

## The Crystal Structure of the Orthorhombic Form of L-(+)-Histidine

BY JOHN J. MADDEN,\* EDWARD L. MCGANDY AND NADRIAN C. SEEMAN†

Crystallography Department, University of Pittsburgh, Pittsburgh, Pennsylvania 15213, U.S.A.

(Received 23 August 1971)

L-(+)-Histidine ( $C_6N_3O_2H_9$ ) crystallizes in the orthorhombic space group  $P2_12_12_1$ , with  $a=5.177$ ,  $b=7.322$ ,  $c=18.87$  Å, and  $Z=4$ . Data were collected with Mo  $K\alpha$  radiation, using balanced filters. The structure was solved by direct phasing methods and refined to a final agreement index of 0.034 for all reflections. The conformation of the molecule is that of the open, extended form, and is stabilized principally by an intramolecular hydrogen bond between the amino nitrogen atom and the adjacent imidazole nitrogen atom. Where this conformation is found in proteins, it is likely to reduce the chemical reactivity of that imidazole group, because one of the imidazole nitrogen atoms is sterically hindered by the peptide backbone.

### Introduction

The amino acid L-histidine (Fig. 1) has been extensively studied because of the ability of its imidazole moiety to act as a proton donor, a proton acceptor, and a nucleophilic reagent. As the free amino acid, histidine catalyzes the degradation of various esters (Rohling & Fox, 1967), while in polypeptides, histidine has been implicated in the mechanism of a number of enzymes, most notably bovine pancreatic ribonuclease (Meadows, Roberts & Jardetsky, 1969) and  $\alpha$ -chymotrypsin (Schoellman, Schoellman & Shaw, 1963). In an effort to relate the conformation of histidine to its reactivity in these systems, its structure has been studied by X-ray crystallography in several metal complexes (Freeman, 1967) and as the protonated hydrochloride (Donohue, Lavine & Rollett, 1956; Donohue & Caron, 1964; Bennett, Davidson, Harding & Morelle, 1970). We have investigated the free base form of L-histidine by X-ray crystallography to determine the conformation of the molecule in the absence of ionic ligands.

### Experimental

#### Preparation of the crystals

Crystals of histidine were prepared by slow evaporation of an aqueous solution of L-histidine (Nutritional Biochemical Corporation), and were then dried over calcium chloride. A large number of the crystals were either twinned or warped, as judged by visual inspection. Of those of acceptable size for data collection, only one was untwinned as determined by precession and Weissenberg photographs. This crystal was cleaved into a plate ( $0.3 \times 0.3 \times 0.1$  mm), and mounted on a 0.05 mm glass fiber along its [121] axis with epoxy resin.

#### Data collection and processing

Table 1 lists the unit-cell parameters, space group, and density as measured on the crystal used for data collection. Intensity data were collected on a Picker four-circle FACS I diffractometer, using  $\omega$  scans of  $0.7^\circ$  taken at a scan rate of  $0.25^\circ \text{ min}^{-1}$  to a  $2\theta$  value of  $55^\circ$ . Molybdenum radiation (Mo  $K\alpha_{\text{avg}}=0.71068$  Å) was used with balanced zirconium and yttrium filters. A bandpass intensity correction (Young, 1966; McGandy, 1969) was applied. 992 independent reflections, excluding systematic absences, were measured, of which 927 were observed and 65 unobserved. A reflection was considered to be unobserved if its measured intensity was less than  $2.50\sigma$  (with  $\sigma^2=N_T+N_{BG1}+N_{BG2}$ , where  $N_T$ =total peak count, and  $N_{BG1}$  and  $N_{BG2}$ =the background counts on either side of the peak). The intensity for unobserved reflections was set arbitrarily to  $1.25\sigma$ , and these reflections were given a weight of zero in the least-squares refinement procedure.

Table 1. Crystal data

L-Histidine $C_6N_3O_2H_9$ . M.W. 155.2 dalton
Space group: orthorhombic, $P2_12_12_1$
Systematically absent reflections: $h00$ , $h=2n+1$ ; $0k0$ , $k=2n+1$ ; $00l$ , $l=2n+1$
$a=5.177$ (5) Å
$b=7.322$ (7)
$c=18.87$ (2)
$\alpha=\beta=\gamma=90.00^\circ$
$Z=4$
$d_{\text{calc}}=1.446$ (7) $\text{g.cm}^{-3}$
$d_{\text{obs}}=1.428$ (7) $\text{g.cm}^{-3}$

#### Solution of the structure

The structure was solved using direct-method programs (Maslen & Hall, 1967) for generating triple products ( $DP3$ ) and for tangent refinement ( $DP5$ ), and a program (Main, 1968) for the selection of origin- and enantiomorph-determining reflections. Normalized structure amplitudes ( $E$ 's) were determined from scale and temperature factors, estimated from a Wilson plot

\* Present address: Biochemistry, Frick Chemical Laboratory, Princeton University, Princeton, N.J. 08540, U.S.A.

† Present address: Biology Department, Schermerhorn Hall, 119th St., Columbia University, New York, N.Y., U.S.A.

(Shiono, 1966), and were renormalized on parity,  $\sin \theta$ , and  $hkl$  groupings. The 117 highest  $E$  reflections were used as input to the  $DP3$  program and the Main program.

