X-ray Study of L-Arginine HCl, L-Cysteine, DL-Lysine and DL-Phenylalanine

By B. Khawas

Division Of Agricultural Physics, Indian Agricultural Research Institute, New Delhi-12, India

Preliminary crystal-structure data for L-arginine HCl,, L-cysteine, DL-lysine and DL-phenylalanine have been derived from an analysis of their powder patterns. L-arginine HCl, crystallized from water, is tetragonal; $a = 13\cdot95$ (4), $c = 10\cdot57$ Å, space group $P422$ and $Z = 8$. L-cysteine and DL-phenylalanine are monoclinic. L-cysteine has $a = 13\cdot26$ (4), $b = 9\cdot65$ (7), $c = 7\cdot14$ (3) Å, $\gamma = 96°11'$, space group $P2$ and $Z = 8$. DL-phenylalanine has $a = 12\cdot93$, $b = 10\cdot03$, $c = 6\cdot42$ (8) Å, $\gamma = 99°12'$, space group $P2$ and $Z = 4$. DL-lysine is orthorhombic; $a = 8\cdot89$, $b = 11\cdot04$ (7), $c = 7\cdot13$ (3) Å, space group $P222$ and $Z = 4$.

Introduction

The crystal structures of amino acids and their complexes have been studied in detail from good single-crystals. A few of them, which do not produce single-crystals, have been studied in this institute. Karle & Karle (1964) reported the crystal structure of L-arginine dihydrate. No work on the crystal structure of L-arginine HCl, L-cysteine, DL-lysine and DL-phenylalanine has yet been reported. L-Arginine HCl can replace L-arginine in artificial dietary mixtures of amino acids in place of protein. L-Cysteine is a constituent of wool keratin. DL-Lysine and DL-phenylalanine are common constituents of proteins. The present paper describes a derivation of the unit cells and probable space groups of these compounds, which may be useful for interpreting X-ray diffraction photographs of animal materials like silk, wool, proteins, *etc.*

Analysis

Pure L-arginine HCl, L-cysteine and DL-phenylalanine, obtained from the British Drug Houses Ltd, England, were each crystallized from water after prolonged evaporation at room temperature. L-Cysteine was obtained as thread-like fibre crystals and the rest were all microcrystalline. Further purification was obtained by recrystallization. Powder patterns were taken on an 11·48 cm diameter camera using Cu $K\alpha$ radiation. Line positions were measured to an accuracy of 0·05 mm and their spacings were calculated after the usual corrections for film shrinkage. A rotation pattern of L-cysteine crystals about the fibre axis was taken. The observed data have been analysed by systematic procedures. The correctness of indexing was checked by de Wolff's (1968) test in which the actual discrepancy was compared with the maximum average discrepancy for an arbitrary solution and a 'figure of merit' was defined:

$$M_{20} = Q_{20}/2\bar{\epsilon}N_{20}, \qquad (1)$$

where $Q_{20} = Q$ value for the 20th observed and indexed line, $\bar{\epsilon} =$ actual discrepancy for these twenty lines, $N_{20} =$ number of different Q values calculated up to Q_{20}. For a correct result, M_{20} should be greater than 10 provided that X_{20}, the number of non-indexed lines up to Q_{20}, does not exceed 2. For comparatively less accurate data, M_{20} may be less than 10.

