hexagonal phases, contains within it empty hexagonal tunnels (Magnéli, 1949*b*). It is interesting to note that this has been identified in needles grown from tungsten metal in an electron-diffraction camera, (Hashimoto, Tanaka, Yoda & Araki, 1958) in which the high vacuum could prevent the formation of an AX_3 type oxide.

We wish to thank Dr L. S. Williams for drawing our attention to the whisker growths.

References

ANDERSSON, G. (1953). *Acta Chem. Scand.* **7**, 154.
HÄGG, G. & MAGNÉLI, A. (1954). *Revs. Pure Appl. Chem.* (*Australia*), **4**, 235.
HASHIMOTO, H., TANAKA, K., YODA, E. & ARAKI, H. (1958). *Acta Met.* **6**, 557.
KEHL, W. L., HAY, R. G. & WAHL, D. (1952). *J. Appl. Phys.* **23**, 212.
MAGNÉLI, A. (1949*a*). *Acta Chem. Scand.* **3**, 88.
MAGNÉLI, A. (1949*b*). *Ark. Kemi*, **1**, 223.
MAGNÉLI, A. (1953). *Acta Chem. Scand.* **7**, 315.
MATTHIAS, B. T. & WOOD, E. A. (1951). *Phys. Rev.* **84**, 1255.
RIECK, G. D. (1943). *Rec. Trav. Chim. Pays-Bas*, **62**, 427.
THOMAS, L. H. & UMEDA, K. (1957). *J. Chem. Phys.* **26**, 293.
VIERVOLL, H. & ØGRIM, O. (1949). *Acta Cryst.* **2**, 277.
WADSLEY, A. D. (1955). *Revs. Pure Appl. Chem.* (*Australia*), **5**, 165.
WESTMAN, S. & MAGNÉLI, A. (1958). *Acta Chem. Scand.* **12**, 363.
WYART, J. & FOËX, M. (1951). *C. R. Acad. Sci., Paris*, **232**, 2459.

Acta Cryst. (1961). **14**, 383

The Position of Anomalous Scatterers in Protein Crystals

BY MICHAEL G. ROSSMANN

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, Cambridge, England

(*Received* 7 *January* 1960 *and in revised form* 18 *May* 1960)

A Patterson summation with $\{|F(hkl)|-|F(\bar{h}\bar{k}\bar{l})|\}^2$ coefficients gives peaks at the end of vectors relating the anomalous scatterers in the structure. Application to two mercury derivatives of horse haemoglobin, using Cu $K\alpha$ diffraction data, shows that the ends of vectors between mercury atoms can be recognized easily, and the mercury positions determined, without resorting to isomorphous replacement. Interactions between the iron atoms of the haem groups and the mercury atoms are lost in the background, but can nevertheless be located by removing the spurious background peaks with Buerger's Minimum Function.

Introduction

When a crystal exhibits an observable departure from Friedel's law, in that $F(hkl)$ is markedly different from $F(\bar{h}\bar{k}\bar{l})$, then two types of Patterson functions can be constructed (Okaya *et al.*, 1955; for a summary of this work and general review of the literature see Buerger, 1959, pp. 76–79). The first of these, the cosine function, is defined by

$$P_c(u,v,w)=\frac{1}{V}\sum_h\sum_k\sum_l|F(hkl)|^2\cos 2\pi(hu+kv+lw)$$

and differs from the normal Patterson only in the detailed shape of the peaks. The second function, defined by

$$P_s(u,v,w)=\frac{1}{V}\sum_h\sum_k\sum_l|F(hkl)|^2\sin 2\pi(hu+kv+lw)$$

gives positive peaks in one direction and negative peaks in the opposite direction at the end of vectors relating the anomalous scatterers with the remainder of the scattering material in the cell. Neither of these functions has so far been of use in the analysis of protein structures, for the interaction between the anomalous scatterers with the rest of the protein is sufficient to obscure the vectors between the anomalous scatterers themselves.