The Main program generated three linearly independent reflections to determine an origin, one reflection to fix the enantiomorph, and an arbitrary fifth reflection, which was necessary to generate phases for all of the other reflections. Phases were generated and refined by program  $DP5$ , and an  $E$  Fourier map was calculated. This map contained five major peaks, which could not be attributed to any particular group in histidine, but which were used to calculate structure

Table 2. *Positional parameters of L-histidine*

The  $x$ ,  $y$ , and  $z$  coordinates for one of the four molecules in the unit cell are given, along with an estimate of the accuracy of each. The three symmetry-related molecules have the coordinates:  $0.5-x$ ,  $-y$ ,  $0.5+z$ ;  $0.5+x$ ,  $0.5-y$ ,  $-z$ ; and  $-x$ ,  $0.5+y$ ,  $0.5-z$ .

The values for the nonhydrogen atoms are multiplied by  $10^4$ , those for the hydrogen atoms by  $10^3$ .

	X	Y	Z
O(1)	2249 (3)	47 (2)	1982 (1)
O(2)	-1759 (3)	90 (2)	2419 (1)
N(1)	4120 (3)	-1788 (2)	3107 (1)
N(2)	5331 (3)	-605 (2)	4470 (1)
N(3)	3743 (3)	-5 (2)	5534 (1)
C(1)	622 (3)	-170 (2)	2463 (1)
C(2)	1647 (3)	-765 (2)	3197 (1)
C(3)	2089 (4)	907 (3)	3674 (1)
C(4)	3082 (4)	398 (2)	4397 (1)
C(5)	5629 (4)	-816 (3)	5163 (1)
C(6)	2092 (4)	774 (3)	5050 (1)
H(1N1)	487 (4)	-186 (3)	353 (1)
H(2N1)	374 (5)	-295 (4)	296 (1)
H(3N1)	524 (5)	-123 (3)	281 (1)
H(N3)	349 (5)	0 (4)	603 (1)
H(C2)	41 (4)	-162 (3)	342 (1)
H(1C3)	44 (5)	152 (3)	371 (1)
H(2C3)	320 (5)	184 (4)	345 (1)
H(C5)	704 (5)	-146 (4)	536 (1)
H(C6)	53 (5)	146 (4)	517 (1)

factors for the high  $E$  reflections. The 10 reflections which gave the best agreement with the observed structure factors and which contained three origin-determining reflections and an enantiomorph one, were recycled into  $DP5$ , and the resultant  $E$  Fourier map calculated. This map contained the 11 nonhydrogen atoms of histidine and gave an initial  $R$  of 0.27 ( $R = \sum ||F_o| - |F_c|| / |F_o|$ ).

After six cycles of isotropic full-matrix least-squares refinement (Busing, Martin & Levy, 1962), the parameter shifts had converged and the  $R$  value was 0.07 for all reflections. A difference Fourier map revealed four peaks, corresponding to four of the hydrogen atoms. Structure factors were recalculated, including these four hydrogen atoms, and a second difference Fourier map was computed. The highest five peaks corresponded to the remaining hydrogen atoms. Two cycles of least-squares refinement, using isotropic temperature factors for the hydrogen atoms and anisotropic ones for the nonhydrogen atoms, reduced the unweighted  $R$  to 0.032 for all reflections and 0.030 for all observed reflections. To improve the agreement between  $F_o$  and  $F_c$  for both the low- and high-angle ( $\theta$ ) reflections, a weighting scheme (Snyder, 1968) was derived such that  $\Delta F/\sigma$  was equal to 1.0 for 15  $\sin \theta$  groups of  $F_o$  containing equal numbers of reflections. Two cycles of full-matrix least-squares refinement, using this weighting scheme, yielded an  $R$  of 0.034 for all reflections and gave the parameters and molecular dimensions shown in Tables 2 to 8. The atomic positions calculated with this empirical weighting scheme were not significantly different from those calculated without a weighting scheme, while the thermal parameters showed slight variations between the two weighting schemes.

## Discussion

The distances and valence angles in the orthorhombic form of L-histidine (Table 5) are not significantly

Table 3. *Thermal parameters for L-histidine*

The anisotropic coefficients are derived from the expression:  $T = \exp [-(h^2\beta_{11} + k^2\beta_{22} + l^2\beta_{33} + 2hk\beta_{12} + 2hl\beta_{13} + 2kl\beta_{23})]$

$\beta$  and  $\sigma\beta \times 10^5$ .

