

mined by the closeness with which the layers can be packed together, and therefore bears a direct relationship to the length of the $X-Al_2$ and Al_1-Al_2 bonds (as the latter two bonds link atoms in the different layers, their lengths are related to the length of the hexad axis). Thus all the significant differences in bond lengths are associated with bonds involving Al_2 atoms, and as all these bonds are longer in the $\varphi(AlMn)$ structure than the $\beta(AlMnSi)$ structure it seems possible (bearing in mind that the atomic radius of aluminium is greater than that of silicon) that the silicon atoms in $\beta(AlMnSi)$ preferentially occupy some of the $12(k)$ sites in the structure (i.e. the sites occupied by Al_2 atoms in $\varphi(AlMn)$).

Conclusion

The structure of Mn_3Al_{10} has been accurately determined and its similarity with that of $\beta(AlMnSi)$ suggests that the silicon atoms in $\beta(AlMnSi)$ do not occupy the origin sites exclusively as had been previously assumed. The significant bond-length differences between the two structures indicate that the occupation of the $12(k)$ sites by silicon in $\beta(AlMnSi)$ is probably greater than the occupation of the $2(a)$ and $6(h)$ sites.

I wish to thank the Department of Scientific and Industrial Research for the award of a Research Studentship. This research forms part of a project supported by the Air Force Office of Scientific Re-

search of the Air Research and Development Command, United States Air Force, through its European Office, under Contract No. AF. 61(052)-50. Some of the equipment used was purchased with the help of a generous award from the Government Grants Committee of the Royal Society.

I am greatly indebted to Prof. N. F. Mott and Dr W. H. Taylor for provision of facilities and for their interest and encouragement. I should like to thank Dr P. Jane Brown for many valuable discussions and Miss Gay Schroeder for computational assistance. I am indebted to Dr M. V. Wilkes and Mr E. N. Mutch, of the Mathematical Laboratory, for permission to use the Edsac Digital Computer. Aluminium Laboratories Ltd., Banbury, kindly gave the super-purity aluminium used in this work.

References

- BLAND, J. A. (1958). *Acta Cryst.* **11**, 236.
 CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65.
 FARQUHAR, M. C. M. & LIPSON, H. (1946). *Proc. Phys. Soc. Lond.* **58**, 200.
 HANSEN, M. (1958). *Constitution of Binary Alloys*. New York: McGraw Hill.
 LIPSON, H. & COCHRAN, W. (1953). *The Crystalline State*, vol. III. London: Bell.
 NICOL, A. D. I. (1953). *Acta Cryst.* **6**, 285.
 PRATT, J. N. & RAYNOR, G. V. (1951). *Proc. Roy. Soc. A*, **205**, 103.
 ROBERTSON, J. M. (1943). *J. Sci. Instrum.* **20**, 175.
 ROBINSON, K. (1952). *Acta Cryst.* **5**, 397.
 VIERVOLL, H. & ÖGRIM, O. (1949). *Acta Cryst.* **2**, 277.

Acta Cryst. (1959). **12**, 396

The Crystal Structure of Hexagonal L-Cystine

BY BERYL M. OUGHTON AND PAULINE M. HARRISON*

Chemical Crystallography Laboratory, South Parks Road, Oxford, England

(Received 9 July 1958 and in revised form 7 October 1958)

The structure of hexagonal L-cystine, space group $P6_122$, has been solved by three-dimensional methods. The structure is essentially composed of glycine-like sheets lying perpendicular to the c -axis with the C-R bonds pointing alternately up and down in successive sheets and linked together in pairs by disulphide bridges. There is a very satisfactory three-dimensional network of hydrogen bonds linking the molecules together. Crystallographic data for the space group $P6_122$ are given.

Introduction

The work described here on the structure of hexagonal L-cystine $[S-CH_2-CH(NH_2)-COOH]_2$ is part of a programme of research on the structures of proteins and peptides which is in progress in this laboratory, and has a particular bearing on the structures of insulin and gramicidin S.

Insulin contains three cystinyl linkages per molecule of 6,000 mW., and a knowledge of the possible conformations of these linkages is of great interest. Gramicidin S, a decapeptide, forms a series of derivatives, N-iodoacetyl, N-chloroacetyl and N-acetyl, which crystallize in the same space group as hexagonal L-cystine, namely $P6_122$ (Schmidt, Hodgkin & Oughton, 1957), and the latter has been used as a model for gaining experience in the analysis of structures crystallizing in this highly symmetrical space group.

* Present address: Biochemistry Department, The University, Sheffield, 10, England.

Experimental details

The crystallization of L-cystine in a form suitable for X-ray analysis presents some difficulties owing to its sparing solubility in water and its insolubility in alcohol. It can, however, be dissolved in alkali and precipitated by organic acids (*Beilstein's Organische Chemie*, 4, p. 507).

A small quantity of L-cystine was dissolved in warm 10% ammonia. Warm dilute acetic acid was added gradually and the volume found which would just bring the solution to the point of precipitation. The procedure was then repeated using the same amounts of reagents and the warm solution was allowed to cool down slowly. The crystals deposited were mainly in highly twinned clusters, but single optically uniaxial, hexagonal plates could also be picked out. The crystal from which the X-ray diffraction data were obtained was a hexagonal plate measuring about 0.3 mm. across by 0.1 mm. in thickness.

As given in a preliminary note (Oughton & Harrison, 1957) the unit cell is hexagonal with dimensions $a = 5.422$ and $c = 56.275$ Å.

$U = 1432.7$ Å³, $mW. = 240.2$, $D_m = 1.677$ g.cm.⁻³,*
 $D_x = 1.671$ g.cm.⁻³, $Z = 6$, $F(000) = 756$, $\mu = 49.3$ cm.⁻¹.

The space group is $P6_122$ with all the atoms in general (12-fold) positions. The two halves of each molecule must therefore be related to one another by one of the two-fold axes of symmetry imposed by the space group.

Three-dimensional (hkl) reflections out to the limit of Cu $K\alpha$ radiation were recorded using a Weissenberg camera and rotating the crystal about the $[10\bar{1}0]$ and $[11\bar{2}0]$ axes. The intensities were estimated visually, Lorentz and polarization factors applied, and the absolute scale and average temperature factor estimated by Wilson's method (Wilson, 1942). No corrections for absorption were made, and the average error in F from absorption effects is approximately 7%.

Derivation and refinement of the structure

In the earlier stages of the work no high-speed computing equipment was available and it was decided to attempt to solve the structure mainly in projection. Later, when we were able to have calculations done on an electronic computer three-dimensional methods were used entirely, and it is of interest to compare the relative success and speed of the two methods of approach.

