Structure and Stereochemistry of Nucleic Acid Components and Their Reaction Products. IV.* Crystal Structure of the Potassium Salt of *N*-(Purin-6-ylcarbamoyl)glycine Monohydrate

By R. Parthasarathy† and Jean M. Ohrt

Center for Crystallographic Research, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, New York 14263, U.S.A.

and Girish B. Chheda

General Clinical Research Center, Roswell Park Memorial Institute, 666 Elm Street, Buffalo, New York 14263, U.S.A.

The potassium salt of *N*-(purin-6-ylcarbamoyl)glycine, a hypermodified base in t-RNA, was crystallized as a monohydrate ($C_8H_7O_3N_6K.H_2O$). The crystals are monoclinic, space group $P2_1/c$ with cell parameters (at $22 \pm 3 °C$): $a = 14.063$ (3), $b = 7.218$ (1), $c = 14.424$ (1) Å, $\beta = 129.12$ (1)° with four molecules per unit cell. Three-dimensional data (2242 reflexions to the limit $2\theta = 160°$ for Cu $K\alpha$ radiation) were collected on a diffractometer. The structure was solved by the multisolution technique and refined to an R of 0.058. Significant structural features are: (i) an internal hydrogen bond involving N(glycine)–H···N(1) adenine, (ii) 'distal' conformation of the substituent on N(6), (iii) the stacking and extensive overlapping of the ureidopurine moiety in a head-to-tail fashion, in planes 3.2 Å apart and (iv) a sevenfold coordination of the potassium ion. All polar hydrogens take part in hydrogen bonding. There is little self association of amino acid or nucleic acid moieties through hydrogen bonding. The stacking of the bases is stabilized by a hydrogen bond between N(9) of one molecule and the carboxyl oxygen belonging to another molecule stacked on top of the first molecule.

Introduction

Various hypermodified nucleic acid bases such as N^6-(Δ^2-isopentenyl)adenine (IPA), its 2-methyl thio analogue (2-MT-IPA) *etc.*, have been found in a position adjacent to the 3'-end of the anticodon in several t-RNA's whose codons begin with uracil (Hall, 1971; Nishimura, Yamada & Ishikura, 1969). *N*-(Purin-6-ylcarbamoyl)-L-threonine (PCT) occupies an analogous position in several t-RNA's which respond to codons beginning with adenine (Chheda *et al.*, 1969; Schweizer, Chheda, Baczynskyj & Hall, 1969; Takemura, Murakami & Miyazaki, 1969; Ishikura, Yamada, Muraro, Saneyoshi & Nishimura, 1969). Although its glycine analogue *N*-(purin-6-ylcarbamoyl)glycine (PCG) has been isolated from enzyme digests of unfractionated yeast (Schweizer, McGrath & Baczynskyj, 1970) it has not been determined whether its position in the anticodon is analogous to that of PCT. The present work is a continuation of our investigations of the relationship of the three-dimensional structure of these modified bases to their biological activity (Parthasarathy, Ohrt & Chheda, 1974*b,c*) and describes the crystal structure of the monohydrate of the potassium salt (I) (PCGK) of PCG; the conformation of the molecule and some biological aspects of the results have been reported elsewhere (Parthasarathy, Ohrt & Chheda, 1974*a*).

---

* Part III: Parthasarathy, Ohrt & Chheda (1974*c*).
† Author to whom correspondence should be addressed.
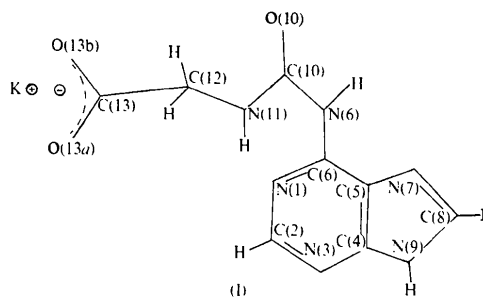
Experimental

