SHORT COMMUNICATIONS

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 *words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible.*

The use of correctly oriented molecular fragments for a modification of E values. By ALBRECHT MESSERSCHMIDT, GÜNTER RECK* and LEO KUTSCHABSKY, *Central Institute of Molecular Biology of the Academy of Sciences of the GDR, Robert-Rössle-Strasse* 10, *DDR*-1115 *Berlin-Buch, German Democratic Republic*

Abstract

If the atoms in a unit cell are not almost randomly distributed difficulties may arise in the normal application of direct methods. Structural information can be taken into account mainly by modification of the probablity distribution of the E values or by modification of the E values themselves. A straighforward method involving the second possibility, applied successfully to the solution of two noncentro-symmetric steroid structures, is described in this paper. The modified E values are calculated by the formula $|E_m(hkl)| = [|E_{obs}(hkl)|^2 - c|E_{frag}(hkl)|^2]^{1/2}$, with $E_m(hkl)$ = modified E value, $E_{obs}(hkl)$ = observed E value, $E_{frag}(hkl)$ = theoretical E value calculated with coordinates of a correctly oriented molecular fragment.

In the case of approximately planar molecules, normal runs of direct-methods programs such as *MULTAN* 78 (Main, Hull, Lessinger, Germain, Declercq & Woolfson, 1978) or *SHELX* 76 (Sheldrick, 1976) without the use of structural information often fail or reveal correctly oriented but incorrectly positioned molecular fragments. Usually the Patterson map shows in such cases a gridlike distribution of peaks. Main (1975) pointed out how to use structural information to improve direct methods. In this paper four different cases are distinguished: 1. randomly positioned atoms; 2. groups of atoms with random position and random orientation; 3. correctly oriented but randomly positioned atomic groups; 4. correctly positioned atoms. An especially important improvement is obtained in cases 3 and 4 because, besides a new normalization of E values, the phase and the variance of the Σ_2 relationships are estimated using the structural information. Information in cases 3 and 4 can be obtained from E maps without the use of structural information or from Patterson vector search methods (*e.g.* Hoppe, 1957*a*, *b*; Huber, 1965; Nordman, 1966). Case 4 is also successfully treated by the application of direct methods for phase extension and refinement of difference structure factors (*DIRDIF*, Beurskens, Bosman, Doesburg, Gould, Van den Hark & Prick, 1980) while case 3 needs the combination of *DIRDIF* with a translation function (*TRADIR*, Beurskens, Doesburg & Beurskens, 1980) or the pure application of a translation function (*e.g.* Karle, 1975).

* To whom all correspondence should be addressed.

In this paper, a new possibility of using structural information for the improvement of direct methods is reported. The improvement is done by a modification of E values. The modification is in dealing with case 3 or 4 and the information of the correctly oriented molecular fragment has to be taken from a normal E map or from Patterson vector rotation search as mentioned above. New E values are calculated according to (1):

$$|E_m(hkl)| = [|E_{obs}(hkl)|^2 - c|E_{frag}(hkl)|^2]^{1/2}, \quad (1)$$

with $E_m(hkl)$ = modified E value, $E_{obs}(hkl)$ = observed E value, $E_{frag}(hkl)$ = E values calculated with coordinates of a correctly oriented fragment, c = subtracting constant.

The $|E_m(hkl)|^2$ values represent the Fourier transform of a Patterson map derived from a Patterson map calculated with the $|E_{obs}(hkl)|^2$ and modified by subtracting a theoretical Patterson function containing the vector set of a correctly oriented fragment of the molecule. It may be expected that in the modified Patterson map $[|E_m(hkl)|^2]$ gridlike distributions of peaks are less prominent than in the original Patterson function $[|E_{obs}(hkl)|^2]$. Therefore, it may be assumed that the probability distribution of the $E_m(hkl)$ deviates less from the theoretical curve for randomly distributed atoms than the distribution of the original E values. The constant c ($c < 1.0$) reduces the effect of the subtraction and a value of 0.2 has proved to be favourable. The modified E values are rescaled to $\langle|E_m(hkl)|^2\rangle = 1$ and entered into the *MULTAN* part of the *MULTAN* 78 system.

The modification was successfully applied to the solution of 3-methoxy-14α,15α-oxido-estra-1,3,5(10)-triene-17α-ol (Messerschmidt & Schubert, 1982), space group $P2_12_12_1$, 22 non-hydrogen atoms per asymmetric unit, and to a modified cardenolide (Messerschmidt & Megges, 1982), space group $P2_12_12_1$, 35 non-hydrogen atoms per asymmetric unit. In both cases, the normal *MULTAN* 78 and *SHELX* runs revealed E maps consisting of fragments containing six-membered rings. The peaks of different heights form grids. Two transfused six-membered rings formed by the highest peaks served as input parameters for the modification. For the calculation of the $E_{frag}(hkl)$ values the scattering curve of carbon was used. A subtraction constant c equal to 0.2 was applied and proved satisfactory. For both structures, only 64 phase sets were generated including 1500 Σ_2 relationships. The first structure (Fig. 1) was found from an E map calculated for the phase set with third-highest combined

