February, 1980] *Bull. Chem. Soc. Jpn.*, ***53***, *383—387 (1980)*

Structures and Properties of 5-Substituted Cytosines. II.** The Crystal Structure of 5-Methylcytosine Hemihydrate

Akio TAKENAKA,* Masuhiro KATO, and Yoshio SASADA*

Laboratory of Chemistry for Natural Products, Tokyo Institute of Technology,
Nagatsuta, Midori-ku, Yokohama 227

(Received July 30, 1979)

The crystal structure of 5-methylcytosine hemihydrate has been determined by an X-ray analysis to investigate the substitution effect on the cytosine moiety. The space group is C2/c, with unit-cell dimensions $a=17.832(4)$, $b=8.261(2)$, $c=10.641(3)$ Å, $\beta=125.12(1)°$, and $Z=8$. The hydrogen bonding scheme is quite different from that of cytosine, its monohydrate, 5-bromocytosine, and its dioxane hemi-solvate. This is ascribed to the packing effect of the methyl group. A comparison with the cytosine structure indicates some deviations in bond lengths and angles, which are due to the steric and electronic effects caused by methylation. The similar values of the C(2)–N(3)–C(4) angle in cytosine and 5-methylcytosine can be related to a resemblance of their pK_a values.

In the course of the studies on the protein: nucleic acid interaction, we have found two different interaction modes of acidic amino acids with cytosine and 5-bromocytosine.[1–4] To interpret this in terms of the molecular structure of the component, crystal structures of 5-substituted cytosines have been determined by X-ray analyses.

Previously it was revealed that the geometrical change of cytosine upon bromination at C(5) is related to its physico-chemical property in solution.[5] The present paper deals with the crystal structure of 5-methylcytosine. Since its substituent has almost the same van der Waals radius[6] as bromine but electronically opposite effect, the comparison of their structures will give some additional supports to the view mentioned in the previous paper.[5]

Experimental and Structure Determination

5-Methylcytosine was purchased from Aldrich Chemical Co. Crystals were obtained from an aqueous solution. Density was measured by flotation in a mixture of cyclohexane and carbon tetrachloride. Weissenberg photographs showed the systematic absences, hkl for $h+k=2n+1$ and $h0l$ for $l=2n+1$, indicating the space group to be C2/c or Cc. A crystal $0.3\times0.3\times0.4$ mm was used for data collection on a Rigaku four-circle diffractometer with graphite-monochromated Mo $K\alpha$ radiation ($\lambda=0.71069$ Å). The unit-cell dimensions were determined with 24 high angle reflexions. Intensities were measured in the ω-2θ scan mode with a scan width of 1.25° (in ω) plus α_1-α_2 divergence and a scan speed of 4° (in 2θ) min^{-1}. Four reference reflexions showed no significant intensity deterioration throughout the data collection. The intensities were corrected for Lorentz and polarization factors but not for absorption effects. Of 850 independent reflexions in the range $2<2\theta<55°$, 93 reflexions with F_o less than $3\sigma_P(F_o)$ were considered zero reflexions, where $\sigma_P(F_o)$ is evaluated by counting statistics; the observational threshold value, $F_{\rm lim}$, was 3.70. The standard deviations were estimated by the equation of $\sigma^2(F_o)=\sigma_P^2(F_o)+qF_o^2$, where q, 3×10^{-5}, was derived from the variations of the monitored reflexions.[7] Crystal data are summarized in Table 1.

