Acta Cryst. (1978). B34, 2556–2565

X-ray Studies of Crystalline Complexes Involving Amino Acids. III. The Structure of the Twinned Pseudosymmetric Crystals of a Complex Between Histidine and Aspartic Acid*

BY T. N. BHAT AND M. VIJAYAN†

Department of Physics and Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012, India

(*Received* 8 *September* 1977; *accepted* 27 *February* 1978)

L-Histidine–L-aspartic acid monohydrate, $C_6H_9N_3O_2.C_4H_7NO_4.H_2O$, crystallizes in the monoclinic space group $P2_1$ with eight formula units in a unit cell of dimensions $a = 5{\cdot}177\ (3)$, $b = 37{\cdot}214\ (20)$, $c = 13{\cdot}672\ (10)$ Å, and $\beta = 90°$. The crystals are twinned about the a or the c axes leading to an apparent orthorhombic symmetry for the diffraction pattern. The X-ray analysis was further complicated by the presence of non-crystallographic 2_1 screw axes and a pseudo translation of $c/2$. The structure was solved in several stages using photographic data and refined to an R value of 0·113 using a modified least-squares procedure. The crystallographic asymmetric unit of each twin component contains four molecules each of histidine, aspartic acid and water. One half of these are related to the other half by non-crystallographic 2_1 screw axes parallel to the a axis. In each half, the two aspartic acid molecules (and the two water molecules) are related to each other by a pseudo translation of $c/2$; the pseudo translation is only approximate for the histidine molecules. Disorder in the stacking of layers, described by a translation of $c/2$, parallel to the b axis also exists in the structure leading to streaks in the diffraction pattern along the row lines parallel to the b^* direction when l is odd. All the aspartic acid molecules in the structure have the same (and rather unusual) conformation with the side-chain carboxyl group *gauche* to both the α-amino and the α-carboxyl groups. Half the histidine molecules in the structure exist in an 'open' conformation whereas the other half exist in the 'closed' conformation. The crystal structure consists of alternating double layers, one double layer containing histidine molecules and the other aspartic acid molecules. The water molecules are sandwiched between the two layers in the aspartic acid double layer. The adjacent double layers are interconnected by hydrogen bonds between the amino and the imidazole groups of the histidine molecules on the one hand and the carboxyl groups of the aspartic acid molecules on the other.

Introduction

We have been carrying out a programme of X-ray analyses of crystalline complexes between amino acids in an attempt to study the atomic details of the non-covalent interactions involving them. As part of this programme, the crystal structures of lysine aspartate (Bhat & Vijayan, 1976) and arginine glutamate monohydrate (Bhat & Vijayan, 1977) have already been reported. The structure analysis of another such complex, histidine–aspartic acid monohydrate, $C_6H_9N_3O_2.C_4H_7NO_4.H_2O$, is presented here.

Experimental

Thin needle-like crystals of the complex were grown by the slow diffusion of acetone into an aqueous solution of histidine and aspartic acid in molar proportion. The symmetry and the unit-cell dimensions of the crystals were determined from X-ray diffraction photographs and the density was measured by flotation in a mixture of benzene and carbon tetrachloride.

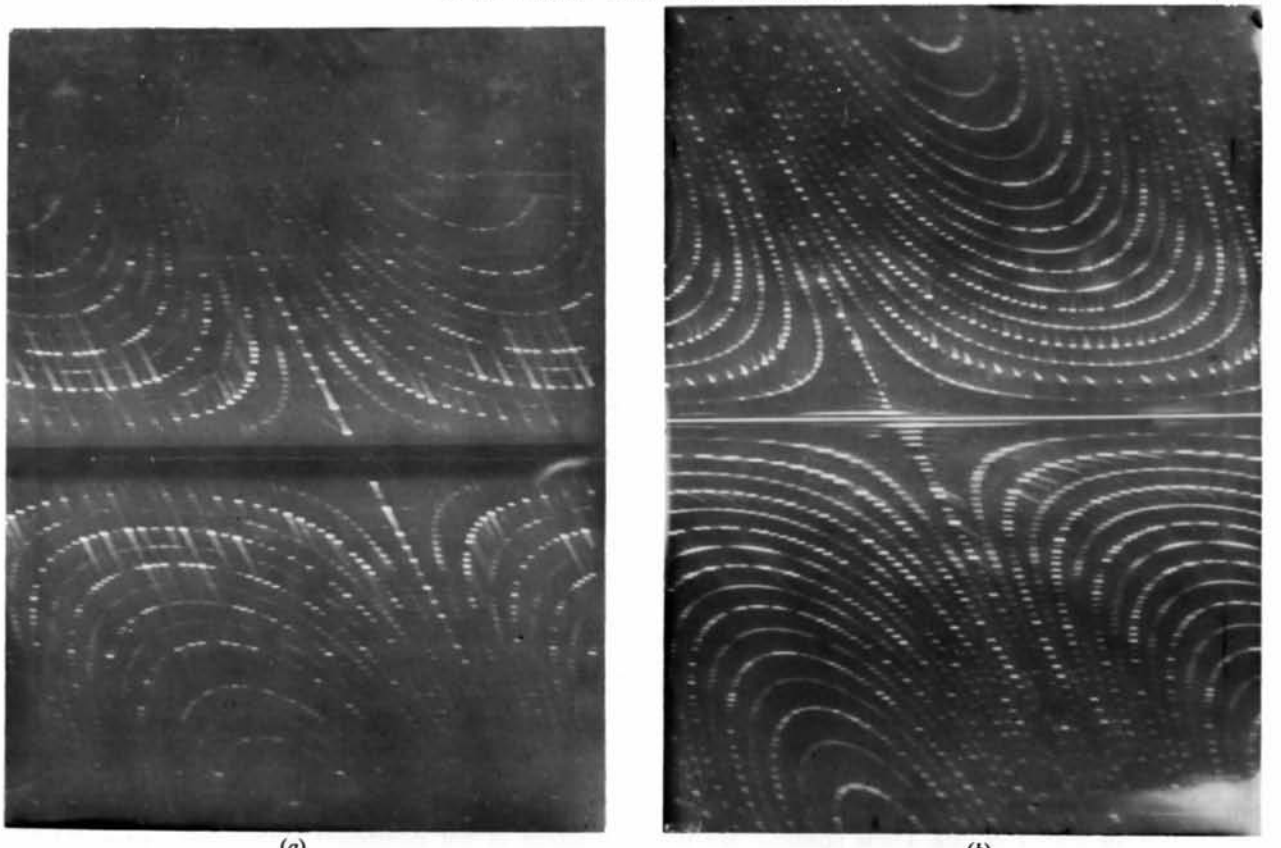
Crystal data

L-Histidine–L-aspartic acid monohydrate, $C_6H_9N_3O_2.C_4H_7NO_4.H_2O$, apparent space group $P2_12_12$, $a = 5{\cdot}177\ (3)$, $b = 37{\cdot}214\ (20)$, $c = 13{\cdot}672\ (10)$ Å, $D_m = 1{\cdot}547\ (8)$, $D_c = 1{\cdot}544$ g cm^{-3}, $Z = 8$, μ(Cu $K\alpha$) = 10·38 cm^{-1}.

