Acta Cryst. (1975). A31, 818

Consistency Test for the Determination of a Correct Phase Set of the Structure Factors

BY J. HAŠEK

Institute of Macromolecular Chemistry, Czechoslovak Academy of Sciences, 16206 Prague 6, Czechoslovakia

(*Received* 17 *February* 1975; *accepted* 28 *May* 1975)

When solving the phase problem by direct methods one usually has to decide which of the many suggested solutions is the correct one. The coefficients K and N defined in this paper can be calculated for each possible solution. A correct solution is expected to be that which gives the lowest values of the above coefficients.

Introduction

In recent years, considerable attention has been devoted to the problem of finding a correct phase set of the structure factors. Schenk (1972) criticized the shortcomings of the existing methods reflected mainly in the investigation of crystal structures in symmorphous space groups. Later, several procedures were published demonstrating how to avoid the above difficulties by using relationships making possible an indication of the negative sign of the structure invariant.

Hašek (1974) showed that these difficulties can also be removed in symmorphous space groups without applying the above special relationships. For the case of centrosymmetric structures he used a procedure in which the calculation of the coefficient K for the tested phase set of the structure factors is virtually reduced to the summation of triplets with the positive and with the negative sign respectively in the invidual groups according to their weights.

In this paper the procedure has been extended to include noncentrosymmetric space groups.

Theoretical relationships

In recent years expressions have been derived for the mean values of cosine of the invariant $\Phi_{111}=\varphi_H+\varphi_K+\varphi_{-H-K}$ (Karle & Gilardi, 1973) and $\Phi_{211}=2\varphi_H+\varphi_{K-H}+\varphi_{-H-K}$ (Giacovazzo, 1974). These expressions can be written as

$$\langle\cos\Phi\rangle_{\text{theor}}=\frac{I_1(w)}{I_0(w)}\simeq[(0{\cdot}0106w-0{\cdot}1304)w+0{\cdot}5658]w. \quad (1)$$

The above approximation holds for the range $0{\cdot}0\leq w\leq 6{\cdot}0$. For $w\geq 6{\cdot}0$, $I_1(w)/I_0(w)\simeq 1$ (Main, 1971). The symbol I_i denotes a Bessel function of imaginary argument of the ith order. For the structure invariant Φ_{111} the weight w is given by (Hauptman, 1972) $w=2\,.\,\sigma_3\,.\,\sigma_2^{-3/2}\,.\,|E_HE_KE_{H-K}|$ or more exactly, (Karle & Gilardi, 1970),

$$w=2\sigma_3\,.\,\sigma_2^{-3/2}\times\frac{|E_HE_KE_{H-K}|}{1+2|U_HU_KU_{H-K}|-|U_H|^2-|U_K|^2-|U_{H-K}|^2}, \quad (2)$$

where $\sigma_i=\sum_{j=1}^{N}f_j^i$, f_j is the scattering factor of the jth atom, U_H is the unitary structure factor and N is the number of atoms in the unit cell. For the structure invariant Φ_{211} the weight w is given by Giacovazzo (1974):

$$w=2\,.\,\sigma_3^2\,.\,\sigma_3^2(2\,.\,|E_K|^2-1)|E_H^2\,.\,E_{K-H}E_{-H-K}|\,. \quad (3)$$

For the variances of the cosine of the invariant Φ_{111} and Φ_{211} relationships have been derived (Hauptman, 1972; Karle & Gilardi, 1973; Giacovazzo, 1974) which can be written as

$$\text{var}\,(\cos\Phi_{\text{theor}})=1-\frac{I_1(w)}{w\,.\,I_0(w)}-\left(\frac{I_1(w)}{I_0(w)}\right)^2 \quad (4)$$

where the weight w is given by (2) for the invariant Φ_{111} and by (3) for the invariant Φ_{211}. It can be expected for structures with a great number of atoms in the unit cell that the mean values $\langle\cos\Phi\rangle_{\text{calc}}$ and the variances $\text{var}\,(\cos\Phi)_{\text{calc}}$ calculated for a correct phase set of the structure factors will be very well approximated by the corresponding values calculated by using (1) and (4). In the case of an incorrect phase set, when on the respective Fourier map there may appear a region with negative values of the 'electron density' or exceedingly strong false peaks, the correspondence between the $\langle\cos\Phi\rangle_{\text{calc}}$ and $\text{var}\,(\cos\Phi)_{\text{calc}}$ values on the one hand and the respective $\langle\cos\Phi\rangle_{\text{theor}}$ and $\text{var}\,(\cos\Phi)_{\text{theor}}$ values on the other will be considerably poorer.

This principle is the basis of the following consistency test.

