C_{14}, assuming an approximate C–H distance of 1·0 Å. The z coordinate of the hy-

Table 2. *The fractional coordinates referred to the monoclinic crystal axes*

	x	y	z	x^*	y^*
C_2	0·2416	0·2225	0·9002	0·2414	0·2225
C_4	0·1003	0·2533	0·1295	0·1019	0·2541
C_5	0·0835	0·1769	0·1944	0·0841	0·1759
C_6	0·1456	0·1151	0·1155	0·1463	0·1143
C_8	0·4798	0·2481	0·3638	0·4801	0·2480
C_{10}	0·2879	0·0841	0·8790	0·2891	0·0832
C_{12}	0·1960	0·3641	0·9209	0·1959	0·3638
C_{14}	0·4533	0·3950	0·4584	0·4536	0·3947
N_1	0·2193	0·1418	0·9735	0·2196	0·1415
N_3	0·1796	0·2764	0·9848	0·1801	0·2769
N_7	0·0009	0·1749	0·3376	0·0020	0·1749
N_9	0·0407	0·3008	0·2440	0·0403	0·3008
O_{11}	0·3057	0·2397	0·7614	0·3063	0·2400
O_{13}	0·1374	0·0404	0·1616	0·1363	0·0404
O_{15}	0·0166	0·4709	0·2705	0·0184	0·4705
H_1	0·413	0·239	0·474	—	—
H_2	0·487	0·438	0·599	—	—
H_3	0·435	0·437	0·278	—	—
H_4	0·395	0·363	0·510	—	—
H_5	0·263	0·362	0·857	—	—
H_6	0·228	0·396	0·105	—	—
H_7	0·142	0·377	0·783	—	—
H_8	0·348	0·100	0·772	—	—
H_9	0·300	0·033	0·022	—	—
H_{10}	0·257	0·060	0·676	—	—

drogen of the iminazole ring was also calculated, assuming a planar configuration with the rest of the molecule. Again the hydrogens of the water molecule could not be located. Inclusion of the contribution of the hydrogen atoms did not reduce the final R factor of 15·8% appreciably. Additional temperature factors

Table 3. *Observed and calculated structure factors*

The structure factors are on an approximately absolute scale for $F_{000} = 448$ and in the table they have been multiplied by 10. The hydrogen atoms are not included in the $\{hkl\}$ values of F_c. '*a*' indicates off film.

