electrostatic potential in the expected N lone-pair position as χ_N is increased incrementally from 0 to 48.5°. These calculations are consistent with the crystallographic evidence that all N···H contacts within the sum of van der Waals radii (2.75 Å) that have $\chi_N \geq 35°$ and $\rho_H \geq 130°$ may be classified as hydrogen bonds. These N···H contacts are closely aligned with the assumed N lone-pair vector and such evidence is normally taken to infer an interaction between H(δ+) and the lone-pair density. However, by analogy with the perpendicular approach of truly 'non-bonded' H to planar Nsp^2 systems (XIV), the lone-pair direction is also the approach direction of minimal steric hindrance to sp^3 hybridized N atoms (XV). Despite this analogy, we feel that the crystallographic results presented here provide tentative evidence that the incoming donor H atom tracks the changing direction of the N lone-pair vector as χ_N increases from 35° to ~60° during the latter phases of the N$sp^2 \rightarrow$ Nsp^3 transition.

We are indebted to Professor G. Gilli and his colleagues at the Universita di Ferrara, Italy, for preprints of their 1993 paper (Ferretti *et al.*, 1993) and for helpful discussions. We also thank referees for constructive suggestions on initial versions of this paper, Dr Neil Stewart (CCDC) for assistance with the illustrations and Dr Olga Kennard OBE FRS for her encouragement of this work.

References

ALLEN, F. H., DAVIES, J. E., GALLOY, J. J., JOHNSON, O., KENNARD, O., MACRAE, C. F., MITCHELL, E., MITCHELL, G. F., SMITH, J. M. & WATSON, D. G. (1991). *J. Chem. Inf. Comput. Sci.* **31**, 187.

ALLEN, F. H., KENNARD, O., WATSON, D. G., BRAMMER, L., ORPEN, A. G. & TAYLOR, R. (1987). *J. Chem. Soc. Perkin Trans.* 2, pp. S1–S19.

BARBIERI, G., BENASSI, R., GRANDI, R., PAGNONI, U. M. & TADDEI, F. (1979). *J. Chem. Soc. Perkin Trans.* 2, pp. 330–336.

BELLUCCI, F., BERTOLASI, V., FERRETTI, V. & GILLI, G. (1985). *Acta Cryst.* C**41**, 544–546.

BERTOLASI, V., BELLUCCI, F., FERRETTI, V. & GILLI, G. (1984). *Acta Cryst.* A**40**, C107.

BOCK, C. W., GEORGE, P. & TRACHTMAN, M. (1986). *Theor. Chim. Acta*, **69**, 235–245.

BONDI, A. (1964). *J. Phys. Chem.* **68**, 441–451.

BROWN, K. L., DAMM, L., DUNITZ, J. D., ESCHENMOSER, A., HOBI, R. & KRATKY, C. (1978). *Helv. Chim. Acta*, **61**, 3108–3135.

BÜRGI, H.-B. & SHEFTER, E. (1975). *Tetrahedron*, **31**, 2976–2981.

BÜRGI, H.-B., DUNITZ, J. D. & SHEFTER, E. (1973). *J. Am. Chem. Soc.* **95**, 5065.

Cambridge Structural Database User's Manual (1992). Version 5.1. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

Cambridge Structural Database User's Manual (1994). *Getting Started with the CSD*. Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge, England.

DUNITZ, J. D. (1979). *X-ray Analysis and the Structure of Organic Molecules*, p. 333. Ithaca, NY: Cornell Univ. Press.

DUNITZ, J. D. & WINKLER, F. K. (1975). *Acta Cryst.* B**31**, 251–263.

FERRETTI, V., BERTOLASI, V., GILLI, P. & GILLI, G. (1993). *J. Phys. Chem.* **49**, 13568–13574.

GILLI, G. & BERTOLASI, V. (1979). *J. Am. Chem. Soc.* **101**, 7704–7711.

GILLI, G., BERTOLASI, V., BELLUCCI, F. & FERRETTI, V. (1986). *J. Am. Chem. Soc.* **108**, 2420–2424.

GUEST, M. F., VAN LENTHE, J. H., KENDRICK, J., SCHOEFFEL, K., SHERWOOD, P. & HARRISON, R. J. (1993). *GAMESS-UK User's Guide and Reference Manual*. Computing for Science Ltd., SERC Daresbury Laboratory, Daresbury, Warrington, England.

KROON, J., KANTERS, J. A., VAN DUIJNEVELDT-VAN DE RIJDT, J. G. C. H., VAN DUIJNEVELDT, F. B. & VLIEGENTHART, J. A. (1975). *J. Mol. Struct.* **24**, 109–129.

LISTER, D. G., TYLER, J. K., HOEG, J. H. & LARSEN, W. (1974). *J. Mol. Struct.* **23**, 253–264.

MACKENZIE, R. & MACNICOL, D. D. (1970). *Chem. Commun.* pp. 1299–1300.

MARTIN, G. L., GOUESNARD, J. P., DORIE, J., RABILLER, C. & MARTIN, M. L. (1977). *J. Am. Chem. Soc.* **99**, 1381–1384.

SIDDALL, T. H., STEWART, W. E. & KNIGHT, F. D. (1970). *J. Phys. Chem.* **74**, 3580–3583.

TIERNEY, L. (1992). *The XLISP-STAT Package*. School of Statistics, Univ. of Minnesota, Minneapolis, USA.

WALLIS, J. D., EASTON, R. J. C. & DUNITZ, J. D. (1993). *Helv. Chim. Acta*, **76**, 1411–1424.

WINKLER, F. K. & DUNITZ, J. D. (1971). *J. Mol. Biol.* **59**, 169–182.

Acta Cryst. (1995). B**51**, 1081–1097

Electrostatic Properties of β-Cytidine and Cytosine Monohydrate from Bragg Diffraction

BY LIRONG CHEN AND B. M. CRAVEN*

Department of Crystallography, University of Pittsburgh, Pittsburgh, PA 15260, *USA*

(*Received* 4 *January* 1995; *accepted* 27 *April* 1995)

Abstract

The charge-density distribution in the crystal structure of the nucleoside β-cytidine at 123 K has been determined from X-ray diffraction data (Ag $K\alpha$, $\lambda = 0.5608$ Å) using all 7233 reflections with $\sin\theta/\lambda \leq 1.14$ Å^{-1}. Maps of electrostatic potential for the cytosine base in cytidine are similar to those derived from previous charge-density studies of cytosine monohydrate and 1-β-D-arabinosylcytosine, after taking chemical differences into account.

* Author for correspondence.

These results were obtained from pseudoatom multipole refinements including κ as a variable to describe the expansion or contraction of the spherical valence shell for each atom type. A new structure refinement of this type has also been carried out for cytosine monohydrate. A survey of κ values for hydrogen indicates that this is a variable which is not well determined experimentally. Variations in κ_H are relayed into the population parameters obtained for other pseudoatoms and can have a small but significant effect on molecular properties, such as the dipole moment. Assuming an average theoretical value of $\kappa_H = 1.24$, the molecular dipole moments calculated from the monopole and dipole pseudoatom population parameters are 17(4) and 8.2(15) Debye for cytidine nucleoside and cytosine in the monohydrate, respectively. Systems of atomic point charges are presented for cytidine, cytosine and water. These are generally satisfactory in reproducing the experimentally determined molecular electrostatic potentials and dipole moments.

Introduction

The crystal structure of β-cytidine (Fig. 1; hereafter, cytidine) was the first to be determined for a nucleoside component of RNA (Furberg, 1950). This was done before the structure of DNA or RNA was determined. Three-dimensional photographic data were used later for refining the structure of cytidine by least-squares methods (Furberg, Petersen & Romming, 1965). The results agree within experimental error with a recent room-temperature redetermination based on diffractometer data (Ward, 1993). Other related crystal structures which have been determined at atomic resolution include α-cytidine (Post, Birnbaum, Huber & Shugar, 1977), in which the ribose has the α-anomeric configuration, 1-β-D-arabinosylcytosine (Tougard & Lefebvre-Soubeyran, 1974), cytosine (Barker & Marsh, 1964; McClure & Craven, 1973) and cytosine monohydrate by X-ray diffraction (Jeffrey & Kinoshita, 1963; McClure & Craven, 1973) and neutron diffraction at 82 K (Weber, Craven & McMullan, 1980). Many crystal complexes involving G–C base pairs have been studied, the most notable being sodium guanylyl-3′,5′-cytidine nonahydrate (Rosenberg, Seeman, Day & Rich, 1976), because this structure first showed minihelical RNA at atomic resolution. Experimental charge-density studies have been carried out for cytosine monohydrate at 97 K (Eisenstein, 1988) and at 82 K (Weber & Craven, 1990), and also for zwitterionic 2′-deoxycytidine-5′-monophosphate monohydrate (Pearlman & Kim, 1985, 1990) and 1-β-D-arabinosylcytosine* (Pearlman & Kim, 1990).

* Pearlman & Kim (1990) describe a charge-density study of two nucleotides and four nucleosides, including cytidine. However, their cytidine crystals are misnamed, being actually the arabinosyl nucleoside with the crystal structure previously determined by Tougard & Lefebvre-Soubeyran (1974).

The charge-density distribution in cytidine has presently been determined in order to derive a set of point charges which reliably reproduce the experimental molecular electrostatic potential. We are studying a series of nucleic acid components in this way. We wish to examine how well such charges are transferable from one crystal environment to another and whether they might be affected by factors such as differences in hydrogen bonding and molecular conformation. In our studies we assume an aspherical pseudoatom model (Stewart, 1976) involving 16 electron-population parameters per pseudoatom. Pearlman & Kim (1990) have already listed point charges derived from their X-ray studies of nucleoside and nucleotide crystal structures. However, their work differs from ours because in their structure refinements they assume a simpler spherical pseudoatom model (Coppens *et al.*, 1979), with only one electron-population parameter and one valence-shell radial parameter per pseudoatom.

Presently, we report an analysis of new X-ray diffraction data for β-cytidine collected at high resolution ($\sin\theta/\lambda \leq 1.14$ Å^{-1}) and reduced temperature (123 K) and also new refinements for cytosine monohydrate using the X-ray data of Weber & Craven (1990). In comparing net atomic charges and molecular dipole moments with those reported for 1-β-D-arabinosylcytosine (Pearlman & Kim, 1990), agreement is significantly dependent on the choice of exponent in the radial density function assumed for the H atoms. We recommend that a theoretical value be assumed for this parameter because it is not well determined experimentally.

Experimental

Crystals of β-cytidine (Sigma Chemical Company, Inc.) were grown from 90% alcohol solution by slow evaporation. No phase transition occurred during cooling from 298 to 110 K. The crystal chosen for data collection measured 0.375 × 0.462 × 0.550 mm, was elongated along the c-axis and showed the forms {110} and

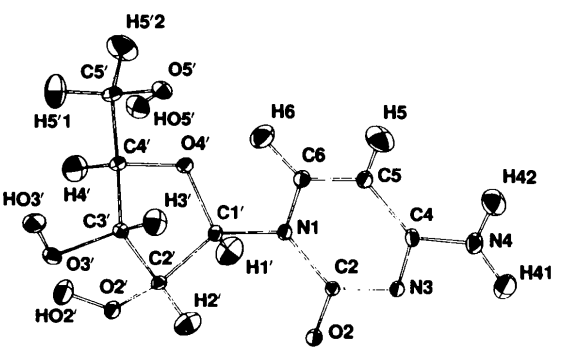


Fig. 1. Molecular structure of β-cytidine showing the atomic nomenclature. Atomic displacement ellipsoids (Johnson, 1976) are at the 50% level of probability for enclosing the atom.

{001}. The crystal was mounted with the *c*-axis close to the ϕ-axis of an Enraf–Nonius CAD-4 diffractometer and was kept at a temperature of 123 (2) K in a stream of nitrogen gas. In order to minimize ice formation on the crystal, the diffractometer was sealed in a box with dried air constantly blown into the box. The temperature was monitored using a thermocouple in the cold stream about 8 mm upstream from the crystal.

The crystal data for cytidine ($C_9H_{13}N_3O_5$) are $M_r = 243.23$, orthorhombic, $P2_12_12_1$, $Z = 4$, $F(000) = 512$, $D_x = 1.542\,\text{mg mm}^{-3}$, $\mu = 0.0728\,\text{mm}^{-1}$. The X-radiation was Pd-filtered Ag $K\alpha$ ($\lambda = 0.5608$ Å). The unit-cell dimensions at 298 and 123 K were obtained from a least-squares fit of $\sin^2\theta$ values for 25 reflections in the range $13 < \theta < 20°$. Each reflection was measured at four equivalent positions (two with positive and two with negative θ). The cell parameters thus found are in good agreement with the room-temperature values previously reported (see Table 1). X-ray intensities were measured by $\omega/2\theta$ scans with ω-scan width $(0.50 + 0.45\tan\theta)°$ and with scan speeds varying from 0.62 to 2.0° min^{-1}. The intensities of three standard reflections (13,00, 444, 1,12,0) were monitored after every 6000 s of data collection. There was a decrease of $\sim 4\%$ in intensity during the collection of the last 2000 reflections, for which compensating scaling was applied. Reflections in the range $0 \le h \le 31$, $0 \le k \le 33$, $0 \le l \le 11$ were measured. Using the computer program of Blessing (1984), integrated intensities were derived from the intensity profile recorded at 96 points for each reflection. An analytical absorption correction (Templeton & Templeton, 1973) gave intensity correction factors in the range 1.029–1.038. A total of 7233 independent reflections were obtained, of which 5564 showed $F^2 > 2.5\sigma(F^2)$.

