

is 1.11 Å and the average bridge B–H distance is 1.27 Å. These results are not accurate enough to confirm or refute the asymmetry of the bridge hydrogen atoms.

The present results establish that simple substitutional derivatives of decaborane are formed which retain the essential symmetry of the isolated decaborane molecule, and confirm the existence of a 1-ethyldecaborane which has been suggested by Blay, Dunstan & Williams (1960) on the basis of chemical and nuclear magnetic resonance evidence.

The author wishes to thank Dr D. Appleman of the U.S. Geological Survey for the computation of the Lorentz and polarization factors, R. D. Mason for assistance with intensity measurements, Dr S. Block for his interest and advice during the course of the investigations, and G. Ross for supplying the sample.

Acta Cryst. (1964). **17**, 338

The Relative Positions of Independent Molecules Within the Same Asymmetric Unit

BY MICHAEL G. ROSSMANN AND D. M. BLOW

M. R. C. Laboratory of Molecular Biology, Hills Road, Cambridge, England

AND MARJORIE M. HARDING* AND ELEANOR COLLER

Chemical Crystallography Laboratory, South Parks Road, Oxford, England

(Received 14 March 1963)

If \mathbf{x} and \mathbf{x}' are position vectors of equivalent points in identical molecules in different parts of the crystallographic asymmetric unit, then the linear relationship between these points may be written as $\mathbf{x}' = [\mathbf{C}]\mathbf{x} + \mathbf{d}$, when $[\mathbf{C}]$ is a rotation matrix and \mathbf{d} a translation vector. A function is derived for determining \mathbf{d} , given $[\mathbf{C}]$, when no knowledge of the phases is available.

Consider a crystal structure which contains two or more molecules, or other identical distributions of electron density, within the crystallographic asymmetric unit. Such distributions of density will be referred to as sub-units. In a previous paper we have shown how the angular relationship between the sub-units may be derived from the Patterson function (Rossmann & Blow, 1962). This depended on the idea that the Patterson vectors within one sub-unit (the 'self-vectors') formed a similar distribution for each sub-unit; so that a rotation can be found which brings the self-vectors from one sub-unit into coincidence with those from the other.

In this paper we shall consider how the 'cross-vectors', or Patterson vectors from one sub-unit to another, can be used to determine the translation required to bring one sub-unit into coincidence with the other, after suitable rotation about a given axis. In other papers (Hodgkin, Harding, Coller & Rossmann, 1964; Blow, Rossmann & Jeffery, 1964) we describe the application of this method to the proteins insulin and chymotrypsin. In a further paper (Rossmann & Blow, 1963) we have suggested how the condition that the sub-units must be identical, once their relative orientation and position are known, may be used to determine phases, thus leading to a solution of the structure.

Although the result of a rotation and translation in three dimensions depends on the position of the

References

- BERGHUIS, J., HAANAPPEL, IJ. M., POTTERS, M., LOOPSTRA, B. O., MACGILLAVRY, C. H. & VEENENDAAL, A. L. (1955). *Acta Cryst.* **8**, 478.
 BLAY, N. J., DUNSTAN, I. & WILLIAMS, R. L. (1960). *J. Chem. Soc.* p. 430.
 BUSING, W. R. & LEVY, H. A. (1959a). *A Crystallographic Least Squares Refinement Program for the IBM 704*. Oak Ridge National Laboratory Report 59-4-37.
 BUSING, W. R. & LEVY, H. A. (1959b). *A Crystallographic Function and Error Program for the IBM 704*. Oak Ridge National Laboratory Report 59-12-3.
 JAMES, R. W. & BRINDLEY, G. W. (1931). *Phil. Mag.* **12**, 81.
 KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
 KASPER, J. S., LUCHT, C. M. & HARKER, D. (1950). *Acta Cryst.* **3**, 436.
 McWEENY, R. (1951). *Acta Cryst.* **4**, 513.
 MOORE, E. B., JR., DICKERSON, R. E. & LIPSCOMB, W. N. (1957). *J. Chem. Phys.* **27**, 209.
 NORDMAN, C. E. & LIPSCOMB, W. N. (1953). *J. Chem. Phys.* **21**, 1856.
 POST, B., SCHWARTZ, R. S. & FANKUCHEN, I. (1951). *Rev. Sci. Instrum.* **22**, 218.