Normalized structure factor statistics strongly indicated the space group to be centrosymmetric, C2/c. The structure was solved by the direct method and its parameters were refined by the full-matrix least-squares method. All the hydrogen atoms, found on a difference map, were included. The quantity minimized was $\sum w(|F_o|-|F_c|)^2$, with $w=1/\sigma^2(F_o)$. The zero reflexions were included in least-squares calculation by assuming $F_o=F_{\rm lim}$ and $w=w(F_{\rm lim})$; those for which $|F_c|<F_{\rm lim}$ were omitted. The final R was 0.066 ($R=0.056$ for $F_o>3\sigma(F_o)$); the maximum shift of parameters in the last cycle was 0.02σ for C, N, and O, and 0.18σ for H atoms. Atomic scattering factors used were taken from International Tables for X-ray Crystallography.[8] Atomic parameters are listed in Table 2, and observed and calculated structure factors in Table 3.[9]

TABLE 1. CRYSTAL DATA

5-Methylcytosine hemihydrate	
$C_5H_7N_3O\cdot1/2H_2O$	
Space group: C2/c	
$a=17.832(4)$ Å	$Z=8$
$b=8.261(2)$	$D_x=1.39$ g cm^{-3}
$c=10.641(3)$	$D_m=1.39$
$\beta=125.12(1)°$	
$U=1282.1(6)$ Å^3	

Results and Discussion

Crystal Structure. Figure 1 shows the crystal structure. Hydrogen bond distances and angles are given in Table 4 (as for atomic numbering, see Fig. 2). Two N(4)–H(4A)···N(3) hydrogen bonds between molecules related by the inversion at (1/4, 1/4, 1/2) constitute a cytosine dimer. This pattern is similar to the dimer component of the cytosine ribbon found in 5-bromocytosine: *N*-tosyl-L-glutamic acid and 5-bromocytosine: *N*-phthaloyl-DL-glutamic acid complexes.[1,2] The carbonyl oxygen, O(2), is an acceptor of hydrogen bond from the remaining N(4)–H(4B) group of the adjacent molecule to form a layer of the dimers. The water molecule, the oxygen, O(W), of which is located on the two-fold axis, links the cytosine layers with two O(W)–H(W)···O(2) and two O(W)···H(1)–N(1) hydrogen bonds in tetrahedral disposition.

The hydrogen bonding scheme found in the present crystal is quite different from that of cytosine,[10] its monohydrate,[11] 5-bromocytosine,[5] and its hemi-dioxane solvate,[12] in which cytosines form a ribbon by the two hydrogen bonds, N(4)–H···O(2) and N(3)···H–N(1), along the 2_1 axis. This difference is ascribed to the

---

** Part I of this series is Ref. 5.

TABLE 2. FINAL POSITIONAL AND THERMAL PARAMETERS
Standard deviations are given in parentheses. The anisotropic thermal factor has the form $\exp[-(\beta_{11}h^2+\beta_{22}k^2+\beta_{33}l^2+\beta_{12}hk+\beta_{13}hl+\beta_{23}kl)]$.

Atom	x*	y*	z*	β_{11}**	β_{22}**	β_{33}**	β_{12}**	β_{13}**	β_{23}**
N(1)	35236(17)	12576(35)	44937(25)	295(11)	1668(44)	580(26)	186(34)	263(28)	335(52)
C(2)	36833(18)	16817(36)	34169(29)	259(11)	1358(41)	647(29)	82(34)	320(30)	−92(54)
N(3)	29618(16)	19130(33)	19417(25)	278(10)	1583(40)	683(27)	56(31)	439(29)	222(50)
C(4)	21085(17)	17728(32)	15676(29)	267(11)	1029(34)	706(29)	49(30)	401(30)	−63(49)
C(5)	19115(18)	13174(36)	26539(30)	304(11)	1279(39)	715(29)	22(35)	517(32)	85(55)
C(6)	26538(20)	10495(37)	40990(30)	361(13)	1373(42)	710(30)	165(36)	618(34)	306(56)
O(2)	44933(14)	18279(34)	38273(26)	259(9)	2121(44)	978(27)	−60(31)	−77(26)	−264(55)
N(4)	14207(17)	20461(38)	1015(28)	265(11)	2032(52)	688(28)	134(37)	376(30)	495(59)
C(7)	9478(22)	10805(51)	22025(41)	342(14)	1941(61)	1234(43)	81(47)	787(42)	461(87)
O(W)	50000	2345(44)	75000	315(13)	1575(52)	864(37)	0	442(38)	0
Atom	x**	y**	z**	B(/Å^2)					
H(1)	4033(36)	1019(61)	5485(60)	2.93(0.93)					
H(4A)	1567(27)	2308(48)	−459(48)	1.30(0.67)					
H(4B)	855(33)	2149(51)	−204(52)	1.99(0.76)					
H(6)	2613(30)	669(51)	4925(51)	1.79(0.74)					
H(7A)	947(33)	780(58)	3117(56)	2.72(0.90)					
H(7B)	582(41)	2088(72)	1688(68)	4.02(1.18)					
H(7C)	669(39)	226(68)	1493(66)	3.35(1.03)					
H(W)	5204(37)	−323(62)	7163(62)	3.09(1.04)					