Although the space group $P6_122$ is non-centrosymmetric, there are three centrosymmetric projections, perpendicular to $[10\bar{1}0]$, $[11\bar{2}0]$ and $[0001]$. The $(10\bar{1}0)$ and $(11\bar{2}0)$ projections belong to the plane group pmg , each equivalent point of the plane group representing three equivalent points of the space-

group. Thus in each projection one obtains a view of the $[-S.CH_2.CH(NH_2).COOH]$ group in three different orientations at 60° to one another. In both the $(10\bar{1}0)$ and $(11\bar{2}0)$ projections, the parameters of symmetry-related atoms forming the asymmetric unit of the plane group, are interrelated (see Appendix), and it is theoretically possible to determine all three parameters of each atom from one projection alone.

This space group contains two sets of two-fold axes, one parallel and the other perpendicular to the three a -axes. The two halves of a cystine molecule may therefore be related to one another by a two-fold axis of symmetry parallel to either $[10\bar{1}0]$ or $[11\bar{2}0]$, and the six S-S' bonds may be either parallel or perpendicular to the a -axes.

The $(10\bar{1}0)$ and $(11\bar{2}0)$ Patterson projections were

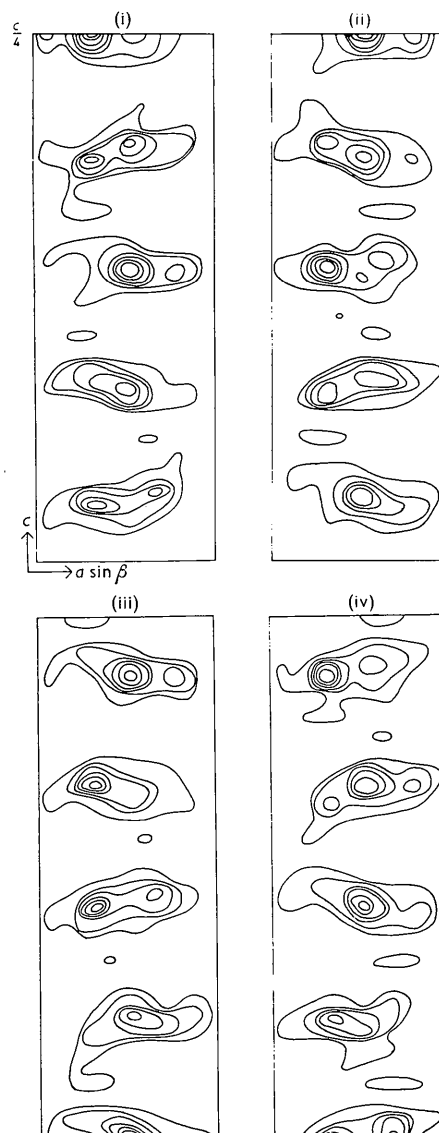


Fig. 1. Four possible $(10\bar{1}0)$ Fourier projections based on direct sign determination.

* As measured by Steinrauf & Jensen (1956).

calculated, but due to overlapping of peaks they gave no indication of the positions of the S atoms.

Direct sign-determining methods (Cochran, 1952; Zachariassen, 1952) were then tried on the $(h0\bar{h}l)$ and $(hh.2\bar{h}.l)$ reflections. Very few signs could be determined absolutely, owing to the rather low average \bar{U} value, but signs of 25 $(h0\bar{h}l)$ reflections with $\bar{U} \geq 0.2$ could be expressed in terms of two unknown signs. The four possible projections obtained by permuting the signs of the two unknowns were calculated (Fig. 1). Two of them (Fig. 1, (i) and (ii)), contain peaks which may be associated with reasonable S positions (Appendix), both with the S-S' bonds parallel to a , but at slightly different positions along the $[11\bar{2}0]$ axis.

In spite of the limitations imposed by the interrelation between parameters, many structures could be fitted to these two, rather badly resolved, projections, and difficulty was experienced in judging their general correctness.

It was therefore decided to calculate a three-dimensional Patterson synthesis in order to determine the S positions with more certainty. The three-dimensional Patterson synthesis for this space group contains five Harker sections (Harker, 1936; Buerger, 1946), three perpendicular to the c -axis, and two parallel to the c -axis. Sections at $z = \frac{1}{6}$, $\frac{1}{3}$ and $\frac{1}{2}$ represent the (0001) projection on a scale of 1, $\sqrt{3}$ and 2 respectively, the section at $z = \frac{1}{3}$ being rotated through 30° about the c -axis relative to the Fourier projection. Owing to the very long c -axis there is far too much overlapping of peaks in these sections to use them as a direct method of structure determination (Fig. 2(a)-(c)), but by comparison of the three sections one can see that most of the atoms are probably concentrated around the c -axis, with relatively little density in the central regions of the unit cell. Fig. 2(d) is the (0001) projection of the actual structure, and it can be seen that the sections do give the correct general distribution of atoms in this projection.

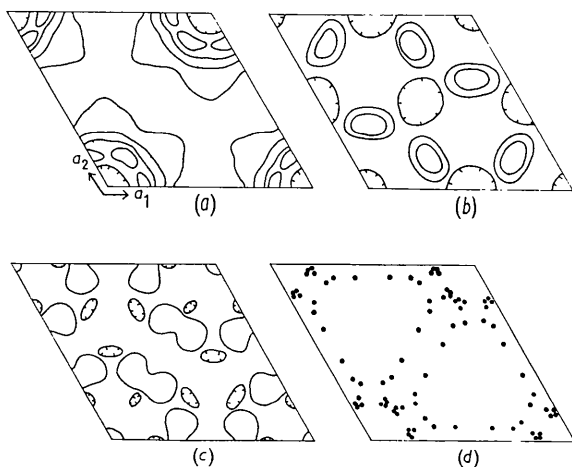


Fig. 2. Sections at (a) $z = \frac{1}{6}$; (b) $z = \frac{1}{3}$; (c) $z = \frac{1}{2}$ through the three-dimensional Patterson synthesis; (d) (0001) projection of the actual structure.

The remaining two Harker sections are at $x_1 = 0$ and $x_1 = -x_2$ and represent the $(10\bar{1}0)$ and $(11\bar{2}0)$ projections respectively on double the actual scale.

Assuming a S-S' bond length of 2.04 Å, a satisfactory set of peaks was found (see Appendix) in these two sections which could be associated with S-S interactions (Fig. 3). They place the S-S' bonds parallel to a , the two halves of each molecule thus being related by a 2-fold axis of symmetry perpendicular to a . As the positions of the peaks selected were not quite self-consistent, the best parameters for S were determined by least squares from the six equations relating the Patterson peak co-ordinates to the Fourier co-ordinates (Appendix), weighting the equations according to the width of the Patterson peak to which they referred.