This paper deals with an attempt to utilize the observed anomalous scattering in order to determine the positions of only those atoms responsible for the deviations from Friedel's law. An application to horse haemoglobin shows that it is possible to determine the position of the heavy mercury atoms from the anomalous data of one compound only, without resorting to any other isomorphous derivative. It was also possible to determine the position of the iron-containing haem groups, without excessive interference from small anomalous scatterers such as sulphur atoms, which are present in fairly large quantities in haemoglobin crystals. This was, therefore, an independent check on the haem positions as found by Perutz *et al.* (1960).

Theory

The scattering factor of an atom can be regarded as having two components:

1. A real component f_0, due to the electron configuration.
2. An additional complex component which is only appreciable when the X-ray wavelength is near an absorption edge.

We can therefore write

$$\begin{aligned} f &= f_0 + (\Delta f' + i\Delta f'') \\ &= f' + if'' \text{ for simplicity.} \end{aligned}$$

Thus, if $\mathbf{h}$ is the reciprocal-lattice vector which joins the point (h, k, l) to the origin, and if $\mathbf{r}_n$ is the position vector of the nth atom in real space, then the structure factor of the (hkl) plane may be written as

$$F(\mathbf{h})=\sum_{i=1}^{I} f_{0i}\exp[2\pi i\mathbf{h}.\mathbf{r}_i]+\sum_{j=1}^{J}(f'_j+if''_j)\exp[2\pi i\mathbf{h}.\mathbf{r}_j]\,.$$

I is the number of atoms in the unit cell that do not have any appreciable anomalous scattering, and J is the number which cause measurable deviation from Friedel's law.

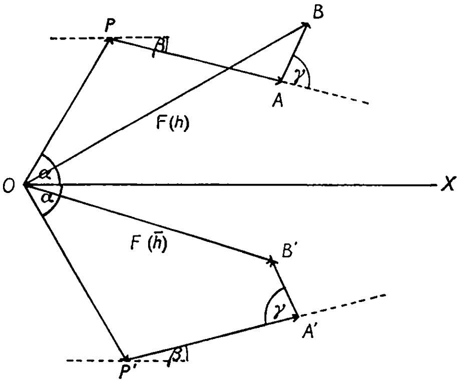


Fig. 1. Vector components of two Friedel related structure factors $F(\mathbf{h})$ and $F(\bar{\mathbf{h}})$. **OP** is the protein contribution, **PA** is the normal and **AB** the anomalous contribution of the anomalous scatterers to $F(\mathbf{h})$.

Let us now construct Fig. 1. OX is the reference line

$$\mathbf{OP}=\sum_{i=1}^{I} f_{0i}\exp[2\pi i\mathbf{h}.\mathbf{r}_i]\,,$$

which represents the contribution of the normal scattering material in the cell to the total structure factor.

$$\mathbf{PA}=\sum_{j=1}^{J} f'_j\exp[2\pi i\mathbf{h}.\mathbf{r}_j]\,,$$

which represents the normal contribution of the anomalous scattering material.

$$\mathbf{AB}=i\sum_{j=1}^{J} f''_j\exp[2\pi i\mathbf{h}.\mathbf{r}_j]\,.$$

Hence

$$F(\mathbf{h})=\mathbf{OP}+\mathbf{PA}+\mathbf{AB}\,.$$

Let us now define the phrase *atoms of the same type* to mean those atoms which have the same ratio $k=f''/f'$ for any one reflection. Hence, if the anomalous scatterers in the unit cell are all of the same type,

$$\begin{aligned} \mathbf{AB} &= ik\sum_{j=1}^{J} f'_j\exp[2\pi i\mathbf{h}.\mathbf{r}_j] \\ &= ik\mathbf{PA}\,. \end{aligned}$$

That is, in this special case, $AB/PA=f''/f'$ and **AB** leads **PA** by a phase angle of $\pi/2$.

