Acta Cryst. (1973). **B29**, 404

Application of Translation Functions to the Direct Methods Solution of an Unknown Structure

By L. T. J. Delbaere and M. N. G. James

Department of Biochemistry, University of Alberta, Edmonton, Alberta, Canada

(*Received* 27 *July* 1972; *accepted* 20 *October* 1972)

The direct methods solution of the crystal structure of 2,5′-anhydro-2′,3′-isopropylidenecyclouridine initially located a portion of the molecule in an incorrect position in the unit cell. The vector shift required to move the partial structure to the correct position was determined by use of a difference function proposed by Karle [*Acta Cryst.* (1972), **B28**, 820–824].

The attempted solution of the crystal structure of 2,5′-anhydro-2′,3′-isopropylidenecyclouridine, $C_{12}H_{14}N_2O_5$ (Fig. 1) by direct methods led to the location of a portion of the molecule in the unit cell correctly oriented with respect to the crystallographic axes but incorrectly positioned with respect to the crystallographic symmetry elements. The vector displacement of this portion of the molecule from its actual position in the unit cell was determined by use of the difference function which was proposed by Karle (1972).

The colorless crystals of 2,5′-anhydro-2′,3′-isopropylidenecyclouridine were supplied by Dr R. U. Lemieux of the Department of Chemistry, University of Alberta. The space group of these crystals is $P2_1$ with cell dimensions $a=10{\cdot}344$, $b=6{\cdot}407$, $c=9{\cdot}077$ Å and $\beta=93{\cdot}69°$. The origin was specified by the following reflections:

h	k	l	E	phase
5	0	7	2·5	0
$\bar{2}$	0	5	2·4	0
$\overline{10}$	1	1	2·6	0

Symbols were assigned to four additional reflections and including these along with the origin set, symbolic addition by hand (Karle & Karle, 1966) on a $\sum_2$ listing was carried out using the relationship

$$\varphi_{\mathbf{h}} \approx \langle \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}} \rangle_{k_r} . \tag{1}$$

By making use of the most commonly occurring relationships between symbols four possible solutions were obtained. One of these symbols was given values of $\pi/4$ and $3\pi/4$ radians, thus fixing the enantiomorph. The four starting sets of seven reflections were each used as input into the phase refinement procedure using the tangent formula (Karle & Hauptman, 1956),

$$\tan \varphi_{\mathbf{h}} = \frac{\sum_{k} |E_{\mathbf{k}} \,.\, E_{\mathbf{h}-\mathbf{k}}| \sin (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}{\sum_{k} |E_{\mathbf{k}} \,.\, E_{\mathbf{h}-\mathbf{k}}| \cos (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})} . \tag{2}$$

Almost all of the 150 reflections with $E_{\min} \geq 1{\cdot}6$ had their phases so determined. The R_{Karle} values (Karle & Karle, 1966) ranged from 0·178 to 0·201 for these four solutions. Phase relationships involving the three principal zonal reflections are often unreliable, probably due to overlap of atoms in projection (Karle,

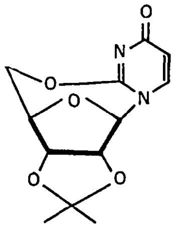


Fig. 1. 2,5′-Anhydro-2′,3′-*O*-isopropylidene cyclouridine.

