

and refinement was usually complete after ten to fifteen cycles.

The refined values of the parameters compared well with the Rossmann programme values, except for the effective atomic numbers, which were uniformly smaller than indicated by the Rossmann programme. The Hart  $h0l$  values were used in the 2 Å Fourier synthesis, although there may be some question remaining about the proper  $Z$  values to use, as evidenced by negative regions in the myoglobin Fourier synthesis at the three heavy-atom sites and some blurring of detail in the immediate neighbourhood. This problem is currently being investigated.

Thanks are due to Dr R. G. Hart, whose heavy-atom refinement programme proved invaluable, and to Dr M. G. Rossmann, whose least-squares refinement provided an independent check on the parameters. We should also like to express our appreciation to Miss Mary Pinkerton, whose assistance was invaluable at all stages of the work.

## References

- BLOW, D. M. (1958). *Proc. Roy. Soc. A*, **247**, 302.  
 BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794.  
 BODO, G., DINTZIS, H. M., KENDREW, J. C. & WYCKOFF, H. W. (1959). *Proc. Roy. Soc. A*, **253**, 70.  
 BRAGG, W. L. & PERUTZ, M. F. (1954). *Proc. Roy. Soc. A*, **225**, 315.  
 DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. E. (1960). *Symposium on Computer Methods and the Phase Problem*, p. 84. Glasgow: Pergamon Press.  
 GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. Roy. Soc. A*, **225**, 287.  
 HARKER, D. (1956). *Acta Cryst.* **9**, 1.  
 KENDREW, J. C., BODO, G., DINTZIS, H. M., PARRISH, R. G., WYCKOFF, H. & PHILLIPS, D. C. (1958). *Nature, Lond.* **181**, 662.  
 KENDREW, J. C., DICKERSON, R. E., STRANDBERG, B. E., HART, R. G., DAVIES, D. R., PHILLIPS, D. C. & SHORE, V. C. (1960). *Nature, Lond.* **185**, 422.  
 ROSSMANN, M. G. (1960). *Acta Cryst.* **13**, 221.

*Acta Cryst.* (1961). **14**, 1195

## The Single Isomorphous Replacement Method

BY D. M. BLOW AND MICHAEL G. ROSSMANN

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

(Received 8 September 1960 and in revised form 31 October 1960)

Despite the phase ambiguity which arises when one isomorphous pair is used to determine phases in a non-centrosymmetric structure, a single pair of compounds can be used to give an interpretable Fourier synthesis. Two conditions must be satisfied: the replacing atoms must themselves form a non-centrosymmetric array, and a sufficient number of terms must be available.

The method has been applied to the crystalline proteins haemoglobin and myoglobin. Examples are given which show the improvements which can be made by the use of weighting functions and by the introduction of anomalous-dispersion data.

### 1. Introduction

The isomorphous-replacement method, when applied in its conventional form to a non-centrosymmetric structure, leads to an ambiguous result for the phase angle (Bokhoven *et al.*, 1951). A general method of removing the ambiguity is to employ a series of compounds with isomorphous replacements at different sites, and this method has been used successfully with the proteins myoglobin and haemoglobin (Kendrew *et al.*, 1960; Perutz *et al.*, 1960). However, the preparation of suitable isomorphous protein derivatives has been a matter of great difficulty, and is likely to remain the most time-consuming step in a protein structure determination. It is therefore important to find methods which use the minimum number of isomorphous derivatives.

Rogers (1951) suggested a procedure applicable

when only one isomorphous pair is available. Although in the form proposed it only applies to different atoms substituting at the same site, it may readily be generalized. A synthesis is calculated in which each term is given the phase of the scattering of the replacing electrons, with a sign chosen according as the intensity is increased or decreased by the replacement. When the replacing electrons form a non-centrosymmetric arrangement, this synthesis tends to be similar to the electron-density function, but with background superimposed. Rogers' function has the unfortunate property of giving strong weight to terms where the isomorphous replacement has little effect on the intensity.

Kartha & Ramachandran (1955) showed how the minimum function (Buerger, 1951) applied to the difference Patterson could in principle reveal a non-centrosymmetric structure under the same circum-

stances. Ramachandran & Raman (1959), in attempting to find a 'reciprocal space' method analogous to the 'vector space' method, proposed several functions with properties similar to that of Rogers (1951). The same limitations on heavy-atom position were assumed, but the generalization is straightforward. One of these functions,  $\beta_{is}^0$ , so generalized, is essentially the one now proposed by us, arrived at on quite independent grounds.

Recently an organic structure has been solved by application of this method (Kartha, private communication, 1960).\*

## 2. The method

The fundamental equation of the isomorphous-replacement method is

$$\mathbf{F}_2 = \mathbf{F}_1 + \mathbf{f}_{12}, \quad (1)$$

where  $\mathbf{F}_1$ ,  $\mathbf{F}_2$  are the structure factors of two isomorphous compounds, and  $\mathbf{f}_{12}$  the calculated structure factor of the scattering matter in structure 2 not present in structure 1.

Neglecting error, a single pair of isomorphous compounds gives two possible solutions for the structure factor  $\mathbf{F}_1$  in (1), which may be called  $\mathbf{F}_a$  and  $\mathbf{F}_b$  (see Fig. 1). In the absence of other information the two solutions are equally probable. It has been shown by Blow & Crick (1959), that the best one can do in such a case is to use the mid-point of the two, namely  $\xi = \frac{1}{2}(\mathbf{F}_a + \mathbf{F}_b)$ , as the structure factor in a Fourier synthesis.

