PRESS, W. & HÜLLER, A. (1973). *Acta Cryst.* A**29**, 252–256.
PRESS, W. & HÜLLER, A. (1978). *J. Chem. Phys.* **68**, 4465–4467.
SAKATA, M., HARADA, J., COOPER, M. J. & ROUSE, K. D. (1980). *Acta Cryst.* A**36**, 7–15.
SARMANOV, O. V. & BRATOEVA, Z. N. (1967). *Teor. Veroyatn. Ee Primen.* **12**, 470–481.
WRIGHT, A. F. & LEHMANN, M. S. (1981). *J. Solid State Chem.* **36**, 371–380.
ZUCKER, U. H., PERENTHALER, E., KUHS, W. F., BACHMANN, R. & SCHULZ, H. (1982). *PROMETHEUS*. Program system for structure refinements. Submitted to *J. Appl. Cryst.*
ZUCKER, U. H. & SCHULZ, H. (1982). *Acta Cryst.* A**38**, 563–568.

Acta Cryst. (1983). A**39**, 158–166

An Empirical Method for Correcting Diffractometer Data for Absorption Effects

BY NIGEL WALKER AND DAVID STUART*†

Chemistry Department, Queen Mary College, Mile End Road, London E1 4NS, England

(*Received* 20 *July* 1982; *accepted* 9 *September* 1982)

Abstract

Absorption effects usually present the most serious source of systematic error in the determination of structure factors from single-crystal X-ray diffraction measurements if the crystal is not ground to a sphere or cylinder. A novel method is proposed for the correction of these effects for data collected on a diffractometer. The method works from the premise that the manifestation of systematic errors due to absorption, unlike most other sources of systematic error, will not be evenly distributed through reciprocal space, but will be localized. A Fourier series in the polar angles of the incident and diffracted beam paths is used to model an absorption surface for the difference between the observed and calculated structure factors. Knowledge of crystal dimensions or linear absorption coefficient is not required, and the method does not necessitate the measurement of azimuthal scans or any extra data beyond the unique set. Moreover, application of the correction is not dependent upon the Laue symmetry of the crystal or the geometry of the diffractometer. The method is compared with other commonly used corrections and results are presented which demonstrate its potential.

1. Introduction

In the course of a single-crystal structure determination, two methods are frequently used for the correction of absorption effects for crystals of arbitrary shape.

The numerical integration method of Busing & Levy (1957) requires that the faces of the crystal are indexed and their distance from a common point within the crystal is accurately determined. Precise measurement of the crystal becomes critical for those with a high linear absorption coefficient [see equation (2.4)]. The presence of an external absorber such as crystal mother liquor or any adhesive used in the mounting of the crystal, or lack of identifiable faces, will also produce serious problems for this approach. Furthermore, without a very powerful computer, considerable time is required for the evaluation of the integral by the Gaussian algorithm of Coppens, Leiserowitz & Rabinovich (1965) to a suitable accuracy.

The semi-empirical correction method for diffractometer data of North, Phillips & Mathews (1968) requires the measurement of azimuths ψ for a single reflection by rotating the crystal about the goniometer-head axis φ (Fig. 1). The resultant transmission curve is normalized and hence provides only a relative correction. Though widely used in small-molecule structure determinations, the method has three serious limitations. An unfavourable crystal mounting for data collection on an Enraf–Nonius CAD4 diffractometer (which utilizes κ rather than Eulerian cradle geometry), can render it impossible to record a complete azimuthal scan for any reflection. This problem can be averted if the crystal is mounted so that a reciprocal-lattice axis is approximately coincident with the φ axis. Not only is it inconvenient to remount a crystal where its morphology did not allow a satisfactory initial alignment to be made, but this restraint negates the advantages of automatic orientation systems present on modern diffractometers. Secondly, an observed change in the intensity of a reflection with azimuth ψ may be due to anisotropic primary extinction (Seiler & Dunitz, 1978) or aniso-

* Supported by: Laboratory of Molecular Biophysics, Zoology Department, South Parks Road, Oxford OX1 3PS, England.
† Present address: Institute of Biophysics, Academia Sinica, Beijing, Peoples Republic of China.

0567-7394/83/010158-09$01.50

tropic secondary extinction (Coppens & Hamilton, 1970) rather than absorption. Thirdly, in practice only a single transmission curve is obtained for the correction of all data and the correction is therefore independent of the other spherical polar angle, μ (Fig. 1), which increases with increasing Bragg angle θ for these reflections. Usually, scans of several reflections at differing Bragg angle are collected, and averaged. For the standard bisecting equatorial geometry used for data collection on a four-circle diffractometer, the maximum possible value of μ is θ_{max}; see equation [A1(iv)]. Since the method was first proposed for the correction of protein data where θ_{max} is often between 7 and 18° (for 6 and 2·5 Å resolution data, respectively, collected using Cu $K\alpha$ radiation); this simplification was acceptable. The range of θ for each protein crystal is even less than θ_{max} if the data are collected in shells of increasing resolution using several crystals. However, data collected from a single crystal for a small-molecule structure determination often extend to a θ_{max} of 30° for Mo $K\alpha$ radiation and a θ_{max} of 70° for Cu $K\alpha$ radiation. Clearly a correction that is dependent on μ as well as φ is required for such data.

