*Acta Cryst.* (1975). A**31**, 543

# Translation Vector Functions Based on a Deconvolution of the Patterson Function Provided by Transform Methods

BY DAVID A. LANGS

*Medical Foundation of Buffalo, 73 High Street, Buffalo, New York 14203, U.S.A.*

(*Received 9 October* 1974; *accepted 24 February* 1975)

In this paper a translation function is derived for which there is a significant reduction in the intensity of the spurious maxima often experienced with the use of alternative methods. An analysis of partial structures indicates that the remaining atoms of a crystal structure may, in principle, be obtained directly from the translation function derivation by Fourier synthesis.

## 1. Introduction

A number of techniques exist for the solution of crystal structures containing molecules whose conformation is in part unambiguously defined by a rigid arrangement of covalent bonds. Many crystal structure analyses may not initially be suitable for treatment by these methods since it is not always possible *a priori* to describe a reference structure. It is known that direct methods, which use no *a priori* structural information, often produce a correctly oriented, but translationally displaced portion of the structure, and Karle (1972) has recently called attention to the applicability of translation functions in augmenting direct-methods solutions of this type.

The methods of crystal structure analysis which incorporate *a priori* structural information may be conveniently divided into four groups: 1. discrepancy functions based on the trigonometric form of the structure factor expressed in terms of an unknown displacement vector (Booth, 1948; Vand & Pepinsky, 1956; Taylor & Morley, 1959); 2. translation and rotation function based on the convolution of the square of the structure amplitudes between the real and reference structures (Tollin & Cochran, 1964; Tollin, 1966; Crowther & Blow, 1967; Karle, 1972); 3. analogous orientation and displacement functions which operate on the Fourier transforms of these intensities (Hoppe, 1957; Nordman & Nakatsu, 1963; Huber, 1965; Braun, Hornstra & Leenhouts, 1969); 4. algebraic procedures which yield probabilistic estimates of crystal structure invariants given the molecular conformation (Hauptman, 1964) or conformation and orientation (Kroon & Krabbendam, 1970) of the structures of the crystallographic asymmetric unit.

Presented herein is a translation function which is basically the Patterson function of the deconvoluted molecular structure. It will be shown that the magnitude and location of spurious maxima in translation vector syntheses are determined by the molecule structure and the particular translation function used. The magnitude of spurious peaks produced by the phase-modulated translation function described in this paper will be shown to be half that produced by translation functions which are both phase and amplitude modulated. It will also be shown that the translation functions, when applied to correctly positioned partial structures, can aid in the location of missing structural fragments *via* a substitution of the coefficients involved in Fourier synthesis.

## 2. Molecular transform notation

Given that the orientation of a chemical structure or molecule has been deduced, one may define its transform in terms of its coordinates relative to an arbitrarily chosen origin as

$$F_{\mathbf{h}_0}=\sum_{j=1}^{N} f_j \exp(2\pi i\mathbf{h}\,.\,\mathbf{r}_{j_0})\,, \tag{1}$$

where $\mathbf{r}_{j_0}$ is the position vector of the $j$th atom with respect to the arbitrary origin. $F_{\mathbf{h}_0}$ is the value of the continuous molecular transform sampled at the reciprocal-lattice point $\mathbf{h}$. This expression is similar to the Ewald transform (Ewald, 1935) of a spatial array of atoms, but differs in that the origin of the transform need not coincide with its true location in the unit cell. Define by $\mathbf{r}_0$ the vector describing the position of the arbitrary origin with respect to a crystallographic origin and by $\mathbf{r}_j$ the position vector of the $j$th atom with respect to that origin. Then

$$\mathbf{r}_j=\mathbf{r}_0+\mathbf{r}_{j_0}\,. \tag{2}$$

It follows that the structure factor, which is

$$F_h=\sum_{j=1}^{N} f_j \exp[2\pi i\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r}_{j_0})]\,, \tag{3}$$

is the product of $F_{\mathbf{h}0}$ and a phase factor which contains the components of the translation vector, $\mathbf{r}_0$, necessary to bring the initial (reference) structure into coincidence with its image correctly placed in the unit cell,

$$F_{\mathbf{h}}=F_{\mathbf{h}_0}\exp(2\pi i\mathbf{h}\,.\,\mathbf{r}_0)\,. \tag{4}$$

If we choose to consider the symmetry of a particular space group, it may be shown that the structure factor can be expressed as a summation of terms, each a product of two factors. The first factor is the transform of the molecule evaluated at the reciprocal-lattice point $hkl$ (or one of its symmetry-equivalent lattice points defined by the crystallographic point group), the second is a phase factor which contains the components of the translation vector. Thus, for example, one finds that the structure amplitude for the space group $P\bar{1}$ may be written as

$$\begin{aligned} F_{hkl} &= F_{hkl_0} \exp\left[2\pi i(hx_0+ky_0+lz_0)\right] \\ &\quad + F_{\bar{h}\bar{k}\bar{l}_0} \exp\left[-2\pi i(hx_0+ky_0+lz_0)\right] \\ &= 2|F_{hkl_0}| \cos\left[2\pi(hx_0+ky_0+lz_0)+\varphi_{hkl_0}\right], \end{aligned} \tag{5}$$

where

$$F_{hkl_0} = |F_{hkl_0}| \exp(i\varphi_{hkl_0})$$

and

$$F_{\bar{h}\bar{k}\bar{l}_0} = |F_{hkl_0}| \exp(-i\varphi_{hkl_0}).$$

Equation (5) leads directly to the coefficient $G_{P\bar{1}}$ defined by

$$\begin{aligned} G_{P\bar{1}} &= \{|F_{hkl}|^2 - 2|F_{hkl_0}|^2\} \\ &= 2|F_{hkl_0}|^2 \cos\left[4\pi(hx_0+ky_0+lz_0)+2\varphi_{hkl_0}\right], \end{aligned} \tag{6}$$

from which the translation vector is determinable as the maximum of the Fourier synthesis

$$\Phi(x,y,z) = \sum_h \sum_k \sum_l G_{P\bar{1}} \cos\left[4\pi(hx+ky+lz)+2\varphi_{hkl_0}\right]. \tag{7}$$

The derivation of (7) from (6) should be seen to be an obvious, foreseeable result for most crystallographers who may choose to reason by the analogy posed by the similarities shared by both $G_{P\bar{1}}$ and its Fourier transform, $\Phi(x,y,z)$, and $F_{hkl}$ and the electron density, $\varrho(x,y,z)$. The algebraic proof that $\Phi(x,y,z)$ possesses a maximum at the value of the true displacement vector, $\mathbf{r}_0$, is given in §3, (15) for the space group $P\bar{1}$.