The argument which shows that the true structure should be obtainable by application of the minimum function to the difference Patterson (Kartha & Ramachandran, 1955) is based on perfect resolution and point atoms. In the next section we consider the conditions required for the method to work in practice. In section 4, we describe in detail its application to protein structures and demonstrate the effect of certain refinements. Section 5 deals with the combination of these results with anomalous scattering data.

## 3. The reciprocal space argument

Any experimental method for the determination of structure factors may be considered to lead to a result  $\mathbf{F}_{exp}$  which differs from the true result  $\mathbf{F}$  by an error  $\epsilon$ :

$$\mathbf{F}_{exp} = \mathbf{F} + \epsilon. \quad (2)$$

Thus in the non-centrosymmetric case, most of the error will arise from uncertainty about the phase.

If the  $\mathbf{F}_{exp}$  are now used to calculate a Fourier transform, the result may be thought of as the true one, with the addition of the transform of the  $\epsilon$ .

\* Note added in proof.— Kartha (1961) has published a paper presenting the same fundamental idea as is expressed in § 2, together with an example from an imaginary structure.

As the  $\epsilon$  are errors, let it be assumed for the time being that they have random phase.

Consider the effect of increasing the number of terms,  $N$ , by increasing the resolution. For simplicity, assume the structure to consist of  $n$  point atoms, distributed at random in the space group  $P1$  according to the postulates of Wilson statistics (1949), so that the mean  $|\overline{\mathbf{F}}|^2 = n f^2$  remains constant as resolution is increased. Assume, further, that  $|\overline{\epsilon}|^2$  remains constant. If  $V$  is the unit-cell volume, the peak heights at the atomic peaks will be  $Nf/V$ , and thus increase proportionately to  $N$ . The background due to error increases, as in the usual random-walk problem, according to the sum of squares, so that its r.m.s. value will be  $(N|\overline{\epsilon}|^2)^{1/2}/V$ , increasing proportionately to  $\sqrt{N}$ . (Series-termination effects are irrelevant to the argument.) We thus have

$$\text{peak/background contrast} = \sqrt{N}f/(\sqrt{N}|\overline{\epsilon}|)^{1/2}. \quad (3)$$

(If a form factor is introduced which applies equally to both  $\mathbf{F}$  and  $\epsilon$ , the terms at higher resolution are associated with weaker intensities. The peak/background ratio does not then increase so rapidly as  $N$  increases, and finally the ratio converges to a constant value as the  $|\mathbf{F}|$ 's become negligible.)

These ideas may now be applied to the single isomorphous-replacement method (Fig. 1). Suppose

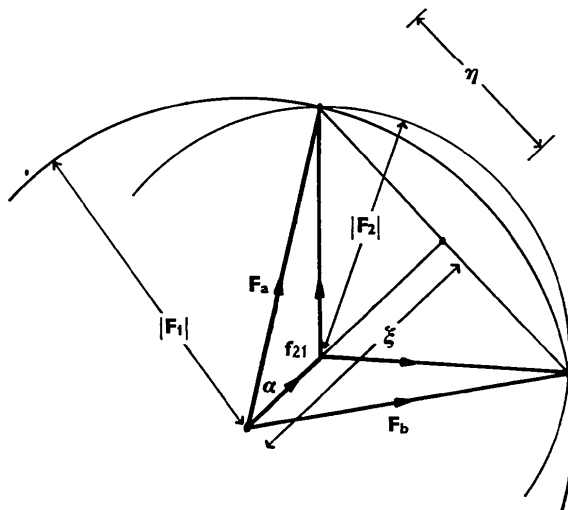


Fig. 1. A single isomorphous replacement in the general non-centrosymmetric case, showing the significance of the vector  $\xi$ .

that  $\mathbf{F}_a$  is the correct value for the structure factor. Then using  $\xi$  in the Fourier summation may be thought of as adding an error term  $\epsilon = -\eta$  to this term, where  $\eta = \frac{1}{2}(\mathbf{F}_a - \mathbf{F}_b)$ . The same is true of all the terms: they consist of 'true' components, which will add coherently to give point peaks of height  $Nf/V$  as before; and in addition 'error' terms  $\epsilon = \pm\eta$  which may be assumed to have random phase.  $\eta = |\mathbf{F}| \sin \alpha$

(Fig. 1), so that  $|\bar{\epsilon}|^2 = \frac{1}{2}|\bar{F}|^2$ , where we have averaged over all values of  $\alpha$ . Using the result  $|\bar{F}|^2 = nf^2$ , this gives  $|\bar{\epsilon}|^2 = \frac{1}{2}nf^2$ , which may be substituted into equation (3) to give

$$\text{peak/background contrast} = (2N/n)^{\frac{1}{2}}.$$

Thus it is clear that in the ideal case of point atoms, the contrast may be made as large as we please by increasing  $N$ . This corresponds to the ideally accurate superposition procedure discussed by Kartha & Ramachandran (1955).

In normal organic structures there are about 20 Å<sup>3</sup> per atom (neglecting hydrogen atoms), and it can be shown to follow that for three-dimensional data  $N/n \approx 35/R^3$ , where  $R$  is the resolution in Å. Thus for organic structures peak/background contrast  $\approx 8.4 R^{-3/2}$ , so that a just-acceptable contrast of 3 could be attained by using a resolution of about 2 Å.