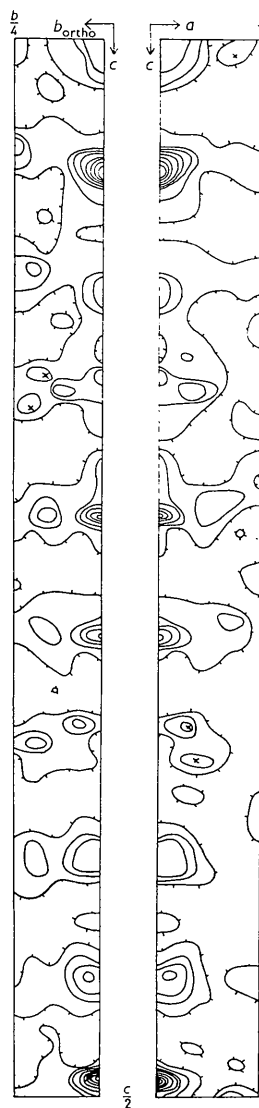


Fig. 3. Sections at $x_1 = 0$ and $x_1 = -x_2$ through the three-dimensional Patterson synthesis. The crosses mark the positions of the S-S interactions.

Table 1. Observed and calculated structure factors for hexagonal L-cystine

(a) Observed reflections

$h k l$	F_o	F_c	$h k l$	F_o	F_c	$h k l$	F_o	F_c	$h k l$	F_o	F_c	$h k l$	F_o	F_c
0 0 6	17.0	15.2	0 2 31	39.1	37.1	0 4 42	5.0	0.1	1 2 3 0	69.6	65.8	1 3 44	15.6	16.0
12	26.7	35.4	32	44.1	47.8	43	15.6	15.0	1	37.1	35.5	44	23.4	24.8
18	261.3	332.9	33	27.8	30.4	44	17.3	16.4	2	44.0	40.7	45	23.4	24.8
24	41.6	37.3	34	44.7	46.3	45	6.3	5.4	3	25.6	22.4	46	7.5	10.4
30	6.5	4.8	35	11.0	11.3	46	13.1	12.1	4	28.8	26.2	47	16.5	13.2
36	128.0	140.3	36	14.5	18.4	47	8.0	7.5	5	24.6	19.1	48	9.8	10.8
42	32.7	35.3	37	2.9	6.5	48	10.6	10.6	6	11.7	11.0	49	5.6	9.0
48	25.7	23.7	38	17.0	15.9	49	5.6	6.4	7	25.2	20.7	50	2.9	3.3
54	23.6	15.2	39	2.9	14.4	50	5.0	5.0	8	4.1	4.7	51	11.4	11.6
60	11.7	12.4	40	32.1	33.7	51	9.7	9.7	9	4.1	4.7	52	4.1	5.2
0 1 7	188.5	174.3	41	17.7	15.2	52	21.9	20.6	10	45.5	41.0	53	8.0	8.0
1	70.6	57.3	42	24.9	25.7	53	11.4	11.7	11	23.6	21.0	54	11.7	10.9
2	36.0	28.3	43	7.5	9.7	54	5	5	12	24.4	23.4	55	4.1	4.9
3	29.7	18.7	44	49	29.2	55	13.9	15.0	13	12.4	11.6	56	19.9	17.8
4	27.7	28.2	45	31.5	30.7	56	8.0	7.5	14	43.5	41.1	57	9.0	10.3
5	26.9	25.0	46	15.1	15.1	57	11.4	13.4	15	45.6	43.5	58	13.9	10.8
6	45.9	36.5	47	5.2	13.5	58	8.0	8.7	16	23.6	23.6	59	9.6	9.6
7	54.7	51.2	48	2.9	6.8	59	14.2	5.9	17	63.8	63.8	60	13.9	9.6
8	9.2	5.3	49	10.6	8.8	60	17.5	20.5	18	19.5	11.4	61	10.8	10.8
9	64.5	34.3	50	5.0	5.0	61	20.1	20.1	19	24.2	24.1	62	15.3	10.8
10	11.4	10.8	51	59	10.4	62	14.9	14.9	20	20.1	19.2	63	15.3	11.7
11	64.5	7.6	52	19.4	17.9	63	14.2	23	21	19.2	19.2	64	10.8	10.8
12	103.5	96.6	53	5.0	4.2	64	9.2	11.3	22	15.2	16.8	65	11.0	11.0
13	99.9	100.7	54	64	12.2	65	2.8	2.2	23	9.5	9.5	66	11.0	14.4
14	17.0	13.9	55	68	18.9	66	5.3	11.6	24	10.1	11.4	67	14.4	13.0
15	34.0	35.1	56	14.3	10.7	67	10.1	10.1	25	11.4	14.3	68	16.0	8.6
16	123.6	121.1	57	5.2	15.0	68	11.4	11.4	26	10.1	11.4	69	15.2	8.6
17	15.0	12.1	58	7.4	5.8	69	8.9	4.2	27	11.4	14.3	70	18.0	16.1
18	14.5	11.1	59	19.7	25.9	70	71.3	79.2	28	15.9	15.7	71	19.9	16.7
19	22.2	17.0	60	4	3.3	71	14.2	15.0	29	15.9	15.7	72	15.0	8.6
20	23	17.5	61	4	29.2	72	34.8	37.1	30	32.0	33.1	73	21	11.7
21	36.5	16.7	62	49	22.2	73	44.5	46.2	31	34.9	18.9	74	23	14.3
22	30.6	20.9	63	11.7	10.4	74	20.1	23.1	32	36.2	36.2	75	10.2	10.1
23	27.1	35.6	64	6	24.5	75	15.6	16.8	33	38.0	22.5	76	12.7	13.1
24	11.4	11.0	65	13.9	13.7	76	6.4	7.5	34	36.0	9.1	77	12.4	11.6
25	41.4	41.8	66	20.7	25.0	77	9	2.9	35	31.0	31.7	78	10.1	10.1
26	9.4	5.7	67	36.1	36.7	78	71.3	79.2	36	22.5	22.4	79	30	13.3
27	31.8	46.0	68	11.2	14.2	79	19.6	19.6	37	15.0	14.8	80	33	15.0
28	86.2	62.6	69	12	23.0	80	24.9	110.8	38	10.2	6.1	81	35	11.0
29	33.7	34.8	70	13	13.3	81	34.8	37.1	39	16.2	16.6	82	33	11.7
30	74.9	104.4	71	15	41.2	82	44.5	46.2	40	16.2	16.6	83	37	13.1
31	28.3	23.7	72	19	31.6	83	15.6	16.8	41	6.3	3.5	84	37	10.0
32	24.3	47.7	73	30.6	32.4	84	23.4	26.5	42	6.3	6.0	85	38	5.0
33	10.6	7.7	74	47.7	50.4	85	20.7	21.7	43	4.1	3.6	86	41	5.9
34	13.9	12.8	75	19.9	23.4	86	15.3	16.2	44	13.9	11.3	87	42	7.5
35	44.4	46.5	76	30.8	31.4	87	28.1	29.9	45	15.8	17.7	88	43	9.3
36	7.5	6.5	77	24	23	88	6.3	6.9	46	8.0	8.0	89	46	6.3