At this point let us consider the similarities between, say, equation (7) and the corresponding location functions given by Vand & Pepinsky (1956),

$$\begin{aligned} T = \sum_h \sum_k \sum_l &\{|E_{hkl}|^2 - 2|E_{hkl_0}|^2\}|E_{hkl_0}|^2 \\ &\times \cos\left[4\pi(hx+ky+lz)+2\varphi_{hkl_0}\right], \end{aligned} \tag{8}$$

Tollin (1966), for an equal atom structure,

$$\begin{aligned} Q(\mathbf{R}_0) = \sum_h \sum_k \sum_l &\{|E_{hkl}|^2 \\ &-1\}|E_{hkl_0}|^2 \cos\left[4\pi(hx+ky+lz)+2\varphi_{hkl_0}\right], \end{aligned} \tag{9}$$

and Karle (1972),

$$\begin{aligned} D_3(\boldsymbol{\delta}) = \sum_h \sum_k \sum_l &\{(|E_{hkl}|^2-1) \\ &-2(|E_{hkl_0}|^2-1)\}[E^2_{hkl_0}-\bar{C}_3] \\ &\times \exp\left[2\pi i(hx+ky+lz)\right], \end{aligned} \tag{10}$$

which are here expressed with normalized structure amplitudes for uniformity. The relationship given by (8) is found to be encompassed in the $T_1(\mathbf{t})$ function of Crowther & Blow (1967). While various terms are defined for the real component of the location function given in the braces of expressions (8) through (10), each function is modulated by both the phase and amplitude of the transform product $E^2_{hkl_0}$.

The location vector functions proposed in this paper emphasize only the importance of the phase of the molecular transform product in the subsequent Fourier synthesis,

$$\begin{aligned} \Phi(x,y,z) = \sum_h \sum_k \sum_l &\{|E_{hkl}|^2 - 2|E_{hkl_0}|^2\} \\ &\times \cos\left[4\pi(hx+ky+lz)+2\varphi_{hkl_0}\right]. \end{aligned} \tag{11}$$

### 3. Relative intensities of spurious maxima

The following analysis is put forth to demonstrate that spurious maxima are an inescapable manifestation of the translation functions referred to above. It will be shown that the spatial location and intensity of spurious maxima in these translation-vector syntheses are primarily determined by the molecular structure and the particular translation function used.

To consider the phase-modulated synthesis (7), one may substitute

$$\{|E_{\mathbf{h}}|^2 - 2|E_{\mathbf{h}_0}|^2\} = 2|E_{\mathbf{h}_0}|^2 \cos\left[4\pi\mathbf{h}\,.\,\mathbf{r}_0 + 2\varphi_{\mathbf{h}_0}\right] \tag{12}$$

to yield

$$\begin{aligned} \Phi(\mathbf{r}) &= 2\sum_h |E_{\mathbf{h}_0}|^2 \cos\left[4\pi\mathbf{h}\,.\,\mathbf{r}_0+2\varphi_{\mathbf{h}_0}\right] \\ &\quad\times \cos\left[4\pi\mathbf{h}\,.\,\mathbf{r}+2\varphi_{\mathbf{h}_0}\right] \\ &= \sum_h |E_{\mathbf{h}_0}|^2 \{\cos\left[4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})\right] \\ &\quad+ \cos\left[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}\right]\}. \end{aligned} \tag{13}$$

Further analysis will become more manageable if, for the time being, we choose to restrict ourselves to an equal-atom problem in this exercise. It may then be shown that

$$|E_{\mathbf{h}_0}|^2 = 1 + \frac{2}{N} \sum_{j>k}^{N} \cos\left[2\pi\mathbf{h}\,.\,(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})\right]. \tag{14}$$

This expression may be substituted into (13), and by averaging over the number of diffraction terms in the synthesis

$$\begin{aligned} \Phi'(\mathbf{r}) = \Bigg\langle \sum_h &\{\cos\left[4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})\right] \\ &+ \cos\left[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}\right]\} \\ &+ \frac{1}{N} \sum_h \sum_{j>k}^{N} \left\{\cos 4\pi\mathbf{h}\,.\left[\mathbf{r}_0-\mathbf{r} \pm \left(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\right)\right]\right. \\ &\left.+ \cos\left[4\pi\mathbf{h}\,.\left[\mathbf{r}_0+\mathbf{r} \pm \left(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\right)\right] + 4\varphi_{h_0}\right]\right\}\Bigg\rangle_h. \end{aligned} \tag{15}$$

It is clear that the phase-modulated synthesis will have a maximum whose intensity is unity in the case that $\mathbf{r}=\mathbf{r}_0$. It also follows that spurious maxima will occur whenever $\mathbf{r}=\mathbf{r}_0\pm(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})/2$. The intensity of these spurious maxima will be roughly proportional to some integer multiple of $1/N$ which would allow for the degree of overlap of certain of these vectors as may be predicted from the molecular structures. Thus, the true translation vector, in addition to being the largest expected maximum of the phase-modulated synthesis, is also an inversion center about which the lesser, structure-dependent, spurious maxima are distributed.