* Present address: Department of Chemistry, University of Edinburgh, West Mains Road, Edinburgh 9, Scotland.

rotation axis, the component of translation in a direction parallel to the axis of rotation is independent

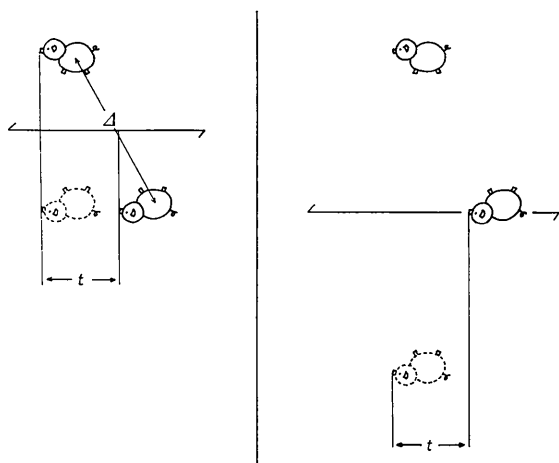


Fig. 1. The position of the twofold rotation axis which relates the two piglets is completely arbitrary. The diagram on the left shows the situation when the translation is parallel to the rotation axis. The diagram on the right has an additional component of translation perpendicular to the rotation axis, but the component parallel to the axis remains unchanged.

of this position (Fig. 1). It is this quantity which can be derived accurately by the method described below. The effect of space group symmetry is to allow a number of such translations to be determined, corresponding to choosing different pairs of sub-units related by the same rotation. Finally the localization in the Patterson projection of cross-vectors related by a particular rotation gives a less accurate indication of the positions of the sub-units in the unit cell.

Relationship between cross-vectors

The relation between two identical arrays of atoms, A and B , may be expressed by a matrix $[C]$ corresponding to a rotation about the origin of coordinates, and a translation vector d , where

$$\mathbf{x}_i^B = [C]\mathbf{x}_i^A + \mathbf{d} \quad (1)$$

for the arrays of N atoms, $i=1, \dots, N$. In a crystal, if $[C]$ does not correspond to a crystallographic rotation, the relation (1) can only apply to a localized group of atoms in the crystal. We define a *sub-unit* A as the array of N atoms at positions \mathbf{x}_i^A which are related to identical atoms at \mathbf{x}_i^B by (1).

It is convenient to choose a reference point S near the centre of each sub-unit, as an origin of a local system of coordinates X , so that

$$\mathbf{X}_i^A = \mathbf{x}_i^A - S^A \quad (2)$$

and

$$|\mathbf{X}_i^A| < r \quad (i=1, \dots, N).$$

(The purpose of choosing S near the centre of the

sub-unit is that r shall be as small as possible.) The reference points are to be equivalent, so that

$$S^B = [C]S^A + \mathbf{d}. \quad (3)$$

It is convenient to define a translation independently of the choice of origin. This requirement is fulfilled by the vector

$$\Delta^{AB} = S^B - S^A. \quad (4)$$

This is related to \mathbf{d} by

$$\Delta^{AB} = ([C] - [I])S^A + \mathbf{d}$$

where I is the identity matrix.

We shall now consider vectors between one sub-unit and another, which may be called 'cross-vectors', and we shall assume that $[C]$ is known (Rossmann & Blow, 1962). Every vector \mathbf{v} from \mathbf{x}_i^A to \mathbf{x}_j^B is related to a vector \mathbf{v}' from \mathbf{x}_i^B to \mathbf{x}_j^A . To the extent that we can demonstrate a relation between them, independent of X_i and X_j , the existence of any cross-vector \mathbf{v} implies the existence of another cross-vector \mathbf{v}' , determined only by $[C]$ and Δ^{AB} . We propose to determine the translation Δ^{AB} by searching the Patterson distribution for appropriately related arrays of vectors.