h	F_o	F_c
	h00	
2	52	-42
4	701	665
6	473	423
8	26	-37
10	162	-181
12	<29	-51
14	<29	15
16	<29	-3
18	51	-45
20	<18	0
	h10	
1	351	341
2	790	764
3	165	-53
4	45	-115
5	435	-420
6	99	-112
7	221	218
8	126	-116
9	63	-32
10	63	-53
11	123	-136
12	60	-50
13	108	105
14	97	61
15	<29	11
16	73	-53
17	<26	5
18	<18	3
	h20	
0	1055	-1195
1	451	409
2	131	21
3	292	-281
4	357	-373
5	265	-277
6	<21	-5
7	413	-413
8	179	199
9	45	24
10	236	-224
11	34	-54
12	128	-113
13	45	-63
14	134	136
15	171	144
16	37	-29
17	26	8
18	18	23
	h30	
1	308	-313
2	58	-58
3	545	-565
4	375	-418
5	79	-74
6	150	-152
7	53	-21
8	142	142
9	252	-260
10	76	113
11	37	-47
12	37	29
13	203	-203
14	<26	0
15	<26	0
16	149	145
17	<26	-10
18	26	26
	h40	
0	291	336
1	231	272
2	227	227
3	26	-18
4	18	-21
5	346	-379
6	131	154
7	183	173
8	53	65
9	81	61
10	97	97
11	63	47
12	79	105
13	95	-82
14	155	-150
15	37	-32
16	68	39
17	26	-26
18	45	-37
	h50	
1	294	-370
2	37	-29
3	144	-95
4	275	281
5	79	-89
6	485	-516
7	26	-21
8	134	154
9	251	278
10	45	-24
11	110	105
12	26	-37
13	134	105
14	58	21
15	63	58
16	133	-92
17	<21	0
18	<18	13
	h60	
0	378	-471
1	155	172
2	278	325
3	53	-68
4	149	97
5	32	-16
6	99	-6
7	204	-183
8	113	129
9	<24	-26
10	<26	16
11	<26	-24
12	84	86
13	<26	-31
14	102	87
15	45	-39
16	<21	5
17	76	45
18	31	18
	h70	
1	223	234
2	270	275
3	171	-165
4	228	251
5	199	202
6	307	329
7	131	-113
8	133	168
9	<23	-5
10	79	115
11	<26	-8
12	63	42
13	<26	-10
14	<26	-10
15	26	-21
16	63	53
17	<21	-11
	h80	
0	265	364
1	92	76
2	303	-320
3	260	302
4	58	-39
5	45	-39
6	37	65
7	53	86
8	95	-95
9	34	13
10	68	71
11	37	-47
12	63	-53
13	<26	-26
14	52	-40
15	<21	-26
16	<21	-5
17	37	-37
	h90	
1	240	-249
2	189	202
3	123	121
4	207	-215
5	103	-144
6	63	-55
7	44	-11
8	163	-131
9	53	-63
10	89	79
11	183	-171
12	63	-53
13	<26	5
14	26	-5
15	37	-37
16	18	11
	h,10,0	
0	226	-234
1	<21	-13
2	147	160
3	116	-89
4	31	50
5	73	-69
6	<24	37
7	123	-113
8	87	-76
9	162	157
10	173	-183
11	45	-55
12	<26	-16
13	63	-53
14	<21	-23
15	<21	-5
16	<16	-8
	h,11,0	
1	160	176
2	105	105
3	97	74
4	97	-102
5	110	115
6	102	105
7	<29	-13
8	168	168
9	63	-63
10	53	73
11	197	199
12	<26	18
13	<21	-31
14	<21	10
15	18	24
16	45	39
	h,12,0	
0	70	-24
1	<24	-3
2	105	113
3	45	50
4	74	79
5	209	-226
6	58	-58
7	<26	-26
8	<26	-3
9	102	-105
10	173	160
11	<26	-31
12	58	47
13	<21	21
14	<18	5
	h,13,0	
1	89	-102
2	50	-76
3	204	212
4	155	153
5	<26	3
6	182	-207
7	37	-37
8	58	-47
9	<26	39
10	58	-58
11	24	-24
12	<21	-5
13	<21	5
14	45	45
15	45	-37
	h,14,0	
0	215	-204
1	37	39
2	249	-267
3	<26	-47
4	81	68
5	87	105
6	63	-58
7	37	31
8	81	-84
9	26	-29
10	24	-3
11	<21	29
12	<21	-26
13	<18	-13
14	16	-18
	h,15,0	
1	82	-82
2	60	-47
3	37	-39
4	45	-45
5	37	37
6	121	131
7	42	-13
8	<26	28
9	<24	29
10	<21	11
11	53	45
12	<18	29
13	<16	5
	h,16,0	
0	<26	13
1	58	-55
2	58	63
3	81	-63
4	<26	-24
5	45	37
6	<26	-13
7	<24	-21
8	<24	-8
9	26	-24
10	52	39