Consistency test

The structure invariants Φ were divided into p groups so that each group contained a sufficiently high number of invariants of approximately the same weight. For each group the mean weight $\bar{w}^{(i)}$ was calculated, and the corresponding mean values of $\langle\cos\Phi\rangle_{\text{theor}}^{(i)}$ and the variances $\text{var}\,(\cos\Phi)_{\text{theor}}^{(i)}$ were determined by using (1) and (4) respectively.

The values of the structure invariants Φ_{calc} were calculated for each phase set of the structure factors tested. Further, for each group of invariants the average value $\langle\cos\Phi\rangle_{\text{calc}}^{(i)}$, an estimate of the variance of

the cosine of the invariant in this group

$$\text{var}\,(\cos\Phi)^{(i)}_{\text{calc}}=\langle(\cos\Phi)^2\rangle^{(i)}_{\text{calc}}-(\langle\cos\Phi\rangle^{(i)}_{\text{calc}})^2$$

and an estimate of the variance of the mean value of the cosine of the invariant var $(\langle\cos\Phi\rangle)^{(i)}_{\text{calc}}$ were also calculated.

It may be expected that the correct phase set of the structure factors will be found among the sets giving the best fit between the values of $\langle\cos\Phi\rangle^{(i)}_{\text{theor}}$ and $\langle\cos\Phi\rangle^{(i)}_{\text{calc}}$. If we assume that $\langle\cos\Phi\rangle^{(i)}_{\text{calc}}$ values are statistically independent, a suitable measure of the agreement between these values is the weighted sum of their deviations squared

$$K=\frac{\sum_{i=1}^{p} r_i(\langle\cos\Phi\rangle^{(i)}_{\text{calc}}-\langle\cos\Phi\rangle^{(i)}_{\text{theor}})^2}{\sum_{i=1}^{p} r_i}, \qquad (5)$$

where

$$r_i=[\text{var}\,(\langle\cos\Phi\rangle)^{(i)}_{\text{calc}}]^{-1}. \qquad (6)$$

At the same time, for the correct phase set of structure factors the coefficient

$$N=\sum_{i=1}^{p}[\text{var}\,(\cos\Phi)^{(i)}_{\text{calc}}-\text{var}\,(\cos\Phi)^{(i)}_{\text{theor}}]^2, \qquad (7)$$

also must assume one of the least values. For the invariant Φ_{111} the coefficients K and N will be denoted by the symbols K_{111} and N_{111}, while for the invariant Φ_{211} the respective notation will be K_{211} and N_{211}.

Thus, the most probable phase system of structure factors can be seen in a system giving the lowest values of the coefficients K_{111} and N_{111} (or K_{211} and N_{211}).

Similarly to the case of the coefficient K (Hašek, 1974), for centrosymmetric space groups, a connexion between the coefficient K_{111} and the absolute figure of merit (Germain, Main & Woolfson, 1971) can also be pointed out. If only one weight interval is taken in the calculation of the coefficient K_{111}, *i.e.* (5) is reduced to

$$K=\langle\cos\Phi\rangle_{\text{calc}}-\langle\cos\Phi\rangle_{\text{theor}},$$

then the coefficient $M=|1-K_{111}|$ will have properties similar to those of the absolute figure of merit.

The reliability of the consistency test depends on the conditions of the validity of (1) and (4). In the case of great overlapping of the Patterson peaks, pseudosymmetry, or if the dominant effect on diffracted intensities has a low number of heavy atoms, it may occur that the correct phase set of the structure factors will not correspond to the lowest values of the coefficients K and N, but only to some of the following values. Unreliably determined values of normalized structure factors can have a similar effect.

Conclusion

There are many routes to the determination of a correct phase set of the structure factors. The advantage of the procedure just outlined compared to the majority of the other methods is seen in its simplicity: calculations using relationships (1), (2), and (5) consist only in a correct classification and averaging of values already calculated by the common direct methods.

Since the procedure used here for the calculation of the coefficients K_{111} and N_{111} – simple as it is – suppresses the frequently criticized shortcomings of the figures of merit of symmorphous space groups, it seems to be a suitable supplement to the majority of the existing direct methods employing triplet relationships for the investigation of crystal structures.

The author thanks Dr K. Huml for his helpful criticism of the manuscript.

References

GERMAIN, G. & MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.

GIACOVAZZO, C. (1974). *Acta Cryst.* A**30**, 522–529.

HAŠEK, J. (1974). *Acta Cryst.* A**30**, 576–579.

HAUPTMAN, H. A. (1972). *Crystal Structure Determination: The Role of Cosine Seminvariants*, pp. 156–157. New York: Plenum Press.

KARLE, J. & GILARDI, R. D. (1973). *Acta Cryst.* A**29**, 401–407.

MAIN, P. (1971). Summer School in York (Lecture notes D1). Physics Department, University of York, Heslington, York YO1, England.

SCHENK, H. (1972). *Acta Cryst.* A**28**, 412–422.