We must now review the assumptions on which this result rests, and consider their validity under various conditions. First, we have assumed that the number of reflexions,  $N$ , is sufficient for various statistical assumptions to be reasonable. This is a good assumption for all structures to which isomorphous replacement is likely to be applied. Even for  $n=10$  non-hydrogen atoms, the required resolution demands about 45 reflexions, which would make the statistical predictions reasonably accurate. The assumption of point atoms is less justified. At 2 Å the unitary scattering factor of carbon has fallen to 0.43 and this would reduce the contrast considerably below the predicted level. In proteins, crystal imperfection causes a much more rapid decay of scattered intensity with  $\sin \theta$ , and at 2 Å the contrast would have approached its limit as the intensities fall towards zero. In both these cases we are, however, justified in sharpening the data, since the inherent errors of the method are much more important than the experimental errors which would be magnified by sharpening.

Proteins gain an advantage over the simple theory because the assumption of a Wilsonian distribution of atoms in the structure is not optimistic enough. This arises, at comparatively high resolutions, because a part of the unit cell is inaccessible to the rigid molecule and is filled with a non-crystalline salt-water mixture. Thus, for myoglobin at 2 Å,  $N=9600$  reflexions for  $n$  = about 1200 atoms and hence leads to a peak/background contrast of 4 instead of 3, as might be expected for a simple organic structure.

An advantage arises also at much lower resolution for the  $\alpha$ -proteins, because the arrangement of the polypeptide backbone into tightly packed helices gives rise to a contrast much greater than that expected by Wilson statistics. Thus, at 6 Å resolution, the  $\alpha$ -helix shows a contrast of about 0.7 e.Å<sup>-3</sup> over the surrounding side chains, while the expected background due to error is  $(\frac{1}{2}N|\bar{F}|^2)^{\frac{1}{2}}$  which for haemoglobin is

about 0.22 e.Å<sup>-3</sup>, giving a just acceptable contrast of 3.

Finally the assumption that the phase of  $\eta$  is random must be considered. The effect of a non-random phase distribution is seen most clearly by reference to the case where the array of replacing atoms is centrosymmetric. In this case  $\xi$  is always real,  $\eta$  always imaginary, and the effect of the errors is to generate the mirror image of the structure as well as the required structure. It is clear that, for any simple distribution of replacing atoms, there will be points corresponding to partial centres of symmetry. If, for instance, there are four atoms in the unit cell, forming a non-centrosymmetric array, there will be a point midway between any two atoms which is a partial centre of symmetry. If this point is taken as origin, the contribution of two of the four atoms to the structure factor  $f_{12}$  will always be real.  $f_{12}$  will usually have larger real parts than imaginary parts, and the same is true of  $\xi$ . Thus the single isomorphous replacement method will tend to generate a weak image of the structure by inversion through each partial centre.

Patterson (1949) showed that a Fourier transform with coefficients

$$f^2 = a^2 - b^2 + 2iab$$

provides a map of the 'centrosymmetriness' of the structure obtained by Fourier transformation of  $f$ . This function has been called the Patterson function of the second kind, to distinguish it from the more familiar function with coefficients  $|f|^2$ . It may be shown that the single isomorphous-replacement method results in the convolution of the true structure with a centrosymmetriness map of the replacing atoms, very similar to the Patterson function of the second kind. A general method of deconvolution has not yet been found. We believe this to be the most serious difficulty of the method in its application to proteins, and, in the case where there are four equal replacing atoms per unit cell, it may give rise to images half as strong as the true structure.

It may be pointed out that exactly the same difficulty arises with the 'heavy-atom' method of structure determination in the non-centrosymmetric case.

To summarize the arguments of the last few paragraphs, in the ideal case of point atoms the method would work well if the resolution is taken beyond the distance between the atoms. For real organic structures it is much more than sufficient to go to the limit of the copper sphere in data collection. For proteins, the special features of the structures cause the method to be applicable (with sharpening) at the limited resolution available, and in structures containing  $\alpha$ -helices it is also applicable at about 6 Å resolution. A major difficulty in all of these applications is the appearance of weaker, spurious images.

#### 4. Straightforward application of the single isomorphous replacement method

Referring to Fig. 1, it is clear that  $\xi$  has the phase of  $\mathbf{f}_{12}$  and magnitude  $F_1 \cos \alpha$ , where  $\alpha$  is given by

$$F_2^2 = F_1^2 + |\mathbf{f}_{12}|^2 + 2F_1|\mathbf{f}_{12}| \cos \alpha.$$

Writing  $\mathbf{f}_{12} = a + ib$  and  $\xi = A + iB$ , it follows that

$$A = |\xi|a/|\mathbf{f}_{12}|, \quad B = |\xi|b/|\mathbf{f}_{12}|, \quad (4)$$

where

$$|\xi| = |(F_2^2 - F_1^2 - |\mathbf{f}_{12}|^2)/2|\mathbf{f}_{12}|,$$

with the exceptions given below [(5)]. In practice, error will sometimes prohibit construction of a closed triangle with sides  $F_1$ ,  $F_2$ ,  $|\mathbf{f}_{12}|$ , and in this case the best that can be done is to give  $\xi$  the magnitude of  $F_1$ . The full conditions are:

$$(a) \quad \text{if } |\mathbf{f}_{12}| < F_1 - F_2, \quad |\xi| = -F_1, \quad (5a)$$

$$(b) \quad \text{if } |\mathbf{f}_{12}| < F_2 - F_1, \quad |\xi| = +F_1, \quad (5b)$$

$$(c) \quad \text{if } |\mathbf{f}_{12}| > F_1 + F_2, \quad |\xi| = -F_1. \quad (5c)$$

The  $\beta_{15}^0$  function (Ramachandran & Raman, 1959) provides the same weighting scheme, except that no special provision is made when the above inequalities apply. They become especially important when  $|\mathbf{f}_{12}|$  is small, and in this case the function of Rogers (1951) is also seriously inadequate. A more consistent way of dealing with this case is described below.