We shall now prove that a Patterson synthesis with $\{|F(\mathbf{h})|-|F(\bar{\mathbf{h}})|\}^2$ coefficients will have positive peaks at the ends of vectors relating anomalous scatterers.*

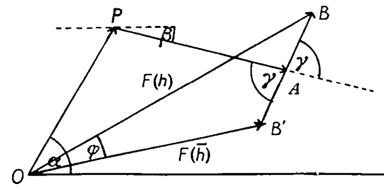


Fig. 2. Argand diagram showing relationship between $F(\mathbf{h})$ and the complex conjugate of $F(\bar{\mathbf{h}})$.

Let **OP** have a phase angle α, let **PA** have a phase angle $-\beta$ and let **AB** lead **PA** by an angle γ (Fig. 1). Similarly for reflection $F(\bar{\mathbf{h}})$, $OP'=OP$ but has a phase angle of $-\alpha$, $P'A'=PA$ but has a phase angle of β, and $A'B'=AB$ leading $P'A'$ by an angle $180-\gamma$. Thus $F(\bar{\mathbf{h}})=\mathbf{OP'}+\mathbf{P'A'}+\mathbf{A'B'}$. We may now construct Fig. 2 from Fig. 1 by reflecting everything below the reference line OX, into the top half of the

---

* The two dimensional form of the anomalous dispersion Patterson with $\{|F(\mathbf{h})|-|F(\bar{\mathbf{h}})|\}^2$ coefficients has been previously used by Blow (1957), in an attempt to find in projection, the positions of the mercury atoms of PCMB-haemoglobin. Due to the use of data restricted to one zone of reciprocal space, the results were only partly convincing, and were never published.

Blow pointed out that $(|F(\mathbf{h})|-|F(\bar{\mathbf{h}})|)^2$ may be regarded as a weighted version of $(F^2(\mathbf{h})-F^2(\bar{\mathbf{h}}))^2$, and that a synthesis using these coefficients may be regarded as the self-convolution of Okaya, Saito & Pepinsky's P_s function (1955). Since the P_s function has peaks at vectors relating normal scatterers to anomalous scatterers, its self-convolution will have strong peaks at the vectors relating anomalous scatterers, and if there are few anomalous scatterers and many normal scatterers, these peaks will dominate the synthesis.

This approach suggests that the background of the anomalous dispersion Patterson is made up of positive and negative peaks at the interactions of other peaks in the P_s function. For this reason, when there is a large number of unresolved atoms, the anomalous dispersion Patterson is likely to have less background then Pepinsky & Okaya's $(P-P_c)$ function (1956), which has a smaller number of peaks, all negative, in the background.

diagram. P' will then coincide with P, and A' with A. AB will make an angle of $(\pi-\gamma)$ with PA in a clockwise direction, while AB' will make an angle γ with PA in an anti-clockwise direction. Thus $BB'=2.AB$. If all the anomalous scatterers are of the same type then $\gamma=\pi/2$, and $BB' \propto PA$. More generally, BB' is proportional to the vector sum of the structure factor contributions of the J anomalously scattering atoms, whose relative weights are given by f_j''.

It can be shown (Rossmann, 1960*a*) that a Patterson with $(BB')^2$ coefficients is equivalent to a Patterson with $\{|OB|-|OB'|\}^2$ coefficients provided the angle $\varphi=BOB'$ is small. For most reflections φ must be very small since $f' \gg f''$, and hence $AB \ll PA$. Therefore the proposed Patterson will have positive peaks at the end of vectors between anomalous scatterers.

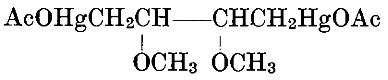
The last step in the argument can be put in a different form. Okaya, Saito & Pepinsky (1955) showed that their P_c function has peaks which would be produced by scatterers with form factors $(f_p'f_q'+f_p''f_q'')$ at the ends of vectors relating the pth and the qth atom. The vector **AB** contains no real scattering-factor component; thus $f'=0$. Hence, in a synthesis with coefficients $(BB')^2$, we find peaks with scattering factors $f_p''f_q''$ only. That is, these are peaks relating the different anomalous scatterers in the unit cell.

