Acta Cryst. (1994). B50, 161–174

Electron-Density and Electrostatic Potential Distributions in Nucleoside Analogs. 1. An Experimental Study of 3′-*O*-Acetyl-2′-deoxy-5-methoxymethyluridine

BY YUNYI WEI,* RICHARD BARTON AND BEVERLY ROBERTSON†

Department of Physics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2

(*Received* 4 *March* 1992; *accepted* 28 *October* 1993)

Abstract

The experimental electron-density distribution and the distribution of electrostatic potential for 3′-*O*-acetyl-2′-deoxy-5-methyoxymethyluridine have been determined from high-resolution X-ray diffraction data collected at 145 (4) K using graphite-monochromated Mo $K\alpha$ radiation to $\sin\theta/\lambda = 1.08$ Å^{-1}. Crystal data: $C_{13}H_{18}N_2O_7$, $M_r = 314.297$, triclinic, $P\bar{1}$, $a = 5.9701$ (7), $b = 6.7357$ (6), $c = 9.5770$ (10) Å, $\alpha = 77.21$ (1), $\beta = 79.95$ (1), $\gamma = 75.45$ (1)°, $V = 360.63$ (6) Å^3, $Z = 1$, $F(000) = 166$, $D_x = 1.447$, $D_m = 1.43$ (2) Mg m^{-3}, $\mu = 0.1278$ mm^{-1}, $\lambda = 0.71069$ Å. The integrated intensities of 15 087 reflections were measured and reduced to 7067 independent reflections with $I \geq 3\sigma(I)$. The partial charges on the atoms of the molecule were determined and are in reasonable agreement with *ab initio* values reported for fragments of the molecule. The acetal bond, O′(4)—C′(1), shows higher electron density than the O′(4)—C′(4) bond does, which is more typical of an ether bond. The presence within a molecule of methyl H atoms in the vicinity of an O atom that might serve as the acceptor of a hydrogen bond appears to significantly perturb the electrostatic potential around the O atom and make it less acceptable for hydrogen bonding. The atoms N(3), O(52) and O(4) combine to create a molecular surface with negative electrostatic potential that extends over part of one face of the 5-methoxymethyluridine ring.

1. Introduction

1.1. *The mathematical model*

The methodology of charge-density crystallography has evolved over the past two decades from the study of simple organic molecules to the study of complexes of transition metals and organic molecules of sufficient size and complexity that they are capable of showing interesting biological properties (Klein, Majeste & Stevens, 1987; Klein, Stevens, Zacharias & Glusker, 1987; Klein & Stevens, 1988*a*; Souhassou *et al.*, 1991). The determination of the accurate distribution of the electron density of bio-organic molecules requires that the integrated intensities of reflections be measured with very high precision and accuracy because the features of interest in the experimental electron density may involve variations in the electron density of a similar order of magnitude as the effects of random errors and possible systematic errors.

Systematic error in the results of electron-density determinations may be introduced by errors in the description of the sample crystal, in the measurement of the integrated intensities, in the method of reduction of integrated intensities to the observed structure amplitudes, and in the structural model used to fit the observed electron density by least-squares minimization of the differences between the observed and calculated structure amplitudes.

The structural model proposed by Stewart (1976), and modified by Hansen & Coppens (1978), conveniently expresses the electron density of atoms as an expansion in terms of multipoles,

$$\rho(\mathbf{r}) = \rho_c(r) + \kappa^3 P_v \rho_v(\kappa r) + \kappa'^3 \sum_l R_l(\kappa' r) \times \sum_{m=-l}^{+l} P_{lm} Y_{lm}(\theta, \vartheta), \quad (1)$$

where P_v and P_{lm} are refinable population parameters, κ and κ' are refinable parameters describing the contraction/expansion of the spherical and non-spherical components of the valence electron density, respectively, and ρ_c and ρ_v are the core electron density and the spherical part of the valence electron density. The radial functions $R_l(\kappa' r)$ may be Slater-type functions (Hansen & Coppens, 1978) or numerically defined functions based on a Hartree–Fock calculation of a spherically averaged valence density.

Traditional use of the results of X-ray diffraction in the study of molecular recognition usually involves matching the geometry of the substrate to that of an active site of the macromolecule. The geometry of the molecule in the crystal is assumed to be similar to that in solution. The usefulness of this approach is limited because the geometry of the

* Present address: Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada T6G 2H7.

† To whom correspondence should be addressed.

substrate is often altered from that in solution before it encounters the active site. Souhassou *et al.* (1991) have discussed further the reasons that the distribution of electrostatic potential of a molecule is important in the early stages of molecular recognition, and why an accurate knowledge of the full-charge-density distribution is needed in order to calculate the electrostatic potential distribution.

In general, a property of the electron distribution can be calculated as a Fourier summation of the structure factors with each term weighted by the Fourier transform of the operator corresponding to that property (Spackman & Stewart, 1981; Klein & Stevens, 1988*b*). In particular, the distribution of electrostatic potential for a molecule in the crystal may be calculated from

$$\vartheta(\mathbf{r}) = \frac{1}{V}\sum(F_{\mathbf{H}}'/d_{\mathbf{H}}^{*2}) \exp(-2\pi i \mathbf{d}_{\mathbf{H}}^{*}.\mathbf{r}), \qquad (2)$$

where $\mathbf{d}_{\mathbf{H}}^{*}$ is the conventional reciprocal lattice vector and $F_{\mathbf{H}}'$ is a calculated structure factor based on a multipole model modified to include the nuclear charge.

$\vartheta(\mathbf{r})$ calculated in this way is affected by series termination, although the $d_{\mathbf{H}}^{*-2}$ factor substantially lowers the effect of missing terms. Of greater significance is the fact that the direct contributions of the charges on neighboring molecules in the crystal are included in the calculation of the electrostatic potential. In a biological system, the neighboring molecules would not be the same species of molecule as the sample molecule, but would instead be mostly water molecules.

Alternatively, the distribution of electrostatic potential can be calculated directly from

$$\vartheta(\mathbf{r}) = \sum_i \frac{Z_i}{|\mathbf{r}_i - \mathbf{r}|} - \int \frac{\rho(\mathbf{r}')\mathrm{d}\mathbf{r}'}{|\mathbf{r}' - \mathbf{r}|}, \qquad (3)$$

where the first term represents the contribution from nuclear charges and the second terms represents the contribution from the electron density. The sum and integration normally include the atoms of only one molecule and, therefore, does not directly include the effects of the charge distribution of neighboring molecules. Nevertheless, the charge distribution of neighboring molecules may influence the charge distribution within the volume of integration and thereby indirectly alter the electrostatic potential of a molecule in a crystal from that of a free molecule or of a molecule in aqueous solution.

1.2. *Nucleoside analogs*

In order to investigate the reliability and value of experimental electron-density and electrostatic potential results as applied to biologically significant organic molecules, it is useful to study a series of similar compounds to test for the consistency of the results. Therefore, we report here the first in a series of studies of the electron-density and electrostatic potential distributions of a series of nucleoside analogs.

Studies of the electron-density distributions of smaller organic molecules have been used in combination with theoretical and computational results to investigate the fundamental character of atomic interactions within molecules. The present program of research is intended to provide similar information, such as the partial charges on the atoms forming important biological functions including the ribose group. The longer term goal of this research will be a better understanding of the interactions between molecules and the specificity of molecular recognition.

Many nucleoside analogs have been shown to have biological activity against viral diseases. They mimic cellular nucleosides and they (or their phosphorylated equivalents) may show significantly different specificity for virus-induced enzymes relative to cellular enzymes. Target enzymes for the use of nucleoside analogs as chemotherapy agents vary from reverse transcriptase, in the case of the human immunodeficiency virus (HIV), to the kinases and DNA polymerase, in the case of the herpes simplex viruses (HSV).

Several nucleoside analogs have been studied by Gupta and coworkers as potential chemotherapy agents against HSV (Gupta *et al.*, 1987). In particular, 2′-deoxy-5-methoxymethyluridine (MMdUrd) is a selective agent against the herpes virus (Ayisi, Gupta & Babiuk, 1985). The compound 3′-*O*-acetyl-2′-deoxy-5-methoxymethyluridine (MMAcdUrd) was synthesized to test the hypothesis that the acetyl derivative of MMdUrd might show increased efficacy as a consequence of its potentially greater lipophilicity (Quail, Tourigny, Delbaere, El-Kabbani, Stuart & Gupta, 1988).

We have undertaken a systematic study of the electron densities and electrostatic potentials of a series of analogs of thymidine because of their fundamental biological significance and their potential use as chemotherapy agents. Because of their relative simplicity, they are also a tractable system for study. We report here the experimentally determined electron-density distribution and electrostatic potential of MMAcdUrd.

2. Experimental

2.1. *Data*

The synthesis and the conventionally determined structure of MMAcdUrd have been reported previously (Stuart, Ayisi, Tourigny & Gupta, 1985;

Quail *et al.*, 1988). The atom-numbering scheme used here is that which is conventionally used for nucleoside analogs and is as shown in Fig. 1. A sample of the compound was provided by G. Tourigny, University of Saskatchewan. It was recrystallized by slow evaporation over a period of 2 weeks from a mixed solvent of acetone and ethyl acetate. A crystal with the dimensions $0.52 \times 0.25 \times 0.36$ mm was chosen for X-ray data collection. The sample crystal was mounted on a modified Picker four-circle diffractometer controlled by the NRCC data-collection system (Gabe, Larson, Lee & Wang, 1979) with the crystallographic *a* axis offset approximately 3° from being parallel to the φ-axis of the diffractometer. The beam spot size was roughly 1.5 mm in diameter. The distribution of the intensity of the X-ray beam was not measured.

The crystal was cooled using a modified Enraf–Nonius gas-flow system, controlled by a small computer. The temperature was monitored with a copper–constantan thermocouple mounted near the front of the gas-delivery nozzle. The difference between the temperature of the thermocouple used for monitoring the temperature, and the temperature of the sample crystal was measured by replacing the sample with a second copper–constantan thermocouple constructed from very thin thermocouple wire. The temperature difference was found to be 20 K. (The separation between the thermocouples was roughly 1.5 cm.) Data were collected with a sample temperature estimated to be 145 (4) K.