	$\beta_{11}$ ( $\sigma\beta$ or $\sigma\beta_{11}$ )	$\beta_{22}$ ( $\sigma\beta_{22}$ )	$\beta_{33}$ ( $\sigma\beta_{33}$ )	$\beta_{12}$ ( $\sigma\beta_{12}$ )	$\beta_{13}$ ( $\sigma\beta_{13}$ )	$\beta_{23}$ ( $\sigma\beta_{23}$ )
O(1)	2424 (15)	2236 (14)	115 (2)	-207 (2)	-15 (1)	125 (1)
O(2)	1870 (11)	1552 (9)	226 (4)	71 (1)	-104 (2)	187 (4)
N(1)	1961 (12)	950 (6)	115 (2)	223 (4)	-7 (1)	1 (1)
N(2)	2430 (16)	1411 (9)	128 (2)	315 (5)	-59 (1)	-6 (1)
N(3)	2932 (18)	1319 (8)	107 (1)	-211 (4)	-23 (1)	-13 (1)
C(1)	2035 (12)	855 (6)	116 (2)	-197 (4)	-105 (2)	24 (1)
C(2)	1615 (9)	848 (6)	105 (2)	57 (1)	-3 (1)	32 (1)
C(3)	2993 (19)	884 (6)	126 (2)	392 (6)	-111 (2)	-21 (1)
C(4)	2316 (14)	816 (5)	121 (2)	26 (1)	-67 (2)	-28 (1)
C(5)	2544 (15)	1375 (8)	141 (3)	179 (4)	-135 (2)	36 (1)
C(6)	2439 (15)	1159 (7)	147 (3)	36 (1)	-11 (1)	-84 (2)

Isotropic temperature factors  $B$  for hydrogen atoms

H(1N1)	H(2N1)	H(3N1)	H(N3)	H(C2)	H(1C3)	H(2C3)	H(C5)	H(C6)
1.65 (5)	3.18 (4)	2.81 (9)	4.27 (10)	1.48 (4)	2.59 (7)	2.69 (8)	3.01 (9)	3.56 (9)

Table 4. Structure factor table for orthorhombic L-histidine

The various columns listed are  $I$ ,  $F_{obs}$ , and  $F_{calc}$ . Unobserved reflections are designated with an asterisk after the  $F_{obs}$ .