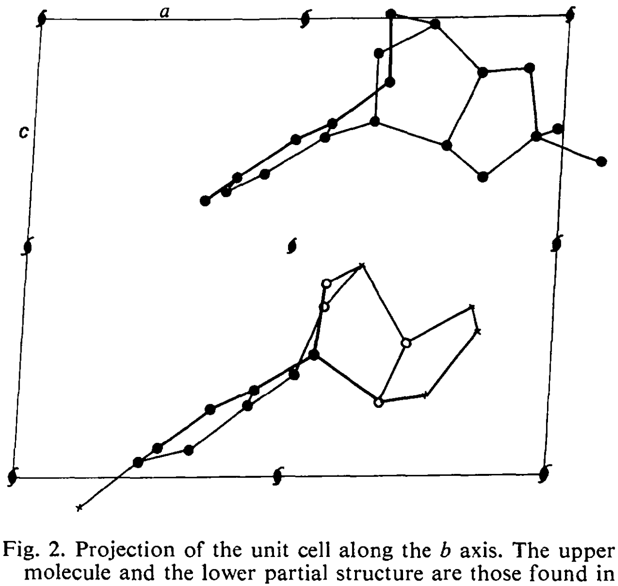


Fig. 2. Projection of the unit cell along the b axis. The upper molecule and the lower partial structure are those found in the E maps for the $R_{\text{Karle}}=0{\cdot}150$ and 0·193 solutions respectively.

1970). Considerable overlap of atoms in projection was probable for the cyclonucleoside; thus such relationships were removed from the $\sum_2$ listing and from the tangent-formula refinement. An examination of the phase values for each solution indicated some tendencies among groups of reflections with the same k index to have approximately the same phase value or π radians from the same phase value. The solution having $R_{\text{Karle}}=0{\cdot}193$ had the most random distribution of phase values and an E map calculated for this solution contained a portion of the molecule (see Fig. 2, lower part); E maps calculated for the other three solutions were not chemically meaningful. The peaks on this map corresponding to the atoms of the uracil ring were well resolved, whereas peaks occurring off the plane containing this ring were diffuse. Nevertheless eleven additional peaks forming a chemically reasonable unit were chosen from the remaining maxima in the map. A structure factor calculation (all of the atoms were given the scattering factor of carbon) produced an R index of 0·516. Attempted least-squares refinement of the structure was unsuccessful. A portion of the molecule had apparently been located in an incorrect position in the unit cell.

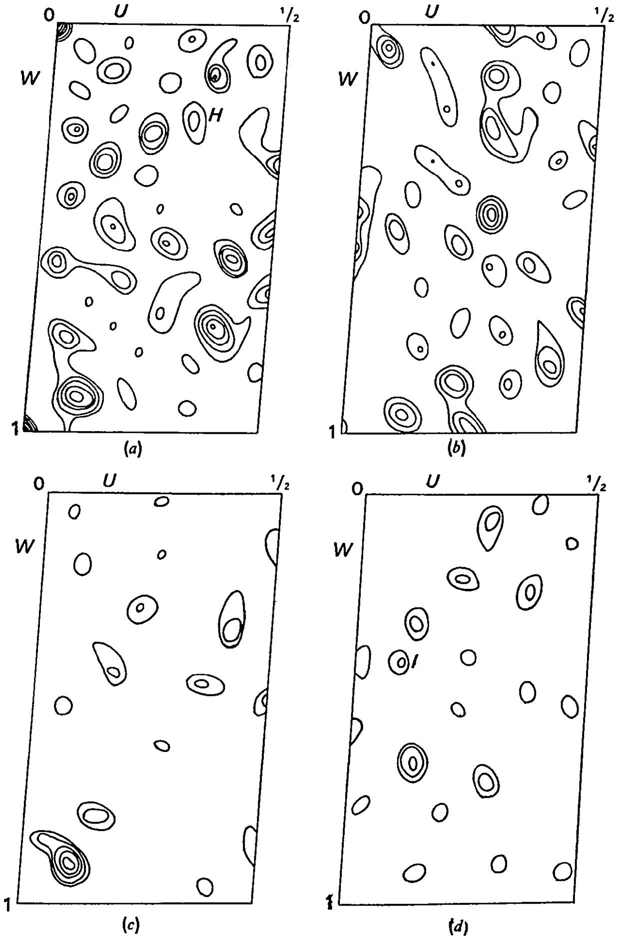


Fig. 3. (a) $v=0$ section of the Patterson function calculated with $(|E_{\mathbf{h}}|^2-1)$ as coefficients. (b) $v=\frac{1}{2}$ section of the Patterson function calculated with $(|E_{\mathbf{h}}|^2-1)$ as coefficients. (c) $v=0$ section of the Patterson function calculated with $(|E_{\mathbf{h},M}(S_1,\ldots S_n)|^2-1)$ as coefficients. (d) $v=\frac{1}{2}$ section of the Patterson function calculated with $(|E_{\mathbf{h},M}(S_1,\ldots S_n)|^2-1)$ as coefficients.