We used the data of Perutz *et al.* (1960) from a haemoglobin compound containing dimercurey acetate and mercury acetate ('DMA + HgAc<sub>2</sub>') to give  $F_2$ ,  $a$  and  $b$ ; the  $F_1$  were taken from the unsubstituted haemoglobin data of Perutz *et al.* The 'DMA + HgAc<sub>2</sub>' compound contains two independent heavy-atom sites in the C2 unit cell. Substitution appears to be incomplete at both sites, but the DMA site has about

1.9 times as much mercury as the HgAc<sub>2</sub>. The available data comprised 1074 independent reflexions, going out to a spacing of 5.8 Å, and at this resolution the two mercury atoms of DMA are not resolved, but appear as an elongated peak. Full details about these compounds are to be published elsewhere (Cullis *et al.*, 1961).

Fig. 2(a) shows one section from the resulting 3-dimensional Fourier synthesis. This section includes one of the two iron atoms in the asymmetric unit of the haemoglobin molecule, and the comparable section calculated from seven isomorphous compounds by Perutz *et al.* (1960) is given for comparison (Fig. 2(b)).

The agreement is striking, but obviously shows room for improvement. Our first thought was to reduce the effect of reflexions in which  $|\mathbf{f}_{12}|$  is small compared to experimental error. When this happens one of the special cases (5a) or (5b) will usually apply, and the equations give  $\xi$  the full magnitude of  $F_1$ , although in fact we have very little knowledge about the phase angle, and it would be more correct to omit the reflexion. It was therefore decided to use the weighting functions proposed by Blow & Crick (1959), which takes care of cases of this sort. Instead of  $|\xi|$  in equation (4) we use a function  $|\xi'|$  defined by

$$|\xi'| = F_1 \frac{\int_0^\pi \cos \alpha \exp \{-\varepsilon^2/2E^2\} d\alpha}{\int_0^\pi \exp \{-\varepsilon^2/2E^2\} d\alpha}, \quad (6)$$

where

$$\varepsilon = F_2 - (F_1^2 + |\mathbf{f}_{12}|^2 + 2F_1|\mathbf{f}_{12}| \cos \alpha)^{\frac{1}{2}}$$

and  $E$  is an estimate of the total r.m.s. error in the observed and calculated structure factors.  $E$  was taken to be  $(50 + 0.05F_1)$  e./unit cell for the calculations on

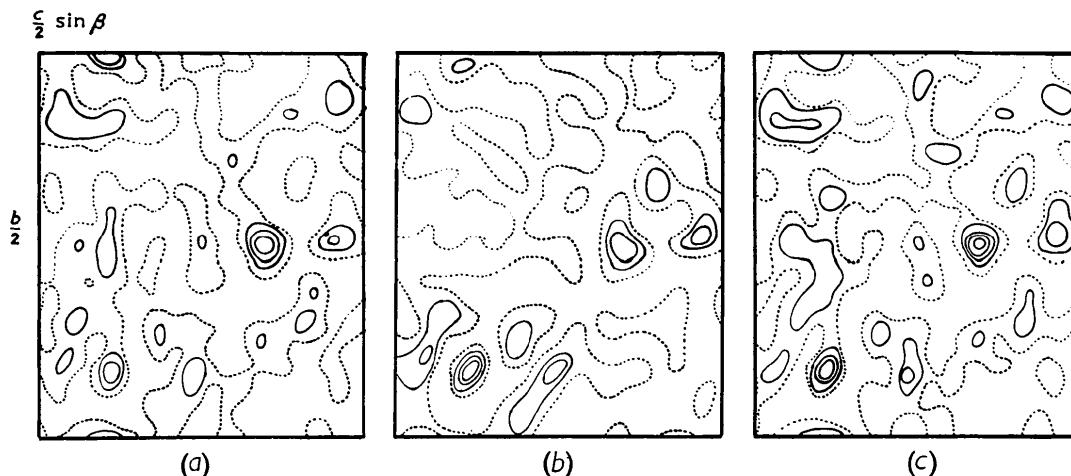


Fig. 2. (a) The section  $y=11/32$  from the synthesis calculated for haemoglobin using the structure factors  $\xi$  from equations (4) and (5), based on a single isomorphous pair. (b) The corresponding section calculated from a series of seven isomorphous derivatives by Perutz *et al.* (1960). (c) The same section from the haemoglobin synthesis using structure factors  $\xi'$  from (4) and (6).