If a series of transmission curves for various values of μ are measured, a difficulty arises in placing each of these normalized curves onto a common scale. Kopfmann & Huber (1968) have developed a similar method to that of North *et al.* (1968) which overcomes this problem by constructing an absorption surface using general reflections as well as axial ones. However, it requires that more time is spent on the determination of an absorption correction, and it is dependent on the availability of sufficiently strong reflections at high Bragg angle. Although their method was originally intended for use on an open Eulerian cradle diffractometer, the method is better suited to a diffractometer with a full-circle Eulerian cradle rather than one using κ geometry, in order that more azimuths can be measured for the general reflections.

As an alternative to these approaches, Katayama, Sakabe & Sakabe (1972) suggested a statistical correction for absorption. They recognized that, due to the smoothly varying nature of absorption effects, a good approximation to the transmission is a Fourier series in the polar angles of both the incident and diffracted beams, where only the low-order terms are important. The values for these coefficients are obtained by the analysis of the symmetry-related reflections within the data set. This requires high lattice symmetry and that equivalent reflections do not have the same absorption factors – conditions often not fulfilled. The measurement of data beyond the unique set required for the structure determination is also necessary. The method only performs a relative correction as it cannot provide any information for a correction within the unique set, and has been formulated for the special case of bisecting equatorial diffraction geometry. Flack (1974) has since proposed a method which combines the features of an empirical correction from azimuthal scans with those of a statistical evaluation, using diffractometer Eulerian angles rather than polar angles in the Fourier series. This method has also only been formulated for the case of bisecting equatorial geometry on a four-circle diffractometer. In both this method and that of Kopfmann & Huber (1968), an approximate θ-dependent correction is provided by the application of a spherical correction obtained from tabulated absorption factors using an estimate of the 'equivalent sphere'.

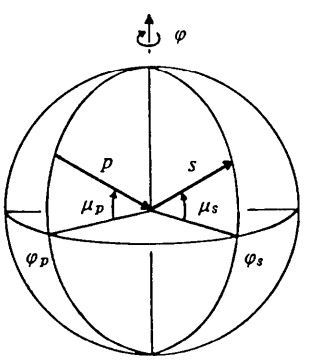

Fig. 1. Schematic representation of the spherical polar angles φ_p, μ_p, φ_s and μ_s, which refer to the primary beam direction **p** and secondary beam direction **s**.

Previous work on an absorption correction for protein data collected on a rotation camera (Stuart & Walker, 1979) has shown that a Fourier series can provide a useful way of modelling absorption differences between two sets of structure factors. In this paper, we present an extension of this method for the correction of data collected on a diffractometer. Differential absorption of the X-ray beam for the different paths taken through a crystal by each Bragg reflection will produce systematic errors that relate to areas in reciprocal space. If the true amplitudes (*i.e.* those corresponding to no absorption) were known, then an absorption surface could be constructed to model the discrepancy between the true and observed amplitudes. While the true amplitudes are, of course, not known, a good first approximation is available in the form of calculated structure factors from atomic positions and thermal parameters obtained using data suffering from absorption effects. Systematic errors that arise from the absence or inaccurate positioning of the scattering atoms used in a structure-factor calculation will tend to be evenly distributed through reciprocal space, unlike those due to differential absorption which will be localized. A Fourier series is used to model this discrepancy between the calculated and observed structure-factor amplitudes and the absorption surface obtained then used to correct the observed data.

2. The method

If F_j^o are the observed structure-factor amplitudes and F_j^c are those obtained from a structure-factor calculation, we modify the values of F_j^o by the following empirical correction:

$$F_j^m = kF_j^o A_{p,s}, \tag{2.1}$$

where k is an overall scale factor obtained from the least-squares refinement, and F_j^m are the modified values of the observed structure-factor amplitudes. The absorption coefficient, $A_{p,s}$, is represented by a Fourier series:

$$A_{p,s} = \sum_n \sum_m P_{n,m}[\sin(n\varphi_p + m\mu_p) + \sin(n\varphi_s + m\mu_s)] + Q_{n,m}[\cos(n\varphi_p + m\mu_p) + \cos(n\varphi_s + m\mu_s)], \tag{2.2}$$

where φ_p, μ_p, φ_s and μ_s are the spherical polar angles defining the incident and diffracted beams, respectively (Fig. 1). $P_{n,m}$ and $Q_{n,m}$ are Fourier coefficients whose values are obtained by minimizing the sum of the squares of the residuals, R_j, where

$$R_j = (||F_j^c| - |F_j^m||)\,\omega_j \tag{2.3}$$

and ω_j is a suitable weighting function.

This empirical function may be seen to be related to a previous model for absorption effects.

The transmission factor for any diffracted beam is:

$$T_{p,s} = \frac{1}{V}\int_V \exp[-M(x_p + x_s)]\,dV, \tag{2.4}$$

where V is the crystal volume, M is the linear absorption coefficient for the crystal and x_p and x_s are the primary and secondary beam path lengths.

