Acta Cryst. (1971). B27, 1780

Multiple Heavy-Atom Sites in Protein Crystals having Centrosymmetric Projections: Interpretation of Vector Maps. I. Modification of the Three-Dimensional Patterson Difference Synthesis.

BY P. C. MOEWS AND C. W. BUNN

The Royal Institution of Great Britain, Davy-Faraday Laboratory, London, England

(*Received* 26 *May* 1970 *and in revised form* 13 *November* 1970)

The three-dimensional Patterson difference synthesis, which is used for locating heavy atoms in protein crystals, gives a vector distribution which is imperfect because for most X-ray reflexions $|F_H|$, the structure factor for the heavy atoms, is not equal to the experimental quantity $||F_{PH}|-|F_P||$, the difference between the structure factors for the heavy-atom-containing crystal and the native crystal. The quantities $||F_{PH}|-|F_H||$ are, for the acentric reflexions, too low by a factor depending on the difference of phase between F_{PH} and F_H; the average value of this factor is 0·64 (average value of cos α). On the other hand, for centric reflexions $||F_{PH}|-|F_P||=|F_H|$ (except for a few weak protein reflexions). By weighting the acentric reflexions upwards by the factor 1/0·64, serious defects of the vector distribution are removed.

Many different heavy-atom-containing substances form adducts with rennin crystals, but in all adducts studied so far, the heavy atoms are in a number of sites, so that it is difficult to locate them by established methods (Bunn, Camerman, Liang Tung-T'sai, Moews & Baumber, 1970). Even the space group is difficult to establish, because $I222$ and $I2_12_12_1$ cannot be distinguished by systematic absences of reflexions or by intensity statistics, and distinction by vector maps (Rogers, 1950) is only straightforward for very simple (preferably single-site) heavy-atom adducts.

Orthorhombic fourfold multiplicity entails a steep increase in the number of heavy-atom vectors with increasing number of sites: for 4 sites there are 36 vectors per lattice point, for 6 sites 78 vectors. In vector maps for the centrosymmetric projections, given by the Patterson difference synthesis, overlapping of peaks in ridges and plateaux leads to ambiguity of interpretation, so that several different interpretations of almost equal merit are possible.

Two different ways of dealing with these difficulties have been tried. One was to formulate the several alternative interpretations of the principal projections of several different adducts; interpretations leading to the same set of protein signs are likely to be the correct ones. This process, repeated for at least two but preferably all three principal projections, would reveal the space group and locate the heavy-atom sites. The second was to try to interpret the results of the three-dimensional Patterson difference synthesis with coefficients $(|F_{PH}|-|F_P|)^2$. The limitation here is the familiar one (Phillips, 1966) that for most of the reflexions (all except those for the principal zones) the phase angles of F_{PH} (for the heavy-atom-containing crystal) and F_P (for the protein crystal) can have any value, and the difference of phase angle between F_{PH} and F_P means that $||F_{PH}|-|F_P||$ does not equal $|F_H|$ (Fig. 1); it is less than $|F_H|$ by an unknown amount which is different for each reflexion (depending on α, the difference between protein and heavy-atom phase angles), and the three-dimensional vector map is inevitably falsified. The heavy-atom vectors are expected to appear at about half weight and there are also spurious features (Phillips, 1966). [If there are appreciable differences between $|F_{hkl}|$ and $|F_{\bar{h}\bar{k}\bar{l}}|$ due to anomalous scattering, the true values of $|F_H|$ and $|\alpha|$ may be calculated (Singh & Ramaseshan, 1966), but such differences are often too small to be useful.]

Neither of these methods, in its original form, was successful with rennin adducts; therefore, new methods or modifications of those mentioned have been sought. The present paper describes a modification of the three-dimensional Patterson difference synthesis, applicable to structures having one or more centrosymmetric projections, since in these circumstances some of the spurious features are systematic and can be minimized by utilizing the different properties of centric and acentric reflexions.