Let

$$\mathbf{v} = \mathbf{x}_j^B - \mathbf{x}_i^A = [C]\mathbf{x}_j^A + \mathbf{d} - \mathbf{x}_i^A = [C]\mathbf{X}_j^A + \Delta^{AB} - \mathbf{X}_i^A$$

by (1) and (2), and similarly let

$$\mathbf{v}' = \mathbf{x}_j^A - \mathbf{x}_i^B = \mathbf{x}_j^A - [C]\mathbf{x}_i^B - \mathbf{d} = \mathbf{X}_j^A - [C]\mathbf{X}_i^A - \Delta^{AB}.$$

Then, eliminating \mathbf{X}_i^A between the expressions for \mathbf{v} and \mathbf{v}'

$$\mathbf{v}' = [C]\mathbf{v} - ([I] + [C])\Delta^{AB} - ([C]^2 - [I])\mathbf{X}_j^A. \quad (5)$$

Rotation of 180°

In three cases already studied we have found a rotation of 180° relating protein sub-units (Rossmann & Blow, 1962; Hodgkin *et al.*, 1964; Blow *et al.*, 1964). Since this appears to be a common mode of association of sub-units, it is an important special case.

If $[C]$ represents a rotation of 180°, $[C]^2 = [I]$ and (5) becomes

$$\mathbf{v}' = [C]\mathbf{v} - ([I] + [C])\Delta^{AB} \quad (6)$$

which is independent of the atomic positions \mathbf{X}_j . This means the cross-vector \mathbf{v} can always be superimposed on \mathbf{v}' by the rotation and translation given by (6). It is possible to define a function $T(\Delta)$ which will have a large value only if Δ has the value given by (6).

$$T(\Delta) = \int_{|\mathbf{v}-\Delta| < 2r} P(\mathbf{v}) \cdot P(\mathbf{v}') d\mathbf{v}'. \quad (7)$$

Here $P(\mathbf{v})$ represents the Patterson function, and \mathbf{v} and \mathbf{v}' are related by (6). The volume of integration

will include all cross-vectors related by (6). (The integration (7) is over a sphere; if more detailed information about molecular shape were available any other suitable shape could be used.)

The expression (7) can readily be expanded in a form suitable for direct calculation from the observed intensities. Putting $\mathbf{u} = \mathbf{v} - \Delta$, one finds

$$T(\Delta) = \int_{|\mathbf{u}| < 2r} P(\mathbf{u} + \Delta) P([\mathbf{C}]\mathbf{u} - \Delta) d\mathbf{u}.$$

This form shows that the translation function depends on comparing a region of the rotated Patterson function around $-\Delta$ with a region of the unrotated Patterson function around $+\Delta$. Expanding in Fourier series

$$\begin{aligned} T(\Delta) &= \int_{|\mathbf{u}| < 2r} \left[\sum_{\mathbf{h}} F_{\mathbf{h}}^2 \exp \{2\pi i(\mathbf{h} \cdot \mathbf{u} + \Delta)\} \right] \\ &\quad \times \left[\sum_{\mathbf{p}} F_{\mathbf{p}}^2 \exp \{2\pi i\mathbf{p}([\mathbf{C}]\mathbf{u} - \Delta)\} \right] d\mathbf{u} \\ &= \sum_{\mathbf{h}} \sum_{\mathbf{p}} F_{\mathbf{h}}^2 F_{\mathbf{p}}^2 \exp \{2\pi i(\mathbf{h} - \mathbf{p}) \cdot \Delta\} \\ &\quad \times \int_{|\mathbf{u}| < 2r} \exp \{(\mathbf{h} + \mathbf{p}[\mathbf{C}]) \cdot \mathbf{u}\} d\mathbf{u}. \end{aligned}$$