0 0 0	17 10 20	7 89 90	3 149 149	13 48 44	1 100 3	3 144 44	1 81 81
2 130 132	18 60 21	7 89 90	4 55 51	14 37 34	2 55 55	4 115 115	1 81 81
4 244 244	18 60 21	7 89 90	5 153 154	15 119 147	3 25 25	5 115 115	2 22 20
6 117 117	18 60 21	7 89 90	6 87 87	16 182 182	4 55 55	6 100 100	4 67 67
8 244 244	2 134 138	11 95 91	7 128 127	17 79 5	5 52 42	7 22 20	8 22 20
10 194 186	1 80 81	12 85 84	8 85 83	18 23 6	6 44 45	8 44 45	10 51 53
12 34 30	6 25 21	13 95 84	9 132 123	19 24 47	7 26 20	9 51 53	12 34 30
14 88 81	5 34 32	14 145 145	10 167 170	20 10 5	8 44 45	10 51 53	14 88 81
16 434 456	9 79 89	15 141 145	11 130 130	21 20 18	9 44 45	11 67 67	16 434 456
18 30 46	7 34 33	16 49 49	12 84 83	22 10 1	10 51 53	12 14 19	18 30 46
20 141 182	8 15 16	17 50 51	13 140 151	23 0 141 141	11 57 50	13 14 19	20 141 182
22 38 37	12 14 12	18 60 61	14 48 46	24 102 127	12 44 45	14 14 19	22 38 37
24 51 44	10 42 44	19 65 61	15 87 90	2 175 172	13 16 16	15 24 20	24 51 44
26 0 0	18 15 15	20 72 72	16 39 43	3 262 262	14 40 40	16 44 45	26 0 0
28 1 1	12 73 72	21 57 56	17 9 9	4 20 14	15 37 35	17 44 45	28 1 1
30 105 110	12 73 72	21 57 56	18 110 110	5 69 70	16 44 45	18 44 45	30 105 110
32 282 292	13 51 49	22 65 64	19 110 110	6 103 103	17 44 45	19 44 45	32 282 292
34 212 224	14 102 107	23 72 72	20 41 39	7 215 212	18 44 45	20 44 45	34 212 224
36 43 42	15 41 38	24 77 77	21 41 39	8 215 212	19 44 45	21 44 45	36 43 42
38 324 340	16 85 89	25 84 84	22 15 15	9 40 36	20 44 45	22 44 45	38 324 340
40 7 7	22 22 21	26 91 91	23 15 15	10 101 101	21 44 45	23 44 45	40 7 7
42 222 214	0 50 50	27 95 95	24 2 2	11 103 103	22 44 45	24 44 45	42 222 214
44 162 185	1 2 1	28 100 100	25 2 2	12 103 103	23 44 45	25 44 45	44 162 185
46 146 145	2 69 70	29 106 106	26 1 1	13 103 103	24 44 45	26 44 45	46 146 145
48 276 277	3 15 15	30 112 112	27 2 2	14 103 103	25 44 45	27 44 45	48 276 277
50 111 112	4 34 34	31 214 214	28 2 2	15 103 103	26 44 45	28 44 45	50 111 112
52 122 120	5 32 33	32 68 66	29 4 4	16 103 103	27 44 45	29 44 45	52 122 120
54 134 137	6 34 34	33 214 214	30 4 4	17 103 103	28 44 45	30 44 45	54 134 137
56 175 178	7 39 39	34 11 11	31 7 7	18 103 103	29 44 45	31 44 45	56 175 178
58 11 28	8 64 64	35 117 117	32 7 7	19 103 103	30 44 45	32 44 45	58 11 28
60 18 76	9 79 89	36 150 150	33 10 10	20 103 103	31 44 45	33 44 45	60 18 76
62 17 9	10 64 63	37 19 19	34 10 10	21 103 103	32 44 45	34 44 45	62 17 9
64 182 185	11 2 1	38 123 123	35 10 10	22 103 103	33 44 45	35 44 45	64 182 185
66 19 36	12 11 2	39 150 150	36 11 11	23 103 103	34 44 45	36 44 45	66 19 36
68 20 23	13 10 13	40 43 45	37 12 12	24 103 103	35 44 45	37 44 45	68 20 23
70 21 20	14 10 13	41 30 30	38 12 12	25 103 103	36 44 45	38 44 45	70 21 20
72 103 103	15 22 27	42 30 30	39 12 12	26 103 103	37 44 45	39 44 45	72 103 103
74 38 38	16 12 14	43 1 1	40 12 12	27 103 103	38 44 45	40 44 45	74 38 38
76 102 103	17 12 14	44 2 2	41 12 12	28 103 103	39 44 45	41 44 45	76 102 103
78 0 0	18 10 6	45 10 6	42 12 12	29 103 103	40 44 45	42 44 45	78 0 0
80 1333 1353	19 4 4	46 10 6	43 12 12	30 103 103	41 44 45	43 44 45	80 1333 1353
82 176 181	6 97 97	47 110 110	44 12 12	31 103 103	42 44 45	44 44 45	82 176 181
84 2 15 25	7 97 97	48 2 2	45 12 12	32 103 103	43 44 45	45 44 45	84 2 15 25
86 1964 2027	8 2 2	49 2 2	46 12 12	33 103 103	44 44 45	46 44 45	86 1964 2027
88 4 501 505	9 236 231	50 5 5	47 12 12	34 103 103	45 44 45	47 44 45	88 4 501 505
90 16 10	10 2 2	51 9 9	48 12 12	35 103 103	46 44 45	48 44 45	90 16 10
92 350 347	11 241 231	52 9 9	49 12 12	36 103 103	47 44 45	49 44 45	92 350 347
94 7 248 253	12 134 143	53 9 9	50 12 12	37 103 103	48 44 45	50 44 45	94 7 248 253
96 18 10	13 9 9	54 9 9	51 12 12	38 103 103	49 44 45	51 44 45	96 18 10
98 251 245	14 191 177	55 10 10	52 12 12	39 103 103	50 44 45	52 44 45	98 251 245
100 283 287	15 115 110	56 10 10	53 12 12	40 103 103	51 44 45	53 44 45	100 283 287
102 13 10	16 115 110	57 10 10	54 12 12	41 103 103	52 44 45	54 44 45	102 13 10
104 13 10	17 104 104	58 10 10	55 12 12	42 103 103	53 44 45	55 44 45	104 13 10
106 54 48	18 52 55	59 10 10	56 12 12	43 103 103	54 44 45	56 44 45	106 54 48