The atomic positional coordinates from Furberg, Petersen & Romming (1965) were assumed as the starting structure model. Using the computer program *POP* (Craven, Weber, He & Klooster, 1993), a conventional least-squares refinement was carried out with minimization of the residual $\sum w(F_c^2 - F_o^2)^2$ and with $w = 1/\sigma(F^2)$. Only reflections with $F^2 > 2.5\sigma(F^2)$ were included. There were 205 variables consisting of the scale factor, positional parameters, anisotropic mean square (m.s.) displacement parameters for non-H atoms and isotropic parameters for H atoms. This refinement gave $R(F^2) = 0.062$, $R_w(F^2) = 0.109$ and $S = 1.341$*.

When carrying out a charge-density study it is desirable to have accurate positional and m.s. displacement parameters for H atoms from neutron diffraction. Unfortunately, we were unable to grow a cytidine crystal large enough for neutron data collection. Therefore, H-atom parameters were derived as follows. First, the CH, CH_2, NH_2 and OH bonds were extended to lengths

Table 1. *Cell parameters* (Å)

	Present study		Ward (1993)	Furberg, Petersen & Romming (1965)
	Ag $K\alpha$, 123 K	298K	Mo $K\alpha$, 298 K	Cu $K\alpha$, 298 K
a	13.925 (1)	13.987 (2)	13.980 (4)	13.991 (2)
b	14.715 (1)	14.775 (2)	14.788 (3)	14.786 (2)
c	5.076 (1)	5.115 (2)	5.119 (3)	5.116 (1)

1.08, 1.10, 1.00 and 0.98 Å, respectively. These bond lengths are in accordance with values obtained for adenosine at 123 K by neutron diffraction (Klooster, Ruble, Craven & McMullan, 1991). Second, anisotropic m.s. displacement parameters for the H atoms were assumed to be the sum of m.s. displacements from the external or lattice vibrations of the molecule and from the internal or intramolecular vibrations of the CH, CH_2 NH_2 and OH groups. The contribution of the external vibrations was obtained by assuming that each H atom is carried rigidly on the heavier atom framework of the molecule. For cytidine, the molecule was assumed to be vibrating as a segmented rigid body (He & Craven, 1993). After trial and error, the best least-squares fit to the U^{ij} values for the atoms of the molecular framework was obtained assuming three rigid segments, namely the cytosine base, the ribose sugar and the C5′—O5′ hydroxyl group (Fig. 2). The variables fitted by least squares were the **T**, **L** and **S** tensor components for overall rigid-body motions, the internal libration (τ_1) about the glycosidic bond and the two bond-angle bending modes θ_1 and θ_2, as defined in Table 2. The internal m.s. displacements for the H atoms were assumed to the same as those derived from neutron

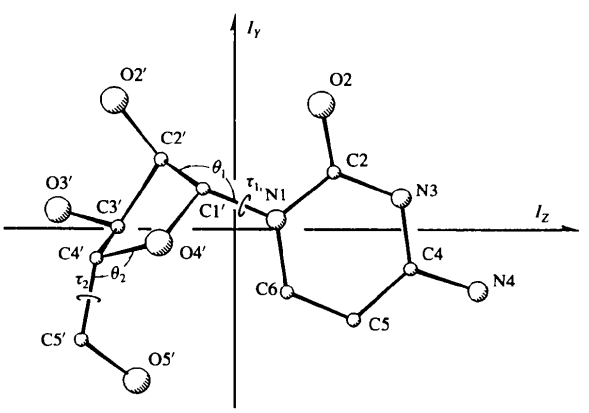


Fig. 2. Thermal vibrational model for β-cytidine. The molecular frame is assumed to consist of three rigid segments: (1) the cytosine base; (2) the ribose with substituents at C2′ and C3′; (3) the C4′ substituent group. Also indicated are the four internal vibrational modes which were considered (torsional vibrations τ_1 and τ_2, and bond-angle bending vibrations θ_1 and θ_2). The molecule is shown with respect to the inertial axes in the projection down I_x (normal to the mass-weighted best plane through the atoms of the frame) with the origin at the center of mass. Note that the internal vibration τ_2 is not included in the final result due to the significant negative of the principal value.

* $R(F^2) = \sum|\Delta|/\sum|F^2|$, $R_w(F^2) = [\sum w\Delta^2/\sum(F^2)^2]^{1/2}$, $S = [\sum w\Delta^2/(m_{obs} - n_{para})]^{1/2}$; $\Delta = F_o^2 - F_c^2$.

Table 2. *Thermal vibrations for cytidine*

The molecular-framework atomic m.s. displacements (H atoms omitted) were fitted with a segmented body model (He & Craven, 1993). The results are with respect to the Cartesian axes corresponding to the principal moments of inertia of the molecule and with the origin at the molecular center of mass (see Fig. 2). The function minimized was $\sum w(U_o^{ij} - U_c^{ij})^2$ with the sum taken over the six thermal components for all C, N and O atoms with $w = 1/\sigma^2(U_{eq})$.

Simple rigid body	$R_w = 0.172$
Torsional vibrations τ_1 and τ_2	$R_w = 0.156$
Vibrations involve τ_1 and bending θ_1	$R_w = 0.142$
Vibrations involve τ_1, τ_2 and θ_1	$R_w = 0.140$
Vibrations involve τ_1, θ_1 and θ_2	$R_w = 0.100$
Goodness-of-fit	3.831

M.s. torsional libration τ_1 (deg^2)	19.6 (2.5)
M.s. bending vibration θ_1 (deg^2)	8.8 (1.5)
M.s. bending vibration θ_2 (deg^2)	9.5 (2.0)

Translational tensor T (Å^2)	0.0085 (4)	−0.0005 (3)	0.0008 (2)
		0.0075 (3)	0.0000 (2)
			0.0075 (3)
Principal values (Å^2)	0.0092	0.0070	0.0076
Librational tensor L (deg^2)	0.6 (2)	−0.1 (1)	−0.6 (2)
		1.2 (2)	−1.3 (1)
			0.8 (5)
Principal values (deg^2)	0.8	1.5	0.7
Cross tensor S (Å × deg)	−0.007 (9)	−0.001 (3)	−0.016 (4)
	−0.017 (3)	−0.016 (9)	−0.018 (6)
	0.007 (5)	0.024 (8)	0.018 (1)

Table 3. *Derived* H-*atom parameters*

(*a*) Positional parameters

The first line gives the adjusted positions from neutron diffraction results and the second gives the positions from the conventional refinement (e.s.d.s in parentheses).

	x	*y*	*z*
H5	0.754 (2)	0.100 (2)	0.198 (5)
	0.747 (2)	0.104 (2)	0.199 (5)
H6	0.620 (1)	0.068 (1)	0.518 (4)
	0.619 (1)	0.075 (1)	0.501 (4)
H41	0.734 (1)	0.283 (1)	−0.280 (4)
	0.735 (1)	0.277 (1)	−0.264 (4)
H42	0.810 (2)	0.206 (2)	−0.134 (5)
	0.799 (2)	0.210 (2)	−0.136 (5)
HO2′	0.258 (2)	0.076 (2)	0.626 (6)
	0.263 (2)	0.086 (2)	0.611 (6)
HO3′	0.352 (2)	−0.113 (2)	0.460 (6)
	0.349 (2)	−0.107 (2)	0.459 (6)
HO5′	0.564 (2)	−0.121 (2)	0.526 (4)
	0.568 (1)	−0.113 (2)	0.544 (4)
H5′1	0.470 (1)	−0.121 (2)	0.865 (6)
	0.473 (2)	−0.117 (2)	0.863 (6)
H5′2	0.549 (2)	−0.042 (2)	1.021 (6)
	0.547 (2)	−0.043 (2)	1.015 (6)
H1′	0.431 (1)	0.225 (1)	0.633 (4)
	0.433 (1)	0.219 (1)	0.627 (4)
H2′	0.374 (1)	0.125 (1)	0.219 (5)
	0.374 (1)	0.123 (1)	0.243 (5)
H3′	0.459 (1)	−0.008 (1)	0.346 (4)
	0.456 (1)	−0.007 (1)	0.353 (4)
H4′	0.382 (1)	0.019 (1)	0.898 (4)
	0.389 (1)	0.019 (1)	0.881 (4)

(*b*) Derived m.s. displacement parameters (Å^2) for the H atoms

U^{ij} values derived from the sum of rigid-body and intramolecular contributions. U_{iso} values were obtained directly from the least-squares structure refinement.

$$U_{eq} = (1/3)\sum_i \sum_j U_{ij} a_i^* a_j^* \mathbf{a}_i.\mathbf{a}_j.$$

	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}	U_{eq}	U_{iso}
H5	0.017	0.026	0.041	0.006	−0.004	−0.002	0.028	0.023 (5)
H6	0.030	0.017	0.025	−0.003	−0.005	0.005	0.024	0.018 (5)
H41	0.025	0.024	0.029	−0.001	0.003	0.008	0.026	0.011 (4)
H42	0.016	0.030	0.040	0.003	0.005	0.004	0.028	0.020 (6)
HO2′	0.022	0.016	0.027	−0.001	0.006	0.006	0.022	0.010 (6)
HO3′	0.019	0.013	0.025	0.003	−0.003	0.002	0.019	0.024 (6)
HO5′	0.020	0.016	0.021	−0.003	−0.004	−0.006	0.019	0.017 (5)
H5′1	0.033	0.018	0.037	−0.002	0.008	−0.004	0.029	0.041 (7)
H5′2	0.029	0.044	0.020	−0.005	−0.008	−0.007	0.031	0.043 (7)
H1′	0.034	0.015	0.025	0.005	0.005	−0.003	0.025	0.012 (4)
H2′	0.028	0.024	0.016	0.004	0.002	0.003	0.027	0.034 (5)
H3′	0.021	0.021	0.025	0.002	0.008	−0.002	0.022	0.031 (5)
H4′	0.024	0.031	0.021	0.002	0.007	0.002	0.025	0.030 (6)

diffraction for the corresponding H atoms in adenosine (see Table 6*b* in Klooster, Ruble, Craven & McMullan, 1991). The resulting total anisotropic m.s. displacements for the H atoms are in Table 3 and are represented by 50% probability ellipsoids in Fig. 1.

The charge-density distribution in the crystal structure was determined from least-squares refinements based on the rigid pseudoatom model of Stewart (1976). Hartree–Fock atomic wavefunctions with Slater-type orbitals and with coefficients as tabulated by Clementi & Roetti (1974) were used to obtain atomic scattering factors. For C, N and O, we assumed invariant Hartree–Fock *K*-shell cores populated by two electrons. For the Hartree–Fock valence shell, we introduce κ as a variable radial contraction or expansion parameter for each atom type (Coppens *et al.*, 1979). Deformations of the valence shell were represented by a complete multipole expansion (up to octapole) about each atom center, each multipole term having a single Slater-type radial function with a standard value for the exponent ($\alpha = 6.50$, 7.37 and 8.50 Å^{-1} for C, N and O, respectively; Hehre, Stewart & Pople, 1969). For the H atoms, the multipole expansion involved monopole through quadrupole terms with a fixed standard radial exponent $\alpha = 4.69$ Å^{-1} (Hehre, Stewart & Pople, 1969). The pseudoatom refinement was carried out with the *POP* computer program (Craven, Weber, He & Klooster, 1993) and involved all 7233 reflections. There were 548 variables which included 153 from the initial conventional refinement. Fixed values from Table 3 were assumed for the H-atom positional and anisotropic m.s. displacement parameters. The new variables consisted of six κ parameters ($\kappa_{C'}$: trigonal bonded; $\kappa_{C''}$: tetrahedral bonded; $\kappa_{N'}$: trigonal; $\kappa_{N''}$: with lone pair; $\kappa_{O'}$: carbonyl; $\kappa_{O''}$: hydroxyl) and 389 electron population parameters. No extinction parameters were introduced. The function minimized by least-squares was $\sum w\Delta^2$, where $w = 1/\sigma^2(F^2)$ and $\Delta = F_c^2 - F_o^2$. This refinement gave final values $R(F^2) = 0.054$, $R_w(F^2) = 0.083$ and $S = 1.630$. The final atomic parameters are listed in Table 4. Interatomic distances and angles from this full multipole refinement are given in Table 5. In the refinement, the number of valence electrons in the molecule was unconstrained with an overall scale factor applied only to the *K*-shell scattering. The neutral cytidine molecule contains 34 *K*-shell and 94

Table 4. *Atomic parameters for cytidine from the multipole refinement*

(*a*) Positional and thermal parameters ($\times 10^4$ for x, y, z; $Å^2 \times 10^2$ for U^{ij}). The anisotropic temperature factors have the form: $T = \exp[-2\pi^2 \sum_i \sum_j h_i h_j a^{i*} a^{j*} U^{ij}]$.

	x	y	z	U^{11}	U^{22}	U^{33}	U^{12}	U^{13}	U^{23}
N1	5359 (1)	1769 (1)	3810 (2)	84 (2)	79 (2)	117 (2)	2 (2)	11 (2)	16 (2)
C2	5267 (1)	2488 (1)	2053 (2)	78 (2)	75 (2)	114 (3)	2 (2)	8 (2)	13 (2)
N3	6000 (1)	2690 (1)	399 (2)	86 (2)	87 (2)	137 (3)	−1 (2)	19 (2)	21 (2)
C4	6778 (1)	2151 (1)	337 (2)	86 (2)	95 (2)	145 (3)	−6 (2)	20 (2)	14 (2)
C5	6892 (1)	1401 (1)	2093 (2)	92 (2)	129 (3)	180 (3)	22 (2)	28 (3)	51 (3)
C6	6166 (1)	1241 (1)	3811 (2)	98 (2)	107 (3)	158 (3)	16 (2)	18 (3)	35 (3)
N4	7450 (1)	2346 (1)	−1449 (2)	110 (3)	154 (3)	189 (4)	10 (2)	54 (3)	44 (3)
O2	4507 (1)	2941 (1)	2081 (2)	94 (2)	97 (2)	152 (3)	20 (2)	15 (3)	25 (3)
C1′	4545 (1)	1592 (1)	5637 (2)	93 (2)	79 (2)	102 (2)	−3 (2)	6 (2)	2 (2)
C2′	3724 (1)	1077 (1)	4254 (2)	88 (2)	79 (2)	91 (2)	9 (2)	2 (2)	10 (2)
C3′	4019 (1)	96 (1)	4791 (2)	89 (2)	76 (2)	95 (2)	6 (2)	−5 (2)	−7 (2)
C4′	4407 (1)	152 (1)	7607 (2)	106 (2)	81 (2)	83 (2)	6 (2)	−3 (2)	6 (2)
O4′	4862 (1)	1036 (1)	7729 (2)	125 (2)	90 (2)	86 (2)	−11 (2)	−17 (2)	1 (2)
O2′	2837 (1)	1312 (1)	5446 (2)	96 (2)	107 (2)	160 (3)	18 (2)	36 (3)	31 (2)
O3′	3251 (1)	−519 (1)	4480 (2)	109 (3)	90 (2)	174 (3)	−4 (2)	−33 (3)	−8 (3)
C5′	5120 (1)	−585 (1)	8356 (2)	142 (3)	104 (3)	115 (3)	26 (2)	−20 (3)	14 (2)
O5′	5846 (1)	−720 (1)	6438 (2)	106 (3)	125 (3)	164 (3)	20 (2)	−10 (3)	−21 (3)