L-*Arginine* HCl

The observed data did not fit into the cubic system and tests (Azároff & Buerger 1958) showed that the crystal system might be tetragonal or hexagonal. The pattern was successfully analysed by Hesse's (1948) method for the tetragonal system. The constant $10^4/a^2 = 51\cdot3$ could be taken to identify the 1st, 2nd, 4th, 5th, 7th, 10th, 13th, 14th and 17th line for $hk0$ indices (Table 1). A value of $10^4/c^2 = 89\cdot5$ from $Q_{006} = 3221$ and $Q_{007} = 4386$ then explains all the observed lines. Value $Q_{20} = 1366$, $\bar{\epsilon} = 3\cdot15$, and $N_{20} = 24$; therefore $M_{20} = 9$ from equation (1), a result that does not indicate complete reliability. Zubunko (1968) suggested that a line-pair of high-order indices should be chosen to calculate either the a or c dimension of a tetragonal crystal such that either $H_i/L_i - H_k/L_k$ or $L_i/H_i - L_k/H_k$ is a maximum, where $H_i = h_1^2 + k_1^2$, $H_k = h_2^2 + k_2^2$, $L_i = l_1^2$ and $L_k = l_2^2$. Two lines of indices, 751 and 406, were chosen to satisfy this condition for the a axis. The unit cell has $a = 13\cdot954$, $c = 10\cdot57$ Å, and if $Z = 8$ the calculated density is $1\cdot355$ g.cm^{-3}. Measured density is $1\cdot33$ (5) g.cm^{-3}. The conditions limiting possible reflexions are nil; the probable space group for an optically active molecule is $P422$.

Table 1. *Powder data with indices for* L-*arginine* HCl

Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
m	9·855 Å	103	103	110
w	6·864	212	205	200
s	5·285	358	358	002
s	4·968	407	409 ⎱	102 ⎱
			410 ⎰	220 ⎰
vw	4·682	456	461 ⎱	112 ⎱
			462 ⎰	300 ⎰
m	4·493	495	500	221
m	4·396	517	513	310
vs(bd.)	4·216	563	563	202
vw	4·023	618	615	212
m	3·850	675	667	320
w	3·652	750	757	321

Table 1 (cont.)

Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
s	3·521	807	806	003
s	3·479	826	821	400
w	3·388	871	871	312
			872	410
vs	3·321	907	908	113
			910	401
vw	3·224	962	962	411
m	3·121	1027	1026	420
m	3·059	1069	1062	213
vw	2·867	1217	1216	223
w	2·706	1366	1372	501
s	2·647	1427	1423	511
			1432	004
m	2·546	1543	1535	114
vw	2·478	1629	1626	403
			1637	204
s	2·409	1724	1729	333
w	2·295	1899	1894	304
			1898	610
w	2·247	1981	1988	601
s	2·209	2049	2052	620
w	2·159	2145	2139	503
			2142	621
m	2·114	2238	2237	005
vw(bd.)	2·023	2443	2443	205
vw	1·961	2600	2603	703
vw	1·937	2665	2667	632
			2668	640
vw	1·835	2970	2975	730
w	1·805	3069	3065	731
			3075	444
w	1·762	3221	3219	651
			3222	006
m	1·722	3372	3371	713
			3373	801
w	1·693	3489	3484	624
			3487	652
			3488	820
vw	1·603	3892	3886	751
			3889	326
m	1·576	4036	4042	406
w	1·510	4386	4386	007

L-Cysteine

The fibre repeat was calculated as $c = 7\cdot01$ Å from measurements of layer-line separations in the rotation pattern. Spacings for all reflexions up to the first layer line were recorded (Table 2). The first two lines were assumed to be second orders of 100 and 010, and suitable values of a^*, b^*, γ^* were determined from powder data (Table 3), which explained all equatorial spots. Value $Q_{100} = 57\cdot5$ was taken from $Q_{500} = 1439$ and $Q_{600} = 2064$. The constants $Q_{010} = 108\cdot5$ and $2(Q_{100} \cdot Q_{010})^{1/2} \cos \gamma^* = 17$ were refined from observed pairs of 320 and 330 reflexions. Value $Q_{001} = 196$ was taken from $Q_{202} = 1014$ and $Q_{022} = 1218$, with which all lines were indexed within discrepancy $Q_{obs} \sim Q_{calc} = 8$ for monoclinic symmetry using formula:

$$Q_{hkl} = \sum h^2 Q_{100} + 2hk(Q_{100} \cdot Q_{010})^{1/2} \cos \gamma^*, \qquad (2)$$