North *et al.* (1968) have suggested that an approximation to the overall transmission factor $T_{p,s}$ can be obtained from transmission factors for the composite primary and secondary beams:

$$T_{p,s} = (T_p + T_s)/2. \tag{2.5}$$

If each is represented by a Fourier series as suggested by Katayama *et al.* (1972) then

$$T_{p,s} = \sum_n \sum_m P_{n,m}\sin(n\varphi_p + m\mu_p) + Q_{n,m}\cos(n\varphi_p + m\mu_p) + \sum_{n'} \sum_{m'} P_{n',m'}\sin(n'\varphi_s + m'\mu_s) + Q_{n',m'}\cos(n'\varphi_s + m'\mu_s). \tag{2.6}$$

However, only a single absorption surface is required to model both primary and secondary beam absorption effects and hence $P_{n',m'} = P_{n,m}$ and $Q_{n',m'} = Q_{n,m}$, therefore

$$T_{p,s} = \sum_n \sum_m P_{n,m}[\sin(n\varphi_p + m\mu_p) + \sin(n\varphi_s + m\mu_s)] + Q_{n,m}[\cos(n\varphi_p + m\mu_p) + \cos(n\varphi_s + m\mu_s)]. \tag{2.7}$$

The only difference between (2.2) and (2.7) is that the former is an absorption factor acting on amplitudes while the latter is a transmission factor acting on intensities. Expression (2.2) is chosen in preference since a Fourier series modelling differences in amplitudes will be more smoothly varying than a series modelling intensities.

If the form of (2.1) provides a good model for absorption effects that are due to the shape and orientation of the crystal, then the corrected amplitudes can be considered to approximate those obtained from a crystal of constant dimensions with respect to the polar angles φ and μ, *i.e.* a sphere. Integration of (2.4) for the different path lengths through a spherical crystal of linear absorption coefficient M and radius r produces an absorption factor for each diffracted beam which is solely dependent upon the Bragg angle, θ. The relative correction between reflections at low θ and those at high θ becomes increasingly severe for crystals where the product of M and r is high, *viz.* where $Mr > 1{\cdot}0$. Although values for this absorption factor are available in a tabulated form (*International Tables for X-ray Crystallography.* 1972) for $Mr \leq 10{\cdot}0$, we are unable to use these for two reasons. Firstly, we do not know the radius of the 'equivalent sphere', and, secondly, some of these effects may have been corrected for by the absorption surface previously evaluated or may be accommodated in the temperature factors assigned to the atoms. Instead, we correct the modified structure-factor amplitudes, F_j^m, using the following empirical expression:

$$F_j^{m'} = F_j^m A_\theta, \tag{2.8}$$

where the spherical absorption coefficient, A_θ, is represented by a Fourier series in θ:

$$A_\theta = \sum_n R_n \sin(n\theta) + S_n \cos(n\theta). \tag{2.9}$$

R_n and S_n are Fourier coefficients whose values are obtained in a similar manner to before; see (2.3).

Since the effects of secondary extinction are of a similar form to those of absorption (Hamilton, 1963), the use of (2.8) may also provide a correction for these systematic errors. The evaluation of $A_{p,s}$ in (2.1) rests on the assumption that beams occurring at similar polar angles have the same path length through the crystal. Any purely θ-dependent absorption effects will invalidate this and so the absorption surface is redetermined using observed structure-factor amplitudes, F_j^o, corrected for A_θ absorption effects.

The approximation of North *et al.* (1968) [(2.5)] may consequently be reformulated as

$$T_{p,s} = [(T_p + T_s)/2].T_{p-s}, \quad (2.10)$$

where $\mathbf{p} - \mathbf{s} = 2\theta$.

3. Implementation

A computer program *DIFABS* evaluates the Fourier series coefficients required to model the absorption surface of the crystal, and applies the correction to the observed data. The program is written entirely in Fortran IV and runs in 23K words of memory on a DEC PDP11/34 mini-computer under the RSX-11M operating system. The program integrates into the Enraf–Nonius *SDP* suite of programs for processing diffractometer data: however, for generality, the program works throughout in Eulerian geometry. The correction of a set of 3000 reflection data takes approximately 30 mins, or 15 mins, if a θ-dependent correction is not performed. This compares favourably with the time taken for the Gaussian method of Coppens *et al.* (1965) on the same computer: 6 h if 1000 grid points are used for a crystal with six faces.

The matrix containing the normal equations is solved by Gauss–Jordan elimination. It was found necessary to make the elements of the array 'double precision', *viz.* four 16-bit words with accuracy to 16 decimal places, to prevent the order of elimination of the variables altering their value.

The correction is evaluated using observed structure factors, F_j^o, corrected for Lorentz–polarization effects and any X-ray exposure-dependent decay in intensity, with systematically absent reflections removed. The correction is calculated and applied to the data set before any symmetry-equivalent or Friedel-pair reflections are averaged. However, the calculated amplitudes, F_j^c, will have been obtained from the refinement of the unique data set. The structure-factor calculation should be performed after isotropic least-squares refinement with as many of the atoms in the molecule(s) that can be found using difference Fourier-map techniques (usually all non-hydrogen atoms). The correction should not be done against structure factors calculated after anisotropic refinement as it is well established that temperature-factor ellipsoids are capable of accommodating some of the gross absorption effects. If difficulty in locating the atoms in crystals with a high linear absorption coefficient is encountered, a preliminary correction against structure factors calculated from just the heavy-atom refined position(s) may prove useful (see § 4).

The weighting function, ω_j, used is $1/\sqrt{(k\sigma_j)}$, where k is the overall scale factor in (2.1). This was chosen in preference to $1/(k\sigma_j)^2$ or $1/(k\sigma_j)$ so as not to weight down unduly the high-angle data which will be relatively weaker. The value of σ_j for each reflection is derived from the counting-statistics standard deviation, σ_j^I:

$$\sigma_j = \sigma_j^I/2F_j^o \text{ Lp}, \quad (3.1)$$

where Lp is the Lorentz–polarization factor.

