J. Biosci., Vol. 20, Number 2, March 1995, pp 225-234 © Printed in India.

Variability in ionization state, stoichiometry and aggregation in histidine complexes with formic acid*

STEPHEN SURESH and M VIJAYAN†
Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India

MS received 21 November 1994; revised 14 March 1995

Abstract. The crystal structures of the complexes of L and DL histidine with formic acid have been determined as part of an effort to define biologically and evolutionarily important interactions and aggregation patterns. In terms of ionization state and stoichiometry they may be described as L-histidine formate formic acid and DL-histidine formate monohydrate respectively. In the L-histidine complex, amino acid molecules arranged in head-to-tail sequences centred around 2_1 screw axes are interconnected by formic acid molecules and formate ions. Histidine-formate interactions in the structure gives rise to a characteristic interaction pattern involving a linear array of alternating imidazole groups and formate ions. In DL-histidine formale monohydrate, head-to-tail sequences involving glide related molecules are interconnected through main chain-side chain interactions leading to amino acid layers. The layers are held together by formate ions and water molecules arranged in strings along which the ion and the molecule alternate. The patterns of amino acid aggregation in histidine complexes exhibit considerably higher variability than those in complexes involving arginine and lysine do.

Keywords. Amino acid complexes; molecular aggregation; amino acid-carboxylic acid interaction; chemical evolution.

1. Introduction

Non-covalent interactions are crucial to the structure, function and assembly of proteins. Our studies of crystalline complexes containing amino acids and peptides (Vijayan 1988; Prasad and Vijayan 1993a; Stephen Suresh *et al* 1994a,b) have not only yielded a wealth of information on biologically important non-covalent interactions, but have also provided a detailed understanding of the well defined patterns of interactions and aggregation involving amino acids and peptides. These patterns provide useful insights into the role of molecular interactions and aggregation in chemical evolution (Vijayan 1980; Suresh and Vijayan 1983b; Suresh and Vijayan 1985; Vijayan 1988). The current focus of the programme is on amino acid and peptide complexes with small organic molecules that are believed to have existed in the prebiotic milieu. The information obtained from these complexes are expected to be relevant to the understanding of the self-assembly processes that might have led to the first self-replicating systems, in addition to its intrinsic interest in relation to molecular association in general. We have already obtained interesting results from the analysis of a number of complexes involving succinic and acetic acids (Suresh and Vijayan 1983a; Soman *et al* 1989; Prasad and Vijayan 1990, 1991,

*X-ray studies on crystalline complexes involving amino and peptides Part XXIX.
†Corresponding author.

1993a,b; Stephen Suresh *et al* 1994a). The complexes of arginine (Stephen Suresh *et al* 1994b) and lysine (Stephen Suresh and Vijayan 1994) with formic acid, the most abundant carboxylic acid produced in the simulated prebiotic synthesis experiments (Miller and Urey 1959; Kvenvolden *et al* 1971; Miller and Orgel 1974) and the smallest of its kind, have also been analysed. Here we present the crystal structures of L- and DL-histidine complexed with formic acid. Histidine is a particularly interesting amino acid because it can act as a proton donor, a proton acceptor, a nucleophilic agent and a ligand to metal ions (Madden *et al* 1972). It plays a crucial role in the catalytic activity of numerous enzymes.

2. Methods

The crystalline complexes in both instances were prepared by vapour diffusion of acetone into aqueous solutions of histidine containing an excess of formic acid. The space groups and the unit cell dimensions of the crystals were determined by Weissenberg photography using nickel filtered copper radiation and the densities were measured by floatation in mixtures of benzene and carbon tetrachloride. The cell parameters were subsequently refined on a CAD-4 diffractometer, which was also used to collect the intensity data. The experimental details and refinement parameters are given in table 1.

Table 1. Crystal data, experimental details and refinement parameters for L-histidine formate formic acid and DL-histidine formate rnonohydrate.

