

- COCHRAN, W. (1954). *Acta Cryst.* **7**, 503.
 COCHRAN, W. (1956). *Acta Cryst.* **9**, 924.
 COX, E. G., CRUICKSHANK, D. W. J. & SMITH, J. A. S. (1955). *Nature, Lond.* **175**, 766.
 CRUICKSHANK, D. W. J. (1949). *Acta Cryst.* **2**, 154.
 GROTH, P. (1906-19). *Chemische Krystallographie*, vol. 3, p. 318. Leipzig: Engelmann.
 HOERNI, J. A. & IBERS, J. A. (1954). *Acta Cryst.* **7**, 744.
 JEFFREY, G. A. & CRUICKSHANK, D. W. J. (1953). *Quart. Rev. Chem. Soc., Lond.* **7**, 335.
 LANGE, J. J. DE, ROBERTSON, J. M. & WOODWARD, I. (1939). *Proc. Roy. Soc. A*, **171**, 398.
 McDONALD, T. R. R. (1956). *Acta Cryst.* **9**, 162.
 MEGAW, H. D. (1954). *Acta Cryst.* **7**, 771.
 PAULING, L. (1940). *The Nature of the Chemical Bond*, p. 175. Ithaca: Cornell University Press.
 PEERDEMAN, A. F. & BIJVOET, J. M. (1956). *Acta Cryst.* **9**, 1012.
 PEERDEMAN, A. F., BOMMEL, A. J. VAN & BIJVOET, J. M. (1951). *Proc. K. Ned. Akad. Wet. B*, **54**, 16.
 ROSS, P. A. (1928). *J. Opt. Soc. Am.* **16**, 433.
 SPRENKELS, A. J. J. (1956). Dissertation, Utrecht.
 STERN, F. & BEEVERS, C. A. (1950). *Acta Cryst.* **3**, 341.
 WHITTAKER, E. J. W. (1953). *Acta Cryst.* **6**, 222.

Acta Cryst. (1958). **11**, 70

The Determination of the Coordinates of Heavy Atoms in Protein Crystals

By W. L. BRAGG

The Royal Institution, 21 Albemarle Street, London W. 1, England

(Received 25 September 1957)

The paper deals with the determination of the coordinates of heavy atoms added to protein crystals. Previous methods have in general been based on Patterson or Fourier series. In the method described here, the changes in amplitude of the individual diffracted beams due to the addition of the heavy atom are plotted on a graph, from which the required coordinate can be deduced by inspection. Examples of the determination of coordinates in actual cases are given.

1. Introduction

The method of determining the phases of $F(hkl)$ values in protein crystals by the addition of heavy atoms has been established by Perutz (Green, Ingram & Perutz, 1954). He was the first to show that certain chemical groups containing a heavy atom such as mercury can be added to the molecules without appreciable alteration of crystal form, and that they produce measurable differences in the values of $I(hkl)$.

In a centrosymmetric projection, where all structure factors are real (plus or minus), the values of $\Delta F = (|F(\text{protein} + \text{heavy atom})| - |F(\text{protein})|)$ can be used to form a 'Difference Patterson'. This is so because $(\Delta F)^2 = (F(H))^2$, $F(H)$ being the structure factor of the heavy atom alone. The peaks in the Patterson indicate the approximate coordinates of the heavy atom, and from these the signs of most of the ΔF values can be ascertained. A 'Difference Fourier' can then be formed with ΔF values which gives more accurate values of the coordinates and clears up doubtful signs. A knowledge of the signs, and of the effect of the heavy atom in increasing or decreasing the intensity of each diffracted beam, enables the signs of $|F(P)|$ to be determined and a Fourier of the protein structure to be formed.

The methods outlined above can be applied only to centrosymmetrical structures. Perutz (1956) has described Fourier-series methods of dealing with the more complex cases without symmetry centres. The

present note describes simple alternative ways of measuring coordinates of added heavy atoms which do not involve the use of Patterson or Fourier series and can be applied to non-centrosymmetric cases. Estimates of the coordinates so obtained are compared with those deduced by other methods. It is important to know these coordinates as accurately as possible when they are used to determine the phases of $F(\text{protein})$ (see, for instance, Harker (1956) and the methods described here have certain advantages as regards accuracy.