(*b*) Electron population parameters for cytidine

All values except P_v are $\times 10$. These correspond to the normalized functions referred to crystal axes, as described in the Appendix of Epstein, Ruble & Craven (1982). p_v values give the valence-shell monopole population and have been normalized to $\sum p_v = 94.0$. Valence shell radial density functions have $\kappa_{C'} = 1.040(7)$, $\kappa_{C''} = 1.039(6)$, $\kappa_{N'} = 1.005(6)$, $\kappa_{N''} = 0.984(7)$, $\kappa_{O'} = 0.978(5)$, $\kappa_{O''} = 0.970(3)$, $\kappa_H = 1.24$ (see text for details).

	p_v	*d*1	*d*2	*d*3	*q*1	*q*2	*q*3	*q*4	*q*5	*o*1	*o*2	*o*3	*o*4	*o*5	*o*6	*o*7
N1	5.23 (6)	0 (2)	−8 (2)	−3 (3)	2 (2)	−1 (2)	2 (3)	3 (3)	−5 (2)	−4 (2)	−15 (2)	17 (3)	0 (5)	−10 (2)	2 (2)	1 (3)
C2	3.66 (6)	−3 (3)	2 (3)	−3 (3)	13 (2)	−15 (3)	−14 (3)	−18 (3)	−6 (3)	2 (3)	21 (3)	−19 (3)	−11 (5)	17 (3)	−15 (3)	−1 (3)
N3	5.40 (6)	3 (2)	−3 (2)	−2 (3)	3 (2)	−10 (2)	−5 (3)	−9 (3)	−1 (3)	−2 (2)	−12 (2)	3 (3)	1 (5)	−11 (3)	−4 (3)	−8 (3)
C4	3.77 (6)	−1 (3)	−2 (3)	5 (3)	2 (2)	−5 (2)	−14 (3)	−11 (3)	−8 (3)	2 (3)	19 (3)	−23 (3)	−16 (5)	19 (3)	−12 (3)	−6 (4)
C5	4.06 (6)	−9 (3)	3 (3)	5 (3)	11 (3)	−2 (3)	−3 (3)	−14 (3)	0 (3)	−8 (3)	−15 (3)	19 (3)	−4 (5)	−16 (3)	13 (3)	−1 (4)
C6	3.89 (6)	3 (3)	−3 (3)	1 (3)	3 (3)	−7 (3)	−14 (3)	−9 (3)	−3 (3)	4 (3)	17 (3)	−7 (3)	−11 (5)	14 (3)	−12 (3)	−4 (3)
N4	5.07 (6)	−3 (3)	−1 (3)	7 (3)	0 (3)	−6 (3)	0 (3)	3 (3)	−2 (3)	−3 (3)	−8 (2)	13 (3)	2 (5)	−11 (3)	7 (3)	2 (3)
O2	6.48 (4)	−2 (2)	0 (2)	−1 (3)	2 (2)	0 (2)	−6 (3)	−5 (3)	2 (3)	−2 (2)	−7 (2)	0 (2)	−4 (5)	−1 (2)	1 (2)	−3 (3)
C1′	3.74 (6)	−19 (3)	−8 (3)	−12 (3)	7 (3)	−1 (3)	4 (3)	8 (3)	−12 (3)	5 (3)	−10 (3)	−16 (3)	−33 (5)	11 (3)	−14 (3)	6 (3)
C2′	3.76 (6)	10 (3)	−8 (3)	−9 (3)	−10 (2)	4 (3)	6 (3)	−1 (3)	−3 (3)	−16 (3)	14 (3)	12 (3)	−7 (5)	−4 (3)	4 (3)	−15 (3)
C3′	3.83 (6)	7 (3)	9 (3)	−9 (3)	−5 (2)	−1 (3)	−5 (3)	−4 (3)	−6 (3)	13 (3)	−17 (3)	−3 (3)	−7 (5)	24 (3)	2 (3)	16 (3)
C4′	3.85 (6)	−10 (3)	−4 (3)	−16 (3)	8 (2)	−12 (3)	3 (3)	−4 (3)	1 (3)	−25 (3)	−2 (3)	1 (3)	−13 (5)	−14 (3)	−4 (3)	−13 (3)
O4′	6.56 (4)	1 (2)	−3 (2)	0 (3)	7 (2)	−7 (2)	−3 (3)	9 (3)	−4 (2)	5 (2)	2 (2)	5 (2)	4 (4)	2 (2)	10 (2)	4 (2)
O2′	6.57 (4)	−3 (2)	5 (2)	−10 (3)	−1 (2)	7 (3)	10 (3)	−1 (3)	5 (3)	5 (2)	−6 (2)	−6 (2)	1 (4)	−4 (2)	2 (2)	4 (2)
O3′	6.59 (5)	−11 (2)	−5 (2)	−7 (3)	7 (2)	−6 (2)	−8 (3)	4 (3)	13 (3)	−11 (2)	9 (2)	−4 (3)	−8 (5)	−6 (2)	6 (2)	1 (3)
C5′	3.63 (6)	−21 (3)	−1 (3)	−6 (3)	−3 (3)	5 (3)	6 (3)	2 (3)	−2 (3)	16 (3)	0 (3)	−8 (3)	11 (5)	19 (3)	5 (3)	15 (3)
O5′	6.53 (5)	2 (2)	7 (2)	−5 (3)	−2 (2)	8 (3)	5 (3)	−9 (3)	−6 (3)	−1 (2)	5 (2)	5 (3)	−9 (4)	−5 (2)	−1 (2)	−3 (3)
H5	0.93 (4)	−18 (3)	10 (3)	5 (3)	1 (3)	−5 (4)	4 (5)	6 (5)	−10 (4)							
H6	0.98 (3)	−10 (3)	9 (3)	−7 (3)	2 (4)	3 (4)	8 (5)	−10 (5)	0 (4)							
H41	0.91 (3)	4 (3)	−19 (3)	21 (3)	−13 (4)	−12 (4)	5 (5)	2 (5)	−11 (4)							
H42	0.74 (3)	−20 (3)	8 (3)	−10 (4)	−8 (4)	−4 (4)	−6 (5)	−4 (5)	−6 (4)							
HO2′	0.64 (3)	9 (3)	25 (3)	−6 (3)	−13 (3)	10 (4)	10 (5)	2 (5)	8 (4)							
HO3′	0.69 (3)	−7 (3)	9 (3)	9 (3)	−16 (3)	−4 (4)	5 (4)	3 (4)	−7 (4)							
HO5′	0.74 (3)	−2 (3)	15 (3)	12 (3)	−11 (4)	4 (4)	10 (4)	20 (5)	−10 (4)							
H5′1	0.84 (4)	11 (3)	5 (3)	7 (4)	6 (4)	11 (4)	−1 (5)	10 (5)	2 (4)							
H5′2	0.96 (4)	−14 (3)	4 (3)	−21 (3)	6 (4)	0 (4)	8 (5)	8 (5)	25 (4)							
H1′	0.93 (3)	8 (3)	−17 (3)	2 (3)	−10 (4)	−21 (4)	2 (5)	−6 (5)	9 (4)							
H2′	1.03 (3)	−14 (3)	−2 (3)	13 (3)	4 (4)	−5 (4)	7 (5)	−11 (5)	22 (4)							
H3′	0.99 (3)	−7 (3)	1 (3)	19 (3)	−2 (4)	7 (4)	−5 (5)	1 (5)	−8 (4)							
H4′	1.00 (4)	6 (3)	−17 (3)	0 (3)	6 (4)	−17 (5)	−13 (5)	12 (5)	−5 (4)							

valence-shell electrons. At the completion of refinement, assuming a total of 128 electrons per molecule, we obtain 34.51 (7) *K*-shell and 93.5 (3) valence-shell electrons. The population parameters in Table 4 have been scaled so that $\sum p_v = 94$. The net atomic charges are in Table 7(*a*). Pairs of κ parameters for the same atom type were found to have the same values within experimental error. Average values are $\kappa = 0.974(2), 0.995(5)$ and $1.040(6)$ for O, N and C, respectively, indicating as expected that the valence shell density is expanded for O pseudoatoms and contracted for C. The difference-Fourier synthesis of residual electron density shows features ranging from 0.13 to $-0.16(8)\,e\,Å^{-3}$, which are no more than marginally significant.*

* Figures showing the residual electron density and the deformation density for cytidine together with tables of least-squares planes data and a list of reflection data have been deposited with the IUCr (Reference: BK0024). Copies may be obtained through the Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

Table 5. *Geometric parameters for cytidine*

Columns are: (1) values from present X-ray diffraction at 123 K, uncorrected; (2) values from (1) with thermal vibration corrections; (3) values from X-ray diffraction at room temperature (Ward, 1993); (4) values from three-dimensional photographic data at room temperature (Furberg, Petersen & Romming, 1965).

(*a*) Bond distances involving non-H atoms (Å)

	(1)	(2)	(3)	(4)
N1—C2	1.390 (1)	1.392	1.393 (4)	1.379 (6)
N1—C1′	1.487 (1)	1.488	1.490 (4)	1.497 (6)
N1—C6	1.367 (1)	1.368	1.363 (4)	1.368 (6)
N3—C2	1.355 (1)	1.356	1.351 (4)	1.361 (6)
N3—C4	1.343 (1)	1.345	1.346 (4)	1.335 (6)
N4—C4	1.334 (1)	1.335	1.319 (4)	1.333 (7)
O2—C2	1.250 (1)	1.252	1.241 (4)	1.246 (5)
O4′—C1′	1.411 (1)	1.412	1.410 (4)	1.412 (6)
O4′—C4′	1.448 (1)	1.450	1.460 (4)	1.445 (5)
O2′—C2′	1.419 (1)	1.420	1.412 (4)	1.425 (5)
O3′—C3′	1.410 (1)	1.412	1.412 (4)	1.410 (5)
O5′—C5′	1.417 (1)	1.419	1.419 (5)	1.404 (6)
C4—C5	1.428 (1)	1.430	1.417 (5)	1.420 (7)
C5—C6	1.356 (1)	1.357	1.334 (5)	1.345 (7)
C1′—C2′	1.541 (1)	1.543	1.534 (5)	1.527 (6)
C2′—C3′	1.525 (1)	1.526	1.518 (5)	1.507 (6)
C3′—C4′	1.531 (1)	1.533	1.515 (5)	1.533 (6)
C4′—C5′	1.519 (1)	1.521	1.507 (5)	1.509 (6)

(*b*) Bond angles involving non-H atoms (°)

C1′—O4′—C4′	110.6 (1)	O4′—C1′—C2′	106.8 (1)
C1′—N1—C2	117.6 (1)	O4′—C1′—N1	109.4 (1)
C1′—N1—C6	121.9 (1)	N1—C1′—C2′	111.5 (1)
C2—N1—C6	120.6 (1)	O2′—C2′—C1′	109.4 (1)
C2—N3—C4	119.6 (1)	O2′—C2′—C3′	112.9 (1)
O2—C2—N1	118.5 (1)	C1′—C2′—C3′	100.7 (1)
O2—C2—N3	121.9 (1)	O3′—C3′—C2′	112.5 (1)
N1—C2—N3	119.7 (1)	O3—C3′—C4′	114.0 (1)
N3—C4—N4	117.1 (1)	C2′—C3′—C4′	102.2 (1)
N3—C4—C5	122.1 (1)	O4′—C4′—C3′	104.1 (1)
N4—C4—C5	120.9 (1)	O4′—C4′—C5′	110.2 (1)
C4—C5—C6	117.9 (1)	C3′—C4′—C5′	115.2 (1)
N1—C6—C5	121.0 (1)	O5′—C5′—C4′	113.2 (1)

(*c*) Torsion angles involving non-H atoms (°)

C6—N1—C2—N3	−2.6 (1)	N1—C1′—C2′—O2′	−150.7 (1)
C6—N1—C2—O2	178.5 (1)	O4′—C1′—C2′—C3′	−29.2 (1)
C1′—N1—C2—N3	178.2 (1)	O4′—C1′—C2′—O2′	89.8 (1)
C1′—N1—C2—O2	−0.7 (1)	C2′—C1′—O4′—C4′	7.8 (1)
C2—N1—C6—C5	−0.5 (1)	C1′—C2′—C3—C4′	38.4 (1)
C1′—N1—C6—C5	178.6 (1)	C1′—C2′—C3′—O3′	161.1 (1)
C2—N1—C1′—C2′	79.2 (1)	O2′—C2′—C3′—C4′	−78.1 (1)
C2—N1—C1′—O4′	−162.9 (1)	O2′—C2′—C3′—O3′	44.6 (1)
C6—N1—C1′—C2′	−100.0 (1)	C2′—C3′—C4′—O4′	−35.1 (1)
C6—N1—C1′—O4′	17.9 (1)	C2′—C3′—C4′—C5′	−155.8 (1)
N1—C2—N3—C4	5.0 (1)	O3′—C3′—C4′—O4′	−156.7 (1)
O2—C2—N3—C4	−176.1 (1)	O3′—C3′—C4′—C5′	82.5 (1)
C2—N3—C4—C5	−4.4 (1)	C3′—C4′—C4′—C1′	17.2 (1)
C2—N3—C4—N4	175.6 (1)	C5′—C4′—O4′—C1′	141.3 (1)
N3—C4—C5—C6	1.3 (1)	C3′—C4′—C5′—O5′	47.1 (1)
N4—C4—C5—C6	−178.4 (1)	O4′—C4′—C5′—O5′	−70.2 (1)
C4—C5—C6—N1	1.2 (1)	N1—C1′—O4′—C4′	−113.1 (1)
N1—C1′—C2′—C3′	90.2 (1)		

(*d*) Distances (Å) and angles (°) for molecular interactions

Values obtained from: (1) current study, X-ray, 123 K; (2) Ward (1993), X-ray, 298 K.