The integral has been discussed by Rossmann & Blow (1962), where it is written as $(U/V)G_{\mathbf{h}\mathbf{p}} \exp(i\Omega_{\mathbf{h}\mathbf{p}})$. In the present case $\Omega_{\mathbf{h}\mathbf{p}} = 0$, as the volume of integration is symmetrical about $\mathbf{u} = 0$. Using the Friedel relationships this gives

$$T(\Delta) = (U/V) \sum_{\mathbf{h}} \sum_{\mathbf{p}} F_{\mathbf{h}}^2 F_{\mathbf{p}}^2 G_{\mathbf{h}\mathbf{p}} \cos 2\pi(\mathbf{h} - \mathbf{p}) \cdot \Delta. \quad (8)$$

This may be recognized as a centrosymmetric Fourier summation with coefficients

$$\left\{ \sum_{\mathbf{p}} F_{\mathbf{H}-\mathbf{p}}^2 F_{\mathbf{p}}^2 G_{\mathbf{H}-\mathbf{p}, \mathbf{p}} \right\}$$

for the term with indices $\mathbf{H} = \mathbf{h} - \mathbf{p}$.

Precise and imprecise parameters

Whereas \mathbf{d} in equation (1) has a value as precise as the values of atomic coordinates, the 'centre' of a sub-unit is a much less precise concept. As shown in Fig. 1, Δ^{AB} has a precise component, \mathbf{t} , parallel to the rotation axis and an imprecise one perpendicular to it; the component \mathbf{t} is always the same, no matter where the rotation axis is placed. The vector $([\mathbf{I}] + [\mathbf{C}])\Delta^{AB}$ in equation (6) is parallel to the rotation axis, and is equal to $2\mathbf{t}$. Thus the vector \mathbf{v}' is superimposed on \mathbf{v} by rotation of the Patterson, followed by translation along the axial direction. The only effect of the component of Δ perpendicular to the rotation axis is to vary the region of the Patterson over which the integral (7) is taken. $T(\Delta)$ will be relatively insensitive to such variation, since the superposable vectors lie in a region $2r$ in diameter.

It is therefore convenient to consider two com-

ponents of Δ . Any peak in the translation function $T(\Delta)$ precisely determines one parameter, \mathbf{t} , representing the translation parallel to the rotation axis between two sub-units. At the same time it determines in a relatively imprecise way the perpendicular (or sideways) component, \mathbf{s} , of the vector separating the sub-units. This is illustrated in two dimensions in Figs. 2, 3 and 4. Fig. 2 shows a postulated structure; Fig. 3 its Patterson function and Figs. 4(a), (b) and (c) show various attempts at matching up the cross

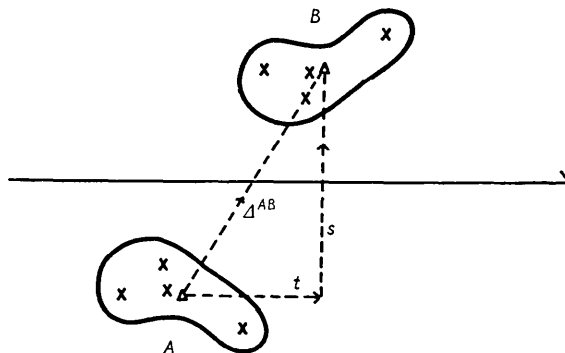


Fig. 2. The crosses represent atoms in a two-dimensional model structure. The triangles are the points chosen as approximate centres of molecules *A* and *B*. Δ^{AB} has components \mathbf{t} and \mathbf{s} parallel and perpendicular to the screw rotation axis, respectively.

