

phases of triple products. Before performing these calculations we had expected that when the heavy-atom positions were unknown we would find an improved formula related to (1) and (3), with a scaling factor depending on the magnitudes of $\tau_{\mathbf{hk}}$ and $\tau_{\overline{\mathbf{hk}}}$. However, no improved formula was obtained; this, we think, is probably due to our use of the assumption $\langle S_i | \tau_{\overline{\mathbf{hk}}} \rangle = 0$.

In the improvement described in the preceding section we employ the magnitudes of the individual structure factors. In view of the excellent results it is expected that more intricate calculations *via* the joint probability distribution of the magnitudes and phases of $F_{\mathbf{h}}$, $F_{\mathbf{k}}$, $F_{\overline{\mathbf{h+k}}}$, $F_{\overline{\mathbf{h}}}$, $F_{\overline{\mathbf{k}}}$ and $F_{\overline{\mathbf{h+k}}}$ will lead to a formula for $\sin \varphi_{\mathbf{hk}}$ which is closely related to our improved formulae (1) and (24).

Surprisingly, for dextimide the results with the improved formula are better than the results with the formula in which the exact values of the cosines are used (Table 1). Therefore it follows that here the estimates compensate for the approximations made to obtain (1) and (2).

APPENDIX

(i) If there are three identical anomalous scatterers one may approximate

$$\tau''_{\mathbf{hk}} \simeq 3f'' [f(\mathbf{h})f(\mathbf{k}) + f(\mathbf{h})f(\overline{\mathbf{h+k}}) + f(\mathbf{k})f(\overline{\mathbf{h+k}})] \\ \times \left\{ 1 + \frac{2}{3} \sum_{i=1}^3 [\cos 2\pi\mathbf{h} \cdot \mathbf{U}_i + \cos 2\pi\mathbf{k} \cdot \mathbf{U}_i + \cos 2\pi(\mathbf{h+k}) \cdot \mathbf{U}_i] \right\}$$

Acta Cryst. (1978). A34, 450–453

Overlapping Patterson Peaks and Direct Methods: The Structure of Prostratin

BY P. E. NIXON

*Department of Chemistry, University of Auckland, Auckland, New Zealand and Astbury Department of Biophysics, University of Leeds, Leeds LS2 9JT, England**

(Received 21 December 1977; accepted 19 January 1978)

A generalized discussion of direct methods of solving the phase problem leads to the suggestion that overlapping vectors in the Patterson function may be the cause of failures in these methods. A simple removal of these overlapping vectors is proposed, which produces modified $|E|$ values. These were used with the program *MULTAN* to solve the structure of a prostratin derivative, $C_{22}H_{28}O_6$, which had otherwise resisted solution.

1. Direct methods

For the discussion of direct methods which follows, an analogy with the heavy-atom method may be useful.

* Present address.

$$\simeq 3f'' [f(\mathbf{h})f(\mathbf{k}) + f(\mathbf{h})f(\overline{\mathbf{h+k}}) + f(\mathbf{k})f(\overline{\mathbf{h+k}})] \\ \times [1 + 2K(|E_{\mathbf{h}}|^2 + |E_{\mathbf{k}}|^2 + |E_{\mathbf{h+k}}|^2 - 3)],$$

where the \mathbf{U}_i are the vectors between the anomalous scatterers and K is given by (23).

(ii) If there are four identical anomalous scatterers with a centrosymmetric configuration one may approximate

$$\tau''_{\mathbf{hk}} \simeq 4f'' [f(\mathbf{h})f(\mathbf{k}) + f(\mathbf{h})f(\overline{\mathbf{h+k}}) + f(\mathbf{k})f(\overline{\mathbf{h+k}})] \\ \times \left\{ 1 + \frac{1}{2} \sum_{i=1}^2 [\cos 2\pi\mathbf{h} \cdot \mathbf{U}_i + \cos 2\pi\mathbf{k} \cdot \mathbf{U}_i + \cos 2\pi(\mathbf{h+k}) \cdot \mathbf{U}_i + \sum_{i=1}^2 [\cos 2\pi\mathbf{h} \cdot \mathbf{V}_i + \cos 2\pi\mathbf{k} \cdot \mathbf{V}_i + \cos 2\pi(\mathbf{h+k}) \cdot \mathbf{V}_i] \right\} \\ \simeq 4f'' [f(\mathbf{h})f(\mathbf{k}) + f(\mathbf{h})f(\overline{\mathbf{h+k}}) + f(\mathbf{k})f(\overline{\mathbf{h+k}})] \\ \times [1 + 3K(|E_{\mathbf{h}}|^2 + |E_{\mathbf{k}}|^2 + |E_{\mathbf{h+k}}|^2 - 3)],$$

where the \mathbf{U}_i are the single vectors between the anomalous scatterers and the \mathbf{V}_i are the double vectors. Again K is given by (23).