A synthetic sample of PCG (Chheda & Hong, 1971) was used as the starting material to prepare the potassium salt of PCG. Five mmol of PCG were suspended in 100 ml of water and $0.1N$ KOH was added until the *p*H was 7.8. The solution was evaporated to dryness. PCG was crystallized as the monohydrate ($C_8H_7O_3N_6K.H_2O$) from a 1:3 mixture of propanol and water. The unit-cell constants were refined (and their standard deviations estimated) by a least-squares refinement of the 2θ values of 41 reflections at large 2θ values such that the peaks from Cu $K\alpha_1$ and Cu $K\alpha_2$ could be distinguished. The relevant crystal data are given in Table 1. Complete three-dimensional intensity data (2242 non-equivalent reflections to the limit $2\theta = 160°$) were measured on a GE XRD-490 automatic diffractometer by the stationary crystal–stationary counter technique (Furnas & Harker, 1955) with a

5° take-off angle. Balanced Ni–Co Ross filters were used for the monochromatization. The 2039 reflections in the above (sin θ/λ) range with their intensities twice the background were used for the structural refinement. The crystal used for the data collection had the dimensions $0.22 \times 0.56 \times 0.23$ mm and was mounted with $-\mathbf{b}^*$ along the φ axis of the goniostat. The difference in absorption as a function of φ was measured for the axial reflections and was used to approximately correct the anisotropy of absorption. This correction was about 15% for most reflections and up to 36% for reflections making a small angle with $\mathbf{a}^*$. The data were processed in the usual way.

Determination and refinement of the structure

All 19 non-hydrogen atoms were located in an E map based on phases determined by the application of a multisolution technique (Karle & Karle, 1966; Germain, Main & Woolfson, 1971) with the program *MULTAN. R* $(\sum||F_o|-|F_c||/\sum|F_o|)$ for this structure was 0.35. The positional and thermal parameters of the 19 non-hydrogen atoms were subjected to several cycles of least-squares refinement, employing a block-diagonal approximation. The nine hydrogen atoms were located from an electron density difference map after R had reached 0.076. Inclusion of their positional and thermal parameters in the least-squares refinement lowered R to 0.058. Refinement was terminated at the point when the shifts of the atomic parameters were a small fraction (of the order of 0.1 for the non-hydrogen atoms and 0.3 for the hydrogen atoms) of the corresponding e.s.d.'s.*

The observations were weighted according to the scheme of Evans (1961) based on counting statistics and the refinement was carried out by minimizing $w[|F_o|-(1/k)|F_c|]^2$. Atomic scattering factors for K^+, O, N and C and the dispersion correction for K^+ were those listed in *International Tables for X-ray Crystallography* (1968). For the hydrogen atoms, the values given by Stewart, Davidson & Simpson (1965) were used.

Discussion of the structure

The final positional and thermal parameters for non-hydrogen atoms are given in Table 2 and those for

---

* A list of structure factors and tables of the least-squares planes and stacking contacts close to the sum of the van der Waals radii have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 31765 (18 pp., 1 microfiche). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

Table 1. *Crystal data*

$C_8H_7O_3N_6K \cdot H_2O$

$a = 14.063$ (3) Å	$V = 1136$ Å³
$b = 7.218$ (1)	$M = 292.30$
$c = 14.424$ (1)	$\varrho_{obs} = 1.71$ g cm⁻³
$\beta = 129.12$ (1)°	$\varrho_{calc} = 1.71$
Space group: $P2_1/c$	$Z = 4$
λ(Cu $K\alpha_1$) = 1.54051 Å	
$\mu = 43.2$ cm⁻¹	
Temperature $22(\pm 3)$°C	

Fig. 1. Bond distances in PCGK. (All figures were drawn with the *ORTEP* program.)