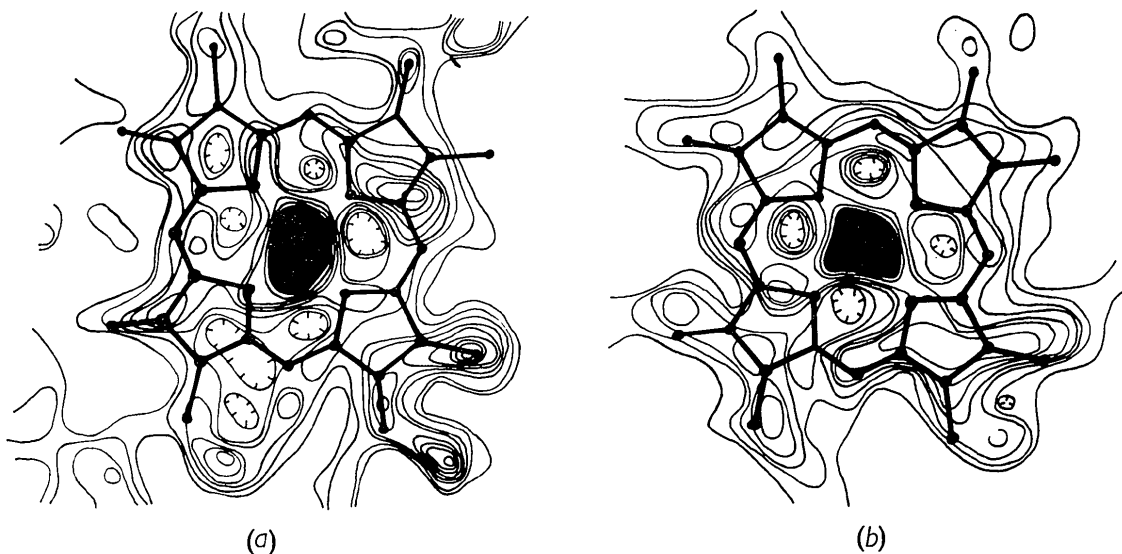


Fig. 3. (a) The haem section of the myoglobin molecule, calculated by using structure factors  $\mathfrak{F}'$  from equations (4) and (6), based on a single isomorphous pair and sharpened. (b) The same section, calculated from a series of five isomorphous derivatives by Kendrew *et al.* (1960), using the same sharpening-factor.

haemoglobin, and 20 e./unit cell for myoglobin. The integrations were done numerically by summation at  $10^\circ$  intervals of  $\alpha$ ,  $\alpha=0^\circ$  and  $\alpha=180^\circ$  being given half weight. In Fig. 2(c), the result for the same section of the haemoglobin Fourier synthesis is shown.

Fig. 3(a) shows a small section of the myoglobin synthesis which includes the porphyrin-ring system of the haem group calculated by the same method. In this calculation 9600 independent reflexions are included, going out to a spacing of 2 Å. The structure factors of the mercury diammine derivative ('HgAm<sub>2</sub>') of myoglobin were taken as  $F_1$ , and of the *p*-chloro mercuribenzene sulphonic acid derivative ('PCMS') as  $F_2$ . Thus  $f_{12}$  was the calculated structure factor of the PCMS mercury atom *minus* that of the HgAm<sub>2</sub> mercury atom. This procedure was needed to obtain a non-centrosymmetric arrangement of scattering matter for the replacing atoms in the  $P2_1$  unit cell. Although two heavy-atom derivatives have been used there is, of course, still only one isomorphous pair. The myoglobin structure factors were calculated according to equation (6) and sharpened by the factor  $\exp\{1.5(\sin\theta/\lambda)^2\}$ , the same as was used in the calculation of Kendrew *et al.* (1960). The same section calculated from their series of five isomorphous compounds, is shown in Fig. 3(b).

The 'fit' of the known shape of the haem group to the results from single isomorphous replacement was nearly as good as in the multiple isomorphous-replacement method. In the case of haemoglobin, the agreement was good enough to encourage us to plot out the whole three-dimensional synthesis. From the single isomorphous-replacement results it was possible to follow the two chains of high electron density, cor-

responding to the  $\alpha$ -helix backbone of the protein, throughout almost all their length. There were a few branch points, where the peaks from one chain coalesce with a neighbouring one, and one or two stretches where the image of the  $\alpha$ -helix faded near to the background level. Without prior knowledge of the structure, most of the  $\alpha$ -helix would have been correctly interpreted, but there would have been a few ambiguities which are adequately removed by the multiple isomorphous-replacement method. One puzzling feature was that the iron atoms at the centre of the haem groups never came much above the level of the polypeptide chains, while in the multiple isomorphous-replacement results at the same resolution they show about double the density.

### 5. The combination of the single isomorphous-replacement method with anomalous-dispersion data

Although this is as far as we have been able to take the single isomorphous-replacement technique, we realize that in practice there would be other important information available, which one ought not to neglect. This is the effect of anomalous scattering by the heavy replacing atoms. Bijvoet (1954) pointed out that this effect could be used to solve the ambiguity of the isomorphous-replacement method. The suggestion has been put into practice by several authors, notably Ramachandran & Raman (1956), who used anomalous scattering and a single isomorphous replacement to verify the structure of L-ephedrine hydrochloride.

Our primary concern is with the application of these techniques to protein structures, and experimental

accuracy is not yet sufficient for the marginally small changes due to anomalous scattering to be measured with a high degree of certainty. For this reason it is particularly important that proper weighting functions should be used. Although this question has been discussed by Blow (1958), a more simple and elegant way of calculating the weights has now been developed.

This method consists of taking the data for  $hkl$  and  $\bar{h}\bar{k}\bar{l}$  separately throughout, so that instead of the two intersecting circles of Fig. 2, there are now, in fact, three intersecting circles of radius  $F_1(hkl)$ ,  $F_2(hkl)$ ,  $F_2(\bar{h}\bar{k}\bar{l})$ . (If there are also anomalous effects in the  $F_1$ , these can be allowed for by treating  $F_1(hkl)$  and  $F_1(\bar{h}\bar{k}\bar{l})$  separately.)