* $\times 10^5$, ** $\times 10^4$.

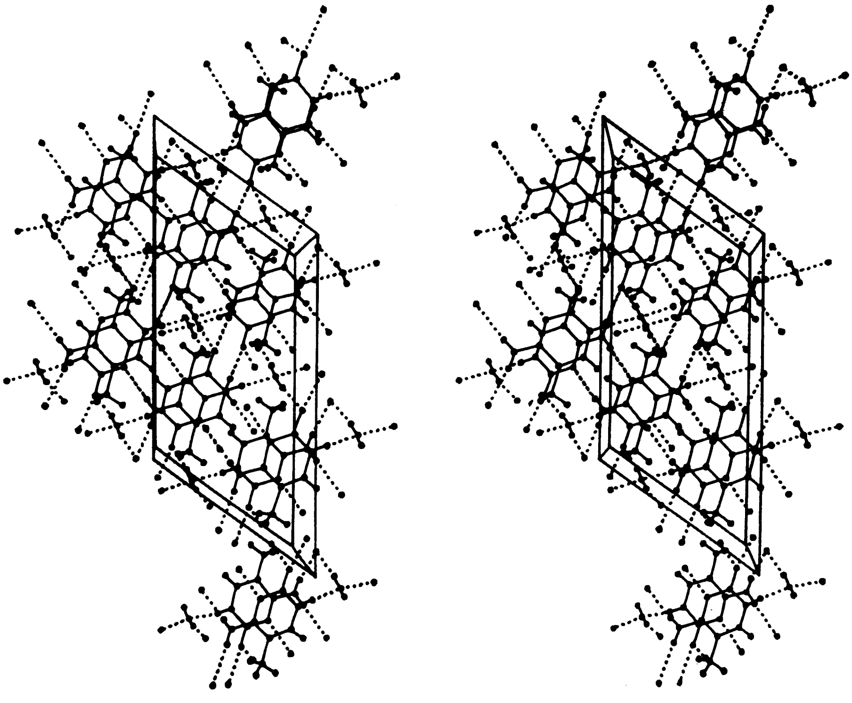

Fig. 1. Stereoscopic diagram showing the molecular packing and the hydrogen-bond network. The origin is at the upper right rear corner, ***b*** points to the viewer, ***a*** downward, and ***c*** leftward.

steric effect of methyl group at C(5) atom. A hypothetical model of 5-methylcytosine crystal was constructed using the parameters of 5-bromocytosine crystal by an interactive computer graphics,[13] and intermolecular atomic distances were calculated by rotating the methyl group around the C(5)–C(7) bond. Such an examination showed that the hydrogen atoms of the methyl group could be put considerably far from the surrounding atoms by an appropriate rotation, owing to the shorter C–CH_3 bond distance than C–Br bond, though van der Waals radii of bromine and methyl group are nearly the same.[6] In this crystal, therefore,