2. The determination of x and z coordinates in a monoclinic crystal

I am indebted to Dr J. C. Kendrew for letting me use some unpublished F measurements for Sperm whale myoglobin Type A, determined in the course of an investigation directed towards a three-dimensional Fourier of the structure. They are measurements of $|F(hk0)|$ and $|F(0kl)|$ for

myoglobin,
 myoglobin + parachloromercuri-benzenesulphonic acid, PCMBS,
 myoglobin + mercury diammine, HgAm_2 ,
 myoglobin + gold chloride, Au.

The space group and cell constants are

$$P2_1, Z = 2, \\ a = 64.6, b = 31.1, c = 34.8 \text{ \AA}, \beta = 105.5^\circ, \\ \text{with atoms at } x, y, z \text{ and } -x, y - \frac{1}{2}, -z.$$

The x and z coordinates of the heavy atoms will be determined for the non-centrosymmetric (hkl) and ($0kl$) projections as an example; they can be compared with the coordinates obtained from the centrosymmetric ($h0l$) projection by Kendrew using Fourier methods. In considering the interpretation of the ΔF values it must be remembered that their accuracy is low; the average value of ΔF is about 100 and the random error ± 30 . It is indeed a feat of X-ray technique to get significant values of ΔF when one atom is added to a molecule, the molecular weight of which is 17,000. We have to deduce the coordinates as best we can from these necessarily imperfect data.

Fig. 1(a) shows the values of $|\Delta F(0kl)|$ for PCMBS

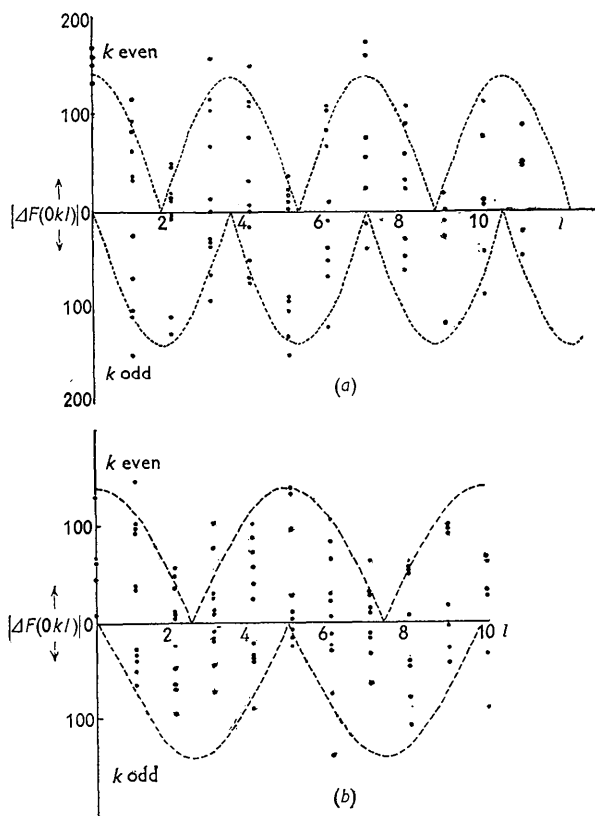


Fig. 1. (a) $|\Delta F(0kl)|$ values for PCMBS, plotted against l .
(b) $|\Delta F(0kl)|$ values for HgAm₂, plotted against l .