Hydrogen-bond interactions*

	X···*Y*		*X*—H		H···*Y*		∠*X*—H···*Y*	
X—H···*Y*	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
N4—H42···O2^{i}	2.913 (2)	2.947 (4)	1.00 (3)	0.84 (4)	2.00 (3)	2.22 (4)	152 (2)	145 (4)
N4—H41···O2^{i}	2.883 (2)	2.922 (5)	1.00 (3)	1.00 (4)	1.97 (2)	1.96 (5)	150 (1)	159 (4)
O3′—HO3′···N3^{ii}	2.835 (2)	2.866 (3)	0.98 (3)	0.91 (4)	1.86 (3)	1.96 (5)	176 (3)	171 (5)
O5′—HO5′···O2^{ii}	2.705 (2)	2.707 (4)	0.98 (2)	0.99 (4)	1.74 (2)	1.74 (5)	169 (2)	168 (4)
O2′—HO2′···O3′^{iii}	2.801 (2)	2.845 (4)	0.98 (3)	0.84 (4)	2.03 (3)	2.13 (4)	134 (2)	143 (3)

Table 5 (*cont.*)

Other short X—H···O distances

	$X\cdots Y$		X—H		H···Y		$\angle X$—H···Y	
X—H···Y	(1)	(2)	(1)	(2)	(1)	(2)	(1)	(2)
O2′—HO2′···O3′^{iv}	2.799 (2)	2.787 (4)	0.98 (3)	0.84 (4)	2.29 (3)	2.41 (4)	112 (2)	108 (3)
C5—H5···O5′^{v}	3.322 (2)	3.341 (5)	1.08 (3)	0.96 (3)	2.30 (3)	2.44 (3)	157 (2)	157 (3)
C6—H6···O5′^{iv}	3.210 (2)	3.245 (5)	1.08 (2)	0.94 (4)	2.21 (2)	2.34 (4)	153 (1)	160 (4)

Symmetry codes: (i) $\frac{1}{2}+x, \frac{1}{2}-y, -z$; (ii) $1-x, y-\frac{1}{2}, \frac{1}{2}-z$; (iii) $\frac{1}{2}-x, -y, \frac{1}{2}+z$; (iv) x, y, z; (v) $\frac{3}{2}-x, -y, z-\frac{1}{2}$.

* H-atoms have the modified positional parameters with e.s.d.'s from unmodified values (Table 3).

We have also carried out a charge-density refinement for cytidine at the monopole level (Coppens *et al.*, 1979) with the same radial Hartree–Fock atomic scattering factors as described above. There was only one electron-population parameter (p_v) per pseudoatom, together with a valence-shell κ-parameter constrained to be the same for each pseudoatom type (C, N, O and H). For the H atoms, positional parameters and an isotropic m.s. displacement parameter were refined. This simplified model is very similar to that used by Pearlman & Kim (1985, 1990), who assumed a more contracted H atom with $\alpha = 5.29$ Å^{-1} for the H atoms. For cytidine, our monopole refinement gave final values $R(F^2) = 0.070$, $R_w(F^2) = 0.102$ and $S = 1.951$. Net atomic charges from this refinement are in Table 7(*b*).

In order to make a more direct comparison of our cytidine results with those for cytosine monohydrate, we have carried out new refinements for cytosine monohydrate using the X-ray data of Weber & Craven (1990). Initially we assumed fixed nuclear positional and m.s. displacement parameters for H determined by neutron diffraction (Weber, Craven & McMullan, 1980). Otherwise, the full multipole model was used as described above for cytidine. This refinement was carried out using all 4034 reflections with minimization of the residual $\sum w(F_c^2 - F_o^2)^2$ and using weights based only on counting statistics, whereas previously Weber & Craven (1990) used only 1432 reflections having $F^2 > 2.5\sigma(F^2)$. Their minimization was for the residual $\sum w(|F_o| - |F_c|)^2$ and weights were estimated from multiple observations of certain reflections. With the m.s. displacement factors for C, N and O as variables, in contrast to the earlier refinement, they were found to be in very good agreement with the values obtained from neutron diffraction. The refinement was completed with fixed nuclear positional and m.s. anisotropic displacement parameters for all nuclei from neutron diffraction (Weber, Craven & McMullan, 1980), so that the variables consisted only of a scale factor for the K-shell scattering, three κ-parameters (C, N and O) and 172 electron-population parameters. A fixed standard value of $\alpha = 4.69$ Å^{-1} was assumed for H. Final values were $R(F^2) = 0.102$, $R_w(F^2) = 0.016$ and $S = 1.18$. After scaling to the actual total number of electrons (68; 18 K-shell and 50 valence-shell) we obtained 17.1 K-shell and 50.9 valence-shell electrons. Scaling to a total of 50 valence-shell electrons gave a net negative charge 0.16 (4) e on the water molecule, with a corresponding positive charge on cytosine. The electron population parameters in Table 6 are scaled so that the cytosine and water molecules are both electrically neutral. This was done in order to allow calculation of molecular dipole moments for water and cytosine. Corresponding atomic net charges are in Table 7(*c*).

We have also carried out a refinement for cytosine monohydrate with a model very similar to that of Pearlman & Kim (1990), in which we made no use of neutron diffraction results. All atomic positional and m.s. displacement parameters (U_{iso} for H) were refined. The charge-density variables consisted of the monopole population parameters only and a single κ-parameter for each pseudoatom type, except for the fixed value of $\alpha = 5.29$ Å^{-1} for H corresponding to the value assumed by Pearlman & Kim (1990). In this refinement, values of U_{iso} for H (or D) were small in terms of estimated errors [0.014 (5) for D1, 0.010 (5) Å^2 for DW1] and became negative, but not significantly so, for D4. Refinement was completed with a small fixed positive value for D4 and gave $R(F^2) = 0.121$, $R_w(F^2) = 0.022$ and $S = 1.62$. Final atomic point charges derived from this refinement are given in Table 7(*d*).

Discussion

Molecular geometry and crystal packing

In Table 5(*a*), the bond lengths presently obtained are shown before and after correction for segmented body thermal libration and are listed together with the room-temperature values given by Ward (1993) and Furberg, Petersen & Romming (1965). Our low-temperature values tend to be longer than the corresponding room-temperature values presumably because the latter are uncorrected for thermal libration effects. The agreement is generally satisfactory, except possibly for the O4′—C4′ bond for which Ward (1993) reports an uncorrected distance already longer [1.460 (4) Å] than ours [1.448 (1) or 1.450 Å, corrected]. The O2=C2 double bond is shorter [1.252 (1) Å] than is observed in cytosine monohydrate from neutron diffraction [1.262 (1) Å; Weber, Craven & McMullan, 1980], whereas the C—N bonds at N1 are longer [1.392 (1)

Table 6. *Electron population parameters for cytosine monohydrate*

All values except p_v are $\times 10$. Populations for the higher multipoles are normalized (Epstein, Ruble & Craven, 1982). The final κ parameters were 1.067 (3), 1.023 (3) and 0.998 (2) for C, N and O atoms, respectively. The fixed κ value for H was 1.24. In this full multipole refinement, all atomic positional and m.s. displacement parameters were fixed at values from neutron diffraction (Weber, Craven & McMullan, 1980).

	p_v	$d1$	$d2$	$d3$	$q1$	$q2$	$q3$	$q4$	$q5$	$o1$	$o2$	$o3$	$o4$	$o5$	$o6$	$o7$
N1	5.39 (5)	−4 (2)	−3 (2)	−10 (2)	4 (3)	2 (3)	−2 (2)	5 (2)	−28 (2)	20 (4)	7 (4)	−14 (3)	−8 (3)	9 (3)	16 (3)	−4 (3)
D1	0.56 (3)	11 (2)	32 (2)	2 (2)	−12 (4)	22 (4)	17 (5)	5 (5)	5 (4)							
C2	3.88 (4)	4 (2)	6 (3)	−6 (2)	10 (3)	3 (3)	−8 (2)	−2 (3)	−21 (2)	−35 (4)	−11 (6)	15 (3)	3 (5)	5 (3)	5 (4)	2 (3)
O2	6.37 (3)	−4 (2)	7 (2)	6 (1)	−7 (2)	2 (3)	1 (2)	−6 (2)	−3 (2)	6 (2)	8 (3)	−3 (2)	8 (4)	−6 (2)	−2 (3)	−3 (2)
N3	5.20 (4)	−11 (2)	3 (2)	7 (2)	−22 (3)	−5 (3)	−4 (2)	−6 (2)	−9 (2)	7 (4)	0 (4)	4 (3)	−9 (3)	16 (3)	−1 (3)	−6 (3)
C4	4.10 (5)	−3 (3)	−8 (3)	8 (2)	7 (3)	−2 (3)	−19 (2)	−7 (3)	−35 (2)	−51 (4)	−3 (5)	18 (3)	3 (4)	−5 (3)	6 (3)	2 (3)
N4	4.92 (5)	−5 (2)	−3 (2)	−4 (2)	1 (2)	1 (3)	2 (2)	3 (3)	5 (2)	13 (3)	−10 (4)	−16 (3)	−6 (3)	−12 (2)	−1 (3)	−4 (3)
D3	0.89 (3)	12 (2)	−30 (3)	−16 (2)	−12 (4)	−7 (4)	4 (5)	16 (5)	−32 (4)							
D4	0.93 (2)	−36 (3)	−13 (2)	16 (2)	20 (3)	11 (4)	−14 (5)	0 (5)	−11 (4)							
C5	3.93 (4)	1 (2)	4 (3)	8 (2)	−5 (3)	3 (4)	2 (2)	−7 (3)	−19 (2)	11 (4)	18 (5)	−20 (3)	19 (4)	2 (3)	−12 (4)	2 (3)
H5	1.02 (3)	−10 (3)	−3 (2)	3 (2)	4 (3)	6 (4)	21 (5)	−26 (5)	−24 (4)							
C6	3.91 (5)	6 (3)	8 (3)	−21 (2)	−19 (3)	−1 (3)	−32 (2)	10 (3)	−12 (3)	−15 (4)	−10 (5)	11 (3)	2 (4)	0 (3)	0 (3)	10 (3)
H6	0.89 (3)	−10 (2)	18 (3)	4 (2)	−7 (4)	−7 (4)	−21 (5)	29 (5)	8 (4)							
OW	6.28 (3)	−11 (2)	2 (1)	−15 (2)	−3 (2)	−3 (2)	1 (2)	3 (2)	−7 (2)	0 (3)	−7 (2)	−8 (3)	−3 (3)	5 (3)	−1 (2)	13 (3)
DW1	0.87 (2)	−7 (2)	7 (2)	−28 (2)	−15 (4)	−5 (4)	16 (4)	0 (5)	2 (4)							
DW2	0.85 (2)	−5 (3)	6 (2)	19 (2)	9 (4)	4 (4)	−9 (5)	13 (5)	18 (4)							

Table 7. *Atomic point charges for the cytosine base*

Values are $q = Z_v - p_v$, where Z_v is the number of valence-shell electrons in the neutral pseudoatom and p_v is the observed number (the monopole population parameter). Columns are:

(*a*) Cytidine. Full multipole refinement assuming $\kappa_H = 1.24$.

(*b*) Cytidine. Monopole refinement assuming $\kappa_H = 1.40$.

(*c*) Cytosine from the monohydrate. Full multipole refinement assuming $\kappa_H = 1.24$.

(*d*) Cytosine from the monohydrate. Monopole refinement assuming $\kappa_H = 1.40$.

(*e*) 1-β-D-Arabinosylcytosine. Monopole refinement assuming $\kappa_H = 1.40$, with average values for the e.s.d.'s as given by Pearlman & Kim (1990).

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)
N1	−0.23 (6)	−0.23 (6)	−0.39 (5)	−0.39 (4)	−0.49 (7)
D1	—	—	+0.44 (3)	+0.36 (3)	—
C2	+0.34 (6)	+0.58 (6)	+0.12 (4)	+0.50 (3)	+0.63 (9)
O2	−0.48 (4)	−0.57 (7)	−0.37 (3)	−0.53 (2)	−0.71 (6)
N3	−0.40 (6)	−0.63 (7)	−0.20 (4)	−0.49 (2)	−0.59 (7)
C4	+0.23 (6)	+0.23 (6)	−0.10 (5)	+0.29 (3)	+0.43 (9)
N4	−0.07 (6)	−0.15 (6)	+0.08 (5)	−0.41 (5)	−0.33 (7)
D3/H41	+0.09 (3)	+0.19 (4)	+0.11 (3)	+0.22 (4)	+0.31 (5)
D4/H42	+0.26 (3)	+0.28 (5)	+0.07 (4)	+0.28 (2)	+0.39 (5)
C5	−0.06 (6)	−0.09 (5)	+0.07 (3)	−0.10 (3)	−0.31 (9)
H5	+0.07 (4)	+0.08 (4)	−0.02 (3)	+0.05 (2)	+0.19 (5)
C6	+0.11 (6)	+0.30 (6)	+0.09 (5)	+0.17 (4)	+0.12 (9)
H6	+0.02 (3)	+0.11 (5)	+0.11 (3)	+0.05 (3)	+0.12 (5)

versus 1.372 (1) Å for C2—N1; 1.367 (1) *versus* 1.357 (1) Å for N1—C6]. These effects are most likely owing to the formation of the glycosidic N1—C1′ bond in cytidine and possibly also the involvement of O2 as acceptor in only two hydrogen bonds as compared with three in cytosine monohydrate.