	L-histidine formate formic acid	DL-histidine formate monohydrate
Chemical formula	$C_6H_{10}N_3O_2^+ \cdot CH_2O_2 \cdot CHO_2^-$	$C_6H_{10}N_3O_2^+ \cdot CHO_2^- H_2O$
Formula weight	247·21	219·20
Crystal system	Monoclinic	Monoclinic
Space group	$P2_1$	$P2_1/c$
Unit cell dimensions		
a	7·261 (1) Å	9·309 (1) Å
b	7·441 (1) Å	9·138 (1) Å
c	11·062 (2) Å	11·741 (1) Å
β	104·93 (1)°	92·10(1)°
Volume	577·5 (1) $Å^3$	998·1 (2) $Å^3$
Z	2	4
Density (calculated)	1·422 Mg/m^3	1·459 Mg/m^3
Density (measured)	1·42 (2) Mg/m^3	1·46 (2) Mg/m^3
Crystal size	0·53 × 0·45 × 0·40 mm	0·53 × 0·45 × 0·25 mm
Wave length	1·5418 Å	1·5418 Å
μ	1·063 mm^{-1}	1·073 mm^{-1}
Max. Bragg angle	75°	75°
Unique reflections	1244	2065
Goodness-of-fit	1·082	1·054
R [I > 2 sigma(I)]	0·0310	0·0412
wR2 (all data)	0·0834	0·1061
$(\Delta/\sigma)_{max}$	0·025	0·032
$(\Delta\rho)_{min}$	− 0·222 $e/Å^3$	− 0·235 $e/Å^3$
$(\Delta\rho)_{max}$	0·161 $e/Å^3$	0·225 $e/Å^3$

Estimated SD given in parenthesis.

The structures were solved using the direct methods program SHELXS-86 (Sheldrick 1986) and refined by the full-matrix least squares method employing a minimization procedure based on F^2 using the program SHELXL-93 (Sheldrick 1993). The hydrogen atoms were located using difference Fourier maps and stereochemical considerations. Non-hydrogen atoms were refined anisotropically and hydrogen atoms isotropically. The positional parameters and the equivalent isotropic thermal parameters of the non-hydrogen atoms in the two structures are given in tables 2 and 3.

Table 2. Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($Å^2 \times 10^4$) in L-histidine formate formic acid.

	x	*y*	*z*	U_{eq}
N(1)	2713(2)	1350(3)	3580(2)	398(6)
O(1)	5261(2)	–995(3)	4843(1)	553(6)
O(2)	6292(3)	–1774(3)	3182(2)	612(6)
C(1)	5286(3)	–887(4)	3731(2)	413(6)
C(2)	3994(3)	495(3)	2898(2)	370(6)
C(3)	5216(3)	1912(4)	2473(2)	437(7)
C(4)	4120(3)	3171(3)	1504(2)	405(7)
N(5)	3068(3)	2604(3)	342(2)	446(7)
C(6)	2348(4)	4046(5)	–320(2)	540(9)
N(7)	2893(4)	5498(4)	360(2)	570(10)
C(8)	4008(4)	4993(4)	1510(3)	508(9)
O(11)	2217(2)	9140(3)	–270(2)	534(6)
O(12)	404(3)	7538(3)	–1796(2)	616(7)
C(11)	997(3)	8973(5)	–1282(2)	512(8)
O(21)	10772(3)	7865(4)	3219(2)	802(9)
O(22)	9095(3)	6391(4)	4327(2)	766(9)
C(21)	10641(4)	6815(6)	4016(3)	693(11)

Table 3. Atomic coordinates ($\times 10^4$) and equivalent isotropic displacement parameters ($Å^2 \times 10^4$) in DL-histidine formate monohydrate.

	x	*y*	*z*	U_{eq}
N(1)	5344(2)	1366(2)	8771(1)	364(5)
O(1)	4597(2)	1115(2)	6588(1)	495(5)
O(2)	5514(2)	3263(2)	6085(1)	466(4)
C(1)	5210(2)	2301(2)	6780(1)	302(4)
C(2)	5670(2)	2628(2)	8031(1)	270(4)
C(3)	7261(2)	3001(2)	8119(1)	362(5)
C(4)	7809(2)	3517(2)	9260(1)	299(4)
N(5)	7414(2)	4826(2)	9730(1)	347(4)
C(6)	8126(2)	5003(2)	10723(2)	395(6)
N(7)	8963(2)	3858(2)	10903(1)	383(5)
C(8)	8782(2)	2911(2)	10002(1)	366(5)
O(11)	2292(2)	1567(2)	9012(1)	566(6)
O(12)	949(2)	1938(2)	7451(1)	587(6)
C(11)	1314(2)	1242(2)	8322(2)	426(6)
W(1)	2096(2)	626(2)	1278(2)	640(6)