37	24.4	21.0	78	24	20.5	89	33.2	32.9	47	6.4	7.2	90	47	8.6
38	11.7	13.6	79	28	19.1	90	44.1	44.0	91	9.6	9.6	1 5 0	17.0	13.1
39	47.5	47.9	80	29	15.0	91	41.2	37.7	92	5.0	6.2	2	8.0	4.1
40	34.5	34.0	81	32	19.4	92	76.6	70.4	93	27.8	26.1	3	13.0	5.4
41	25.5	23.0	82	33	16.3	93	9.0	9.6	94	10.3	8.3	4	9.8	6.4
42	50.9	47.9	83	30	10.6	94	18	26.3	95	10.3	8.3	5	13.9	9.3
43	25.5	23.0	84	33	26.6	95	20	38.3	96	5.6	4.8	6	12	15.8
44	30.8	27.1	85	34	8.9	96	40.6	42.4	97	13	15.6	7	15.6	8.9
45	12.0	10.4	86	38	9.0	97	22	23.6	98	12	15.6	8	15.6	8.9
46	30.2	26.2	87	6.3	6.1	98	23	23.9	99	15	7.5	9	15.6	9.7
47	18.0	18.0	88	12.7	13.0	99	22.3	22.8	1 3 0	11.7	12.5	10	9.4	8.2
48	8.1	8.1	89	13.0	13.6	100	5.0	5.2	1	11.7	11.1	11	18.0	15.6
49	11.7	11.6	90	15.6	14.8	101	22.3	27.5	2	23.4	26.6	12	25.5	5.9
50	30.0	24.2	91	12.0	9.9	102	30.0	33.6	3	11.7	11.1	13	28	7.0
51	8.0	13.1	92	46.5	43.3	103	11.4	16.0	4	10.2	10.3	14	28	3.5
0 2 31	39.1	37.1	93	13.0	13.6	104	11.4	16.0	5	29.2	31.0	15	29	7.3
1 1 7	188.5	174.3	94	9.0	9.8	105	5.6	7.0	6	23.6	23.5	16	30	24.6
1 2 3 0	69.6	65.8	95	11.7	14.5	106	5.6	7.0	7	18.2	19.1	17	31	20.0
1 3 44	15.6	16.0	96	13.9	15.0	107	11.7	14.5	8	18.2	19.1	18	32	18.0
2 3 5 1	19.9	17.8	97	14.2	15.3	108	13.9	15.0	9	18.2	19.1	19	33	16.0
2 4 6 0	19.9	17.8	98	14.2	15.3	109	13.9	15.0	10	18.2	19.1	20	34	14.0
2 5 7 0	19.9	17.8	99	14.2	15.3	110	13.9	15.0	11	18.2	19.1	21	35	12.0
2 6 8 0	19.9	17.8	100	14.2	15.3	111	13.9	15.0	12	18.2	19.1	22	36	10.0
2 7 9 0	19.9	17.8	101	14.2	15.3	112	13.9	15.0	13	18.2	19.1	23	37	8.0
2 8 10 0	19.9	17.8	102	14.2	15.3	113	13.9	15.0	14	18.2	19.1	24	38	6.0
2 9 11 0	19.9	17.8	103	14.2	15.3	114	13.9	15.0	15	18.2	19.1	25	39	4.0
2 10 12 0	19.9	17.8	104	14.2	15.3	115	13.9	15.0	16	18.2	19.1	26	40	2.0
2 11 13 0	19.9	17.8	105	14.2	15.3	116	13.9	15.0	17	18.2	19.1	27	41	1.0
2 12 14 0	19.9	17.8	106	14.2	15.3	117	13.9	15.0	18	18.2	19.1	28	42	0.5
2 13 15 0	19.9	17.8	107	14.2	15.3	118	13.9	15.0	19	18.2	19.1	29	43	0.2
2 14 16 0	19.9	17.8	108	14.2	15.3	119	13.9	15.0	20	18.2	19.1	30	44	0.1
2 15 17 0	19.9	17.8	109	14.2	15.3	120	13.9	15.0	21	18.2	19.1	31	45	0.0
2 16 18 0	19.9	17.8	110	14.2	15.3	121	13.9	15.0	22	18.2	19.1	32	46	0.0
2 17 19 0	19.9	17.8	111	14.2	15.3	122	13.9	15.0	23	18.2	19.1	33	47	0.0
2 18 20 0	19.9	17.8	112	14.2	15.3	123	13.9	15.0	24	18.2	19.1	34	48	0.0
2 19 21 0	19.9	17.8	113	14.2										

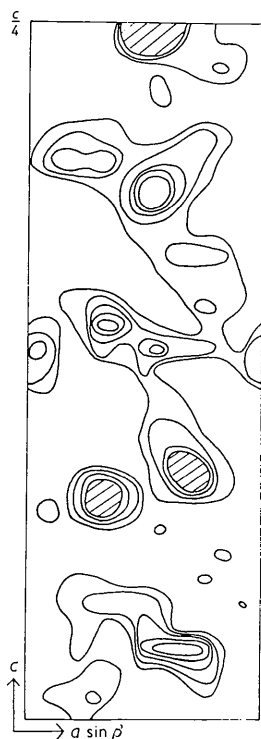


Fig. 4. $(10\bar{1}0)$ Fourier projection phased on the S position alone.

Structure factors were calculated for the $(h0\bar{h}l)$ and $(hh.2\bar{h}.l)$ reflections up to $h_{\max.} = 4$, $l_{\max.} = 36$ using the S positions alone. The $(10\bar{1}0)$ (Fig. 4) and $(11\bar{2}0)$ projections were then calculated using these signs. It may be seen that the second of the four projections (Fig. 1) calculated on the basis of the direct sign-determining method gives S parameters similar to those determined from the three-dimensional Patterson synthesis and has the same general appearance as the S-phased $(10\bar{1}0)$ projection. Subsequent comparison with the structure factors for the final, refined parameters now shows that 23 out of 25 of the signs put into projection (ii) are correct, and 84% of the signs for all the $(h0\bar{h}l)$ reflections are correctly determined by the S contributions.