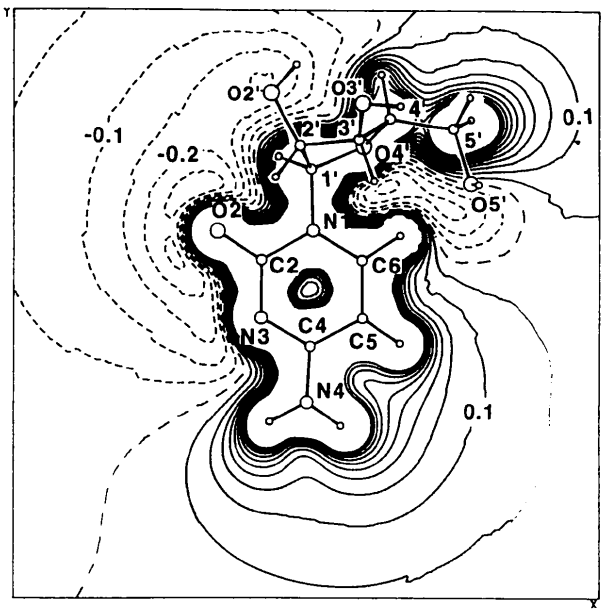
The cytosine base including the ring substituents N4 and O2 is almost planar (see deposition footnote). The mean atomic displacement from the best least-squares plane is 0.016 Å, except N3 which is 0.054 Å from the plane. The displacement of N3 is towards HO3′ in the intermolecular hydrogen bond O3′—HO3′···N3. Here, HO3′···N3 forms an angle of 46° with the cytosine ring plane. The ribose ring is nonplanar, with C3′ displaced 0.60 Å away from the plane formed by the other four atoms. Thus, the ribose in cytidine has the C3′-*endo* conformation. The planes of the cytosine and ribose rings (including C3′) make a dihedral angle of 79.9°. The relative orientation of cytosine base and ribose corresponds to an *anti* conformation of the nucleoside with a torsion angle χ(C6—N1—C1′—O4′) of 17.9 (1)°.

Each cytidine molecule forms conventional hydrogen bonds involving every N—H and O—H H atom (see Fig. 2 in Furberg, Petersen & Romming, 1965). The hydrogen-bonding distances at 123 K (Table 5*d*) are shorter than at room temperature by average values of 0.023 Å for O···O and 0.35 Å for O···N, indicating that the intermolecular interactions are slightly stronger at the lower temperature. The weak C—H···O5′ interactions are also shorter (0.027 Å). One of these is intermolecular (H5···O5′, 2.30 Å) and the other is intramolecular (H6···O5′, 2.21 Å). Jeffrey & Saenger (1991) in their Fig. 17.31 point out that in the structure determined by Furberg, Petersen & Romming (1965), HO3′ may be involved in an intramolecular HO2′···O3′ interaction as well as the more conventional intermolecular hydrogen bond, HO2′···O3′. The distances and angles presently obtained (Table 5*d*) indicate that the intramolecular HO2′···O3′ distance is longer (2.29 *versus* 2.03 Å) and the intramolecular angle O2′—HO2′···O3′ is more severely bent (112 *versus* 134°) than in the intermolecular hydrogen bond. In Figs. 3–6, we show maps of electrostatic potential in small clusters of cytidine molecules. Map sections are through regions involving the hydrogen bonds and other short interatomic distances.

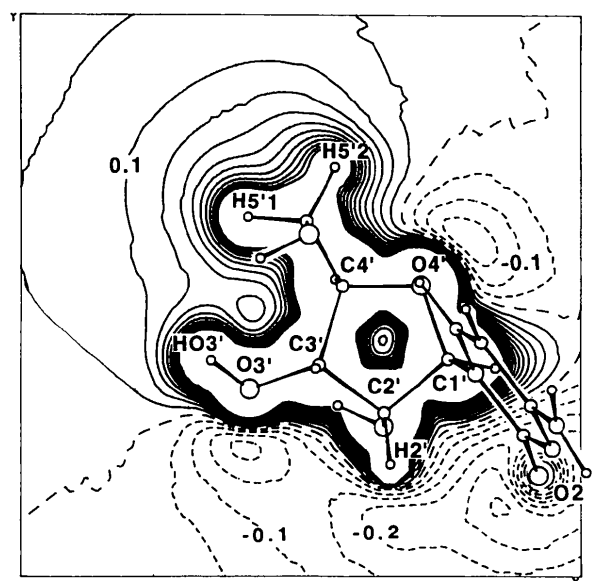
Molecular electrostatic potential for cytidine

Maps of the electrostatic potential for a single molecule (Fig. 3) and for clusters for three cytidine molecules isolated from the crystal structure (Figs. 4, 5

and 6) were derived from the complete set of pseudoatom multipole parameters (Table 4) using the procedure of Stewart & Craven (1993). It is emphasized that although these molecules have been removed from the crystal environment, they retain the configuration in the crystal and remain polarized by crystal environmental effects



(*a*)



(*b*)

Fig. 3. Electrostatic potential of β-cytidine for a complete single molecule removed from the crystal structure but with atoms at rest. Contours are at intervals of 0.05 e Å^{-1} with electronegative contours as dashed lines. This is the total electrostatic potential derived from the full pseudoatom model (Stewart & Craven, 1993). (*a*) Section in the plane of the cytosine base, (*b*) section in the best plane through the atoms of the ribose ring.

such as hydrogen bonding. Regions of electronegativity are indicated by dashed contours.

Fig. (3*a*) shows the electrostatic potential in the plane of the cytosine base. It can be seen that there is an extensive region of electronegativity around O2 and N3 and an electropositive region around the N4 amino group and C5. Similar features are shown in the map for cytosine isolated from the monohydrate structure (Fig. 3*a* from the earlier refinement by Weber & Craven, 1990; and here in Fig. 7 from the new refinement). In cytidine, electronegativity is concentrated around O2, rather than being almost uniformly distributed over O2 and N3, as for cytosine in the monohydrate. There are differences in the crystal environment of this electronegative region. In both structures, N3 forms one hydrogen bond, but the number of hydrogen bonds at O2 is reduced from three in cytosine monohydrate to two in cytidine, owing to steric hindrance by the sugar. However, in cytidine the electronegativity around O2 is enhanced by contributions from the sugar O4′ and O2′ and by the absence of the electropositive H1.

Fig. 4 shows the electrostatic potential for a cluster of three cytidine molecules which have the same configuration as in the crystal structure. The section is chosen to illustrate the potential in the region of hydrogen bonding at O2 and also the hydrogen bond N4—H41···O2′. In previous studies of the potential in molecular clusters, as in the case of phosphorylethanolamine (Swaminathan & Craven, 1984) and urea (Stewart, 1991), the hydrogen bonds were characterized by a weak bridge of electropositivity along the H···O separation. This is the result of superposing regions of electropositivity from H and electronegativity from O, which largely cancel each other. The effect in cytidine is shown in Figs. 4(*c*) and (*d*). The resultant weakly electropositive bridging occurs in each of the O—H···O and N—H···O hydrogen bonds (Fig. 4*b*).

The electrostatic potential in Fig. 5 is for a different cluster of three molecules, namely those involved in the hydrogen bonding at O2′. The section is through O2′ and O3′ of the same molecule (I) and through O3′ of molecule (II) which forms the intermolecular interaction O2′—HO2′···O3′. Molecule (III) provides the N—H group which also hydrogen bonds with O2′. The electrostatic potential for molecule (I) only is shown in Fig. 5(*c*). In the HO2′···O3′ region of this map, the overlap of the positive potential from HO2′ and the negative potential from O3′ gives a net electronegative potential. This indicates that HO2′ is inefficient in satisfying the capacity of O3′ as an intramolecular hydrogen-bond acceptor. When the cluster is formed, a weakly electropositive bridge is developed between HO2′ of molecule (I) and O3′ of molecule (II) (Fig. 5*b*). This is consistent with the existence of an intermolecular hydrogen bond. However, compared with the potential at HO2′···O3′ in the isolated molecule (I) (Fig. 5*c*), the presence of molecule (II) further increases

the electronegative potential near O3′ (Fig. 5*b*), indicating a weakening of the HO2′···O3′ intramolecular interaction.

In Fig. 6, the electrostatic potential is shown for a three-molecule cluster consisting of the molecules closest to the region of the weak C—H···O5′ interactions. On the basis of the distances and angles involved (Table 5*d*), these might be described as weak hydrogen-bonding interactions (Taylor & Kennard, 1982; Jeffrey & Saenger, 1991). It is of interest that the intramolecular interaction C6—H6···O5′ has the shorter H···O distance (2.21 Å), but the H···O region is electronegative (seen in both Fig. 6 and for an isolated single molecule in Fig. 3). The intermolecular interaction C5—H5···O5′ has a longer H···O distance (2.30 Å), but it exhibits a weakly electropositive bridge which is a common feature for conventional hydrogen bonds. The C—H groups at C5 and C6 lack the strong polarity exhibited by the

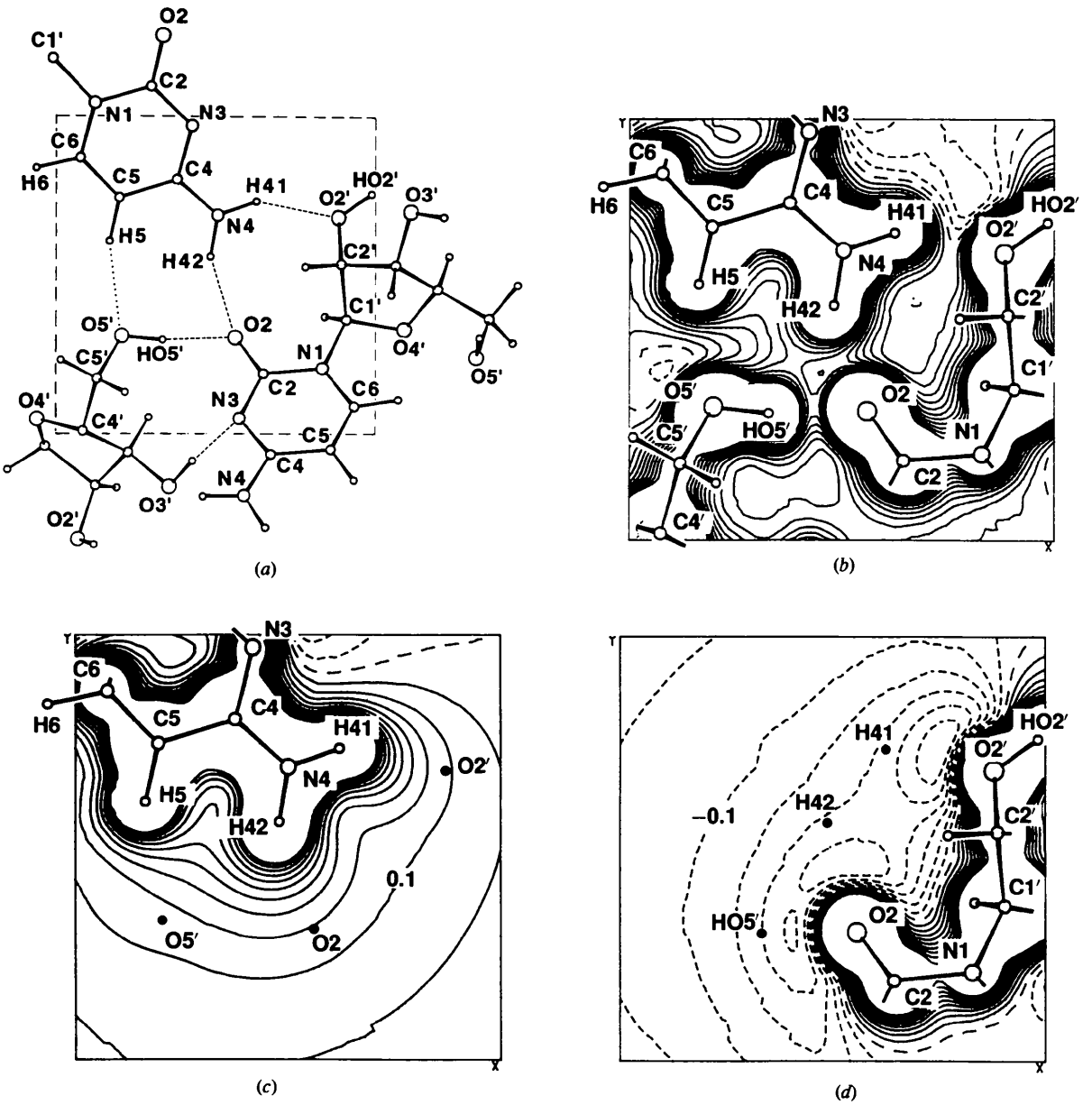


Fig. 4. The electrostatic potential for a cluster of three isolated β-cytidine molecules in the same configuration as in the crystal structure. The potential is derived from the full pseudoatom model and is contoured as in Fig. 3. The section is through the atom O2 and through N4 and O5′, which form hydrogen bonds with O2. This section is also close to the atoms involved in the weak C5—H5···O5′ intermolecular interaction. (*a*) The projection of the cluster onto the plane of the section; (*b*) the total electrostatic potential in the section; (*c*) the contribution to the potential from the single molecule at the top left in (*a*); (*d*) the contribution to the potential from the molecule at the right in (*a*).