11	<18	0
12	<16	0
	h,17,0	
1	87	89
2	<26	3
3	74	71
4	63	61
5	<24	13
6	39	37
7	26	26
8	26	-31
9	<18	0
10	<16	-5
11	<10	0
	h,18,0	
0	53	-63
1	<24	-3
2	<24	-10
3	<24	24
4	45	-50
5	58	47
6	21	-3
7	26	31
8	<18	5
9	<16	-8
10	45	-45
	h,19,0	
1	37	42
2	<21	-5
3	<21	-8
4	<21	-8
5	<18	-13
6	<13	0
7	<13	0
8	<13	0
	h,20,0	
0	21	-3
1	<18	-13
2	<18	13
3	<16	5
4	<16	-5
5	<13	-5
6	<13	-8
	h,21,0	
1	18	-8
2	21	-21
3	18	8
	h01	
2	702	-990
4	143	-138
6	38	-52
8	<25	-45
10	<28	28
12	55	-10
14	33	38
16	28	-52
	h11	
0	a	-407
1	113	-98
2	638	-873
3	198	-208
4	120	118
5	88	95
6	133	-143
7	125	140
8	95	-118
9	<25	10
10	<28	12
11	33	-5
12	43	-60
13	55	77
14	35	-40
15	30	-18
16	60	53
17	33	-33
	h21	
0	203	-254
1	364	-376
2	251	251
3	138	120
4	53	62
5	98	-138
6	160	195
7	<23	3
8	<25	0
9	115	-125
10	93	-105
11	<28	-23
12	<30	15
13	<28	23
14	52	43
15	73	-95
16	28	-28
17	33	38
	h31	
0	100	-85
1	128	63
2	198	200
3	60	50
4	65	60
5	62	50
6	30	-50
7	53	65
8	45	-60
9	133	-136
10	40	30
11	53	53
12	85	80
13	<28	-17
14	35	40
15	72	70
16	53	-57
17	28	30
	h41	
0	20	10
1	158	-158
2	53	-43
3	85	55
4	28	-13
5	143	135
6	73	-103
7	113	-103
8	45	25
9	45	48
10	<28	-33
11	113	135
12	<30	-10
13	68	70
14	<25	-25
15	30	43
16	55	43
	h51	
0	62	28
1	100	96
2	201	-238
3	60	58
4	23	10
5	25	8
6	115	105
7	40	-30
8	173	-178
9	88	80
10	<28	58
11	65	-55
12	<25	-15
13	<23	-13
14	<20	-10
15	35	-35
16	33	43
	h61	
0	80	-90
1	188	170
2	173	200
3	205	225
4	190	-170
5	75	55
6	77	-24
7	75	63
8	125	-130
9	77	-50
10	53	-53
11	55	-43
12	35	-35
13	33	-40
14	43	-43
15	28	-30
16	35	-18
17	25	-25
	h71	
0	360	-455
1	140	-133
2	83	123
3	93	-95
4	40	-55
5	<20	3
6	<20	3
7	90	-95
8	92	80
9	<25	15
10	<28	-40
11	<25	-27
12	<23	20
13	<23	13
14	28	-28
15	<18	3
16	35	-40
	h81	
0	140	160
1	155	-148
2	125	-135
3	70	-82
4	140	147
5	173	230
6	133	120
7	50	-82
8	80	58
9	<28	35
10	93	90
11	<27	-3
12	<25	5
13	33	-43
14	35	33
15	<18	-10
16	18	8
	h91	
0	40	35
1	68	68
2	203	-200
3	143	125
4	187	226
5	75	65
6	55	-75
7	33	58
8	33	-58
9	50	50
10	70	-55
11	<28	-15
12	<25	-13
13	72	53
	h,10,1	
0	127	-137
1	120	100
2	117	95
3	163	195
4	30	-57
5	33	-25
6	45	-33
7	33	57
8	55	-20
9	35	-38
10	55	-40
11	77	75
12	50	45
13	70	63
	h,11,1	
0	93	-110
1	48	-58
2	<25	13
3	50	-40
4	32	43
5	85	-90
6	58	-65
7	<28	35
8	<25	3
9	<25	-28
10	78	58
11	90	98
12	80	85
13	63	-63
	h,12,1	
0	118	-120
1	65	-48
2	<28	10
3	88	-123
4	<28	30
5	107	117
6	115	-127
7	113	-117
8	77	95
9	<28	-5
10	68	75
11	<25	3
12	38	-28
13	38	-40
	h,13,1	
0	135	-115
1	68	92
2	75	-82
3	82	65
4	143	-173
5	87	-68
6	73	55
7	83	-35
8	75	-68
9	<28	5
10	48	-28
11	<23	-25
12	50	-45
	h,14,1	
0	88	123
1	<28	-33
2	85	75
3	<28	25
4	117	110
5	78	-107
6	<28	-10
7	33	58
8	60	-60
9	<25	15
10	<23	-3
	h,15,1	
0	136	145
1	<28	13
2	70	68
3	80	85
4	53	58
5	33	30
6	40	-40
7	<25	-20
8	53	58
9	35	23
	h,16,1	
0	<28	-35
1	100	93
2	<28	15