The vector displacement required to shift the molecule to its correct position was determined by the calculation of the following difference functions (Karle, 1972):

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

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

where $(|E_{\mathbf{h}}|^2-1)$ are the observed normalized structure amplitudes; $(|E_{\mathbf{h},M}(S_1,\ldots S_n)|^2-1)$ are the normalized structure amplitudes computed for the partial structure; S_i represents one of the n symmetry equivalents of the space group.

The $v=0$ and $v=\frac{1}{2}$ sections of the Patterson function computed with $(|E_{\mathbf{h}}|^2-1)$ as coefficients are given in Fig. 3(a) and (b) respectively; the corresponding sections of the Patterson function with $[|E_{\mathbf{h},M}(S_1,\ldots S_n)|^2-1]$ as coefficients are in Fig. 3(c) and (d) respectively. The $v=0$ sections of the functions calculated from equations (3) and (4) are in Fig. 4(a) and (b) respectively and Fig. 4(c) is the $v=\frac{1}{2}$ section calculated from equation (3). It is of importance to note the relative peak population on the $v=0$ and $v=\frac{1}{2}$ sections of the $(|E|^2-1)$ Patterson map. Even though the correct space group is $P2_1$ the $v=0$ section is more densely populated with vector peaks than is the Harker section ($v=\frac{1}{2}$).

An atom at (x,y,z) in the partial structure is located at $(x+m,y,z+p)$ in the actual structure, where m and p are the shifts required in the x and z directions in order to move the atom in the partial structure to its correct position. The difference between the Harker peak $(2x+2m,\frac{1}{2},2z+2p)$ for an atom in the actual structure and its Harker peak $(2x,\frac{1}{2},2z)$ in the partial structure is $(2m,0,2p)$. These differences for various atoms should contribute to the same peak in the $v=0$ section of the difference functions calculated from either equation (3) or (4). This peak is denoted by G in Fig. 4(a) and (b). Since there are special relationships between the locations of atoms in the structure, a similar peak, denoted by J in Fig. 4(c), is also present on the $v=\frac{1}{2}$ section calculated from equation (3). This is due to pairs of atoms having approximately the same x and z coordinates with the y coordinates differing by $\frac{1}{2}$. The vector peak between one such atom and the symmetry equivalent of the other member of the pair is $(2x+2m,0,2z+2p)$. The Harker peak for these atoms is $(2x,\frac{1}{2},2z)$ and the difference between these two peaks is $(2m,\frac{1}{2},2p)$. A pair of oxygen atoms in the structure satisfies this condition. The peak H in Fig. 3(*a*) and the peak I in Fig. 3(*d*) correspond to the peaks on the $v=0$ and $v=\frac{1}{2}$ sections of the Patterson functions and their difference is the peak J [Fig. 4(*c*)]. That the peak J is due to intramolecular vectors is shown by its absence on the $v=\frac{1}{2}$ section calculated from equation (4). It is interesting to note that the strongest and best resolved of the peaks due to the correct vector shift occurs on the $v=\frac{1}{2}$ section of the difference function calculated from equation (3). The correct shift which was obtained from this peak is $0{\cdot}112x/a+0{\cdot}394z/c$.