plotted against l , all even k values being plotted above the axis and all odd k values below the axis. When k is even, the amplitude of $F(H)$ is $2f_1 \cos 2\pi lz_1$, when k is odd it is $2f_1 \sin 2\pi lz_1$. The $|\Delta F|$ values will in general be less than $|F(H)|$ because of the phase angle between $F(P)$ and $F(H)$, but in a reasonable proportion of the cases, they may be expected to approach the maximum (for instance in 30% of the cases on the average $|\Delta F|$ will be within 10% of $|F(H)|$). We can

therefore adjust an envelope to enclose all the points, which has a $|\cos \theta|$ form above the axis and a $|\sin \theta|$ form below the axis, with due allowance for the random errors. Such an envelope is drawn in Fig. 1(a). It shows that $F(\text{max.})$ is about 140 units, indicating the amount of heavy atom in the unit cell. The positions of the maxima and minima show that z_1 must be very closely $1/7$, or 0.143 , since the $|\cos \theta|$ curve has maxima at $l = 0, 3.5, 7, 10.5$ and the $|\sin \theta|$ curve has minima at these points. The envelope is definitely a poorer fit to the points if drawn for $z_1 = 1/6.75 = 0.148$ or $1/7.25 = 0.138$. It appears reasonable to estimate that

$$z_1 = 0.143 \pm 0.003.$$

Fig. 1(b) shows a similar procedure applied to the ΔF values for HgAm₂. The figures are not so consistent as in the case of PCMBS. There are probably systematic errors due to a lack of complete isomorphism between the two crystals, or to the mercury not being wholly in one site. The maxima and minima must be close to $l = 2.5, 5, 7.5, 10$, possibly at slightly lower values, and it is estimated that

$$z_2 = 0.102 \pm 0.005.$$

Each set of curves gives the distance of the mercury atom from the nearest twofold axis, but it remains to be decided whether taking z_1 to be 0.143 , we should take z_2 to be 0.102 or $0.5 + 0.102$. A decision can be made by considering the values of $F(00l)$ when l is odd, and noting in each case whether they are altered by PCMS and HgAm₂ in the same or in opposite sense. It is clear that the latter value for z_2 is the correct one.

The values are compared below with those deduced by Kendrew from the centrosymmetric projection:

$$z_1 = 0.143 \pm 0.003, \quad z_1 = 0.144 \text{ (Kendrew)}, \\ z_2 = 0.602 \pm 0.005, \quad z_2 = 0.598 \text{ (Kendrew)}.$$

3. Plot of $\Delta F(00l)$ values

It is possible, of course, to plot only the $\Delta F(00l)$ values against l . In this case $|\Delta F| = |F(H)|$ and the points should all be on a $|\cos \theta|$ curve. It is of advantage, however, to plot all the ΔF values because the

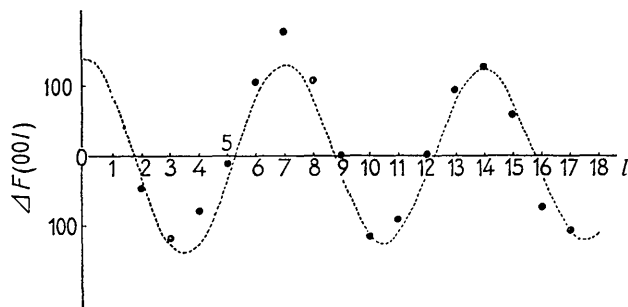


Fig. 2. $|\Delta F(00l)|$ values for PCMBS, plotted against l . The ΔF values have been plotted with the appropriate plus and minus signs.

random errors then have less effect on the accuracy of the estimate.

Fig. 2. shows an example of plotting $F(00l)$ values for PCMBS. This is a favourable case, because the crystals appear to be closely isomorphous and the mercury to be attached to a single site. I am indebted to Dr Kendrew for extending the measurements out to the 17th order, so as to test how accurately the z coordinate could be deduced. Fig. 2 illustrates an advantage of the present method. It would be a formidable task to extend a survey of $(0kl)$ spectra to such a high order, but by concentrating on the $(00l)$ series one gets the accuracy corresponding to a high resolving power while only a limited number of spectra need be measured.