Table 2. *Positional and thermal parameters* ($\times 10^4$)

$T_F = \exp[-(b_{11}h^2 + b_{22}k^2 + b_{33}l^2 + b_{12}hk + b_{13}hl + b_{23}kl)]$. Estimated standard deviations are in parentheses.

	x	y	z	b_{11}	b_{22}	b_{33}	b_{12}	b_{13}	b_{23}
K	−608 (1)	1776 (1)	7226 (1)	33 (1)	109 (2)	51 (1)	6 (1)	51 (1)	2 (2)
O(W)	−181 (4)	4050 (5)	6044 (3)	109 (4)	170 (7)	80 (3)	117 (9)	145 (6)	93 (8)
N(1)	5037 (3)	2077 (4)	9464 (3)	33 (2)	95 (6)	41 (2)	1 (6)	51 (4)	0 (6)
C(2)	6204 (4)	2660 (6)	10166 (3)	34 (3)	106 (7)	52 (3)	7 (7)	57 (5)	5 (7)
N(3)	6845 (3)	3512 (5)	11216 (3)	34 (2)	101 (6)	48 (2)	−8 (6)	47 (4)	1 (6)
C(4)	6164 (3)	3773 (5)	11561 (3)	34 (3)	68 (6)	40 (3)	4 (7)	38 (5)	13 (6)
C(5)	4962 (3)	3214 (5)	10946 (3)	36 (3)	75 (6)	37 (2)	5 (7)	44 (4)	8 (6)
C(6)	4383 (3)	2347 (5)	9834 (3)	32 (3)	67 (6)	35 (2)	3 (6)	39 (4)	8 (6)
N(6)	3192 (3)	1791 (4)	9162 (3)	28 (2)	109 (6)	41 (2)	−12 (6)	49 (4)	−20 (6)
N(7)	4548 (3)	3688 (5)	11571 (3)	47 (3)	105 (3)	46 (2)	−12 (7)	63 (4)	−10 (6)
C(8)	5495 (4)	4516 (6)	12517 (4)	53 (3)	108 (7)	50 (3)	−8 (8)	70 (5)	−10 (7)
N(9)	6498 (3)	4608 (5)	12579 (3)	40 (3)	96 (6)	41 (2)	−13 (6)	38 (4)	−14 (6)
C(10)	2428 (3)	928 (4)	8035 (3)	26 (3)	67 (6)	39 (2)	1 (6)	33 (4)	8 (6)
O(10)	1410 (2)	422 (4)	7649 (3)	31 (3)	131 (6)	61 (2)	−32 (6)	56 (4)	−37 (6)
N(11)	2882 (3)	688 (5)	7459 (3)	27 (2)	117 (6)	38 (2)	−17 (6)	42 (4)	−7 (6)
C(12)	2122 (4)	73 (5)	6227 (3)	43 (3)	98 (7)	36 (3)	−15 (7)	48 (5)	−6 (7)
C(13)	1641 (3)	1657 (5)	5327 (3)	30 (3)	104 (7)	38 (2)	−16 (7)	43 (4)	−12 (7)
O(13a)	1562 (3)	3248 (4)	5652 (2)	49 (2)	101 (5)	44 (2)	19 (6)	47 (4)	−4 (5)
O(13b)	1354 (3)	1275 (4)	4343 (2)	64 (3)	142 (6)	38 (2)	−6 (7)	60 (4)	−12 (6)

hydrogen atoms are given in Table 3. The bond lengths and angles involving non-hydrogen atoms are illustrated in Figs. 1 and 2. Bond distances and angles involving the hydrogen atoms fall in the usual range for X-ray determinations. The average e.s.d.'s in the bond distances and angles involving non-hydrogen atoms are 0·006 Å and 0·4° respectively. When one of the atoms involved is a hydrogen, the average e.s.d.'s are 0·06 Å and 3° respectively.