108 150 133	19 115 110	60 10 10	57 12 12	44 103 103	55 44 45	57 44 45	108 150 133
110 250 257	20 165 165	61 10 10	58 12 12	45 103 103	56 44 45	58 44 45	110 250 257
112 13 10	21 67 67	62 10 10	59 12 12	46 103 103	57 44 45	59 44 45	112 13 10
114 13 10	22 67 67	63 10 10	60 12 12	47 103 103	58 44 45	60 44 45	114 13 10
116 20 20	23 67 67	64 10 10	61 12 12	48 103 103	59 44 45	61 44 45	116 20 20
118 25 26	24 67 67	65 10 10	62 12 12	49 103 103	60 44 45	62 44 45	118 25 26
120 151 152	25 67 67	66 10 10	63 12 12	50 103 103	61 44 45	63 44 45	120 151 152
122 17 10	26 67 67	67 10 10	64 12 12	51 103 103	62 44 45	64 44 45	122 17 10
124 180 182	27 67 67	68 10 10	65 12 12	52 103 103	63 44 45	65 44 45	124 180 182
126 32 27	28 67 67	69 10 10	66 12 12	53 103 103	64 44 45	66 44 45	126 32 27
128 53 50	29 67 67	70 10 10	67 12 12	54 103 103	65 44 45	67 44 45	128 53 50
130 0 0	30 67 67	71 10 10	68 12 12	55 103 103	66 44 45	68 44 45	130 0 0
132 37 35	31 67 67	72 10 10	69 12 12	56 103 103	67 44 45	69 44 45	132 37 35
134 1 1	32 67 67	73 10 10	70 12 12	57 103 103	68 44 45	70 44 45	134 1 1
136 39 42	33 67 67	74 10 10	71 12 12	58 103 103	69 44 45	71 44 45	136 39 42
138 387 383	34 67 67	75 10 10	72 12 12	59 103 103	70 44 45	72 44 45	138 387 383
140 5 130 126	35 67 67	76 10 10	73 12 12	60 103 103	71 44 45	73 44 45	140 5 130 126
142 239 236	36 67 67	77 10 10	74 12 12	61 103 103	72 44 45	74 44 45	142 239 236
144 138 139	37 67 67	78 10 10	75 12 12	62 103 103	73 44 45	75 44 45	144 138 139
146 110 100	38 67 67	79 10 10	76 12 12	63 103 103	74 44 45	76 44 45	146 110 100
148 149 149	39 67 67	80 10 10	77 12 12	64 103 103	75 44 45	77 44 45	148 149 149
150 38 43	40 67 67	81 10 10	78 12 12	65 103 103	76 44 45	78 44 45	150 38 43
152 41 44	41 67 67	82 10 10	79 12 12	66 103 103	77 44 45	79 44 45	152 41 44
154 331 316	42 67 67	83 10 10	80 12 12	67 103 103	78 44 45	80 44 45	154 331 316
156 138 88	43 67 67	84 10 10	81 12 12	68 103 103	79 44 45	81 44 45	156 138 88
158 46 87	44 67 67	85 10 10	82 12 12	69 103 103	80 44 45	82 44 45	158 46 87
160 120 120	45 67 67	86 10 10	83 12 12	70 103 103	81 44 45	83 44 45	160 120 120
162 94 10	46 67 67	87 10 10	84 12 12	71 103 103	82 44 45	84 44 45	162 94 10
164 94 10	47 67 67	88 10 10	85 12 12	72 103 103	83 44 45	85 44 45	164 94 10
166 31 31	48 67 67	89 10 10	86 12 12	73 103 103	84 44 45	86 44 45	166 31 31
168 94 4	49 67 67	90 10 10	87 12 12	74 103 103	85 44 45	87 44 45	168 94 4
170 107 103	50 67 67	91 10 10	88 12 12	75 103 103	86 44 45	88 44 45	170 107 103
172 33 27	51 67 67	92 10 10	89 12 12	76 103 103	87 44 45	89 44 45	172 33 27
174 100 101	52 67 67	93 10 10	90 12 12	77 103 103	88 44 45	90 44 45	174 100 101
176 11 15	53 67 67	94 10 10	91 12 12	78 103 103	89 44 45	91 44 45	176 11 15
178 11 15	54 67 67	95 10 10	92 12 12	79 103 103	90 44 45	92 44 45	178 11 15
180 303 297	55 67 67	96 10 10	93 12 12	80 103 103	91 44 45	93 44 45	180 303 297
182 1 01 82	56 67 67	97 10 10	94 12 12	81 103 103	92 44 45	94 44 45	182 1 01 82
184 252 242	57 67 67	98 10 10	95 12 12	82 103 103	93 44 45	95 44 45	184 252 242
186 17 13	58 67 67	99 10 10	96 12 12	83 103 103	94 44 45	96 44 45	186 17 13
188 235 248	59 67 67	100 10 10	97 12 12	84 103 103	95 44 45	97 44 45	188 235 248
190 5 62 62	60 67 67	101 10 10	98 12 12	85 103 103	96 44 45	98 44 45	190 5 62 62
192 548 344	61 67 67	102 10 10	99 12 12	86 103 103	97 44 45	99 44 45	192 548 344
194 7 12 12	62 67 67	103 10 10	100 12 12	87 103 103	98 44 45	100 44 45	194 7 12 12
196 104 104	63 67 67	104 10 10	101 12 12	88 103 103	99 44 45	101 44 45	196 104 104
198 7 12 12	64 67 67	105 10 10	102 12 12	89 103 103	100 44 45	102 44 45	198 7 12 12
200 185 183	65 67 67	106 10 10	103 12 12	90 103 103	101 44 45	103 44 45	200 185 183
202 11 74 76	66 67 67	107 10 10	104 12 12	91 103 103	102 44 45	104 44 45	202 11 74 76
204 11 74 76	67 67 67	108 10 10	105 12 12	92 103 103	103 44 45	105 44 45	204 11 74 76
206 44 44	68 67 67	109 10 10	106 12 12	93 103 103	104 44 45	106 44 45	206 44 44
208 50 52	69 67 67	110 10 10	107 12 12	94 103 103	105 44 45	107 44 45	208 50 52
210 67 76	70 67 67	111 10 10	108 12 12	95 103 103	106 44 45	108 44 45</	