4. Determination of the x coordinates

Fig. 3 illustrates the application of the envelope method to the $(hk0)$ projection, for PCMBS and HgAm₂. The envelope for PCMBS has been drawn in a somewhat different way. It is clear by inspection of the ΔF values that x_1 for PCMBS is nearly but not quite 0.25. A value of exactly 0.25 would give maxima for h even and zero for h odd when k is even, and vice versa when k is odd. When h is small, the observed values follow this rule, but as h increases the rule is reversed. An envelope with minima at alternate orders is difficult to outline, and it is more convenient to put $x_1 = 0.25 - x'_1$ and to determine x'_1 .

The structure factors become

$$\begin{aligned} k \text{ even: } & 2f_1 \cos h(\frac{1}{2}\pi - \theta), \\ k \text{ odd: } & 2f_1 \sin h(\frac{1}{2}\pi - \theta), \end{aligned}$$

where $\theta = 2\pi x'_1$. We then have

$$\begin{aligned} |\Delta F| &\leq |2f_1 \cos 2\pi h x'_1|, & h+k \text{ even,} \\ |\Delta F| &\leq |2f_1 \sin 2\pi h x'_1|, & h+k \text{ odd,} \end{aligned}$$

and we plot $(h+k)$ even above the axis, $(h+k)$ odd below the axis, as drawn in Fig. 3(a). It will be seen that $2\pi h x'_1 = \pi$ when $h \approx 18$, so $x'_1 \approx 1/36 = 0.028$. It is estimated that

$$x_1 = 0.25 - x'_1 = 0.222 \pm 0.002.$$

The results for HgAm₂ are again less consistent, as in the $(0kl)$ series. The envelope at lower values of l indicates $x_2 = 1/13 = 0.077$. The higher orders are more consistent with $x_2 = 1/14 = 0.071$. It is clear however that up to the 10th order it would be safe in calculating phase to assume the former value with $f(\text{max.}) = 140$.

The ΔF values for Au are erratic, as is to be expected since their average value is only about 50 as compared with a random error of 30. The values for $\Delta F(h00)$ are plotted in Fig. 3(c) with signs derived by comparison with the ΔF values for PCMBS and HgAm₂. They determine x_3 as approximately 0.10. When ΔF values are plotted for all $(hk0)$, as in Fig.

3(a) and (b), however, there are a number of anomalies at higher orders which suggest that Au is not attached wholly to one site.