All reflections, except those which are unobserved, are corrected for absorption, but only those which fulfil the following criteria are used in the evaluation of the Fourier coefficients in (2.2) and (2.9):

$$\text{(i) } F_j^o \geq 3\sigma_j$$
$$\text{(ii) } F_j^c \geq 3k\sigma_j \quad (3.2)$$
$$\text{(iii) } F_j^c \leq 2kF_j^o.$$

Criterion (iii) is to prevent the use of any reflection likely to be suffering from severe extinction effects.

The Fourier series in (2.2) and (2.9) were constructed, by experimental determination, from the minimum number of terms commensurate with a good model as judged by the r.m.s. residual, R_a, defined as

$$R_a = \left[\sum_j ||F_j^c| - |F_j^m||^2 / \sum_j |F_j^c|^2\right]^{1/2} \times 100. \quad (3.3)$$

R_a is calculated using only those reflections which satisfy the criteria in (3.2).

The coefficients used are $P_{n,m}$ and $Q_{n,m}$, where (*a*) $n = 1, 2, 3, 4, 5, 6$ and $m = 0$ for pure φ terms, (*b*) $n = 0$ and $m = 1, 2, 3, 4$ for pure μ terms and (*c*) $n = 1, 2, 3, 4$ and $m = 1$, and $n = m = 2, 3, 4$ for cross terms: a total of 34. The coefficients R_n and S_n used for the θ-dependent correction are $n = 1, 2, 3, 4, 5, 6$: a total of 12. The zero-order terms of both series are set to unity and if taken outside the summation, then (2.1) and (2.8) can be re-written, respectively, as

$$F_j^m = kF_j^o(1 + A'_{p,s}) \quad (3.4)$$

$$F_j^{m'} = F_j^m(1 + A'_\theta). \quad (3.5)$$

The range for μ is dependent on the maximum Bragg angle of the data and the mounting of the crystal, the maximum Bragg angle being dependent on the maximum resolution of the data and the wavelength of the radiation used. Thus for a crystal mounted about a reciprocal-lattice axis and illuminated with Mo $K\alpha$ radiation the range may be as little as 0 to 20°. The range for φ is dependent on the crystal-lattice system (a full 360° for the unique data set in triclinic or monoclinic cells). Since the range for μ is limited when compared to that for φ, the cross terms chosen are weighted towards φ dependence. It is important that sufficient pure μ terms are included to enable an accurate modelling in this direction, but without introducing too much flexibility at the extremes where the function is less well conditioned. Reflections with a high μ value will be those at high scattering angles and will therefore be relatively weaker and hence given a lower weighting, ω_j. In order to provide a better modelling at high μ, the distribution of μ is adjusted by taking the square root of μ for this polar angle in (2.2). The sign is retained after taking the root of the

modulus. Since μ is now much smaller, this results in the cross terms becoming dominated by φ. The signed square root is therefore scaled by a constant so that μ_{max} approximately returns to its original value: we have found that a wavelength-dependent constant (7λ) works well. The overall result of this process is simply to redistribute the reflections into more equal groups with respect to ranges of μ, making the function better determined for the high-angle data.

In calculating the polar angles for the primary and secondary beam, the method does not make the assumption on the equivalent absorbence of a reverse beam which is made by the method of North *et al.* (1968); namely that $A_p = A_{-p}$, $A_p = A_s$ and $A_p = A_{-s}$. Huber & Kopfmann (1969) recognized that in the presence of an external absorber the only relationship between A_p and A_s to remain valid is that $A_p = A_{-s}$. Our method only requires the use of this latter relationship in order that a single absorption surface may be used to model both primary and secondary beam absorption effects.

Formulae for the calculation of the polar angles from the Eulerian diffractometer angles are given in the Appendix for the case of bisecting equatorial geometry used with a four-circle diffractometer.

4. Tests of the method

Table 1 summarizes the results obtained from nine crystals on which the method has been tested. The crystals used represent a cross section of crystal morphology, linear absorption coefficient, number of data, lattice symmetry and source of radiation, commonly experienced in organic and inorganic structure determinations. The choice of Fourier coefficients described in (§ 3) was used throughout.

Table 2 shows the results of least-squares refinement with data before and after absorption correction. The R factors R_1 and R_2, are defined as follows:

$$R_1 = \left[\sum_j \| F_j^c | - | F_j^o \| / \sum_j | F_j^o | \right] \times 100 \qquad (4.1)$$

$$R_2 = \left[\sum_j \| F_j^c | - | F_j^o \|^2 / \sum_j | F_j^o |^2 \right]^{1/2} \times 100. \qquad (4.2)$$

R factors were calculated following isotropic least-squares refinement using all non-hydrogen atoms for the uncorrected data and the data after correction by (i) the numerical method of Coppens *et al.* (1965) using 1000 grid points, (ii) the semi-empirical method of North *et al.* (1968), and (iii) using *DIFABS*. Use of the numerical method was not possible in many cases because equi-dimensional crystals were selected in preference to ones with more 'perfect' morphology so as to reduce the relative absorption effects and hence allow a better correction to be made with the semi-empirical method. It will be noticed that, in all cases, correction with *DIFABS* produced the greatest reduction in the R factors. The 'final' R factors for each structure were calculated with *DIFABS* corrected data using anisotropic thermal parameters for the non-hydrogen atoms and isotropic parameters for all the hydrogen atoms, except in the cases of the triosmium crystals where not all of the hydrogen-atom positions could be found. Full-matrix least-squares refinement was performed using all reflections greater than $3\sigma_I$ [see (3.1)], and unit weights. After applying the absorption

Table 1. *Data used to test the method*

The third crystal dimension is in the direction of the φ rotation axis. The linear absorption coefficient given is for the $K\alpha$ of the radiation in the previous column. μ is defined in Fig. 1. The r.m.s. residual, R_a, is defined in the text.