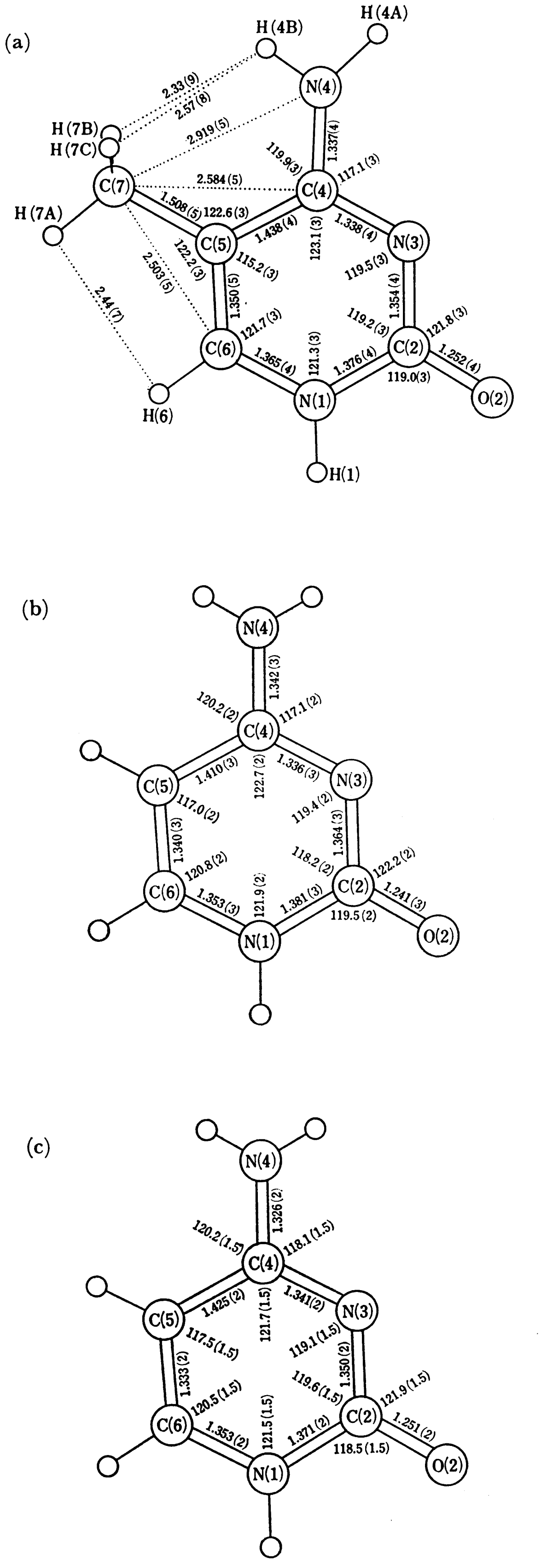

Fig. 2. Bond lengths (l/Å) and angles (ϕ/°), and their standard deviations in parentheses. (a) The present work, 5-methylcytosine. Some non-bonded distances (l/Å) from the methyl group are also shown. (b) Cytosine. (c) Cytosine monohydrate.

TABLE 4. HYDROGEN BOND DISTANCES AND ANGLES
Standard deviations are given in parentheses.

Distances (l/Å)	
$N(4)\cdots N(3)^a$	3.072(4)
$H(4A)\cdots N(3)^a$	2.27(4)
$N(1)\cdots O(W)$	2.861(3)
$H(1)\cdots O(W)$	1.93(5)
$N(4)\cdots O(2)^b$	3.026(4)
$H(4B)\cdots O(2)^b$	2.19(4)
$O(2)\cdots O(W)^c$	2.680(3)
$O(2)\cdots H(W)^c$	1.90(5)