An important observation to make at this point is that the cosine terms in (15) which contain the argument $4\varphi_{\mathbf{h}_0}$ may in fact be eliminated if one executes the following synthesis

$$\begin{aligned}\Phi''(\mathbf{r})=&\langle 2\sum_h |E_{\mathbf{h}_0}|^2 \sin(4\pi\mathbf{h}\,.\,\mathbf{r}_0\\ &+2\varphi_{\mathbf{h}_0})\sin(4\pi\mathbf{h}\,.\,\mathbf{r}+2\varphi_{\mathbf{h}_0})\rangle_h\\ =&\Big\langle\sum_h\{\cos 4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})\\ &-\cos[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}]\}\\ &+\frac{1}{N}\sum_h\sum_{j>k}^{N}\Big\{\cos\Big[4\pi\mathbf{h}\,.\Big[\mathbf{r}_0-\mathbf{r}\pm\Big(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\Big)\Big]\Big]\\ &-\cos\Big[4\pi\mathbf{h}\,.\Big[\mathbf{r}_0+\mathbf{r}\pm\Big(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\Big)\Big]+4\varphi_{\mathbf{h}_0}\Big]\Big\}\Big\rangle_h,\end{aligned}\tag{16}$$

which upon adding to (15) yields

$$\begin{aligned}\Phi'''(\mathbf{r})=&\Phi''(\mathbf{r})+\Phi'(\mathbf{r})\\ =&\Big\langle\sum_h\Big\{\cos 4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})\\ &+\frac{1}{N}\sum_{j>k}^{N}\cos 4\pi\mathbf{h}\,.\Big[\mathbf{r}_0+\mathbf{r}\pm\Big(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\Big)\Big]\Big\}\Big\rangle_h.\end{aligned}\tag{17}$$

An inconvenience arises in that although one can take measures to insure that $|\cos(4\pi\mathbf{h}\,.\,\mathbf{r}_0+2\varphi_{\mathbf{h}_0})|$ does not exceed a value of one by simply resetting the value of $[|E_\mathbf{h}|^2-2|E_{\mathbf{h}_0}|^2]$ in equation (12), one does not *a priori* know the mathematical sign of $\sin(4\pi\mathbf{h}\,.\,\mathbf{r}_0+2\varphi_{\mathbf{h}_0})$, since it is only known that

$$\begin{aligned}&|\sin(4\pi\mathbf{h}\,.\,\mathbf{r}_0+2\varphi_{\mathbf{h}_0})|\\ &\quad=\{1-[\cos(4\pi\mathbf{h}\,.\,\mathbf{r}_0+2\varphi_{\mathbf{h}_0})]^2\}^{1/2}\,.\end{aligned}\tag{18}$$

One satisfactory way to resolve this dilemma is to select those vectors which correspond to maxima in synthesis (15) and determine the tentative sign of the sine term directly from the trial values of $\mathbf{r}_0$ and the known phase $\varphi_{\mathbf{h}_0}$. Peaks in the vector map which arise from those cosine terms involving $4\varphi_{\mathbf{h}_0}$ should vanish for the correct solution, and are not expected to cancel for the erroneous solutions.

A parallel analysis for the function (8) leads to an expression similar to (13):

$$\begin{aligned}T=\sum_h |E_{\mathbf{h}_0}|^4\{&\cos[4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})]\\ &+\cos[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}]\}\,,\end{aligned}\tag{19}$$

and since

$$\begin{aligned}|E_{\mathbf{h}_0}|^4=&\Big(2-\frac{1}{N}\Big)+\frac{8}{N}\Big(1-\frac{1}{N}\Big)\\ &\times\sum_{j>k}^{N}\cos 2\pi\mathbf{h}\,.\,(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})\\ &+\frac{2}{N^2}\sum_{j>k}^{N}\cos 4\pi\mathbf{h}\,.\,(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})\\ &+\frac{4}{N^2}\sum_{j\neq k}\sum_{j\neq l}\cos 2\pi\mathbf{h}\,.\,(2\mathbf{r}_{j_0}-\mathbf{r}_{k_0}-\mathbf{r}_{l_0})\\ &+\frac{4}{N^2}\sum_{j\neq k\neq l\neq m}\cos 2\pi\mathbf{h}\,.\,(\mathbf{r}_{j_0}-\mathbf{r}_{k_0}+\mathbf{r}_{l_0}-\mathbf{r}_{m_0})\end{aligned}\tag{20}$$

it follows that

$$\begin{aligned}T'=&\Big\langle\Big(2-\frac{1}{N}\Big)\sum_h\cos 4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r})\\ &+\cos[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}]\\ &+\frac{4}{N}\Big(1-\frac{1}{N}\Big)\sum_h\sum_{j>k}\cos 4\pi\mathbf{h}\,.\Big[\mathbf{r}_0-\mathbf{r}\pm\Big(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\Big)\Big]\\ &+\cos\Big[4\pi\mathbf{h}\,.\Big(\mathbf{r}_0+\mathbf{r}\pm\Big(\frac{\mathbf{r}_{j_0}-\mathbf{r}_{k_0}}{2}\Big)\Big)+4\varphi_{\mathbf{h}_0}\Big]\\ &+\frac{1}{N^2}\sum_h\sum_{j>k}\cos 4\pi\mathbf{h}\,.\,[\mathbf{r}_0-\mathbf{r}\pm(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})]\\ &+\cos\{4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r}\pm(\mathbf{r}_{j_0}-\mathbf{r}_{k_0})]+4\varphi_{\mathbf{h}_0}]\\ &+\frac{2}{N^2}\sum_h\sum_{j\neq k}\sum_{j\neq l}\cos 4\pi\mathbf{h}\,.\Big[\mathbf{r}_0-\mathbf{r}\pm\Big(\mathbf{r}_{j_0}-\Big(\frac{\mathbf{r}_{k0}+\mathbf{r}_{l0}}{2}\Big)\Big)\Big]\\ &+\cos\Big[4\pi\mathbf{h}\,.\Big[\mathbf{r}_0+\mathbf{r}\pm\Big(\mathbf{r}_{j_0}-\Big(\frac{\mathbf{r}_{k0}+\mathbf{r}_{l0}}{2}\Big)\Big)\Big]+4\varphi_{\mathbf{h}_0}\Big]\\ &+\frac{2}{N^2}\sum_h\sum_{j\neq k\neq l\neq m}\cos 4\pi\mathbf{h}\,.\Big[\mathbf{r}_0-\mathbf{r}\\ &\pm\Big(\frac{\mathbf{r}_{j0}-\mathbf{r}_{k0}+\mathbf{r}_{l0}-\mathbf{r}_{m0}}{2}\Big)\Big]\\ &+\cos\Big[4\pi\mathbf{h}\,.\Big(\mathbf{r}_0+\mathbf{r}\pm\Big)\frac{\mathbf{r}_{j0}-\mathbf{r}_{k0}+\mathbf{r}_{l0}-\mathbf{r}_{m0}}{2}\Big)\Big)+4\varphi_{\mathbf{h}_0}\Big]\Big\rangle_h.\end{aligned}\tag{21}$$