3	53	53
4	33	-47
5	<28	5
6	<25	-18
7	30	28
	h,17,1	
0	<25	0
1	<25	-18
2	<25	-13
3	45	-45
4	43	-30
5	35	-40
6	<23	23
	h,18,1	
0	<23	-13
1	<23	-10
2	<23	8
3	<23	-5
4	<23	-8
5	28	-15
6	35	30
$\bar{h}$	F_o	F_c
	$\bar{h}$01	
2	228	-256
4	241	-234
6	89	-98
8	45	-63
10	105	93
12	60	-3
14	53	-10
16	<29	37
18	89	-92
	$\bar{h}$11	
1	a	73
2	272	-353
3	<16	-12
4	92	-79
5	525	-525
6	196	196
7	51	29
8	303	-240
9	101	95
10	92	92
11	149	-157
12	41	-35
13	92	54
14	44	19
15	<35	22
16	<31	35
17	<38	0
18	<25	0
	$\bar{h}$21	
1	66	35
2	177	161
3	301	282
4	79	-86
5	165	123
6	<22	-3
7	196	-155
8	278	335
9	60	72
10	158	133
11	117	-82
12	197	-158
13	167	-177
14	187	165
15	<35	-19
16	<32	16
17	<25	-28
18	44	44
	$\bar{h}$31	
1	300	256
2	342	364
3	<13	-3
4	244	227
5	105	151
6	212	-218
7	228	208
8	183	105
9	44	69
10	173	142
11	123	117
12	123	-129
13	167	-136
14	86	-101
15	<35	-16
16	60	-38
	$\bar{h}$41	
1	288	171
2	404	406
3	341	300
4	257	-247
5	107	-107
6	63	51
7	237	206
8	101	-89
9	86	101
10	92	-101
11	310	297
12	143	161
13	41	35
14	253	-218
15	132	165
	$\bar{h}$51	
1	54	-51
2	183	-123
3	136	133
4	341	310
5	139	-127
6	212	-223
7	287	-326
8	<29	0
9	136	-123
10	86	32
11	89	-107
12	189	133
13	54	-51
14	67	57
15	51	32
	$\bar{h}$61	
1	<19	13
2	94	28
3	173	193
4	486	440
5	104	83
6	70	-79
7	208	-193
8	133	89
9	67	-41
10	<35	-44
11	101	-63
12	70	-95
13	<35	-16
14	127	82
15	79	-67
	$\bar{h}$71	
1	323	313
2	253	-133
3	224	-231
4	189	-208
5	257	253
6	222	-224
7	<28	-9
8	<32	-56
9	<32	-9
10	129	-102
11	<35	13
12	54	-46
13	<32	-3
14	66	-60
15	47	-38
	$\bar{h}$81	
1	54	-79
2	258	-243
3	31	-44
4	215	-221
5	<26	-28
6	101	-86
7	89	-60
8	63	-48
9	158	-73
10	<35	10
11	<35	35
12	<31	44
13	<31	-13
13	41	-6
	$\bar{h}$91	
1	256	-266
2	174	177
3	54	57
4	44	-69
5	76	-107
6	111	89
7	92	127
8	262	-271
9	142	123
10	133	114
11	83	-25
12	<35	0
13	57	66
14	<27	0
	$\bar{h}$,10,1	
1	117	105
2	<28	-19
3	231	233
4	70	117
5	114	79
6	234	183
7	174	-127
8	123	111
9	256	266
10	<35	31
11	<35	16
12	66	48
13	41	44
14	<28	0
	$\bar{h}$,11,1	
1	76	-51
2	139	133
3	129	-105
4	88	-48
5	57	35
6	67	-69
7	79	-57
8	183	149
9	189	-180
10	82	54
11	95	104
12	41	28
13	<31	0
14	47	-47
	$\bar{h}$,12,1	
1	41	13
2	41	28
3	127	120
4	123	-126
5	<35	13
6	105	-63
7	92	101
8	244	-244
9	98	-73
10	<35	-44
11	<31	6
12	<29	28
13	75	-67
	$\bar{h}$,13,1	
1	165	-196
2	136	-136
3	158	193
4	123	139
5	120	101
6	60	-54
7	<35	-3
8	<35	-20
9	<31	46
10	52	-63
11	<29	-25
12	57	-57
13	41	31
	$\bar{h}$,14,1	
1	69	-35
2	187	-189
3	69	-72
4	206	248
5	<35	6
6	76	79
7	<31	35
8	<28	13
9	<25	0
	$\bar{h}$,15,1	
1	<35	56
2	<35	-13
3	<35	-48
4	82	-79
5	<32	32
6	60	35
7	<28	-13
8	<25	24
	$\bar{h}$,16,1	
1	111	120
2	41	78
3	51	-57
4	<35	-31
5	<31	28
6	<28	-25
	$\bar{h}$,17,1	
1	<28	-3
2	64	-69
3	<28	22
4	<25	-16
5	<25	13
	$\bar{h}$,18,1	
1	47	-60
2	<28	-22
3	44	-13
4	44	-28
5	<25	0