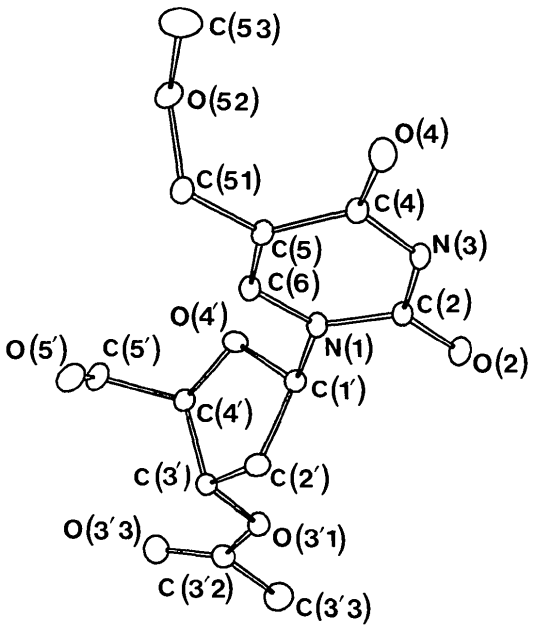


Fig. 1. View of the molecule showing atomic labeling. H atoms have been omitted for clarity.

During data collection the goniostat of the diffractometer was surrounded by a closed chamber filled with dry nitrogen.

The lattice constants at 145 K were determined from a least-squares fit to the observed setting angles of 50 reflections with $42 < 2\theta < 73°$ using the NRCC programs (Gabe *et al.*, 1979). The diffraction peaks corresponding to Mo $K\alpha_1$ radiation ($\lambda = 0.70926$ Å) were used for the determination of the setting angles. Intensity measurements were collected using a θ–2θ scan mode, a highly oriented graphite monochromator and the peak profile method as contained in the NRCC diffractometer control system (Gabe *et al.*, 1979). The scan width for each reflection was set at $(0.8 + 0.692\tan\theta + 0.8°)$. This expression reflects only the dispersion in the diffraction from the sample. Dispersion from diffraction by the monochromator is in a plane orthogonal to the plane of the sample dispersion.

The intensities of three standard reflections were measured after each 47 reflections. The peak profile of each reflection was analyzed to give its integrated intensity and the estimated standard deviation (e.s.d.) of the integrated intensity based on counting statistics. The data were rescaled to account for the decrease in the scattering power of the crystal as manifested in the 2% decrease in the intensities of the standard reflections over the period of the data collection. The scatter in the total counts used to measure the intensities of the standard reflections was analyzed to determine the extent to which it exceeded that predicted from counting statistics. It was assumed that the excess scatter was proportional to I^2 and that the intensities of the individual reflections contained a similar excess scatter, and their e.s.d.'s were modified accordingly.

The intensities of all reflections with $\sin\theta/\lambda \leq 1.08$ Å^{-1} were collected, giving a total of 15087 integrated intensities. The data were corrected for Lorentz and polarization effects using the *XTAL*2.4 system of crystallographic software (Stewart, Hall, Alden, Olthof-Hazekamp & Doherty, 1983). The data were also corrected for absorption using the Gaussian numerical integration method. The minimum and maximum corrections were 1.029 and 1.050, respectively.

The internal agreement index, $\sum|I - \langle I\rangle|/\sum I$, where $\langle I\rangle$ is averaged over Freidel equivalent reflections, was 0.014. The variance of the averaged intensity was set as $\sum\sigma^2(I)/N$, where N is the number of times a reflection was measured (normally twice), unless the quantity $\sum(I^2 - \langle I^2\rangle)^2/N(N-1)$ indicated a larger value. Using variances estimated in this way, $\sum\sigma(I)/I$ was calculated to be 0.021. This is larger than the internal agreement index. The available computing capacity was limited. Therefore, it was assumed that the differences between the intensities of Freidel

equivalent reflections were sufficiently small that their intensities could be averaged. The data were merged to give 7586 independent reflections. The scattering amplitudes and their e.s.d.'s were calculated from the intensities and their e.s.d.'s. Only those reflections for which $I > 3\sigma(I)$, of which there were 7067, were used for refinement.

2.2. *Refinement*

The starting coordinates for least-squares refinement were taken from Quail *et al.* (1988). A conventional refinement was first carried out with the *CRYSLQ* routine in the *XTAL* system (Stewart *et al.*, 1983), using the data with $\sin\theta/\lambda \leq 0.60$ Å^{-1} and scattering factors from Cromer & Mann (1968). The scattering factors included corrections for the real part of the anomalous dispersion for the non-H atoms. The refinement was based on $|F|$, with weights calculated as $w = \sigma^{-2}(F)$. The coordinates and anisotropic displacement parameters, U_{ij}, for the non-H atoms were refined next using only those reflections for which $\sin\theta/\lambda \geq 0.80$ Å^{-1}, for which the X-ray scattering is mainly from the core electrons. The details of this refinement are given as Ref. 1 in Table 1. The coordinates and isotropic displacement parameters of the H atoms were next refined using all the data. The H atoms were then moved along their bond vectors to a point which gave C—H, N—H and O—H distances equal to those for similar bonds in other compounds, as determined by neutron diffraction (Allen *et al.*, 1987).

The structure was then refined using data with $\sin\theta/\lambda \geq 0.80$ Å^{-1} and the program *POP*, written by Craven, Weber & He (1987). The positions and thermal parameters of the H atoms were constrained, but the refinement allowed variation of third- and fourth-order displacement parameters, C_{ijk} and D_{ijkl}. None of the C_{ijk} or D_{ijkl} refined to values different from zero by more than three standard deviations and it was assumed that all such terms could be set to zero. They were not varied in subsequent refinements.

2.3. *Partial charges*

The least-squares program *AMANDA* (Stevens, 1990) was used to determine the partial charges on the atoms of the molecule. The model used in the refinement was based on Stewart's (1976) rigid pseudo-atom model, with the addition of expansion/contraction parameters κ and κ', as defined in (1). The functions describing the radial dependence of the non-spherical multipole terms are significantly different from zero at the nuclei of the atoms bonded to the atom which serves as the origin of the radial function. Electron density and its charge may then be incorrectly assigned to a neighboring atom. Therefore, in the refinements used to determine the partial charges, only P_v and κ were varied. All P_{lm} values were set at zero. The κ values of similar atoms were constrained to be equal. The scattering of the core and valence electron density was calculated from the orbital scattering factors given by Cromer & Waber (1974), with the valence scattering power normalized to one electron. The full data set was used with weights calculated as $w = \sigma^{-2}$. The adjustment of the scale factor was based on constraint of the total charge on the molecule to neutrality, rather than a fit of $|F_c|$ to $|F_o|$, as in a conventional least-squares refinement of a crystal structure. The partial charges were then calculated as $q_i = -(P_{vi} - N_{vi})$, where N_{vi} is the number of valence electrons in a neutral atom. The details of the refinement are given in Table 1 as Ref. 2. The values of P_{vi}, q_i and values from *ab initio* calculations of Singh & Kollman (1984) are listed in Table 2.

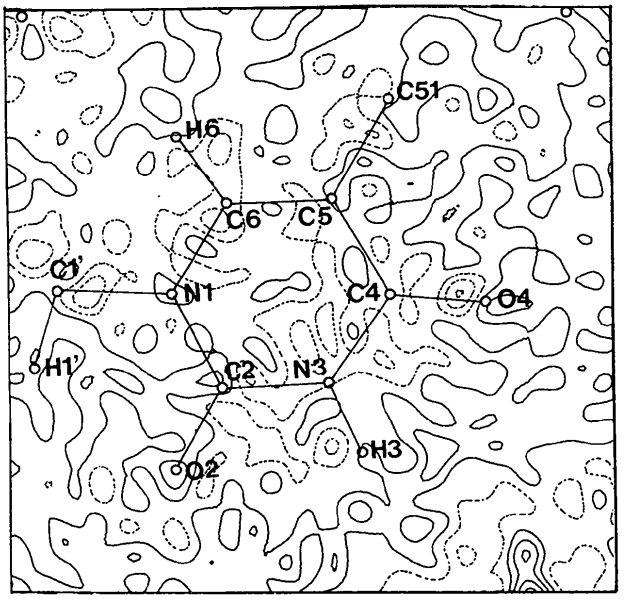
2.4. *Electron-density distributions*

A number of choices are available for the implementation of the multipole model proposed by Stewart (1976). These have been reviewed by Lecomte (1991). In particular, the program *POP* (Craven *et al.*, 1987) assumes that $\kappa = \kappa' = 1$. The multipole model is rotationally invariant if no constraints are placed on P_{lm}, which permits a common local coordinate system to be used for all atoms in the program *POP*. This facilitates the calculation of the distribution of electrostatic potential directly from P_{lm}, but the absence of constraints based on assumed local molecular symmetry leads to a proliferation of least-squares variables as l increases and a consequent decrease in the ratio of the number of observations to the number of variables. The procedure has the obvious advantage that the absence of assumed constraints means an absence of possible bias associated with those constraints.

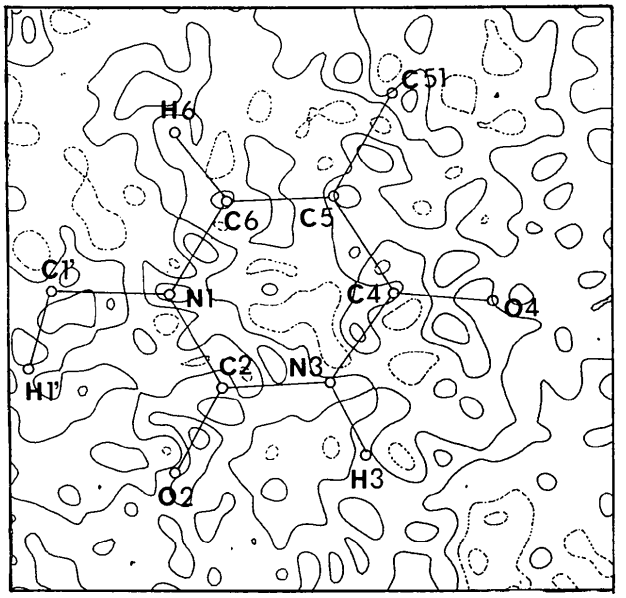
The radial functions $R_l(r)$ are chosen as $N_l r^{n_l}\exp(-\xi r)$, where n_l and ξ may be set at values such as those given by Hehre, Stewart & Pople (1969) and N_l is a normalizing factor. The exponent ξ may also be refined. The quantity $\rho_c(\mathbf{r})$ may be used to represent only the core electrons, as discussed earlier, or it may represent the free-atom density, in which case P_v is the number of excess electrons on the atom.