Table 3. *Final positional* (× 10^3) *and thermal parameters* (× 10) *for hydrogen atoms*

	x	y	z	B
H(C2)	662 (4)	242 (7)	982 (4)	27 (9)
H(N6)	285 (4)	210 (7)	941 (4)	23 (9)
H(C8)	545 (5)	502 (8)	1304 (4)	32 (11)
H(N9)	721 (4)	524 (6)	1317 (4)	18 (8)
H(N11)	359 (4)	115 (6)	782 (4)	15 (8)
Ha(C12)	139 (5)	−48 (7)	608 (4)	29 (10)
Hb(C12)	263 (4)	−70 (6)	613 (4)	21 (9)
Ha(OW)	−73 (5)	504 (8)	543 (5)	36 (12)
Hb(OW)	30 (5)	381 (7)	598 (4)	50 (10)

The bond distances show that the carboxyl group of the glycine moiety in PCGK is ionized as expected, and the N(9) rather than the N(7) position of the adenine moiety is protonated. Other bond distances worth noting are: (i) the increase in the length of the C(6)–N(6) bond (1·362 Å) found in several N(6) substituted adenines; (ii) the inequality of the two C(sp^2)–N(sp^2) bonds in the ureido group, the C(10)–N(6) bond nearer the base being 0·07 Å longer than the C(10)–N(11) bond. A similar situation has been found in several other ureidopurines and has been pointed out elsewhere (Parthasarathy, Ohrt & Chheda, 1974c).

The adenine base, the ureido group and the carboxyl group are all planar to the limits of accuracy of our data; the planarity of the adenine portion extends through the ureido link but the C^α atom [C(12)] of glycine lies 0·44 Å out of this plane. The plane of the carboxyl group is almost perpendicular (89·9°) to the plane through the base. The twist of the carboxyl group (ψ) (IUPAC–IUB Commission on Biochemical Nomenclature, 1970) about the C^α–C′ [C(12)–C(13)] bond is ±156·4°.

The ureido substituent at N(6) is distal (*trans*) to the imidazole ring and restricts the conformation of the molecule about the two C(sp^2)–N(sp^2) bonds [N(16)–C(10) and C(10)–N(11)], a feature common to ureido purines as noted earlier (Parthasarathy, Ohrt & Chheda, 1974c).

The molecular stacking, hydrogen bonding and potassium coordination are illustrated in Fig. 3. The distances and angles involved in the hydrogen bonding are given in Table 4. The conformation of the substituent about N(6) orients the amino nitrogen in such a way that an intramolecular hydrogen bond is formed from N(11)–H(N11) of glycine to the N(1) of adenine, resulting in the formation of a planar six-membered ring. The formation of such intramolecular hydrogen bonding is common in organic compounds (Hamilton & Ibers, 1968; Tichy, 1965) and, in particular, has also been observed in other nucleic acid components (Parthasarathy, Ohrt & Chheda, 1973).

There are two polar hydrogens on the base, one on N(6) and the other on N(9). The hydrogens, H(N6) and H(N9), are hydrogen-bonded to O(13b) and O(13a) of the carboxyl groups of neighboring molecules (Table 4). N(3) does not partake in any hydrogen bonding.

The water molecule takes part in only two hydrogen bonds; it donates its two hydrogens to O(13a) of carboxyl groups of two different molecules.

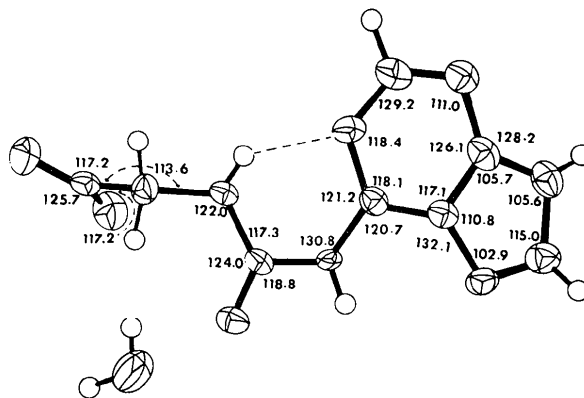The K$^+$ ion is seven-coordinated (Fig. 4) to two

Fig. 2. Bond angles in PCGK.