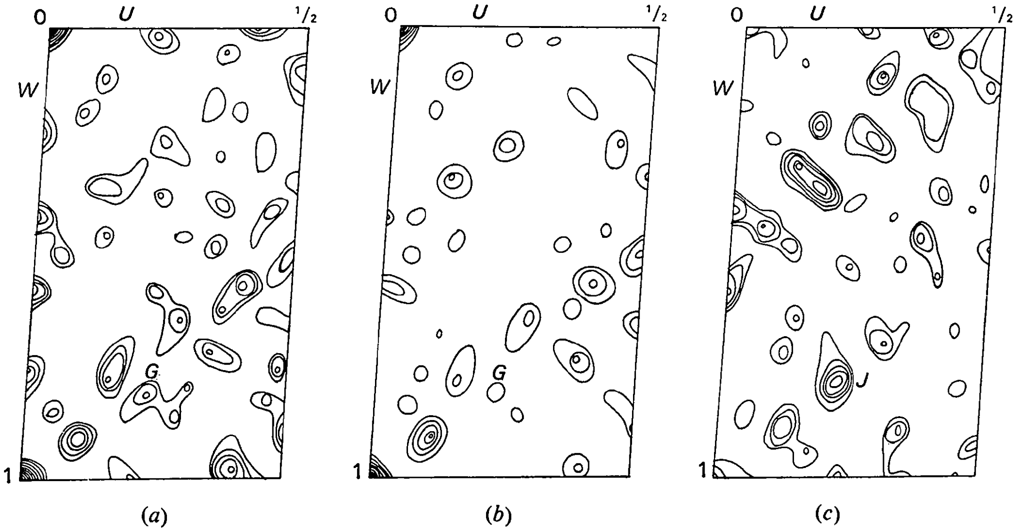
This apparently fortuitous appearance of the correct shift peak on the $v=\frac{1}{2}$ section computed using equation (3) is thus a direct result of this particular structure which contains considerable overlap in the (010) projection. The correct shift peak is present on the $v=0$ sections as seen in Fig. 4(*a*) and (*b*) but its relatively small peak height would have precluded solution of the structure for a considerable time.

The initial R index on the shifted fragment was 0·511 (all atoms considered as carbon). The phases of fifty two reflections having a minimum E value of 1·7 and also a calculated structure amplitude for this shifted moiety of at least one-half of the observed structure amplitude were used as the starting set in the tangent formula (2). The phase values of the starting set were fixed for the first cycle (five iterations per cycle) and two further cycles generated the phases of 174 of the 176 reflections having a minimum E value of 1·55. The R_{Karle} was 0·150 for this solution. The strongest peaks on the resulting E map correspond to all of the non-hydrogen atoms of the structure. Initial structure-factor calculations with all of these atoms denoted as their correct species, produced an R index of 0·187. After three cycles of full-matrix least-squares refinement, the hydrogen atoms were located from a difference Fourier synthesis. Two further cycles of refinement on the parameters of all of the atoms (isotropic temperature factors) produced the present R index of 0·078.

Fig. 2 is a projection of the unit cell along the b axis. The upper molecule is the correctly located molecule as found in the E map based on the solution with $R_{\text{Karle}}=0{\cdot}150$. The lower partial structure represents those features found in the E map based on the $R_{\text{Karle}}=0{\cdot}193$ solution – the solid circles denote atoms which, after the vector shift, are in correct positions, open circles signify those which, after the correct displacement, have an incorrect y coordinate and the ×'s are the peaks which, when shifted, do not correspond to atoms in the actual structure.

The tangent-formula refinement with the three two-dimensional reflection phase relationships removed calculated 64 phase values within 10 degrees of their final values whereas the same starting set with all the triple relationships included in the tangent formula, produced only 54 phases which had values within 10 degrees of their final values. In the crystal structure of 2,5′-anhydro-2′,3′-*O*-isopropylidenecyclouridine, there is considerable overlap of atoms in projection, particularly along the b axis of the unit cell. Thus, in such a case, the removal of the three sets of zonal reflections *hk*0, *h*0*l*, 0*kl* from the tangent formula generates more accurate phase values.

We would like to thank the referee for pointing out our initial misinterprctation of the difference Patterson



(*a*) (*b*) (*c*)

Fig. 4. (*a*) $v=0$ section of the difference function calculated from equation (3). (*b*) $v=0$ section of the difference function calculated from equation (4). (*c*) $v=\frac{1}{2}$ section of the difference function calculated from equation (3).

maps computed with equation (3). The NRC Crystallographic programs (Ahmed, Hall, Pippy & Huber, 1966) and *ORFLS* were used in the computations. The data were collected on a Picker FACS-1 System with molybdenum radiation (graphite monochromator). This work was supported in part by grant MA-3406 to M.N.G.J. from the Medical Research Council of Canada and in part by the National Research Council of Canada grant (A-172) to R.U. Lemieux for support to L.T.J.D.