Hoy, 1972; Edington, 1969). There are, however, significant differences between the neutral histidines and the protonated forms, *e.g.* L-histidine.HCl.H<sub>2</sub>O

Table 7. *Hydrogen bonding distances and angles in L-histidine*

Atom *A* is covalently bonded to the hydrogen (H) and *B* is hydrogen-bonded to atom *A*. The number in parentheses after the atom name refers to the unit cell and symmetry operation relating this atom to the atoms listed in Table 2. 555 refers to the origin cell (as per the *ORTEP* conventions), while 1, 2, 3, and 4 refer to the symmetry operators listed in Table 2. If no such information is listed, the atom is in the origin unit cell, with a symmetry operation of 1 (Johnson 1965).

Atom		Distances and angles		
<i>A</i>	<i>B</i>	<i>A</i> - <i>B</i>	<i>B</i> -H	∠ <i>AHB</i>
N(1)	N(2)	2.783 Å	1.990 Å	143.8°
N(1)	O(2) (545,4)	2.773	1.925	159.2
N(1)	O(2) (655,1)	2.851	1.981	167.6
N(3)	O(1) (555,2)	2.781	1.866	176.5

Close contacts		
N(1)	O(1)	2.693
N(1)	O(1) (645,4)	2.988

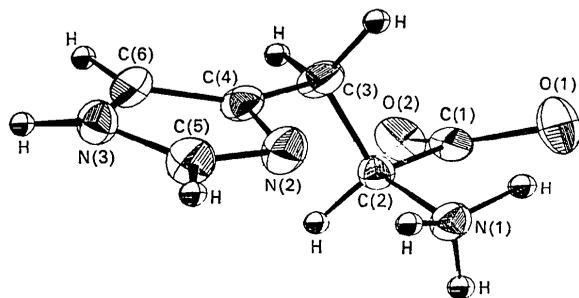


Fig. 1. An *ORTEP* plot (Johnson 1965) of L-histidine, as found in this study, demonstrating the numbering system used.

(Donohue & Caron, 1964) and DL-histidine.HCl.2H<sub>2</sub>O (Bennett *et al.*, 1970). Orthorhombic L-histidine has an extended alanine backbone as does the DL hydrochloride salt, while the L hydrochloride salt folds back on itself around C(2)-C(3). This folding causes the imidazole residue to be *gauche* to both the carboxyl and primary amino groups, while in the extended conformation the imidazole is *gauche* only to the amino group, and *trans* to the carboxyl. Thus, as noted by Bennett *et al.* (1970), the angles C(1)-C(2)-C(3), C(2)-C(3)-C(4), and to a lesser extent N(1)-C(2)-C(3), are significantly larger in L-histidine hydrochloride than in the orthorhombic free base and the DL-hydrochloride (Table 6).

The angles around the N(2), C(5) and N(3) atoms in the imidazole residue are also significantly different in the protonated and unprotonated forms. In the protonated hydrochloride salts, the three angles centered on these atoms are approximately equal to 109°, a fact indicative of the aromaticity of the ring. In the unprotonated orthorhombic compound, the ring angles at N(2) and N(3) are both significantly compressed

Table 8. *Planarity of the imidazole ring plus the adjacent carbon C(3)*

The plane calculated for the imidazole ring plus the carbon C(3) by least-squares analysis is given by the equation:

$$0.5317A + 0.8463B + 0.0324C = 1.3633.$$

	Distance from plane
N(2)	0.0015 Å
N(3)	0.0012
C(3)	-0.0004
C(4)	0.0002
C(5)	-0.0004
C(6)	-0.0020

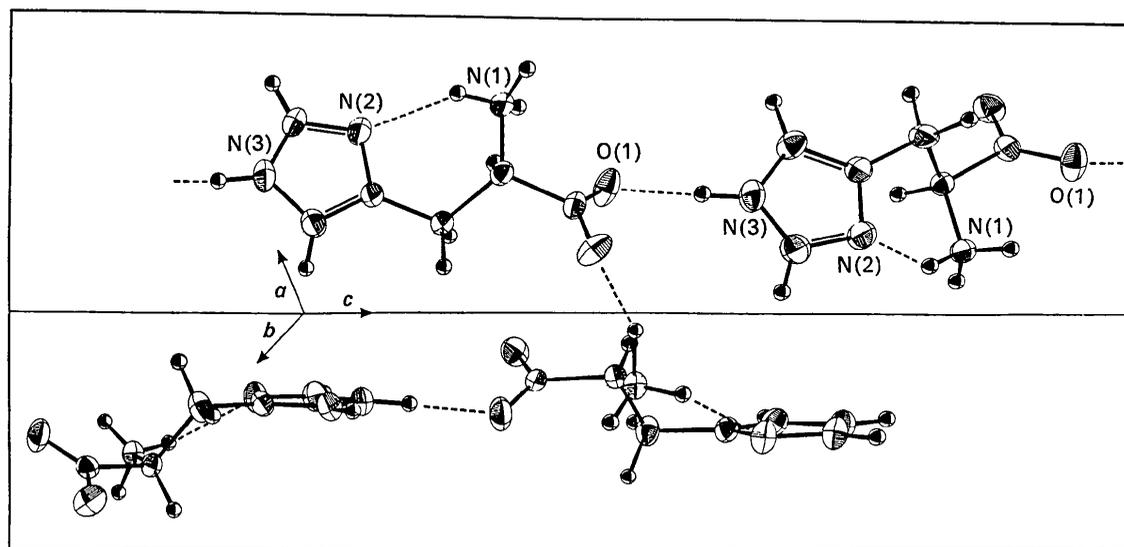


Fig. 2. An *ORTEP* plot of the contents of the unit cell of orthorhombic L-histidine. The dotted lines represent hydrogen bonds found in this structure.

(104.9 and 106.9° respectively) while the N(2)–C(5)–N(3) angle is widened to 112°, which is consistent with an increase in the  $sp^2$  character of the bond hybridization at the ring carbon, C(5).

A series of three intermolecular hydrogen bonds determines the packing of the carboxyl and amino groups, and links the head of one molecule [N(3)] to the tail of the next [O(1)] (Fig. 2). A weak intramolecular hydrogen bond with an NH...N distance of 1.99 Å also occurs between the N(1) and N(2) atoms, contrary to the assertion of Kier (1968) that this type of bond could probably not form in a similar compound (histamine). N(1) also approaches O(1) of the same molecule to a distance of 2.69 Å, but none of the N(1) hydrogen atoms is closer than 2.42 Å to O(1), so that there is no hydrogen bond formed (Table 7). This type of interaction is identical to the electrostatic interaction described by Sasisekharan (1971) for free amino acids.