Application to horse haemoglobin

Horse haemoglobin crystallizes in space group $C2$ with cell dimensions

$$a=109{\cdot}2,\ b=63{\cdot}2,\ c=54{\cdot}7\ \text{Å},\ \beta=110{\cdot}7^\circ.$$

A number of different isomorphous heavy-atom compounds have been prepared for the three-dimensional Fourier synthesis of haemoglobin (Perutz, Rossmann, Cullis, Muirhead, Will & North, 1960). For three of these, the intensities of some *hkl* and $\bar{h}\bar{k}\bar{l}$ reflections were measured separately. A nearly complete coverage of all reflections in the 5·8 Å sphere, using Cu $K\alpha$ radiation, was made for the compound containing 4 moles of $HgCl_2$ per mole of haemoglobin, and for the compound with 2 moles of



$$\underset{\displaystyle OCH_3}{AcOHgCH_2\overset{}{C}H}\text{——}\underset{\displaystyle OCH_3}{CHCH_2HgOAc}$$

per mole of haemoglobin .

The latter will be referred to here as the 'Baker mercurial'. The $HgCl_2$ compound has mercury attached to two sites in the asymmetric unit, but full substitution was achieved at only one of these. The Baker mercurial has full substitution at the end which is attached to a sulphur atom in the protein, but at the other end of the mercurial only half the molecules seem to be substituted with mercury. The site which was fully substituted in the Baker mercurial coincided with the site which was only partially substituted in the $HgCl_2$ compound. The two mercury atoms in the Baker mercurial were not resolved, and could therefore be regarded as equivalent to a single, heavily substituted site per half molecule. Significant anomalous dispersion $(|I(h)-I(\bar{\mathbf{h}})/I_{\text{mean}}|>0{\cdot}1)$ was found in 268 and 258 of the 1100 reflections in the 5·8 Å sphere for the $HgCl_2$ and Baker mercurial compounds

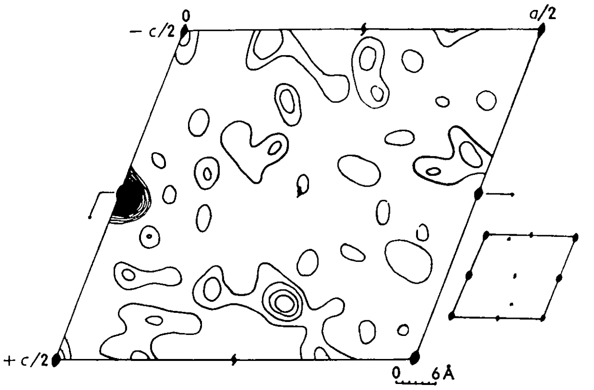


Fig. 3. Harker section through anomalous dispersion Patterson of $HgCl_2$ compound. Controus at equal but arbitrary intervals.

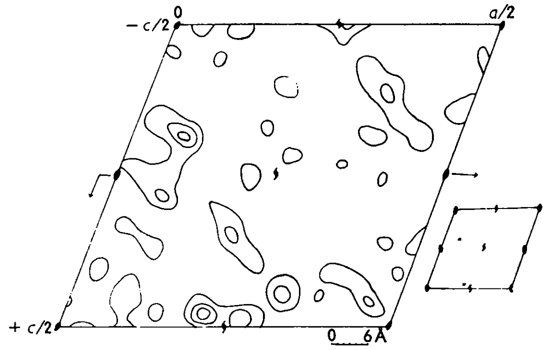


Fig. 4. Section $y=(6/32)b$ through anomalous dispersion Pattersom of $HgCl_2$ compound. Contours at equal but arbitrary intervals.

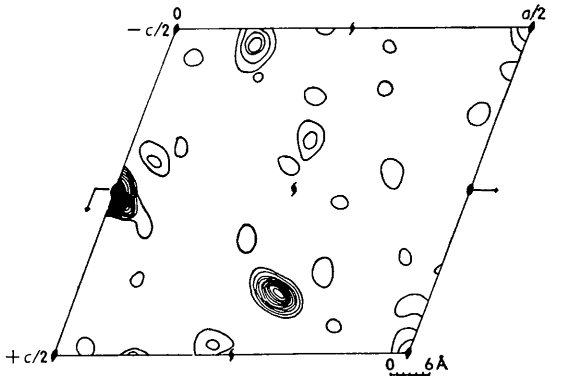


Fig. 5. Harker section through correlation function of $HgCl_2$ compound and unsubstituted proteins. Contours at equal but arbitrary intervals.