figure of merit. This E map contained all 22 non-hydrogen atoms to be found. For the modified cardenolide structure, the E map calculated for the phase set with the second-highest combined figure of merit showed 31 atoms of the total 35 non-hydrogen atoms of the molecule. It is important to remark that the translations of the correctly positioned molecules did not coincide with the translations of the structural fragments selected from the normal E map. The modification of E values is easily incorporated in existing
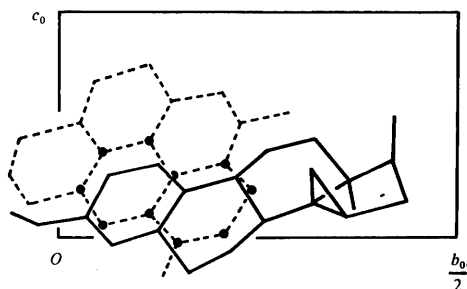
direct-methods programs and is an interesting alternative to the above-mentioned methods.

Fig. 1. Comparison of E maps of 3-methoxy-14α,15α-oxido-estra-1,3,5(10)-triene-17α-ol. The peaks of the best E map calculated with the $E_{obs}(hkl)$ are connected by dotted lines representing a grid. The solid black circles mark the peaks which were used for the calculation of the $E_{frag}(hkl)$. The solid lines connect the 22 highest peaks in the E map with the third-highest combined figure of merit calculated with the $E_m(hkl)$. All these peaks correspond to atoms.

References

BEURSKENS, P. T., BOESBURG, H. M. & BEURSKENS, G. (1980). *TRADIR*. Strengthened translation functions. Univ. of Nijmegen, the Netherlands.

BEURSKENS, P. T., BOSMAN, W. P., DOESBURG, H. M., GOULD, R. O., VAN DEN HARK, TH. E. M. & PRICK, P. A. J. (1980). *DIRDIF*. Direct methods for difference structures. Univ. of Nijmegen, the Netherlands.

HOPPE, W. (1957a). *Acta Cryst.* 10, 750–754.

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

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

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

MAIN, P. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 97–105. Copenhagen: Munksgaard.

MAIN, P., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1978). *MULTAN 78. A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univ. of York, England.

MESSERSCHMIDT, A. & MEGGES, R. (1982). *Acta Cryst.* Submitted.

MESSERSCHMIDT, A. & SCHUBERT, S. (1982). *Acta Cryst.* Submitted.

NORDMAN, C. E. (1966). *Trans. Am. Crystallogr. Assoc.* 2, 65–74.

SHELDRICK, G. M. (1976). Program for crystal structure determination. Univ. of Cambridge, England.

On the estimated standard deviation of the atom-to-plane distance.

By TETSUZO ITO, *The Institute of Physical and Chemical Research, Wako-shi, Saitama 351, Japan*

Abstract

An expression is given for the estimated standard deviation of the atom-to-plane distance of an atom defining the least-squares plane.

Two least-squares methods have been proposed to determine the best-plane parameters and their error matrix (variance–covariance matrix): one is a method of undetermined multiplier (Ito, 1981a, hereinafter called paper I), and the other is an elimination method (Ito, 1981b, paper II). Although the two methods give identical results in a general case, the elimination method has an advantage that it can deal with a plane defined with only three atoms.

In paper I, an expression is given for the estimated standard deviation of the atom-to-plane distance of the *i*th atom:*

* As will be explained in the following, σ_1 of (11) and σ_2 of (1) are the estimated standard deviations of the atom-to-plane distance of an atom defining and not defining the plane, respectively.

$$\sigma_2^2(d_i) = (^d\mathbf{M}_0)_{ii} + \tilde{\mathbf{y}}_i\,{}^n\mathbf{M}\mathbf{y}_i, \tag{1}$$

where $^d\mathbf{M}_0$ is the error matrix for the atom-to-plane distances originating from the atomic positional errors, $\tilde{\mathbf{y}}_i$ (transpose of $\mathbf{y}_i$) is an atomic coordinate matrix:

$$\tilde{\mathbf{y}}_i \equiv (x_i\,y_i\,z_i - 1), \tag{2}$$

and $^n\mathbf{M}$ is the error matrix for the four plane parameters,

$$\tilde{\mathbf{n}} \equiv (m_1\,m_2\,m_3\,d_0), \tag{3}$$

which is obtained through the best-plane calculations. Expression (1) has been derived under the assumption that the atomic coordinates and the plane parameters are not correlated. Therefore, it is applicable only to atoms not defining the plane. The assumption is not valid for the plane-defining atoms because their coordinates are used to determine the plane parameters. Particularly in the case of a three-atom plane, the correlation is perfect; since the three atoms necessarily lie on the plane, their atom-to-plane distances of zero should have zero standard deviations, which is evidently in contradiction with (1).