3. Results and discussion

3.1 *Stoichiometry and ionization state*

In both the complexes, the α-amino and the imidazole groups are protonated and positively charged while the α-carboxyl groups are deprotonated and negatively charged. The molecules are thus zwitterionic and carry a net positive charge each. In the L-histidine complex, the asymmetric unit contains one histidinium cation, a formate anion and a formic acid molecule. The situation is somewhat similar to that in L-histidine semisuccinate trihydrate (Prasad and Vijayan 1993b), where the asymmetric unit contains one histidinium cation and a singly negatively charged semisuccinate ion with one protonated, neutral carboxyl group and one deprotonated negatively charged carboxylate group. The similarity in ionization state and stoichiometry extends to the corresponding DL-histidine complexes as well. DL-histidine succinate dihydrate (Prasad and Vijayan 1993b) contains, as in the formate complex, one carboxylate group for each histidinium ion. The two forms of the complex of L-histidine with acetic acid (Stephen Suresh *et al.* 1994a) however contain, unlike in those with succinic: and formic acids, only one carboxylate group (and no carboxyl group) for every histidinium ion. A complex between DL-histidine and acetic acid is yet to be crystallized. Among the three carboxylic acids, formic acid is the most acidic with a *pK* of 3·75, while succinic acid has *pK* values of 4·16 and 5·61 and acetic acid has an intermediate value of 4·75 (CRC Handbook of Chemistry and Physics 1980). The observed stoichiometry and ionization states in the formic, acetic and succinic acid complexes cannot be explained in a satisfactory manner on the basis of the *pK* values.

3.2 *Molecular conformation*

The bond lengths and angles in the two structures are normal. As is well known, the conformation of the histidine molecule is defined by χ^1 and χ^{21} (IUPAC-IUB Commision on Biochemical Nomenclature 1970). χ^1, which defines the disposition of the side chain with respect to the main chain, can take values in the neighbourhood of –60°, 60° or 180° corresponding to the open conformation *I* (g^-), closed conformation (g^+) and open conformation II (*t*) respectively (Bhat and Vijayan 1978; Krause *et al* 1991). Though the preferred values of χ^{21} are -90° and +90°, it often deviates from these ideal values due to interactions of the imidazole with other groups in the crystal. In both the structures, the molecule adopts the sterically most favourable open conformation I' with $\chi^1 \sim$ –60°. In fact this is the most Frequently observed conformation of histidine, occurring in 12 of the 19 independent histidine molecules observed in crystal structures so far (Prasad and Vijayan 1993b; Stephen Suresh *et al* 1994a). In the L-histidine complex, χ^{21} takes a value of –60·9° and in the DL complex it takes a value of –67·5°. The carboxyl hydrogen in the formic acid molecule in L-histidine formate formic acid adopts the energetically favourable *cis* conformation (Miyazawa and Pitzer 1959).