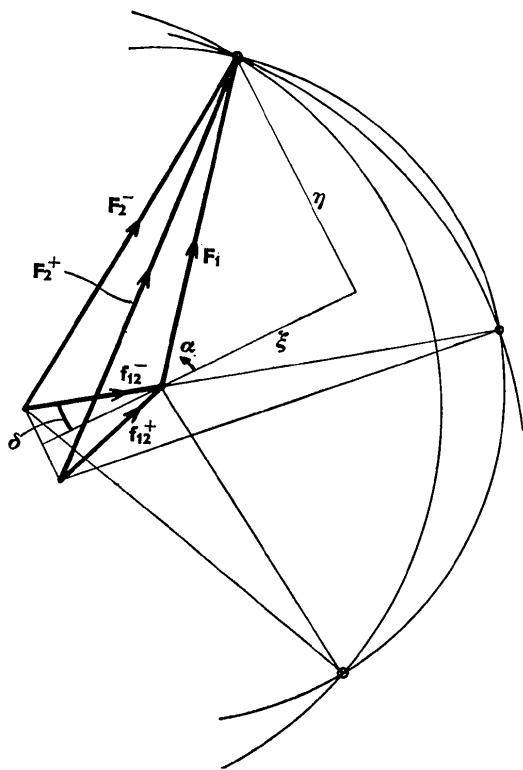


Fig. 4. The anomalous scattering effect generates two different triangles, and eliminates the ambiguity. The + superscript indicates structure factors for the reflexion  $hkl$ ; those with a - indicate the complex conjugates of structure factors for the reflexion  $\bar{h}\bar{k}\bar{l}$ .

Fig. 4 shows the data for  $hkl$  and  $\bar{h}\bar{k}\bar{l}$ , plotted in the manner of Fig. 2, the  $\bar{h}\bar{k}\bar{l}$  structure factors being plotted as the complex conjugates. For brevity, the superscripts + and - are used to refer to  $hkl$  and  $\bar{h}\bar{k}\bar{l}$ . We denote  $f_{12}(hkl)$  by  $f_{12}' + if_{12}''$  and  $f_{12}^*(\bar{h}\bar{k}\bar{l})$  by  $f_{12}' - if_{12}''$ . There are now two distinct triangles, giving two values of  $\varepsilon$ :

$$\varepsilon^+ = F_2^+ - [F_1^{+2} + |f_{12}|^2 + 2F_1^+|f_{12}| \cos(\alpha + \delta)]^{\frac{1}{2}}$$

and

$$\varepsilon^- = F_2^- - [F_1^{-2} + |f_{12}|^2 + 2F_1^-|f_{12}| \cos(\alpha - \delta)]^{\frac{1}{2}},$$

where  $\tan \delta = |f_{12}''|/|f_{12}'|$ . In this work we assumed  $|F_1^+| = |F_1^-|$ . In Fig. 4 the symmetry of the diagram about the line of  $f_{12}$  is now lost, and in addition to the term  $\xi$  parallel to this line, there is a contribution in the perpendicular direction. In place of (6) we have to use

$$|\xi''| + i|\eta''| = F_1 \frac{\int_0^{2\pi} e^{i\alpha} \exp\{-(\varepsilon^{+2} + \varepsilon^{-2})/2E^2\} d\alpha}{\int_0^{2\pi} \exp\{-(\varepsilon^{+2} + \varepsilon^{-2})/2E^2\} d\alpha} \quad (7)$$

(4) is replaced by

$$A = (|\xi''|a - |\eta''|b)/|f_{12}'|; \quad B = (|\xi''|b + |\eta''|a)/|f_{12}'|, \quad (8)$$

where

$$f_{12}' = a + ib.$$

A complete three-dimensional, 5.8 Å resolution Fourier synthesis of haemoglobin was calculated from the anomalous dispersion data of the '4HgCl<sub>2</sub>' compound together with the unsubstituted compound, according to expression (8). The '4HgCl<sub>2</sub>' compound contains roughly one mercury atom at each of the two independent sites per asymmetric unit. The synthesis was first calculated using anomalous-scattering data only from those reflexions where a 'significant' anomalous scattering effect was observed, namely if

$$|(F_2^+)^2 - (F_2^-)^2| > 0.05\{(F_2^+)^2 + (F_2^-)^2\}.$$

The 256 reflexions which gave such large anomalous effects produced a maximum of 5% change in the density. The calculation was therefore repeated using the procedure described above for every reflexion, whether or not the departure from Friedel's Law could be regarded as significant. The total effect of all these contributions produced a maximum change of 20% in the density. This is a striking example of the way in which a large number of contributions, each in itself of little significance, can result in important effects when combined as in a Fourier transform.

A model of the asymmetric unit (which contains two polypeptide chains) was built by cutting out the density shapes above a chosen level. The final model thus represents a solid whose density is higher than the surrounding space. This model is compared with a model built in a similar manner, using the results of seven isomorphous replacements, by Perutz *et al.* (1960), in Fig. 5.