The charge density of each of the atoms in histidine was estimated using an INDO approximation (Pople, Beveridge & Dobosh, 1967), giving the results shown in Table 9. The unprotonated nitrogen atom of histidine has a large negative charge, as is the case for the unprotonated nitrogen atom of  $\beta$ -(pyrazolyl-3)-L-alanine (Seeman, 1970) with which it is compared. These ring charges can be explained by postulating the resonance forms shown in Fig. 3. The contributions of resonance form (II), in which the unprotonated nitrogen atom has formal negative charges, is further confirmed by the fact that the N(3)–C(5) bond in histidine

and the N(3)–C(5) bond in pyrazole–alanine are considerably shortened (1.339 and 1.332 Å respectively) compared to the expected N–C single bond length (about 1.47 Å), and even shorter than those of N(2)–C(4) and N(3)–C(6) (1.382 and 1.374 Å, respectively) of the histidine imidazole.

Table 9. Charge density on the nonhydrogen atoms of L-histidine and  $\beta$ -(pyrazolyl-3)-L-alanine (Seeman, 1970) as determined by INDO (Pople *et al.*, 1967)

The numbers given represent the number of electrons found on each atom, such that a charge of 7.00 would be neutral for nitrogen, highly negative for carbon, and highly positive for oxygen.

	Electronic charge	
	L-Histidine	$\beta$ -(Pyrazolyl-3)-L-alanine
O(1)	8.58	8.58
O(2)	8.54	8.54
N(1)	6.95	7.16
N(2)	7.31	7.25
N(3)	7.05	6.95
C(1)	5.53	5.51
C(2)	5.99	5.97
C(3)	5.96	5.98
C(4)	5.93	5.89
C(5)	5.79	5.92
C(6)	5.99	6.06

While the *trans* conformation about C(2)–C(3) places the bulkiest groups (carboxyl and imidazole) on opposite sides of the molecule, thereby reducing steric interference, there are several other conformations of equally low energy (Kistenmacher & Marsh, 1971). More stability of the type described by Ponnuswamy & Sasisekharan (1970), is obtained by the intramolecular, zwitterionic interaction between the O(1) and N(1) atoms, and also from the intramolecular hydrogen bond between the N(1) and N(2) atoms. In this conformation, N(2), the unprotonated nitrogen atom, is sterically hindered so that it cannot be approached by an electrophilic reagent without first a rotation of 120° around either C(2)–C(3) or C(3)–C(4). While such a rotation is energetically allowed for the free amino acid in solution, in a polypeptide, steric hindrance from spatially neighboring groups could prevent such a rotation, thereby chemically inactivating the imidazole residue.

We wish to thank Professor G. A. Jeffrey for the use of the facilities of the Crystallography Department. Thanks also go to Dr R. Shono for programming assistance and to the University of Pittsburgh Computing Center for the use of its IBM 7090. We also thank Dr Robert Stewart, of Carnegie-Mellon U., for helping one of us (NCS) to perform the INDO calculations. This research is supported by Training Grant No. GM-01728 of the U.S. Public Health Service, National Institutes of Health, and by Grant Number L-71 of the Pittsburgh Health Research and Services Foundation.

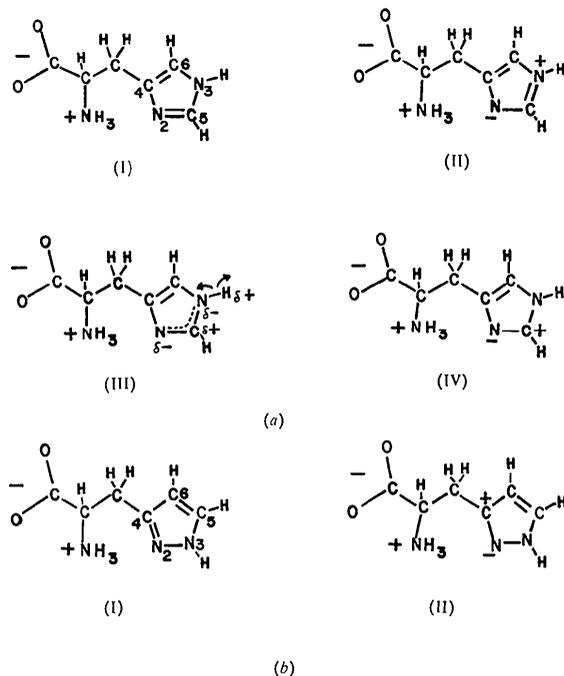


Fig. 3. The primary resonance contributors to the structures of (a) L-histidine and (b)  $\beta$ -(pyrazolyl-3)-alanine (Seeman, 1970).