3.3 *Hydrogen bonding and molecular aggregation*

The crystal structures of the two complexes are illustrated in figures 1 and 2,

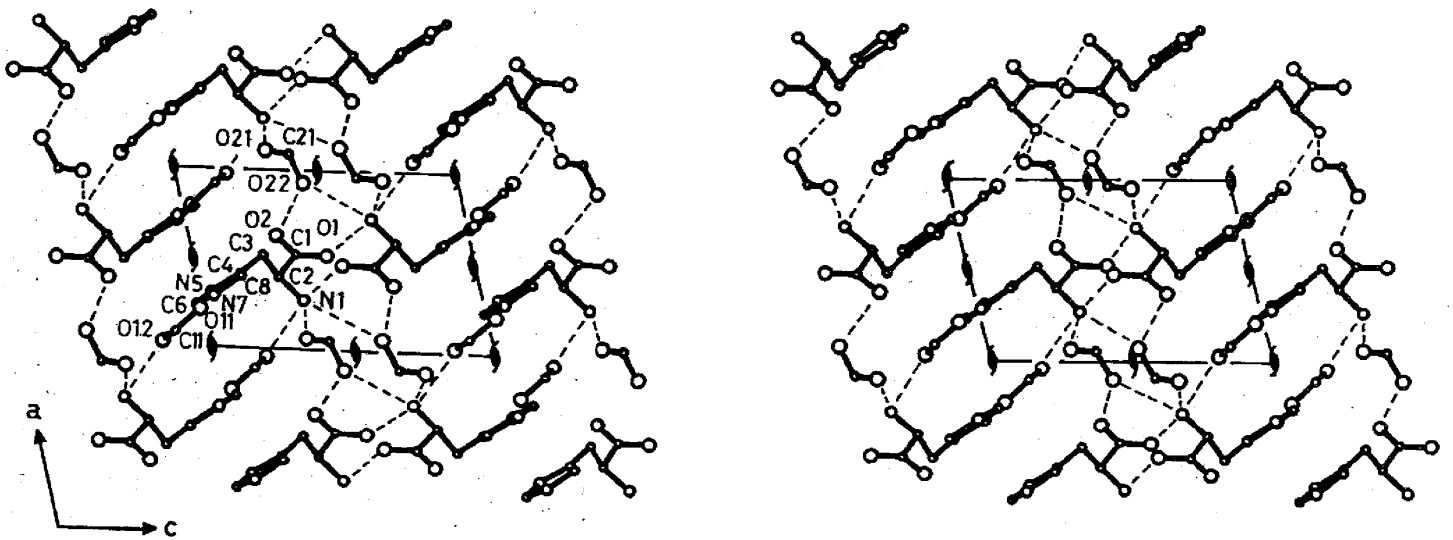

Figure 1. Crystal structure of L-histidine formate formic acid. In this and subsequent figures carbon, nitrogen and oxygen atoms are indicated by circles of increasing size, and broken lines represent hydrogen bonds. All the figures were drawn using PLUTO (Motherwell and Clegg 1978).