A comparison of those terms common to both (15) and (21) clearly indicates that the latter produces spurious maxima which are twice as large as those anticipated in the phase-modulated synthesis (15). It should also be noted that the remaining three summations in (21) contain approximately $N^2/2$, $N^3/2$, and $N^4/4$ contributors respectively. It will be recognized that only a small percentage of these contributors in the latter summations need be vectors that overlap in order to bring about a situation where a spurious maximum may exceed the intensity of the true location-vector maximum. Similar analyses of the $Q(\mathbf{R}_0)$ and $D_3(\boldsymbol{\delta})$ synthesis are seen to involve $|E_{\mathbf{h}_0}|^4$ and have spurious background characteristics similar to (21). One of the referees examining this manuscript stressed that much of the spurious background in the $Q(\mathbf{R}_0)$ function is attributable to the overriding role played by a few unfortunate, large terms which appear in the synthesis. The referee offered that one physical interpretation of the reduction in the background noise of the phase-modulated synthesis might be that many more terms provide a significant contribution to the function and minimize the effect of series termination in the subsequent synthesis.

It should be added that whereas insufficient time has elapsed for the Karle translation function to gain as general an acceptance as the $Q$ function, a recent application (Delbaere & James, 1973) leads one to conclude that although the noise level of spurious maxima by either method may be considerable, it is often possible to eliminate intelligently certain of these spurious maxima as probable solutions.

The spurious background can be eliminated by reformulating (11)

$$\begin{aligned}\Phi''(\mathbf{r}) &= \left\langle \sum_h \left\{ \frac{|E_\mathbf{h}|^2-2|E_{\mathbf{h}_0}|^2}{|E_{\mathbf{h}_0}|^2} \cos\,[4\pi\mathbf{h}\,.\,\mathbf{r}+2\varphi_{\mathbf{h}_0}] \right\rangle_h \\ &= \langle \sum_h \{\cos 4\pi\mathbf{h}\,.\,(\mathbf{r}_0-\mathbf{r}) \\ &\quad + \cos\,[4\pi\mathbf{h}\,.\,(\mathbf{r}_0+\mathbf{r})+4\varphi_{\mathbf{h}_0}]\}\rangle_h\,. \end{aligned} \tag{22}$$

The theoretical advantages of this form are compromised by certain computational problems. In practice one will encounter difficulties with space groups which incorporate more than one symmetry operation, and in applications to partial structures. For a partial structure one must take precautions to insure that a quotient of the form $\{|E_\mathbf{h}|^2-2|E_{\mathbf{h}_0}|^2/2|E_{\mathbf{h}_0}|^2\}$ is well behaved in that the cosine term in (12) assumes reasonable values. This might suggest the removal of those terms from (22) for which $|E_{\mathbf{h}_0}|^2$ is smaller than some fixed threshold value. Trial calculations have shown that the evaluation of (22) over those terms for which $|E_{\mathbf{h}_0}|^2$ is large may greatly reduce the size of the Fourier synthesis but the end result is usually no better than the normal phase-modulated synthesis (11).

Similar difficulties are encountered in higher space-group symmetry for complete structures. The presence of four distinct cosine terms given in equation (30) in the Appendix implies that $G_{P2_1/c}$ is at best a probabilistic estimate of any one cosine term when $h,l\neq 0, k\neq 0$. However, the restriction of (30) to either $h0l$ or $0k0$ data makes it possible to express $G_{P2_1/c}$ as a function of a single cosine term which may be precisely estimated.

## 4. Analysis of partial structures

It may be worth while to note that one need not be restricted to crystallographic problems which contain one molecule in the asymmetric unit, nor do these molecules need to be completely described. In general, the structure-factor equations may be formulated in $P1$ as

$$F_\mathbf{h} = \sum_{j=1}^{N} F_{\mathbf{h}_{0_j}} \exp\,(2\pi i\mathbf{h}\,.\,\mathbf{r}_j) + \sum_{j=1}^{M} f_j \exp\,(2\pi i\mathbf{h}\,.\,\mathbf{r}_j) \tag{23}$$

where the $F_{\mathbf{h}_{0_j}}$ are transforms of correctly oriented molecules or fragments thereof and the $f_j$ are the atomic scattering factors of any residual atoms in the unit cell. All displacements are relative to the origin defining any chosen transform designated as $F_{\mathbf{h}_{0_j}}$. It follows that