Once again, in spite of the improvement in the appearance of the S-phased projections, it was found that several structures could be fitted to them, having R values between 35 and 45%. R had dropped to 30% for one of the trial structures after four rounds of refinement but progress was slow and erratic and the essential correctness of the structure was still in doubt.

At this point, the use of an electronic computer became available, and a three-dimensional Fourier synthesis was calculated, using all the observed reflections phased according to the S positions alone (see Fig. 1, Oughton & Harrison, 1957). This gave a completely unambiguous picture of the whole struc-

ture. The atoms appeared with the following peak heights:

S	32 e.Å ⁻³	C ₁	3 e.Å ⁻³	O ₁	7½ e.Å ⁻³
N	4½	C ₂	5½	O ₂	5
		C ₃	4		

and there were no spurious peaks above the 2 e.Å⁻³ level. The N position alone was rather badly defined, and co-ordinates were selected on the elongated peak, which placed the N atom approximately in the C₂C₃O₁O₂ plane.

Using the atomic parameters so derived, F_c values for all reflections with $F_0 \neq 0$ were calculated, and gave $R = 21\%$. It is interesting to note that this structure is essentially the same as the one derived from projection, the average shift in parameters from those of the structure giving $R = 30\%$ being about 0.5 Å. It is clear, however, that three-dimensional methods lead more directly and unambiguously to the correct answer. A further round of refinement in three dimensions brought R down to 15.6%.

A three-dimensional $(F_0 - F_c)$ synthesis was then calculated. It indicated minor shifts in the parameters of all the atoms and adjustments in the temperature factors of some of them. The major temperature-factor errors were in the S atom which showed considerable anisotropy of motion, and in the oxygen atom, O₂, which has a higher thermal motion than the other light atoms. In addition, peaks of approximately 0.7 e.Å⁻³ in height were found in positions which could be associated with the six hydrogen atoms.

Table 2. Final positional and thermal parameters

Atomic positions						
(x, y, z) as defined in <i>Int. Tab.</i> for space group $P6_122$						
	x	$\sigma(x)$	y	$\sigma(y)$	z	$\sigma(z)$
S	0.19866	0.000506	0.03138	0.000550	0.41278	0.000042
C ₁	0.0886	0.00237	0.7310	0.00237	0.39270	0.000171
C ₂	0.0515	0.00208	0.7836	0.00208	0.36697	0.000166
C ₃	0.7967	0.00195	0.8301	0.00195	0.36231	0.000148
N ₁	0.3165	0.00175	0.0337	0.00175	0.35686	0.000175
O ₁	0.8493	0.00150	0.0732	0.00150	0.35613	0.000148
O ₂	0.5586	0.00171	0.6200	0.00171	0.36578	0.000199
H ₁	0.223		0.675		0.3923	
H ₂	0.926		0.562		0.3961	
H ₃	0.010		0.620		0.3563	
H ₄	0.493		0.015		0.3538	
H ₅	0.383		0.215		0.3667	
H ₆	0.283		0.985		0.3390	

Thermal parameters

For S

$$T = \exp - (0.04139h^2 + 0.04555k^2 + 0.000166l^2 + 0.04555hk) \quad (\varepsilon = \eta = 0)$$

For the light atoms $T = \exp - [B (\sin^2 \theta / \lambda^2)]$

Atom	B
C ₁	3.2 ₃
C ₂	3.2 ₀
C ₃	3.2 ₀
N ₁	3.0 ₆
O ₁	3.3.
O ₂	3.8 ₁

The constants $\alpha-\eta$ in the expression for an anisotropic temperature factor $-\ln T = \alpha h^2 + \beta k^2 + \gamma l^2 + \delta hk + \epsilon kl + \eta lh$ were estimated from the values of $\delta^2 D/\delta x^2, \dots$ (Cochran, 1951; Dunitz & Rollett, 1956) for the S atom, and individual isotropic temperature factors were estimated for the light atoms (Table 2).

In the third cycle, the parameter shifts, temperature factor modifications and hydrogen atom contributions were included (Table 1). The R value dropped to 12.3%. It was decided that, on consideration of

the complexity of the space group and the accuracy of the data, further refinement would not be profitable. The results of the last three-dimensional Fourier synthesis is shown in Fig. 5. The final parameters are given in Table 2.

The standard deviations of the S, C, N and O co-ordinates were determined from Cruickshank's (1949) equation:

$$\sigma(x_n) = 2\pi \{ \sum h^2 (F_0 - F_c)^2 \}^{1/2} \left(a V \left(\frac{\partial^2 \rho_c}{\partial x^2} \right) \right).$$

The positions of the hydrogen atoms and the thermal vibration parameters were determined from the one three-dimensional difference synthesis without further refinement, and therefore are not of the same order of accuracy as the positional parameters.

Discussion of the structure

A description of the conformation of the molecule, its packing in the crystal lattice, and its relationship to the conformation of the $[-S.-CH_2.CH.NH.CO_2]_2$ group in L-L'-diglycyl-L-cystine dihydrate (Yakel & Hughes, 1954) and to cystine hydrobromide (Corsmit, Schuyff & Feil, 1956) has already been given (Oughton & Harrison, 1957). As stated there, the structure is essentially composed of glycine-like sheets lying perpendicular to the c -axis, with the C-R bonds pointing alternately up and down in successive sheets which are linked together in pairs by disulphide bridges.

The dimensions of the disulphide bridge (Table 3) are reasonable, although the S-C₁-C₂ angle is slightly large. The S-C₁ bond, 1.82 Å, is not unusually long, as it is in L-L'-diglycyl-L-cystine dihydrate. The C₁-C₂ bond makes an angle of 106° 10' with the plane of the (CN-COO) group, and the separation produced by the bridge between the glycine-like sheets is about 6.3 Å. The C₂N₁.C₃O₁O₂ group is nearly planar, with the C₂N₁ bond 4° out of the plane of the C₂, C₃, O₁, O₂ atoms.