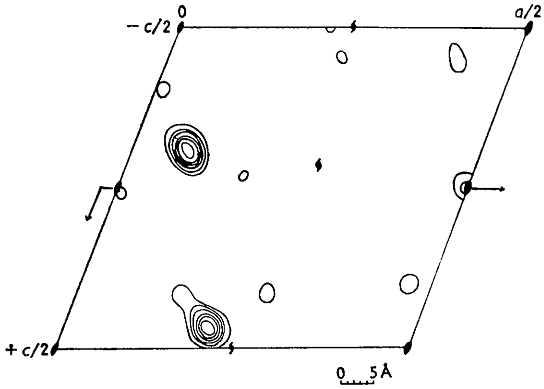


Fig. 6. Section $y=(6/32)b$ through correlation function for $HgCl_2$ compound and unsubstituted protein. Contours at equal arbitrary intervals.

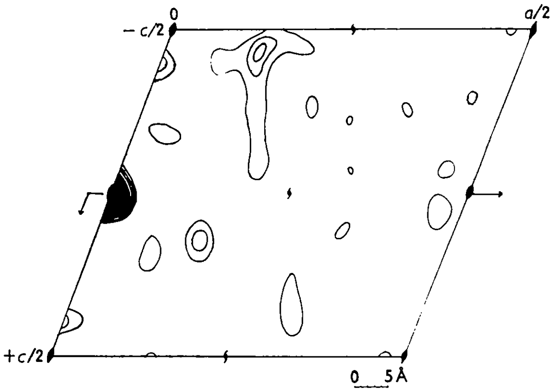


Fig. 7. Harker section through anomalous dispersion Patterson mercurial compound. Contours at equal but arbitrary intervals.

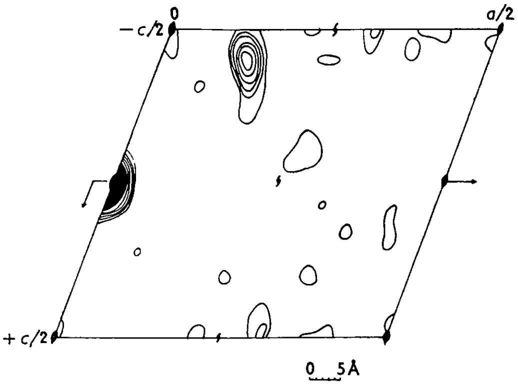


Fig. 8. Harker section through correlation function for Baker mercurial compound and unsubstituted protein. Contours at equal but arbitrary intervals.

respectively. Fig. 3 shows the Harker section, $y=0$, and Fig. 4 the section $y=\frac{6}{32}b$ on which both general peaks appear in the three-dimensional anomalous-dispersion Patterson functions of the $HgCl_2$ compound. These should be compared with Figs. 5 and 6 which show the same sections for the corresponding $(|F_{\text{heavy}}|-|F_{\text{unsubstituted}}|)^2$ correlation function (Rossmann, 1960*a*) which determines the position of the mercury atoms with the help of the unsubstituted isomorphous compound. The greatest difference in peak positions due to Hg · · · Hg interactions between the anomalous-dispersion synthesis and the isomorphous-replacement synthesis is 1·5 Å. Figs. 7 and 8 are respectively the Harker sections of the anomalous-dispersion synthesis and of the correlation function for the 'Baker mercurial' compound. Table 1 compares the relative peak heights of Hg · · · Hg interactions.