## References

- BENNETT, I., DAVIDSON, A. G. H., HARDING, M. M. & MORELLE, I. (1970). *Acta Cryst.* B26, 1722.
- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS*. Report ORNL-TM-305. Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- DONOHUE, J. & CARON, A. (1964). *Acta Cryst.* 17, 1187.
- DONOHUE, J., LAVINE, L. R. & ROLLETT, J. S. (1956). *Acta Cryst.* 9, 655.
- EDINGTON, P. (1969). Ph. D. Dissertation. Univ. of Edinburgh.
- FREEMAN, H. C. (1967). *Advanc. Protein Chem.* 22, 257.
- JOHNSON, C. K. (1965). *ORTEP*. Report ORNL-3794. Oak Ridge National Laboratory, Oak Ridge, Tennessee.
- KIER, L. B. (1968). *J. Med. Chem.* 11, 441.
- KISTENMACHER, T. D. & MARSH, R. E. (1971). *Science*, 172, 945.
- MADDEN, J. J., MCGANDY, E. L., SEEMAN, N. C., HARDING, M. M. & HOY, A. (1972). *Acta Cryst.* B28, 2382.
- MAIN, P. (1968). *A Non-centrosymmetric Direct Method Program*. Crystallography Department, Univ. of Pittsburgh, Pa. 15213.
- MASLEN, T. & HALL, S. (1967). *DP3 and DP5 - Direct Method Programs*. Crystallography Department, Univ. of Pittsburgh, Pittsburgh, Pa. 15213.
- MCGANDY, E. L. (1969). *PIKOT - A Program for Processing X-ray Single Crystal Diffractometer Data*. Crystallography Department, Univ. of Pittsburgh, Pittsburgh, Pa. 15213.
- MEADOWS, D. H., ROBERTS, G. C. K., & JARDETSKY, O. (1969). *J. Mol. Biol.* 45, 491.
- PONNUSWAMY, A. & SASISEKHARAN, V. (1970). *Int. J. Protein Res.* 2, 37.
- POPLE, J., BEVERIDGE, D. L. & DOBOSH, P. D. (1967). *J. Chem. Phys.* 47, 2026.
- ROHLING, D. L. & FOX, S. W. (1967). *Arch. Biochem. Biophys.* 118, 122.
- SASISEKHARAN, V. (1971). Private communication.
- SCHOELLMAN, G., SCHOELLMAN, E. & SHAW, E. (1963). *Biochemistry*, 2, 252.
- SEEMAN, N. C. (1970). Ph. D. Dissertation. Univ. of Pittsburgh, Pittsburgh, Pa. 15213.
- SHIONO, R. (1966). *A Wilson Plot Program for the I.B.M. 1130, with the Option to Renormalize the Calculated E's*. Crystallography Department, Univ. of Pittsburgh, Pa. 15213.
- SNYDER, R. (1968). *WEIGHT - a Program for Producing a Least-Squares Weighting Scheme*. Crystallography Department, Univ. of Pittsburgh, Pittsburgh, Pa. 15213.
- YOUNG, R. A. (1966). *Background Intensities in Single-Crystal Diffractometry*. Technical Report No. 2, Project No. A-389, ASTIA 262912, Dept. of Physics, Georgia Inst. of Technology, Atlanta, Georgia 30332.

*Acta Cryst.* (1972). B28, 2382

## The Crystal Structure of the Monoclinic Form of L-Histidine

BY J. J. MADDEN,\* E. L. MCGANDY & N. C. SEEMAN†

*Department of Crystallography, University of Pittsburgh, Pittsburgh, Pa. 15213, U.S.A.*

AND MARJORIE M. HARDING & A. HOY

*Department of Chemistry, University of Edinburgh, Edinburgh 9, Scotland*

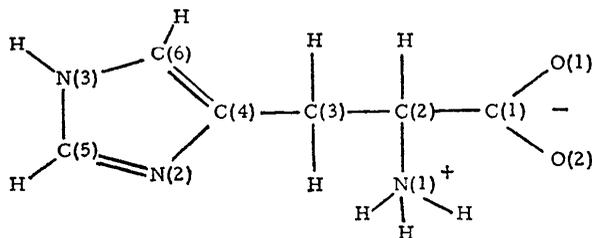
(Received 4 October 1971)

L-Histidine ( $C_6N_2O_2H$ ) crystallizes from ethanol in the monoclinic space group  $P2_1$ , with  $a=5.172$ ,  $b=7.384$ ,  $c=9.474$  Å,  $\beta=97.162^\circ$  and  $Z=2$ . The structure was solved simultaneously by independent investigations using the tangent formula and from a trial solution based on the structure of the orthorhombic form. The crystals show lamellar twinning, which arises from faults in the stacking of the imidazole residues such that there are two possible orientations of the unit cells. The structures could not be refined below an  $R=0.10$ , but a comparison of the bond distances and angles with those of other free-base histidines shows no significant differences.

### Introduction

Histidine, pictured below, and some of its isostructural analogs, have now been examined as free bases in a series of compounds which includes orthorhombic L-histidine (Madden, McGandy & Seeman, 1972), D,L-histidine (Edington, 1970),  $\beta$ -(pyrazoyl-3)-L-alanine (Seeman, McGandy & Rosenstein, 1972), and mono-

clinic L-histidine which is described in this paper and was studied independently and simultaneously at Edinburgh and Pittsburgh.



\* Present address: Biochemistry, Frick Chemical Laboratory, Princeton University, N.J. 08540, U.S.A.

† Present address: Biology Department, Schermerhorn Hall, 119th Street, Columbia University, New York, N.Y., U.S.A.