necessary input for the calculation. According to (36) and (49), only the phase invariants of the structure-factor triplets and quartets affect the calculation.

Both the reflection intensities and mode excitations have peak values at the corresponding critical angles. This is due to the term $[\sin\varphi+(\sin^2\varphi-\sin^2\theta_i)^{1/2}]^{-1}$ involved in the wavefield amplitudes.

Small changes in φ_N cause large variations in the reflection intensities, as has already been shown in Figs. 6, 7 and 8. Experimentally, in order to detect this variation, strictly parallel and intense radiation sources with wavelength tunability are required. Use of synchrotron radiation is indispensable for carrying out the GIXD experiments (*e.g.* Cowan, Brennan, Jach, Bedzyk & Materlik, 1986; Sakata & Hashizume, 1988; Durbin & Gog, 1989).

In conclusion, we have derived analytical expressions for reflection intensity, wavefield amplitude and accommodation for two-beam and symmetric N-beam ($N>2$) GIXD. A new geometric scheme has also been provided in this study to reveal the excitation of the dispersion surface. Numerical calculation is straightforward for symmetric N-beam GIXD and should be modified for general asymmetric cases for which the obtaining of analytical expressions for intensities and wavefield amplitudes is not guaranteed.

The authors thank the National Science Council for financial support. One of the authors (HHH) is grateful to the same organization for providing a graduate fellowship.

References

AFANAS'EV, A. M. & MELKONYAN, M. K. (1983). *Acta Cryst.* **A39**, 207-210.

ALEKSANDROV, P. A., AFANAS'EV, A. M. & STEPANOV, S. A. (1984). *Phys. Status Solidi A*, **86**, 143-154.

COWAN, P. L. (1985). *Phys. Rev. B*, **32**, 5437-5439.

COWAN, P. L., BRENNAN, S., JACH, T., BEDZYK, M. J. & MATERLIK, G. (1986). *Phys. Rev. Lett.* **57**, 2399-2402.

DURBIN, S. M. & GOG, T. (1989). *Acta Cryst.* **A45**, 132-141.

EWALD, P. P. (1917). *Ann. Phys. (Leipzig)*, **54**, 519-597.

FUOSS, P. H., LIANG, K. S. & EISENBERGER, P. (1989). In *Synchrotron Radiation Research: Advances in Surface Science*, edited by R. Z. BACHRACH. New York: Plenum.

HOCHE, H. R., BRUMMER, O. & NIEBER, J. (1986). *Acta Cryst.* **A42**, 585-587.

International Tables for X-ray Crystallography (1974). Vol. IV. Birmingham: Kynoch Press. (Present distributor Kluwer Academic Publishers, Dordrecht.)

LAUE, M. VON (1931). *Ergeb. Exakten Naturwiss.* **10**, 133-158.

MARRA, W. C., EISENBERGER, P. & CHO, A. Y. (1979). *J. Appl. Phys.* **50**, 6927-6933.

SAKATA, O. & HASHIZUME, H. (1987). Report RLEMTIT, **12**, 45-57. Research Laboratory of Engineering Materials, Tokyo Institute of Technology, Japan.

SAKATA, O. & HASHIZUME, H. (1988). *Jpn. J. Appl. Phys.* **27**, L1976-L1979.

VINEYARD, G. H. (1982). *Phys. Rev. B*, **26**, 4146-4159.

Acta Cryst. (1989). **A45**, 833-839

Determination of Crystal Structures from Poor-Quality Data Using Patterson Methods

By C. C. Wilson

Neutron Science Division, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, England

(*Received* 13 *March* 1989; *accepted* 19 *July* 1989)

Abstract

It is found to be possible to solve the structures of fairly simple materials from very poor-quality diffraction data by the use of reciprocal-space Patterson methods. Data sets assessed include those from a high-resolution neutron powder diffractometer, data sets with inaccurate randomized $|F_{\mathbf{h}}|$ values, very small data sets (as few as ten reflections) and data sets with no estimates of $|F_{\mathbf{h}}|$ values given. While refinement is not always feasible from such data it is found possible to obtain atomic positions and consequent structural information with reasonable accuracy. Reasons for using Patterson rather than direct methods in such cases are discussed.