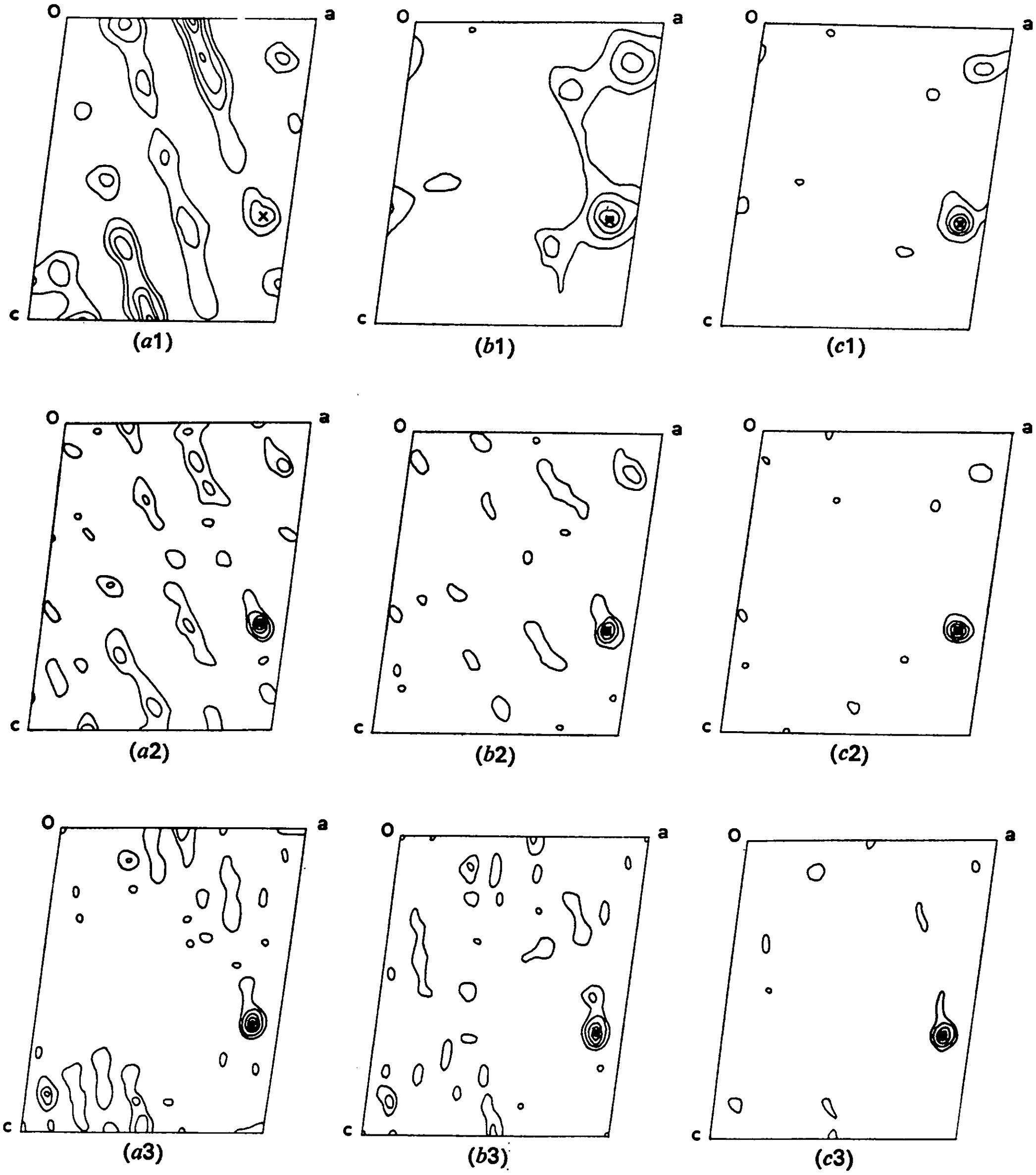
$$\begin{aligned} G_\mathbf{h} &= \{|F_\mathbf{h}|^2 - \sum_{j=1}^{N} |F_{\mathbf{h}_{0_j}}|^2 - \sum_{j=1}^{M} f_j^2\} \\ &= 2 \sum_{i\neq}^{N} \sum_{j}^{N} |F_{\mathbf{h}_{0i}} F_{\mathbf{h}_{0j}}| \cos\,[2\pi\mathbf{h}\,.\,(\mathbf{r}_i-\mathbf{r}_j)+\varphi_{\mathbf{h}_{0i}}-\varphi_{\mathbf{h}_{0j}}] \\ &\quad + 2 \sum_{i}^{N} \sum_{j}^{M} |F_{\mathbf{h}_{0i}} f_j| \cos\,[2\pi\mathbf{h}\,.\,(\mathbf{r}_i-\mathbf{r}_j)+\varphi_{\mathbf{h}_{0i}}] \\ &\quad + 2 \sum_{i\neq}^{M} \sum_{j}^{M} f_i f_j \cos\,[2\pi\mathbf{h}\,.\,(\mathbf{r}_i-\mathbf{r}_j)]\,. \end{aligned} \tag{24}$$

If we consider the three distinct parts of equation (24) it will be recognized that the Fourier transform of $G_\mathbf{h}$ *via* the trigonometric form $\cos\,[2\pi\mathbf{h}\,.\,\mathbf{r}+\varphi_{\mathbf{h}_{0i}}-\varphi_{\mathbf{h}_{0j}}]$ yields the Patterson function of the $i$th and $j$th molecular fragments. A similar exercise regarding the second term prescribes the synthesis of $G_\mathbf{h} \cos\,[2\pi\mathbf{h}\,.\,\mathbf{r}+\varphi_{\mathbf{h}_{0i}}]$ which produces the crystal structure of the residual atoms relative to the origin of the $i$th molecular transform. The last term of equation (24) indicates that the Patterson function based on $G_\mathbf{h} \cos\,(2\pi\mathbf{h}\,.\,\mathbf{r})$ will simply be the Patterson function of the residual atoms. Given that the cross vectors between molecular fragments may be anticipated to produce the largest vector maxima, since they are proportional to $\langle|F_{\mathbf{h}_{0i}} F_{\mathbf{h}_{0j}}|\rangle_h$, it follows that structure elucidation should begin with Fourier syntheses to determine these vectors whenever it is possible.

## 5. Test calculations

Data were obtained for the crystal structure of *p*-methoxybenzenesulfon-*p*-anisidide, $C_{14}H_{15}NO_4S$ (Pokrywiecki, Weeks & Duax, 1973) which crystallizes in the space group $P2_1/c$ with four molecules in the unit cell. A number of preliminary test calculations were performed in order to compare the results given by equation (32) given in the Appendix with those obtained by the appropriate $Q(\mathbf{R}_0)$ and $T_1(\mathbf{t})$ functions referred to above. Random, normalized Gaussian displacements were imparted to the refined atomic coordinates of the non-hydrogen atoms of the structure so that the root-mean-square error in the relative atomic positions of the molecule was 0·11 Å. This was done to simulate the magnitude of error one might obtain in a trial model constructed from Patterson vectors. A provisional structure for the location vector synthesis was obtained by shifting this molecule to an arbitrary position within the unit cell. The scale factor used to place the structure amplitudes on an absolute scale was that obtained from the refinement; an overall isothermal temperature factor of 3·5 $Å^{-2}$ was used in the computations. Only data for which $(\sin\theta)/\lambda \lesssim$ 0·30 $Å^{-1}$ were included in the syntheses. Care was taken to remove the $0k0$ reflections which only affect the background of the maps and make direct comparisons of questionable value. The $h0l$ data were also subsequently ignored in an affort to minimize the satellite peaks they produce as a consequence of the glide component. The calculations were repeated for normalized structure amplitude data and the corresponding sharpened location-vector maps were obtained.