The intensity data were recorded from a needle-like specimen of mean radius 0·3 mm using the multiple-film equi-inclination Weissenberg method (Cu $K\alpha$ radiation) for reciprocal levels Hkl, $H = 0$ through 4; the intensities were estimated visually. Of the 3499 possible independent reflections in the Cu sphere, 3029 were recorded, of which 1998 were in the measurable range. The data were corrected for Lorentz and polarization factors and for spot shape, but not for absorption ($\mu r = 0{\cdot}3$). The crystals were too thin and brittle to be cut along any axis other than the a axis,

* Contribution No. 113 from the Molecular Biophysics Unit, Indian Institute of Science, Bangalore, India.

† Author to whom correspondence should be addressed at the Molecular Biophysics Unit.



(*a*) (*b*)

Fig. 1. (*a*) 0*kl* and (*b*) 2*kl* Weissenberg photographs.

and hence cross-level data could not be collected. Therefore, absolute scale factors and overall temperature factors were determined individually for each level using Wilson's statistics. The temperature factors thus determined were nearly the same for all levels and their average value was taken as the overall temperature factor for the entire data. Normalized structure factors were calculated using this temperature factor, and the absolute scales were redetermined by setting $\langle E^2 \rangle = 1$ for each level.

As indicated earlier, the apparent space group of the crystals was $P2_12_12$. In addition to the systematic absences required by the space group, 00*l* reflections with $l = 4n + 2$ were also found to be absent. Normally, additional absences of this type point to the presence of features such as twinning. However, even after a detailed examination of the intensity data and the careful observation of the crystals under a polarizing microscope, no satisfactory twinned arrangement could be postulated to account for the symmetry of, and the systematic absences in, the diffraction pattern. Yet another, and the most striking, feature of the diffraction pattern was the difference in intensity between those reflections with even *l* values and those with odd. The latter were systematically weaker than the former, as can be seen from Fig. 1. In fact, the mean square *E* value of *l*-even reflections was 1·6 and that of *l*-odd reflections 0·3. The extreme weakness of *l*-odd reflections appeared to indicate the presence of a pseudo translation of $c/2$ within the unit cell, which was conceivable as there were two sets of chemically identical molecules in the asymmetric unit. Assuming this pseudo translation to be perfect as a first approximation and neglecting the *l*-odd reflections, one can arrive at an averaged half cell with a cell translation of $c/2$ along the *z* direction. Because of the additional absences among 00*l* reflections, the half cell had a space-group symmetry of $P2_12_12_1$.

Structure analysis

Determination of the average structure in the half cell

Several attempts were made to solve the structure in the full cell ($P2_12_12$) using the program *MULTAN* (Germain, Main & Woolfson, 1971). None of these attempts, in which $|E|$'s for *l*-odd and *l*-even reflections were partially or completely renormalized, were successful. The solution of the average structure in the half cell (*l*-odd reflections neglected; $l' = l/2$, $c' = c/2$; space group $P2_12_12_1$) was then attempted using *MULTAN*. The *E* map corresponding to the best solution was not readily interpretable. However, peaks corresponding to all the atoms in the structure, other than those belonging to the imidazole ring, could be identified with some difficulty on the basis of geometrical considerations. A difference Fourier map phased on these

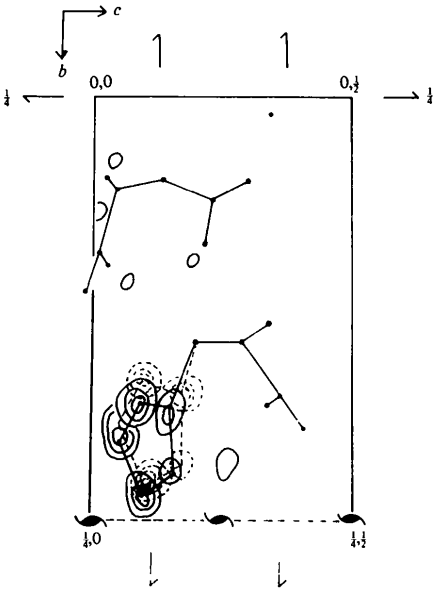


Fig. 2. Difference Fourier map phased on the atoms obtained from the E map. Full contours and broken contours represent different imidazole rings. The contours are spaced equally but start at an arbitrary value.

atoms was then calculated (Fig. 2). This map showed two sets of peaks corresponding to two different positions of the imidazole ring. These positions are also indicated in Fig. 2. The imidazole ring in the half cell was formally treated as 'disordered' and the ten atomic positions were each tentatively assigned occupancy factors of 0·5. The positional and isotropic thermal parameters were then refined using the full-matrix structure-factor least-squares method to an R value of 0·24.

The calculation of bond lengths and angles at this stage showed that the C^{β}–C^{γ} bond connecting one of the imidazole rings to the rest of the histidine molecule was abnormally long. Also, the atoms in the histidine molecule had rather high temperature factors. These features appeared to indicate that not only the imidazole ring but the entire histidine molecule was 'disordered' in the half cell. Therefore, each atom in the histidine molecule was assigned two positions by small displacements (determined by geometrical considerations) from the refined (and presumably mean) position. The two positions for each atom were assigned occupancy factors of 0·5. The subsequent refinement with single positions for the atoms in the aspartic acid molecule and the water O atom, and two positions of 0·5 occupancy for each of the atoms in the histidine molecule, converged at an R factor of 0·168. These positions are indicated in Fig. 3. There was no evidence, in terms of bond lengths, bond angles and temperature factors, to suggest any major 'disorder' in the aspartic acid molecule. Also a few cycles of block-diagonal structure-factor least-squares calculations were carried out to refine the occupancy factors of the atoms in the two 'disordered' histidine molecules. In these cycles, the temperature factors, which had comparable values in the two histidine molecules, were held constant. All the occupancy factors refined to values close to 0·5, indicating that the two sets of positions have, as originally assumed, equal occupancies.

Having determined the average structure in the half cell it might appear that the solution in the full cell could be achieved in a reasonably straightforward manner by placing two half cells side by side along the z direction and making sure that, in each row along the z direction, histidine molecule A was placed in one of the two half cells and molecule B in the other. Several such arrangements were possible as there were four rows of histidine molecules parallel to the z direction. However, none of these arrangements had $P2_12_12$ symmetry. Thus, the position at this stage of the analysis appeared to be the following: (*a*) The geometries of the histidine and the aspartic acid molecules in the crystals had been determined. All the aspartic acid molecules have nearly the same structure. There are two types of histidine molecules, one of type A and the other of type B (Fig. 3). (*b*) The average structure in the half cell has the space-group symmetry $P2_12_12_1$ with a 'disordered' histidine molecule. (*c*) The structure is not consistent with space group $P2_12_12$ even though the symmetry of the diffraction pattern and the systematic absences appeared to indicate the space group to be $P2_12_12$.

