L'auteur tient enfin à remercier ici d'une part tous les Cristallographes qui ont bien voulu lui communiquer leurs programmes pour les calculs et d'autre part le Professeur Raymond Weiss de l'Institut de Chimie de Strasbourg et ses collaborateurs avec l'aide desquels cette structure a été commencée.

Références

FAURE, R. & LOISELEUR, H. (1972). *Acta Cryst.* **B28**,811.
International Tables for X-ray Crystallography (1962). Vol. III. Birmingham: Kynoch Press.
LOISELEUR, H., THOMAS, G., CHEVRIER, B. & GRANDJEAN, D. (1967). *Chem. Commun.* p. 182.
MOORE, F. H. (1963). *Acta Cryst.* **16,** 1169.

Acta Cryst. (1972). B**28**, 820

Translation Functions and Direct Methods

BY JEROME KARLE

Naval Research Laboratory, Washington, D.C. 20390, *U.S.A.*

(*Received* 3 *June* 1971)

Translation functions are useful in direct methods of structure analysis when the initial result is a properly oriented but misplaced atomic grouping. Some alternative forms for translation functions are presented. In test calculations, it is shown that spurious indications may be eliminated and the reliability of the translation functions enhanced by basing the computations on several alternative placements of the atomic groupings. The computations appear to be often benefited by the use of approximately the outer one-half of the data rather than the full set.

Introduction

In applying direct methods of phase determination to the analysis of crystal structures, it has been found that the initial structural information obtained may, on occasion, not be complete, particularly in the case of noncentrosymmetric crystals. The locations of atoms obtained at first may be limited to only a portion of the structure. As has been shown in numerous investigations, knowledge of a partial structure can be readily developed into a complete structure by use of the tangent formula (Karle, 1968), so that obtaining only a fragment of the structure from the phase determination offers no essential difficulty.

A more serious problem arises when a properly oriented structural fragment is translated with respect to a permitted origin of the unit cell. Under such circumstances, it is usually not possible to develop the partial structure into a complete structure by means of the tangent formula without first placing it properly with respect to a permitted origin. The problem of determining the correct position of an oriented molecule, or part of a molecule, has been the subject of many investigations and several methods have been proposed: Vand & Pepinsky (1956), Hoppe (1957), Nordman & Nakatsu (1963), Huber (1965), Rossmann, Blow, Harding & Coller (1964), Tollin & Cochran (1964), Tollin (1966) and Crowther & Blow (1967). The last four investigations propose calculations involving the X-ray intensities, whereas the others are concerned with calculations on their Fourier transform, the Patterson function. In all cases, the methods depend upon the fact that the intensity data contain information concerning the relative locations of symmetry equivalents in the unit cell, and the extraction of this information permits the correct placement of an oriented molecular fragment with respect to a proper origin of the space group. Due to the common basis for the various translation analyses, there are many similarities among them which may not be immediately apparent, because of different notation or different reference points in the unit cell. For example, Tollin (1969) showed that a translation function of Crowther & Blow (1967) and the one proposed by Tollin (1966) are essentially the same.

There is a twofold purpose in writing this paper. The first is to call attention to the applicability of translation functions to direct methods of structure determination when a misplaced, oriented structure or partial structure is obtained. The second purpose is to suggest some alternative translation functions which are closely related to those of Tollin (1966) and Crowther & Blow (1967). The interpretation of the proposed translation functions is different from previous ones, namely, they represent a vector indicating twice the difference between a proper origin for the space group and the origin with respect to which the structure factor for the misplaced partial structure is calculated. The computations involved are simple to perform.

Various translation functions differ in the origins with respect to which the structure factors are computed and in the inclusion or absence of the complex conjugates for partial structure factors or of equivalent atomic groupings not associated with the symmetry element of interest. It is not obvious that such differences should materially affect the relative effectiveness of the various translation functions. Nevertheless, tests carried out indicate that some differences do occur which are due to variations in the maxima that cause ambiguous interpretations of the resulting Fourier maps.