In the version of the *POP* program used in these refinements, the condition of charge neutrality for the molecule as a whole is achieved by adjustment by the user of the scale factor multiplying the observed structure amplitude, using a trial and error procedure. The program *AMANDA* (Stevens, 1990) is a variation of the program *MOLLY*, written by

Hansen & Coppens (1978). It permits the use of constraints based on assumed local symmetry to limit the number of least-squares variables, in which case care must be exercised to avoid imposing false local symmetry on individual atoms. A normal difference map, based on the multipole model, may be used to reveal any errors in the structural model,



(*a*)



(*b*)

Fig. 2. A comparison of the residual maps in the plane of the base ring as obtained from the refinements using (*a*) *AMANDA* and (*b*) *POP*. Contour intervals are at 0.05 e Å $^{-3}$. Positive contours are solid and negative contours are dashed. The zero contour is omitted.

Table 1. *Least-squares refinements*

See text for an explanation of each refinement.

	Ref. 1 *XTAL2.4* (*CRYLSQ*)	Ref. 2 *AMANDA* (κ ref.)	Ref. 3 *AMANDA* (multipole)	Ref. 4 *POP* (multipole)
Variables, non-H atoms	r_j, U_{ij}	κ_j, P_{vj}	κ_j, P_{vj}, κ'_j, r_j, U_{ij}, to hexadecapole	ΔP_{vj}, r_j, U_{ij}, to octapole
Variables, H-atoms		κ_j, P_{vj}	κ_j, P_{vj}, κ'_j, U_{iso} to dipole	ΔP_{vj}, U_{iso}, to dipole
No. of variables	172	51	488	638
No. of observations	3994	7067	7067	7067
R	0.024	0.028	0.017	0.018
wR	0.022	0.039	0.019	0.016
S	0.96	2.61	1.32	1.08
$\sin\theta/\lambda$ (Å $^{-1}$)	0.80–1.08	0.0–1.08	0.0–1.08	0.0 1.08

including those from incorrect constraints. *AMANDA* also differs from *POP* in that the parameters κ and κ' may be refined, and the function $\rho_v(r)$ describing the dependence of the valence electron density on distance from the atom center for the multipole terms may be derived from Hartree–Fock atomic orbitals (*International Tables for X-ray Crystallography*, 1974, Vol. IV) or from Slater functions. *AMANDA* uses electrical neutrality as a constraint condition, although small molecular charges may be generated by the accumulation of round-off error in successive cycles of least squares. The electrical neutrality condition may be relaxed.

In order to determine the distribution of electron density, the structure was refined using the program *AMANDA* (Ref. 3, Table 1). The multipole occupancies up to and including the hexadecapole terms were allowed to vary in the refinement and all orders of the multipoles showed significant occupancies. The function chosen for $\rho_v(r)$ was of the Hartree–Fock type. The exponential coefficients in the radial function for the higher order multipoles were taken from Hehre *et al.* (1969), *i.e.* 3.44, 3.90, 4.50 and 2.48 Bohr^{-1} for C, N, O and H atoms, respectively. The quantities η_l were taken from Stevens (1981). The number of least-squares variables was restricted by applying constraints to the multipole occupancy parameters P_{lm}, consistent with local symmetry.

The electron density was also determined using the program *POP* (Ref. 4, Table 1). A common coordinate systems was used and the multipolar expansion was limited to octapoles for the non-H atoms. The monopole term P_v for an atom represented only the deviation of the atom from neutrality, *i.e.* $\rho_c(r)$ represented the neutral atom, not the core electrons as in the *AMANDA* refinement. Further, the radial function $\rho_v(r)$ was the same Slater function as used for the higher order multipoles for the same atom. In both the *POP* and *AMANDA* refinements, all parameters except the coordinates of the H atoms were varied in the final cycles of least squares.

Table 2. *Refinement of partial charges*

Experimental charges are from Ref. 1. The *ab initio* charges are from Singh & Kollman (1984).

	κ*	P_v	q(exp) (e)	q(*ab initio*) (e)
N(1)	0.964 (4)	5.66 (7)	−0.66	−0.23
C(2)	1.030 (3)	3.18 (7)	0.82	0.85
O(2)	0.959 (4)	6.53 (4)	−0.53	−0.49
N(3)	0.981 (4)	5.58 (8)	−0.58	−0.85
C(4)	1.030 (3)	3.48 (7)	0.52	0.81
O(4)	0.959 (4)	6.59 (4)	−0.59	−0.46
C(5)	0.975 (5)	4.50 (7)	−0.50	−0.18
C(6)	1.030 (3)	3.83 (7)	0.17	0.03
C(51)	1.010 (4)	3.68 (7)	0.32	—
O(52)	0.987 (3)	6.22 (4)	−0.22	—
C(53)	1.002 (4)	4.44 (8)	−0.44	—
C(1′)	1.010 (4)	3.80 (7)	0.20	0.56
C(2′)	1.010 (4)	4.16 (8)	−0.16	−0.31
C(3′)	1.010 (4)	4.12 (7)	−0.12	0.23
C(4′)	1.010 (4)	4.09 (7)	−0.09	0.04
O(4′)	0.987 (3)	6.46 (4)	−0.46	−0.37
C(5′)	1.010 (4)	3.98 (7)	0.02	0.12
O(5′)	0.987 (3)	6.33 (4)	−0.33	−0.40
O(3′1)	0.987 (3)	6.24 (4)	−0.24	—
C(3′2)	1.030 (3)	3.15 (7)	0.85	—
O(3′3)	0.959 (4)	6.77 (4)	−0.77	—
C(3′4)	1.002 (4)	4.42 (8)	−0.42	—
H(3)	1.028 (12)	0.91 (4)	0.09	0.36
H(6)	1.028 (12)	0.89 (3)	0.11	0.13
H(511)	1.028 (12)	0.92 (3)	0.08	—
H(512)	1.028 (12)	0.92 (3)	0.08	—
H(531)	1.028 (12)	0.70 (3)	0.30	—
H(532)	1.028 (12)	0.70 (3)	0.30	—
H(533)	1.028 (12)	0.70 (3)	0.30	—
H(5′)	1.028 (12)	0.89 (4)	0.11	0.01
H(2′1)	1.028 (12)	0.79 (4)	0.21	0.08
H(2′2)	1.028 (12)	0.79 (3)	0.21	0.08
H(3′)	1.028 (12)	0.85 (4)	0.15	0.03
H(3′31)	1.028 (12)	0.79 (3)	0.21	—
H(3′32)	1.028 (12)	0.79 (3)	0.21	—
H(3′33)	1.028 (12)	0.79 (3)	0.21	—
H(4′)	1.028 (12)	0.88 (3)	0.12	0.06
H(5′1)	1.028 (12)	1.00 (3)	0.00	0.00
H(5′2)	1.028 (12)	1.00 (3)	0.00	0.04
H(5′0)	1.028 (12)	0.49 (4)	0.51	0.28

* κ for similar atoms were constrained to common values and P_v for similar H atoms were constrained to common values.

The final difference maps in the plane of the base ring are given in Fig. 2 for refinements 3 and 4. Although the results are similar the *POP* refinement shows slightly less residual electron density. This is confirmed by the significantly lower value of the goodness-of fit parameter S for the *POP* refinement, as given in Table 1. The full results from Ref. 4 are given in Tables 3 and 4. Important bond lengths, bond angles and torsional angles from Ref. 4 are given in Tables 5, 6 and 7, respectively.* The bond lengths in Table 5 are not corrected for the effects of thermal motion. Given the level of accuracy quoted, corrections for thermal motion may lead to significant changes in the bond lengths.

* A list of structure amplitudes with the calculated values corresponding to Ref. 4 has been deposited with the British Library Document Supply Centre as Supplementary Publication No. SUP 71591 (66 pp.). Copies may be obtained through The Technical Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England. [CIF reference: CR0428]

Table 3. *Final least-squares parameters for multipole refinement using POP*

(*a*) Atomic coordinates and U_{eq} ($\times 10^4$)

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

	x	y	z	U_{eq}
N(1)†	0.43462	0.67838	0.32312	131 (1)
C(2)	0.32128 (9)	0.86771 (6)	0.36109 (5)	130 (1)
O(2)	0.37378 (10)	1.03386 (6)	0.30032 (6)	186 (2)
N(3)	0.14636 (9)	0.85902 (5)	0.47397 (4)	140 (1)
C(4)	0.08249 (8)	0.68093 (6)	0.55869 (5)	137 (1)
O(4)	−0.07606 (11)	0.69601 (7)	0.65889 (6)	215 (2)
C(5)	0.22067 (8)	0.48572 (5)	0.51939 (6)	127 (1)
C(6)	0.38720 (9)	0.49401 (6)	0.40423 (5)	132 (1)
C(51)	0.17782 (20)	0.28265 (7)	0.60718 (5)	145 (1)
O(52)	−0.02924 (9)	0.23676 (5)	0.57760 (5)	155 (1)
C(53)	−0.20898 (12)	0.23075 (9)	0.69621 (8)	260 (2)
C(1′)	0.62589 (8)	0.67091 (6)	0.20360 (5)	126 (1)
C(2′)	0.86653 (9)	0.58457 (5)	0.25328 (5)	141 (1)
C(3′)	1.00799 (9)	0.49001 (7)	0.12670 (5)	131 (1)
O(3′1)	1.10101 (8)	0.65519 (6)	0.02823 (4)	150 (1)
C(3′2)	1.26601 (9)	0.59406 (6)	−0.07817 (5)	158 (1)
O(3′3)	1.32439 (12)	0.41601 (9)	−0.09820 (8)	205 (2)
C(3′3)	1.36050 (14)	0.77430 (11)	−0.16790 (9)	235 (2)
C(4′)	0.82608 (9)	0.42698 (6)	0.05923 (5)	133 (1)
O(4′)	0.59957 (9)	0.53185 (7)	0.11966 (5)	150 (1)
C(5′)	0.83767 (11)	0.19417 (7)	0.08973 (6)	191 (2)
O(5′)	0.81924 (12)	0.11189 (9)	0.24044 (7)	228 (2)
H(3)	0.0666‡	1.0011	0.4968	225 (29)
H(6)	0.4972	0.3588	0.3658	62 (22)
H(511)	0.3269	0.1591	0.5802	219 (28)
H(512)	0.1771	0.2834	0.7212	208 (28)
H(531)	−0.3359	0.1743	0.6613	264 (31)
H(532)	−0.2637	0.3662	0.7429	362 (38)
H(533)	−0.1373	0.1153	0.7815	416 (57)
H(1′)	0.5942	0.8311	0.1429	98 (21)
H(2′1)	0.9366	0.7107	0.2710	241 (31)
H(2′2)	0.8543	0.4620	0.3484	208 (29)
H(3′1)	1.1514	0.3553	0.1498	186 (27)
H(3′31)	1.2334	0.8615	−0.2357	394 (40)
H(3′32)	1.3800	0.8705	−0.0990	504 (46)
H(3′33)	1.5174	0.7235	−0.2338	475 (44)
H(4′1)	0.8422	0.4802	−0.0574	221 (27)
H(5′1)	1.0018	0.1105	0.0389	278 (31)
H(5′2)	0.6931	0.1648	0.0469	299 (33)
H(5′0)	0.6650	0.0845	0.2724	372 (34)