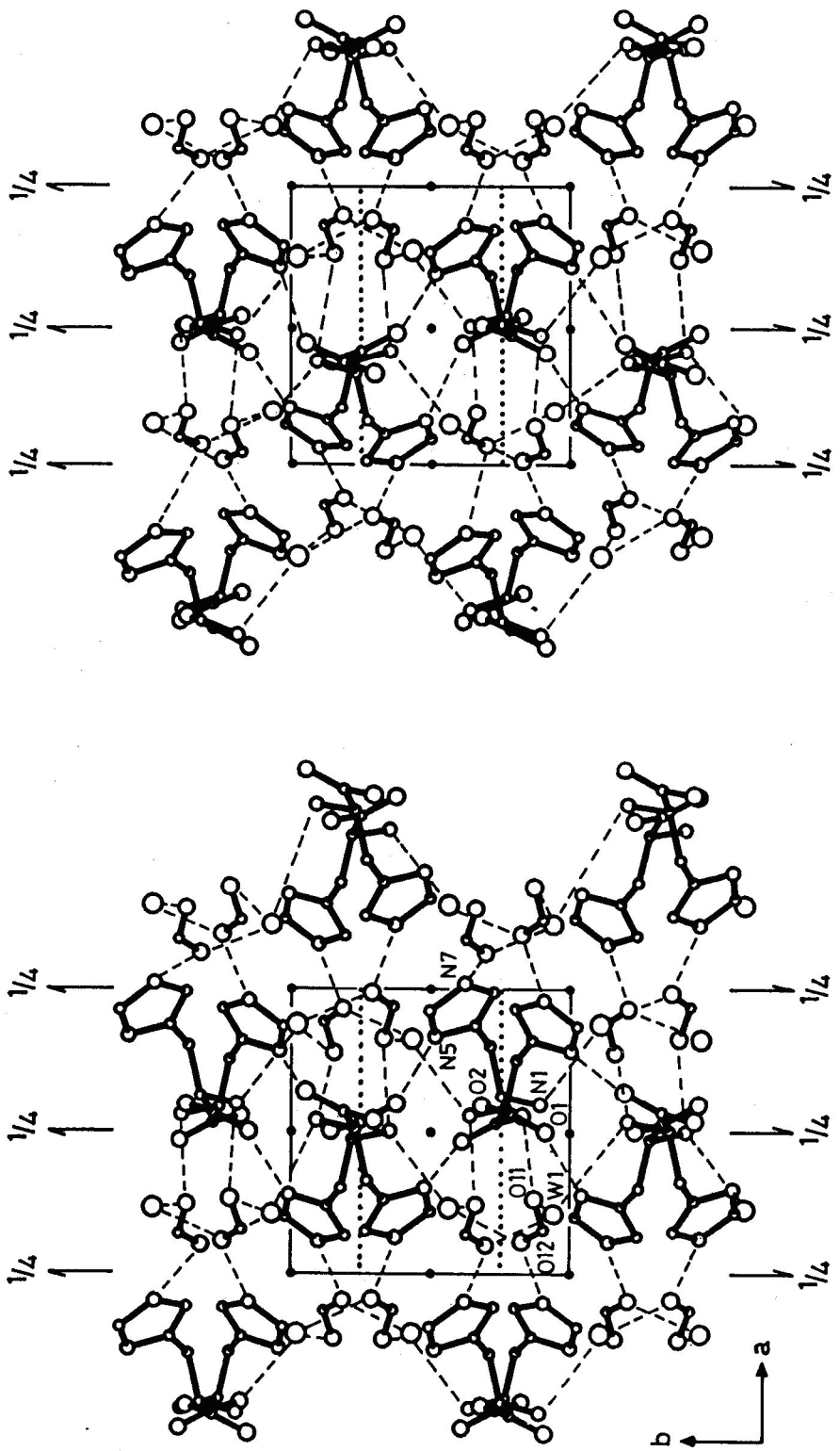

Figure 2. Crystal structure of DL-histidine formate monohydrate. Only nitrogen and oxygen atoms are numbered. The hydrogen bond between N1 of one molecule and 02 of a glide related molecule has not been shown, as the atoms partially overlap.

while the parameters of the hydrogen bonds in them are listed in table 4. In the L-histidine complex, the α-amino group takes part as proton donor in four hydrogen bonds, of which two constitute a bifurcated hydrogen bond involving 021 of a formic acid molecule and 022 of a symmetry equivalent of the same molecule. The remaining two involve a formate ion and an *α*-carboxylate group. The imidazole group interacts exclusively with formate ions through two N–H . . . O hydrogen bonds. The remaining hydrogen bond is a short O–H . . . O type interaction

Table 4. Hydrogen bond parameter.

A–H . . . B	A . . . B (Å)	A–H . . . B (°)
L-histidine formate formic acid		
N1–H1N1 . . . 022^{b}	2·941(3)	122(2)
N1–H1N1 . . . 021^{c}	2·929(3)	124(2)
N1–H2N1 . . . 01^{d}	2·790(3)	167(2)
N1–H3N1 . . . 012^{e}	2·738(2)	165(3)
N5–H1N5 . . . 011^{f}	2·696(3)	176(3)
N7–H1N7 . . . 011^{a}	2·810(4)	171(4)
022–H022 . . . 02^{g}	2·507(3)	169(5)
Symmetry Codes: (*a*) + X, + Y, + Z; (*b*) –X + 1, + Y–1/2, –Z + 1; (*c*) + X–1, + Y–1, + Z; (*d*) –X + 1, + Y+ 1/2, –Z + 1; (*e*) –X, + Y–1/2, –Z; (*f*) + X, + Y–1, + Z; (*g*) + X, + Y + 1, + Z.		
DL-histidine formate monohydrate		
N1–H1N1 . . . 02^{b}	2·736(2)	174(2)
N1–H1N1 . . . 011^{a}	2·871(2)	146(2)
N1–H3N1 . . . W1^{c}	3·001(3)	155(3)
N5–H1N5 . . . 01^{d}	2·659(2)	162(2)
N7–H1N7 . . . 012^{e}	2·646(2)	162(3)
W1–H1W1 . . . 012^{f}	2·846(3)	175(3)
W1–H2W1 . . . 011^{g}	2·809(3)	178(3)
Symmetry Codes: (*a*) X, Y, Z; (*b*) + X, –Y + 1/2, + Z + 1/2; (*c*) –X + 1, –Y, –Z + 1; (*d*) –X + 1, + Y + 1/2, –Z + 1/2 + 1; (*e*) +X + 1, –Y + 1/2, +Z + 1/2; (*f*) +X, –Y + 1/2, +Z–1/2; (*g*) + X, +Y, +Z–1.		