Crystal	Compound formula	Crystal system: space group in parentheses	Approximate crystal dimensions (mm)	Radiation	Linear absorption coefficient (cm^{-1})	θ_{max} (°)	μ_{min} (°)	μ_{max}	Total number of reflections: non-zero in parentheses	Number of reflections used in evaluation of Fourier coefficients	R_a Before *DIFABS* correction (%)	R_a After *DIFABS* correction (%)
SROAC	$[Sr(CH_3COO)_2]_2 .H_2O$	Triclinic ($P\bar{1}$)	0·2 × 0·15 × 0·6	Mo	77·3	25	−1	23	2553 (2383)	2148	7·2	4·6
ICIOD	$C_{22}H_{26}N_2O_2$	Monoclinic ($P2_1/c$)	0·2 × 0·05 × 0·4	Cu	6·3	60	−5	55	2809 (2342)	1685	12·7	10·8
INDISI	$In_2Si_4N_2(CH_3)_{16}$	Monoclinic ($C2/c$)	Crystal not available	Mo	17·9	25	−21	24	2388 (2314)	2097	7·8	5·4
BROXO	$BrC_8H_{13}O_2$	Monoclinic ($C2/c$)	Crystal not available	Mo	48·8	25	−1	23	1539 (1319)	1047	14·7	7·4
TRIOS	$HOs_3(CO)_{10}$ $NCONHN(CH_3)_2$	Monoclinic ($P2_1/n$)	0·35 × 0·25 × 0·45	Mo	172·7	25	−7	23	3891 (3474)	2903	10·8	6·3
OSONE	$H_2Os_3(CO)_8$ $(C_5NH_4)_2$	Monoclinic ($P2_1/c$)	0·3 × 0·15 × 0·2	Mo	173·5	25	−8	25	3862 (3446)	2846	14·7	6·4
IRINS	$IrCl_4[S(CH_3)_2]_2$ $(CH_3)_4N$	Orthorhombic ($Fmmm$)	0·45 × 0·15 × 0·6	Mo	80·0	25	1	25	473 (473)	473	24·9	6·5
OSBPY	$H_2Os_3(CO)_8$ $(C_5NH_3CH_2C_6H_5)_2$	Orthorhombic ($P2_12_12_1$)	0·2 × 0·1 × 0·25	Mo	118·1	25	0	24	3596 (3317)	2701	12·7	6·8
OSFME	$Os_3(CO)_{10}$ (CHCHO)	Orthorhombic ($Aba2$)	0·25 × 0·1 × 0·3	Mo	227·9	30	0	29	2531 (2296)	1769	18·4	8·1

Table 2. *Results of least-squares refinement from data used to test the method*

Crystallographic R factors, R_1 and R_2, are defined in the text. The R factors for the uncorrected data and data after absorption correction were obtained from isotropic least-squares refinement. The 'final' R factors were from least-squares refinement of *DIFABS* corrected data using anisotropic thermal parameters for the non-hydrogen atoms and isotropic parameters for the hydrogen atoms.

	Before correction		After correction of Coppens *et al.* (1965)		After correction of North *et al.* (1968)		After correction with *DIFABS*		'Final'	
Crystal	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2	R_1	R_2
SROAC	6·0	7·2	4·6	5·5	4·4	5·3	4·1	4·6	2·3	2·9
ICIOD	13·5	12·8			13·4	12·7	12·1	11·0	5·4	5·0
INDISI	7·3	7·8					4·9	5·5	2·2	2·5
BROXO	14·4	14·7			12·6	12·7	7·2	7·4	3·5	3·7
TRIOS	9·1	10·9			7·1	7·9	5·7	6·3	3·5	4·1
OSONE	12·0	14·6	6·2	7·2	6·4	7·6	5·4	6·3	4·0	5·0
IRINS	20·9	24·2			7·4	9·3	5·2	6·0	3·8	4·6
OSBPY	11·2	13·2			5·9	7·4	5·8	7·2	4·2	5·7
OSFME	15·5	19·0			8·7	9·8	6·5	7·8	5·1	6·4

correction as described, and re-refining the atomic positions and isotropic temperature factors, no further correction to the data could be obtained by re-running *DIFABS* on the corrected data.

Results from crystals with a low linear absorption coefficient show very little decrease in the r.m.s. residual R_a or crystallographic R factors after absorption correction with *DIFABS*. Crystal ICIOD with a linear absorption coefficient of 6·3 cm^{-1} (Cu $K\alpha$) caused the residual to decrease less than 2% (12·7 to 10·8%, see Table 1) even though the crystal was a very thin plate. Other crystals tested which have a linear absorption coefficient of less than 1 cm^{-1} (crystals of organic compounds collected using Mo radiation) have given reductions in R_a of less than 1%. These results demonstrate that the modification of the structure-factor amplitudes by the method is due to a modelling of absorption effects, when they exist, and that the method does not just 'rescale' the observed data so as to be more similar to the calculated structure factors, thus reducing the R factor.