Table 1. *Peak heights of* Hg · · · Hg *interactions on arbitrary scales*

Interactions	Anomalous dispersion Patterson		Isomorphous replacement correlation	
	$HgCl_2$	Baker mercurial	$HgCl_2$	Baker mercurial
$Hg_1 \cdots Hg_1$	39	—	42	—
$Hg_2 \cdots Hg_2$	24	32	21	32
$Hg_1 \cdots Hg_2$ (*A*)	26	—	25	—
$Hg_1 \cdots Hg_2$ (*B*)	27	—	26	—
Origin	194	105	117	130
Highest background peak	22	23	10	10

Table 2. *Comparison of anomalous scattering for one* Fe *and one* Hg *atom with* Cu $K\alpha$ *radiation*

	Z	$\Delta f'$	$\Delta f''$	(f'/f'') max.
Fe	26	−1·1	3·4	7·3
Hg	80	−5	8	9·4

Table 2 gives values for the anomalous dispersion of iron and mercury, according to the calculations of Dauben & Templeton (1955). It appears that f'' for iron is only just under half the corresponding value of mercury. This opens the possibility of determining the position of the two iron containing haem groups in the asymmetric unit. However, so many peaks only slightly less significant than the Hg · · · Hg peaks were found (16 peaks between the arbitrary heights 18 and 22 for the $HgCl_2$ compound) that a straightforward inspection of the anomalous dispersion synthesis failed to locate the haem groups.

The use of the Buerger minimum function to find the haem groups

Provided the Fe · · · Hg vectors show up at all, the problem of determining the haem positions from the anomalous dispersion synthesis is one of selecting a set of consistent vector peaks corresponding to the distribution of iron and mercury atoms. It is to be hoped that all the other background peaks caused by imperfect data, and approximations made in the theory of the method, are sufficiently random for only one solution to be possible. The Buerger minimum function (Buerger, 1951) seems to present itself as a