O—H groups in cytidine. Thus, the p_v-values for H5 and H6 (0.93 and 0.94 e) correspond to atoms which are almost neutral when compared with those for the electron-depleted hydroxyl H atoms (0.64, 0.69 and 0.74 e; Table 4*b*). The electrostatic potential near H5 and H6 will not be greatly affected by the H atoms themselves, but could be influenced by other nearby polar groups. In the case of cytidine, we hesitate to characterize the C—H···O interactions as hydrogen bonds.

Experimental studies of the charge distribution in hydrogen-bonded systems are valuable because the hydrogen bond is largely electrostatic in nature. Maps of the electrostatic potential for clusters of cytidine molecules show how the long-range effects of polar groups on hydrogen bonding can be important. However, such maps should not be overinterpreted. It must be remembered that the potential within a cluster will change as molecules are added or deleted.

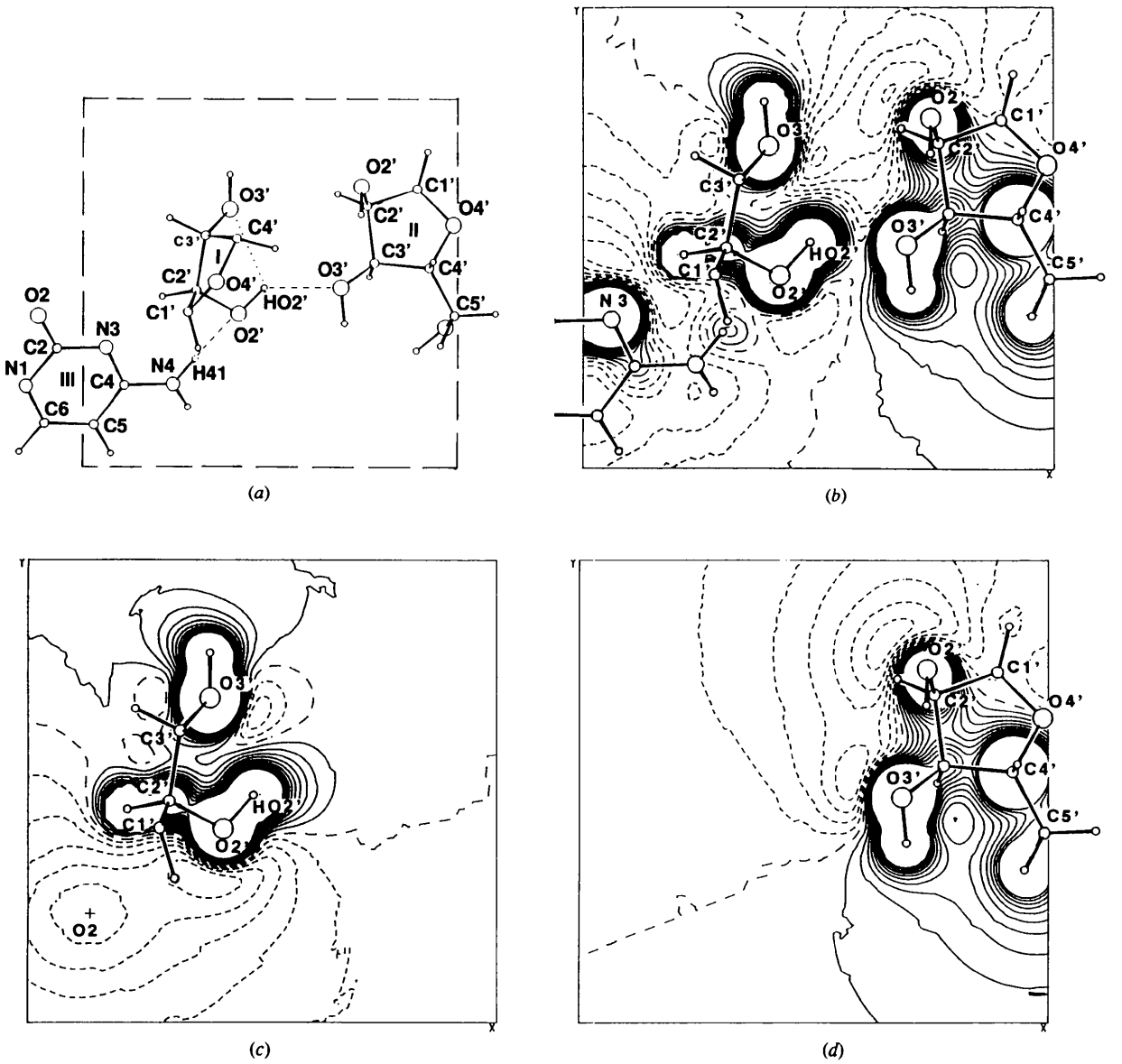


Fig. 5. The electrostatic potential for a cluster of three isolated β-cytidine molecules. The section is through O2′—HO2′ and O3′ of the same molecule (I), and through the O3′ atom of molecule (II), which forms an intermolecular hydrogen bond and an intramolecular interaction. The O2′ atom from (I) also accepts another intermolecular hydrogen bond from molecule (III), which is 0.7–2.2 Å below the plane. (*a*) The projection of the cluster onto the plane of the section; (*b*) the total electrostatic potential in the section; (*c*) the contribution to the potential from the single molecule (I) at the center in (*a*). The negative potential at the left bottom is from the O2 atom which is 1.2 Å above the plane; (*d*) the contribution to the potential from molecule (II).

Beyond the van der Waals radius, the electrostatic potential for an atom is given well by q/r, where q is the net atomic charge and r is the distance from the nucleus (Stewart & Craven, 1993). Thus, the contribution of an atom to the electrostatic potential is spherically symmetric about its nucleus and remains significant at considerable distances. For example, the electrostatic interaction energy of two charges separated by 2.8 Å changes by only 20% when the distance increases to 3.5 Å. Even for strong hydrogen bonds H···X, in which H and X carry large charges and are close together, other charged groups further away may make significant contributions to the interaction. The present study of cytidine helps to establish the notion of multicenter hydrogen bonds as described for example by Jeffrey & Saenger (1991).

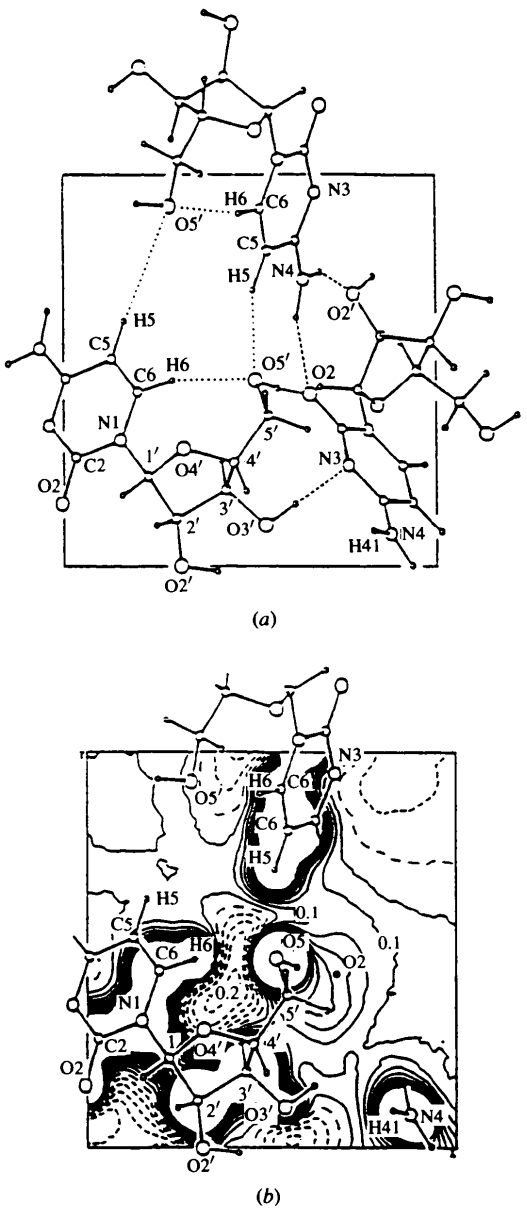


Fig. 6. The electrostatic potential for a cluster of three isolated β-cytidine molecules in the same configuration as in the crystal structure. The section is through the O5′ atom and through the atoms H5 and H6 involved in the two weak C—H···O5′ interactions. (*a*) The projection of the cluster onto the plane of the section; (*b*) the total electrostatic potential in the section.

The effect of κ_H on the electrostatic properties of cytosine and water

Compared with the electron density for an isolated H atom, the formation of a covalent bond leads to an appreciable contraction of the spherically averaged distribution (Stewart, Davidson & Simpson, 1965). From a consideration of *ab initio* wavefunctions for ten small molecules, including H_2O and CH_4, Hehre, Stewart & Pople (1969) derived a single Slater-type radial function for bonded hydrogen of the form $\exp(-\kappa_H\alpha_0 n)$, where $\alpha_0 = 2\,\text{bohr}^{-1} = 3.78\,\text{Å}^{-1}$ and the contraction parameter κ_H has values in the range from 1.16 (for CH_4) to 1.31 (for HF). They recommended an average value of $\kappa_H = 1.24$, which gives $\alpha = \kappa_H\alpha_0 = 4.69\,\text{Å}^{-1}$. It should be noted that these values are based on the optimized energy for the various molecular wavefunctions rather than an optimized fit to the charge-density distribution, as in our present diffraction studies. Nevertheless, we assumed a fixed standard value of $\kappa_H = 1.24$ in our initial refinements of cytidine and cytosine monohydrate. The hydrogen pseudoatom assumed by Pearlman & Kim (1990; $\kappa_H = 1.40$) is more contracted and therefore takes on a smaller fraction of the total molecular charge-density distribution. In compensation, neighboring bonded pseudoatoms expand to take a greater share. Thus, a variation in κ_H is expected to cause differences in the monopole populations (pseudoatom charges) of all the pseudoatoms. This is not a serious concern unless these somewhat different sets of atomic parameters differ in the efficiency with which they fit the total molecular charge density. It is important to determine whether variation in κ_H might lead to significant differences in molecular electrostatic properties. In order to investigate the effect of κ_H on molecular properties, we have carried out further full multipole refinements for the structure of cytosine monohydrate. In the initial refinement described above, with $\kappa_H = 1.24$ as a fixed parameter, we obtained $R_w(F^2) = 0.016$, $R(F^2) = 0.101$ and $S = 1.18$. When κ_H was introduced as a variable with the model otherwise unchanged, the hydrogen pseudoatoms expanded [$\kappa_H = 1.08\,(1)$] with a very small improvement in agreement [$R_w(F^2) = 0.015$, $R(F^2) = 0.101$ and $S = 1.16$]. With $\kappa_H = 1.40$ as a fixed parameter, as assumed by Pearlman & Kim (1990), the agreement deteriorated slightly [$R_w(F^2) = 0.016$, $R(F^2) = 0.102$ and $S = 1.22$]. The electrostatic potential for isolated molecules of cytosine and water obtained assuming fixed

values of $\kappa_H = 1.08$, 1.24 and 1.40 are shown in Fig. 7. It appears that with contraction of the hydrogen pseudoatoms (κ_H increasing from 1.08 to 1.40), there is a progressive increase in the polarity of both the cytosine and water molecules. The effect can also be seen as an increase in the molecular dipole moment, which for cytosine increases by 3.0σ from 5.4(16) to 11.7(13) Debyes, whilst for water the increase is 1.3σ from 1.88(28) to 2.36(25) Debyes. There are no significant changes in the direction of these molecular dipoles which, within experimental error, point along the C=O bond for cytosine and the twofold axis for water. We conclude that indeed changes in κ_H may cause significant changes in molecular electrostatic properties.

Stewart (1991) has carried out a series of full multipole structure refinements which effectively included κ_H as a variable. His refinements all made use of X-ray data collected in our laboratory. The hydrogen nuclei were all assigned fixed positional and m.s. displacement parameters, as determined by neutron diffraction. Values for κ_H were 1.47(5) for urea,* 1.23(3) for imidazole, 1.19(2) for benzene and 1.12(2) for 9-methyladenine. With the inclusion of $\kappa_H = 1.08(1)$ presently obtained for cytosine monohydrate and 1.40(3) obtained for glycylglycine (Coppens *et al.*, 1979), there is a range for κ_H which exceeds the estimated error and which is difficult to rationalize on a chemical basis. These estimates also exceed the range 1.16–1.31 reported in the theoretical study by Hehre, Stewart & Pople (1969). It is believed that κ_H is not reliably determined from structure refinements. It must be remembered that in the thermally-averaged total electron density, which is all that can be observed experimentally, the hydrogen pseudoatom makes a very small contribution which is not spatially resolved from the much larger contribution by its covalently bonded partner. Thus, κ_H will be easily

* There is a typographical error in Table 2 of Stewart (1991). The value $\alpha = 5.946$ Å^{-1} should be 5.546 Å^{-1} (R. F. Stewart, private communication).