Complete structure solution

The successful refinement of the average structure implied that the arrangement of the aspartic acid and the water molecules could not be very different from

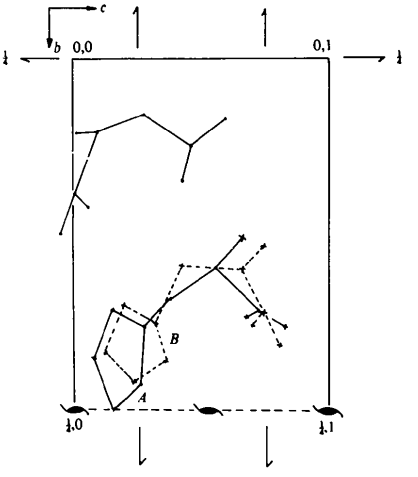


Fig. 3. The double histidine molecule and the single aspartic acid molecule observed in the half cell.

that obtained by a spatial extension of the arrangement shown in Fig. 3. However, the presence of two different types of histidine molecules introduced more than one possibility for the arrangement of these molecules in the crystal structure. These possibilities can be understood with reference to Fig. 4, which shows the distribution of equivalent positions in the full cell consistent with the space-group symmetry ($P2_12_12_1$) of the half cell. The coordinates of the equivalent positions are as follows: (1) x,y,z; (2) $x, y, z+\frac{1}{2}$; (3) $x+\frac{1}{2}, -y+\frac{1}{2}, -z+\frac{1}{2}$; (4) $x+\frac{1}{2}, -y+\frac{1}{2}, -z$; (5) $-x, y+\frac{1}{2}, -z+\frac{3}{4}$; (6) $-x, y+\frac{1}{2}, -z+\frac{1}{4}$; (7) $-x+\frac{1}{2}, -y, z-\frac{1}{4}$; (8) $-x+\frac{1}{2}, -y, z+\frac{1}{4}$. Each position could be occupied by a molecule of type A or of type B, or by both with partial occupancies. If the occupancies of molecule A and molecule B at position n are designated by P_{An} and P_{Bn}, the fact that molecules A and B have equal (0·5) occupancies in the average structure in the half cell requires that

$$P_{A1} + P_{A2} = P_{B1} + P_{B2} = 1$$

$$P_{A3} + P_{A4} = P_{B3} + P_{B4} = 1$$

$$P_{A5} + P_{A6} = P_{B5} + P_{B6} = 1$$

and

$$P_{A7} + P_{A8} = P_{B7} + P_{B8} = 1.$$

In order to leave no vacant space in any unit cell, it is also reasonable to assume that

$$P_{An} + P_{Bn} = 1$$

at every position. Thus, the number of independent parameters defining the occupancies reduced to four. Structure factors of a set of randomly selected reflections were then calculated for different possible combinations of occupancies. A careful examination of the results of these calculations indicated that the correct solution was likely to correspond to one of the following four independent combinations:

$$(1)\ P_{A1} = P_{A3} = P_{A5} = 1, \quad P_{A7} = 0$$

$$(2)\ P_{A1} = P_{A5} = P_{A7} = 1, \quad P_{A3} = 0$$

$$(3)\ P_{A1} = P_{A5} = 1, \quad P_{A3} = P_{A7} = 0$$

$$(4)\ P_{A1} = P_{A7} = 1, \quad P_{A3} = P_{A5} = 0.$$

These combinations will be referred to below as 1358, 1457, 1458 and 1467 respectively. It should also be noticed that these combinations do not involve fractional occupancies and hence correspond to ordered structures.

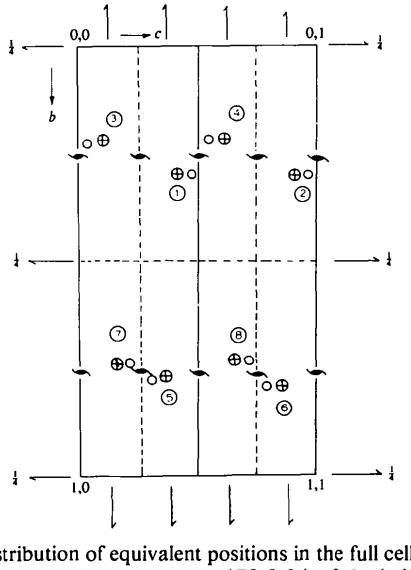


Fig. 4. Distribution of equivalent positions in the full cell consistent with the space-group symmetry ($P2_12_12_1$) of the half cell. Each position can be occupied by either (or both with fractional occupancy) of the two histidine molecules.

An examination of the distribution of equivalent positions in the four combinations showed the corresponding structures to be monoclinic, space group $P2_1$, with the same cell parameters as those of the full cell. In 1358 and 1457, the crystallographic 2_1 screw axis was parallel to the b axis. In the former, (1) was related to (5), and (3) was related to (8) by this 2_1 screw axis. In addition, (1) was related to (3) and (5) was related to (8) by a non-crystallographic 2_1 screw axis parallel to a. Identical relations existed among the molecules of type B at positions (2), (4), (6) and (7). A similar symmetry relationship existed for 1457. 1458 and 1467 were also monoclinic, space group $P2_1$, but with the 2_1 screw axis parallel to the a axis. Again, non-crystallographic 2_1 screw axes were present, but they were parallel to the b axis. In fact, the distributions of molecules in 1458 and 1467 were closely analogous to those in 1358 and 1457, except for the interchange of the directions of crystallographic and non-crystallographic 2_1 screw axes.

In the X-ray diffraction pattern from monoclinic crystals $|F_{hkl}| \neq |F_{hk\bar{l}}|$, when b or a is the unique axis. When they are equal, as in the present case, the equality is likely to have been caused by non-crystallographic symmetry. Therefore, hkl and $hk\bar{l}$ reflections were treated as formally independent (except in the case of those zonal reflections for which monoclinic symmetry required them to be equivalent) in the subsequent structure-factor least-squares calculations though the magnitudes of the observed structure factors were the same for a pair of reflections hkl and $hk\bar{l}$.

Structure-factor least-squares refinement of atomic parameters was then carried out for all four combinations. In these calculations the non-crystallographic symmetry was not made use of, and the positional coordinates and the isotropic temperature factors of all 84 atoms in the asymmetric unit (of the respective monoclinic cell) were refined. The refinement converged at $R = 0{\cdot}23$ for 1358; the final R values for the other three combinations varied from 0·25 to 0·28. These results indicated that 1358 was perhaps mar-

ginally superior to the other combinations. No further reduction of R was, however, possible.