Angles (ϕ/°)	
$N(4)–H(4A)\cdots N(3)^a$	177(4)
$N(4)–H(4B)\cdots O(2)^b$	161(5)
$C(4)–N(4)\cdots N(3)^a$	114.3(2)
$C(4)–N(4)\cdots O(2)^b$	128.2(2)
$N(3)^a\cdots N(4)\cdots O(2)^b$	112.0(1)
$C(2)–N(3)\cdots N(4)^a$	111.8(2)
$C(2)–N(3)\cdots H(4A)^a$	111(1)
$C(4)–N(3)\cdots N(4)^a$	128.6(2)
$C(4)–N(3)\cdots H(4A)^a$	129(1)
$C(2)–O(2)\cdots N(4)^d$	166.3(2)
$C(2)–O(2)\cdots H(4B)^d$	161(1)
$C(2)–O(2)\cdots O(W)^c$	115.2(2)
$C(2)–O(2)\cdots H(W)^c$	113(2)
$N(4)^d\cdots O(2)\cdots O(W)^c$	77.9(1)
$H(4B)^d\cdots O(2)\cdots H(W)^c$	85(2)
$N(1)–H(1)\cdots O(W)$	172(5)
$C(2)–N(1)\cdots O(W)$	120.5(2)
$C(6)–N(1)\cdots O(W)$	117.3(2)
$O(W)–H(W)\cdots O(2)^c$	171(6)
$N(1)\cdots O(W)\cdots N(1)^e$	145.6(1)
$N(1)\cdots O(W)\cdots O(2)^c$	88.3(1)
$N(1)\cdots O(W)\cdots O(2)^f$	113.9(1)
$O(2)^c\cdots O(W)\cdots O(2)^f$	101.1(1)
$H(1)\cdots O(W)–H(W)$	92(5)
$H(1)\cdots O(W)–H(W)^e$	111(4)
$H(1)\cdots O(W)\cdots H(1)^e$	141(2)

Symmetry codes
(a) $1/2-x,\ 1/2-y,\ -z$
(b) $-1/2+x,\ 1/2-y,\ -1/2+z$
(c) $1-x,\ -y,\ 1-z$
(d) $1/2+x,\ 1/2-y,\ 1/2+z$
(e) $1-x, y, 3/2-z$
(f) $x,\ -y,\ 1/2+z$

the packing of 5-methylcytosine molecules is thought to be looser than that of 5-bromocytosine.† Such a hypothetical loose packing of 5-methylcytosine is actual-

---

† Crystal structure of 5-bromocytosine closely resembles that of cytosine. In both crystals, the molecules form ribbons with hydrogen bonds along the two-fold screw axis. The only difference is the lacking of inter-ribbon hydrogen bonds in the former due to bromine substituents.[5] It is argued, therefore, that the former crystal is less stable than the latter one. This is supported by the fact that recrystallization of 5-bromocytosine from a water–dioxane mixture gives a complex with dioxane, in which the dioxane molecule links the 5-bromocytosine ribbons.[12] The screwed ribbon structure is thought to be barely allowed to be taken in the 5-bromocytosine crystal which was the basis of the present simulation.

TABLE 5. BOND LENGTHS AND ANGLES INVOLVING HYDROGEN ATOMS
Standard deviations are given in parentheses.

Distance (l/Å)		Angle(ϕ/°)	
N(1)-H(1)	0.94(6)	C(2)-N(1)-H(1)	117(4)
N(4)-H(4A)	0.81(5)	C(6)-N(1)-H(1)	121(4)
N(4)-H(4B)	0.87(5)	C(4)-N(4)-H(4A)	116(3)
C(6)-H(6)	0.97(5)	C(4)-N(4)-H(4B)	123(4)
C(7)-H(7A)	1.01(6)	H(4A)-N(4)-H(4B)	120(5)
C(7)-H(7B)	1.00(7)	C(5)-C(6)-H(6)	123(3)
C(7)-H(7C)	0.94(7)	N(1)-C(6)-H(6)	115(3)
O(W)-H(W)	0.79(6)	C(5)-C(7)-H(7A)	111(3)
		C(5)-C(7)-H(7B)	109(4)
Symmetry code		C(5)-C(7)-H(7C)	109(4)
(e) $1-x, y, 3/2-z$		H(7A)-C(7)-H(7B)	112(5)
		H(7A)-C(7)-H(7C)	107(5)
		H(7B)-C(7)-H(7C)	108(6)
		H(W)-O(W)-H(W)^{e}	109(5)

ly forbidden and replaced by a different hydrogen-bond network involving the water in order that the molecules pack more effectively.