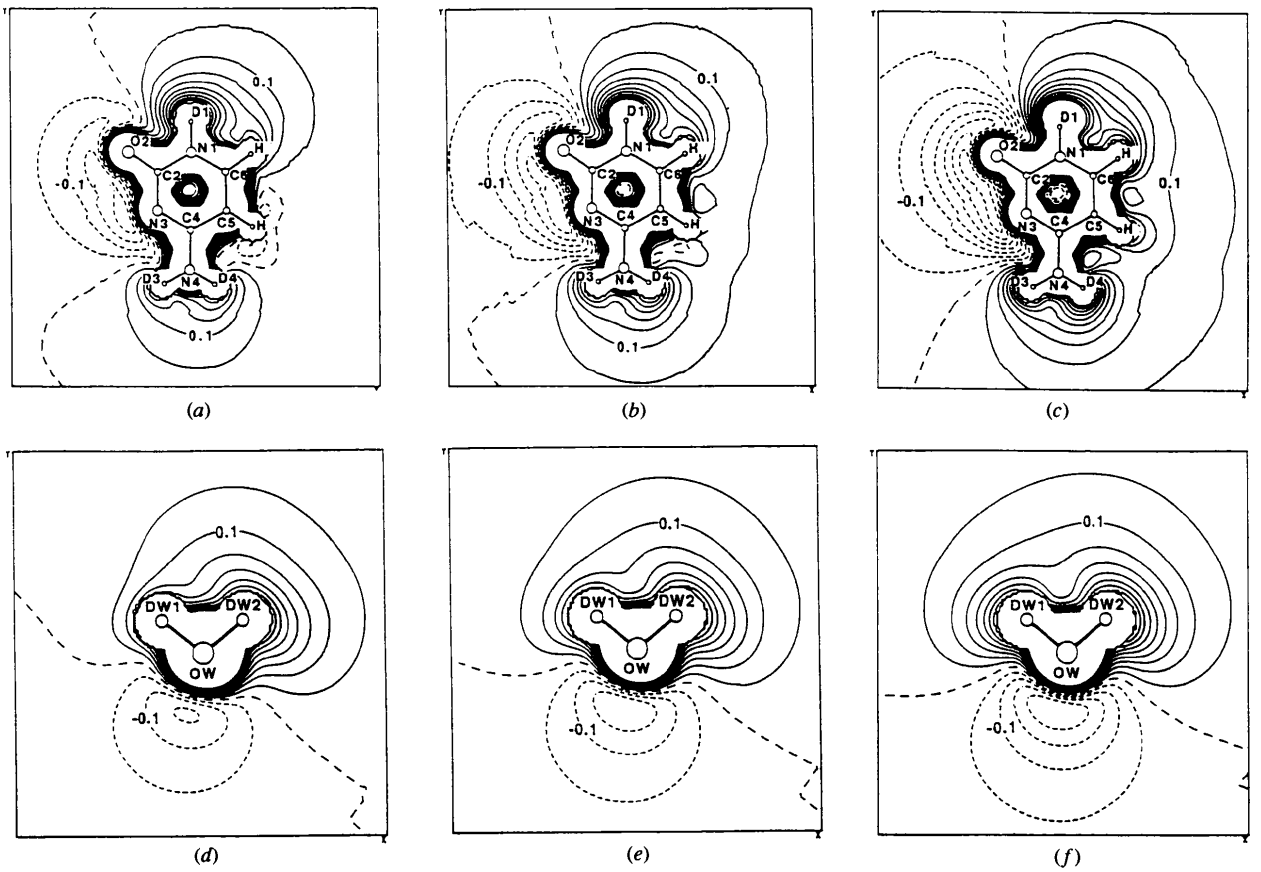


Fig. 7. Total electrostatic potential for isolated molecules from the crystal structure of cytosine monohydrate. The potential is derived from the full multipole model with three different κ parameters and is contoured as in Fig. 3. The three maps for each molecule show the effect of changes in the radial density function assumed for the hydrogen pseudoatoms. The radial density function for the hydrogen monopole has the form $r^2 \exp(-\kappa\alpha r)$ with $\alpha = 3.78$ Å^{-1}. For each molecule, the map at the left is for $\kappa = 1.08$ (H atom most expanded), in the center for $\kappa = 1.24$ and at the right for $\kappa = 1.40$ (H atom most contracted).

biased by systematic errors in the experimental data and will also be affected by how reliably the charge distribution is deconvoluted from the thermal vibrations of the H atom. Following Hehre, Stewart & Pople (1969), we prefer to assume a fixed value of $\kappa_H = 1.24$, which is close to the weighted average (1.25) of the above experimental values.

As will be discussed below, a more contracted H atom may be advisable when the least-squares structure refinement is carried out with the simple spherical pseudoatom model and in the absence of neutron diffraction data. For cytosine monohydrate, the spherical atom model with $\kappa_H = 1.40$ gives reasonably good agreement with the results from the full multipole model with $\kappa_H = 1.24$. However, further studies with additional structures are required in order to determine an optimum value to be used for κ_H in spherical atom refinements.

Comparison of the riboside and arabinoside of cytosine

The crystal structure of β-D-arabinosyl-cytosine (Tougard & Lefebvre-Soubeyran, 1974) is quite different from that of cytidine, although the molecules differ only in the configuration at C2′. The torsion angle at the glycosidic bond is similar in both structures but the sugar conformation is different. With the furanose ring in the C2′-*endo* configuration, the hydroxyl group O2′ in the arabinoside forms an intramolecular hydrogen bond O2′—H···O5′ and this requires a twist of C5′—O5′ about the bond C4′—C5′ from its position in cytidine. The arabinoside thus has one less H atom available for intermolecular hydrogen bonding and as a consequence, the cytosine ring nitrogen N3 is not hydrogen bonded. The twist of C5′—O5′ also removes the possibility of the short intramolecular interaction C5—H5···O5′ which occurs in cytidine.

It is of interest to compare the molecular electrostatic potential of the arabinoside which we have mapped using Pearlman & Kim's charge parameters (Fig. 8) with that of cytidine (Fig. 3). It is emphasized that the model used by Pearlman & Kim (1990) in their study of the arabinoside involved a best fit of charged spherical pseudoatoms to the observed density, whereas our fit with cytidine involved aspherical pseudoatoms with multipole terms up to octapole in order to obtain a more detailed description of chemical bonding effects. Nevertheless, since the predominant contribution to the electrostatic potential comes from the spherical component of the pseudoatom charge distribution, it is expected that the potential obtained with the simpler spherical pseudoatom model should be a good approximation at distances greater than the molecular van der Waals envelope. For γ-aminobutyric acid, Stewart & Craven (1993) found that the contributions of the higher multipoles to the total electrostatic potential were indeed small at such distances. However, they also found that when the higher multipoles were omitted from the structure refinement model, as in the studies by Pearlman & Kim (1990), the resulting spherical-atom population parameters were not so effective in deriving the potential as the monopole populations from a full multipole refinement.

For the arabinoside (Fig. 8*a*), the electrostatic potential in the electropositive region consisting of the C4 amino group, C5 and C6 is very similar to that obtained for cytosine (Fig. 3*a*). Outside the contour at $-0.10\,e\,Å^{-1}$, the electronegative potential spanning N3 and O2 is also very similar. However, at short range the minimum in electronegativity is deeper for the arabinoside than in cytidine or cytosine in the monohydrate, except when the

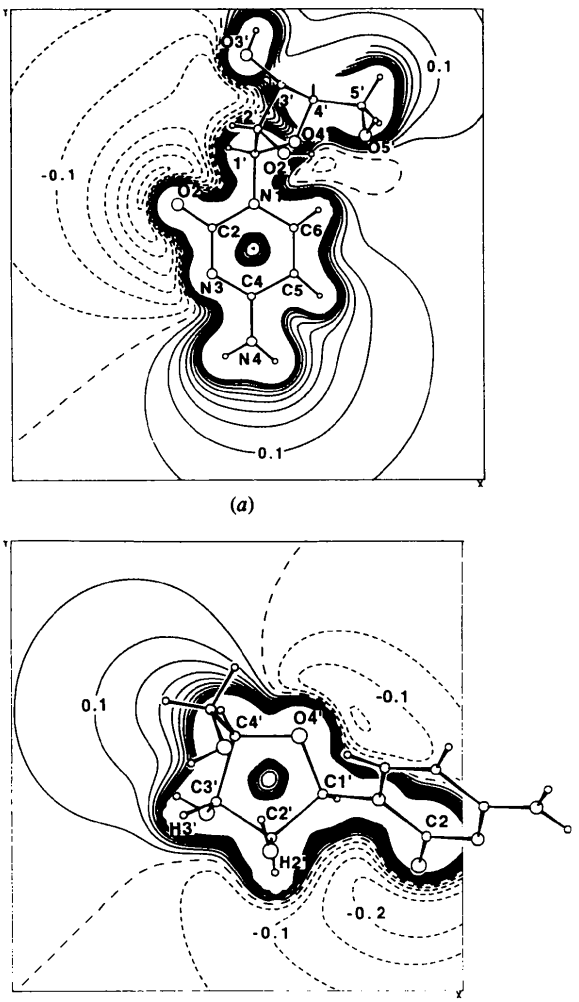


Fig. 8. Total electrostatic potential for an isolated molecule of 1-β-D-arabinosylcytosine. The potential is derived from the spherical atom refinement with population parameters from Pearlman & Kim (1990) and with atomic positional parameters from Toughard & Lefebvre-Soubeyran (1974). Contours are as in Fig. 3. (*a*) Section in the plane of the cytosine base; (*b*) section in the best plane through atoms of the arabinose ring. Note that O2′ is an equatorial substituent, whereas in ribose (Figs. 1 and 3*b*) O2′ is axial.

cytosine potential is derived assuming $\kappa_H = 1.40$ (Fig. 7). Possibly the agreement in this case is a consequence of Pearlman & Kim's assumption that $\kappa_H = 1.40$. As noted above, the assumption of a more contracted hydrogen may be appropriate when refinement is carried out with spherical pseudoatoms. The potential near the top in Figs. 3(*a*) and 8(*a*) is different, as expected, because of the differences in sugar conformations.

Molecular dipole moments

The molecular dipole moment for cytidine nucleoside calculated from the full multipole refinement (that is, from the monopole and dipole pseudoatom population parameters; Stewart, 1972) is 17 (4) Debye (see Table 8). This agrees with the value of 15 (4) Debye derived from the spherical atom refinement with $\kappa_H = 1.40$. For the cytosine base together with the C1′ substituent, the dipole moment from the full multipole model is 11 (2) Debye. For cytosine from cytosine monohydrate, the new full multipole refinement gives a dipole moment of 8.2 (15) Debye, in agreement with the previously reported value [8.0 (14) Debye; Spackman, Weber & Craven, 1988] and with theoretical values (Spackman, 1992). The presently reported dipole moments make angles of 4 (6) and 12 (13)° with the C2—C5 direction for the cytidine nucleoside and the cytosine base, respectively. For the arabinoside, Pearlman & Kim (1990) report a dipole moment of 15.6 Debye.

Table 8. *Molecular dipole moments* (Debye)

Columns are:

(*a*) Cytidine. Dipole moment derived from the monopole and dipole populations parameters.

(*b*) Cytidine. Values derived from the spherical atom refinement assuming $\kappa_H = 1.40$.

(*c*) Cytosine from the monohydrate. Values derived from the new full multipole refinement assuming $\kappa_H = 1.24$.

(*d*) Cytosine from the monohydrate. Values derived from the spherical atom refinement assuming $\kappa_H = 1.40$.

(*e*) 1-β-D-Arabinosylcytosine. Values derived from the spherical atom refinement assuming $\kappa_H = 1.40$ (Pearlman & Kim, 1990).

(*f*) Cytosine. Values derived from semiempirical quantum chemical methods (Tasi, Pálinkó, Nyerges, Fejes & Förster, 1993).

	(*a*)	(*b*)	(*c*)	(*d*)	(*e*)	(*f*)
Cytidine (nucleoside)	17.2 (41)	15.3 (44)	—	—	15.6	—
Cytosine (including C1′ or D1)	11.1 (21)	11.0 (21)	8.2 (15)	7.6 (15)	14.1	7.03

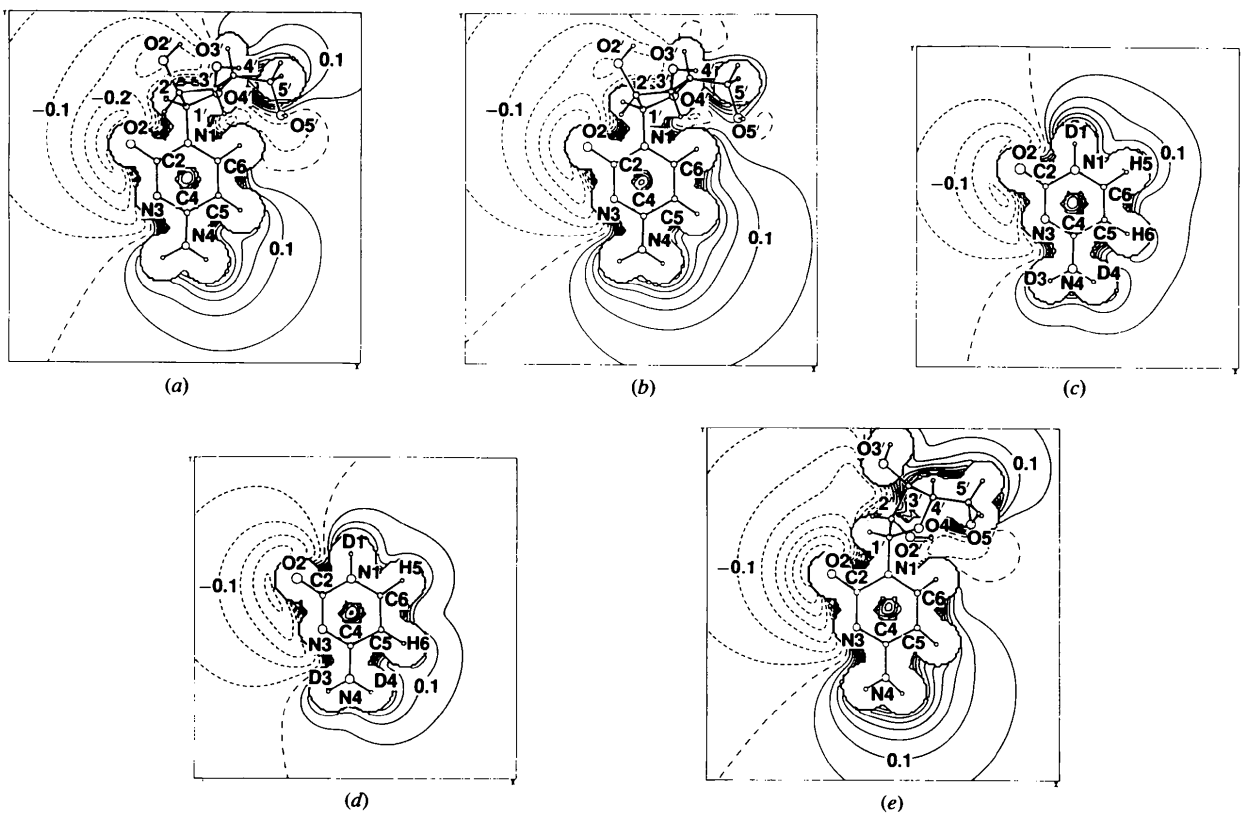


Fig. 9. Maps of the molecular electrostatic potentials derived according to the simplified procedure of Stewart & Craven (1993). They are based on the systems of point charges from the columns in Table 7. Contours are as in Fig. 3. (*a*) Cytidine, point charges from Table 7*a*; (*b*) cytidine, point charges from Table 7*b*; (*c*) cytosine from cytosine monohydrate, point charges from Table 7*c*; (*d*) cytosine from cytosine monohydrate, point charges from Table 7*d*; (*e*) 1-β-D-arabinosylcytosine, point charges from Table 7*e*.