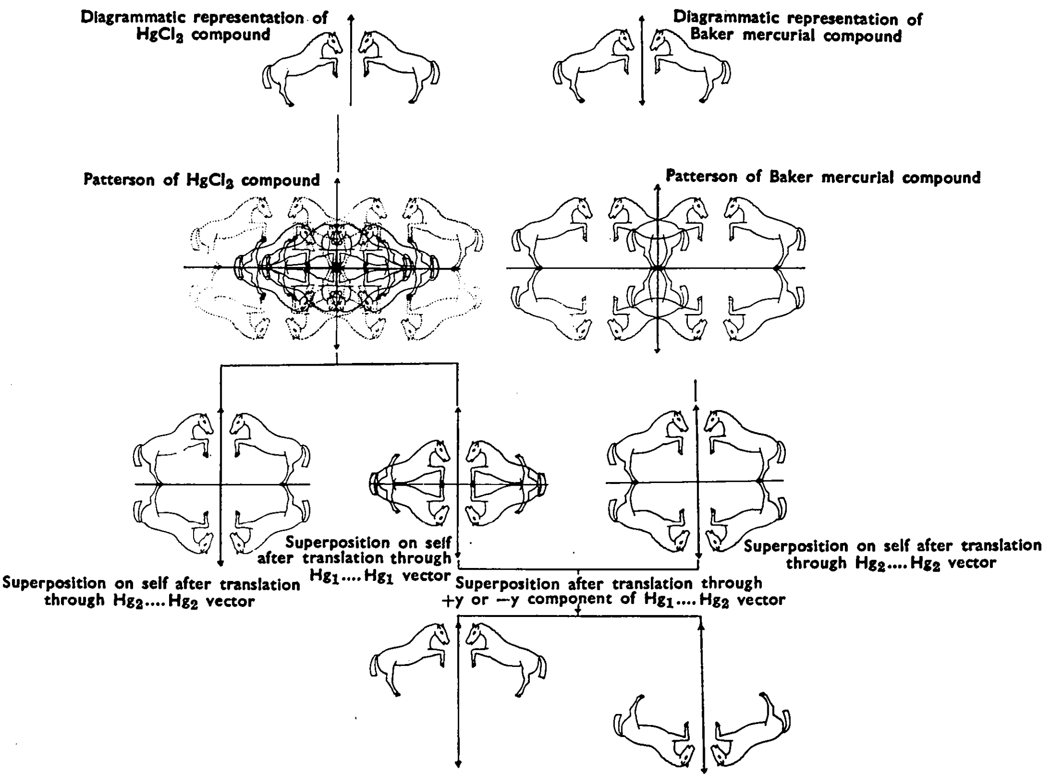


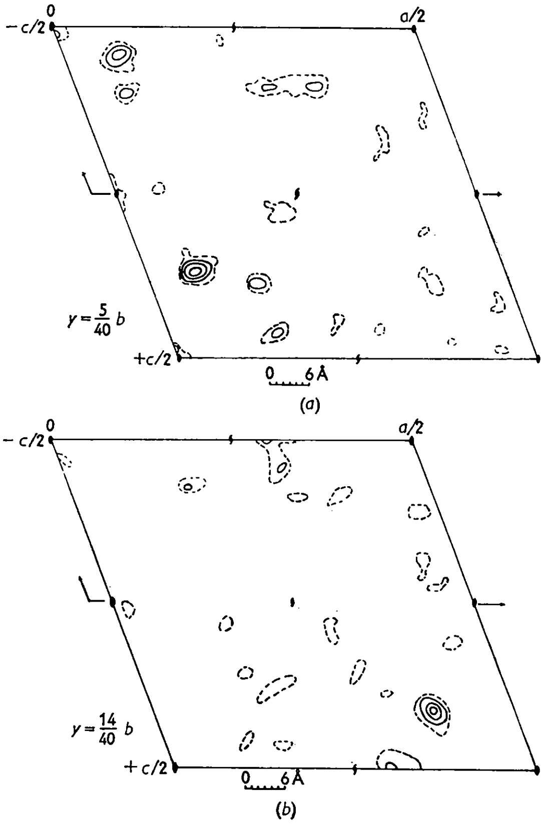
Fig. 9. Diagrammatic representation of the superposition technique used to determine the position of the iron atoms from the anomalous dispersion Pattersons of the $HgCl_2$ and Baker compounds. One horse represents the asymmetric motif, which, in the present situation, is the relative arrangement of iron and mercury atoms.

useful device to achieve this object. The method is identical when employed to solve a structure with difference Patterson synthesis between isomorphous compounds (Kartha & Ramachandran, 1955). The application to proteins will be dealt with elsewhere (Rossmann, 1960*b*) and hence only a brief account follows below.

Let us take a horse (Fig. 9) to represent the asymmetric unit of horse haemoglobin. Further, let us simplify the argument by assuming space group $P2$, rather than $C2$. The hooves represent the heavy mercury atoms. The amount of black in the hooves shows diagrammatically the proportion of substitution. Each horse represents the asymmetric motif which is, in the present case, the arrangement of the iron and mercury atoms. It is immaterial whether the motif is a continuous line or volume, or discrete points or volumes.

If the atoms in the outline of the horse contain H electrons, and a certain heavy atom contains Z electrons, then the Patterson synthesis will have peaks containing ZH (electrons)^2. These vectors will follow faithfully the outline of each horse in the unit cell. Thus each heavy atom produces an 'image' of the unit cell in Patterson space with the Patterson origin on the heavy-atom site. In addition we have the centrosymmetrically related image. We may, therefore, describe a Patterson synthesis as being the sum of the images of the unit cell seen through each heavy atom when it is at the Patterson origin, plus the centrosymmetrically related images. Further, the strength of each image (ZH) depends on the weight of the heavy atom (Z). Thus a fully substituted mercury atom gives a stronger image than a partially substituted mercury atom while the images in iron atoms are quite negligible. (Fig. 9 shows strong images in heavy outline.) The problem of interpreting the Patterson is thus reduced to sorting out a single image.