The molecules are very satisfactorily hydrogen-

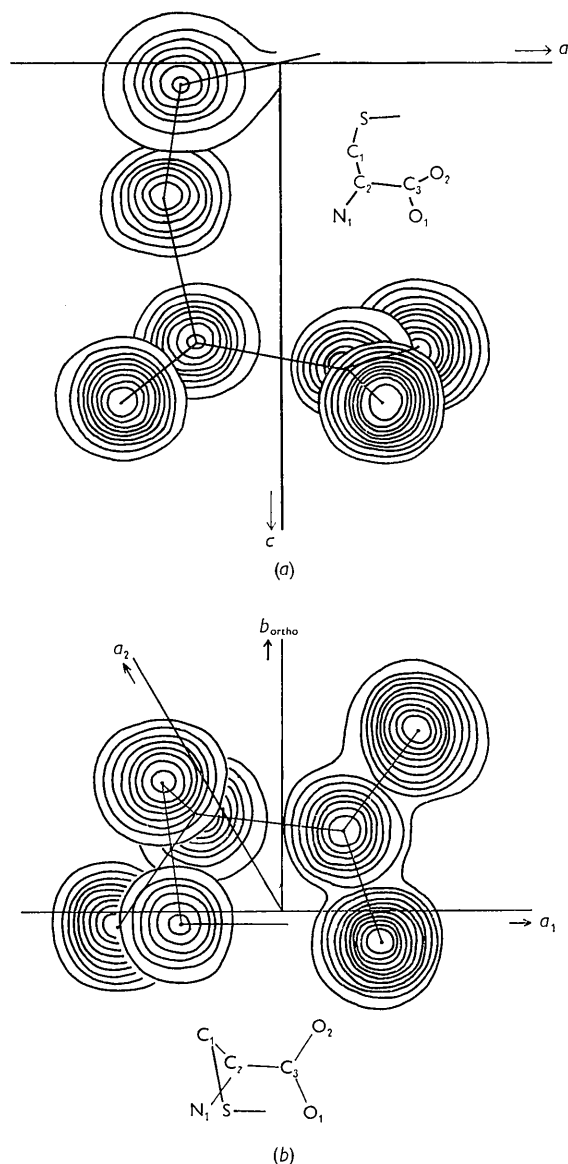


Fig. 5. The conformation of half the cystine molecule viewed (a) down the two-fold axis parallel to [1120]; (b) down [0001]. The contours are those obtained from the final three-dimensional Fourier synthesis. Contours begin at the 1 e.Å⁻³ level and are at 1 e.Å⁻³ intervals except for the S atom, which is contoured from the 5 e.Å⁻³ level in 5 e.Å⁻³ intervals.

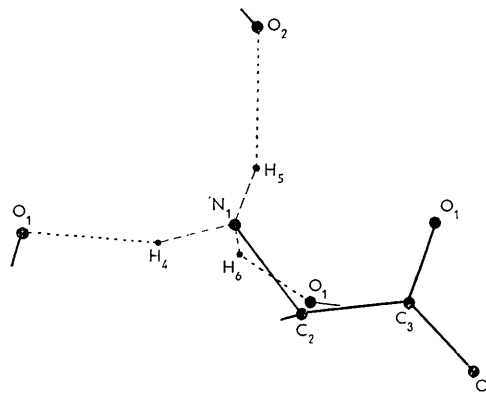


Fig. 6. Diagram showing the environment of the NH₃⁺ group, viewed down the c -axis.

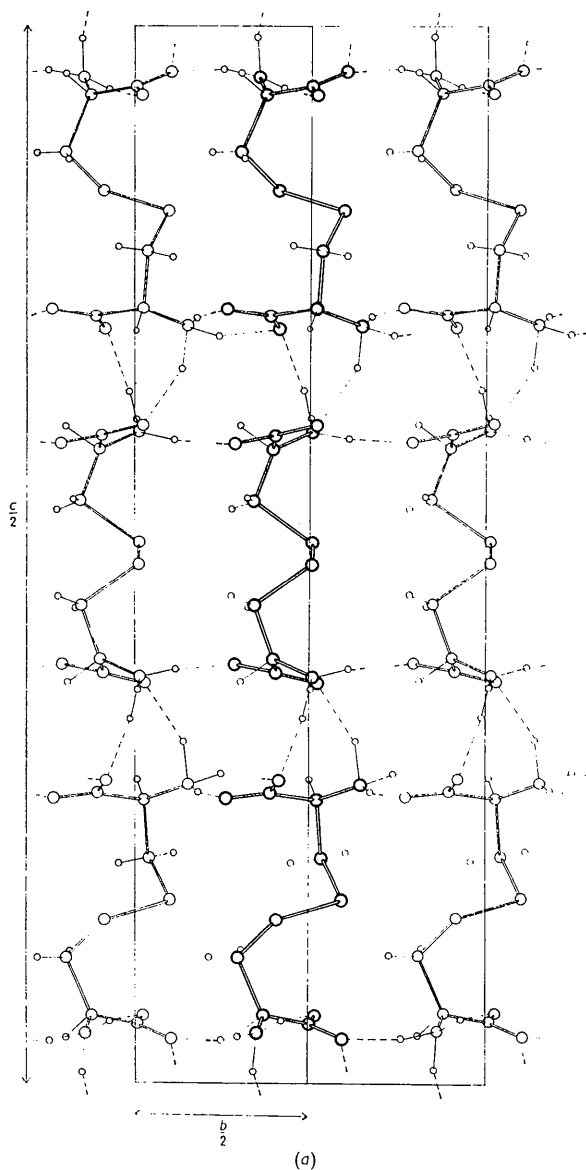


Fig. 7. The structure viewed down (a) $[10\bar{1}0]$ and (b) $[0001]$ showing the packing of the molecule and the disposition of the hydrogen bonds.

Table 3. Bond lengths and bond angles in L-cystine

Bond lengths	
S-S'	2.032 ± 0.003 Å
S-C ₁	1.820 ± 0.012
C ₁ -C ₂	1.509 ± 0.016
C ₂ -C ₃	1.543 ± 0.014
C ₂ -N ₁	1.511 ± 0.014
C ₃ -O ₁	1.250 ± 0.013
C ₃ -O ₂	1.238 ± 0.014
Bond angles	
S'-S-C ₁	$114^\circ 30' \pm 24'$
S-C ₁ -C ₂	$116^\circ 12' \pm 53'$
C ₁ -C ₂ -C ₃	$114^\circ 13' \pm 52'$
C ₁ -C ₂ -N ₁	$111^\circ 54' \pm 53'$
N ₁ -C ₂ -C ₃	$108^\circ 32' \pm 47'$
C ₂ -C ₃ -O ₁	$117^\circ 54' \pm 53'$
C ₂ -C ₃ -O ₂	$115^\circ 20' \pm 55'$
O ₁ -C ₃ -O ₂	$126^\circ 48' \pm 1^\circ$
Hydrogen bonds	
N ₁ ...O ₁	2.789 ± 0.013 Å
N ₁ ...O ₂	2.809 ± 0.014
N ₁ ...O ₁ '	2.865 ± 0.014
N ₁ -H ₄	1.0 ₄ Å
N ₁ -H ₅	1.0 ₂
N ₁ -H ₆	1.0 ₇
C ₂ -N ₁ -H ₄	119°
C ₂ -N ₁ -H ₅	112°
C ₂ -N ₁ -H ₆	100°
O ₁ -N ₁ -H ₄	16°
O ₂ -N ₁ -H ₅	21°
O ₁ -N ₁ -H ₆	21°
Dihedral angle	
C ₁ '-S'-S-C ₁	$106^\circ \pm 1^\circ$