A comparison of the location-vector maps, apart from indicating the correct displacement vector by their largest maxima, readily revealed similarities in



(*a*1) (*b*1) (*c*1)
(*a*2) (*b*2) (*c*2)
(*a*3) (*b*3) (*c*3)

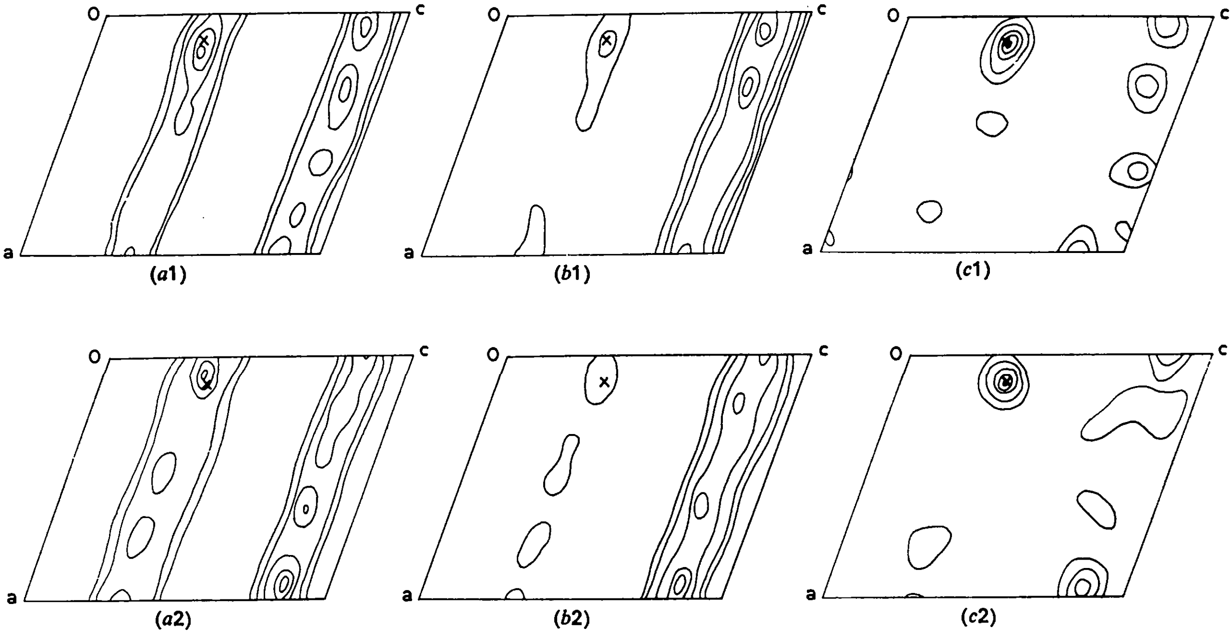
Fig. 1. Location vector syntheses for *p*-methoxybenzenesulfon-*p*-anisidide: (*a*) $Q(\mathbf{R}_0)$, (*b*) $T_1(\mathbf{t})$, Section $y=\frac{1}{2}$, and (*c*) $\Phi(x,z)$. Postscripts denote (1) $F$ synthesis, (2) $E$ synthesis, and (3) $E$ synthesis with $h0l$ data omitted. Contour levels are plotted for relative densities of 25, 50, 75, 90 and 100; the position of the anticipated location vector is marked with ×.

the positions of the larger spurious maxima. Whereas the relative number and the magnitude of spurious maxima were comparable for the $Q(\mathbf{R}_0)$ and $T_1(\mathbf{t})$ syntheses, the phase-modulated synthesis, $\Phi(x,z)$, showed a significant reduction in the magnitude of this background as is illustrated in Fig. 1.

Further trial calculations were performed in order to determine whether this degree of improvement could be anticipated for the less ideal situations in which only partial structural information is available. Data obtained from the crystal structure of estriol (Cooper, Norton & Hauptman, 1969) were used to this end. Estriol crystallizes in the space group $P2_1$ with two molecules in the asymmetric unit; the ($h0l$) projection of the structure is severely overlapped. Relative atomic coordinates were generated for each of the two molecules in the manner previously described. Data were extended to include 828 terms for which $(\sin\theta)/\lambda \lesssim 0{\cdot}40$ Å$^{-1}$, the $0k0$ reflections were ignored as before. Sharpened $Q(\mathbf{R}_0)$, $T_1(\mathbf{t})$ and $\Phi(x,z)$ syntheses were executed on the basis of the partial structural information provided by either of the two molecules with the exclusion of the other. The appropriate $\Phi(x,z)$ synthesis for the space group $P2_1$ is defined by equation (29) in the Appendix. The real components of the molecular deconvolution corresponding to the quantities in the braces of equations (28) were redefined for a partial structure as is indicated by the left-hand side of equation (24). The resultant location-vector maps are presented in Fig. 2. The $Q(\mathbf{R}_0)$ and $T_1(\mathbf{t})$ functions are seen to produce 'sausage-shaped' maxima which are indicative of the limited ability of these functions to locate the correct molecular image in the crystal structure provided by the example. This discrimination appears to be significantly better for the $\Phi(x,z)$ function in that the background noise of the map is lower and the distortion of the location-vector maxima is less pronounced.

Whereas the initial molecular coordinates were obtained by adding the same displacement vector to the refined coordinates of both contiguous molecules, in a practical application these vectors would seldom be similar. In the particular problem one is free to define the position of the first molecule with respect to any of the four equivalent screw axes, but then a fourfold ambiguity must be resolved with respect to the correct placement of the second molecule. The map contours in Fig. 2 are decidedly dimensionless for the $E$ modulus syntheses but would have units of electrons-squared if $F$ values were used.