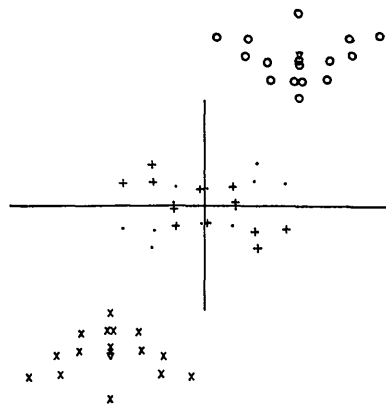


Fig. 3. The vectors arising from the structure in Fig. 2. The self-vectors of molecules *A* and *B* are represented by + and ·; the cross-vectors from molecules *A* to *B* and *B* to *A* by x and o. Triangles mark the position of $+\Delta^{AB}$ and $-\Delta^{AB}$.

Patterson vectors. In Fig. 4(a) the Patterson within the circle centred at $-\Delta^{AB}$ has been rotated about the known rotation axis through the origin and translated by $2\mathbf{t}$ parallel to this axis. This gives the best match within the circle. A translation other than $2\mathbf{t}$ produces no overlap at all (Fig. 4(b)). If the circle is placed at a centre Δ with the same component \mathbf{t} parallel to the rotation axis, but an altered perpendicular component, not all the vector coincidences

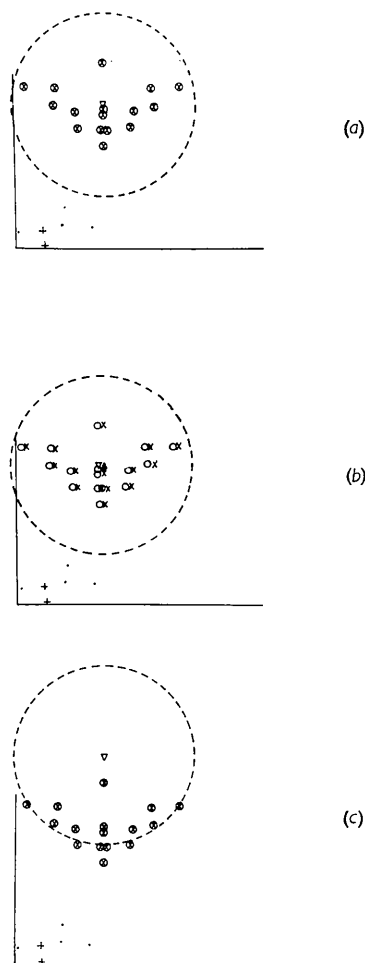


Fig. 4. (a) Superposition of vectors around $+\Delta^{AB}$ and $-\Delta^{AB}$ within a sphere of suitable radius, after rotating the latter about the rotation axis through the Patterson origin and translating by $2\mathbf{t}$ parallel to the axis. (b) Similar superposition with a wrong choice of the precise parameter \mathbf{t} producing no significant vector coincidences. (c) Similar superposition but with the imprecise parameter \mathbf{s} chosen badly, producing good vector coincidences although some fall outside the sphere of integration.

will lie within the circle, and a reduced value of $T(\Delta)$ will result (Fig. 4(c)).

General rotation

In the case of a 180° rotation $([\mathbf{C}]^2 - [\mathbf{I}]) = 0$, which gives (5) a simple form. A similar simplification may be made in the case of n sub-units related by rotations of $2\pi/n$, by considering three of these sub-units A, B, C , and letting $\mathbf{v} = \mathbf{x}_j^B - \mathbf{x}_i^A$ while $\mathbf{v}' = \mathbf{x}_i^B - \mathbf{x}_j^C$. For a completely general rotation there is no way of eliminating \mathbf{X}_j^A from (5). This means that no rotation and translation of the Patterson function can lead to exact superposition of all cross-vectors between the two sub-units. Nevertheless, especially when a pair

of sub-units is related by rotation through an angle near to 0° or 180° , all cross-vectors can be brought close to superposition.