are listed in Table 1, but the smooth scaling curve between observed and calculated structure factors cannot be represented by a Gaussian one. This is to be expected since it takes into consideration many factors including partial correction for spot shape. The contribution of the water molecule to this projection was estimated as 0·4 from its peak height in the difference generalized projections. The final *hk*l projection was used to evaluate another set of *x* and *y* coordinates, and both sets were combined as for theophylline. The average change in *x* and *y* from the final *hk*0 coordinates is 0·0007 and 0·0004 respectively. Bond lengths and bond angles were calculated from the *hk*0 coordinates since these had probably been more accurately refined. The final coordinates are listed in Table 2, x^* and y^* represent the weighted *x* and *y* coordinates from the *hk*0 and *hk*l projections. Observed and calculated structure factors are given in Table 3; the overall *R* factor for the 602 observed reflexions is 14·6%.

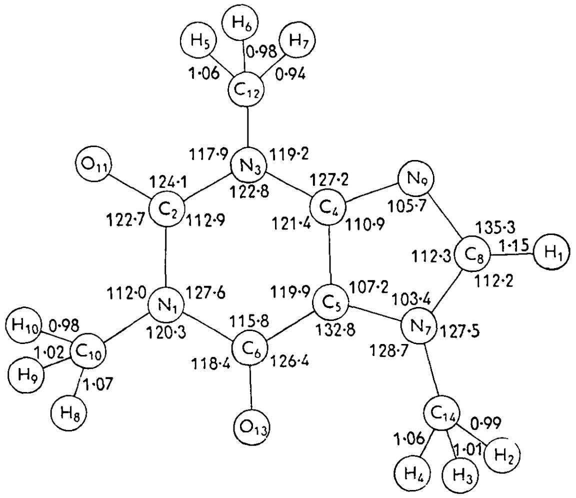
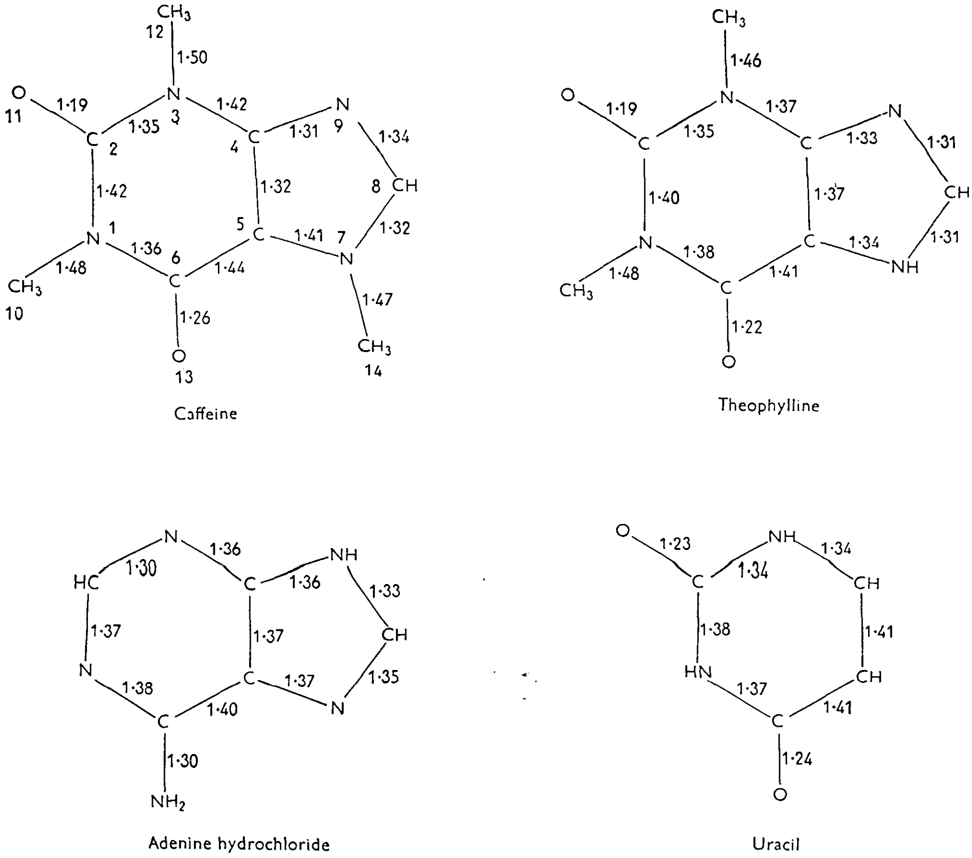