Analysis

An oriented molecular fragment is arbitrarily placed in the unit cell, and structure factors are computed based on the fragment and its symmetry equivalents. A Patterson map computed from these structure factors would contain sets of peaks in the Harker sections which are similarly displaced from the properly placed corresponding peaks, as contained in the measured intensities. There would be the same displacement vector for each atom in the structural fragment and, given a calculation which could reveal these displacement vectors, it would be expected that a significant peak would be obtained owing to the reinforcement of the similar vectors. A suitable calculation is now proposed.

The coefficients of an origin-removed Patterson function, based on the observed normalized structure factors, are $(|E_{\mathbf{h}}|^2-1)$. A corresponding Patterson map in the same space group can be computed for a molecular fragment from the coefficients $(|E_{\mathbf{h},M}(S_1, \ldots, S_n)|^2-1)$, where the subscript M distinguishes structure factors computed from the atomic positions of a molecular fragment from those obtained from the observed data. The symbols S_i represent the n symmetry equivalents of the molecular fragment, *e.g.*, S_1 is the identity. A Patterson map containing corresponding sets of properly placed and misplaced sets of vectors characteristic of the Harker sections may be computed from:

$$P_1(\mathbf{r}) = \sum_{\mathbf{h}} \{(|E_{\mathbf{h}}|^2-1)+[|E_{\mathbf{h},M}(S_1, \ldots, S_n)|^2-1]\} \cos 2\pi\mathbf{h}\,.\,\mathbf{r}\,. \quad (1)$$

The differences between the corresponding sets of properly placed and misplaced vectors are contained in sections of a map computed from the square of the coefficients in equation (1):

$$P_2(\mathbf{r}) = \sum_{\mathbf{h}} \{(|E_{\mathbf{h}}|^2-1)+[|E_{\mathbf{h},M}(S_1, \ldots, S_n)|^2-1]\}^2 \cos 2\pi\mathbf{h}\,.\,\mathbf{r}\,. \quad (2)$$

By carrying out the square in equation (2), it is seen that the differences of interest are contained in the cross term, suggesting the calculation of appropriate sections of the difference function:

$$D(\boldsymbol{\delta}) = \sum_{\mathbf{h}} (|E_{\mathbf{h}}|^2-1)\,[|E_{\mathbf{h},M}(S_1, \ldots, S_n)|^2-1] \cos 2\pi\mathbf{h}\,.\,\boldsymbol{\delta}\,, \quad (3)$$

where the desired value for $\boldsymbol{\delta}$ would give the vector difference between the incorrectly placed vectors in a Harker section computed from the oriented molecular fragment, and the corectly placed vectors as contained in the observed intensity data. The vectors of interest in the Harker section are, of course, those between corresponding atoms in the molecular fragment and its symmetry equivalent. The value of $\boldsymbol{\delta}$ obtained from equation (3) can be immediately interpreted in terms of a translation of the fragment which would place it properly in the unit cell.

The translation function of Rossmann *et al.* (1964), $T(\boldsymbol{\Delta})$, is interpretable in terms of the translation required to bring one molecular unit into coincidence with another after suitable rotation. Their particular interest concerned independent molecules within the same asymmetric unit. The translation function of Tollin (1966), $Q(\mathbf{R}_0)$, determines a proper shift $\mathbf{R}_0$ for a set of atomic coordinates misplaced with respect to an appropriate origin by making use of the properties of the sum function for Patterson superposition. The $Q(\mathbf{R}_0)$ function is based on the sum function defined in terms of the initial atomic coordinates and a symmetry operation characteristic of the space group of interest. The translation function of Crowther & Blow (1967), $T(\mathbf{t})$, provides the translation vector between a molecule or a molecular fragment and an equivalent properly rotated unit appropriate to a particular symmetry element of the unit cell.

A variety of alterations of equation (3) might be considered. Following the suggestion of Crowther & Blow (1967) that intramolecular vectors can be eliminated from the observed data, $D(\boldsymbol{\delta})$ can be written

$$D_1(\boldsymbol{\delta}) = \sum_{\mathbf{h}} \Big\{\Big[(|E_{\mathbf{h}}|^2-1)-k\sum_{i=1}^{n}(|E_{h,M}(S_i)|^2-1)\Big] \times\Big[(|E_{h,M}(S_1, \ldots, S_n)|^2-1)-\frac{1}{n}\sum_{i=1}^{n}(|E_{h,M}(S_i)|^2-1)\Big]-\bar{C}_1\Big\} \cos 2\pi\mathbf{h}\,.\,\boldsymbol{\delta}\,, \quad (4)$$