(*b*) Anisotropic mean-square atomic displacement parameters ($\times 10^4$)

	U_{11}	U_{22}	U_{33}	U_{12}	U_{13}	U_{23}
N(1)	137 (2)	102 (1)	138 (1)	−28 (1)	27 (1)	−24 (2)
C(2)	129 (1)	101 (1)	152 (1)	−31 (1)	15 (1)	−26 (1)
O(2)	187 (2)	108 (1)	238 (2)	−54 (1)	49 (1)	−21 (1)
N(3)	144 (1)	100 (1)	166 (1)	−31 (1)	32 (1)	38 (1)
C(4)	137 (1)	116 (1)	151 (1)	−37 (1)	29 (1)	−37 (1)
O(4)	217 (2)	165 (1)	231 (2)	−56 (1)	110 (2)	−59 (1)
C(5)	135 (1)	102 (1)	138 (1)	−35 (1)	14 (1)	−25 (1)
C(6)	141 (1)	102 (1)	141 (1)	−28 (1)	21 (1)	−27 (1)
C(51)	154 (1)	112 (1)	157 (2)	−41 (1)	1 (1)	−10 (1)
O(52)	165 (1)	153 (1)	157 (1)	−72 (2)	28 (1)	−49 (1)
C(53)	178 (2)	377 (3)	243 (2)	−107 (2)	74 (1)	−120 (2)
C(1′)	129 (1)	121 (1)	122 (1)	−30 (1)	9 (1)	−24 (1)
C(2′)	135 (1)	166 (1)	128 (1)	−51 (1)	4 (1)	−31 (1)
C(3′)	113 (1)	124 (1)	143 (1)	−30 (1)	7 (1)	−16 (1)
O(3′1)	130 (1)	130 (1)	170 (2)	−41 (1)	32 (1)	−17 (1)
C(3′2)	132 (1)	175 (2)	151 (1)	−44 (1)	20 (1)	−18 (1)
O(3′3)	196 (2)	204 (2)	207 (2)	−33 (1)	36 (1)	−77 (2)
C(3′3)	218 (2)	251 (2)	202 (2)	−102 (2)	36 (1)	19 (1)
C(4′)	134 (1)	134 (1)	133 (1)	−40 (1)	15 (1)	−39 (1)
O(4′)	120 (1)	192 (2)	155 (1)	−31 (1)	−3 (1)	−77 (1)
C(5′)	220 (2)	143 (1)	220 (2)	−66 (1)	12 (1)	−60 (1)
O(5′)	251 (2)	182 (2)	249 (2)	−102 (2)	−48 (1)	26 (1)

† Origin defining.

‡ H-atom positions set at C—H, N—H and O—H distances taken from neutron diffraction results.

Table 4. *Multipole occupancy parameters from POP refinement*

The reference coordinate system consists of the *a*, *b**, *axb** axes. (ΔP_v is multiplied by 100 and the higher order terms are multiplied by 10.)

	ΔP_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
N(1)	28 (2)	−2 (2)	−2 (1)	−1 (1)	−2 (1)	3 (1)	6 (1)	2 (1)	3 (1)	7 (1)	1 (1)	−18 (1)	−1 (1)	7 (1)	−1 (1)	−15 (2)
C(2)	−13 (3)	8 (2)	−1 (1)	−2 (2)	−16 (1)	7 (1)	−12 (1)	−2 (1)	0 (1)	−23 (1)	5 (1)	21 (1)	−2 (1)	−13 (1)	−9 (1)	11 (1)
O(2)	26 (2)	5 (1)	9 (1)	−5 (1)	−5 (1)	−1 (1)	0 (1)	2 (1)	6 (1)	−2 (1)	1 (1)	−7 (1)	−5 (1)	0 (1)	−2 (1)	−7 (1)
N(3)	27 (3)	−3 (2)	−1 (1)	2 (1)	−3 (1)	3 (1)	8 (1)	2 (1)	2 (1)	8 (1)	−4 (1)	−17 (1)	−5 (1)	6 (1)	1 (1)	−13 (1)
C(4)	−8 (3)	−8 (2)	−1 (1)	3 (2)	−7 (1)	−3 (1)	−19 (1)	7 (1)	4 (1)	−15 (1)	3 (1)	18 (1)	−7 (1)	−17 (1)	1 (1)	7 (1)
O(4)	27 (2)	−15 (2)	7 (1)	0 (1)	−11 (1)	3 (1)	0 (1)	−1 (1)	−2 (1)	2 (1)	0 (1)	−14 (1)	0 (1)	−8 (1)	0 (1)	−5 (1)
C(5)	15 (3)	2 (2)	−3 (2)	−1 (1)	−5 (1)	1 (1)	−6 (1)	1 (1)	2 (1)	14 (1)	3 (1)	−20 (1)	−4 (1)	3 (1)	−3 (1)	−16 (1)
C(6)	−9 (3)	2 (2)	−10 (1)	8 (1)	−6 (1)	−6 (1)	−14 (1)	6 (1)	2 (1)	−17 (1)	−2 (1)	16 (1)	−4 (1)	−10 (1)	−4 (1)	10 (1)
C(51)	−6 (3)	5 (2)	7 (2)	13 (2)	−5 (1)	0 (1)	3 (1)	3 (1)	8 (1)	−15 (1)	−10 (1)	−1 (1)	−4 (1)	6 (1)	1 (1)	18 (1)
O(52)	27 (2)	−1 (2)	−2 (1)	−9 (1)	−10 (1)	−3 (1)	6 (1)	−1 (1)	1 (1)	1 (1)	1 (1)	2 (1)	−5 (1)	−10 (1)	−4 (1)	−10 (1)
C(53)	22 (4)	−19 (2)	14 (2)	−4 (1)	−7 (1)	5 (2)	13 (1)	−6 (1)	6 (1)	11 (2)	−11 (1)	−6 (1)	−6 (1)	19 (1)	−6 (1)	8 (1)
C(1′)	−12 (3)	8 (2)	5 (2)	−1 (2)	−6 (1)	−5 (1)	10 (1)	−3 (1)	−4 (1)	4 (1)	−12 (1)	19 (1)	−18 (1)	−21 (1)	−6 (1)	−3 (1)
C(2′)	8 (3)	−2 (2)	−5 (1)	−4 (1)	−7 (1)	1 (1)	6 (1)	−2 (1)	4 (1)	−4 (1)	13 (1)	−11 (1)	11 (1)	7 (1)	−11 (1)	−12 (1)
C(3′)	4 (3)	−4 (2)	−10 (2)	5 (1)	0 (1)	−8 (1)	9 (1)	9 (1)	2 (1)	−17 (1)	−5 (1)	−8 (1)	−23 (1)	−11 (1)	12 (1)	8 (1)
O(3′1)	27 (2)	−4 (1)	5 (1)	1 (1)	−3 (1)	0 (1)	5 (1)	5 (1)	2 (1)	2 (1)	−2 (1)	−7 (1)	6 (1)	0 (1)	−5 (1)	−10 (1)
C(3′2)	−9 (4)	−1 (2)	−12 (2)	−12 (1)	−22 (1)	6 (1)	−13 (1)	9 (1)	−3 (1)	−8 (1)	16 (1)	12 (1)	−20 (1)	−13 (1)	8 (1)	8 (1)
O(3′3)	26 (2)	−6 (2)	−11 (1)	−4 (1)	−4 (1)	5 (1)	−7 (1)	0 (1)	4 (1)	1 (1)	−1 (1)	−5 (1)	1 (1)	−3 (1)	−1 (1)	−7 (1)
C(3′3)	−21 (2)	−4 (2)	−5 (2)	−4 (2)	−2 (1)	1 (1)	3 (1)	5 (1)	7 (1)	12 (1)	4 (1)	−4 (1)	5 (1)	−1 (1)	1 (1)	−14 (1)
C(4′)	−6 (3)	6 (2)	−6 (2)	−2 (2)	−6 (1)	6 (1)	11 (1)	0 (1)	5 (1)	6 (1)	20 (1)	6 (1)	−8 (1)	−4 (1)	12 (1)	−22 (2)
O(4′)	26 (2)	−8 (1)	−4 (1)	0 (1)	−6 (1)	3 (1)	5 (1)	−8 (1)	5 (1)	−5 (1)	−3 (1)	−9 (1)	4 (1)	2 (1)	1 (1)	−1 (1)
C(5′)	14 (4)	−3 (2)	3 (2)	−8 (2)	−2 (1)	3 (1)	0 (1)	6 (1)	−3 (1)	−7 (1)	−16 (1)	−15 (1)	−1 (1)	−1 (1)	−22 (1)	13 (1)
O(5′)	17 (2)	3 (2)	1 (1)	4 (1)	−5 (1)	3 (1)	5 (1)	4 (1)	−3 (1)	−3 (1)	−4 (1)	−3 (1)	4 (1)	3 (1)	2 (1)	−9 (1)
H(3)	−18 (2)	−2 (2)	−17 (2)	−2 (1)												
H(6)	−21 (1)	−10 (1)	7 (1)	11 (1)												
H(511)	−7 (2)	−14 (2)	11 (2)	7 (1)												
H(512)	−12 (2)	−15 (2)	1 (2)	−16 (2)												
H(531)	−18 (2)	4 (2)	11 (2)	8 (2)												
H(532)	−15 (2)	6 (2)	−14 (2)	−15 (2)												
H(533)	−13 (2)	−7 (2)	−3 (2)	−10 (2)												
H(1′)	−11 (2)	12 (1)	−15 (1)	1 (1)												
H(2′1)	−9 (2)	−13 (2)	−18 (2)	−1 (1)												
H(2′2)	−10 (2)	6 (2)	12 (2)	−7 (1)												
H(3′1)	−9 (2)	−4 (2)	17 (2)	7 (1)												
H(3′31)	−9 (2)	12 (2)	−3 (2)	12 (1)												
H(3′32)	−6 (2)	−3 (2)	4 (2)	−10 (2)												
H(3′33)	−9 (2)	−15 (2)	−1 (2)	12 (2)												
H(4′1)	−7 (2)	2 (2)	−4 (2)	14 (2)												
H(5′1)	−8 (2)	−11 (2)	14 (2)	9 (2)												
H(5′2)	−11 (2)	16 (2)	10 (2)	−1 (2)												
H(5′0)	−18 (2)	16 (2)	7 (2)	−11 (2)												

† ΔP_v is the monopole term, d_1–d_3 are the dipole terms, q_1–q_5 are the quadrupole terms and o_1–o_7 are the octapole terms.