Introduction

It is not always possible to collect good-quality structure-factor data from a crystallographic experiment. Problems can arise for various experimental reasons, owing, for example, to small poor-quality crystals, or to the existence of only a powder sample. The resolution of such experimental difficulties is beyond the scope of this work but rather the potential for the extraction of as much information as possible about the structure under such unfavourable circumstances will be discussed. It is in the area of garnering structural information from poor-quality data that Patterson methods have a significant advantage over direct methods, especially when the data are very scarce.

0108-7673/89/120833-07$03.00

The advantages of Patterson methods over direct methods for very poor data

By the very nature of the phase-determining processes undertaken in direct methods of structure determination, a fairly considerable subset of reflections is required. If this is not available it is very difficult to establish a suitable phase pathway even using multisolution methods. The nature of the triplet and other relations means that one in general requires a large pool of reflections within which to combine indices in order to establish the invariant relationships. It follows that in order to exploit direct methods a fairly comprehensive list of reflections must be measured. On the other hand, model-fitting Patterson methods require only that the reflections available be in some way 'representative' of the diffraction from the sample. The number of reflections required for Patterson-method structure determination tends therefore to be somewhat smaller than for direct methods.

In addition direct methods, being fundamentally a statistical procedure based to a large extent (through probability expressions) on the measured intensities (and hence $|E_{\mathbf{h}}|$ values), require a fairly accurate set of diffracted intensities. In particular, if a weak reflection is accidentally given a large $|E_{\mathbf{h}}|$ value (as can happen in powder patterns where there are overlapped reflections) then the mistaken incorporation of this reflection early in the phase-determining procedure can seriously prejudice subsequent phases and hence lead to an incorrect map and to a failure to solve the structure. With Patterson calculations such an error on an individual reflection will not be propagated through the procedure in the same way and is liable merely to lead to inaccuracy rather than incorrectness in the final Fourier map.

There is also the possibility of obtaining very low-resolution Patterson maps which contain basically the correct structure but not necessarily at atomic resolution. This deconvolution of a correct answer from a very broadly distributed Patterson function is made possible by the provision of model stereochemical information. The phase-determination formulae of direct methods are not as applicable to this type of deconvolution of structural information from low-resolution data. The application of structural knowledge in direct methods can be used either to weight normalization and modify probabilities for invariants (Main, 1976) or in interpretation of the final E map. Both of these use the structural information in a weaker way than the *a priori* application of this information in the Patterson techniques.

For an unknown structure, the Patterson function does not rely on the missing phase information, and in this sense the calculated Patterson map is more meaningful in relation to the true solution than an equivalent $|F|$ map calculated with no phases. The relative strength and robustness of Patterson methods in the case of a poor data set arises fundamentally from the fact that the Patterson calculations are rooted firmly in the measurement and manipulation of intensities. Direct methods, on the other hand, rely to a far greater extent on the more critical phases which may be unavailable, especially for poor data sets. This leads to less stability in the calculations.

Structure determination methods

The direct-methods calculations to be described were carried out using the *MITHRIL* program (Gilmore, 1984), while the Patterson-methods calculations used *PATMET* (Wilson & Tollin, 1986). The use of these programs for neutron data has been discussed elsewhere (Wilson, 1987).

The *PATMET* techniques are described in detail elsewhere (Wilson, 1988) but a very brief outline will be given here to clarify results presented later. *PATMET* normally uses sharpened squared structure factors, denoted $|F^s_{\mathbf{h}}|^2$. The basic stages of the *PATMET* calculations are:

(i) *$I(\theta,\varphi)$ function.* If there is a planar group present in the structure, the $I(\theta,\varphi)$ function (Tollin & Cochran, 1964) locates this, in terms of the two spherical polar angles θ and φ;

(ii) *$\sigma(\theta_1,\theta_2,\theta_3)$ function.* This rotation function (Tollin & Cochran, 1964) finds the best fit of input model to the data, performing rotations described by the Eulerian angles θ_1, θ_2, θ_3. If the σ-function calculation follows planar-group orientation determination, only a 1D rotation-function calculation is required, in the plane found for this group by the $I(\theta,\varphi)$. This 1D rotation function is denoted $\sigma(\theta_1)$ here;

(iii) *Q functions.* These translation functions (Tollin, 1966) position an oriented group within the cell with respect to each individual symmetry element.

It is convenient to define an agreement index between the coordinates found using *PATMET* and the refined set for the known structure. This parameter, termed the coordinate R factor R_C, is defined as

$$R_C=(1/3N)\sum_{i=1}^{N}\{[(x_{ir}-x_{ip})/x_{ir}]^2+[(y_{ir}-y_{ip})/y_{ir}]^2+[(z_{ir}-z_{ip})/z_{ir}]^2\}^{1/2},$$

where the summation is over the N atoms i in the model; x_{ir}, x_{ip} *etc.* refer to the refined and *PATMET* coordinates respectively.