(Ramanadham and Chidambaram 1978; Mitra and Ramakrishnan 1981) involving formic acid and the α-carboxylate group. In DL-histidine formate monohydrate, the α-amino group forms hydrogen bonds with an α-carboxylate group, a formate ion and a water molecule. Unlike in the L-histidine complex, the imidazole group in the complex of the racemate interacts with a formate ion as well as an α-carboxylate group. The lone water molecule in the structure interacts as proton donor exclusively with formate ions.

It has been shown earlier that crystal structures containing amino acids can often be described in terms of head-to-tail sequences in which the α-amino and the α-carboxylate groups are brought into periodic hydrogen bonded proximity (Suresh and Vijayan 1983b). When the sequence contains screw related molecules that are connected through N1 ... 01 hydrogen bonds, it is referred to as a Z1 type

head-to-tail sequence. In L-histidine formate formic acid (figure 1) the histidine molecules from ribbons along *b* involving such sequences formed by a Nl . . . 01 hydrogen bond and its screw equivalents. The Z1 sequences involving 01 are rarer than the Z2 sequences involving 02. They occur with around half the frequency of Z2 sequences (Suresh and Vijayan 1983b). In fact L-histidine formate formic acid is the only crystal structure observed so far in which histidine forms a Zl sequence. The Zl ribbons related by an *a* translation are linked by formic acid molecules to form a layer in the *ab* plane. The layers are then interconnected by formate ions that link screw related histidine molecules from adjacent layers. The formate ions also serve to lend further stability to the Zl ribbons by linking alternate histidine molecules, i.e., those related by a *b* translation in the Zl sequence. In fact the histidine molecules and formate ions in the structure form a characteristic interaction pattern involving a linear chain of alternating imidazole rings and formate ions (figure 3), which is similar to the ones containing amino and carboxylate groups observed earlier (Vijayan 1988 Soman *et al* 1988).

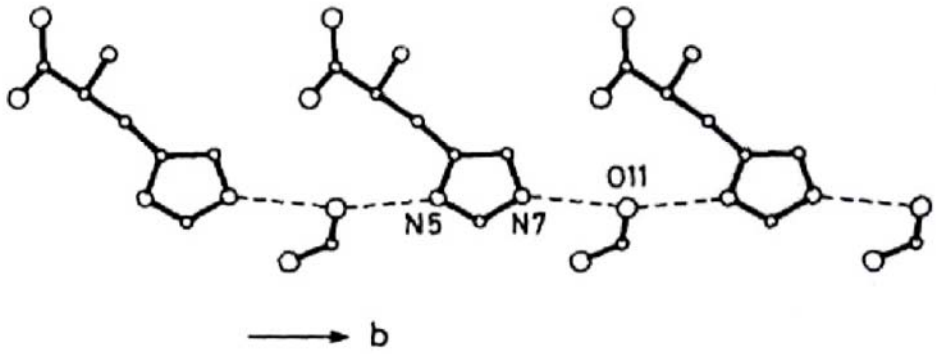

Figure 3. The characterise interaction pattern observed in the L-histidine complex.

The amino acid molecules in DL-histidine formate monohydrate (figure 2) form columns along *c,* through DL2 type head-to-tail sequences (Suresh and Vijayan 1983b) involving a Nl . . . 02 hydrogen bond and its glide equivalents. The DL2 sequence has earlier been observed in the structures of DL-histidine (Edington and Harding 1974) and its hydrochloride (Bennet *et al* 1970). In the crystals of DL-histidine, where the imidazole group is uncharged, this sequence is further stabilized by an intermolecular Nl . . . N5 interaction, while in the hydrochloride it is stabilized by an intermolecular N5 . . . 01 hydrogen bond. A N5 . . . 01 hydrogen bond exists in DL-histidine formate monohydrate also, but it is between screw related molecules and serves to connect screw related ribbons to form a corrugated sheet in the *bc* plane. Adjacent sheets slacked along *a* arc interconnected by formate ions and water molecules. In fact, each of the channels created by the stacking of amino acid sheets is filled by a hydrogen bonded zigzag siring containing alternate formate ions and water molecules (figure 4).