The results were used to calculate dynamic model deformation maps using equation (4),

$$\delta\rho(\mathbf{r}) = \frac{1}{V}\sum_{\mathbf{H}}[|F_{\text{mul}}|\exp(i\vartheta_{\text{mul}}) - |F_{\text{sph}}|\exp(i\vartheta_{\text{sph}})]\exp(-2\pi i\mathbf{H}.\mathbf{R}), \quad (4)$$

where 'mul' refers to quantities calculated from the full multipole model, accounting for the displacement parameters, and 'sph' refers to quantities calculated from a model with neutral spherical atoms placed at the same locations as the atoms in the molecule, as determined in Ref. 4. The spherical atom model which is subtracted is the independent atom model of Stewart (1991) and the promolecule of Hirshfeld (1971).

2.5. *Electrostatic potential distributions*

The results from the POP refinement are in a form that facilitates the calculation of the electrostatic potential, by summing over the contributions from all multipoles, the spherical electron-density distribution and the nuclear charge of each atom, using a method proposed by Stewart (1982). The method assumes no thermal motion of the molecule. Distributions of the electrostatic potential of the isolated molecule of MMAcdUrd were calculated using (3) with Stewart's method and the program *MOLPOT*89 (He, 1984). For the study of the electrostatic potentials of hydrogen bonds, the contribution of both molecules involved in the hydrogen bonds were included.

3. Discussion

3.1. *Partial charges*

The experimentally determined partial charges are compared in Table 4 with values determined by *ab initio* quantum mechanics computations by Singh & Kollman (1984). The computed values are obtained

Table 5. *Bond distances for MMAcdUrd* (Å)

N(1)—C(2)	1.3792 (4)	N(1)—C(6)	1.3779 (4)
N(1)—C(1′)	1.4687 (4)	C(2)—O(2)	1.2303 (6)
N(3)—C(2)	1.3680 (6)	N(3)—C(4)	1.3905 (6)
C(4)—O(4)	1.2268 (7)	C(4)—C(5)	1.4549 (6)
C(5)—C(6)	1.3512 (7)	C(5)—C(51)	1.4949 (6)
C(51)—O(52)	1.4342 (10)	O(52)—C(53)	1.4211 (10)
C(1′)—C(2′)	1.5247 (7)	C(1′)—O(4′)	1.4141 (8)
C(2′)—C(3′)	1.5175 (7)	C(3′)—O(3′1)	1.4487 (6)
C(3′)—C(4′)	1.5412 (8)	O(3′1)—C(3′2)	1.3453 (6)
C(3′2)—O(3′3)	1.2104 (8)	C(3′2)—C(3′3)	1.4979 (9)
C(4′)—C(3′)	1.5412 (8)	C(4′)—O(4′)	1.4459 (7)
C(4′)—C(5′)	1.5162 (6)	C(5′)—O(5′)	1.4225 (8)
H(6)—C(6)	1.077	H(511)—C(51)	1.093
H(512)—C(51)	1.092	H(531)—C(53)	1.058
H(532)—C(53)	1.058	H(533)—C(53)	1.062
H(1′)—C(1′)	1.094	H(2′1)—C(2′)	1.092
H(2′2)—C(2′)	1.091	H(3′1)—C(3′)	1.092
H(3′31)—C(3′3)	1.059	H(3′32)—C(3′3)	1.060
H(3′33)—C(3′3)	1.059	H(4′1)—C(4′)	1.090
H(5′1)—C(5′)	1.092	H(5′2)—C(5′)	1.094
H(5′0)—O(5′)	0.968	H(3)—N(3)	1.009

Table 6. *Bond angles for MMAcdUrd* (°)

C(2)—N(1)—C(6)	121.03 (3)	C(2)—N(1)—C(1′)	119.52 (3)
C(6)—N(1)—C(1′)	119.13 (2)	N(1)—C(2)—O(2)	123.16 (4)
N(1)—C(2)—N(3)	115.32 (3)	O(2)—C(2)—N(3)	121.51 (4)
C(2)—N(3)—C(4)	127.14 (3)	N(3)—C(4)—O(4)	120.23 (4)
N(3)—C(4)—C(5)	114.54 (4)	O(4)—C(4)—C(5)	125.22 (4)
C(4)—C(5)—C(6)	118.47 (4)	C(4)—C(5)—C(51)	119.75 (4)
C(6)—C(5)—C(51)	121.78 (4)	N(1)—C(6)—C(5)	123.20 (3)
C(5)—C(51)—O(52)	112.05 (4)	C(51)—O(52)—C(53)	114.52 (7)
N(1)—C(1′)—C(2′)	113.42 (4)	N(1)—C(1′)—O(4′)	107.91 (4)
C(2′)—C(1′)—O(4′)	105.84 (3)	C(1′)—C(2′)—C(3′)	102.70 (4)
C(2′)—C(3′)—O(3′1)	106.62 (4)	C(2′)—C(3′)—C(4′)	103.85 (4)
O(3′1)—C(3′)—C(4′)	110.54 (4)	C(3′)—O(3′1)—C(3′2)	115.59 (4)
O(3′1)—C(3′2)—O(3′3)	123.25 (5)	O(3′1)—C(3′2)—C(3′3)	111.24 (4)
O(3′3)—C(3′2)—C(3′3)	125.50 (6)	C(3′)—C(4′)—O(4′)	106.52 (4)
C(3′)—C(4′)—C(5′)	114.15 (4)	O(4′)—C(4′)—C(5′)	108.74 (4)
C(1′)—O(4′)—C(4′)	109.90 (5)	C(4′)—C(5′)—O(5′)	111.53 (5)

Table 7. *Selected torsion angles for MMAcdUrd* (°)

C(2)—N(1)—C(1′)—C(2′)	105.45 (4)
C(6)—N(1)—C(1′)—C(2′)	−68.21 (5)
C(2)—N(1)—C(1′)—O(4′)	−137.65 (4)
C(6)—N(1)—C(1′)—O(4′)	48.70 (5)
O(2)—C(2)—N(1)—C(1′)	−0.77 (7)
N(1)—C(1′)—C(2′)—C(3′)	151.93 (3)
N(1)—C(1′)—O(4′)—C(4′)	−147.37 (3)
C(1′)—C(2′)—C(3′)—C(4′)	−28.68 (4)
C(1′)—C(2′)—C(3′)—O(3′1)	88.11 (5)
C(2′)—C(3′)—C(4′)—C(5′)	−105.42 (4)
C(2′)—C(3′)—C(4′)—O(4′)	14.60 (4)
C(3′)—C(4′)—C(5′)—O(5′)	55.21 (6)
C(3′)—C(4′)—O(4′)—C(1′)	6.85 (5)
O(4′)—C(4′)—C(5′)—O(5′)	−63.55 (6)
C(2′)—C(1′)—O(4′)—C(4′)	−25.65 (5)
O(3′1)—C(3′)—C(4′)—C(5′)	140.57 (4)
O(3′1)—C(3′)—C(4′)—O(4′)	−99.41 (4)
C(6)—C(5)—C(51)—O(52)	−104.58 (6)
C(4)—C(5)—C(51)—O(52)	76.13 (7)

by first calculating the electrostatic potential for points within a shell surrounding the molecule directly, using an STO-3G basis set. Point charges are then placed on the atom centers and are varied by least squares. The least-squares calculation fits the electrostatic potential at the points in the shell arising from those point charges to the electrostatic potential determined from the computational method. This was performed for two fragments comprising major portions of MMAcdUrd. Those fragments are 1-CH_3 thymine and 1-NH_3 deoxyribose. Therefore, the atoms N(1), C(5), C(1′) and C(3′) are substituted differently in the experimental and computational models. The role of substitution is demonstrated by the fact that in unsubstituted thymine the partial charge on N(1) is calculated by Singh & Kollman (1984) to be −0.62 e, and in uracil the partial charge on C(5) is calculated to be −0.52 e. Both values are in agreement with the experimentally determined values for MMAcdUrd but differ significantly from the computational values for 1-CH_3 thymine, which are −0.23 and −0.18 e, respectively.

Given the wide range of published values of computationally determined partial charges for related compounds, the quality of agreement between the computed results of Singh & Kollman (1984) and the experimentally determined values is interesting.

3.2. *Electron-density distribution*

The dynamic model deformation-density map in the plane of the base ring as determined from the *POP* refinement is shown in Fig. 3. Several features are of interest. The lone pairs on O(2) are not resolved in this plane or in any other plane. The lone pairs on O(4) are clearly resolved in the plane of the ring. The molecule packing of MMAcdUrd is relatively simple, with $Z = 1$. Adjacent and identically oriented molecules are hydrogen bonded to form columns along the b axis. The O(2) atom is the acceptor of a hydrogen bond from O(5′) on an adjacent molecule, and N(3) is the donor of a hydrogen bond to O(5,2) on the same adjacent molecule. The O(4) atom is not involved in any hydrogen bonds and shows higher anisotropic thermal motion. Therefore, the difference between the electron-density distribution of O(2) and O(4) is not caused by thermal smearing.

The single peak in the electron density in the region of the lone pair of O(2) is on the same side of the C(2)—O(2) axis as is the donor atom O(5′).* The lone pair on O(4) which is closer to H(3) shows greater electron density than the lone pair closer to C(51). The two peaks subtend an angle of 130° at the O(4) position. The plane of the two peaks and the O(4) position make an angle of 12° with the plane of the base ring, with the weaker peak below the plane by a distance of 0.09 Å and the stronger peak 0.06 Å above the plane. The O(4) atom itself is 0.038 (1) Å below the plane and the C(4)—O(4) bond is bent 2.5°

* This is on the same side of the base plane as the atom C(2′) (see Fig. 1). This side will be designated as the 'above' side of the base plane and the other side will be designated the 'below' side of the base plane.

away from C(51), suggesting a significant effect of the C(5) substituent on the position and electron density of O(4).