Atomic point charges

It is convenient for computations involving large molecules such as polynucleotides to represent the charge distribution as a system of atomic point charges such as those given in Table 7. These point charges are obtained simply as Z_v–p_v, where Z_v is the number of valence-shell electrons in a neutral pseudoatom and p_v is the experimental value for the pseudoatom monopole population parameter. Before adopting this approach, it must be assured that the atomic point charges are efficient in representing the electrostatic properties of the molecule and also that the point charges are transferable among closely related chemical groups.

Allowing for chemical differences at N1, there is good agreement between the point charges obtained from the full multipole refinements for cytosine in cytidine (Table 7*a*) and in the monohydrate (Table 7*c*). Similarly, the point charges obtained from the spherical atom refinements agree with each other (Table 7*b* and *d*) and with the point charges for cytosine in the arabinoside (Table 7*e*). There are some significant differences in point charges obtained with the two different structure models, but this is to expected since the refinements lead to different atomic positional and thermal parameters, particularly for the H atoms. The important criterion is whether there is agreement in the molecular properties obtained from the two models.

In Fig. 9, the molecular electrostatic potential is mapped assuming the simplified model of Stewart & Craven (1993), which makes use of atomic point charges together with Gaussian terms semiempirically introduced for better modeling of the potential close to the van der Waals envelope. It can be seen that maps derived using point charges from the monopole terms of the full multipole refinement are very similar to those obtained from the charge distribution of the full pseudoatom model [Fig. 9*a versus* Fig. 3*a*; Fig. 9*c versus* Fig. 7 (top center)]. Potential maps with point charges derived from spherical pseudoatom refinements with $\kappa_H = 1.40$ are shown in Figs. 9(*b*) and (*d*) for cytidine and cytosine monohydrate, respectively. These maps also show general agreement with the potential from the full multipole model. The major difference for cytidine is near the sugar residue where there are significant differences in atomic charges for H4′, H5′1 and H5′2 atoms, which changed from being electropositive to slightly electronegative. These effects may be due to the relaxation of hydrogen positional and displacement parameters from the adjusted values given in Table 3. The molecular electrostatic potential map for 1-β-D-arabinosyl-cytosine derived from point charges (Fig. 9*e*) is very similar to that obtained from the spherical pseudoatom model (Fig. 8*a*).

We conclude that a simplified model involving experimentally determined point charges holds considerable promise for deriving electrostatic properties for cytosine and cytidine when these molecules are extracted from their crystal environment and incorporated in other systems.

This work was supported by a grant GM-39513 from the National Institutes of Health. We are grateful to Professor Robert F. Stewart for discussions and to Dr John Ruble and Mrs Joan Klinger for technical assistance.

References

BARKER, D. L. & MARSH, R. E. (1964). *Acta Cryst.* **17**, 1581–1587.

BLESSING, R. H. (1984). *Computer Programs for Profile Data Procession*. Medical Foundation of Buffalo, New York.

CLEMENTI, E. & ROETTI, C. (1974). *At. Data Nucl. Data Tables*, **14**, 177–478.

COPPENS, P., GURU ROW, T. N., LEUNG, P., STEVENS, E. D., BECKER, P. J. & YANG, Y. W. (1979). *Acta Cryst.* A**35**, 63–72.

CRAVEN, B. M., WEBER, H.-P., HE, X. & KLOOSTER, W. T. (1993). *The POP Procedure: Computer Programs to Derive Electrostatic Properties From Bragg Reflections*. Technical Report TR-93-1. Crystallography Department, Univ. of Pittsburgh, Pittsburgh, PA.

EISENSTEIN, M. (1988). *Acta Cryst.* B**44**, 412–426.

EPSTEIN, J., RUBLE, J. R. & CRAVEN, B. M. (1982). *Acta Cryst.* B**38**, 140–149.

FURBERG, S. (1950). *Acta Cryst.* **3**, 325–335.

FURBERG, S., PETERSEN, C. S. & ROMMING, C. (1965). *Acta Cryst.* **18**, 313–320.

HE, X. & CRAVEN, B. M. (1993). *Acta Cryst.* A**49**, 10–22.

HEHRE, W. J., STEWART, R. F. & POPLE, J. A. (1969). *J. Chem. Phys.* **51**, 2657–2664.

JEFFREY, G. A. & KINOSHITA, Y. (1963). *Acta Cryst.* **16**, 20–28.

JEFFREY, G. A. & SAENGER, W. (1991). *Hydrogen Bonding in Biological Structures*. Berlin: Springer-Verlag.

JOHNSON, C. K. (1976). *ORTEP*II. Report ORNL-5138. Oak Ridge National Laboratory, Tennessee, USA.

KLOOSTER, W. T., RUBLE, J. R., CRAVEN, B. M. & MCMULLAN, R. K. (1991). *Acta Cryst.* B**47**, 376–383.

MCCLURE, R. J. & CRAVEN, B. M. (1973). *Acta Cryst.* B**29**, 1234–1238.

PEARLMAN, D. A. & KIM, S.-H. (1985). *Biopolymers*, **24**, 327–357.

PEARLMAN, D. A. & KIM, S.-H. (1990). *J. Mol. Biol.* **211**, 171–187.

POST, M. L., BIRNBAUM, G. I., HUBER, C. P. & SHUGAR, D. (1977). *Biochim. Biophys. Acta*, **479**, 133–142.

ROSENBERG, J. M., SEEMAN, N. C., DAY, R. O. & RICH, A. (1976). *J. Mol. Biol.* **104**, 145–167.

SPACKMAN, M. A. (1992). *Chem. Rev.* **92**, 1769–1797.

SPACKMAN, M. A., WEBER, H.-P. & CRAVEN, B. M. (1988). *J. Am. Chem. Soc.* **110**, 775–782.

STEWART, R. F. (1972). *J. Chem. Phys.* **57**, 1664–1668.

STEWART, R. F. (1976). *Acta Cryst.* A**32**, 565–574.

STEWART, R. F. (1991). In *The Application of Charge Density Research to Chemistry and Drug Design*, edited by G. A. JEFFREY & J. F. PINIELLA, pp. 63–110. New York: Plenum Publishing Corporation.

STEWART, R. F. & CRAVEN, B. M. (1993). *J. Biophys.* **65**, 998–1005.

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

SWAMINATHAN, S. & CRAVEN, B. M. (1984). *Acta Cryst.* B**40**, 511–518.

TASI, G., PÁLINKÓ, I., NYERGES, L., FEJES, P. & FÖRSTER, H. (1993). *J. Chem. Inf. Comput. Sci.* **33**, 296–299.

TAYLOR, R. & KENNARD, O. (1982). *J. Am. Chem. Soc.* **104**, 5063–5070.

TEMPLETON, L. K. & TEMPLETON, D. H. (1973). *Abstr. Am. Cryst. Assoc. Meeting*, p. 143. Storrs, Connecticut.

TOUGARD, P. & LEFEBVRE-SOUBEYRAN, O. (1974). *Acta Cryst.* B**30**, 86–89.
WARD, D. L. (1993). *Acta Cryst.* C**49**, 1789–1792.
WEBER, H.-P. & CRAVEN, B. M. (1990). *Acta Cryst.* B**46**, 532–538.
WEBER, H.-P., CRAVEN, B. M. & MCMULLAN, R. K. (1980). *Acta Cryst.* B**36**, 645–649.

SHORT COMMUNICATIONS

Acta Cryst. (1995). B**51**, 1097–1102

Structure of 2,2,3,3-tetranitrobutane

BY L. L. KOH, K. Y. SIM, H. H. HUANG,* Y. L. LAM and E. P. LIANG, *Department of Chemistry, National University of Singapore, Lower Kent Ridge Road, Singapore* 0511, *Singapore*

(*Received* 23 *December* 1994; *accepted* 20 *July* 1995)

Abstract

In two crystalline polymorphs, 2,2,3,3-tetranitrobutane adopts a conformation in which the two methyl groups are *gauche* to each other and the nitro groups are staggered to minimize steric interactions in the crystal lattice.

Introduction

From dipole moment data, supported by IR and Raman spectroscopic measurements, it was concluded that 2,2,3,3-tetranitrobutane exists as a mixture of 66% of the *trans* and 34% of the *gauche* rotamer in carbon tetrachloride solution at 298 K (Tan, Chia & Huang, 1986). It was also suggested that some of the observed weak spectral bands of the solid could be attributed to the presence of the *trans* rotamer in the solid state in addition to the *gauche* form. The suggested presence of the *trans* rotamer was based mainly on the fact that no additional bands were seen on dissolving the solid in organic solvents. Furthermore, three bands were observed in the solid-state IR spectra at 847 (weak) 860 (strong) and 877 (weak) cm^{-1} in the region where C—N stretching vibrations are expected (Lin-Vien, Colthup, Fately & Graselli; 1991). In his IR study of solid 2,2,3,3-tetranitrobutane, Diallo (1974) assumed that only two IR-active bands due to the C—N stretching modes of the *trans*-rotamer would appear in the 800–900 cm^{-1} region. Since three bands attributable to the C—N stretching vibrations instead of two were observed, Diallo concluded that 2,2,3,3-tetranitrobutane exists as a mixture of *gauche* and *trans* conformers in the solid state, the additional band being due to the *gauche* conformer. However, our X-ray crystal diffraction study of two different crystalline forms of the compound show that in both cases, 2,2,3,3-tetranitrobutane has a structure in which the two methyl groups are *gauche*. Although this does not rule out the possibility of other existing forms of the solid which may contain *trans* rotamers, it appears that the *gauche* rotamer is the preferred conformer in the crystalline state. Consequently, the earlier inference that both the *gauche* and *trans* rotamers are present in the solid state needs to be reviewed. On the assumption that the crystals chosen for our X-ray studies are representative of the bulk material, all the IR and Raman bands in the solid-state spectra must therefore be assigned to vibrations of the *gauche* rotamer. We have also re-examined the IR and Raman spectra of the dry compound taken from the same batch of carefully dried crystals that was used for our X-ray study and measured some depolarization ratios, paying particular attention to the 800–900 cm^{-1} region. The measurements were carried out using the JASCO NR-1800 Laser Raman spectrophotometer (resolution ± 0.15 cm^{-1}) and the Perkin–Elmer 1725X spectrophotometer (resolution ± 1 cm^{-1}). We believe the spectra thus obtained are more accurate than our earlier measurements in terms of observed frequencies. We have also remeasured the IR spectra in solution and in the vapour phase up to a temperature of 433 K to investigate the possibility of occluded solvent. However, the new IR spectra still remained the same, exhibiting the same three bands with practically the same relative intensities at the frequencies observed earlier. Furthermore, the new Raman spectra also show that the three bands of interest occur at frequencies of 845 (weak), 862 (weak) and 874 (strong) cm^{-1}. Although these bands are slightly different from the earlier Raman observations, they are now virtually identical to the IR frequencies; all three are clearly depolarized. Thus,. we can now see that these results are incompatible with the existence of the *trans* (C_{2h}) rotamer in the solid state for two reasons. (1) The three Raman shifts clearly coincide with the three IR bands observed. This correspondence is incompatible with the *trans* (C_{2h}) conformation, but is expected for the *gauche* (C_2) rotamer. (2) The fact that all three Raman lines are depolarized and the same lines are IR active suggest that they belong to vibrations of the *B*-type of *gauche* molecules. This conclusion is based on the selection rules shown in Table 1. Here we are assuming that the *gauche* molecules have strictly C_2 symmetry, although the X-ray structure shows that the *gauche* molecules have only approximate C_2 symmetry as the nitro groups are staggered in such a way as to lower the symmetry of the molecule as a whole.

Nevertheless, we feel that these rules provide a useful guide for drawing some general conclusions. We have also reconsidered the number of bands expected for the C—N stretching vibrations. A careful analysis shows that in theory, there can be four types of such vibrations for the *gauche* C_2 molecule and also four for the *trans* C_{2h} molecule. In fact a vibrational frequency computation based on SAM1 parametrization gives the following results in cm^{-1}.

trans	ν_{CN}(sym)	A_g 790.7	B_u 839.4
conformer	ν_{CN}(asym)	A_u 770.7	B_g 820.3
gauche	ν_{CN}(sym)	*A* 727.4	*B* 819.9
conformer	ν_{CN}(asym)	*A* 779.6	*B* 795.4

* Author to whom correspondence should be addressed.