Solution parameters and models used

The three crystal structures examined here are the following:

(I) 3-deazauracil (Low & Wilson, 1983; Fig. 1). $C_5H_5NO_2$, orthorhombic, $P2_12_12_1$, $a=8{\cdot}638$, $b=5{\cdot}279$, $c=11{\cdot}220$ Å, $Z=4$. The expected orientation of the planar group is at (θ, φ) values of (44°, 53°). Three Q functions can be calculated, positioning the group with respect to the three screw axes. The model used for the structural solution is shown in Fig. 1;

(II) 2′,3′,5′-tri-*O*-acetylguanosine (Wilson, Low & Tollin, 1985; Fig. 2). $C_{16}H_{19}N_5O_8$, monoclinic, $P2_1$, $a = 7{\cdot}414$, $b = 11{\cdot}491$, $c = 11{\cdot}790$ Å, $\beta = 99{\cdot}47°$, $Z = 2$. The planar-group orientation is at $(\theta, \varphi) = (54°, 180°)$. A Q function defines the position of the group with respect to the 2_1 axis, the other translational position being undefined in this space group. The model used is shown in Fig. 2. This structure was chosen for these analyses since the planar model group is a minor part of the overall structure;

(III) 3,4-dihydroxycyclobut-3-ene-1,2-dione, squaric acid (Semmingsen, Hollander & Koetzle, 1977; Fig. 3). $C_4H_2O_4$, monoclinic, $P2_1/m$, $a = 6{\cdot}143$, $b = 5{\cdot}286$, $c = 6{\cdot}148$ Å, $\beta = 86{\cdot}96°$, $Z = 2$. The molecule lies exactly in the $y = 1/4$ plane, $(\theta, \varphi) = (90°, 90°)$. The model is positioned within the cell by the 2D $Q(X_0, Z_0)$ function with respect to the 2_1 axis and by the 1D $Q(Y_0)$ function with respect to the mirror plane.

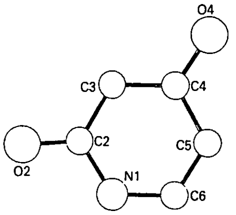

Fig. 1. The planar molecule of 3-deazauracil (excluding H atoms). The model used in the *PATMET* solution consists of all of these atoms.

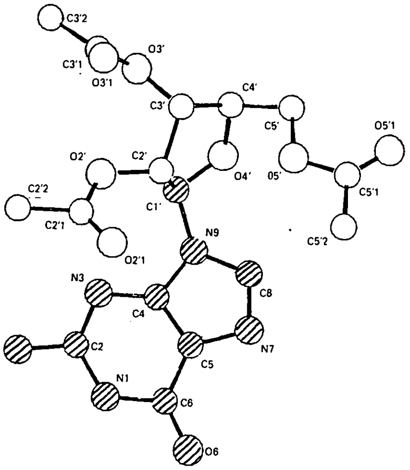

Fig. 2. The molecule of 2′,3′,5′-tri-*O*-acetylguanosine (excluding H atoms). The model atoms used are shaded.

Structural solution examples

(*a*) *High-resolution neutron powder data*

The so-called *ab initio* determination of crystal structures from very high-resolution powder data has been a focus of much recent interest (David, Johnson & Wilson, 1988; and references therein), using both X-ray synchrotron and neutron data. The data sets to be assessed were collected on the High-Resolution Powder Diffractometer, HRPD (Johnson & David, 1985), at the ISIS Spallation Neutron Source.

(i) 3,4-*Dihydroxycyclobut*-3-*ene*-1,2-*dione*: *squaric acid.* Data were collected from a very high-quality sample of squaric acid (Nelmes, Tun, David & Harrison, 1987) in the high-resolution ($\Delta d/d \sim 5 \times 10^{-4}$) instrument configuration. This material was being studied by high-resolution powder diffraction primarily to examine the small monoclinic distortion from tetragonal symmetry.