where $Q_{20} = 1657$, $\bar{\varepsilon} = 3\cdot2$, $N_{20} = 23$; thus $M_{20} = 11$, indicating that the indexing is essentially correct. Cell dimensions, calculated from these constants, are $a = 13\cdot264$, $b = 9\cdot657$, $c = 7\cdot143$ Å and $\gamma = 96°\,11'$. These dimensions are reduced by Buerger's (1957) procedure and correspond to a primitive unit cell. Measured density is $1\cdot74$ (5) g.cm^{-3}; the density calculated for eight molecules per unit cell is $1\cdot767$ g.cm^{-3}. The conditions limiting possible reflexions are nil; the probable space group is $P2$.

Table 2. Data on the rotation pattern of L-cysteine

	Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
Zero-layer line	m	6·75 Å	220	230	200
	vs	4·84	427	434	020
	m	4·08	601	596	$2\bar{2}0$
	vs	3·86	671	677	310
	vs	3·67	742	732	220
	w	3·41	860	850	$3\bar{2}0$
	s	2·86	1222	1218	$4\bar{2}0$
	vs	2·64	1435	1438	500
	w	2·48	1626	1631	510
	vs	2·45	1666	1657	330
	w	2·34	1826	1830	$2\bar{4}0$
First-layer line	s	6·15	264	254	101
	s	5·33	352	345	$1\bar{1}1$
	vs	4·91	415	424	201
	vs	4·16	578	569	211
	vs	3·40	865	873	311

Table 3. Powder data with indices for L-cysteine

Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
w	6·707 Å	222	230	200
w	6·176	262	254	101
w	5·330	352	345	$1\bar{1}1$
m	4·843	426	424	201
			434	020
w	4·467	501	501	$2\bar{1}1$
m	4·206	565	569	211
vw	4·072	603	596	$2\bar{2}0$
w	3·842	677	677	310
vs(bd.)	3·675	740	732	220
vvw	3·382	874	873	311
vw	3·274	933	928	221
			933	$1\bar{1}2$
m	3·140	1014	1014	202
vw	3·080	1054	1046	$3\bar{2}1$
			1054	320
w	2·939	1158	1157	212, $4\bar{1}1$
m	2·865	1218	1218	022, $4\bar{2}0$
m	2·724	1348	1341	$3\bar{3}0$
m	2·636	1439	1438	500
vw	2·576	1507	1505	231
vw	2·544	1545	1537	$3\bar{3}1$
vw	2·457	1657	1657	330
m	2·435	1687	1686	421
vw	2·377	1770	1764	003
vw	2·337	1831	1827	511
			1830	$2\bar{4}0$
w	2·262	1954	1947	113
w	2·222	2025	2026	$2\bar{4}1$
vvw	2·201	2064	2056	141
			2069	213
			2070	600
vvw	2·180	2104	2101	430
			2102	240
vvw	2·153	2157	2159	$5\bar{3}0$
vw	2·017	2458	2458	340
vvw	1·956	2614	2614	$2\bar{4}2$
vw	1·898	2776	2773	$2\bar{5}0$
vw	1·845	2938	2937	$6\bar{3}1$
			2943	$5\bar{3}2$

Table 3 (cont.)

Intensity	d	10⁴/d²(obs)	10⁴/d²(calc)	hkl
vw	1·809	3056	3051 ⎱ 3064 ⎰	151 ⎱ 612 ⎰
vvw	1·770	3192	3194	104
vw	1·712	3412	3411	333
vw	1·623	3796	3798	7$\bar{2}$2
w	1·594	3936	3932	2$\bar{6}$0