where

$$k = \sum_{j=1}^{P} Z_j^2 \Big/ \sum_{j=1}^{N} Z_j^2\,, \quad (5)$$

p is the number of atoms in the molecular fragment, N is the number of atoms in the entire unit cell and n is the symmetry number of the space group. Note that N atoms contribute to $|E_{\mathbf{h}}|$, p atoms contribute to $|E_{\mathbf{h},M}(S_i)|$ and np atoms contribute to $|E_{\mathbf{h},M}(S_1, \ldots, S_n)|$. The factors k and $1/n$ arise in equation (4) from the need to have both terms in each of the square brackets on the same scale. The term $\bar{C}_1$ represents the average value of a coefficient obtained by summing $D_1(\boldsymbol{\delta})$ at the origin and dividing by the number of terms. It removes the large peak at the origin and its possible interference with maxima close by.

Instead of including all the symmetry elements at once, a pair representing the molecular fragment and one symmetry equivalent could be considered, giving

$$\begin{aligned} D_2(\boldsymbol{\delta}) = \sum_{\mathbf{h}} \{ & [(|E_{\mathbf{h}}|^2-1) - k \sum_{i=1}^{n} (|E_{\mathbf{h},M}(S_i)|^2-1)] \\ & \times [(|E_{\mathbf{h},M}(S_1,S_2)|^2-1) \\ & - \tfrac{1}{2} \sum_{i=1}^{2} (|E_{\mathbf{h},M}(S_i)|^2-1)] - \bar{C}_2 \} \cos 2\pi \mathbf{h} \,.\, \boldsymbol{\delta} \,. \quad (6) \end{aligned}$$

An alternative formulation eliminates the centrosymmetric character of the structure factors computed from the molecular fragment, giving:

$$\begin{aligned} D_3(\boldsymbol{\delta}) = \sum_{\mathbf{h}} \{ & [(|E_{\mathbf{h}}|^2-1) - k \sum_{i=1}^{n} (|E_{\mathbf{h},M}(S_i)|^2-1)] \\ & \times [E_{\mathbf{h},M}(S_1) E^*_{\mathbf{h},M}(S_2)] - \bar{C}_3 \} \exp(-2\pi i \mathbf{h} \,.\, \boldsymbol{\delta}) \,. \quad (7) \end{aligned}$$

The functions $D_1(\boldsymbol{\delta})$ and $D_3(\boldsymbol{\delta})$ are comparable to the functions $T_2(\mathbf{t})$ and $T(\mathbf{t})$, respectively, of Crowther & Blow (1967). The distinction between them lies mainly in the nature of the calculated structure factors and the subsequent interpretation of the results. The function $D_3(\boldsymbol{\delta})$ also bears a close relationship to the work of Tollin & Cochran (1964) and Tollin 1966).

The calculation of $D_1(\boldsymbol{\delta})$ is carried out on appropriate sections of the Patterson space group corresponding to the structure, and $D_2(\boldsymbol{\delta})$ and $D_3(\boldsymbol{\delta})$ are computed on appropriate sections of space groups $P\bar{1}$ and $P1$, respectively. With all these functions, when a correct value for $\boldsymbol{\delta}$ is determined, the proper shift to apply to the misplaced fragment is $-\boldsymbol{\delta}/2$.

Test calculations

Test calculations have been performed with equations (4) and (7). In one example, they were based on a misplaced structural fragment, which appeared as one of four ambiguous possibilities in a structure determination by the symbolic addition procedure of a radiation product of mescaline (Karle & Karle, 1970). A properly placed partial structure was also obtained as another of the ambiguous possibilities and had been used in the original structure determination. The correct structure in space group *Aba*2 is shown in Fig. 1 and the misplaced partial structure comprising 9 of the total of 20 independent atoms is shown in Fig. 2.