It was noticed in all four cases that the calculated structure factors of reflections with even l very nearly retained the orthorhombic symmetry ($|F_{chkl}| = |F_{chk\bar{l}}|$, in addition to the other relations characteristic of the monoclinic cell) of the observed diffraction pattern. In fact, the calculated structure factors were nearly the same for these reflections in the four sets. In each set, the R factor was in the region of 0·15 when l-even reflections alone were considered. On the other hand, the orthorhombic symmetry of the observed pattern was not retained by the calculated structure factors of l-odd reflections. Also, the R factors for these reflections alone were in the region of 0·35. Thus, no possible ordered arrangement of molecules could be found to explain the orthorhombic symmetry of the entire diffraction pattern. Extensive calculations were also made for several disordered structures (with partial occupancies of histidine molecules at different positions), but these did not yield satisfactory results either.

Twinning

A careful examination of the different sets of calculations revealed that the relationship

$$|F_{chkl}|^2 + |F_{chk\bar{l}}|^2 = 2|F_{ohkl}|^2$$

occurred systematically among the structure factors for the combination 1358, but not among those for the other possible combinations. This relation could be readily explained if the monoclinic crystals with a structure corresponding to the combination 1358 were

Table 1. *Final positional parameters (in fractional coordinates* $\times 10^4$*) with their standard deviations in parentheses*

	x	y	z		x	y	z
O(1)	9533 (24)	3072 (4)	2096 (10)	O(1)*S*	14475 (18)	1937 (3)	5382 (7)
O(2)	13240 (21)	3136 (3)	2893 (8)	O(2)*S*	18189 (19)	1873 (3)	4596 (9)
N(1)	7314 (25)	3657 (4)	2863 (10)	C(1)*S*	15805 (23)	1781 (3)	4762 (10)
C(1)	10846 (21)	3217 (3)	2764 (9)	N(1)*S*	12349 (20)	1339 (3)	4612 (7)
C(2)	9511 (23)	3491 (3)	3406 (9)	C(2)*S*	14563 (20)	1517 (2)	4075 (7)
C(3)	8454 (21)	3293 (4)	4330 (9)	C(3)*S*	13355 (23)	1713 (3)	3131 (9)
C(4)	10332 (16)	3047 (2)	4821 (7)	C(4)*S*	15233 (24)	1953 (3)	2696 (8)
N(5)	12243 (24)	3155 (4)	5381 (9)	N(5)*S*	17303 (23)	1834 (3)	2101 (10)
C(6)	13675 (22)	2884 (3)	5733 (8)	C(6)*S*	18673 (21)	2129 (3)	1785 (8)
N(7)	12693 (17)	2577 (2)	5280 (7)	N(7)*S*	17719 (17)	2422 (2)	2216 (7)
C(8)	10537 (16)	2676 (2)	4743 (7)	C(8)*S*	15606 (27)	2324 (4)	2736 (11)
O(11)	3173 (21)	3806 (3)	6278 (8)	O(11)*S*	8186 (22)	1186 (3)	1220 (9)
O(12)	−920 (23)	3961 (3)	6018 (9)	O(12)*S*	4089 (21)	1043 (3)	1478 (8)
N(11)	5071 (18)	4472	6102 (7)	C(11)*S*	6536 (23)	974 (3)	1435 (9)
C(11)	1506 (18)	4028 (3)	6038 (7)	N(11)*S*	10087 (22)	528 (3)	1405 (9)
C(12)	2206 (21)	4411 (3)	5808 (8)	C(12)*S*	7259 (21)	574 (3)	1696 (8)
C(13)	1695 (16)	4544 (2)	4786 (7)	C(13)*S*	6695 (23)	459 (4)	2724 (9)
C(14)	3511 (21)	4401 (3)	3986 (8)	C(14)*S*	8523 (20)	606 (3)	3504 (8)
O(15)	5249 (17)	4184 (2)	4142 (7)	O(15)*S*	10251 (19)	813 (3)	3336 (7)
O(16)	3262 (17)	4542 (3)	3113 (7)	O(16)*S*	8263 (20)	461 (3)	4366 (8)
W(1)	−1521 (19)	4844 (3)	2764 (7)	*W*(1)*S*	3462 (25)	140 (4)	4731 (11)
O(21)	8715 (19)	2952 (3)	7251 (7)	O(21)*S*	13751 (33)	2038 (5)	280 (12)
O(22)	12563 (23)	3057 (3)	7882 (9)	O(22)*S*	17614 (24)	1931 (3)	−417 (8)
C(21)	10238 (18)	3146 (3)	7710 (8)	C(21)*S*	15289 (23)	1841 (3)	−202 (9)
N(21)	6632 (22)	3590 (3)	7763 (10)	N(21)*S*	11634 (22)	1413 (3)	−280 (9)
C(22)	9322 (18)	3493 (3)	8106 (7)	C(22)*S*	14317 (21)	1511 (3)	−649 (8)
C(23)	9459 (33)	3499 (5)	9261 (14)	C(23)*S*	14508 (25)	1510 (4)	−1770 (10)
C(24)	7718 (19)	3208 (3)	9696 (7)	C(24)*S*	12659 (22)	1784 (3)	−2205 (9)
N(25)	5695 (22)	3301 (4)	10325 (10)	N(25)*S*	10719 (28)	1694 (4)	−2828 (11)
C(26)	4359 (21)	3000 (2)	10671 (6)	C(26)*S*	9367 (26)	1996 (4)	−3171 (10)
N(27)	5356 (15)	2721 (3)	10202 (8)	N(27)*S*	10450 (22)	2281 (3)	−2691 (8)
C(28)	7524 (21)	2844 (3)	9627 (9)	C(28)*S*	12554 (26)	2156 (4)	−2124 (10)
O(31)	3190 (25)	3843 (4)	11408 (11)	O(31)*S*	8160 (22)	1161 (3)	−3930 (8)
O(32)	−927 (18)	3993 (3)	11120 (7)	O(32)*S*	4035 (25)	1012 (3)	−3640 (10)
N(31)	5050 (19)	4489 (3)	11137 (8)	C(31)*S*	6410 (19)	952 (3)	−3680 (8)
C(31)	1454 (19)	4059 (3)	11188 (8)	N(31)*S*	10031 (27)	510 (4)	−3642 (11)
C(32)	2315 (21)	4439 (3)	10863 (8)	C(32)*S*	7209 (18)	570 (3)	−3356 (6)
C(33)	1703 (22)	4537 (3)	9812 (10)	C(33)*S*	6679 (25)	468 (4)	−2295 (10)
C(34)	3492 (19)	4367 (3)	8975 (8)	C(34)*S*	8390 (25)	642 (3)	−1470 (10)
O(35)	4902 (17)	4115 (2)	9211 (7)	O(35)*S*	9960 (25)	879 (4)	−1694 (11)
O(36)	3326 (21)	4510 (4)	8120 (9)	O(36)*S*	8275 (24)	498 (4)	−623 (9)
W(2)	−1472 (18)	4836 (3)	7776 (8)	*W*(2)*S*	3477 (26)	169 (4)	−269 (10)

assumed to be twinned about the a or c axes. In the twinned crystals, one component would then be related to the other by a twofold rotation axis parallel to the x or z directions. A monoclinic angle of 90° meant that the hkl reflection from one component superposed exactly over the $hk\bar{l}$ (or $\bar{h}kl$) reflection from the other. When the two components have equal weights, the observed diffraction pattern would have complete orthorhombic symmetry, as found in the present structure. Therefore, it appeared that twinning of the type outlined above was present in the crystals of histidine–aspartic acid monohydrate. The subsequent successful refinement of the twinned structure confirmed that this was indeed the case.