Another problem needs to be discussed here. In the initial allocation of co-ordinates to the heavy atoms, an arbitrary choice has to be made about their absolute configuration, while the reciprocal lattice is conventionally indexed on a right-handed system. These two choices may or may not be consistent, and an inconsistent choice corresponds to reversing the sign of  $|\eta''|$  in (8). The proper way to proceed is to calculate Fourier syntheses from (7) in two parts. The first

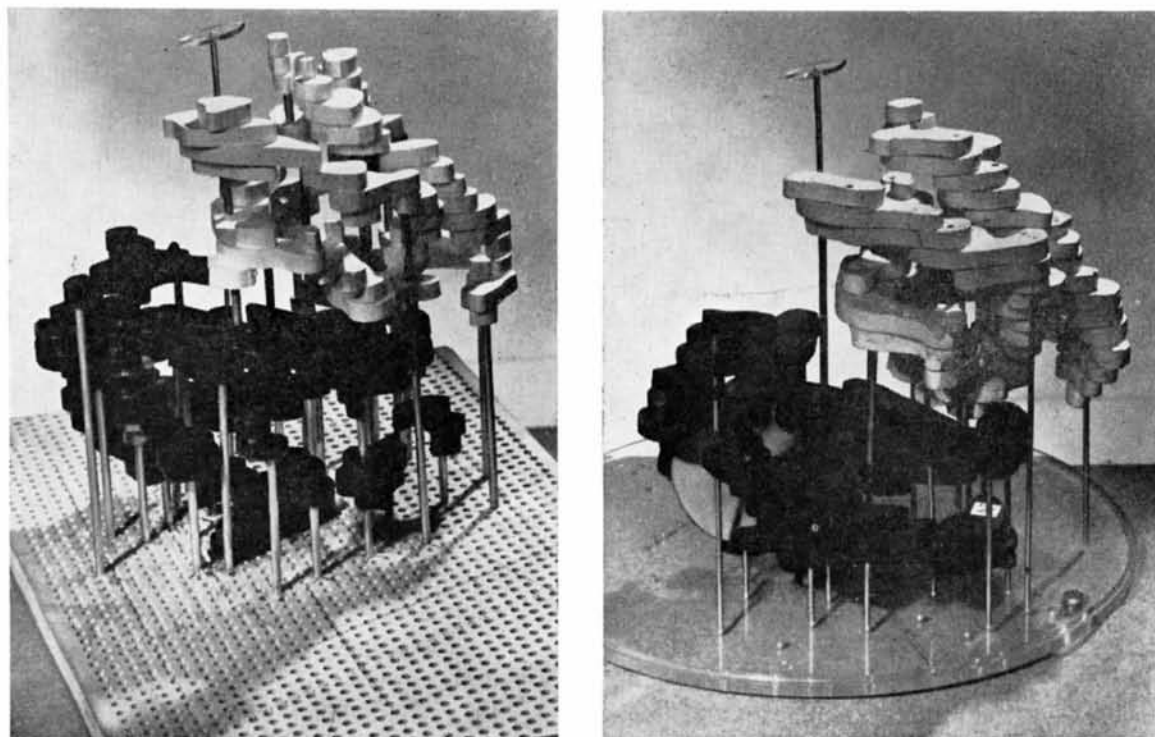


Fig. 5. (a) A model of one half haemoglobin molecule (the asymmetric unit) built from a Fourier synthesis calculated from equation (7) and (8) using a single isomorphous pair. (b) An analogous model from the data of Perutz *et al.* (1960) using a series of seven isomorphous compounds.

(corresponding to the single isomorphous-replacement method alone, and dependent only on the choice of the configuration of the heavy atoms) uses structure factors

$$\xi'' = |\xi''|(a + ib)/|f_{12}|;$$

the second corresponds to the anomalous-scattering method alone, using

$$\eta'' = |\eta''|(-b + ia)/|f_{12}|.$$

The comparative weakness of the anomalous-scattering method will normally make the  $\eta''$  Fourier weaker than the  $\xi''$ , but there should be a significant correlation between them, under the conditions in which the single isomorphous-replacement method is applicable. The correlation will be either positive or negative, and will indicate whether the syntheses need to be added or subtracted, in order to combine the two methods. A negative correlation will indicate that the wrong enantiomorph of the heavy atoms was chosen, and that the correct structure will be the mirror image of the synthesis obtained by using structure factors ( $\xi'' - \eta''$ ).

It may be noted that the  $\eta''$  synthesis is closely analogous to the straightforward single isomorphous-replacement synthesis, and is of use in the study of somewhat simpler heavy-atom compounds when no isomorphous compound is available. It can be used when the anomalous scatterers form a non-centro-

symmetric arrangement, as well as when they are centrosymmetrically arranged (Raman, 1959).

Although in the study of non-centrosymmetric structures by the isomorphous replacement method a series of three or more compounds is a desirable asset, it is clear that in principle there is no stringent requirement for more than two. A single isomorphous pair may be expected to give a considerable amount of useful information, provided the replacing atoms are heavy enough and arranged in a non-centrosymmetric manner.

The calculations described in this paper were made possible by the free access to their unpublished data granted us by Dr M. F. Perutz and Dr J. C. Kendrew. We also wish to thank them for allowing us to publish some of their results here, for comparison with our own. A great deal of computation was involved, for which we must thank the Unit computing staff, and the Director of the University Mathematical Laboratory for the use of EDSAC 2. Mrs Margaret Allen performed the formidable task of extracting the anomalous-scattering data from the original lists of photographic intensities.

We are grateful to Dr Aaron Klug for a stimulating discussion on the centro-symmetry problem, and Dr W. Cochran for looking at some of our theory and for drawing our attention to the Patterson function of the second kind. The concept of the 'contrast' was intro-

duced to us some years ago by Dr F. H. C. Crick. The referee made some very useful comments.