Table 4. *Hydrogen bonds*

D–H···A	D–H	D···A	H···A	D–H···A
O(W)–Ha(OW)···O(13a^{vii})	1·01 Å	2·751 Å	1·74 Å	174·2°
O(W)–Hb(OW)···O(13a)	0·76	2·901	2·15	175·3
N(6)—H(N6)···O(13b^v)	0·80	3·092	2·35	154·1
N(9)—H(N9)···O(13a^i)	0·93	2·751	1·83	169·6
N(11)–H(N11)···N(1)	0·85	2·738	2·03	140·8

Symmetry code (reference molecule x,y,z)

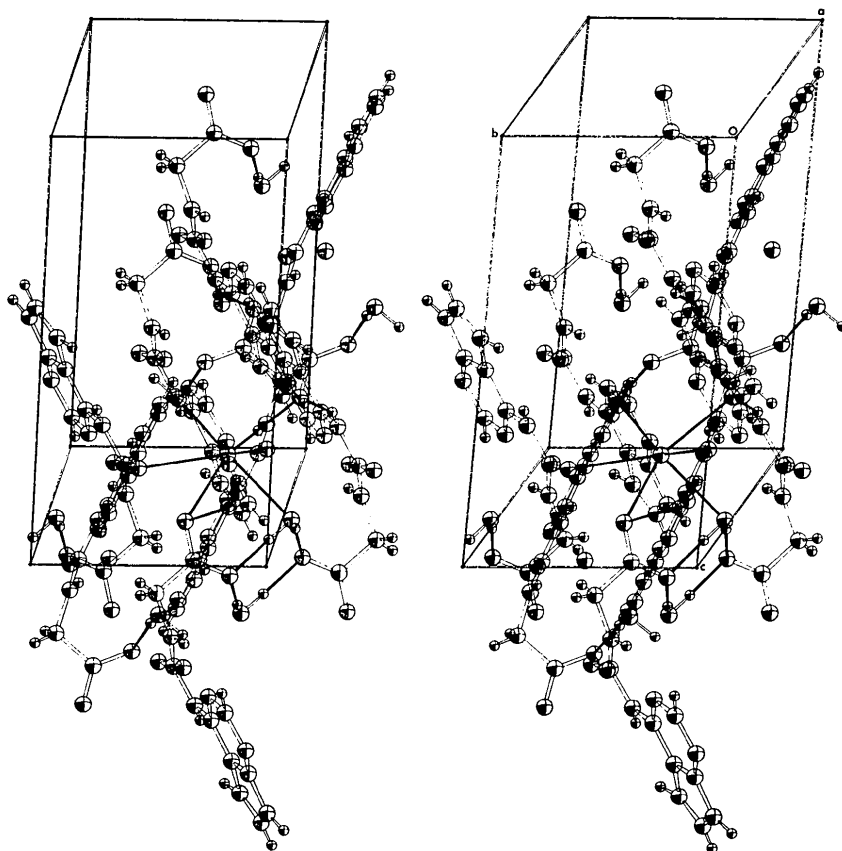| Superscript			Superscript			
i	1−x,	1−y, 2−z	v		x, ½−y,	½+z
ii	x,	−½+y, ³⁄₂−z	vi	−1+x,	½−y,	−½+z
iii	x,	½+y, ³⁄₂−z	vii		x̄, 1−y,	1−z
iv	x̄,	ȳ, 1−z	viii	1−x,	ȳ,	2−z

Fig. 3. A stereo view of the molecular stacking, hydrogen bonding and K⁺ coordination.

water oxygens, the keto oxygen O(10) on two ureido groups, O(13b) of two different carboxyl groups and N(3). The coordination distances range from 2·68 to 2·19 Å. The next closest contact of 3·50 Å is to C(2) at $\bar{x}, \bar{y}, 1-z$. The sevenfold coordination of potassium cannot be described in terms of simple geometry. Seven cation coordination has also been observed in several other structures (see, for example, Carrell, 1973; Einspar, Marsh & Donohue, 1972; Lanfredi, Pellinghelli & Tiripichio, 1972; Glusker, Orehowsky, Casciato & Carrell, 1972; Messmer & Palenik, 1971).
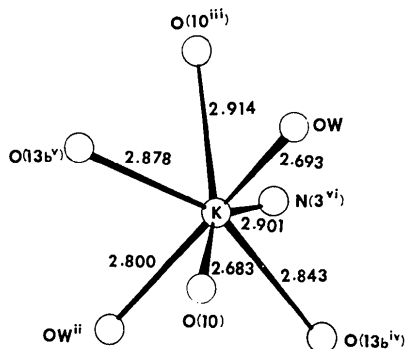
Fig. 4. The sevenfold coordination of the K⁺ ion.