The electron densities in the bonds forming the base ring have peaks with similar maximum deformation densities, but not necessarily similar volumes. As shown in Fig. 3, within the plane of the ring the formal C(5)=C(6) double bond is intermediate in peak height between that of N(1)—C(2) and that with the minimum density, C(2)—N(3). This is only a section through the electron density of the base ring and does not necessarily reflect the total electron density of the bonds.

Fig. 4. shows the dynamic model deformation density in the ribose ring calculated from the *POP* refinement. It is constructed by first calculating the map of the dynamic density in the plane formed by each atom with its neighboring ring atoms. These maps were joined at the midpoints of the ring bonds. Thus, the contours in Fig. 4 are from planes through each atom and the angles between adjacent planes are the bond torsional angles associated with the bond between the same atoms. Both C(4′) and O(4′) show lobes of electron density external to the ring.

Fig. 5 shows a section through O(4′) and perpendicular to the C(2′)—C(3′) bond as calculated from the *POP* refinement. Two lobes of electron density are clearly visible. This, together with the value of 109.90 (5)° for the C(1′)—O(4′)—C(4′) angle, indicate nearly ideal sp^3 hybridization of O(4′). The central electron density of the C(2′)—C(3′) bond is nearly circular in cross section and its bond length, 1.5175 (7) Å, is consistent with the bond not being affected by delocalization.

It should be noted that the appearance of lone pairs in the deformation model dynamic-density map is highly model dependent. The *AMANDA* refinement gives one broad continuous lobe of electron density for the O(4′) lone pairs. The *AMANDA* refinement uses a different radial distribution function than *POP* for the deformation electron-density hexadecapole population parameters, while *POP* does not, and uses constraints imposing mirror symmetry on O(4′) while *POP* does not. It would appear that the hexadecapole population parameters do not contribute significanlty to the resolution of the electron density and that constraints must be used with great caution.

The reason for the low density in the O(4′)—C(1′) and O(4′)—C(4′) bonds has been discussed by Coppens (1989) and Souhassou *et al.* (1991). It reflects the fact that for atoms on the right side of the periodic table, the spherical atom which has been subtracted from the experimentally determined bonding electron density formally contains more

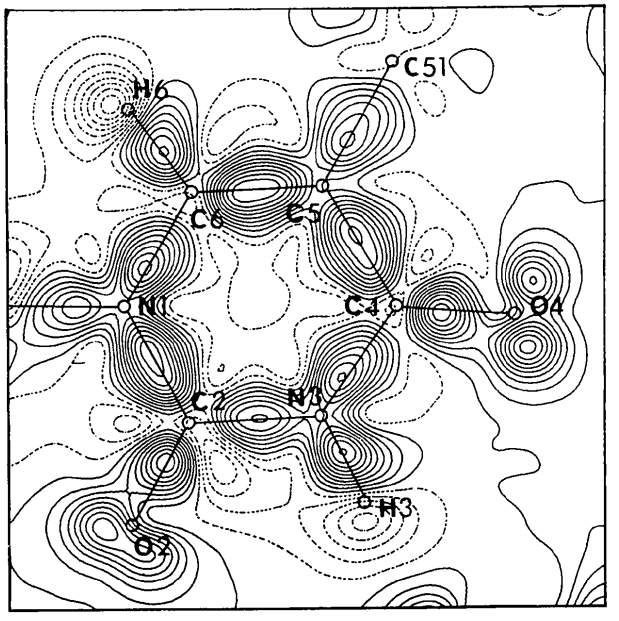


Fig. 3. Dynamic model deformation density in the plane of the base ring using the results from the *POP* refinement. Contours as in Fig. 2.

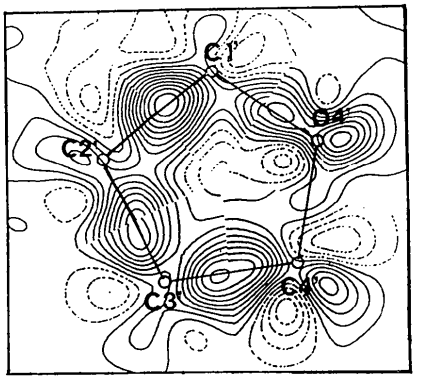


Fig. 4. Composite dynamic model deformation density of the deoxyribose ring using the results from the *POP* refinement. Contours as in Fig. 2.

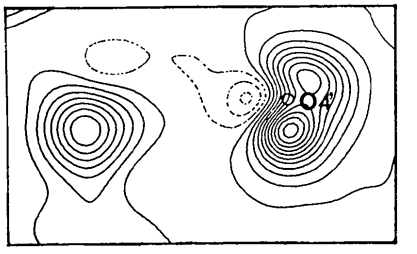


Fig. 5. Dynamic model deformation density in a plane through O(4′) and the C(2′)—C(3′) bond, normal to the latter, using the results from the *POP* refinement. The cross section of the C(2′)—C(3′) bond is the lobe of electron density on the left. Contours as in Fig. 2.

electron density than would be placed there by the hybridized orbitals. For nitrogen, 5/4 electrons are subtracted per *L*-shell orbital and for oxygen, 6/4. For both oxygen and nitrogen, we find positive deformation density at their atom-origin site and a tendency for the deformation bonding density to be polarized toward their atom site. The deformation bond density of N—C ring bonds appears similar to that for C—C bonds and O—C bonds show significantly less deformation bond density than do C—C bonds.

The O(4′)—C(1′) bond is an acetal bond produced by the condensation of an aldehyde and an alcohol, while the O(4′)—C(4′) bond is an ether bond. The latter is 0.032 (1) Å longer and has less electron density than the former, as shown in Fig. 4. The difference is demonstrated more clearly in Fig. 6, which gives the O(4′)—C(1′) and O(4′)—C(4′) bonds in cross section. The O(4′)—C(4′) bond also appears elongated, perpendicular to the plane of the ribose ring. This is characteristic of a partial double bond, as observed with the resolution used in this refinement (Pearlman & Kim, 1985).

Fig. 5 shows that the electron densities of the lone pairs on O(4′) are not equal. That which is *cis* to the C(1′)—N(1) bond has approximately 0.15 e Å^{-3} greater electron density. Although one might expect electron–electron repulsion between the lone-pair electron density and electron density in the C(1′)—N(1) bond to promote electrons to the *trans* lone pair from the *cis* lone pair, Wolfe (1972) has shown that electron–nucleus interactions involving lone-pair electrons and polar bonds are also important and consideration of the total conformational energy usually shows the *gauche* conformation to have the minimum energy. The C(2′)—C(1′)—O(4′)—C(4′) torsional angle is −25.65 (4)°, as dictated by the overall conformation of the deoxyribose ring. This implies that the lone pairs and the N(1) and H(1′) atoms are close to an eclipsed conformation. Nevertheless, the greater electron density of the *cis* lone pair is consistent with the aforementioned *gauche* effect.

Taylor & Kennard (1984) have studied the role of lone pairs in determining the direction of hydrogen bonds and concluded that although spectroscopic

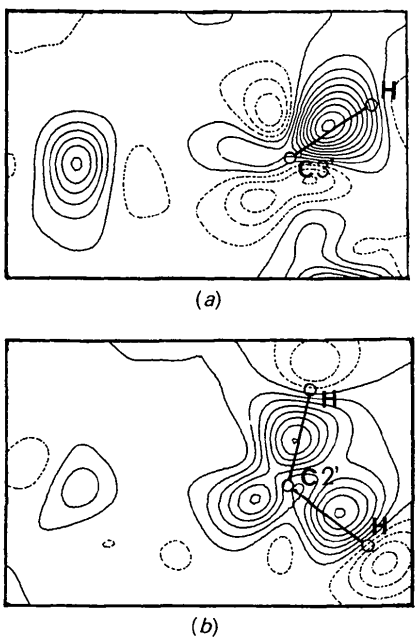


Fig. 6. Dynamic model deformation density of (*a*) the O(4′)—C(1′) bond and (*b*) the O(4′)—C(4′) bond using results from the *POP* refinement. In each case the plane is defined by the normal to the bond and the atom of the deoxyribose ring opposite the bond. The cross section of the bond is the lobe of electron density on the left of the figure. Contours as in Fig. 2.

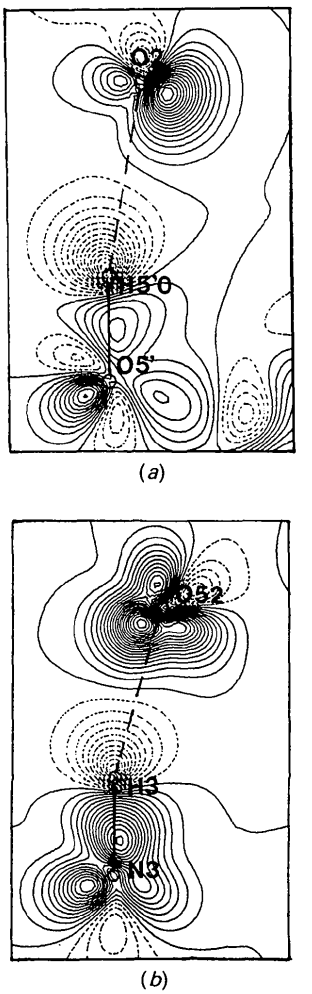


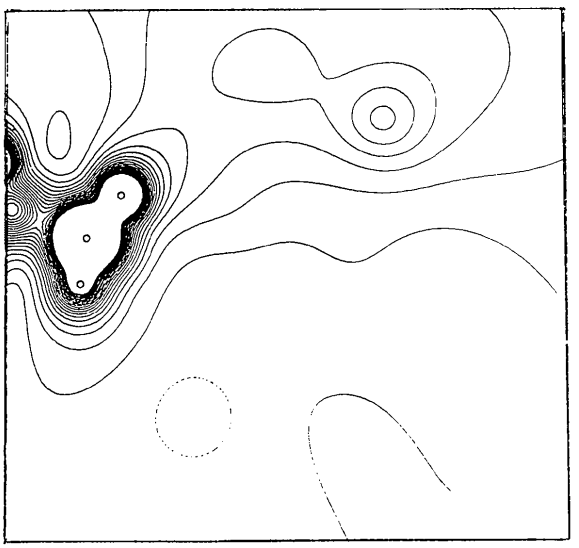
Fig.7. Static multipole density in the planes of the hydrogen bonds using the results from the *POP* refinement: (*a*) O(5′)—H(5′0)···O(2) and (*b*) N(3)—H(3)···O(52). Contours as in Fig. 2.

results show a preference for hydrogen bonds in the direction of lone pairs, the energies involved are small in comparison with crystal-packing effects and no correlations of hydrogen-bond directions with lone-pair directions are found in crystals.