From the profile refinement of the data, a total of 193 $|F_{\mathbf{h}}|$ values was extracted. To simulate a true *ab initio* test, overlapping reflections were averaged at a $\Delta d/d$ resolution of 5×10^{-4}. Attempts to solve the structure using direct methods in *MITHRIL* (Gilmore, 1984) failed - probably because of the extreme planarity of the molecule (all atoms sit on $y = 1/4$). Traditionally, direct methods have always been at their weakest in such cases, *i.e.* where a planar moiety is a dominant fraction of the scattering power of the cell. In this case the planar group is the whole molecule and hence direct methods can be in considerable difficulty. This is demonstrated by the fact that Karle (1968) recycling using the whole published molecule as model fails to produce adequate phasing. Problems like this with planar groups suggest recourse to Patterson methods, which should be ideal for this problem.

In the Patterson-methods calculations, the top hundred sharpened $|F^s_{\mathbf{h}}|^2$ values were used. The structure was solved using the $I(\theta, \varphi)$, $\sigma(\theta_1)$ and $Q(X_0, Z_0)$ and $Q(Y_0)$ functions. The $I(\theta, \varphi)$ peak occurs at (90°, 90°). Using a C_4 square as model, we found the rotation-function peak at $\theta_1 = 160°$, the $Q(X_0, Z_0)$ peak at (0·18, 0·4) and the $Q(Y_0)$ peak at (0·25). The resulting solution coordinates for the four atoms in the model gave an R_C of 0·025.

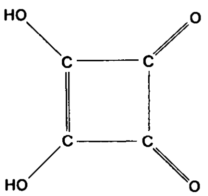

Fig. 3. A schematic representation of the structure of the planar squaric acid molecule. In this case the model used consists of the four C atoms. The Patterson-method results from using the C_4O_4 fragment as model are similar.

These coordinates were used in *MITHRIL* for Fourier recycling, which revealed the four O atoms as the next four peaks in the resulting map, and the structure is thus solved. Similar results were found using the C_4O_4 group as model and the structure was also solved using a C_6 ring as model, together with Fourier recycling. This latter example of Patterson methods orienting and positioning a six-membered ring with sufficient accuracy to give phases to solve a four-membered ring structure by Fourier methods is quite novel and an indication that the model supplied does not have to be very precise.

The procedure of Fourier recycling the final oriented and positioned model from Patterson methods to verify structural solution has been used throughout this work.

(ii) 3-*Deazauracil.* The data from this somewhat poorer-quality sample were collected in the lower-resolution configuration of HRPD ($\Delta d/d \sim 5\times10^{-3}$) (Wadsworth, Wilson & David, 1989), and consisted of 339 reflections. The intensities were extracted from the final profile fit and hence reflect the model used in the refinement. In order to reintroduce the resolution function and simulate genuine $|F_{\mathbf{h}}|$ extraction, reflections were averaged at the 5×10^{-3} resolution level. From this process, some 30 doublets and seven triplets were found to be unresolved, meaning that 81 of the 339 intensities were modified. The best method for the extraction of intensities from powder patterns of unknown structures will not be discussed here, since no truly unknown structures are being examined. This matter is considered elsewhere (David & Wilson, in preparation; Wilson & Wadsworth, in preparation).

The 300 largest $|F_{\mathbf{h}}^{s}|^2$ were used in the structural solution. The structure was found to solve with ease by both Patterson and direct methods. The results of *PATMET* were as follows: $I(\theta, \varphi)$ peak at (44°, 53·2°), $\sigma(\theta_1)$ peak at 122°, $Q(X_0, Y_0, Z_0)$ solution at (0·187, 0·324, 0·212). These parameters led to a set of coordinates which solve the structure, with an R_C of 0·048.

(*b*) *Inaccurate single-crystal data*

If for some reason one has only been able to collect a limited or inaccurate single-crystal data set it is still possible to obtain reasonable structural information from Patterson methods. The ability of *PATMET* to solve structures from very few reflections is discussed below. To investigate the insensitivity of the program to errors in the experimental structure factors, a series of calculations were performed using 'randomized' data, *i.e.* data where the structure factors have been altered in some random fashion. The summary below gives the results for a series of randomized data set calculations. The degree of randomness is measured by the residual R_R defined as

$$R_R = (1/N)\sum_{\mathbf{h}} \big||F_{\mathbf{h}}^R| - |F_{\mathbf{h}}|\big|/|F_{\mathbf{h}}|,$$

where there are N reflections in the data set, $|F_{\mathbf{h}}|$ is the observed structure factor, and $|F_{\mathbf{h}}^R|$ the randomized structure factor of reflection $\mathbf{h}$. For the unmodified observed data, $|F_{\mathbf{h}}^R| = |F_{\mathbf{h}}|$, and $R_R = 0$. In general, the program appears to be robust to fluctuations of up to ~100% in the structure factors used, represented by $R_R \sim 1$. Randomized data-set calculations were carried out on the test structures (I) and (II).