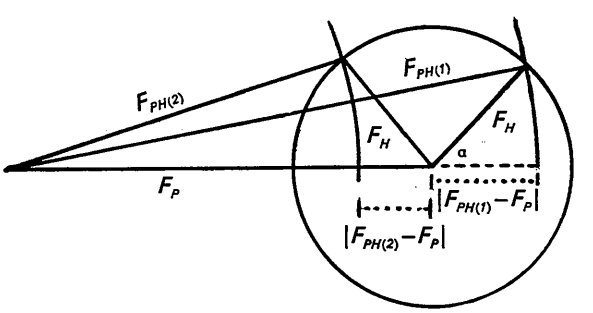

Fig. 1. Phase diagram. Owing to the phase differences α between F_P and F_{PH}, $||F_{PH}|-|F_P||$ does not equal $|F_H|$; it is approximately equal to $|F_H|\cos\alpha$ when $|F_H|$ is small compared with $|F_{PH}|$.

Spurious features in three-dimensional Patterson difference maps

A three-dimensional Patterson difference synthesis for the $YbCl_3$ adduct of cross-linked rennin, which at the time was thought to be among the more promising ones, gave a map containing not only many peaks but also a strong, almost continuous column of vector density along one axis, and less continuous columns along the other axes. If these were genuine it could mean, among other things, a number of sites with cross vectors approximately parallel to the axes; alternatively, the fact that such columns are also present in the three-dimensional vector map of cross-linked rennin itself suggested lack of molecular isomorphism. This problem led to a consideration of the nature of the spurious features to be expected in such maps.

The reflexions for which $||F_{PH}|-|F_P||$ is less than $|F_H|$ are the general *hkl* reflexions; those from the principal zones do not suffer from this error. This indicates an origin for some of the spurious features and suggests that such features could be reduced by suitably weighting the contributions of the two types of reflexions. The acentric (*hkl*) structure amplitudes are too low: they are approximately equal to $|F_H|\cos\alpha$ when F_H is small compared with F_P, as indeed it is in heavy-atom adducts of protein crystals (Fig. 1). Although each *hkl* reflexion is reduced by a different amount, on the average reflexions are multiplied by a factor that is expected to approximate the average value of $\cos\alpha$, which is $2/\pi=0{\cdot}6366$. Therefore, by weighting the acentric reflexions up by the factor $\pi/2$ with respect to the principal zone (centric) reflexions, or alternatively weighting the centric reflexions down by the factor $2/\pi$ with respect to the acentric reflexions, one can expect an improvement in the map.