References

AHMED, F. R., HALL, S. R., PIPPY, M. E. & HUBER, C. P. (1966). N.R.C. Crystallographic Programs for the IBM/360 System, *World List of Crystallographic Computer Programs*, 2nd edition. Appendix p. 52.

KARLE, I. L. (1970). *Crystallographic Computing*, p. 25. Edited by F. R. AHMED. Copenhagen: Munksgaard.

KARLE, J. (1972). *Acta Cryst.* **B28**, 820–824.

KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.

Acta Cryst. (1973). B**29**, 407

A New Direct Method for Characterizing Structures with Stacking Faults, Built up from Translationally Equivalent Layers.

I. Faults in Stackings of Three and Four Layers

BY M. FARKAS-JAHNKE

Research Institute for Technical Physics of the Hungarian Academy of Sciences, Budapest, Hungary

(*Received* 5 *September* 1972; *accepted* 8 *October* 1972)

The direct method for determination of the stacking sequences of periodic polytypes has been successfully applied to the calculation of structural characteristics of stacking-faulted lattices built up of translationally equivalent layers. From the intensity distribution along row-lines of indices $h-k \neq 3n$ on oscillational X-ray patterns $\pi'(m,p)$ sets were calculated, which give the relative rate of occurrence of the related stacking vectors. Formulae are derived for calculating cyclicity, hexagonality and the relative rate of occurrence of four-layer stackings using the $\pi'(m,p)$ values. It is shown that this method may be used also to determine directly the α and β fault parameters used by Jagodzinski.

Introduction

Because of the practical importance of materials with structures built up of translationally equivalent layers, several theories and methods have been worked out since the early days of X-ray diffraction methods to make possible the characterization of their faulted structures (Warren, 1941; Hendricks & Teller, 1942; Gevers, 1952, 1954; Kakinoki & Komura, 1952; Paterson, 1952; Johnson, 1963; Allegra, 1964; Sato, 1966, 1969; Kakinoki, 1967; Lele, Anantharaman & Johnson, 1967; Holloway, 1969; Lele, 1969; Lele, Prasad & Anantharaman, 1969; Lele & Rama Rao, 1970; Prasad & Lele, 1971). These methods, however, are all indirect, assuming random distribution of stacking faults. Moreover, most theories suppose the presence of only one type of stacking fault in the lattice. One method without the latter restriction is that of Jagodzinski (1949*a*, *b*, *c*). Assuming a random stacking-fault distribution and an interaction range of three interlayer spacings ('Reichweite = 3'), he determined the effect of this type of disorder on the intensity distribution of scattered X-rays. His method was partly based on earlier work of Landau (1937) and Wilson (1942). Jagodzinski's two-parameter model was further developed and applied to some practical cases, mostly for the characterization of faulted ZnS structures by Müller (1952) and Singer & Gashurow (1963), who achieved, however, a reasonably good fit between calculated and photometrically measured intensity curves for random stacking fault distribution only. Even in these cases the use of the method was rather tiresome since, being an indirect method the calculation of a set of master curves was needed to find the best fitting curve whose parameters may be characteristic of the structure.

We encountered the problem of characterizing lattices with stacking faults when investigating the structure of a great number of ZnS crystals. These crystals have been widely investigated because of their interesting polymorphic modifications. But, as has been shown by many authors (Müller, 1952; Brafman, Shachar & Steinberger, 1965; Verma & Krishna, 1966), these crystals (the natural ones and also those grown by different methods) only seldom have a completely regular structure. Besides the high-temperature hexagonal, the low-temperature cubic, and the numerous polytype modifications, the structure of the majority of the crystals contains many stacking faults. In such regions the Zn–S double layers of hexagonal symmetry are stacked perpendicularly to the hexagonal *c* axis (*i.e.* the cubic [111] direction) in such a manner that neighbouring