If $[\mathbf{C}]$ represents a rotation κ about some axis, $[\mathbf{C}]^2$ represents a rotation 2κ about the same axis and $([\mathbf{C}]^2 - [\mathbf{I}])\mathbf{X}_j$ is bound to lie in a plane perpendicular to the axis of rotation. Bearing in mind that $|\mathbf{X}_j^A| < r$, it follows that $([\mathbf{C}]^2 - [\mathbf{I}])\mathbf{X}_j^A$ must lie on a disc, centred at the origin of coordinates, perpendicular to the axis of rotation, of radius $|2r \sin \kappa|$. Equation (5) thus indicates a rotation and translation of the vector \mathbf{v}' which will cause the vector between it and \mathbf{v} to lie on such a disc. In this case the 'translation function' must evaluate the correlation between the Patterson density at \mathbf{v} and that in the disc of possible positions for \mathbf{v}' :

$$T(\Delta) = \int_{|\mathbf{u}| < 2r} P(\mathbf{v}) \left\{ \int_{|\mathbf{X}_j| < r} P(\mathbf{v}') d\mathbf{X}_j \right\} d\mathbf{u},$$

where \mathbf{v}' is related to \mathbf{v} and \mathbf{X}_j by (5), and $\mathbf{u} = \mathbf{v} - \Delta$ as before. Making these substitutions

$$T(\Delta) = \int_{|\mathbf{u}| < 2r} P(\mathbf{u} + \Delta) \times \left\{ \int_{|\mathbf{X}| < r} P[\mathbf{C}]\mathbf{u} - \Delta - ([\mathbf{C}]^2 - [\mathbf{I}])\mathbf{X} d\mathbf{X} \right\} d\mathbf{u}.$$

The expansion in Fourier series yields

$$T(\Delta) = \sum_{\mathbf{h}} \sum_{\mathbf{p}} F_{\mathbf{h}}^2 F_{\mathbf{p}}^2 G_{\mathbf{h}\mathbf{p}} w_{\mathbf{p}} \cos 2\pi(\mathbf{h} - \mathbf{p}) \cdot \Delta, \quad (9)$$

where

$$w_{\mathbf{p}} = \int_{|\mathbf{X}| < r} \exp \{ -2\pi i \mathbf{p} \cdot ([\mathbf{C}]^2 - [\mathbf{I}])\mathbf{X} \} d\mathbf{X}$$

is a weighting function which gives full weight to reflexions for which the vector \mathbf{p} lies along the rotation axis, and reduced weight elsewhere. It is, in fact, the Fourier transform of a weighted disc, orientated with its normal along the axis of $[\mathbf{C}]$, and is similar to the function $(1/2\pi r R) J_1(2\pi r R)$ which is the transform of a uniform disc.

The summation (9) is seen to be a centrosymmetric Fourier summation with coefficients

$$\left\{ \sum_{\mathbf{p}} F_{\mathbf{H}-\mathbf{p}}^2 F_{\mathbf{p}}^2 w_{\mathbf{p}} G_{\mathbf{H}-\mathbf{p}, \mathbf{p}} \right\}$$

for the term with indices $\mathbf{H} = \mathbf{h} - \mathbf{p}$.

Effect of space group symmetry

So far, only an isolated pair of sub-units has been considered. In space group $P1$ the choice of origin is entirely arbitrary, and may be taken as the position of \mathbf{S}^A , so that $\mathbf{S}^A = 0$, $\mathbf{S}^B = \Delta^{AB}$.

In all cases of higher symmetry, the positions of crystallographic symmetry elements provide reference points in the unit cell from which the positions of sub-units must be measured. In these cases there are a number of distinct ways of choosing pairs of sub-

units, so that each of these pairs will be associated with a different translation vector Δ . The relation between these values of Δ will usually allow the positions of the sub-units to be related to those of the symmetry elements.

A special difficulty arises in some cases where the non-crystallographic rotation [C] is about an axis perpendicular to a crystallographic symmetry axis, because the sub-unit as defined may become infinite in its extent in the direction parallel to the crystallographic axis. These points are illustrated in detail in other papers (Hodgkin *et al.*, 1964; Blow *et al.*, 1964), which describe application of the translation function to space groups $R3$ and $P2_1$.