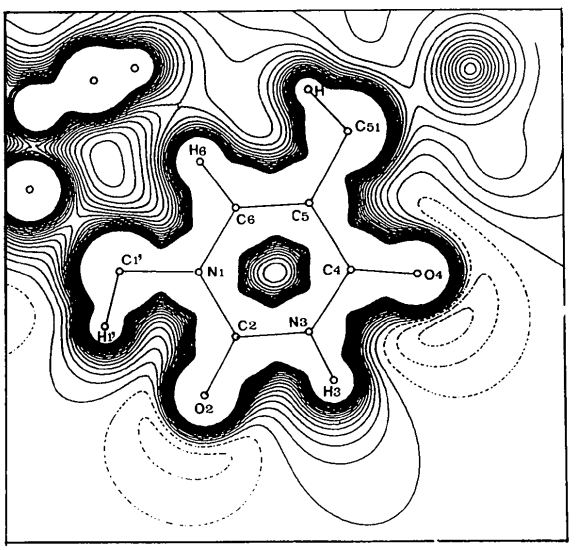
The electron-density distributions, which are contoured in Fig. 7, are calculated directly by summing the contribution of each multipole, as determined from the *POP* refinement, at points in the plane. The method does not account for the effects of thermal motion or for random errors in the observed structure amplitudes. This leads to the highest possible resolution, but the effects of any systematic errors will be obscured. In the case of the O(5′)—H(5′O)···O(2) bond, the lone pairs on O(2) are clearly not pointing in the direction of the donor H atom. However, O(2) is removed 0.085 (1) Å from the plane of the phenyl ring toward the donor atom O(5′), and the single peak in the lone-pair electron density on O(2) points in the general direction of O(5′). The O(5′)—O(2) distance is 2.7775 (10) Å and the angle subtended at H(5′0) is 170.3°.

The plane of Fig. 7(*b*) forms angles of 77.0 (2) and 76.3 (2)° with the plane of C(51), O(52) and C(53) atoms and the plane fitted to the base ring, respectively. This tilts the plane into the bonds from O(52) and N(3) to neighboring atoms in the methoxy substituent and in the ring. The electron densities of these bonds appear as pairs of lobes pointing away from the hydrogen bond. Fig. 7(*b*) shows that the region of depleted electron density on H(3) of the N(3)—H(3)···O(52) hydrogen bond is directly adjacent to a lobe of electron density on O(52). The N(3)—O(52) distance is 2.8235 (8) Å and the N(3)—H(3)···O(52) angle is 164.6°.

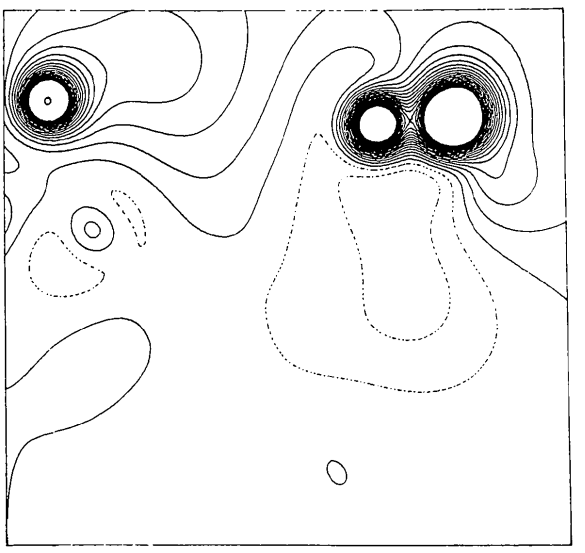
The deformation densities, as illustrated in Fig. 7, show polarization of the electron density of the hydrogen-bonded H atom toward the acceptor. Fig. 3 indicates that this polarization is not significantly greater than for the non-hydrogen-bonded H(6) atom.



(*a*)



(*b*)



(*c*)

Fig. 8. Electrostatic potential distribution of the base ring using the results from the *POP* refinement: (*a*) 1.75 Å above the plane of the ring [toward the donor O(5′) of the hydrogen bond to O(2)], (*b*) in the plane of the ring and (*c*) 1.75 Å below the plane of the ring. The contour interval is 0.05 e $Å^{-1}$ with positive contours solid and negative contours dotted. The zero counter is omitted. The orientation of the plane is as in Fig. 3.

3.3. *Electrostatic potential distributions*

Fig. 8 shows the electrostatic potential distribution (*a*) in a plane 1.75 Å above the plane of the base ring, (*b*) in the plane of the ring and (*c*) 1.75 Å below the plane of the ring as determined from the *POP* refinement. The large extension of the positive electrostatic potential around H(3) and regions of negative electrostatic potential around O(2) and O(4) are apparent in Fig. 8(*b*). The latter is highly asymmetric, reflecting the asymmetry in the electron densities of the lone pairs on O(4) and the strong influence of the positive charges on the neighboring methoxymethyl substituent, residing primarily on its H atoms (see Table 2). The asymmetry places the peak of negative potential both toward the H(3) atom and below the plane of the base ring.

The region of negative electrostatic potential associated with O(4) extends to the lower face of the 5-methoxymethyluridine ring as is shown in Fig. 8(*c*). There it is joined by N(3) and O(52) in creating a large region of slightly negative electrostatic potential on the ring face. The contribution by O(52) to negative electrostatic potential at points near the exterior of the molecule is consistent with the STO-5G computations by Politzer & Murray (1993) showing strong negative electrostatic potential on the surface of the H_3COCH_3 molecule around the vicinity of the O atom.

Politzer & Murray (1990) have carried out extensive computations of the electrostatic potential on the surface faces of planar molecules, which they choose as being 1.75 Å from the central plane. Regions of negative electrostatic potential are particularly important because they are favorable for electrophilic attack. The greater concentration of the positive charge of a molecule means that most of the surface of a neutral molecule will show positive electrostatic potential. Politzer & Murray (1990) have also discussed the tendency for the nominally stronger bonds in unsaturated systems to show negative electrostatic potential at the surface in simple molecules. However, the electron-withdrawing power of substituents such as a chloro substituent may cause regions of negative electrostatic potential to disappear from a molecule. Out-of-plane substituents such as the 5-methoxymethyl substituents would be expected to have even more profound effects on the regions of negative electrostatic potential. The region above the nominally strong C(5)—C(6) bond in MMAcdUrd actually has a strongly positive electrostatic potential relative to other bonds in all planes between the central plane and the surface planes.

Although the electron density of the lone pairs on O(2) is highly asymmetric, the effect of the close proximity of atom H(1′) to the lone pairs is to create a more symmetric distribution of electrostatic potential than for O(4). The peak of electrostatic potential near O(2) is located slightly above the plane of the base toward the donor H atom on an O(5′1) atom of a neighboring molecule.

The electrostatic potential distribution in the plane of the 3′-*O*-acetyl group as determined from the *POP* refinement is illustrated in Fig. 9. The carbonyl O(3′3) atom has a much more pronounced region of negative electrostatic potential than does the ether O(3′1) atom. Furthermore, the electrostatic potential distribution around the O(3′3) atom is highly asymmetric, reflecting the influence of the positive charges on the methyl H atoms on C(3′3). It is noteworthy that the negative electrostatic potential distributions on O(4) and O(3′3) are similarly influenced by neighboring methyl H atoms and both do not accept hydrogen bonds.

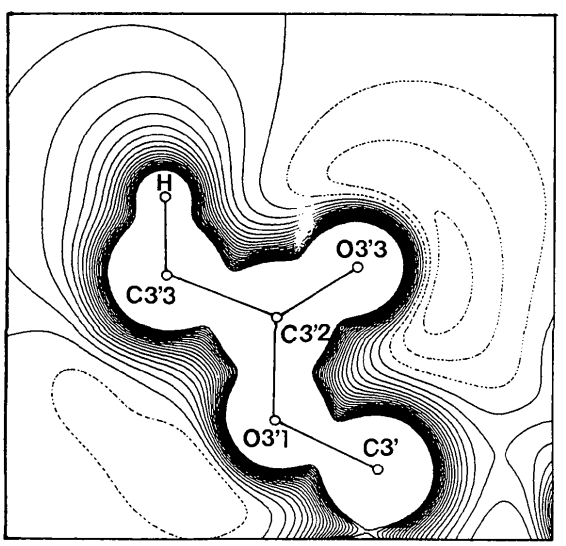


Fig. 9. Electrostatic potential distribution of the 3′-*O*-acetyl group using the results from the *POP* refinement. Contours as in Fig. 8.

Fig. 10 shows for each of the two hydrogen bonds in the plane defined by the three atoms of the hydrogen bond (*a*) the electrostatic potential of the acceptor as contributed by the charge distribution of the acceptor molecule, (*b*) the electrostatic potential in the region of the hydrogen bond as contributed by the charge distribution of both the donor and acceptor molecules and (*c*) the electrostatic potential of the donor as contributed by the charge distribution of the donor molecule.