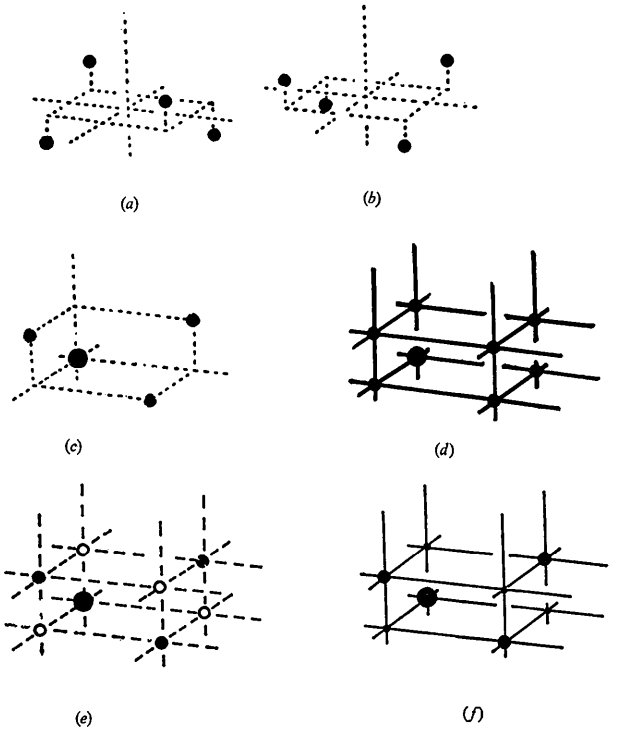

Fig. 2. (*a*) and (*b*) Left- and right-handed tetrahedral groups of atoms. (*c*) Vectors between the atoms, which are exactly the same for both groups (only those in positive directions are shown). (*d*) Vector map given by centric reflexions only; additional peaks appropriate to a doubled racemic structure, and positive columns through the peaks. (*e*) Vector map given by acentric reflexions only (● indicates positive peaks, ○ negative peaks and --- negative columns). (*f*) Vector map given by all reflexions without selective weighting; additional peaks and columns weaker than in (*d*). Selective weighting removes the additional peaks and columns, giving a map in which the only substantial vector densities are at the correct positions shown in (*c*).

The magnitude of the weighting factor was confirmed by calculations for the following artificial structure. The protein structure was simulated by 50 randomly placed atoms of equal diffracting power in the rennin cell with space group *I*222, and structure amplitudes, 'F_p', were calculated for all reflexions from planes having spacings greater than 5·5 Å. A single 'heavy atom' was added, with a diffracting power to give realistic changes in structure amplitude as defined by the formula:

$$\text{average \% change in } F = 100\,\frac{\sum||F_{PH}|-|F_P||}{\sum|F_P|}.$$

A single mercury atom in a protein of molecular weight 30,000 per asymmetric unit gives $\frac{\sum||F_{PH}|-|F_P||}{\sum|F_P|}=20\%$ (calculated from the figures for haemoglobin by Green, Ingram & Perutz, 1954). Two sets of calculations of F_{PH} for changes of 9% and 23% were made. The average value of $||F_{PH}|-|F_P||$ for the acentric reflexions was found to be 0·64 $|\overline{F_H}|$ for a change of 9%, and 0·63 $|\overline{F_H}|$ for a change of 23%. These figures are very close to $2/\pi$. For the centric reflexions, on the other hand, the values of $||F_{PH}|-|F_P||$ were very nearly equal to F_H, except for a few cases where there was a sign cross-over.

Weighted Patterson difference synthesis

Three-dimensional Patterson difference syntheses were then calculated for the same artificial structures, first for centric reflexions only, second for acentric reflexions only, third for all reflexions with $||F_{PH}|-|F_P||$ values unaltered, and finally for down-weighted centric reflexions of value 0·64 ($||F_{PH}|-|F_P||$) together with unaltered acentric reflexions. The results are illustrated in Fig. 2.

A single site in space group *I*222 means a tetrahedral group of four atoms as in Figs. 2(*a*) and 2(*b*). The correct vector peaks should appear on the cell faces as in Fig. 2(*c*); the left- and right-handed structures in

Figs. 2(*a*) and 2(*b*) give exactly the same vectors. (In $I222$ and $I2_12_12_1$, self vectors – those between symmetry-related atoms – appear only on cell faces and at half-way planes, while cross vectors are mostly in the body of the vector cell.) Centric reflexions by themselves give extra vector peaks as in Fig. 2(*d*), together with weaker columns of vector density passing through the peaks. The extra peaks are those appropriate to a doubled (racemic) structure with both left- and right-handed groups of atoms. The columns may be regarded in the following way. The centric reflexions are those for the principal projections; each projection map may be imagined as extending through the cell (the *hk*0 map, for instance, extended along *c*), so that peaks on the projection map become columns; where the three sets of columns intersect, the three-dimensional maxima are built up. Inevitably, they build up maxima appropriate to the doubled racemic structure.