Fig. 3. Bond angles and C–H bond lengths in the caffeine molecule.

Discussion of the structure

Bond lengths and bond angles within the molecule are shown in Figs. 3 and 4. The standard deviation of the *x* and *y* coordinates was calculated by the method of Cruickshank (1949), with the following results:

$\sigma(x_r) = \sigma(y_r) = 0\cdot0081$ Å for the *hk*0 projection,
$\sigma(x_r) = \sigma(y_r) = 0\cdot0082$ Å for the *hk*l projection.



Caffeine

Theophylline

Adenine hydrochloride

Uracil

Fig. 4. Bond lengths in caffeine, theophylline, adenine hydrochloride and uracil.

$\sigma(z_r)$ cannot be estimated by this method, but a value which is probably an over-estimate can be obtained in the following way. The equation of the molecular plane can be represented by $z = -1{\cdot}800x - 0{\cdot}408y + 1{\cdot}423$. Observed z coordinates deviate from the plane by a mean of 0·026 Å, the deviations of N_9 (0·070 Å) and O_{11} (0·051 Å) being the largest. The standard deviation of z coordinates from this plane is then 0·032 Å, and this value was used for $\sigma(z_r)$. In Table 4, the standard deviation in each bond length is given, for several of the bond lengths are only slightly dependent on the difference in z coordinates.

Table 4. *Bond lengths and their standard deviations*

Bond	Bond length (Å)	Standard deviation (Å)
C_2–N_1	1·42	0·014
C_2–N_3	1·35	0·016
C_4–N_3	1·42	0·021
C_4–C_5	1·32	0·014
C_5–C_6	1·44	0·015
C_6–N_1	1·36	0·021
C_4–N_9	1·31	0·019
C_8–N_9	1·34	0·019
C_8–N_7	1·32	0·012
C_5–N_7	1·41	0·021
C_{10}–N_1	1·48	0·016
C_{12}–N_3	1·50	0·013
C_{14}–N_7	1·47	0·018
C_2–O_{11}	1·19	0·023
C_6–O_{13}	1·26	0·013

The standard deviation in a bond angle was found to be about 1·0°, from the equation of Ahmed & Cruickshank (1953).

In a discussion of the bond lengths in caffeine, it is convenient to compare them with other purines and a pyrimidine. For this purpose, adenine hydrochloride (Cochran, 1951), theophylline, and uracil (Parry, 1954) have been chosen, since they have been refined to a high degree of accuracy. Use has been made of the significance tests of Cruickshank & Robertson (1953) and any difference greater than, or equal to, 2·5σ has been taken as significant. Table 5 lists these differences, and Fig. 4 shows the bond lengths in these compounds.