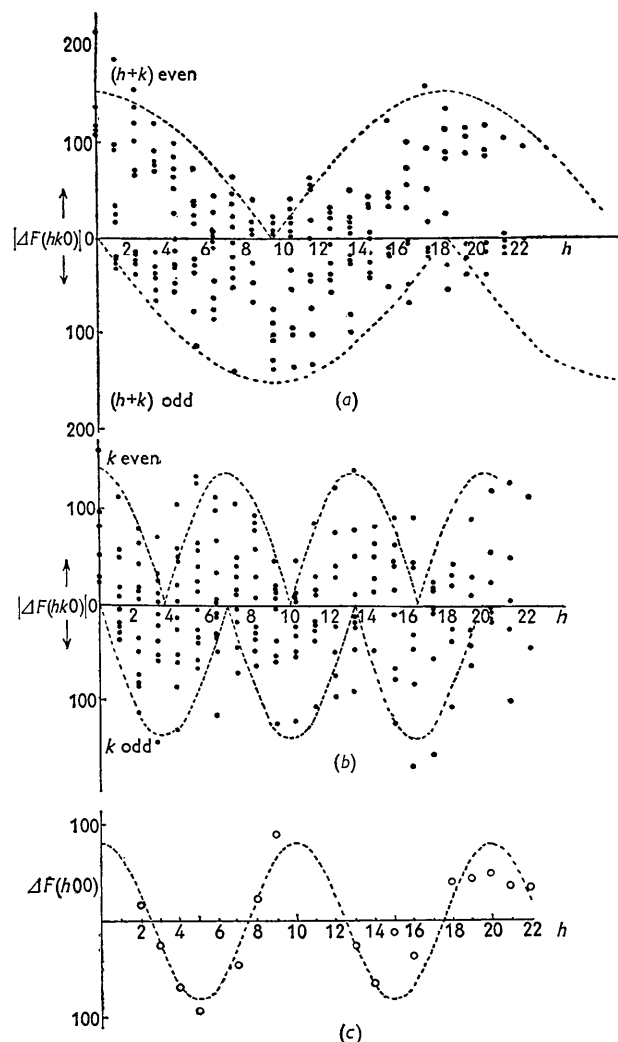


Fig. 3. (a) $|\Delta F(hk0)|$ values for PCMBS, plotted against h . (b) $|\Delta F(hk0)|$ values for HgAm₂, plotted against h . (c) $\Delta F(h00)$ values for Au, plotted against h . In the cases of $h = 6, 10, 11, 12, 17$ the $F(h00)$ spectra are so weak that reliable estimates of ΔF cannot be made.

Adopting the origin chosen by Kendrew, the coordinates compare as follows:

$$\begin{aligned} \text{PCMBS: } & x_1 = 0.278 \pm 0.002, & x_1 &= 0.275 \quad (\text{Kendrew}), \\ \text{HgAm}_2: & x_2 = 0.423 \pm 0.005, & x_2 &= 0.426 \quad (\text{Kendrew}), \\ \text{Au: } & x_3 = 0.400 \pm 0.005, & x_3 &= 0.4004 \quad (\text{Kendrew}). \end{aligned}$$

The agreement is good, the more so when it is considered that the present coordinates, which were estimated independently, were determined for the non-centrosymmetric projections whereas Kendrew's coordinates were based on the centrosymmetric $(h0l)$ projection.

5. Determination of y coordinates in ox haemoglobin

I am indebted to Dr D. W. Green and Dr A. C. T. North for letting me use the figures on which the next example is based. They are $F(0k0)$ values for the orthorhombic crystal of ox haemoglobin ($P2_12_12_1$, $Z = 4$, $a = 64.7$, $b = 160.0$, $c = 55.2$ Å). Two mercury atoms can be added to sulphur groups in each molecule.

In the horse haemoglobin crystals studied by Perutz (Green *et al.*, 1954) the molecules lie on twofold axes, and the added mercury atoms are symmetrically placed with respect to this axis. In ox haemoglobin the molecules are in general positions, but they appear to have, or to approximate to, the same twofold symmetry with their axes slightly inclined to the b axis, so that the pairs of heavy atoms lie as in Fig. 4, where

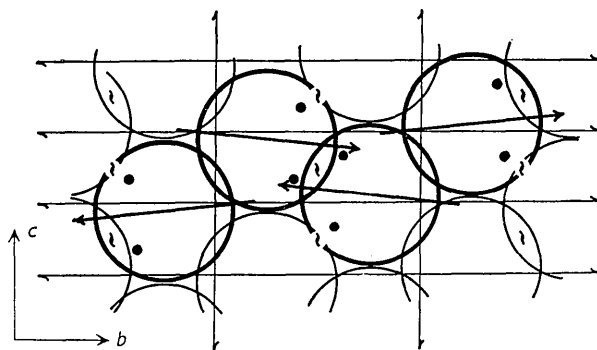


Fig. 4. Diagram showing position of molecules in the 'a' projection of ox haemoglobin (Green & North).

the tilt is exaggerated. The y coordinates of a pair are thus only slightly different, and the double peaks which represent the atoms in the Patterson are barely resolvable. It would appear, however, that the difference in y coordinates can be rather accurately measured by the method described in this paper.

The $\Delta F = |F(P+H)| - |F(P)|$ values for the $(0k0)$ spectra are plotted against k in Fig. 5 up to the 54th order. Since initially the sign of ΔF is unknown, they are plotted above and below the axis and are represented by lines which indicate the possible errors of measurement. The first two loops show that the mean coordinate of the pair is about $1/20$, or $y = 0.05$, and that $F(\text{max.})$ is about 600. It is clear, however, that the values of ΔF cannot be fitted to a simple cosine curve, such as is shown as a broken line in the figure, for there is a run of values near the 30th order which are small, whereas the amplitude at higher k values becomes large again. This must be due to the effect of two atoms with slightly different y coordinates producing two cosine curves which 'beat' with each other. The best fit is obtained by assuming that there are Hg atoms at $(0.050 \pm 0.0084)b$, which gives the full line in Fig. 5. The difference in y values is $b/60$, or 2.66 Å.