DL-*Lysine*

Tests (Azároff & Buerger, 1958) of the observed data (Table 4) showed that the crystal system was not cubic, tetragonal, or hexagonal. The pattern was successfully analysed by Lipson's (1949) method for orthorhombic symmetry. Constants $10^4 \cdot \sin^2 \theta_{100} = 75 \cdot 2$ and $10^4 \cdot \sin^2 \theta_{010} = 48 \cdot 7$ were refined to identify $h00$ and $0k0$ reflexions for $h = 1, 2, 3, 4$ and $k = 2, 3, 4, 7$. The value $10^4 \cdot \sin^2 \theta_{001} = 116 \cdot 8$ was taken from the difference diagram and lines were indexed satisfactorily, where $Q_{20} = 1770$, $\bar{\varepsilon} = 2 \cdot 36$, $N_{20} = 22$, $X_{20} = 1$; thus $M_{20} = 17$ from equation (1) indicates correct indexing. The unit cell has $a = 8 \cdot 89$, $b = 11 \cdot 047$, $c = 7 \cdot 133$ Å, and, if $Z = 4$, the calculated density is $1 \cdot 386$ g.cm⁻³. Measured density is $1 \cdot 36$ (5) g.cm⁻³. The conditions limiting possible reflexions are nil; the probable space group is $P222$ for an organic molecule.

Table 4. *Powder data with indices for* DL-*lysine*

Intensity	d	(10⁴sin²θ)	(10⁴sin²θ)	hkl
m	8·876 Å	75	75	100
vw	7·139	117	117 ⎱ 124 ⎰	001 ⎱ 110 ⎰
w	5·985	166	166	011
vw	5·581	191	192 ⎱ 195 ⎰	101 ⎱ 020 ⎰
s	4·659	274	270	120
w	4·444	301	301	200
vs	4·119	350	350	210
m(bd.)	3·911	389	387	121
w	3·685	438	438	030
vs	3·479	491	496	220
s	3·394	516	514	130
w(bd.)	3·321	539	542	102
w	3·257	559	555	031
vw	3·071	630	630	131
w	3·958	679	677	300
w	2·867	723	726	310
w	2·761	779	779	040
vs	2·697	817	817	212
w	2·612	871	872	320
w	2·550	914	–	–
vw	2·377	1052	1051	003
s	2·324	1101	1100	013
w	2·222	1203	1203 ⎱ 1206 ⎰	400 ⎱ 232 ⎰
m	2·175	1256	1252	410
vw	2·126	1315	1320	401
w	2·096	1353	1352	203
w(bd.)	2·050	1414	1410	151
w	2·001	1485	1490	033
w	1·981	1514	1515 ⎱ 1518 ⎰	421 ⎱ 250 ⎰
vvw	1·937	1583	1582	332
vw	1·900	1646	1642	430
w	1·829	1776	1777	313

Table 4 (cont.)

Intensity	d	(10⁴sin²θ)	(10⁴sin²θ)	hkl
w	1·787	1861	1865	422
w(bd.)	1·734	1977	1982	440
vw	1·692	2076	2075	520
vw	1·656	2167	2167 ⎱ 2170 ⎬ 2171 ⎰	333 ⎱ 204 ⎬ 261 ⎰
vw	1·635	2222	2218 ⎱ 2220 ⎰	214 ⎱ 062 ⎰
vw	1·604	2310	2307	034
w	1·578	2387	2382 ⎱ 2386 ⎰	134 ⎱ 070 ⎰
vw(bd.)	1·542	2500	2503	071
vw	1·517	2582	2578	171
w	1·493	2649	2648	044
w	1·477	2726	2723	144