The images caused by two symmetry related atoms come into coincidence if we displace two identical Patterson maps by the known vector distance between the two images. If a new map could be constructed which shows only high areas in those regions where both the superimposed maps possess high areas, then we are essentially sorting out the images which have been brought in coincidence. Buerger suggested taking the density of the particular Patterson synthesis which has the lower density of the two superimposed maps, in a point-by-point comparison. This procedure can be carried out once for each pair of symmetry-related atoms. Each new minimum function will be a picture of the original cell plus its enantiomorph (Fig. 9). The original cell, or its enantiomorph can be found by superimposing thees results after translation through the y-component (diad axis is in y direction) of the difference of the y co-ordinates of the independent atoms. A choice between the alternative results (Fig. 9) can be made by selecting the one with the correct absolute configuration of the heavy atoms, provided that the latter is known.



Figs. 10(*a*) and (*b*). Sections through the iron containing haem groups of the final superimposed anomalous dispersion Pattersons. Fig. 10(*a*) shows section $y=(5/40)b$, and Fig. 10(*b*) shows section $y=(14/40)b$.

Superposition of the anomalous-dispersion synthesis of the $HgCl_2$ compound on the more completely substituted mercury atoms (Hg_1) showed the iron atoms, although some 'background' peaks were still larger than the iron atoms themselves. Superposition of the same anomalous-dispersion Patterson on the lesser-substituted mercury atom (Hg_2) did not give any good evidence as to the position of the iron atoms. It was therefore assumed that the image in the latter was too weak to be picked up from amongst the background 'noise'. If both sites in the asymmetric unit of the $HgCl_2$ compound had been fully substituted, it would have been possible to find the iron atom positions without the help of any other compound. Instead the strong image in Hg_2 of the 'Baker mercurial' compound had to be combined with the image in Hg_1 of the $HgCl_2$ compound (Fig. 9). The final result shows the iron atoms at a height of 29 arbitrary units and the Hg_2 atom came up with 47 units. Table 2 suggests that the relative peak heights of Fe:Hg are as 3·4:8 in an anomalous-dispersion synthesis; hence a ratio 29:47 is very acceptable when differential degrees of atomic randomness are also considered. The Hg_1 atom was eliminated by use of the Minimum Function, since the Baker mercurial compound cannot give an image of the Hg_1 atom for it only contains Hg_2 atoms. The highest background peak was 27 arbitrary units. The position of the iron atoms agreed to within 1·0 Å of the positions determined by Perutz *et al.* (1960) using isomorphous-replacement techniques. Figs. 10(*a*), (*b*) shows the sections through the final superposition map which contain the iron atoms.

The anomalous dispersion data used in this work were collected by Dr M. F. Perutz, Miss H. Muirhead and Dr A. C. T. North as an aid to the determination of the phase angles of horse haemoglobin. I should like to thank them for making their observations available to me for the present study. I should also like to thank the University of Cambridge Mathematical Laboratory for the very considerable help in the preparation of the electronic computer EDSAC 2 to the requirements of the programmes which were necessary for the work described in this paper. Finally, I should like to acknowledge the very able assistance of Miss B. Davies.

References

Blow, D. M. (1957). Ph.D. thesis. University of Cambridge.

Buerger, M. J. (1951). *Acta Cryst.* **4**, 531.

Buerger, M. J. (1959). *Vector Space.* New York: Wiley & Sons.

Cullis, A. F., Dintzis, H. M. & Perutz, M. F. (1957). *Conference on Haemoglobin*, p. 50. NAS-NRC Publication 557, Washington.

Dauben, C. B. & Templeton, D. H. (1955). *Acta Cryst.* **8**, 841.

Kartha, G. & Ramachandran, G. N. (1955). *Acta Cryst.* **8**, 195.

Okaya, Y., Saito, Y. & Pepinsky, R. (1955). *Phys. Rev.* **98**, 1857.

Pepinsky, R. & Okaya, Y. (1956). *Proc. Nat. Acad. Sci., Wash.* **42**, 286.

Perutz, M. F., Rossmann, M. G., Cullis, A. F., Muirhead, H., Will, G. & North, A. C. T. (1960). *Nature, Lond.* **185**, 416.

Rossmann, M. G. (1960*a*). *Acta Cryst.* **13**, 221.

Rossmann, M. G. (1960*b*). Conference on Computing and the Phase Problem, Glasgow.