Removal of self-vectors

It is important to remember that the self-vectors can be made to superimpose by the rotation [C], with any lattice translation (including zero) giving rise to large origin peaks in $T(\Delta)$. These peaks can easily swamp a peak due to superposition of cross-vectors, especially when the translation Δ is small. It is therefore important to remove the self-vectors.

We have attempted to do this by modifying the Patterson coefficients so as to produce a region of zero density within a volume U' around each origin. The modifying function

$$M(\mathbf{u}) = \sum_{\mathbf{p}} m_{\mathbf{p}} \exp \{2\pi i \mathbf{p} \cdot \mathbf{u}\} = 0 \text{ if } \mathbf{u} \text{ is within } U' \\ = 1 \text{ if } \mathbf{u} \text{ is outside } U'.$$

Then the modified Patterson function

$$P'(\mathbf{u}) = P(\mathbf{u}) \cdot M(\mathbf{u}) = \sum_{\mathbf{h}} Q_{\mathbf{h}} \exp \{2\pi i \mathbf{h} \cdot \mathbf{x}\}$$

where

$$Q_{\mathbf{h}} = \sum_{\mathbf{p}} F_{\mathbf{h}-\mathbf{p}}^2 m_{\mathbf{p}}$$

by the convolution theorem. From the definition of $M(\mathbf{u})$ it follows that

$$m_{\mathbf{p}} = \frac{1}{V} \left[\int_V \exp \{-2\pi i \mathbf{p} \cdot \mathbf{u}\} d\mathbf{u} \right. \\ \left. - \int_{U'} \exp \{-2\pi i \mathbf{p} \cdot \mathbf{u}\} d\mathbf{u} \right],$$

where V is the volume of the unit cell. Hence

$$m_0 = 1 - U'/V$$

and

$$m_{\mathbf{p}} = -(U/V)G_{\mathbf{p},0} \text{ (where } \mathbf{p} \neq 0 \text{),}$$

$G_{\mathbf{p},0}$ having the same significance as before. Hence the modified coefficients are

$$Q_{\mathbf{h}} = F_{\mathbf{h}}^2 \left(1 - \frac{U'}{V}\right) - \frac{U'}{V} \sum_{\mathbf{p} \neq \mathbf{h}} F_{\mathbf{h}-\mathbf{p}}^2 G_{\mathbf{p},0}. \quad (10)$$

If there is overlap between the volumes U' centred on neighbouring origins, the effective value of U'/V in (10) is altered. In this case an empirical value was calculated by ensuring that the modified Patterson density is zero at the origin. If k is the effective value of U'/V

$$(1-k) \sum_{\mathbf{h}} F_{\mathbf{h}}^2 - k \sum_{\mathbf{h}} \sum_{\mathbf{p} \neq \mathbf{h}} F_{\mathbf{h}-\mathbf{p}}^2 G_{\mathbf{p},0} = 0.$$

k seldom differed by more than 5% from U'/V , once all mistakes concerning the multiplicities had been eliminated from the program.

We thank Prof. Dorothy Hodgkin for constant guidance. We are indebted to the University of Cambridge Mathematical Laboratory for their assistance in running our work on the EDSAC 2 computer, and to IBM United Kingdom Ltd. for a gift of free time on the IBM 7090. We also wish to thank Miss Jill Collard and Miss Angela Campbell for assistance in all stages of the work. The program ERFR2 by W. G. Sly, D. P. Shoemaker & J. H. Van den Hende was used for some of the Fourier syntheses.

References

- BLOW, D. M., ROSSMANN, M. G. & JEFFERY, B. A. (1964). *J. Mol. Biol.* **8**, 65.
 HODGKIN, D., HARDING, M., COLLIER, E. & ROSSMANN, M. G. (1964). In preparation.
 ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24.
 ROSSMANN, M. G. & BLOW, D. M. (1963). *Acta Cryst.* **16**, 39.