The first test was carried out with equation (4). In space group *Aba*2, the position of the origin along the *c* axis is arbitrary. A displacement of a structural fragment with respect to a proper origin would involve a possible shift with respect to the *a* or *b* axes or both. Equation (4) is computed in the appropriate Patterson space group, *Ammm* in this case. The $z=0$ section is of interest, and the proper interpretation of the maxima in this section may be derived from a consideration of the differences of corresponding maxima in the Harker sections of a Patterson function, which would also be computed in space group *Ammm* for the correctly placed and misplaced structural fragments, equation (1). The results of such an analysis are summarized in Table 1. Since an *A* centered cell is being considered, identical values would be obtained for all points shifted by $0, \frac{1}{2}, \frac{1}{2}$, but this is of no significance in the interpretation of the displacement of the structural fragment. It is apparent that for equation (4), the symmetry of the 2-dimensional space group *pmm* generates a fourfold ambiguity. Additional ambiguities may arise from spurious peaks in the calculation of equation (4). The ambiguities arising both from the symmetry of the plane group and from the spurious peaks may be resolved by making more than one computation changing the location of the molecular fragment in each. Since the location of the fragment is known for each computation, the change of location of the main maxima in $D_1(\boldsymbol{\delta})$ may be interpreted in terms of a proper placement of the partial structure in the unit cell.

A computation of $D_1(\boldsymbol{\delta})$, based on the structural fragment of the rearrangement product of mescaline, Fig. 2,

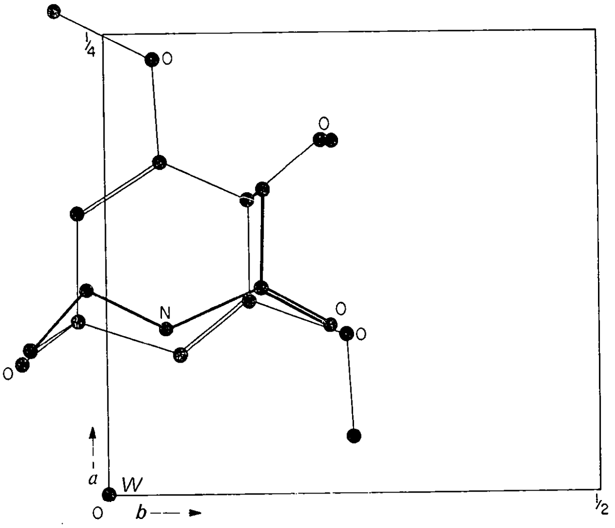

Fig. 1. Structure of photorearrangement product of mescaline.

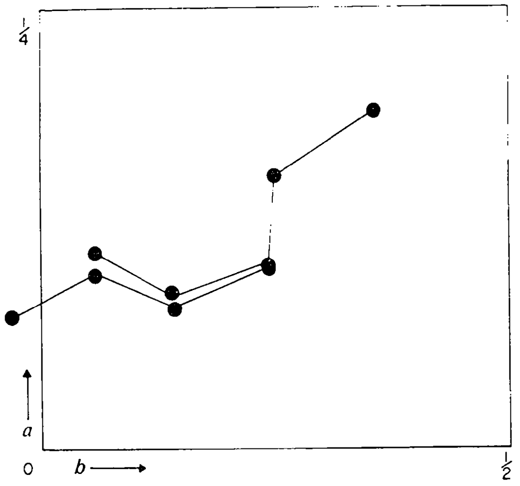

Fig. 2. Misplaced, but properly oriented, partial structure of photorearrangement product of mescaline.

indicated the correct shift among several other possibilities as implied by the interpretation of equation (4) given in Table 1. The additional spurious maxima interfered with the attainment of a unique result. It was found that some improvement in the quality of the map was effected by employing only the outer half of the data in the calculation. Three additional arbitrary shifts of the nine-atom fragment were made and $D_1(\boldsymbol{\delta})$ was computed for each. An analysis of the three maps eliminated most of the spurious peaks and resulted in the correct shift dominating in the maps, as measured by the sum of the values of the corresponding peaks appearing in each map. Additional trial shifts could be expected to enhance the correct answer even further.

Table 1. *Positions of significant peaks for equations* (4), (6), *and* (7) *in terms of the correct coordinates* x, y, *as contained in the observed data* $|E_{\mathbf{h}}|$, *and the misplaced coordinates* x', y', *as contained in* $E_{\mathbf{h},M}$, *given a structure in space group Aba*2

The contributors to Δx, for example, are sets of $(x'-x)$ in which the individual values of the primed and unprimed coordinates vary, but the differences $(x'-x)$ are constant and give a reinforced contribution to the maxima representing $\boldsymbol{\delta}$ or its individual components.