The Difference Patterson calculated by Green & North gave a mean value $y = 0.048$ without resolving

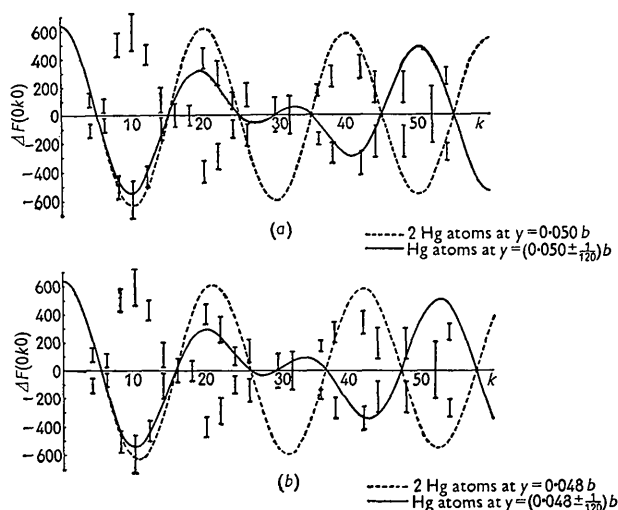


Fig. 5. $\Delta F(0k0)$ values for Hg added to ox haemoglobin, plotted against k (Green & North).

y_1 from y_2 . Fig. 5(b) shows the curve drawn for a mean value of $y = 0.048$; the fit seems clearly poorer, and it may be taken that the mean value of y is 0.050 ± 0.002 . We thus have

$$\begin{aligned} y_1 &= 0.058 \pm 0.002, \\ y_2 &= 0.042 \pm 0.002. \end{aligned}$$

6. The determination of y coordinates in myoglobin

The accurate determination of relative y coordinates is more difficult than that of determining x or z coordinates in the space group $P2_1$ because phase angles must be estimated. The origin lies on a twofold axis, but its position along this axis can be arbitrarily chosen. A method will now be described of measuring the difference between the y coordinates of two heavy atoms by inspection of ΔF values.

Fig. 6 shows a vector diagram, P being the protein vector and S_1 , S_2 the structure-factor vectors of the heavy atoms. To a close approximation, when P is larger than S_1 and S_2 , we may take it that

$$\Delta F_1 \approx S_1 \cos \psi_1, \quad \Delta F_2 \approx S_2 \cos \psi_2,$$

where $\Delta F = |F(\text{protein} + \text{heavy atom})| - |F(\text{protein})|$. The relation becomes inaccurate in the infrequent cases

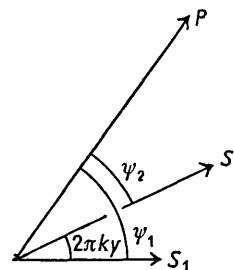


Fig. 6.