References

- HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1977). *Acta Cryst.* A33, 873–878.
KROON, J., SPEK, A. L. & KRABBENDAM, H. (1977). *Acta Cryst.* A33, 382–385.

After the first stage (finding the heavy atoms) one has quite accurate knowledge of small parts of the electron density, complete ignorance concerning the remaining electron density, and approximations to the phases throughout reciprocal space. With a direct method of phase determination, after assigning values for a few

phases, one has quite accurate knowledge of the phases in a few parts of reciprocal space, complete ignorance of the remaining phases, and some sort of approximation to the electron density throughout the unit cell. Although this approximate electron density is not normally calculated, it is implicit in the phases known at that stage. One proceeds by extending the known region in each case.

It is usual to regard the structure factor amplitudes as the observed data; the problem is regarded as solved when a set of atomic parameters has been found which, by Fourier transformation, generates a corresponding set of structure factor amplitudes which agree satisfactorily with the data. However, the Patterson function may instead be regarded as the observed datum, since it is readily obtained from the X-ray observations. The problem is solved when a plausible set of atomic parameters generates by self-convolution a function which agrees satisfactorily with the observed datum.

A direct method of phase determination may be seen then as the postulation of a very approximate electron density throughout the unit cell, followed by its refinement; the refinement is subject to two constraints. First, the implicit 'calculated' Patterson function must be made to agree more closely with the 'observed' Patterson function. This is achieved simply by increasing the number of structure factors to which phases have been assigned (or, if weights are used, by making all the weights approach unity). The phases assigned to the structure factors do not influence the 'calculated' Patterson function. The second constraint is that the electron density implicit in the phases and weights known at an early stage should become more atomic in character. New phases are generated and old ones improved by the use of relations which have been derived from the assumed atomic nature of the electron density.

The number of atoms is not built into the phase relations in a very direct way; *i.e.* not in such a way as to encourage strongly the solution to come up with approximately the correct number of atoms.

2. What can go wrong during this procedure?

The observed Patterson function may contain a few large peaks, caused by the accidental overlap of several interatomic vectors. An approximate structure may be able to throw up a few large spurious peaks which when self-convoluted generate these prominent features of the Patterson map. We can identify those structure factors which contribute most to the overlapping peaks on the Patterson map, and delay their introduction into the set of phased reflexions. Since the magnitude of a normalized structure factor influences the likelihood of its introduction into the set of phased reflexions, we may weight down these potentially troublesome structure factors.

3. Other treatments of the problem

If a large peak appears in the observed Patterson function at *e.g.* $\frac{1}{2}, 0, 0$, then structure factors with h even will be large, and those with h odd will be small. Hauptman & Karle (1959) give an actual example (more complicated than above) and overcome this rational dependence of the atomic positions by renormalizing sets of reflexions.

In the simple example above one would scale so that $\langle |E|^2 \rangle_{h \text{ even}} = 1$ and separately $\langle |E|^2 \rangle_{h \text{ odd}} = 1$. Block & Perloff (1963) used a computer program to search reciprocal space for sets of reflexions with an abnormal $\langle |E|^2 \rangle$.

If $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = \mathbf{0}$, and we know the phases φ_1, φ_2 , of $\mathcal{E}(\mathbf{h}_1)$ and $\mathcal{E}(\mathbf{h}_2)$, a measure of the accuracy of the prediction of φ_3 is $\cos(\varphi_1 + \varphi_2 + \varphi_3)$ (= 1 for a completely accurate prediction). Hauptman & Karle (1962) derived an expression for $|\mathcal{E}_1 \mathcal{E}_2 \mathcal{E}_3| \cos(\varphi_1 + \varphi_2 + \varphi_3)$ in terms of the Patterson function; if individual Patterson peaks could be identified, and overlap was suspected, the value of $|\mathcal{E}_1 \mathcal{E}_2 \mathcal{E}_3| \cos(\varphi_1 + \varphi_2 + \varphi_3)$ could be reduced.

Karle & Hauptman (1964) pointed out that the Patterson function should not have any negative regions and that if the minimum bond length is d , $P(\mathbf{r}) = 0$, $|\mathbf{r}| < d$ (if the origin peak is removed). They suggested modifying the Patterson to fulfil these requirements if they are not already met, and then calculating the Fourier transform. The Fourier coefficients yield new values of the normalized structure factors.

4. Removal of overlapping peaks

The remedy suggested in this paper is simply to remove the large peaks from the Patterson function, Fourier transform to obtain modified structure factors, and then proceed normally. This will reduce those structure factors which contribute most to the overlapping peaks without explicitly identifying them. The 'calculated' Patterson function will be forced to conform to the majority of the 'observed' Patterson function, and not just to its outstanding features.