The crystal structure exhibits stacking and extensive overlapping of the planar portions of the molecules in a head to tail fashion (Figs. 5 and 6). Molecules $A(1-x, \bar{y}, 2-z)$ and $C(1-x, 1-y, 2-z)$ are stacked at distances of 3·27 and 3·18 Å respectively from the reference molecule B. The six-membered rings of molecules B and C lie almost on top of each other and, to a small extent, on the five-membered rings. The glycine residue forms a hydrogen bond from N(9) through H(N9) of one molecule to the oxygen O(13a) of the carboxyl group of another, resulting in the formation of a close-stacked 'dimer'. The type of stacking observed between molecules B and C is very similar to the stacking observed in PCTK (Parthasarathy, Ohrt & Chheda, 1974c). Molecules A and B (Figs. 5 and 6) are stacked in such a way that the ureido portion of one molecule overlaps the six-membered ring of another.

The stacked rings associate more closely than has usually been observed in crystals or polymers (Bugg, Thomas, Sundaralingam & Rao, 1971). Short contacts less than the accepted van der Waals interatomic contacts (3·4 for C···C, 3·2 for C···N and 3·1 Å for C···O; Pauling, 1960) are as follows: C(2)···C(5ⁱ), 3·289; C(4)···C(6ⁱ), 3·246; and C(2)···C(10ᵛⁱⁱⁱ), 3·293 Å. The stacking of the bases dominates the non-polar
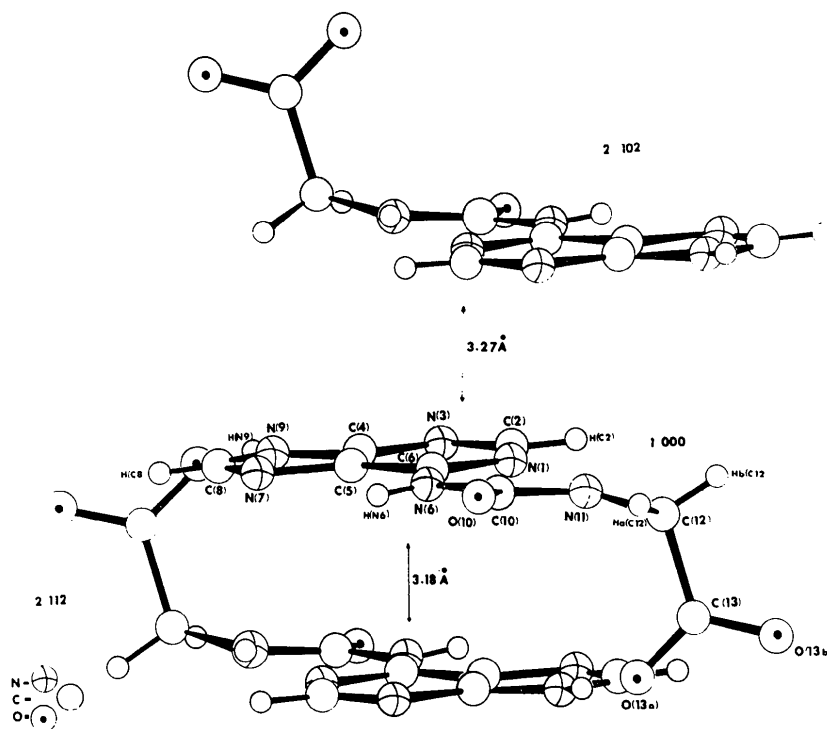
Fig. 5. Stacking of the bases seen edgewise.