For crystal OSONE, a subsequent test was made using a correction against structure factors calculated from just the isotropically refined three osmium positions. R_1 and R_2 decreased from 15·2 and 19·2% for the uncorrected data to 10·9 and 13·4%, respectively, after correction with *DIFABS*. The remaining twenty-eight non-hydrogen atoms were all clearly visible in a difference Fourier map, and when included in isotropic refinement, decreased the R factors to 6·0 and 6·7%, respectively. These R factors are only 0·5% higher than those obtained after a *DIFABS* correction where all the non-hydrogen atoms were used in the calculation of structure factors. The non-hydrogen-atom peak heights in the difference Fourier were between 7·3 and 4·0 e Å^{-3} while 'ghost' peaks around the three osmium atoms were of an average density of 3·8 e Å^{-3}. Only one atom had a peak height less than those of the heavy-atom ghost peaks. In contrast, without any absorption correction a difference Fourier produced peaks at the atomic positions of a similar height (7·1 to 3·6 e Å^{-3}), however these were all lower than the ripple peaks around the osmium atoms (average height 10·5 e Å^{-3}) and the atom corresponding to the weakest peak in the map with *DIFABS* corrected data was not visible at all. Application of a numerical absorption correction similarly reduced the height of these ripples to an average of 4·7 e Å^{-3}. This demonstrates that these ripples are due to absorption errors rather than representing a residual (removable by *DIFABS*) when isotropic temperature factors are used to model true anisotropic thermal motion.

To facilitate the location of the hydrogen atoms in each structure, the non-hydrogen atoms were refined using anisotropic thermal parameters with absorption-corrected data. All three methods of correction produced R factors after this refinement which were very similar. However, for BROXO, where the correction of North *et al.* (1968) did not make a significant improvement (as judged from R factors after isotropic refinement), the thermal ellipsoids from anisotropic refinement were considerably elongated along one major axis (the axis coincident with the φ axis), compared with those from *DIFABS* corrected data. If the thermal-ellipsoid parameters for the atoms in a molecule are capable of removing absorption effects, then, conversely, any true 'overall' anisotropy present in a molecule may be masked by the use of *DIFABS*. Thermal motion of this nature may have been present in crystal BROXO since the crystal underwent X-ray exposure decay in intensity during data collection (at 263 K) and the prolate ellipsoids for the carbon atoms of the eight-membered ring of this molecule were approximately perpendicular to the plane of the ring, and therefore physically reasonable. Unfortunately, the decomposition of the crystal after irradiation prevented its measurement and hence a comparison with a numerical absorption correction could not be made. Hydrogen atoms could be seen in difference Fourier maps in all structures after absorption correction of the

data, and the peak heights were similar with all three methods of correction. The most notable improvement in the difference maps after anisotropic refinement with data corrected using *DIFABS* was the absence of any significant ripple peaks around heavy atoms.

Test using generated data

The method has also been tested using a set of calculated data to which absorption errors, evaluated using the numerical method of Coppens *et al.* (1965), were applied. The CALC data used were calculated structure-factor amplitudes from the coordinates and isotropic temperature factors of the non-hydrogen atoms of crystal SROAC. This crystal is triclinic, space group $P\bar{1}$, with cell dimensions $a = 7{\cdot}23$, $b = 9{\cdot}93$, $c = 10{\cdot}64$ Å, $\alpha = 83{\cdot}6$, $\beta = 82{\cdot}0$ and $\gamma = 74{\cdot}0°$, and has a linear absorption coefficient of 71·7 cm^{-1} (Mo $K\alpha$) for the non-hydrogen atoms. The crystal shape generated was an irregular six-sided polygon of dimensions 0·35 × 0·30 × 0·15 mm along the direct crystal axes. A grid of 14 × 12 × 6 was used in the Gaussian integration and the maximum and minimum transmission factors (on intensity) applied to the data were 53·0 and 25·9%, respectively. After scaling the TEST data suffering absorption effects to the CALC data, the r.m.s. residual, R_a, was 10·4%. This residual dropped to 1·1% after correction with *DIFABS* for just an absorption surface and 0·9% with the inclusion of the θ-dependent correction. To simulate the situation that would be encountered during the application of this method in the course of a structure determination where the CALC amplitudes are not known, a further test was performed. The atom positions and isotropic thermal parameters were refined using the TEST data and a correction applied using structure factors calculated from these new parameters. During the course of this refinement, the atomic positions moved with an r.m.s. shift of 0·015 Å, and R factors R_1 and R_2 were 9·1 and 10·3%, respectively. The r.m.s. residual between the TEST data and the new calculated structure factors was again initially 10·4% but decreased to 1·2% after correction (the r.m.s. residual between these corrected data and the original CALC data was only 1·0%). Isotropic refinement using these corrected data shifted the atomic coordinates back to within an r.m.s. deviation of 0·002 Å from their original correct positions. The final R factors from this refinement were 0·7 and 0·9% for R_1 and R_2 respectively. By comparison, if the TEST data were refined anisotropically, not only did the coordinates shift by a similar r.m.s., 0·015 Å, but the residual between the TEST data and the data calculated from the new atomic parameters fell to 7·4%. *DIFABS* correction using these calculated data only allowed the r.m.s. residual to drop to 5·3% (5·0% if the corrected data were compared with the CALC data). The use of anisotropic refinement before absorption correction clearly inhibited an accurate modelling and correction of the absorption effects. Isotropic refinement did not have this effect (a residual of 1·0 compared to 0·9% without refinement), and the shift in the coordinates caused by refinement against data suffering absorption effects did not prevent a correction being made that reversed this error.