The electrostatic potential in both bonds is everywhere positive. In both cases, the positive electrostatic potential around the donor points toward the acceptor and, in the case of the O(5′)—H(51)···O(2) hydrogen bond, the negative electrostatic potential points toward the donor. For the N(3)—H(3)···O(52)

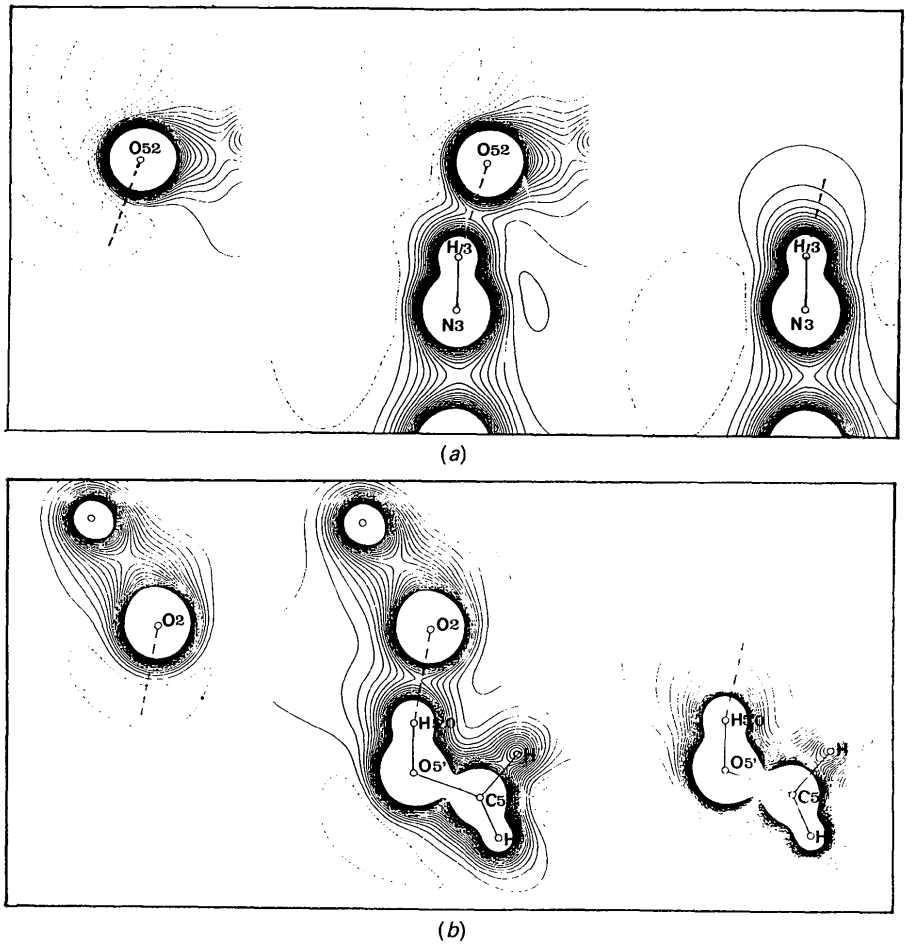


Fig. 10. Electrostatic potential distribution of the hydrogen bonds using the results from the *POP* refinement: (*a*) N(3)—H(3)···O(5'2), (*b*) O(5')—H(5'1)···O(2). Contours as in Fig. 8.

hydrogen bond, the negative electrostatic potential points away from the direction of the donor. The large negative peak in the electrostatic potential distribution around O(52) is associated with the large region of negative electrostatic potential generated by the negative charges on atoms O(4), N(3) and O(52). In summary, it would appear that both molecular-packing considerations and the distributions of electrostatic potential near hydrogen-bond acceptors affect the geometry of hydrogen bonds.

The authors thank the University of Regina for provision of computing time and for the provision of operating funds, and the reviewers for constructive comments. This research was supported by a postdoctoral research fellowship award to YW by the Saskatchewan Health Research Board.

References

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

AYISI, N. K., GUPTA, V. S. & BABIUK, L. A. (1985). *Antiviral Res.* **5**, 13–27.

COPPENS, P. (1989). *J. Phys. Chem.* **93**, 7979–7984.

CRAVEN, B. M., WEBER, H. P. & HE, X. M. (1987). *The POP Refinement Procedure.* Technical Report. Department of Crystallography, Univ. of Pittsburgh, PA, USA.

CROMER, D. T. & MANN, J. B. (1968). *Acta Cryst.* A**24**, 321–325.

CROMER, D. T. & WABER, J. T. (1974). *International Tables for X-ray Crystallography*, Vol. IV, edited by J. A. IBERS & W. C. HAMILTON, pp. 103–146. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

GABE, E. J., LARSON, A. C., LEE, F. L. & WANG, Y. (1979). *The NRCC PDP-8a Crystal Structure System.* Ottawa: National Research Council of Canada.

GUPTA, V. S., TOURIGNY, G., STUART, A. L., DECLERCQ, E., QUAIL, J. W., EKIEL, I., EL-KABBANI, O. A. L. & DELBAERE, L. T. J. (1987). *Antiviral Res.* **7**, 69–77.

HANSEN, N. K. & COPPENS, P. (1978). *Acta Cryst.* A**34**, 909–921.

HE, X. M. (1984). PhD dissertation. Univ. of Pittsburgh, PA, USA.

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

HIRSHFELD, F. L. (1971). *Acta Cryst.* B**27**, 769–781.

KLEIN, C. L., MAJESTE, R. J. & STEVENS, E. D. (1987). *J. Am. Chem. Soc.* **109**, 6675–6681.

KLEIN, C. L. & STEVENS, E. D. (1988*a*). *Acta Cryst.* B**44**, 50–55.

KLEIN, C. L. & STEVENS, E. D. (1988*b*). *Mol. Struct. Energy*, **7**, 25–64.
KLEIN, C. L., STEVENS, E. D., ZACHARIAS, D. E. & GLUSKER, J. P. (1987). *Carcinogenesis*, **8**, 5–18.
LECOMTE, C. (1991). *The Application of Charge Density Research to Chemistry and Drug Design*, edited by G. A. JEFFREY & J. F. PINIELLA, pp. 121–153. New York: Plenum Press.
PEARLMAN, D. A. & KIM, S.-H. (1985). *Biopolymers*, **24**, 327–357.
POLITZER, P. & MURRAY, J. S. (1990). *Theoretical Biochemistry and Molecular Biophysics*, edited by D. L. BEVERIDGE & R. LAVERY, pp. 165–191. New York: Adenine Press.
POLITZER, P. & MURRAY, J. S. (1993). *Supplement E: The Chemistry of the Hydroxyl, Ether and Peroxide Groups*, Vol. 2, ch. 1, edited by S. PATAI. Chichester, England: John Wiley.
QUAIL, J. W., TOURIGNY, G., DELBAERE, L. T. J., EL-KABBANI, O. A. L., STUART, A. L. & GUPTA, S. V. (1988). *Acta Cryst.* C**44**, 150–154.
SINGH, U. C. & KOLLMAN, P. A. (1984). *J. Comp. Chem.* **5**, 129–145.
SOUHASSOU, M., LECOMTE, C., BLESSING, R. H., AUBRY, A., ROHMER, M.-M., WIEST, R., BÉNARD, M. & MARRAUD, M. (1991). *Acta Cryst.* B**47**, 253–266.
SPACKMAN, M. A. & STEWART, R. F. (1981). *Chemical Applications of Atomic and Molecular Electrostatic Potentials*, edited by P. POLITZER & D. G. TRUHLER, pp. 407–426. New York: Plenum Press.
STEVENS, E. D. (1981). *J. Am. Chem. Soc.* **103**, 5087–5095.
STEVENS, E. D. (1990). *The AMANDA Least-Squares Refinement Program. A Modified Version of MOLLY*. Personal communication.
STEWART, J. M., HALL, S. R., ALDEN, R. A., OLTHOF-HAZEKAMP, R. & DOHERTY, T. M. (1983). Editors. *The XTAL System for Crystallographic Programs*. Technical Report TR-1364.2. Computer Science Center, Univ. of Maryland, College Park, Maryland, USA.
STEWART, R. F. (1976). *Acta Cryst.* A**32**, 565–574.
STEWART, R. F. (1982). *God. Jugosl. Cent. Kristallogr.* **17**, 1–24.
STEWART, R. F. (1991). *The Application of Charge Density Research to Chemistry and Drug Design*, edited by G. A. GEFFREY & J. F. PINIELLA, pp. 63–101. New York: Plenum Press.
STUART, A. L., AYISI, N. K., TOURIGNY, G. & GUPTA, V. S. (1985). *J. Pharm. Sci.* **74**, 246–249.
TAYLOR, R. & KENNARD, O. (1984). *Acc. Chem. Res.* **17**, 320–326.
WOLFE, S. (1972). *Acc. Chem. Res.* **5**, 102–111.

Acta Cryst. (1994). B**50**, 174–181

X-ray Diffraction Study of the Crystal Structure of the π-Molecular Compound Pyrene···Pyromellitic Dianhydride at 19 K

BY F. H. HERBSTEIN†

Department of Chemistry, Technion-Israel Institute of Technology, Haifa 32000, *Israel*

AND R. E. MARSH AND S. SAMSON

Noyes Laboratories of Chemical Physics, California Institute of Technology, Pasadena, CA 91125, *USA*

(*Received* 2 *August* 1993; *accepted* 21 *October* 1993)

Abstract

The crystal structure of the pyrene···pyromellitic dianhydride (PMDA) π-molecular compound [($C_{16}H_{10}$:$C_{10}H_2O_6$); PYRPMA] has been refined from intensities measured at 19 K using the low-temperature accessory designed by Samson, Goldish & Dick [*J. Appl. Cryst.* (1980), **13**, 425–432] for a four-circle diffractometer. Earlier results for the ordered structure [Herbstein & Snyman (1969). *Philos. Trans. R. Soc. London Ser. A*, **264**, 635–666] are confirmed and extended; at 19 K, a = 13.664 (3), b = 9.281 (2), c = 14.420 (3) Å, β = 91.80 (2)°, space group $P2_1/n$, Z = 4, with two sets of pyrenes at independent centres of symmetry and the four PMDAs at general positions. The geometrical structures of the two components are in good agreement with quantum mechanical calculations. Analyses of thermal motion and packing show that one set of pyrenes is more tightly packed than the other; the principal interactions in the crystal are π-π* plane-to-plane interactions between pyrene and PMDA and >CH···O=C<, between pyrene and PMDA, and between PMDAs.

1. Introduction

The mixed-stack π-π* charge-transfer molecular compound pyrene···pyromellitic dianhydride (PMDA) [($C_{16}H_{10}$:$C_{10}H_2O_6$); Refcode PYRPMA] undergoes a dis-order-to-order transition at ~160 K. Earlier (photographic) X-ray studies (Herbstein & Snyman, 1969) showed that the space group was $P2_1/a$ (Z = 2) above ~160 K, with pyrenes and PMDAs at independent centres of symmetry [present measurements give a = 13.94 (1), b = 9.34 (1), c = 7.31 (1) Å, β = 93.65 (9)° at 295 K]. Below ~160 K, the c axis was found to be doubled and the space group changed to $P2_1/n$ (Z = 4), with two pairs of pyrenes at independent centres and the four PMDA molecules in the cell at general

† Experimental work performed while on sabbatical leave at Caltech.