In the subsequent calculations, the origin of the unit cell was shifted by $-\frac{1}{8}c$ to make it lie on a 2_1 screw axis, as is conventionally done in monoclinic cells.

The positional and thermal parameters of the 84 non-hydrogen atoms were refined, first isotropically and then anisotropically, to an R factor of 0·113 for 3549 hkl and $hk\bar{l}$ reflections using the following procedure (Sudarsanan, Young & Donnay, 1973) and a block-diagonal *SFLS* program originally written by Dr R. Shiono and modified by one of us (TNB) to deal with twinned crystals. Structure factors were first calculated for hkl and $hk\bar{l}$ reflections and the observed structure factors were then assigned the values

$$|F_{ohkl}| = [(2|F_{chkl}|^2|F_{ohkl}|^2)/(|F_{chkl}|^2 + |F_{chk\bar{l}}|^2)]^{1/2}$$

and

$$|F_{ohk\bar{l}}| = [(2|F_{chk\bar{l}}|^2|F_{ohkl}|^2)/(|F_{chkl}|^2 + |F_{chk\bar{l}}|^2)]^{1/2}.$$

These values were then treated as the observed structure factors of hkl and $hk\bar{l}$ reflections. After each cycle of refinement, $|F_{ohkl}|$ and $|F_{ohk\bar{l}}|$ were re-estimated using the calculated structure factors based on the shifted parameters. The procedure was continued until convergence was reached. All the shifts were less than the corresponding standard deviations in the final cycle. The anisotropic thermal parameters were of the form $\exp[-(b_{11}h^2 + b_{22}k^2 + b_{33}l^2 + 2b_{12}hk + 2b_{13}hl + 2b_{23}kl)]$. The weighting scheme used in the final cycles had the form $1/(a + b|F_o| + c|F_o|^2)$, where $a = 2{\cdot}02$, $b = 0{\cdot}0614$ and $c = 0{\cdot}0094$. The atomic scattering factors used were those given by Cromer & Waber (1965). The final positional coordinates are given in Table 1.*

Disorder in stacking

It should be mentioned that the solution and the refinement of the structure were carried out on the assumption that the crystals were completely ordered. However, streaks along the row lines parallel to the b^* direction when l is odd observed in the diffraction pattern (see Fig. 1) indicated the presence of disorder in the stacking of layers parallel to the b axis. As can be seen from Fig. 5, the arrangement of molecules about the ac plane at $b = 0$ and $b = \frac{1}{2}$ is such as to permit a translation of $c/2$ of either one-half ($b = 0$ to $\frac{1}{2}$, or $\frac{1}{2}$ to 1) or the whole of a unit cell without substantial rearrangement in the rest of the structure, a situation analogous in part to that found in the structures of 1,8-diazacyclotetradecane-2,9-dione (Northolt & Alexander, 1971) and some adducts of diphenylmercury with bidentate ligands (Canty & Gatehouse, 1972). The disorder does not affect row lines when l is even, as it is described by a translation of $c/2$. A detailed quantitative discussion of the disorder is beyond the scope of the present paper. Also, it is believed that the disorder in stacking described above would not materially affect the structural features presented here.

Discussion

Description of the structure

The arrangement of molecules in one of the twin components is given in Fig. 5. The space group is $P2_1$ and the crystallographic asymmetric unit consists of four molecules each of histidine, aspartic acid and water. One half of the asymmetric unit, consisting of His A, His B, Asp A, Asp B, $W(1)$ and $W(2)$, is related to the other half, consisting of His AS, His BS, Asp AS, Asp BS, $W(1)S$ and $W(2)S$, by non-crystallo-

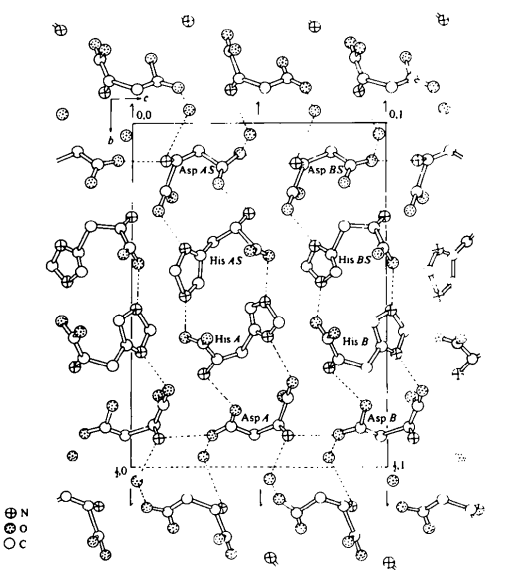


Fig. 5. Arrangement of molecules in the unit cell corresponding to the combination 1358.

* Lists of structure factors and thermal parameters have been deposited with the British Library Lending Division as Supplementary Publication No. SUP 33467 (20 pp.). Copies may be obtained through The Executive Secretary, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

graphic 2_1 screw axes parallel to the *a* axis at $y = \frac{1}{4}$, $z = \frac{3}{8}$ and $y = \frac{1}{4}$, $z = \frac{7}{8}$. A pseudo translation of $c/2$ also exists in the structure. The pseudo translation is nearly exact for the aspartic acid and water molecules; it is approximate for part of the histidine molecules and non-existent for the imidazole rings.

Conformation of histidine

The conformation of the histidine side chain can be described by two torsion angles, χ^1 and χ^{21} or χ^{22} (IUPAC–IUB Commission on Biochemical Nomenclature, 1970). Because of the planarity of the imidazole group, χ^{21} and χ^{22} differ by 180°. Steric considerations and energy calculations suggest a preferred value of $\pm 90°$ for χ^{21} and χ^{22} (Ponnuswamy & Sasisekharan, 1971). The δ N atom, however, is capable of taking part in metal coordination, hydrogen bonding and electrostatic interactions (when protonated). Therefore, considerable departures from the ideal values of $\pm 90°$ for χ^{21} and χ^{22} may be expected to occur in order to accommodate these interactions wherever possible. The torsion angle χ^1, which defines the disposition of the imidazole ring with respect to the rest of the molecule, can assume values in the neighbourhood of −60, 180 and 60° corresponding to the three staggered arrangements. Two of them lead to 'open' or 'extended' conformations whereas the third leads to a 'closed' or 'folded' conformation. In 'open conformation I', which is sterically the most favourable, the imidazole group is *trans* to the carboxyl group and *gauche* to the amino N atom ($\chi^1 \sim -60°$) whereas in 'open conformation II' the imidazole group is *trans* to the amino N atom and *gauche* to the carboxyl group ($\chi^1 \sim 180°$). In the sterically least favourable 'closed' conformation, however, the imidazole group is *gauche* to both the amino N atom and the carboxyl group ($\chi^1 \sim 60°$). Now, corresponding to each value of χ^1, there are two possible values of χ^{21}, one in the neighbourhood of 90° and the other in the neighbourhood of −90°, thus leading to a total of six possible conformations.