interactions while the hydrogen bonding to water and K$^+$ ion coordination dominate the polar interactions. The stacking of the bases is stabilized by the hydrogen bond between N(9) of one molecule and the carboxyl oxygen belonging to another molecule stacked on top of the first molecule. This feature is also found in the structures of PCTK. The orientation of the carboxyl group with respect to the 'amino' nitrogen N(11) facilitates such hydrogen-bonding interaction between molecules stacked on top of each other. This orientation of the carboxyl group ($\psi \sim 180°$) is found (Marsh & Donohue, 1967; Lakshminarayanan, Sasisekharan & Ramachandran, 1967) to be remarkably similar to that observed in amino acids.
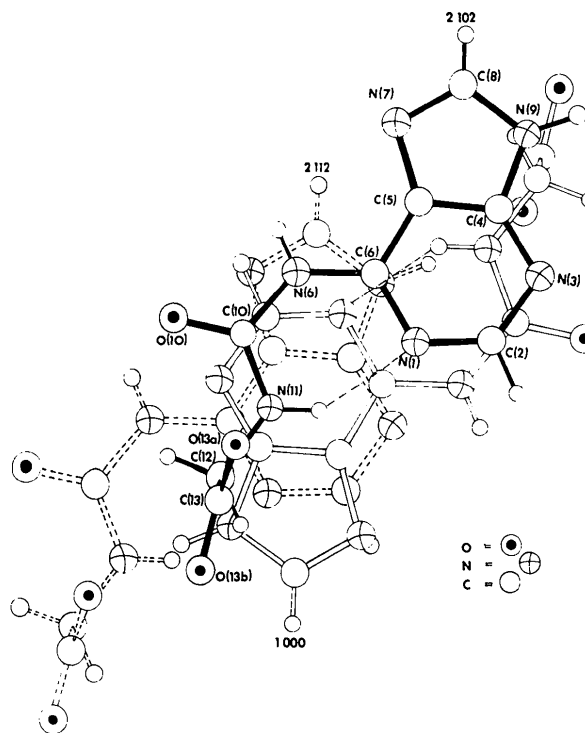
Fig. 6. Stacking of the bases viewed normal to the plane of the bases.

References

BUGG, C. E., THOMAS, J. M., SUNDARALINGAM, M. & RAO, S. T. (1971). *Biopolymers*, 10, 175–219.

CARRELL, H. H. (1973). *Acta Cryst.* B29, 2082–2088.

CHHEDA, G. B., HALL, R. H., MAGRATH, D. I., MOZEJKO, J., SCHWEIZER, M. P., STASIUK, L. & TAYLOR, P. R. (1969). *Biochemistry*, 8, 3278–3282.

CHHEDA, G. B. & HONG, C. I. (1971). *J. Med. Chem.* 14, 748–753.

EINSPAR, H., MARSH, R. E. & DONOHUE, J. (1972). *Acta Cryst.* B28, 2194–2198.

EVANS, H. T. JR (1961). *Acta Cryst.* 14, 689–690.

FURNAS, T. C. JR & HARKER, D. (1955). *Rev. Sci. Instrum.* 26, 449–453.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A27, 368–376.

GLUSKER, J. P., OREHOWSKY, W. JR, CASCIATO, C. A. & CARRELL, H. L. (1972). *Acta Cryst.* B28, 419–425.

HALL, R. H. (1971). *The Modified Nucleosides in Nucleic Acids.* New York: Columbia Univ. Press.

HAMILTON, W. C. & IBERS, J. A. (1968). *Hydrogen Bonding in Solids*, pp. 178–181. New York: Benjamin.

International Tables for X-ray Crystallography (1968). Vol. III, pp. 202–203. Birmingham: Kynoch Press.

ISHIKURA, H., YAMADA, Y., MURARO, K., SANEYOSHI, M. & NISHIMURA, S. (1969). *Biochem. Biophys. Res. Commun.* 37, 990–995.

IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *J. Mol. Biol.* 52, 1–17.

KARLE, J. & KARLE, I. (1966). *Acta Cryst.* 21, 849–859.

LAKSHMINARAYANAN, A. V., SASISEKHARAN, V. & RAMACHANDRAN, G. N. (1967). *Conformation of Biopolymers*, Vol. I, pp. 61–82. New York: Academic Press.

LANFREDI, A. M. M., PELLINGHELLI, M. A. & TIRIPICCHIO, A. (1972). *Acta Cryst.* B28, 1822–1826.

MARSH, R. E. & DONOHUE, J. (1967). *Advanc. Protein Chem.* 22, 234–265.

MESSMER, G. G. & PALENIK, G. J. (1971). *Acta Cryst.* B27, 314–321.

NISHIMURA, S., YAMADA, Y. & ISHIKURA, H. (1969). *Biochem. Biophys. Acta*, 179, 515–517.

PAULING, L. (1960). *The Nature of the Chemical Bond.* Ithaca: Cornell Univ. Press.

PARTHASARATHY, R., OHRT, J. M. & CHHEDA, G. B. (1973). *J. Amer. Chem. Soc.* 95, 8141–8146.

PARTHASARATHY, R., OHRT, J. M. & CHHEDA, G. B. (1974a). *Biochem. Biophys. Res. Commun.* 57, 649–653.

PARTHASARATHY, R., OHRT, J. M. & CHHEDA, G. B. (1974b). *Biochem. Biophys. Res. Commun.* 60, 211–218.

PARTHASARATHY, R., OHRT, J. M. & CHHEDA, G. B. (1974c). *J. Amer. Chem. Soc.* 96, 8087–8094.

SCHWEIZER, M. P., CHHEDA, G. B., BACZYNSKYJ, L. & HALL, R. H. (1969). *Biochemistry*, 8, 3283–3289.

SCHWEIZER, M. P., McGRATH, K. & BACZYNSKYJ, L. (1970). *Biochem. Biophys. Res. Commun.* 40, 1046–1052.

STEWART, R. F., DAVIDSON, E. R. & SIMPSON, W. T. (1965). *J. Chem. Phys.* 42, 3175–3187.

TAKEMURA, S., MURAKAMI, M. & MIYAZAKI, M. (1969). *J. Biochem. Tokyo*, 65, 553–566.

TICHY, M. (1965). *Advanc. Org. Chem.* 5, 115–285.

The Crystal and Molecular Structure of the 3,5-Dinitrobenzoate Derivative of Toxisterol2-*A*

BY CHARLES H. CARLISLE AND PETER F. LINDLEY

*Department of Crystallography, Birkbeck College, University of London, Malet Street, London WC1E 7HX, England*

Crystals of the 3,5-dinitrobenzoate derivative of toxisterol2-*A* are monoclinic, space group $P2_1$, with $a = 6.358$ (1), $b = 21.808$ (3), $c = 12.259$ (2) Å, $\beta = 102.81$ (1)°. The structure has been refined to a residual of 0·056 with 1856 independent significant reflexions measured on an automated four-circle diffractometer. In toxisterol2-*A*, which is derived from ergosterol, ring *B*, a five-membered ring, is linked to ring *C via* a spiro carbon atom at the C8 position.

Introduction

Toxisterols are substances of reputed biological activity which were originally described some fifty years ago (Westerhof & Keverling Buisman, 1956, and references therein). However, only one, toxisterol2-*A* (I), was properly characterized; the subscript 2 indicates that the compound is derived from ergosterol (II) rather than cholesterol (subscript 3).

(I)   (II)
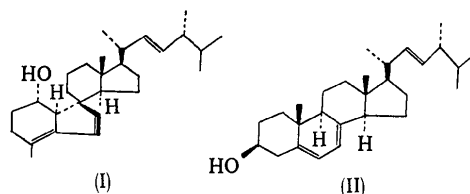
Recently the isolation, characterization and structure determination of three such compounds, toxi-