(i) 3-*Deazauracil.* A series of calculations was performed on this structure using data sets with various degrees of randomization. The randomization was based on fluctuations at some chosen level, using a random number generator to choose the sign of the fluctuation and modify its magnitude. The results of the calculations were as follows (summarized in Table 1).

For 20% ($R_R \sim 0{\cdot}2$) randomization the results obtained from *PATMET* were essentially unchanged from those with the correct data set. At 50% ($R_R \sim 0{\cdot}5$), the peak in the $I(\theta, \varphi)$ function has shifted by some 6°, but the rotation function finds the correct orientation as before. The correct shift is only indicated as the second-highest peak in the $Q(Y_0, Z_0)$ function but the combined map (Wilson & Tollin, 1988) gives the correct (X_0, Y_0, Z_0) shift as the highest ranked solution. The structure is thus solved easily with these data. When 100% ($R_R \sim 1$) randomization is imposed, the degree of inaccuracy of the data is such that there are some problems in obtaining the solution. The rotation function finds an orientation of $\theta_1 = 178°$, implying a rotation of the six-membered ring by 60°. The importance of using only the most basic molecular framework when very poor data are being used is stressed by this orientational ambiguity. Under such circumstances, incorrectly oriented side groups can cause problems in later calculations. For this reason, only a six-membered ring was used for the 100% (and 200%) randomized calculations, excluding the exocyclic atoms from the model. Using this model in the present case, the correct shifts were indicated as the second-ranked solution in the combined Q map, which is an indication that the solution is beginning to be less well indicated. By the stage of 200% ($R_R \sim 2$) randomization, the program is in serious difficulty in attempting to solve the structure. The $I(\theta, \varphi)$ peak has shifted further to some 10° from the correct value and while the rotation function gives a peak at $\theta_1 = 120°$, indicating that the orientation can still be approximately obtained, the Q functions fail to locate the fragment in the cell. The correct shifts are not listed among the combined Q-map solutions and the conclusion has to be that at this stage *PATMET* has failed to solve the structure. The $|F_{\mathbf{h}}^R|$ data by this stage, however, bear very little resemblance to the original $|F_{\mathbf{h}}|$ values, so perhaps this is not too surprising.

(ii) 2′,3′,5′-*Tri-O-acetylguanosine*. For 50% randomization the solution is obtained easily (see Table 1), as it is for 100% randomization. For 200% randomization, the solution is obtained, but with some difficulty, using the second peak in the rotation function and the fifth in the Q map. The interpretation of the latter map can be improved by using correlation of maps calculated using different data sections (Wilson & Tollin, 1988).

The results from these data sets suggest that Patterson methods are rather robust to inaccuracies, in the form of *random* errors, in the reflection data obtained from the experiment. It is of interest to test the behaviour of direct methods with the same randomized data. The results from these calculations can be briefly summarized as follows:

(i) 3-*Deazauracil*. $R_R = 0{\cdot}47$: Solution found by default calculations in *MITHRIL*. $R_R = 0{\cdot}83$: No solution obtained using quartets and examining top six E maps, or by other, more exhaustive, attempts.

(ii) 2′,3′,5′-*Tri-O-acetylguanosine*. Correct data ($R_R = 0$): Solution in default calculations. $R_R = 0{\cdot}46$: No solution obtained using quartets and examining six E maps. Once again more exhaustive solution attempts failed.

As a check on the ability of the randomized data to produce the solution under ideal circumstances, the published model for 2′,3′,5′-tri-*O*-acetylguanosine was used in the direct-methods calculations in a Karle recycling scheme (Karle, 1968), along with the $R_R = 0{\cdot}46$ data. In this case the solution was obtained. Attempts to use this model as a known (oriented and positioned) group (Main, 1976) in a standard direct-methods phasing run did not lead to a solution.