Molecular Structure. Bond lengths and angles are shown in Fig. 2 together with those of cytosine[14] and its monohydrate[14] for comparison. The bond lengths and angles involving the hydrogen atoms are given in Table 5. The deviations from the cytosine values may be attributed to the steric and electronic effects caused by methylation. After an interpretation of benzene ring deformation on the basis of the valence-shell electron-pair repulsion theory,[15] the electron-releasing group attached to C(5) atom makes the C(4)-C(5)-C(6) angle smaller. In addition, the lengthening of C(5)-C(4) and C(5)-C(6) bonds, which are due to repulsions between C(7) and C(4) or C(6) atoms, may allow to narrow the C(4)-C(5)-C(6) angle accompanied by the slight expansions of C(5)-C(4)-N(3) and C(5)-C(6)-N(1) angles. Although the distance between the C(7) and N(4) atoms looks rather short, there are no abnormal contacts between their hydrogen atoms. Steric hindrance between the methyl and amino groups is relieved by an appropriate rotation of the former around the C(5)-C(7) bond, so that the deviation of C(5)-C(4)-N(4) angle becomes small. The steric effect of the methyl group thus is buffered and does not extend to N(3), while that of bromine in 5-bromocytosine makes the C(2)-N(3)-C(4) angle slightly expand. The C(2)-N(3)-C(4) angle, which is varied little by the methyl substitution, suggests no significant change of the effective charge of lone-pair lobe at N(3), and this seems to be related to similar $\mathrm{p}K_\mathrm{a}$ values of cytosine and 5-methylcytosine [4.58 and 4.6, respectively].[16] Table 6 shows the atomic net charges calculated by CNDO approximation[17] for cytosine and its 5-methyl derivative. Invariant charge at N(3) upon methylation supports the above interpretation.

As compared with cytosine structure, the shortening of the C(2)-N(3) and C(4)-N(4) bonds, and the lengthening of the C(2)-O(2) bond accompanied by the increase of N(1)-C(2)-N(3) angle are observed. These changes, though small, suggests the increase of the contribution from the canonical formula in which the amino nitrogen and the carbonyl oxygen carry positive and negative charges, respectively. Such an increase may be induced by the strong hydrogen bond from water molecule to the carbonyl oxygen. Similar trend is also observed in cytosine monohydrate, in which the water molecule participates in the same way.[11,14]

TABLE 6. NET CHARGE DISTRIBUTION

Cytosine		5-Methylcytosine	
N(1)	−0.182	N(1)	−0.183
C(2)	0.417	C(2)	0.415
O(2)	−0.425	O(2)	−0.426
N(3)	−0.337	N(3)	−0.334
C(4)	0.322	C(4)	0.308
N(4)	−0.259	N(4)	−0.259
C(5)	−0.172	C(5)	−0.111
C(6)	0.187	C(6)	0.160
H(5)	0.135	C(7)	−0.024
H(1)	0.031	H(1)	0.134
H(4A)	0.135	H(4A)	0.147
H(4B)	0.136	H(4B)	0.135
H(6)	0.001	H(6)	0.002
		H(7A)	0.009
		H(7B)	0.013
		H(7C)	0.014

TABLE 7. LEAST-SQUARES PLANE FOR PYRIMIDINE RING
X, Y, and Z are in Å along $\boldsymbol{a}^*$, $\boldsymbol{b}$, and $\boldsymbol{c}$, respectively. Asterisks denote atoms defining the plane. Standard deviations are given in parentheses.