DL-*Phenylalanine*

The observed data (Table 5) did not fit into any of the higher crystal systems and Ito's (1950) method for low-symmetry crystals was tried. The first line was assumed to be 100 and $Q_{100} = 61 \cdot 4$ was taken from $Q_{600} = 2212$ and $Q_{700} = 3009$. Value $Q_{010} = 102$ was taken from $Q_{040} = 1632$, and the 2nd and 3rd Q values were found to be symmetrical about $Q_{100} + Q_{010}$. These two lines were given indices $1\bar{1}0$ and 110. Equation (2) was written in the form: $Q_{hkl} = T_i + l^2 Q_{001}$, and the tetrad (100, 010, $1\bar{1}0$, 110) of order $n = \pm 1$ satisfied Zachariasen's (1963) condition: $2(T_1 + T_2) = T_3 + T_4$. This indicated that all lines could be explained for a monoclinic cell, provided a suitable value of Q_{001} was found. Value $Q_{001} = 242$ was taken from $Q_{302} = 1521$ and $Q_{013} = 2280$. The constant $2(Q_{100} \cdot Q_{010})^{1/2} \cos \gamma^* = 25 \cdot 3$ was refined from high-order indices 333 and 442, so that all lines were indexed satisfactorily, using equation (2). Value $Q_{20} = 1521$, $\bar{\varepsilon} = 3 \cdot 45$ and $N_{20} = 26$; thus $M_{20} = 8 \cdot 5$ from equation (1), which does not indicate complete reliability and is probably for less accurate data. The cell dimensions calculated from these constants are: $a = 12 \cdot 93$, $b = 10 \cdot 03$, $c = 6 \cdot 428$ Å and $\gamma = 99° 12'$. These dimensions are reduced by Buerger's (1957) procedure and correspond to a true primitive cell. Assuming $Z = 4$, the calculated density is $1 \cdot 333$ g.cm⁻³ and measured density is $1 \cdot 34$ (5) g.cm⁻³. No systematic absences of observed reflexions could be found; therefore, the probable space group for the phenylalanine molecule is $P2$.

Table 5. *Powder data with indices for* DL-*phenylalanine*

Intensity	d	10⁴/d²(obs)	10⁴/d²(calc)	hkl
s	13·45 Å	55	61	100
w	8·347	144	138	1$\bar{1}$0
m	7·116	197	189	110
s	6·380	246	242 ⎱ 246 ⎰	001 ⎱ 200 ⎰
m	5·441	338	344	011
s	5·178	373	380	1$\bar{1}$1
s(bd.)	4·772	439	431	111
w	4·386	520	520	120
s	4·246	555	552 ⎱ 553 ⎰	2$\bar{2}$0 ⎱ 300 ⎰

Table 5 (cont.)

Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
vw(bd.)	3·952	640	640	211
m	3·623	762	755 }	220 }
			762 }	121 }
m	3·485	823	821	341
m	3·337	898	904	$1\bar{3}0$
w	3·291	923	918	030
vw	3·079	1055	1051 }	321 }
			1055 }	130 }
s	2·957	1144	1146	131
s(bd.)	2·812	1265	1265	212
vw	2·752	1320	1315	230
s	2·639	1436	1428 }	411 }
			1430 }	$4\bar{2}1$ }
vw	2·654	1521	1520 }	$2\bar{2}2$ }
			1521 }	302 }
w	2·511	1586	1592	420, $1\bar{4}0$
w	2·475	1632	1632	040
w	2·423	1703	1698 }	330 }
			1699 }	312 }
vw	2·368	1783	1777	510, $3\bar{2}2$
vw	2·332	1839	1834 }	$1\bar{4}1$ }
			1835 }	421 }
m	2·248	1979	1980	$2\bar{3}2$
s	2·196	2074	2075 }	{ $5\bar{3}0$
			2080 }	{ 240, 322
w	2·126	2212	2204 }	{ 430
			2210 }	{ 600, $4\bar{4}0$
			2211 }	{ $3\bar{3}2$
m	2·094	2280	2280 }	013 }
			2283 }	232 }
w	2·038	2361	2367	113
vw	2·034	2417	2424	203

Table 5 (cont.)