It seems that the proximity of the methyl group in the iminazole ring and the keto group in the pyrimidine ring has caused a repellent force between these two, with the result that these atoms, together with the ones to which they are attached, have moved away from each other; the bond angles in caffeine support this view. This explains the lengthening of the C_5–N_7 bond in caffeine and a significant difference between the C_6–N_1 and C_5–C_6 bonds, which does not occur in theophylline or adenine hydrochloride. As in theophylline, uracil and adenine hydrochloride, the C_2–N_3 bond in caffeine is again short. Although a comparison of theophylline with uracil showed that the formation of an iminazole ring with atoms corresponding to C_4 and C_5 produces a difference in the C_4–C_5 bond which is only just significant, a comparison of caffeine with uracil shows that differences in the C_4–N_3 and C_4–C_5 bonds now become appreciable, and hence a substituted iminazole ring causing steric hindrance with the pyrimidine ring does produce changes in the pyrimidine bonds at and near the point of attachment.

Table 5. *Significantly different bond lengths in a comparison of caffeine with theophylline, adenine hydrochloride and uracil*

Caffeine with theophylline

Bond	Significant difference (Å)	Actual difference (Å)
C_4–C_5	0·04	0·05
C_5–N_7	0·05	0·07

Caffeine with uracil

Bond	Significant difference (Å)	Actual difference (Å)
C_4–N_3	0·06	0·08
C_4–C_5	0·05	0·09

Caffeine with adenine hydrochloride

Bond	Significant difference (Å)	Actual difference (Å)
C_2–N_1	0·05	0·05
C_2–N_3	0·05	0·05
C_4–N_3	0·06	0·06
C_4–C_5	0·05	0·05
C_4–N_9	0·05	0·05

As in theophylline, there is a significant difference between C–N bond in the pyrimidine and purine rings. If the short C_2–N_3 bond is again excluded, together with the shortened C_6–N_1 and lengthened C_5–N_7 bonds, then C_2–N_1 and C_4–N_3 are significantly longer than C_4–N_9, C_8–N_7 and C_8–N_9. Also as in theophylline, the average of the C–N bonds outside the ring (1·48 Å) is in agreement with the C–N single bond of 1·47 Å.

The difference between the two C=O bonds is only just significant. The C_2–O_{11} bond of 1·19 Å is in agreement with the C=O length of 1·20 Å (Vaughan & Donohue, 1952) and is similar to the C=O bonds in theophylline.

The C_4–C_5 bond of 1·32 Å is equivalent to the true C=C bond of 1·33 Å.

As in theophylline and adenine hydrochloride, the lengths of the C–C and C–N bonds in the rings cannot be explained in terms of resonance contributions. The presence of pure double and single bonds outside the rings (with the exception of C_6–O_{13}) limits the number of resonance forms, and the C_4–C_5 double bond precludes plausible resonance forms which explain the double-bond character of the C–N bonds in the iminazole ring. The structure analysis has shown that in any theoretical calculation of bond lengths the steric effects of groups of atoms must be taken into account, for this is critical in relatively complicated ring molecules like caffeine and (to a lesser extent) theophylline.

The intermolecular distances

The water molecules form zigzag chains throughout the crystal structure, and in the chains each oxygen atom is linked to each of its two nearest neighbours by a hydrogen bond through a centre of symmetry. The only other possible hydrogen bond is between the water molecule and N_9 of the iminazole ring (see Fig. 2). The lengths of these bonds, calculated from the final coordinates, are

$$O_{15}\text{–}H \cdots N_9,\ 2{\cdot}85\ \text{Å};\quad O_{15}\text{–}H \cdots O'_{15},\ 2{\cdot}36\ \text{Å};$$
$$O_{15}\text{–}H \cdots O''_{15},\ 2{\cdot}18\ \text{Å}\,.$$