Equation

$$-0.0110(13)X+0.96519(33)Y+0.2613(12)Z-1.2394(54)=0$$

Deviations (l/Å) of atoms from the plane

N(1)*	0.012	C(7)	−0.035
C(2)*	0.005	H(1)	−0.05
N(3)*	−0.016	H(6)	−0.08
C(4)*	0.011	H(4A)	0.03
C(5)*	0.006	H(4B)	0.17
C(6)*	−0.017	H(7A)	−0.02
O(2)	0.005	H(7B)	0.73
N(4)	0.016	H(7C)	−0.83

The least-squares plane of 5-methylcytosine together with the deviations of atoms from the plane is listed in Table 7. Although the pyrimidine ring with amino nitrogen N(4) and carbonyl oxygen O(2) is planar within 0.017 Å, the small deviations of atoms show the slight bending of the ring around the N(3)–C(6) axis. The methyl C(7) atom deviates on the opposite side of the plane to the amino group. This deviation is rather small as compared with that of bromine in 5-bromocytosine because of the relief of short contacts by an appropriate rotation of the methyl group mentioned above.

Figure 1 was drawn by TSD: XTAL, which is a computer-graphics interactive modelling programme for the NOVA 3 computer.[13)]

The present work was partially supported by a Grant-in-Aid for Scientific Research from the Ministry of Education and by the Kawakami Foundation, to which the authors' thanks are due.

References

1) M. Ohki, A. Takenaka, H. Shimanouchi, and Y. Sasada, *Bull. Chem. Soc. Jpn.*, **49**, 3493 (1976).

2) M. Ohki, A. Takenaka, H. Shimanouchi, and Y. Sasada, *Bull. Chem. Soc. Jpn.*, **50**, 90 (1977).

3) A. Takenaka, M. Ohki, and Y. Sasada, 36th National Meeting of the Chemical Society of Japan, Osaka, April 1977, Abstr. I, 4M10.

4) T. Hata, M. Yoshikawa, S. Sato, and C. Tamura, *Acta Crystallogr., Sect. B*, **31**, 312 (1975).

5) M. Kato, A. Takenaka, and Y. Sasada, *Bull. Chem. Soc. Jpn.*, **52**, 49 (1979).

6) L. Pauling, "The Nature of Chemical Bond," 3rd ed, Cornel Univ. Press, New York (1960), p. 260.

7) L. E. McCandlish and G. H. Staut, *Acta Crystallogr., Sect. A*, **31**, 245 (1975).

8) "International Tables for X-Ray Crystallography," Kynoch Press, Birmingham (1974), Vol. IV, p. 71.

9) Table 3 has been deposited as Document No. 8008, at the Office of the Bulletin of the Chemical Society of Japan.

10) D. L. Barker and R. E. Marsh, *Acta Crystallogr.*, **17**, 1581 (1964).

11) G. A. Jefferey and Y. Kinoshita, *Acta Crystallogr.*, **16**, 20 (1963).

12) M. Kato, A. Takenaka, and Y. Sasada, unpublished.

13) A. Takenaka and Y. Sasada, Annual Meeting of the Crystallographic Society of Japan, Hiroshima, November 1978, Abstr. p. 32.

14) R. J. McClure and B. M. Craven, *Acta Crystallogr., Sect. B*, **29**, 1234 (1973).

15) A. Domenicano, A. Vaciago, and C. A. Coulson, *Acta Crystallogr., Sect. B*, **31**, 1630 (1975).

16) A. Albert, "Synthetic Procedures in Nucleic Acid Chemistry," ed by W. W. Zorbach and R. S. Tipson, John Wiley and Sons, Inc., New York (1973), p. 1.

17) J. A. Pople and D. L. Beveridge, "Approximate Molecular Orbital Theory," McGraw-Hill Book Company, New York (1970).