Intensity	d	$10^4/d^2$(obs)	$10^4/d^2$(calc)	hkl
m	2·004	2490	2485 }	$1\bar{5}0$ }
			2488 }	340 }
vw	1·890	2799	2792	051
m	1·823	3009	3009	700
vw	1·743	3292	3288 }	710 }
			3290 }	323 }
			3291 }	251 }
vw	1·723	3368	3364	413
w	1·682	3535	3530	711
w	1·649	3678	3672 }	060 }
			3673 }	540 }
vw	1·606	3877	3872 }	004 }
			3876 }	333 }
w	1·584	3986	3988	442
w	1·547	4178	4172 }	801 }
			4174 }	$7\bar{4}1$ }

References

AZÁROFF, L. V. & BUERGER, M. J. (1958). *The Powder Method*, pp. 118–123. New York: McGraw Hill.
BUERGER, M. J. (1957). *Z. Kristallogr.* **109**, 42.
HESSE, R. (1948). *Acta Cryst.* **1**, 200.
ITO, T. (1950). *X-ray Studies On Polymorphism.* p. 157. Tokyo: Maruzen Co. Ltd.
KARLE, I. L. & KARLE, J. (1964). *Acta Cryst.* **17**, 835.
LIPSON, H. (1949). *Acta Cryst.* **2**, 43.
WOLFF, P. M. DE (1968). *J. Appl. Cryst.* **1**, 108.
ZACHARIASEN, W. H. (1963). *Acta Cryst.* **16**, 572.
ZUBUNKO, V. V. (1968). *Kristallographia*, **13** (5), 776.

Kristallstrukturen von Säurehydraten und Oxoniumsalzen. VI.*
Dioxonium-dihydrogenhypophosphat, (H₃O⁺)₂[H₂P₂O₆²⁻]†

VON DIETRICH MOOTZ‡ UND HORST ALTENBURG

Abteilung für Röntgenstrukturanalyse, Gesellschaft für Molekularbiologische Forschung m.b.H., 3301 Stöckheim über Braunschweig, Deutschland

(*Eingegangen am* 12. *Februar* 1970)

Hypophosphoric acid dihydrate (*Pccn*; $a=6·557$, $b=11·634$, $c=9·464$ Å; $Z=4$) is not a true acid hydrate, $H_4P_2O_6.2H_2O$, but an acid oxonium salt, $(H_3O^+)_2[H_2P_2O_6^{2-}]$, dioxonium dihydrogen hypophosphate. This is suggested by a loose isomorphism to the corresponding ammonium salt, $(NH_4)_2[H_2P_2O_6]$, and confirmed by the hydrogen atom positions, P–O bond lengths and hydrogen bonding scheme, as resulting from a complete structure analysis with three-dimensional X-ray counter data. The final R value for 683 observed independent reflexions was 0·043.

Die kristallinen Verbindungen von Wasser mit schwachen Säuren sind wahrscheinlich generell als echte Säurehydrate (HA.H₂O) und die mit starken Säuren als Oxoniumsalze (im einfachsten Fall: $H_3O^+A^-$) einzuordnen. Hypophosphorsäure, $(HO)_2OP–PO(OH)_2$, ist eine mittelstarke Säure ($K_1=0,64\times10^{-2}$; $K_2=0,15\times10^{-2}$) und bildet mit zwei Mol Wasser eine kristalline Verbindung, die bei 62°C schmilzt. Die hier zunächst offene Entscheidung zwischen einem echten Säurehydrat und einem Oxoniumsalz geschah mit Hilfe einer Kristallstrukturanalyse, deren Durchführung und Ergebnisse in folgenden beschrieben werden.

---

* Mitteilung V: Mootz & Fayos (1970).

† Die Ergebnisse dieser Arbeit wurden in einem grösseren Zusammenhang auf dem VIII International Congress of Crystallography in Stony Brook, New York, U.S.A., vorgetragen (Mootz, Altenburg, Fayos & Wunderlich, 1969).

‡ Sonderdrucke von D. Mootz; gegenwärtige Adresse: Max-Planck-Institut für Festkörperforschung, 7 Stuttgart 1, Heilbronner Str. 69, Deutschland.