Estriol molecule (II) was used as a partial structure from which the crystal structure of the remaining atoms was computed by (24). It should be noted that the appropriate syntheses are similar to unweighted difference electron-density syntheses except that the latter use $|E_{\mathbf{h}}| - |E_{\mathbf{h}_0}|$ in place of the corresponding value of $G_{\mathbf{h}}$ given by the left-hand side of equation (24). In trial (II), these calculations were repeated for a smaller eleven-atom fragment of the estriol molecule which represented less than 25% of the electron density within the unit cell. This eleven-atom fragment included atoms O(3) and C(1) through C(10) which define the $A$ and $B$ rings of the steroid molecule. The root-mean-square displacement of these atoms from



(a1) (b1) (c1)

(a2) (b2) (c2)

Fig. 2. Sharpened location vector syntheses for estriol: ($a$) $Q(\mathbf{R}_0)$, ($b$) $T_1(\mathbf{t})$, Section $y=\frac{1}{2}$, and ($c$) $\Phi(x,z)$. Postscripts (1) and (2) refer to analyses based on the molecular transform provided by either estriol molecule (I) or (II). Contour levels are plotted for relative densities of 25, 50, 75, 90 and 100; the position of the anticipated location vector is marked with ×.

their refined positions was 0·10 Å as previously noted. Lastly, in trial (III), this calculation was executed with the refined coordinates of this eleven-atom fragment. In all instances the corresponding unweighted difference electron-density syntheses were assembled for purposes of comparison.

Each of the Fourier syntheses was examined with the object of recognizing the structure of estriol molecule (I). Molecular packing considerations restrict the search to approximately half the volume of the asymmetric unit. The observations summarized in Table 1 suggest that the synthesis of the residual structure given by (24) is probably better than the corresponding difference electron-density synthesis with regard to both the frequency and intensity of spurious maxima. The accuracy of the atomic coordinates of the residual molecule appears to be comparable. The similarities of trials (II) and (III) tend to indicate that these syntheses are not as sensitive to errors in the description of the partial structure as one may have anticipated. The results alluded to in Table 1 may be misleading, however, in that the correct maxima from trials (II) and (III) define a molecule whose bond lengths vary from 1·0 to 1·9 Å in either extreme. Under such circumstances there is a great temptation to utilize the spurious maxima in constructing molecules which are chemically more acceptable.

Table 1. *Summary of the observations*

| Trial no. from text | Rank no. spurious maxima among 21 largest map peaks | | R.m.s. displacement of correct maxima from refined positions | |
|---|---|---|---|---|
| | (*a*) | (*b*) | (*a*) | (*b*) |
| (I) | 21 | 20, 21 | 0·24 Å | 0·24 Å |
| (II) | 20 | 11, 19 | 0·31 | 0·31 |
| (III) | 20, 21 | 7, 10, 16 | 0·31 | 0·29 |

(*a*) Calculation using equation (24), (*b*) corresponding unweighted difference electron-density synthesis.

## 6. Note on the $D_3(\boldsymbol{\delta})$ function

A note is added in explanation for the omission of the $D_3(\boldsymbol{\delta})$ function in the comparison of translation functions cited at the beginning of the last section. The author submits that the parameter $\bar{C}_3$ given in (10) can serve the purpose of eliminating a cumulative origin peak in the $D_3(\boldsymbol{\delta})$ function; however, no evidence exists to suggest that the $D_3(\boldsymbol{\delta})$ function contains such an origin peak. This may be demonstrated by averaging (10) over the Friedel pairs in the diffraction data to arrive at

$$\begin{aligned} D_3'(\boldsymbol{\delta}) = \sum_h \sum_k \sum_l \, & [(|E_{hkl}|^2-1)-2(|E_{hkl_0}|^2-1)]|E_{hkl_0}|^2 \\ & \times \cos\,[2\pi(hx+ky+lz)+2\varphi_{hkl_0}] \\ & -\sum_h \sum_k \sum_l \, [(|E_{hkl}|^2-1)-2(|E_{hkl_0}|^2-1)]\bar{C}_3 \\ & \times \cos\,[2\pi(hx+ky+lz)]\,. \end{aligned} \tag{25}$$

After further factoring the first term in (25) and borrowing from equations (12) and (8) we arrive at

$$\begin{aligned} D_3'(\boldsymbol{\delta}) = T(2\boldsymbol{\delta}) + \tfrac{1}{2} \sum_h \sum_k \sum_l \, & [|E_{hkl}|^2 - 2|E_{hkl_0}|^2] \\ & -\sum_h \sum_k \sum_l \, [(|E_{hkl}|^2-1)-2(|E_{hkl_0}|^2-1)]\bar{C}_3 \\ & \times \cos\,[2\pi(hx+ky+lz)]\,. \end{aligned} \tag{26}$$

A complete lack of terms of the sort $\cos\,(4\pi\mathbf{h}\,.\,\mathbf{r})$ in (21) insures that $T(2\boldsymbol{\delta})$ possesses no predictable origin peak. The second term of (26) uniformly adds a constant value to all densities in the $D_3(\boldsymbol{\delta})$ function produced by the $T(2\boldsymbol{\delta})$ component. While the last term of (26) may be adjusted to remove any cumulative density at the origin, this activity is clearly unrelated to any improvement of the $D_3(\boldsymbol{\delta})$ function to yield information beyond that which is supplied by the $T(2\boldsymbol{\delta})$ component of the function. In fact the value of $\bar{C}_3$ is immaterial if one performs Karle's multiple-shift procedure to the differential limit. The integral average of all differentially shifted maps summed back onto a common reference map is the single map simply produced by neglecting the last term in (26).

## 7. Summary

The phase-modulated translation function has been tentatively shown to be superior to similar functions which are modulated by both the phase and amplitude of molecular transform products. Its adoption is urged as an improved alternative to those translation functions which are in popular usage.

The translation function of a partial structure has been analyzed as a difference density function which is comparable to the difference electron-density synthesis. The limited scope of the trial examples regarding the comparison of the difference electron densities produced by these two functions should strongly persuade one to exercise caution in drawing too many conclusions at this time.