(*c*) *Structural solution using very small data sets*

(i) 3-*Deazauracil*. The results from calculations using many fewer reflections to solve this structure are also very encouraging (Table 2). The solution using 30 $|F_{\mathbf{h}}^s|^2$ values was perfectly routine. The $I(\theta, \varphi)$ peak has shifted slightly by some 3° but this is not significant for structural solution. For a data set consisting of just the top 15 reflections, the $I(\theta, \varphi)$ peak shifts again, by some 6° from the correct orientation. In this case there is some degradation in the Q-function results, with the correct solution not being the highest peak in every map. However, the combined map ranks the correct shifts as the highest solution. Once one has reduced the data set used to a mere ten reflections (the original single-crystal data set consists of 606), a few more problems arise. In this case the $I(\theta, \varphi)$ is as for 15 reflections but the rotation function gives a peak at $\theta_1 = 0°$ as opposed to $\theta_1 = 122°$ in the earlier runs. This means the molecule as found is rotated by 60° from the 'correct' solution, but due to the molecular symmetry the solutions are almost equivalent. It should again be noted that when one is using such paltry data sets it is inadvisable to include detail in the model beyond the most basic molecular shape, in this case the six-membered ring portion. The Q functions calculated using the resulting model gave the correct shifts as the second-ranked combined map (combined height of 285 as opposed to 288 for the highest solution). As can be seen from these results the solution with ten reflections is somewhat less reliable but the fact that it can be found at all is a promising indicator of the power of Patterson methods working on very limited data sets.

Table 1. *Summary of randomized data solution attempts*

3-Deazauracil (300 reflections) - *PATMET*

R_R	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Y_0, Z_0)$	R_C
0·19	44, 53	122	0·177, 0·324, 0·217	0·034
0·47	42, 47	120	0·171, 0·324, 0·217	0·029
0·95	42, 47	178*	0·185, 0·331, 0·206 (2nd)	0·020
1·90	47, 42	112*	No solution	

2′,3′,5′-Tri-*O*-acetylguanosine (300 reflections) - *PATMET*

R_R	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Z_0)$	R_C
0·46	54, 180	10	0·327, 0·330	0·079
0·92	54, 180	10	0·327, 0·331	0·078
1·85	54, 180	10 (2nd)	0·342, 0·343 (16th)	0·058

* No solution for eight-atom model, exocyclic atoms excluded.

Table 2. *Very small data set solutions in PATMET*

3-Deazauracil

N_{refs}	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Y_0, Z_0)$	R_C
30	46, 55	122	0·185, 0·318, 0·224	0·036
15	50, 54	122	0·184, 0·318, 0·223	0·095
10	50, 54	0*	0·320, 0·460, 0·360 (2nd)	0·018

2′,3′,5′-Tri-*O*-acetylguanosine

N_{refs}	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Z_0)$	R_C
300	54, 180	14	0·342, 0·341	0·067
50	54, 180	12	0·343, 0·341 (2nd)	0·056
25	54, 180	8	0·362, 0·362	0·114
15	56, 180	6	0·361, 0·362	0·114
10	56, 180	12	0·362, 0·362 (5th)	0·071
5	56, 180	12	No solution	

Squaric acid

N_{refs}	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Y_0, Z_0)$	R_C
100	90, 90	22	0·400, 0·250, 0·180	0·010
50	90, 90	24	0·400, 0·250, 0·200	0·024
25	90, 90	114	0·300, 0·250, 0·400	0·017

* No solution for eight-atom model, exocyclic atoms excluded.

Table 3. *Reflection sets used, for equal* $|F_{\mathbf{h}}|$ *estimate data*

3-Deazauracil (all $|F_{\mathbf{h}}| = 1{\cdot}0$)

h k l	*h k l*	*h k l*	*h k l*	*h k l*	
1 1 3	1 1 2	0 2 2	3 2 0	1 1 1	20-reflection data set
2 0 2	0 1 5	4 1 0	2 1 4	2 3 1	
3 0 2	4 0 3	2 2 4	3 1 2	2 1 1	
1 2 2	2 1 3	0 1 4	1 0 5	3 0 6	
0 2 5	1 1 4	4 1 4	2 2 5	2 0 7	
5 2 0	1 2 1	2 2 6	4 3 6	2 1 0	
4 0 1	4 2 2	2 0 4	1 1 6	0 1 7	
3 2 7	0 2 1	1 3 1	7 0 2	3 2 2	
1 0 3	1 2 4	4 3 5	2 0 6	2 1 9	
0 3 3	2 0 5	6 0 7	3 3 7	0 0 14	