The program *MULTAN* (Germain, Main & Woolfson, 1971) was used because it is almost automatic in operation. This was considered to give a more objective test of the proposed modification than a calculation which required a crystallographer's judgement. This was felt to be particularly important in the choosing of origin- and enantiomorph-fixing and starting reflexions (described by Germain, Main & Woolfson, 1970).

5. Prostratin

Prostratin is an extract from *Pimelea prostrata* which proved difficult to characterize. An MnO_2 oxidation product eventually produced crystals of poor quality.

Crystal data: $C_{22}H_{28}O_6$, $M_r = 388.6$, space group $P2_12_12_1$, $a = 9.901$ (2), $b = 11.739$ (3), $c = 17.879$ (3) Å, crystal size $ca\ 0.3 \times 0.18 \times 0.12$ mm. X-ray measurements were made on a Hilger & Watts four-circle diffractometer, using Cu $K\alpha$ radiation (Ni filter). 715 reflexions had intensities $I > 3\sigma$ (σ calculated from counting statistics): 880 had $I > 2\sigma$. Absorption corrections were not applied. Various methods were tried to solve the phase problem, including the *MULTAN* program package, but without success.

A sharpened Patterson function was calculated from all the normalized structure factors [$E(000)$ was not included]. This showed several peaks whose height was about 7% of the origin peak. With four molecules of $C_{22}H_{28}O_6$ in the unit cell, the highest single vector (O—O) should have a height 1.3% of the origin peak, whereas C—C vectors should be 0.7% of the origin peak height. This indicates that the highest point on the Patterson map corresponds to 5–10 overlapping vectors.

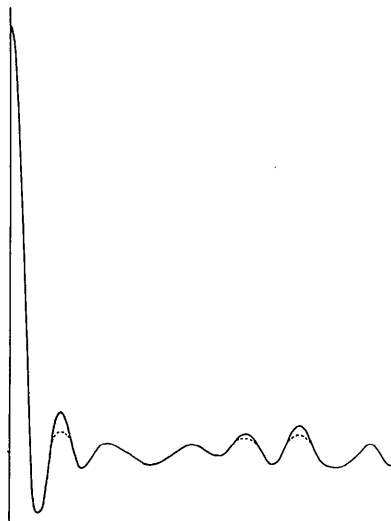


Fig. 1. The sharpened Patterson function (continuous line) of prostratin and the modification made to it (dotted line). The section is taken through the origin and the largest non-origin peak.

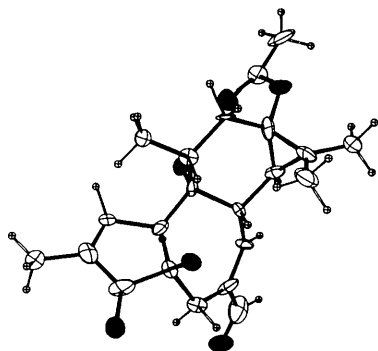


Fig. 2. An ORTEP diagram of prostratin. Oxygen atoms are shown filled, but hydroxyl hydrogen atoms are not shown.

A cut-off level, c , was chosen arbitrarily in the first trial as 5% of the origin peak. A simple modifying function would be

$$P' = c, P \geq c, \\ P' = P, P < c.$$

In fact a smoother modifying function was used, and the origin peak was not changed:

$$P'(\mathbf{x}) = c \tanh [P(\mathbf{x})/c], P > 0 \text{ and } |\mathbf{x}| > 2 \text{ \AA}, \\ P'(\mathbf{x}) = P(\mathbf{x}), P \leq 0 \text{ or } |\mathbf{x}| < 2 \text{ \AA}.$$

Fig. 1 shows a section through the Patterson function passing through the origin and the highest non-origin peak, together with the modifications made according to this cut-off procedure.

This modified Patterson function P' was Fourier transformed to yield a set of modified $|E|^2$ values, and by taking the square root a set of modified $|E|$ values was obtained. Although the Fourier coefficients of this modified Patterson are not necessarily positive, the smallest $|E|^2$ was in fact -0.18 , and none of the E 's above 0.57 had $|E|^2 < 0$ after modification. Only those $|E|$'s > 1.2 were used in the *MULTAN* program. The r.m.s. change in $|E|$ for those structure factors with $|E| \geq 1.2$ was 0.03.