### References

- BIJVOET, J. M. (1954). *Nature, Lond.* **173**, 888.  
 BLOW, D. M. (1958). *Proc. Roy. Soc. A*, **247**, 302.  
 BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794.  
 BOKHOVEN, C., SCHOONE, J. C. & BIJVOET, J. M. (1951). *Acta Cryst.* **4**, 275.  
 BUERGER, M. J. (1951). *Acta Cryst.* **4**, 531.  
 CULLIS, A. F., MUIRHEAD, H., NORTH, A. C. T., PERUTZ, M. F. & ROSSMANN, M. G. (1961). *Proc. Roy. Soc. A*. (In press.)  
 KARTHA, G. & RAMACHANDRAN, G. N. (1955). *Acta Cryst.* **8**, 195.  
 KARTHA, G. (1961). *Acta Cryst.* **14**, 680.  
 KENDREW, J. C., DICKERSON, R. E., STRANDBERG, B. E., HART, R. G., DAVIES, D. R., PHILLIPS, D. C. & SHORE, V. C. (1960). *Nature, Lond.* **185**, 422.  
 OKAYA, Y., SAITO, Y. & PEPINSKY, R. (1955). *Phys. Rev.* **98**, 1857.  
 PATERSON, A. L. (1949). *Acta Cryst.* **2**, 318.  
 PERUTZ, M. F., ROSSMANN, M. G., CULLIS, A. F., MUIRHEAD, H., WILL, G. & NORTH, A. C. T. (1960). *Nature, Lond.* **185**, 416.  
 RAMACHANDRAN, G. N. & RAMAN, S. (1956). *Current Science*, **25**, 348.  
 RAMACHANDRAN, G. N. & RAMAN, S. (1959). *Acta Cryst.* **12**, 957.  
 RAMAN, S. (1959). *Z. Kristallogr.* **111**, 301.  
 ROGERS, D. (1951). *Research*, **4**, 296.  
 WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.

## Short Communications

*Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible. Publication will be quicker if the contributions are without illustrations.*

*Acta Cryst.* (1961). **14**, 1202

### Donnees cristallographiques sur l'antraquinone 1-4 et sur quelques derives substitues.

Par Mm. M. ALLEAUME, R. DARROUY et J. HOUSTY, *Laboratoire de Minéralogie et de Rayons X, Faculté des Sciences de Bordeaux, France*

(Reçu le 21 avril 1961)

#### Anthraquinone 1-4 C<sub>14</sub>H<sub>8</sub>O<sub>2</sub>

L'antraquinone 1-4 se présente sous deux formes suivant le solvant de cristallisation.

Par mise en solution dans le benzène, on obtient de belles aiguilles rouges, allongées suivant [001].

Les paramètres de la maille sont les suivants:

$$a = 13,83, b = 9,65, c = 7,35 \text{ \AA}; \beta = 96^\circ.$$

Ce cristal est de symétrie monoclinique, et la maille contient 4 molécules. Densité calculée  $d = 1,41 \text{ g.cm.}^{-3}$ . Groupe spatial  $P2_1/m$ .

Par mise en solution dans l'acétate d'éthyle, on obtient de fines aiguilles jaunes, allongées suivant [010].

La maille monoclinique a pour paramètres:

$$a = 8,40 \pm 0,01, b = 5,93 \pm 0,01, c = 19,82 \pm 0,02 \text{ \AA}; \\ \beta = 99^\circ \pm 30'.$$

Nombre de molécules par maille: 4.  
 Densité calculée:  $1,52 \text{ g.cm.}^{-3}$ .  
 Groupe spatial:  $P2_1/c$ .

#### Chloro 2-antraquinone 1-4

La Chloro 2-antraquinone 1-4 cristallise dans le système orthorhombique sous forme de plaquettes jaunes allongées suivant la direction [001].

La maille cristalline est caractérisée par les paramètres suivants:

$$a = 21,74 \pm 0,05, b = 5,80 \pm 0,02, c = 8,74 \pm 0,02 \text{ \AA}.$$

Densité calculée:  $1,32 \text{ g.cm.}^{-3}$ .  
 Nombre de molécules dans la maille: 4.  
 Groupe spatial:  $P2_12_12_1$  ou  $P2_12_12$ .

#### Dichloro 2-3 anthraquinone 1-4

Ce composé se présente sous forme de plaquettes jaunes allongées suivant la direction [010].

La maille monoclinique possède les paramètres suivants:

$$a = 22,49 \pm 0,05, b = 8,68 \pm 0,02, c = 5,88 \pm 0,02 \text{ \AA}; \\ \beta = 94^\circ \pm 1^\circ.$$

Densité calculée:  $1,36 \text{ g.cm.}^{-3}$ .  
 Nombre de molécules par maille: 4.  
 Groupe spatial:  $P2_1/c$  ou  $P2/c$ .

#### Dibromo 2-3 anthraquinone 1-4

La dibromo 2-3 anthraquinone 1-4 cristallise dans le système monoclinique sous forme de plaquettes brunes.

Paramètres cristallins:

$$a = 20,50 \pm 0,05, b = 5,76 \pm 0,02, c = 9,48 \pm 0,03 \text{ \AA}; \\ \beta = 92^\circ \pm 1^\circ.$$

Densité calculée:  $1,56 \text{ g.cm.}^{-3}$ .  
 Nombre de molécules dans la maille: 4.  
 Groupe spatial:  $P2_1/c$ .

#### Donnees cristallographiques sur la phenanthrene quinone 9-10

Cristallise sous forme de prismes orangés allongés suivant la direction [010].

Système cristallin: monoclinique.  
 Paramètre de la maille:

$$a = 12,60 \pm 0,03, b = 10,44 \pm 0,02, c = 14,20 \pm 0,03 \text{ \AA}; \\ \beta = 92^\circ \pm 1^\circ.$$

Densité calculée:  $1,47 \text{ g.cm.}^{-3}$ .  
 Nombre de molécules par maille: 8.  
 Groupe spatial:  $C2/c$ .