The modes of amino acid aggregation show substantial invariance in the complexes involving the other two basic amino acids (Suresh and Vijayan 1983a,c; Soman

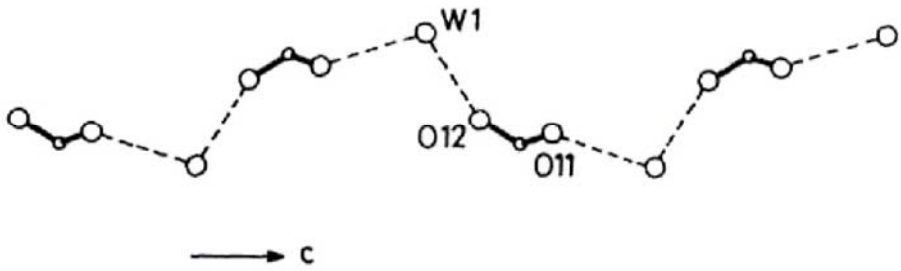

Figure 4. The string of alternating formate ions and water molecules in the DL-histidine complex.

et al. 1989; Prasad and Vijayan 1990, 1991: Stephen Suresh *et al* 1994b; Stephen Suresh and Vijayan 1994). This is particularly true of lysine complexes. In some of them, the interactions and aggregation patterns remain nearly the same, irrespective of changes in the nature and the size of the molecules or ions present in the system (Stephen Suresh and Vijayan 1994). The situation is somewhat different in histidine complexes, with considerable variability among them in the aggregation patterns. The present structures exemplify this situation.

Acknowledgements

The authors would like to thank ISRO, Department of Space, New Delhi for financial support. One of us (SS) is a Senior Research Fellow of the Council of Scientific and Industrial Research, New Delhi. The computations were done in the Supercomputer Education and Research Centre at the Institute.

References

Bhat T N and Vijayan M 1978 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; *Acta Crystallogr.* **B34** 2556–2565

Bennell I, Davidson A G H. Harding M M and Morelle I 1970 The crystal structure of DL-histidine hydrochloride dihydrale; *Acta Crystallogr.* **B26** 1722–1729

CRC Handbook of Chemistry and Physics 1980 (eds) R C Weast und M J Astle (Florida. CRC Press)

Eddington P and Harding M M 1974 The crystal structure of DL-histidine:. Äcta *Crystallogr* **B30** 204-206

lUPAC-IUB Commission on Biochemical Nomenclature 1970 Abbreviations and symbols for the description of the of the conformation of polypeptide chains; *J. Mol. Biol.* **52** 1–17

Krause J A. Baures P W and Eggleston D S 1991 Histidyl conformations and the short N–H . . .N hydrogen bonds: Structure of D,L-histidyl-L, D-histidine Pentahydrate; *Acta Crystallogr.* **B47** 506-511

Kvenvolden K A, Lawless J G and Ponnamperumu C 1971 Non-protein amino acids in the Murchison meteorite; *Proc. Natl. Acad. Sci. USA* **68** 486–490

Madden J J. McGandy E L and seeman N C 1972 The crystal structure of the othorhombic form of L -(+)histidine; *Acta Crystallogr.* **B28** 2377-23R2

Miller S L and Orgel L E 1974 *The origins of life on earth* (New Jersey: Prentice-Hall)

Miller S L and Urey H C 1959 Organic compound synthesis on the primative earth; *Science* **130** 245-251

Mitra J and Ramakrishnan C 1981 Studies on hydrogen bonds; *Int. J. Peptide Protein Res.* **17** 401-411

Miyazawa T and Pitzer K S 1959 Internal rotation and infrared spectra of formic acid monomer and a normal coordinate treatment of out-of-plane vibrations of monomer, dimer and polymer; *J. Chem. Phys.* **30** 1076-1078

Motherwell W D S and Clegg W 1978 *PLUTO Program for plotting molecular and crystal structures* (England: Univ of Cambridge)

Prasad G S and Vijayan M 1990 X-ray studies on crystalline complexes involving amino acids and peptides. XIX. Effects of change in chirality in the complexes of succinic acid with DL- and L-arginine; *Int. J. Peptide Protein Res.* **35** 357-364