The modified $|E|$'s greater than 1.2 were input to the program *MULTAN* and eight solutions obtained. The most likely of these to be the correct one was used to calculate an E map which revealed 27 of the 28 non-hydrogen atoms in the structure. The remaining atoms were found by difference maps and the structure was refined in the usual way, using the original structure factor amplitudes. A preliminary account of the structure has been published (McCormick, Nixon & Waters, 1976), and some of its interesting chemical and biological features discussed. Fig. 2 shows an ORTEP diagram of the structure.

6. Sodium cupric oxalate

This method was also used successfully to solve the structure of sodium cupric oxalate dihydrate, $Na_2Cu(C_2O_4)_2 \cdot 2H_2O$ (Braithwaite, Nixon & Waters, 1978), but in retrospect this may have been using a sledgehammer to crack a nut.

I should like to express my thanks to Professor T. N. Waters for his encouragement and advice on this work, and to Dr D. A. D. Parry for the fast Fourier transform program, used for performing the transformation on the Patterson function.

References

- BLOCK, S. & PERLOFF, A. (1963). *Acta Cryst.* **16**, 1233–1238.
 BRAITHWAITE, A., NIXON, P. E. & WATERS, T. N. (1978). To be published.

- GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* B26, 274–285.
- GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A27, 368–376.
- HAUPTMAN, H. & KARLE, J. (1959). *Acta Cryst.* 12, 846–850.
- HAUPTMAN, H. & KARLE, J. (1962). *Acta Cryst.* 15, 547–550.
- KARLE, J. & HAUPTMAN, H. (1964). *Acta Cryst.* 17, 392–396.
- MCCORMICK, I. R. N., NIXON, P. E. & WATERS, T. N. (1976). *Tetrahedron Lett.* (1976). 20, 1735–1736.

Acta Cryst. (1978). A34, 453–459

Optical Selected-Area Diffraction Patterns of High-Resolution Electron-Microscope Images for Crystal Analysis

BY T. TANJI AND H. HASHIMOTO

Department of Applied Physics, Osaka University, Suita 565, Japan

(Received 17 December 1976; accepted 25 November 1977)

Optical selected-area diffraction patterns made from high-resolution electron micrographs of crystals have been used as a source of diffraction information from areas as small as a single unit cell of the crystal. The intensities of the electron diffraction pattern of the specimen crystal and the optical diffraction patterns of high-resolution electron-microscope images have been discussed by electron optical image formation theory taking account of spherical aberration and defocusing of the objective electron lens and it is concluded that the optical diffraction pattern may be identical with the electron diffraction pattern if the electron micrograph is photographed under optimum conditions. Optical diffraction patterns from areas of 80, 30 and 10 Å in diameter of labradorite feldspar have been taken and the orientation of two adjoining grains, 30 Å in diameter, has been determined. The diffraction pattern from a unit-cell area has also been taken and compared with the calculated intensity.

1. Introduction

The selected-area diffraction technique in electron microscopy serves to provide a diffraction pattern from a small finite area of a crystalline specimen under electron-microscope observation and can be used for identifying materials. Since, however, the objective lens of the electron microscope has a certain amount of spherical aberration, some diffracted beams from the region adjoining the field-limiting aperture can pass through the objective aperture and therefore influence the diffraction pattern. The error of correspondence between the specimen area selected and the area actually producing the diffraction pattern is expressed by $C_s\alpha^3$ as Riecke (1961) has shown, where α is the scattering angle of the electron waves and C_s is the spherical aberration coefficient of the objective lens. For example, when the diffraction pattern of a lattice plane with 1.13 Å spacing (Al_{222}) is taken at 100 kV with a lens having a spherical aberration coefficient $C_s = 3.6$ mm, the error in the selected specimen area is 1260 Å. Therefore, the correspondence between image and diffraction pattern will not hold if the area limited by the field aperture is small. However, for

accurate structure analysis, the uncertainty in the correspondence should be avoided.

Uyeda, Dupouy, Perrier, Ayroles & Bousquet (1963) have pointed out that the error of the area limited by the field-limiting aperture becomes small at high voltages because then the scattering angle α for the electrons is small. Making use of this relation, Koreeda, Okamoto, Shimizu & Katsuta (1971) have shown that the minimum accurate area for taking the diffraction pattern of lattice planes with 1.0 Å spacing at 500 kV becomes 80 Å and diffraction patterns may be taken from an area 250 Å in diameter for crystals containing precipitates and lattice imperfections. Koike & Ueno (1973) and Geiss (1975) have used a small probe of the scanning transmission electron microscope to take diffraction patterns from areas as small as 50 Å in diameter of crystalline materials. Since in these techniques the field-limiting aperture limits the intensity of the electron beams which form the electron diffraction pattern, exposures of 1 to 5 min or more are necessary to take diffraction patterns. During these long exposures radiation damage, specimen contamination and drift of the specimen occur and thus there may be poor correspondence between image and diffraction