2′,3′,5′-Tri-*O*-acetylguanosine (all $|F_{\mathbf{h}}| = 1{\cdot}0$)

h k l	*h k l*	*h k l*	*h k l*	*h k l*
−2 0 2	−2 0 1	−1 0 2	−2 1 1	−1 1 2
−1 0 1	−3 0 3	2 0 1	−4 0 4	0 2 0
1 2 2	−2 0 3	2 1 1	−1 2 1	−2 1 3
1 2 1	−3 0 1	2 3 1	−1 1 3	1 1 1
−2 2 1	0 3 4	0 5 3	1 2 5	1 5 0

(ii) 2′,3′,5′-*Tri-O-acetylguanosine.* As can be seen from Table 2, the solution is indicated unambiguously in *PATMET* for data sets limited to 300, 50, 25 and 15 reflections. In the ten-reflection case the orientation is found easily but the correct shifts are indicated as fifth-highest peak in the *Q* map. This is still an impressive performance. For a five-reflection data set the orientation of the planar group is still found, but there are now insufficient data to calculate the translation function.

As a comparison, it was found impossible to obtain a convergence map in *MITHRIL* from the data set containing 50 reflections, due to the extreme paucity of phase relationships within this group. For a 300-reflection data set the direct-methods calculations were not found to reveal the structure.

(iii) *Squaric acid* (*neutron data*). In *PATMET*, the solutions for data sets of as few as 25 reflections were found to be straightforward (Table 2). Below this (for 15 and 10 reflections), there are problems with the $I(\theta, \varphi)$ function, in which the correct solution is only indicated as the second-highest peak. These problems could be related to the fact that the Patterson-method calculations for neutron data are not yet optimized (Wilson, 1987).

(*d*) *Data comprising a subset of reflections given equal* $|F_{\mathbf{h}}|$ *values*

(i) 3-*Deazauracil.* For ten reflections listed in Table 3, a subset of the full data set chosen to contain the strongest reflections, all $|F_{\mathbf{h}}|$ values were estimated to be equal. Using this type of data, with essentially no attempt made to estimate $|F_{\mathbf{h}}|$ values beyond choosing 'strong' reflections, *PATMET* was found to solve the structure easily for both 50 and 20 reflections (Table 4). The solution for ten reflections in this case is more obscure.

Table 4. *Summary of results, equal* $|F_{\mathbf{h}}|$ *estimate data*

3-Deazauracil

N_{refs}	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Y_0, Z_0)$	R_C
50	48, 56	122	0·185, 0·319, 0·227	0·048
20	50, 54	118	0·173, 0·313, 0·211	0·033
10	No solution			

2′,3′,5′-Tri-*O*-acetylguanosine

N_{refs}	$I(\theta, \varphi)$ (°)	$\sigma(\theta_1)$ (°)	$Q(X_0, Z_0)$	R_C
25	54, 180	10 (2nd)	0·379, 0·383 (2nd)	0·156
15	No solution			

(ii) 2′,3′,5′-*Tri-O-acetylguanosine.* For the 25-reflection data set shown in Table 3, the solution is obtained but the correct orientation in the rotation function is the second-highest peak (Table 4).

Concluding remarks

In general it has been found that Patterson-methods calculations are remarkably robust with regard to the quality of the $|F_{\mathbf{h}}|$ data provided. The performance of *PATMET* in solving crystal structures from powder data, inaccurate data and sparse data has been illustrated by the above examples. Owing to the nature of direct methods, these cannot be as competitive in this field of structure determination from very poor data. Direct-methods solution was found to be untenable in several of the illustrative cases where the Patterson-method solution was fairly routine.

Just how useful these results are to the structure investigator is another matter. Undoubtedly it is always satisfying and valuable to 'solve' a crystal structure, that is to obtain the atomic positions to a reasonable degree of accuracy. Unfortunately, for many of the data sets examined above, such approximate determination of atomic coordinates will be the limit of the crystal-structure investigation. For small and/or poor data sets, refinement of all but the basic structural framework will often be out of the question. Certainly one is very unlikely to obtain accurate bonding or thermal parameters from such data.

The fact remains, however, that the general disposition of the molecule, or some important group within the molecule, can be discovered from even very poor data. The value of such information in conformational studies, stacking analyses or hydrogen-bonding network studies, for example, can be great. The strength of Patterson methods to yield information in these circumstances is available to be exploited. By provision of sensible model information in this way even poor data can be manipulated to reveal useful structural parameters.

References

DAVID, W. I. F., JOHNSON, M. W. & WILSON, C. C. (1988). *Mater. Sci. Forum*, **27/28**, 49-62.