Among the histidine molecules in the present structure, His *A* and His *AS* occur in 'open conformation II' and His *B* and His *BS* in the 'closed' conformation, as can be seen from Fig. 6(*a*) and (*b*). The torsion angles that define the conformation of the histidine molecules are listed in Table 2, along with those in other crystal structures containing histidine. The torsion angle χ^{21} is positive in His *A* and His *AS* whereas it is negative in His *B* and His *BS*. The two sets of histidine molecules also differ substantially in the values of ψ. The values of ψ' in His *A* and His *AS* are −26 and −32° respectively; the corresponding values in His *B* and His *BS* are −8 and −12°.

As can be seen from Table 2, the structures of a number of crystals containing histidine have been determined and it is instructive to compare the conformations observed in them. Metal coordination imposes severe restrictions on the conformational possibilities and hence the metal complexes are not considered here. Of the six possible conformations, only four, namely, 'open conformation I' with $\chi^{21} \sim 90°$, 'open conformation I' with $\chi^{21} \sim -90°$, 'open conformation II' with $\chi^{21} \sim 90°$ and the 'closed' conformation with $\chi^{21} \sim -90°$, have so far been observed in the crystal structures of histidine, its salts and derivatives.

It is interesting to note that the sterically least

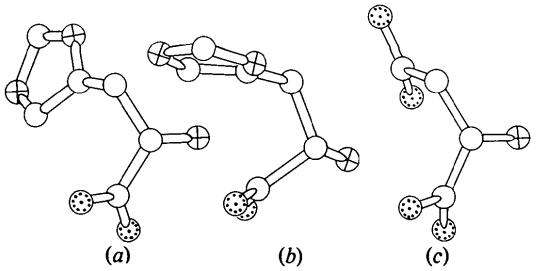


Fig. 6. Conformation of histidine and aspartic acid in the crystal structure: (*a*) 'closed' conformation of histidine (see text), (*b*) 'open' conformation of histidine (see text), (*c*) conformation of aspartic acid ($\chi^1 \simeq 60°$, $\chi^{21} \simeq 0°$).

Table 2. *Side-chain conformational angles* (°) *of histidine and aspartic acid observed in crystal structures other than metal complexes*

Compound	χ^1	χ^{21}	Reference
L-His ($P2_1$)	−57	53	(*a*)
L-His ($P2_12_12_1$)	−59	57	(*b*)
DL-His	−87	−68	(*c*)
L-*N*-Acetyl-His.H_2O			
Molecule (I)	80	−76	(*d*)
Molecule (II)	−63	−87	(*d*)
L-His.HCl.H_2O	71	−120	(*e*)
DL-His.HCl.$2H_2O$	−62	−71	(*f*)
L-His.2HCl	−53	−75	(*g*)
L-Pyro-Glu.L-His methyl ester	−180	57	(*h*)
L-His.L-Asp.H_2O			
His *A*	167	76	(*i*)
His *B*	65	−120	(*i*)
His *AS*	168	78	(*i*)
His *BS*	61	−120	(*i*)
L-Asp	−62	−51	(*j*)
DL-Asp	−62	3	(*k*)
L-Lys.L-Asp	−70	−38	(*l*)
L-His.L-Asp.H_2O			
Asp *A*	53	2	(*i*)
Asp *B*	51	14	(*i*)
Asp *AS*	49	3	(*i*)
Asp *BS*	54	6	(*i*)

References: (*a*) Madden, McGandy & Seeman (1972). (*b*) Lehmann, Koetzle & Hamilton (1972). (*c*) Edington & Harding (1974). (*d*) Kistenmacher, Hunt & Marsh (1972). (*e*) Oda & Koyama (1972). (*f*) Bennett, Davidson, Harding & Morelle (1970). (*g*) Kistenmacher & Sorrell (1973). (*h*) Cotrait & Allard (1973). (*i*) Present study. (*j*) Derissen, Endeman & Peerdeman (1968). (*k*) Rao (1973). (*l*) Bhat & Vijayan (1976).

favourable form ('closed' conformation) has not so far been found to occur when the imidazole group is neutral as in orthorhombic and monoclinic L-His, DL-His and L-pyro-Glu.His methyl ester. In L-His, in both crystalline modifications, an internal hydrogen bond exists between the δ N atom in the imidazole ring and the α N atom, with the latter as the proton donor and the former as the acceptor. In DL-His, however, the imidazole ring is rotated about the C^{α}–C^{β} bond by about 180° and the χ^{21} angle is nearly −68°. Thus, instead of the δ N atom, the δ C atom now faces the α-amino group and, hence, the internal hydrogen bond cannot be formed. The δ N atom is, however, involved in an intermolecular hydrogen bond in crystals of DL-His. Therefore, the energies gained from an internal hydrogen bond and an intermolecular hydrogen bond appear to be of similar magnitude. The two types of hydrogen bonds are mutually exclusive for the neutral imidazole ring of histidine and the choice between them is apparently dictated by the environment of the molecule.

The imidazole ring carries a positive charge in the hydrochlorides and in the two crystallographically independent molecules in the structure of *N*-acetyl-histidine monohydrate. The same is also probably true in the histidine molecules in the present structure. The observed conformations in these structures cannot be rationalized in a simple manner as they are likely to be determined not only by steric considerations but also by the intra- and intermolecular electrostatic interactions involving imidazole rings. Indeed, both the 'open' conformations and the 'closed' conformation are observed in these structures.

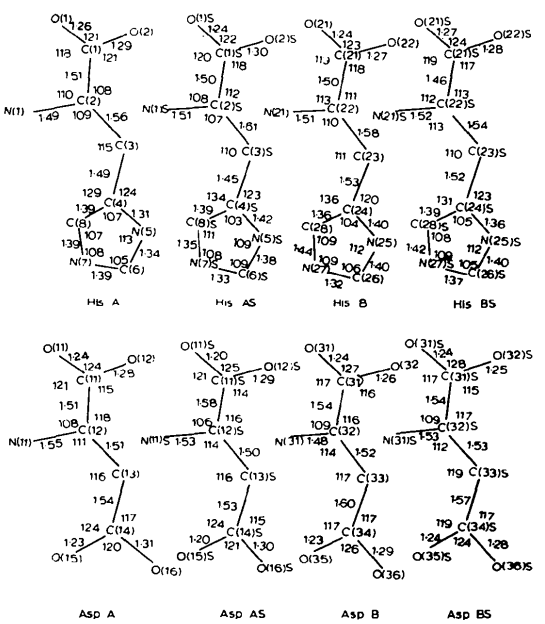


Fig. 7. Bond lengths (Å) and angles (°) in the structure. The estimated standard deviations in bond lengths and angles are approximately 0·02 Å and 1° respectively.