bonded together (Figs. 6, 7), with three hydrogen atoms arranged approximately tetrahedrally around each N atom, and lying close to the $N_1 \cdots O_1$, $N_1 \cdots O_2$ and $N_1 \cdots O_1'$ directions. The nearly equal lengths of the C₃-O₁ and C₃-O₂ bonds, and the disposition of the H atoms about the N atom clearly indicate a zwitterion form.

An interesting feature of the structure is the close van der Waals contacts between each sulphur atom and its non-bonded neighbouring atoms (Table 4). These are all shorter than the sum of the van der Waals radii of the atoms as given by Pauling (1940),

Table 4. Close van der Waals contacts made by the sulphur atoms

	Observed	Calculated from Pauling's radii	Calculated from Donohue's radius for S
S...N (intra-molecular)	3.21 Å	3.35 Å	3.15 Å
S...C ₁	3.67	3.85	3.65
S...S'	3.47	3.70	3.30

and are in agreement with Donohue's (1950) suggestion that the true van der Waals radius of sulphur should be approximately 1.65 Å, i.e. 0.2 Å less than Pauling's value.

APPENDIX

Crystallographic data for the space group *P6₁22*

At many points in the structure analysis it was found simpler to use the related, face-centred orthorhombic unit cell, instead of the hexagonal cell. The relationships used are given below.

Data for the related, face-centred orthorhombic unit cell

$$a_{\text{ortho.}} = a_{1\text{ hex.}}, \quad b_{\text{ortho.}} = (a_1 + 2a_2)_{\text{hex.}}, \quad c_{\text{ortho.}} = c_{\text{hex.}}$$

If these orthogonal axes are used, then the structure can be described in terms of the eight equivalent points of the space group *C222₁*. It must be emphasized that this is *not* the true space group of the structure which is *P6₁22*. In terms of the space group *C222₁*, the asymmetric unit is three times the size of the true asymmetric unit, because the 6₁ axis is now effectively being treated as a 2₁ axis, and the three equivalent points $(x, y, z)_{\text{hex.}}$, $(x, x-y, \frac{1}{6}-z)_{\text{hex.}}$ and $(x-y, x, \frac{1}{6}+z)_{\text{hex.}}$ are all represented by $(x, y, z)_{\text{ortho.}}$. Let

$$\begin{aligned} (x, y, z)_{\text{hex.}} &\equiv (x_1, y_1, z_1)_{\text{ortho.}} \\ (x, x-y, \frac{1}{6}-z)_{\text{hex.}} &\equiv (x_2, y_2, z_2)_{\text{ortho.}} \\ (x-y, x, \frac{1}{6}+z)_{\text{hex.}} &\equiv (x_3, y_3, z_3)_{\text{ortho.}} \end{aligned}$$

Then the orthorhombic co-ordinates are related to one another as follows:

$$\begin{aligned} x_1 + x_2 &= x_3 \\ y_2 + y_3 &= y_1 \\ z_2 &= \frac{1}{6} - z_1 \\ z_3 &= \frac{1}{6} + z_1 \end{aligned}$$

and

$$\begin{aligned} x_2 &= \frac{1}{2}x_1 + \frac{3}{2}y_1, \quad y_2 = \frac{1}{2}x_1 - \frac{1}{2}y_1 \\ x_3 - \frac{1}{2}x_1 - \frac{3}{2}y_1, \quad y_3 &= \frac{1}{2}x_1 + \frac{1}{2}y_1. \end{aligned}$$

Thus (x_1, y_1, z_1) can be determined from the (x, z) or (y, z) projection alone.

Co-ordinates of vectors between equivalent points in terms of the orthorhombic unit cell

$$\begin{aligned} x = 0 \text{ section } (x_1 = 0)_{\text{hex.}} \\ 0, \pm 2y_1, \pm 2z_1 \\ 0, \pm (x_1 - y_1), \pm (\frac{1}{3} - 2z_1) \\ 0, \pm (x_1 + y_1), \pm (\frac{1}{3} + 2z_1). \end{aligned}$$

$$\begin{aligned} y = 0 \text{ section } (x_1 = -x_2)_{\text{hex.}} \\ \pm 2x_1, 0, \pm (\frac{1}{3} - 2z_1) \\ \pm (x_1 + 3y_1), 0, \pm (\frac{1}{6} + 2z_1) \\ \pm (x_1 - 3y_1), 0, \pm (\frac{1}{6} - 2z_1). \end{aligned}$$

The orthorhombic cell was used in finding the *S* parameters from the three-dimensional Patterson; for examining the Fourier projections and in calculating all the three-dimensional Fourier syntheses. The hexagonal cell was used to calculate $(hki\bar{l})$ structure factors.

In the space group *P6₁22*,

$$I_{hk\bar{l}l} = I_{hk\bar{l}\bar{l}} = I_{k\bar{l}hl} = I_{i\bar{l}kl} = I_{h\bar{l}kl} = I_{\bar{l}h\bar{k}l}$$

and the set of independent reflections is only $\frac{1}{24}$ of the whole reciprocal lattice. Symmetry-related reflections have phase angles which differ from one another by $n\pi/3$ for hexagonal, rhombohedral and trigonal space groups, instead of by $n\pi/2$ as for all other space groups. Thus the *A* and *B* values for the whole reciprocal lattice cannot be derived from one unique set simply by taking $A' = \pm A$ or $\pm B$; $B' = \pm B$ or $\pm A$ for symmetry-related reflections. Instead they take the form

$$\begin{aligned} A' &= \pm \frac{1}{2}A \pm \frac{1}{2}\sqrt{3}B \\ B' &= \pm \frac{1}{2}B \pm \frac{1}{2}\sqrt{3}A. \end{aligned}$$

In *International Tables* (p. 479), only the relationships between the phase angles $\alpha_{hki\bar{l}}$, $\alpha_{\bar{l}h\bar{k}l}$ and $\alpha_{h\bar{k}i\bar{l}}$ are given, leaving three others to be determined.