I wish to thank Miss M. Tugac and Mr. T. Crump for their assistance in preparing the figure illustrations. Suggestions by Drs H. Hauptman and E. Green to improve the readability of this manuscript are gratefully acknowledged. This work was supported in part by U.S.P.H. Grant No. RR 05716 and NSF Grant No. MPS73-04992.

## APPENDIX

In the space group $P2_1$ one may readily show that

$$\begin{aligned} F_{hkl} = F_{hkl_0} & \exp\,[2\pi i(hx_0+lz_0)] \\ & +(-1)^k F_{\bar{h}k\bar{l}_0} \exp\,[-2\pi i(hx_0+lz_0)]\,, \end{aligned} \tag{27}$$

where it should be noted that $F_{hkl_0}$ and $F_{\bar{h}k\bar{l}_0}$ are not symmetry related. This may be simply done by associating $F_{hkl_0}$ and its expressed displacement factor with the initial equipoint $(x,y,z)$ and stepwise performing all of the necessary rotational, inversion, mirror and crystallographic translation operations on the molecular transform triple and its displacement factor to generate the remaining equivalent positions of the space group.

Next multiply $F_{hkl}$ by its complex conjugate and transpose the separable squared moduli to the left to get

$$\begin{aligned} G_{P2_1} &= (-1)^k\{|F_{hkl}|^2 - |F_{hkl_0}|^2 - |F_{\bar{h}\bar{k}\bar{l}_0}|^2\} \\ &= 2|F_{hkl_0}F_{\bar{h}\bar{k}\bar{l}_0}| \cos\,[4\pi(hx_0 + lz_0) \\ &\quad + (\varphi_{hkl_0} + \varphi_{h\bar{k}l_0})]\,. \end{aligned} \tag{28}$$

The Fourier transform of $G_{P2_1}$ may be cleanly expressed as

$$\begin{aligned} \Phi(x,z) &= \sum_h \sum_k \sum_l G_{P2_1} \\ &\quad \times \cos\,[4\pi(hx + lz) + (\varphi_{hkl_0} + \varphi_{h\bar{k}l_0})]\,. \end{aligned} \tag{29}$$

A similar analysis in the space group $P2_1/c$ will help illustrate the increasing number of permissible syntheses for space groups containing additional symmetry. One may readily show that

$$\begin{aligned} G_{P2_1/c} &= \left\{\frac{|F_{hkl}|^2}{2} - |F_{hkl_0}|^2 - |F_{h\bar{k}l_0}|^2\right\} \\ &= |F_{hkl_0}|^2 \cos\,[4\pi(hx_0 + ky_0 + lz_0) + 2\varphi_{hkl_0}] \\ &\quad + |F_{hkl_0}|^2 \cos\,[4\pi(hx_0 - ky_0 + lz_0) + 2\varphi_{h\bar{k}l_0}] \\ &\quad + (-1)^{k+l}|F_{hkl_0}F_{h\bar{k}l_0}| \\ &\quad \times \{\cos\,[4\pi(hx_0 + lz_0) + \varphi_{hkl_0} + \varphi_{h\bar{k}l_0}] \\ &\quad + \cos\,[4\pi ky_0 + \varphi_{hkl_0} - \varphi_{h\bar{k}l_0}]\} \end{aligned} \tag{30}$$

which suggests that the following three syntheses

$$\begin{aligned} \Phi(x,y,z) &= \sum_h \sum_k \sum_l G_{P2_1/c} \\ &\quad \times \{\cos\,[4\pi(hx + ky + lz) + 2\varphi_{hkl_0}] \\ &\quad + \cos\,[4\pi(hx - ky + lz) + 2\varphi_{h\bar{k}l_0}]\}\,, \end{aligned} \tag{31}$$

$$\begin{aligned} \Phi(x,z) &= \sum_h \sum_k \sum_l (-1)^{k+l} G_{P2_1/c} \\ &\quad \times \cos\,[4\pi(hx + lz) + (\varphi_{hkl} + \varphi_{h\bar{k}l_0})]\,, \end{aligned} \tag{32}$$

and

$$\begin{aligned} \Phi(y) &= \sum_h \sum_k \sum_l (-1)^{k+l} G_{P2_1/c} \\ &\quad \times \cos\,[4\pi ky + \varphi_{hkl_0} - \varphi_{h\bar{k}l_0}] \end{aligned} \tag{33}$$

may prove useful for determining the molecular displacement vector relative to the inversion center, screw axis, and glide plane respectively.

## References

BOOTH, A. D. (1948). *Fourier Technique in X-ray Organic Structure Analysis.* Cambridge Univ. Press.

BRAUN, P. B., HORNSTRA, J. & LEENHOUTS, J. I. (1969). *Philips Res. Rep.* **24**, 85–118.

COOPER, A., NORTON, D. A. & HAUPTMAN, H. (1969). *Acta Cryst.* B**25**, 814–828.

CROWTHER, R. A. & BLOW, D. M. (1967). *Acta Cryst.* **23**, 544–548.

DELBAERE, L. T. J. & JAMES, M. N. G. (1973). *Acta Cryst.* B**29**, 404–407.

EWALD, P. P. (1935). *Z. Kristallogr.* **90**, 493–494.

HAUPTMAN, H. (1964). *Acta Cryst.* **17**, 1421–1433.

HOPPE, W. (1957). *Z. Elektrochem.* **61**, 1076–1083.

HUBER, R. (1965). *Acta Cryst.* **19**, 353–356.

KARLE, J. (1972). *Acta Cryst.* B**28**, 820–824.

KROON, J. & KRABBENDAM, H. (1970). *Acta Cryst.* B**26**, 312–314.

NORDMAN, C. E. & NAKATSU, K. (1963). *J. Amer. Chem. Soc.* **85**, 353–354.

POKRYWIECKI, S., WEEKS, C. M. & DUAX, W. L. (1973). *Cryst. Struct. Commun.* **2**, 63–66.

TAYLOR, C. A. & MORLEY, K. A. (1959). *Acta Cryst.* **12**, 101–105.

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

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

VAND, V. & PEPINSKY, R. (1956). *Z. Kristallogr.* **108**, 1–14.