5. Discussion

The application of the method described has shown that it can make a significant improvement in the accuracy of structure determinations for a wide variety of crystals, not only when compared with results from uncorrected data but also when compared with other absorption correction methods. Results using generated test data suggest that errors in calculated structure factors, due to inaccurate atomic parameters from least-squares refinement with data suffering absorption effects, do not prevent the method from providing a valid correction. The validity is judged by the reduction in R factors and the movement of atomic coordinates to their true positions upon refinement using the corrected data. The method possesses the advantage of not having to make any extra intensity measurements, index faces or measure the crystal dimensions, and requires little computing time. Since the method uses only extant data, the unavailability of the crystal used for data collection does not prevent an absorption correction being made to the data, on reflection. The major disadvantage of the method lies in the correlation of absorption effects with the modelling of thermal motion of the atoms, and without the addition of any extra information these effects cannot be separated. However, the results that have been obtained with the use of *DIFABS* demonstrate that for structure determinations where details of accurate bond lengths and angles are of prime importance, the use of *DIFABS* provides a quick and easy solution to the problem of absorption. Where information about the anisotropic motion of the atoms is of interest then the use of *DIFABS* may still provide a valid method of correction for absorption, but the results obtained with crystal BROXO suggest that it should be used with more caution. In the case of data suffering from extreme absorption effects a better correction may be obtained if a semi-empirical correction can be determined and applied to the data before *DIFABS* is used. Any errors introduced due to the lack of μ dependency of this method will be smoothly varying and hence can be removed with *DIFABS*. However, with severely absorbing crystals these data will provide a far better starting data set for initial isotropic refinement and calculation of structure factors for use in *DIFABS* than an uncorrected data set. Crystal OSFME with a linear absorption coefficient of 227·9 cm^{-1} (Mo $K\alpha$) gave

azimuthal scans where a decrease of 77% in the transmitted intensity was observed. R factors R_1 and R_2 after isotropic refinement with data corrected with this combined method were 6·0 and 6·5% respectively, a significant improvement when compared to the results obtained when either method was used alone (see Table 2). A similar improvement was obtained for the 'final' R factors, 4·4 and 5·0%. Finally, where very accurate structure factors are required, *e.g.* for the study of the deformation electron density of chemical bonds, and a numerical absorption correction has been made using a large number of grid points, *DIFABS* may provide a useful way of correcting for small errors introduced by inaccuracies in the measurement of the crystal or any external absorber that may have been present.

Although the method has been developed and tested using small-molecule crystals, the method could equally well be applied to protein data collected on a diffractometer in the advanced stages of a structure elucidation when a structure-factor calculation can be performed. Since a complete data set for these macromolecules is usually collected using more than one crystal, the method will serve to place all such subsets onto a common scale as well as providing an absorption correction. The frequent presence of an asymmetric external absorber, *viz.* crystal mother liquor, will also be properly accounted for during correction.

The method is not limited to correcting diffractometer data; indeed a similar version of this approach has been used for the specialized case of protein data collected on a rotation camera (Stuart & Walker, 1979). The method simply requires that the diffraction geometry for the primary and secondary beams is translated into a spherical polar coordinate system.

A possible extension of this method is to incorporate the determination of the Fourier coefficients in (2.2) and (2.9) as extra variables in the least-squares refinement procedure for evaluation during isotropic refinement.

The crystals used in the course of this study were provided by users of the SERC Chemical Crystallographic Service at Queen Mary College. We are grateful to the SERC for both financial support and the provision of the facilities.

APPENDIX
Calculation of polar angles from diffractometer settings

For the standard bisecting position in equatorial geometry used with four-circle diffractometers, the polar angles of the incident and diffracted beam (Fig. 1) can be evaluated from the diffractometer Eulerian angles (θ,χ,φ) as follows:

$$\begin{aligned}
&\text{(i)}\ \mu_p = \mu_s = \mu \\
&\text{(ii)}\ \varphi_p = \varphi + \Delta\varphi \\
&\text{(iii)}\ \varphi_s = \varphi + 180 - \Delta\varphi \\
&\text{(iv)}\ \sin\mu = \sin\theta\sin\chi \\
&\text{(v)}\ \cos\Delta\varphi = (1 + \tan^2\theta\cos^2\chi)^{-1/2}.
\end{aligned} \qquad (A1)$$

The method is neither dependent on these relationships, nor the type of diffractometer or scan method used. For other diffraction geometries, formulae for the polar angles φ_p, μ_p, φ_s and μ_s should similarly be derived from the diffractometer axial system. Equations $(A1)$ are essentially the same as those given by Katayama *et al.* (1972), except that their value for $\Delta\varphi$ is equivalent to $90 - \Delta\varphi$ in $(A1)$. However, their equation for the evaluation of $\Delta\varphi$:

$$\tan\Delta\varphi = \cos\theta/(\cos\chi\sin\theta), \qquad (A2)$$

is indeterminate if $\chi = 90°$.

An approximation commonly used for $\Delta\varphi$, instead of equation $[A1(v)]$, in the application of the method of North *et al.* (1968), is

$$\Delta\varphi = \theta\cos\chi. \qquad (A3)$$

We have investigated the validity of this assumption and Fig. 2 shows a plot of $\Delta\varphi$ against χ for various Bragg angles, θ. It can be seen that the approximation in $(A3)$ is acceptable at low θ but at high θ leads to inaccuracies, except for reflections with values of χ near 0 or 90°. The approximation is therefore suitable for use with the correction of North *et al.* (1968), but not with *DIFABS*.