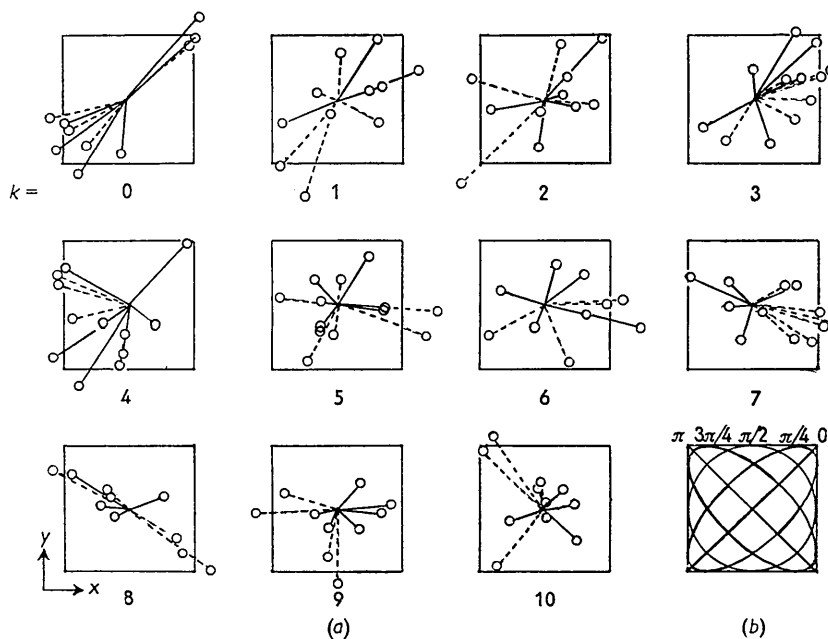


Fig. 7. The determination of the relative displacement of PCMBs and HgAm₂ in the *b* direction.

when *P* is small, but such spectra are hard to measure accurately because of the background intensity, and little is lost by omitting them from consideration.

Fig. 7(a) shows a series of diagrams for successive values of *k*. Each point has as coordinates

$$X = \Delta F_1/S_1 \approx \cos \psi_1, \quad Y = \Delta F_2/S_2 \approx \cos \psi_2,$$

where $S(hk0) = 2f \cos 2\pi hx$, $S(0kl) = 2f \cos 2\pi lz$,
for *k* even;

$S(hk0) = 2f \sin 2\pi hx$, $S(0kl) = 2f \sin 2\pi lz$,
for *k* odd.

If *S* in the denominator is so small as to be comparable with the error of measurement, the corresponding point is omitted because the possible errors in *X* or *Y* are then large.

The phase angle $\psi_1 - \psi_2$ between the *S* vectors is $2\pi ky$, where $y = y_1 - y_2$, the relative displacement of the heavy atoms in the *b* direction. When $2\pi ky$ is zero, *X* and *Y* are equal and have the same sign. When $2\pi ky = \pi$, they are equal but with opposite signs. As $\psi_1 - \psi_2$ increases, the *X*, *Y* points lie first on an ellipse with its major axis at 45° , then (when $\psi_1 - \psi_2 = \pi/2$) on a circle, then on an ellipse with major axis at 135° . They lie on curves like Lissajous' figures, inscribed in a square with a range from +1 to -1 in each direction, as shown in Fig. 7(b).

The 'stars' in Fig. 7(a) represent the plots for $k = 0, 1, 2, \dots, 10$. The full lines represent the (*hk*0) values, the broken lines the (0*kl*) values. The results are irregular, as is to be expected from the large random error. However, a general trend is evident. $2\pi ky$ must be approximately π when $k = 8$. The 'star' for $k = 7$ shows a closer approximation to a line at 135°

than the 'star' for $k = 9$, and it may be estimated that $2\pi y$ lies between $\pi/7.5$ and $\pi/8$. Therefore

$$y = 0.064 \pm 0.003.$$

Kendrew's estimate for the separation, based on Patterson diagrams, is

$$y = 0.058,$$

which would correspond to $2\pi y = \pi/8.5$. There is thus fairly good agreement, but the 'stars' in Fig. 7(a) seem to show rather definitely that the estimate formed from them is a closer approximation.

7. Conclusion

There is a difference in principle between the Fourier methods of determining coordinates, and the methods used here. In a Fourier method all the data are incorporated in one diagram and the position of the atom is denoted by a peak. In the present method, individual spectra are plotted and the coordinate is deduced from the relations between them.