Prasad G S and Vijayan M 1991 X-ray studies on crystalline complexes involving amino acids and peptides. XXIII. Variability in ionization state, conformation and molecular aggregation in the complexes of succinic acid with DL- and L-lysine; *Acta Crystallogr.* **B47** 927-935

Prasad G S and Vijayan M 1993a X-ray studies on crystalline complexes involving amino acids and peptides. XXV. Structures of DL-proline hemisuccinic acid and glycly-L-histidinium semisuccinate monohydrate and a comparative study of amino acid and peptide complexes of succinic. acid; *Acta Crystallogr* **B49** 348–356

Prasad G S and Vijayan M 1993h X-ray studies on crystalline complexes involving amino acids and peptides. XXIV, Different ionization states and novel aggregation patterns in the complexes of succinic acid with DL- and L-histidine; Biopolymers **33** 283-292

Ramanadham M and Chidambaram R 1978 Amino acids: Systematics of molecular structure, conformation and hydrogen bonding; in *Advances in crystallography (ed.)* R Srinivasan (New Delhi: Oxford and IBH) pp 81-103

Sheldrick G M 1986 SHELXS-86. *Program for the solution of crystal structures* (Germany: Univ. of Gottingen)

Sheldrick G. M 1993 SHELXL-93. *Program for refinement of crystal structures* (Germany: Univ. of Gottingen)

Soman J. Suresh C G and Vijayan M 1988 X-ray studies on crystalline complexes involving amino acids and peptides. XV. Crystal structures of L-lysine D-glutamate and L-lysine D-aspartate monohydrate and the effect of chirality on molecular aggregation: *int. J. Peptide Protein Res.* **32** 352-360

Soman J. Rao T, Radhakrishnan R and Vijayan M 1989 X-ray studies on crystalline complexes involving amino acids and peptides. Part XX. Crystal structures of DL-arginine acetate monohydrate and DL-lysine acetate and a comparsion with the corresponding L-amino acid complexes; *J. Bimol Struct. Dynamics* **7** 269-277

Stephen Suresh, Prasad G S and Vijayan M 1994a X-ray studies on crystalline complexes involving amino acids and peptides. XXVI. Crystal structures of two forms of L-histidine acetate and a comparative study of the amino acid complexes of acetic acid; *Int. J. Peptide Protein Res.* **43** 139-145

Stephen Suresh. Padmanabhan S and Vijayan M 1994b X-ray studies on crystalline complexes involving amino acids and peptides. XXVII. Effect of chirality. specific interactions and characteristic aggregation patterns in the structures of arginine and its complexes with formic acid; *J. Biomol Struct, Dynamics* **6** 1425-1435

Stephen Suresh and Vijayan M 1994 X-ray studies on crystalline complexes involvin$_g$ amino acids and peptides, XXVIII. Recurrence of characteristic aggregation and interaction patterns in the crystal structures of DL- mid L-lysine formate; *Acta Crystallogr,* Sect. B (in press)

Suresh C G and Vijayan M 1983a X-ray studies on crystalline complexes involving amino acids and peptides. VII. Head-lo-tail arrangement and a specific interaction in the crystal structure of L-arginine acetate; *Int. J. Peptide Protein Res.* **21** 223-226

Suresh C G and Vijayan M 1983b Occurence and geometrical features of head-to-tail sequences involving amino acids in crystal structure; *Inst. J. Peptide Protein Res.* **22** 129-143

Suresh C G and Vijayan M 1983c X-ray studies on crystalline complexes involving amino acids and peptides. X. Head-to-tail sequences in the crystal structure of L-lysine acetate; *Int J. Peptide Protein Res.* **22** 617–621

Suresh C G and Vijayan M 1985 Head-to-tail sequences and other patterns of peptide aggregation in the solid state; *Int. J. Peptide Protein Res.* **26** 311-328

Vijayan M 1980 Prebiotic evolution; An approach to polymerisation based on crystal structures; *FEBS Lett.* **112** 135–137

Vijayan M 1988 Molecular interactions and aggregation involving amino acids and peptides and their role in chemical evolution; *Prog, Biophys. Mol. Biol.* **52** 71–99s