Acentric reflexions by themselves give peaks in the correct positions for the single (left- or right-handed) structure, but with *negative* columns and peaks [Fig. 2(*e*)] in the same places as the positive columns and peaks in Fig. 2(*d*). When both centric and acentric reflexions are included, without selective weighting, the extra peaks and positive columns due to the centric reflexions are reduced but not completely removed [Fig. 2(*f*)], as would be expected in view of the considerations given earlier. When centric reflexions are down-weighted by the factor 0·64, but acentric reflexions are used with full weight, extra peaks and columns are removed, and the vector peaks are in the correct positions for the single structure as in Fig. 2(*c*). There are, of course, still some imperfections, but the principal imperfections of the three-dimensional map have been removed by reducing the contribution of the centric reflexions.

Following this result on an artificial structure, the same weighting procedure was used in a recalculation of the three-dimensional Patterson difference synthesis of the $YbCl_3$ adduct of rennin. In the resulting map, the columns of vector density along the axes were no longer present, and certain peaks were removed or weakened, while many others suffered little change. Some of the removed peaks are related to persistent peaks in the same way that false racemic peaks were related to genuine peaks in the artificial structure. The positions of some principal persistent peaks on the cell faces are consistent with space group $I222$ rather than $I2_12_12_1$; however, the map is still complex, suggesting 10 or 12 sites, and a comprehensive interpretation has not been achieved. Nevertheless, the weighting procedure has certainly removed some of the spurious features. This procedure will now be applied to the data for other adducts which may be simpler.

The straightforward $(||F_{PH}|-|F_P||)^2$ synthesis without selective weighting has been used effectively for some protein structures (Kartha, Bello, Harker & De Jarnette, 1963), but for the most part these were monoclinic structures with very few sites. The proportion of centric reflexions for a monoclinic structure is small, and spurious columns would stem from only one face of the cell, so that there is no crossing to create spurious peaks; it is for orthorhombic structures, or any that have more than one centric projection, that the weighting procedure is desirable to reduce the contribution of the centric reflexions. It is especially desirable for multi-site adducts, where interpretation of the vectors would be difficult enough even if the map were perfect, and so every effort should be made to remove as many sources of confusion as possible. A further improvement would be to weight the individual values of $||F_{PH}|-|F_P||$ by quantities based on the probability of sign cross-overs for centric reflexions and the analogous situation for acentric reflexions. Expressions for the two types of reflexion are given by Srinivasan (1968). The affected terms are those for which $|F_{PH}|$ or $|F_P|$, or both, are small (comparable with $\overline{F_H}$).

Three-dimensional vector maps as well as projection maps will certainly be necessary for determining heavy-atom positions in multi-site adducts of orthorhombic crystals, because even if it is possible to find the coordinates of all the sites in projections (see paper II), the chirality (left- or right-hand character) of each site is not determined; all three principal projection maps are equally compatible with all possible combinations of chiralities for the several sites. The chiralities of the sites with respect to each other must be determined from cross vectors in three dimensions, which appear mostly in the body of the vector cell, and it is most important that the peaks in the body of the cell should not be false racemic peaks.

We wish to thank the Medical Research Council and the United States National Institutes of Health for financial support which made this investigation possible.

References

BUNN, C. W., CAMERMAN, N., LIANG TUNG-T'SAI, MOEWS, P. C. & BAUMBER, M. E. (1970). *Phil. Trans.* B**257**, 153.

GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. Roy. Soc.* A**225**, 287.

KARTHA, G., BELLO, J., HARKER, D. & DE JARNETTE, F. E. (1963). In *Aspects of Protein Structure*. Ed. G. N. RAMACHANDRAN, p. 13. London: Academic Press.

PHILLIPS, D. C. (1966). *Advances in Structure Research by Diffraction Methods*. Ed. R. BRILL & R. MASON, Vol. 2, p. 75. Interscience Publishers.

ROGERS, D. (1950). *Acta Cryst.* **3**, 455.

SINGH, A. K. & RAMASESHAN, S. (1966). *Acta Cryst.* **21**, 279.

SRINIVASAN, R. (1968). *Z. Kristallogr.* **126**, 175.