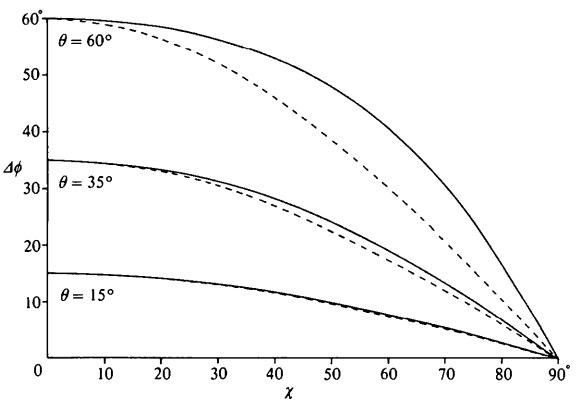

Fig. 2. Graph showing the relationship of $\Delta\varphi$ and χ at $\theta = 15$, 35 and 60° for (*a*) $\cos\Delta\varphi = (1 + \tan^2\theta\cos^2\chi)^{-1/2}$ (solid lines) and (*b*) $\Delta\varphi = \theta\cos\chi$ (dotted lines).

References

BUSING, W. R. & LEVY, H. A. (1957). *Acta Cryst.* **10**, 180–182.
COPPENS, P. & HAMILTON, W. C. (1970). *Acta Cryst.* A**26**, 71–83.
COPPENS, P., LEISEROWITZ, L. & RABINOVICH, D. (1965). *Acta Cryst.* **18**, 1035–1038.
FLACK, H. D. (1974). *Acta Cryst.* A**30**, 569–573.
HAMILTON, W. C. (1963). *Acta Cryst.* **16**, 609–611.
HUBER, R. & KOPFMANN, G. (1969). *Acta Cryst.* A**25**, 143–152.
International Tables for X-ray Crystallography (1972). Vol. II, 3rd ed., pp. 291–312. Birmingham: Kynoch Press.
KATAYAMA, C., SAKABE, N. & SAKABE, K. (1972). *Acta Cryst.* A**28**, 293–295.
KOPFMANN, G. & HUBER, R. (1968). *Acta Cryst.* A**24**, 348–351.
NORTH, A. C. T., PHILLIPS, D. C. & MATHEWS, F. S. (1968). *Acta Cryst.* A**24**, 351–359.
SEILER, P. & DUNITZ, J. D. (1978). *Acta Cryst.* A**34**, 329–336.
STUART, D. & WALKER, N. (1979). *Acta Cryst.* A**35**, 925–933.

Acta Cryst. (1983). A**39**, 166–170

A Method for Refinement of Partially Interpreted Protein Structures Including a Procedure for Scaling Between a Model and an Electron-Density Map

BY T. N. BHAT AND D. M. BLOW

Blackett Laboratory, Imperial College of Science and Technology, London SW7 2AZ, England

(*Received* 22 *February* 1982; *accepted* 26 *August* 1982)

Abstract

A method for the structure-factor least-squares refinement of a poor and incomplete atomic model of a protein molecule is suggested. Usually structure-factor least squares treat calculated phases as an observable to be associated with observed amplitudes of the complete model and assume that the difference between observed and calculated structure factors arises from the inaccuracies of the model. It is proposed that the assumption may be made valid for an incomplete model by including, in the calculation of the structure factors, electron-density features observed in the region outside the existing model. Structure factors obtained from the atomic model and from electron density may be scaled by fitting the radial distribution function in reciprocal space and their relative weights may be determined by an *R*-factor search procedure.

Introduction

Often an electron-density map for a protein, obtained by the multiple isomorphous replacement technique (MIR), can only be partly interpreted in terms of a molecular model. We recently described a density-modification procedure for the cyclic improvement of a partially interpreted electron-density map (Bhat & Blow, 1982). This procedure makes the map more readily interpretable but it does not include any method of adjusting the model coordinates so as to give a better fit to the improved map. (The only change it makes to the model is to change the 'occupancy' of atoms, depending on the values of the electron density in the vicinity of the atom site.)

It seems very desirable to introduce a procedure for refinement of the atomic model (which is likely to be tentative and incomplete) within each cycle of the procedure. In this paper we suggest how this can be done, increasing the power of the method, and reducing the amount of manual intervention. This provides a general method for refinement of an incompletely interpreted structure which is superior to refinement methods previously used.

The parameters of a diffraction pattern, in an ideal sense, are the amplitudes and the associated phase angles. Because the phases are not observed in X-ray diffraction, a number of different approaches to refinement of atomic parameters are possible, depending on the treatment of phases between cycles. There are three main choices,

(1) to keep the phases fixed, say at values obtained from the data [*e.g.* multiple isomorphous replacement (MIR)] – a phase-locked procedure;

(2) to recompute phases between two refinement cycles based on the refined parameters – a phase-free procedure;

(3) to restrain the recomputed phases towards previously obtained values – a phase-restrained procedure.

A general discussion of these types of approach is given by Diamond (1976).

The phase-locked refinement may be visualized as a technique of moving the model to achieve a better fit to the adjacent electron density. The electron density which defines the target positions for the model remains unchanged because of fixed phase angles. Early application of real-space refinement demonstrated a definite limit to the improvement which can be obtained

0567-7394/83/010166-05$01.50