Conformation of aspartic acid

All four crystallographically independent aspartic acid molecules in the structure have the same conformation (Fig. 6*c*). The side-chain conformational angles, along with those observed in the other structures containing aspartic acid, are listed in Table 2. An interesting feature of the conformation in the present structure is that χ^1 has values in the neighbourhood of 60° with the side-chain carboxyl group *gauche* to the α-amino and the α-carboxyl groups. So far, a conformation of this type has only been observed in the metal complexes of aspartic acid where the requirement of tridentate chelation is satisfied only when $\chi^1 \sim 60°$. For a free molecule, this conformation is unfavourable, though not disallowed, sterically and electrostatically as it brings the two carboxyl groups into close proximity. Therefore, intermolecular interactions appear to be largely responsible for forcing the molecules into this conformation in the present structure.

Bond lengths and angles, possible disorder in the carboxyl groups, and the ionization state

The bond lengths and valency angles in the structure are given in Fig. 7. A detailed discussion of these lengths and angles is not warranted because their standard deviations (0·02 Å for bond lengths and 1° for valency angles) are rather high. On the whole these dimensions are comparable, within experimental error, with those found in the other structures containing histidine and aspartic acid.

One feature, which persisted throughout the course of the refinement of the structure, was the rather high temperature factors of the carboxyl O atoms compared with those of the other atoms in the structure. The equivalent isotropic temperature factors (Hamilton, 1959) of these O atoms vary between 3·0 and 5·7 $Å^2$ whereas the average value of the temperature factors of the other atoms in the histidine and aspartic acid molecules is 3·0 $Å^2$. In addition, the dimensions of many of the carboxyl groups were found to be between those expected in the neutral and deprotonated states, even after allowing for the high standard deviations of the bond lengths and angles in the structure. Thus it appeared that the carboxyl O atoms were probably disordered. The disorder was likely to have arisen from any given carboxyl group being neutral in some unit cells and deprotonated in others. The conformation, especially that of aspartic acid, and the distribution of the molecules in the structure also make such a situation plausible.

The low accuracy of the structure, the probable

disorder of the carboxyl O atoms and the non-availability of the positions of the H atoms make it difficult to unambiguously determine the ionization state of the molecules in the structure. However, as will be seen later, geometrical considerations strongly indicate that both the imidazole N atoms in all the histidine molecules are hydrogen-bonded to O atoms belonging to α-carboxyl groups (Fig. 5). Each of these O atoms is *cis* to the corresponding amino N atom and hence would remain unprotonated irrespective of the ionization state of the carboxyl group to which it belongs. The hydrogen bonds should thus be of the N–H···O type. Therefore, it appears to be highly probable that the imidazole groups are protonated.

Crystal structure and hydrogen bonding

The crystal structure of histidine–aspartic acid monohydrate is shown in Fig. 5. Table 3 lists the contact distances between atoms involved in what appear to be hydrogen bonds on geometrical considerations. In the absence of the availability of the positions of the H atoms and in view of the probable disorder in their positions, donors and acceptors are not specified in the table. The number of probable hydrogen bonds listed perhaps corresponds to a conservative estimate as, in almost all cases, a contact distance was called a hydrogen bond only if all the relevant geometrical criteria were satisfied even though considerable departures from ideal geometries are known to occur in many structures. All the hydrogen-bond parameters, except those involving the ε N atoms of histidine molecules, are normal. The $N^{\varepsilon}\cdots O$ bonds which connect the histidine molecules related by the non-crystallograpic screw axes are found to be extremely short.

The crystal structure can be described as consisting of alternating double layers, one double layer consisting of histidine molecules and the other of aspartic acid molecules, which sandwich water molecules. Both the double layers are parallel to the *ac* plane with the former centred on $b/4$ and $3b/4$, and the latter centred on 0 and $b/2$. In the histidine double layer, the *A* molecules form an infinite right-handed helix around a non-crystallographic 2_1 screw axis with pitch *a* and unit twist 180°. In this helix, successive histidine molecules are connected by N(7)–O(1) hydrogen bonds. The helix is further stabilized by N(1)–O(2) hydrogen bonds between molecules related by a translation *a* along the helix axis. Similarly, the *B* molecules form another helix around the other non-crystallographic 2_1 screw axis with the same helical parameters and hydrogen bonds. The latter is, however, left-handed. In the aspartic acid double layer, the two layers are not directly interconnected. Each layer forms an infinite sheet stabilized by N···O hydrogen bonds, parallel to the *bc* plane. The two sheets are then interconnected through the water molecules to form a sandwiched double layer. The histidine and the aspartic acid double layers are interconnected by hydrogen bonds involving the amino and the δ N atoms of the histidine molecules on the one hand and the carboxyl groups of the aspartic acid molecules on the other.

It is interesting to note that the basic and the acidic amino acids first aggregate independently into separate layers as in the case of lysine aspartate and arginine glutamate monohydrate. The layers then come together in an alternating fashion to form the crystal. The structures of lysine aspartate and arginine glutamate monohydrate consisted of alternating single layers whereas the present structure is made up of alternating double layers.

Table 3. *Hydrogen-bond contact distances* (Å)

A···*B*		*A*···*B*		*A*···*B*	
N(1)···O(15)$^{(a)}$	2·84	N(1)···O(32)$^{(d)}$	2·84	N(1)···O(2)$^{(b)}$	2·87
N(11)···O(12)$^{(c)}$	2·82	N(11)···*W*(1)*S*$^{(g)}$	2·84	N(11)···O(36)$^{(a)}$	2·91
N(21)···O(22)$^{(b)}$	2·90	N(21)···O(35)$^{(a)}$	2·92		
N(31)···O(32)$^{(c)}$	2·78	N(31)···O(16)$^{(f)}$	2·86	N(31)···*W*(2)*S*$^{(g)}$	2·90
N(5)···O(11)$^{(c)}$	2·76	N(7)···O(1)*S*$^{(a)}$	2·56	N(25)···O(31)$^{(a)}$	2·82
N(27)···O(21)*S*$^{(e)}$	2·68	*W*(1)···O(16)$^{(a)}$	2·76	*W*(2)···O(36)$^{(a)}$	2·80
N(1)*S*···O(15)*S*$^{(a)}$	2·84	N(1)*S*···O(2)*S*$^{(b)}$	2·93	N(1)*S*···O(32)*S*$^{(k)}$	2·82
N(11)*S*···O(12)*S*$^{(c)}$	2·82	N(11)*S*···*W*(2)$^{(h)}$	2·90	N(11)*S*···O(36)*S*$^{(a)}$	2·93
N(21)*S*···O(22)*S*$^{(b)}$	2·84	N(21)*S*···O(35)*S*$^{(a)}$	2·91		
N(31)*S*···O(32)*S*$^{(c)}$	2·79	N(31)*S*···O(16)*S*$^{(f)}$	2·88	N(31)*S*···*W*(1)$^{(j)}$	2·86
N(5)*S*···O(11)*S*$^{(c)}$	2·73	N(7)*S*···O(1)$^{(c)}$	2·60	N(25)*S*···O(31)*S*$^{(a)}$	2·82
N(27)*S*···O(21)$^{(i)}$	2·65	*W*(1)*S*···O(16)*S*$^{(a)}$	2·80	*W*(2)*S*···O(36)*S*$^{(a)}$	2·81