There may be a somewhat larger error in the z coordinate of O_{15} since the exact amount of its contribution to the generalized projections is difficult to ascertain. If the z coordinate is taken as 0·2575, the two O–H ··· O bonds are equal (2·27 Å) and O_{15}–H ··· N_9 is 2·86 Å. The intermolecular distance of 2·27 Å assumes all the water-molecule sites can be occupied and that these molecules are regularly arranged according to the space group. In this structure determination it was shown that 80% of the water molecule was present in the *hk*0 projection, while only 40% was present in the *hk*1 projection. It seems unlikely that the conditions necessary for the stability of the 5/6 hydrate were unknowingly maintained during photography and so these values represent an average of the water present during this period, for there is every reason to believe that efflorescence was occurring at the time. Work was always started on a freshly prepared crystal but there is a certain time lag between setting the crystal and the commencement of the first photograph. This time lag was only a few hours, but this may be just sufficient for the monohydrate to change to a lower hydrate with a different crystalline form. There is no evidence of change in cell dimensions in any of the photographs, and oscillation photographs obtained within 1½ hr. of removing freshly crystallized caffeine from its mother liquor (water) indicate that the c-axis spacing is the same. In an attempt to prove whether freshly prepared caffeine crystals are an unstable monohydrate, the density determinations described in the experimental section were carried out. As a standard, the density of theophylline was measured, using the same liquids, and thus the error in the molecular-weight determination is of the same order as for caffeine; for corresponding unit-cell dimensions had been measured in the same way, and from photographs taken on the same cameras. In both cases, values corresponding to the monohydrate were consistently reproduced, which lends some support to the existence of the monohydrate, and this form has been shown to have the same c-axis spacing as the average hydrate whose structure has been determined.

If the oxygen atoms are not arranged according to the space group but are 2·7 Å apart, corresponding to a 5/6 hydrate, they should give rise to extra layer lines on c-axis photographs. There is no evidence of extra layer lines, and furthermore it has been shown that the water molecules contribute to the *hk*1 reflexions.

To be in accordance with the space group, the hydrogen atom in the O–H ··· O bond must be situated at the centre, and this is possible if the bond is short. If, however, this bond is 2·7 Å, the hydrogen atom cannot be centrally located and the arrangement of hydrogen atoms in these bonds cannot be according to the space group $P2_1/a$, for then centres of symmetry would direct two hydrogen atoms towards each other and every alternate bond would have no hydrogen atom participating in it. An O–H ··· O bond of normal length would require all the hydrogen atoms in a particular chain to be directed upwards or downwards and the space group would then be $P2_1$. However, if such an arrangement does occur the departure from $P2_1/a$ may be too small to be discernible.

In support of disordering along c, is the difference in chemical properties of caffeine and theophylline, for theophylline does not effloresce even though there are chains of water molecules running through the structure. If the water molecules are disordered in their arrangement along c, only a fraction of them would be hydrogen bonded to the caffeine molecule and thus they would tend to come off readily, in contrast to theophylline where each water molecule is firmly linked to a theophylline molecule.

The evidence at present is in favour of this short bond, but it is proposed to remeasure the unit-cell dimensions using a Geiger-counter spectrometer in order to obtain a smaller error in the molecular-weight determination, and to redetermine the z coordinate(s) of O_{15} by the use of three-dimensional methods.

I should like to thank Dr W. Cochran for helpful discussions, and Dr M. V. Wilkes for his permission to use the EDSAC. I am grateful to Mrs S. Osmond for help with intensity measurements and computations, and for preparing some of the diagrams. For financial support, I am indebted to the University of New Zealand, Newnham College and Girton College.

References

Ahmed, F. R. & Cruickshank, D. W. J. (1953). *Acta Cryst.* **6**, 385.
Cochran, W. (1951). *Acta Cryst.* **4**, 81.
Cochran, W. & Dyer, H. B. (1952). *Acta Cryst.* **5**, 634.
Cruickshank, D. W. J. (1949). *Acta Cryst.* **2**, 65.
Cruickshank, D. W. J. & Robertson, A. P. (1953). *Acta Cryst.* **6**, 698.
Hoerni, J. A. & Ibers, J. A. (1954). *Acta Cryst.* **7**, 744.
Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.
Parry, G. S. (1954). *Acta Cryst.* **7**, 313.
Sutor, D. J. (1956). *Acta Cryst.* **9**, 969.
Sutor, D. J. (1958). *Acta Cryst.* **11**, 83.
Vaughan, P. & Donohue, J. (1952). *Acta Cryst.* **5**, 530.
Waters, K. L. & Beal, G. D. (1946). *J. Amer. Pharm. Ass.* **35**, 12.