GILMORE, C. J. (1984). *J. Appl. Cryst.* **17**, 42-46.

JOHNSON, M. W. & DAVID, W. I. F. (1985). Report RAL-85-112. Rutherford Appleton Laboratory, Didcot, Oxon, England.

KARLE, J. (1968). *Acta Cryst.* B**24**, 182-186.

LOW, J. N. & WILSON, C. C. (1983). *Acta Cryst.* C**39**, 1688-1690.

MAIN, P. (1976). In *Crystallographic Computing Techniques*, edited by F. R. AHMED, K. HUML & B. SEDLACEK, pp. 97-105. Munksgaard: Copenhagen.

NELMES, R. J., TUN, Z., DAVID, W. I. F. & HARRISON, W. T. A. (1987). In *Chemical Crystallography with Pulsed Neutrons and Synchrotron Radiation*, edited by M. A. CARRONDO & G. A. JEFFREY, p. 479. Proc. of NATO ASI.

SEMMINGSEN, D., HOLLANDER, F. J. & KOETZLE, T. F. (1977). *J. Chem. Phys.* **66**, 4405.

TOLLIN, P. (1966). *Acta Cryst.* **21**, 613-614.

TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* **17**, 1322-1324.

WADSWORTH, J. W., WILSON, C. C. & DAVID, W. I. F. (1989). *Nucleosides Nucleotides*, **8**, 537-547.

WILSON, C. C. (1987). Report RAL-87-060. Rutherford Appleton Laboratory, Didcot, Oxon, England.

WILSON, C. C. (1988). Report RAL-88-087. Rutherford Appleton Laboratory, Didcot, Oxon, England.

WILSON, C. C., LOW, J. N. & TOLLIN, P. (1985). *Acta Cryst.* C**41**, 1123-1125.

WILSON, C. C. & TOLLIN, P. (1986). *J. Appl. Cryst.* **19**, 411-412.

WILSON, C. C. & TOLLIN, P. (1988). *Acta Cryst.* A**44**, 226-230.

Acta Cryst. (1989). A**45**, 839-851

Three-Beam and Many-Beam Theory in Electron Diffraction and its Use for Structure-Factor Phase Determination in Non-centrosymmetric Crystal Structures

BY J. M. ZUO

Department of Physics, Arizona State University, Tempe, AZ 85287, USA

R. HØIER

Department of Physics, University of Trondheim-NTH, 7034 Trondheim, Norway

AND J. C. H. SPENCE

Department of Physics, Arizona State University, Tempe, AZ 85287, USA

(*Received* 28 *February* 1989; *accepted* 1 *August* 1989)

Abstract

Closed-form expressions for three-beam dynamical transmission electron diffraction are compared. These are used as a guide to determine the best experimental conditions for the determination of structure-factor phases by convergent-beam electron diffraction in the general non-systematic case. The validity domains of Kambe's [*J. Phys. Soc. Jpn* (1957), **12**, 1-13] 'strong coupling' approximation and Bethe's [*Ann. Phys.* (*Leipzig*) (1928), **87**, 55-129] second approximation are compared, and these approximations reconciled. A comparison of many-beam calculations with experimental non-systematic CBED patterns is used to determine a three-phase invariant for CdS with an accuracy of ±5° in the electron structure-factor phase. If it is assumed that two of the phases are known exactly, the error in the third $(00\bar{2})$ X-ray structure-factor phase would be ±0·75°. The accuracy of the method for determining phases, atomic position parameters and bonding charge distributions is discussed.

1. Introduction

In the 60 years which have passed since the discovery of the diffraction of electrons by crystals there have been many attempts to extract crystal structure and bonding information from electron diffraction patterns [for a review, see Cowley (1981)]. Unlike X-rays, except for special cases, the much stronger interaction of electrons with matter renders the intensity of the multiply scattered electron beams very sensitive to the phases of crystal structure factors. The two-beam dynamical theory, however, does not preserve structure-factor phase information. In his classic study of three-beam dynamical electron diffraction, K. Kambe showed that the dynamical intensity depends on the sum of the phases of the three structure factors involved in the interaction, which is called the three-phase structure invariant (Kambe, 1957). Fourteen years later it was shown for the same non-systematic three-beam case that a degeneracy point exists at which the intensity is zero for centrosymmetric crystals (Gjønnes & Høier, 1971). The position

0108-7673/89/120839-13$03.00