Symmetry code

(*a*) *x*,*y*,*z*
(*b*) $x-1, y, z$
(*c*) $x+1, y, z$
(*d*) $x+1, y, z-1$
(*e*) $x-1, y, z+1$
(*f*) $x, y, z+1$
(*g*) $-x+1, y+\frac{1}{2}, -z+1$
(*h*) $-x+1, y-\frac{1}{2}, -z+1$
(*i*) $x, y, z-1$
(*j*) $-x+1, y-\frac{1}{2}, -z$
(*k*) $x+1, y, z+1$.

The authors thank Professor G. N. Ramachandran for his interest. Financial support from the University Grants Commission, India, and the Department of Science and Technology, India, is gratefully acknowledged. One of us (TNB) thanks the department of Atomic Energy, India, for the award of a research fellowship.

References

BENNETT, I., DAVIDSON, A. G. H., HARDING, M. M. & MORELLE, I. (1970). *Acta Cryst.* B**26**, 1722–1729.
BHAT, T. N. & VIJAYAN, M. (1976). *Acta Cryst.* B**32**, 891–895.
BHAT, T. N. & VIJAYAN, M. (1977). *Acta Cryst.* B**33**, 1754–1759.
CANTY, A. J. & GATEHOUSE, B. M. (1972). *Acta Cryst.* B**28**, 1872–1888.
COTRAIT, M. & ALLARD, M. (1973). *C. R. Acad. Sci. Sér. C*, **276**, 1671–1674.
CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104–109.
DERISSEN, J. L., ENDEMAN, H. J. & PEERDEMAN, A. F. (1968). *Acta Cryst.* B**24**, 1349–1354.
EDINGTON, P. & HARDING, M. M. (1974). *Acta Cryst.* B**30**, 204–206.
GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.
HAMILTON, W. C. (1959). *Acta Cryst.* **12**, 609–610.
IUPAC–IUB COMMISSION ON BIOCHEMICAL NOMENCLATURE (1970). *J. Mol. Biol.* **52**, 1–17.
KISTENMACHER, T. J., HUNT, D. J. & MARSH, R. E. (1972). *Acta Cryst.* B**28**, 3352–3361.
KISTENMACHER, R. J. & SORRELL, T. (1973). *Cryst. Struct. Commun.* **2**, 673–676.
LEHMANN, M. S., KOETZLE, T. F. & HAMILTON, W. C. (1972). *Int. J. Pept. Protein Res.* **4**, 229–239.
MADDEN, J. J., MCGANDY, E. L. & SEEMAN, N. C. (1972). *Acta Cryst.* B**28**, 2377–2382.
NORTHOLT, M. G. & ALEXANDER, L. E. (1971). *Acta Cryst.* B**27**, 523–530.
ODA, K. & KOYAMA, H. (1972). *Acta Cryst.* B**28**, 639–642.
PONNUSWAMY, P. K. & SASISEKHARAN, V. (1971). *Int. J. Protein Res.* **3**, 9–18.
RAO, S. T. (1973). *Acta Cryst.* B**29**, 1718–1720.
SUDARSANAN, K., YOUNG, R. A. & DONNAY, J. D. H. (1973). *Acta Cryst.* B**29**, 808–814.

Acta Cryst. (1978). B**34**, 2565–2570

Order–Disorder Phenomena in Structures of Carboxylic Acids: The Structures of Fluoromalonic Acid and Hydroxymalonic Acid at 20 and −150°C

BY G. ROELOFSEN, J. A. KANTERS, J. KROON, H. M. DOESBURG AND T. KOOPS

Structural Chemistry Group, Rijksuniversiteit Utrecht, Transitorium 3, Padualaan 8, Utrecht, The Netherlands

(*Received* 31 *January* 1978; *accepted* 21 *March* 1978)

Fluoromalonic acid ($C_3H_3FO_4$) crystallizes in space group *Pnam* ($Z = 4$) with $a = 4{\cdot}593$ (1), $b = 8{\cdot}329$ (2), $c = 11{\cdot}219$ (2) Å at room temperature, and $a = 4{\cdot}576$ (1), $b = 7{\cdot}874$ (1), $c = 11{\cdot}321$ (2) Å at liquid-nitrogen temperature. At room temperature the carboxyl group is disordered (C–O 1·253 and 1·257 Å; C–C–O 116·6 and 116·8°), and this is associated with a splitting of the electron density of the carboxylic H atom. This disorder markedly decreases on cooling (C–O 1·292, C=O 1·221 Å; C–C–O 113·6, C–C=O 119·8°) and the electron-density splitting diminishes. At liquid-nitrogen temperature the C=O bond is antiplanar with the C–F bond. Hydroxymalonic acid (tartronic acid, $C_3H_4O_5$) crystallizes in space group $P2_12_12_1$ ($Z = 4$) with $a = 4{\cdot}494$ (1), $b = 8{\cdot}819$ (2), $c = 10{\cdot}882$ (3) Å at room temperature and $a = 4{\cdot}399$ (1), $b = 8{\cdot}731$ (2), $c = 10{\cdot}905$ (3) Å at liquid-nitrogen temperature. The molecular geometries of the structures at both temperatures are equal within experimental error. The C=O bonds are synplanar with the aliphatic C–O bond.

Introduction

As a result of crystallographic study, the C–O distances and C–C–O angles in fluoromalonic acid at room temperature (RT) were found to be equal within experimental error (Roelofsen, Kanters, Kroon & Vliegenthart, 1971). The equality of C–O distances and C–C–O angles has been found in many free carboxylic acids (*e.g.* Manojlović & Speakman, 1967; Housty, 1968; Filipakis, Leiserowitz, Rabinovich & Schmidt, 1972). These findings suggest that the packing of the carboxyl groups is disordered and imply that at least two conformations (syn- and antiplanar) are present (Leiserowitz, 1976). From an analysis of X-ray