Equation	Components of $\boldsymbol{\delta}$				2-Dimensional space group
(4)	$2\Delta x$	$2\Delta y$	0	plus *mm* symmetry for each set of components	*pmm*
	$2\Delta x$	0	0		
	0	$2\Delta y$	0		
(6)	$2\Delta x$	$2\Delta y$	0		*p*2
	$-2\Delta x$	$-2\Delta y$	0		
(7)	$2\Delta x$	$2\Delta y$	0		*p*1

The actual and observed shifts of the fragment are listed in Table 2. Relating the second and third shifts to the first, and taking the average, would imply a shift of 24/120, 9·2/120 as compared to the correct value of 24/120, 9/120. This corresponds to an error of 0·00, 0·02 Å, respectively, as computed from the lattice parameters.

Table 2. *Location of peaks associated with three different shifts of the nine-atom fragment of a mescaline photoproduct*

Maps were computed with equation (4) in space group *Ammm* appropriate to the space group *Aba*2 for the crystal. The intervals are given in $^1/_{120}$.

Expected	Observed	Observed	Observed
$(2\Delta x, 2\Delta y)$	$(2\Delta x, 0)$	$(0, 2\Delta y)$	$(2\Delta x, 2\Delta y)$
48, 18	48·4, 0	0, 17·5	48·8, 18·0
24, 8·4	23·9, 0	0, 9·2	22·5, 7·5
−24, 6	−24·0, 0	0, 7·9	−23·2, 8·2

Function $D_3(\boldsymbol{\delta})$ was also computed with the nine-atom fragment of the mescaline derivative. Using the complete set of data, the largest peak obtained with a value of 638 was spurious. The second largest maximum at 568 was the correct peak. There were two additional spurious maxima at 566 and 557 and six above 475. Using the outer half of the data, the correct peak was the largest one at 261, with the nearest spurious one at 218, and only one other above 200. The spurious peak with a value of 638, obtained from the entire set of data, had the relatively low value of 135 in the second calculation based on the outer half of the data.

Another test calculation was carried out on a misplaced molecular fragment consisting of 8 of the total of 15 atoms in a molecule of cyclo-L-prolyl-L-leucyl (Karle, 1972). The substance crystallizes in space group $P2_12_12_1$. Calculations were made using equation (4) in the three 2-dimensional sections of space group *Pmmm* whose maxima could afford information concerning the components of $\boldsymbol{\delta}$ in each section: $2\Delta x$, $2\Delta y$, 0; $2\Delta x$, 0, $2\Delta z$; and 0, $2\Delta y$, $2\Delta z$, where Δx, for example, represents the difference of sets of x coordinates for corresponding atoms in the misplaced and properly placed structural fragments. A useful redundancy occurs, inasmuch as the shift in a particular coordinate appears in two of the sections, and facilitates the elimination of spurious peaks. Even when a suitable maximum is chosen to represent the shift, it occurs with an eightfold ambiguity due to the *mmm* symmetry. Ambiguities arising both from symmetry and spurious peaks may be resolved by making additional computations, as mentioned before, changing the location of the molecular fragment in each.

Table 3. *Positions of the main maxima and their values in the three sections computed with the outer half of the experimental data for equation* (4) *in space group Pmmm, appropriate to cyclo-*L*-prolyl-*L*-leucyl crystallizing in space group* $P2_12_12_1$

The partial structure, consisting of 8 of the 15 nonhydrogen atoms, was shifted by 5/60, 5/60, 5/60, so that the result expected from equation (4) is 10/60, 10/60, 10/60. The average result derived from the sections is 9·4/60, 10·5/60, 11·1/60. The strong peak in the $y=0$ section is an extended ellipse with high values of 137 and 160 at $z=10$ and 11, respectively, diminishing the apparent discrepancy in the shift of the z coordinate.