They are:

$$\begin{aligned} \alpha_{k\bar{l}hl} &= \alpha_{hki\bar{l}} + \frac{1}{3}(4\pi l) \\ \alpha_{i\bar{l}kl} &= \alpha_{hki\bar{l}} + \frac{1}{3}(2\pi l) \\ \alpha_{h\bar{k}i\bar{l}} &= -\alpha_{hki\bar{l}} + \pi l. \end{aligned}$$

From these six equations α , and therefore *A* and *B*, for all reflections can be derived from one unique set.

If the equation for $\rho_{(r,y,z)}$, given in *International Tables*, p. 474 is used, the summation must be made over $\frac{1}{4}$ of the reciprocal lattice, i.e. for $h = -\infty$ to $+\infty$, $k = 0$ to $+\infty$, $l = 0$ to ∞ . If, however, one transforms to orthogonal axes and uses the face-centred orthorhombic unit cell, having the space-group *C222₁*, the summation is over only $\frac{1}{8}$ of the reciprocal lattice (*International Tables*, p. 386), thus reducing the size of the summation by a factor of two.

The electron density need be calculated only for a volume $a_{\text{ortho.}} \times \frac{1}{2}b_{\text{ortho.}} \times \frac{1}{12}c_{\text{ortho.}}$, which contains two asymmetric units of the hexagonal cell.

By transforming to the related orthorhombic cell, simplified expressions for the electron density involving fewer terms in the summation may be derived for all hexagonal space-groups containing 2-fold axes or planes of symmetry, i.e. *P6₁22*, *P6₂22*, *P6₃22*, *P6₄22*, *P6₅22*, *P6mm*, *P6cc*, *P6₃cm*, *P6₃mc*, *P6m2*, *P62m*, *P62c*, *P6/mmm*, *P6₃mcm* and *P6₃/mmc*.

We wish to thank Mrs D. Hodgkin for suggesting this problem and for frequent encouragement, and also Prof. G. M. J. Schmidt who did some of the preliminary work.

We are also much indebted to Dr J. S. Rollett and to members of the staff of the National Physical

Laboratory who computed the three-dimensional Fourier syntheses and structure factors on 'Deuce', the N.P.L. electronic computer.

References

- BUERGER, M. J. (1946). *J. Appl. Phys.* **17**, 579.
 COCHRAN, W. (1951). *Acta Cryst.* **4**, 408.
 COCHRAN, W. (1952). *Acta Cryst.* **5**, 65.
 CORSMIT, A. F., SCHUYFF, A. & FEIL, D. (1956). *Proc. K. Ned. Akad. Wet. B*, **59**, 470.
 CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 65.
 DONOHUE, J. D. (1950). *J. Amer. Chem. Soc.* **72**, 2701.
 DUNITZ, J. D. & ROLLETT, J. S. (1956). *Acta Cryst.* **9**, 327.

- HARKER, D. (1948). *Amer. Min.* **33**, 764.
 HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70.
 OUGHTON, B. M. & HARRISON, P. M. (1957). *Acta Cryst.* **10**, 479.
 PAULING, L. (1940). *Nature of the Chemical Bond*, p. 176. Ithaca: Cornell University Press.
 SAYRE, D. M. (1952). *Acta Cryst.* **5**, 60.
 SCHMIDT, G. M. J., HODGKIN, D. C. & OUGHTON, B. M. (1957). *Biochem. J.* **65**, 744.
 STEINRAUF, L. K. & JENSEN, L. H. (1956). *Acta Cryst.* **9**, 539.
 WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 152.
 YAKEL, H. L., JR. & HUGHES, E. W. (1954). *Acta Cryst.* **7**, 291.
 ZACHARIASEN, W. H. (1952). *Acta Cryst.* **5**, 68.

Acta Cryst. (1959). **12**, 404

A Unified Program for Phase Determination, Type 1P

BY J. KARLE AND H. HAUPTMAN

U.S. Naval Research Laboratory, Washington 25, D.C., U.S.A.

(Received 7 September 1958)

The new probability approach, in which the crystal structure is fixed and the Miller indices range uniformly but not independently over the integers, has been developed to yield formulas with universal application to all the space groups. The formulation is general and includes with equivalent rigor both the unequal and equal atom cases. If a requirement of rational independence of atomic coordinates is fulfilled, the formulas have exact validity. On the other hand, the formulas should still be useful even if the condition of rational independence is only partially fulfilled. In the application to a particular space group, use is made of the space group symmetries to obtain relationships among the structure factors, by means of which the general formula is suitably specialized.

The present paper is concerned with the space groups comprising type 1P. A detailed procedure for phase determination in this type is described.

1. Introduction

The concept of the joint probability distribution of several structure factors, obtained by treating the atomic coordinates as random variables, was introduced in our Monograph I (Hauptman & Karle, 1953) for the purpose of determining phase directly from the observed X-ray intensities. Independently, Kitaigorodsky (1954), derived the joint probability distribution for three structure factors which constituted a study of the Sayre relation (1952) from the probability point of view. The results obtained from these joint probability distributions were confirmed by Bertaut (1955) and Klug (1958) who employed alternative manipulative techniques which were however mathematically equivalent to the formulation in Monograph I (1953).* It is to be emphasized that all these distribu-

tions were obtained on the basis that \mathbf{h} is fixed and the atomic coordinates range uniformly and independently throughout the asymmetric unit. This is in marked contrast to our more recent work (1958) which treats the crystal structure as fixed and permits the indices to range uniformly, but not independently, throughout reciprocal space. In this way, considerably improved formulas have been obtained.

variances of individual contributors to phase determining formulas. However in a previous paper by us (1956), of which Klug is apparently unaware, measures of statistical significance based on the appropriate variances had already been derived and their application was discussed.

The method of Monograph I utilizes the unusually large E values and their interactions in several phase determining formulas to build an internally consistent set of signs. Thus it should be clear that more is required in a theoretical evaluation than a knowledge of the variance of single terms in individual formulas and the number and quality of the data. Quite fortunately, phase determining theories can be readily confronted by experiment and it is our opinion that the final judgment regarding the applicability of a method is most propitiously made in the laboratory.

* In his paper, Klug (1958) summarizes the theory and attempts to evaluate the status of probability methods based on atomic coordinates as random variables. Klug reiterates limitations stated by others in the past and offers, as new evidence, measures of statistical significance derived from the