The Fourier method has the advantage that it can deal with a large number of coordinates simultaneously, and as most crystals now being analysed are highly complex, its use is almost universal. In the early days of X-ray analysis, however, it was possible to determine coordinates to a high precision without using Fourier series. This was done by concentrating on spectra of high order, the calculated values of which were very sensitive to a small change in a coordinate, and it was a feasible method because the number of variables in the structure was so small. In the present case, we can use similar methods

because generally we are only trying to find the coordinates of one atom, and they can be determined one at a time. We can concentrate on the spectra which can be measured accurately, and neglect such as are uncertain because a weak intensity is hard to measure against the background. Yet such intensities may correspond to F values which affect a Fourier considerably. Termination of a Fourier series may influence the accuracy of the peak position, but in comparing spectra it is of no importance. A selected set can be measured out to a high order without undue labour, and so a high resolution can be attained, as in the example of an $(0k0)$ series for ox haemoglobin given above.

If the crystals were very closely isomorphous, if the added heavy atom were on one definite site, if there were no error in scaling the measurements of the protein against those of the protein with added heavy atom, if all spectra could be included up to an order where they become negligibly small, and if the random errors were small compared with the difference produced by the addition of the heavy atom, even for the

spectra which are faint against the background, any method of calculation would lead to accurate positions for the heavy atom. Actually, however, we are dealing with necessarily imperfect and incomplete data because of all these sources of error. Evidence must be weighed and judgement exercised in arriving at the best estimates of the coordinates. It is of first importance to fix these coordinates as accurately as possible, because they will eventually be needed to determine the phases of the protein structure factors. It is the purpose of this note to draw attention to the possible advantages of a method of representing the observations, which enables one to weigh the evidence of all the most significant and trustworthy data plotted in each case on the one chart.

References

- GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. Roy. Soc. A*, **225**, 287.
 HARKER, D. (1956). *Acta Cryst.* **9**, 1.
 PERUTZ, M. F. (1956). *Acta Cryst.* **9**, 867.

Acta Cryst. (1958). **11**, 75

The Crystal Structure of Iridium Diselenide

BY LUISA BRAHDE BARRICELLI

Kjemisk Institutt A, Universitetet i Oslo, Blindern, Norway

(Received 31 January 1957 and in revised form 24 June 1957)

The structure of IrSe_2 has been determined, using single-crystal data. The unit cell is orthorhombic with $a = 20.94$, $b = 5.93$, $c = 3.74$ Å. There are eight iridium and sixteen selenium atoms in the unit cell. The space group is $Pnam$. The iridium atom has octahedral selenium environment, the coordination distances are three Ir–Se = 2.44 Å and three Ir–Se = 2.52 Å. Half of the selenium atoms are surrounded by three iridium atoms at three corners of a distorted tetrahedron, and by a selenium atom at a distance Se–Se = 2.57 Å at the fourth corner. The other half of the selenium atoms are coordinated in the same way, but the distance Se–Se = 3.27 Å in this case, which implies that very little bond character is left. An explanation of the divergence in the Se–Se distances and of the difference in the Ir–Se octahedral bond lengths is attempted.

Introduction

The iridium selenides have been studied in this institute by Søvold (1954). He confirmed *inter alia* the existence of the compound IrSe_2 , which was first prepared by Wöhler, Ewald & Krall (1933).

Søvold succeeded in preparing a few crystals of IrSe_2 and used a prismatic one with cross-section 0.07×0.07 mm.² and length 0.1 mm. along the shortest axis to take Weissenberg photographs about the three axes, using $\text{Cu } K\alpha$ radiation ($\lambda = 1.542$ Å). With these photographs he determined the unit-cell dimensions, and from the systematic absences of reflexions the two possible space groups $Pnam$ and Pna , with $Pnam$ as the more probable.

By combining the X-ray data and the pycnometrically determined densities Søvold concluded that the diselenide possibly is a non stoichiometric compound, with a selenium content corresponding to the formula IrSe_{2-x} , where $x \approx 0.1$. The calculations here will be based on the composition IrSe_2 .

Moreover, Søvold measured the magnetic susceptibility of iridium diselenide and found a weak and nearly temperature-independent paramagnetism.

Determination of the structure

In the calculation of the Fourier projections it seemed safer to use the space group $Pna2_1$, because only one of the special positions of the space group $Pnam$ proved