Section	Main maxima (× 1/60)	Value
$x=0$	0, 10·6, 9·2	123
	0, 15·0, 24·5	123
	0, 14·5, 14·5	120
	0, 22·0, 4·5	107
	0, 7·0, 13·5	104
$y=0$	9·4, 0, 13·0	179
	21·6, 0, 21·0	120
$z=0$	9·3, 10·4, 0	151
	18·0, 3·7, 0	142
	21·5, 21·5, 0	140
	3·0, 26·0, 0	129

The test calculation on the cyclic dipeptide was carried out with the outer half of the measured reflections. Use of the complete set of data with equation (4) resulted in maps which had many spurious maxima. A significant improvement in the quality of the results was obtained when the inner half of the data was omitted. The test shift of the 8-atom fragment was 5/60, 5/60, 5/60, and the results of computing appropriate sections with equation (4) are shown in Table 3. The shift implied by these results is 4·7/60, 5·3/60, 5·6/60 which would be in error by 0·05, 0·10, 0·06 Å, respectively, as computed from the lattice parameters. In practice, at least one additional calculation with a shift of position for the molecular fragment would be required to eliminate the ambiguity due to the mirror symmetry of space group *Pmmm*.

Function $D_3(\boldsymbol{\delta})$ was computed for the eight-atom fragment of the cyclic dipeptide in the $z=0$ plane to test a shift in the x and y coordinates. The map, using the complete set of data, gave the correct maximum the highest value of 680. The next largest maximum occurred at 506. With the outer half of the data, the highest value at 358 again was associated with the correct peak. The next largest peak occurred at 285. In this example, the relative difference in the values of the peaks favors, somewhat, the use of the entire set of data.

Concluding remarks

It is apparent from the tests which have been carried out and the previous experience of other authors (Young, Tollin & Sutherland, 1968; Young, Tollin & Wilson, 1969) that translation functions can play a useful role in direct methods when the phase determination produces an oriented molecular fragment improperly placed in the unit cell. It may be preferable to employ such functions rather than to repeat the phase determination following a different path through the stepwise development of phases. Once the oriented fragment has been properly placed, the complete structure may be developed by means of the tangent formula (Karle, 1968).

The test calculations do not indicate that either $D_3(\boldsymbol{\delta})$ or $D_1(\boldsymbol{\delta})$ is clearly preferable. The computation of $D_1(\boldsymbol{\delta})$ is somewhat simpler to perform, but the maps contain ambiguities because of symmetry which requires additional calculation for resolution. It is advantageous to have a variety of translation functions available to help identify and eliminate the spurious peaks which may interfere with the correct interpretation of the maps. Spurious peaks are also eliminated by making several trial shifts of the molecular fragment. This removes, in addition, the ambiguities which arise with $D_1(\boldsymbol{\delta})$ and $D_2(\boldsymbol{\delta})$ because of the higher symmetry of the plane groups in which the calculations are performed. The quality of the maps is improved by removing the large maximum at the origin and its accompanying ripple, and by removing the intramolecular vectors. In several tests, it was advantageous to compute the maps using approximately the outer half of the experimental data.

I wish to thank Mr S. Brenner for his fine cooperation in writing the computer programs and carrying out the computations.

References

Crowther, R. A. & Blow, D. M. (1967). *Acta Cryst.* **23**, 544.
Hoppe, W. (1957). *Z. Elektrochem.* **61**, 1076.
Huber, R. (1965). *Acta Cryst.* **19**, 353.
Karle, I. L. (1972). *J. Amer. Chem. Soc.* In the press.
Karle, I. L. & Karle, J. (1970). *Acta Cryst.* B**26**, 1276.
Karle, J. (1968). *Acta Cryst.* B**24**, 182.
Nordman, C. E. & Nakatsu, K. (1963). *J. Amer. Chem. Soc.* **85**, 353.
Rossmann, M. G., Blow, D. M., Harding, M. M. & Coller, E. (1964). *Acta Cryst.* **17**, 338.
Tollin, P. (1966). *Acta Cryst.* **21**, 613.
Tollin, P. (1969). *Acta Cryst.* A**25**, 376.
Tollin, P. & Cochran, W. (1964). *Acta Cryst.* **17**, 1322.
Young, D. W., Tollin, P. & Sutherland, H. (1968). *Acta Cryst.* B**24**, 161.
Young, D. W., Tollin, P. & Wilson, H. R. (1969). *Acta Cryst.* B**25**, 1423.
Vand, V. & Pepinsky, R. (1956). *Z. Kristallogr.* **108**, 1.