Acta Cryst. (1974). A30, 740

Data Collection in Protein Crystallography: Capillary Effects and Background Corrections*

BY M. KRIEGER†, J. L. CHAMBERS, G. G. CHRISTOPH‡, R. M. STROUD§
Crellin Laboratory of Chemistry

AND B. L. TRUS¶
Norman W. Church Laboratory of Chemical Biology, California Institute of Technology, Pasadena, California 91109, *U.S.A.*

(*Received* 16 *January* 1974; *accepted* 24 *May* 1974)

In protein crystallography, observed diffraction intensities must be corrected for background radiation due to scatter from air and scatter and absorption by capillary, crystal and mother liquor. A systematic study shows that a major contribution to background intensity is air scatter arising from the air intercepted by the direct X-ray beam as 'seen' by the receiving-counter aperture. As a result there is a first-order dependence of background on the 2θ angle. The second-order variations in this function are principally due to absorption of the direct beam or air-scattered radiation by the capillary and to diffraction by the glass in the direct beam. To reduce data collection time and crystal exposure, individual background measurements may be approximated by interpolation from empirical background curves or, alternatively, by collecting background intensities for short times and fitting these data with a multidimensional function. If isotropic interpolation is used, *i.e.*, if background is considered to be a function of 2θ alone, systematic errors of up to about 30% can be introduced into the interpolated backgrounds. Methods of accounting for the anisotropy in the background are derived and shown to reduce this error to 1–2%.

Introduction

Because protein crystals are quite susceptible to radiation damage, crystallographers have been exploring different methods for accurate data collection which minimize the X-ray exposure of the crystal. With automated diffractometers, one widely used method of data collection involves counting the background on one or both sides of each measured reflection. There are several methods of reducing the time of data collection and thereby increasing the number of reflections collected per crystal (*e.g.*, Wyckoff, Tsernoglou, Hanson, Know, Lee & Richards, 1970; Watson, Shotton, Cox & Muirhead, 1970). In one method, the observation of individual backgrounds is omitted and background corrections for the measured intensities are calculated from an empirical curve of background *versus* $\sin\theta/\lambda$ measured for each crystal.

If the backgrounds do not vary significantly during the time of collection, almost all of the crystal exposure time can be devoted to intensity data collection, and empirical background curves can be measured after the intensity data have been collected. Empirical background curves have generally been obtained from background values measured with long counting times for points along one lattice row, and applied using the approximation that the background depends only on the Bragg angle and not on the other setting angles (Matthews, Levine & Argos, 1972; Jensen, 1972; Salemme, Freer, Xuong, Alden & Kraut, 1973; Wyckoff, Doscher, Tsernoglou, Inagami, Johnson, Hardman, Allewell, Kelly & Richards, 1967). This approximation can give rise systematic errors of about 30% between interpolated and observed backgrounds.

Hill & Banaszak (1973) have reported observing an additional, 2θ-invariant φ dependence. Our experiments show that the background radiation can vary systematically with φ and χ in addition to 2θ, and furthermore that such φ and χ dependences are functions of 2θ. These dependences can be accounted for by additional components in a simple interpolative procedure. In contrast to Hill & Banaszak, we conclude that the variation with φ arises primarily from capillary absorption and scattering rather than absorption of background radiation by the protein crystal and its mounting to the capillary. There is substantial improvement in interpolated backgrounds if the φ and χ-dependent variations are accounted for.

A second approach to streamlining data collection involves observing very short backgrounds with each reflection, and fitting, by least squares, a function of the diffractometer setting angles to all of the data. Such a function must include cross terms between χ and 2θ, and φ and 2θ, and can vary in complexity depending upon the range of 2θ and the experimental conditions. By pooling the data in this manner, the resulting backgrounds are more accurate than the individual short measurements from which the background function was constructed.

* Contribution No. 4726. Supported by U.S. Public Health Service Grants GM-19984 and GM-12121.
† Danforth Foundation Fellow.
‡ National Institutes of Health Postdoctoral Fellow.
§ National Institutes of Health Career Development Awardee, U.S. Public Health Service Grant No. GM-70469.
¶ Jane Coffin Childs Memorial Fellow.

It should be noted that systematic variations between real and interpolated values for the background can give rise to systematic errors in phase determination, in ΔF terms used in calculating difference maps, and in scaling. Phases and ΔF's free from such errors are crucial to accurate descriptions of molecular shifts from difference maps (Dickerson, Kopka, Varnum & Weinzierl, 1967; Henderson & Moffat, 1971; Krieger, Kay & Stroud, 1973). Furthermore, with protein crystals whose reflection intensities are small, a reflection which is only twice as intense as the background may be ten standard deviations above background. Therefore, well determined reflection intensities ($\gtrsim 3\sigma$) may be in error by as much as 30% if anisotropy in the background is not accounted for.

As the background intensity is independent of crystal absorption, the backgrounds, estimated or measured, should be subtracted from the observed diffraction intensities before applying the standard absorption and Lp corrections.

In this paper we describe two methods for obtaining background corrections of similar quality to individually measured backgrounds, but at significant savings in the time spent on observing backgrounds. Although it is desirable to minimize the background present (*i.e.*, improve the signal-to-noise characteristics of the experiments), we are primarily concerned here with developing background correction methods whose use would be applicable to a wide variety of existing four-axis diffractometers, without modification of the experimental hardware. As a result of this analysis, and the consideration of the origins of components in the background intensity, experimental methods of reducing the background, and thereby increasing the signal-to-noise ratio, are suggested.

Experimental

Measurements were made on several different diffractometers both with and without a monochromator. Details of each experimental arrangement used are described in Table 1. Each machine will henceforth

Table 1. *Characteristic data for diffractometers used in this study*

In all cases the basic four-circle machines have been rebuilt and modified; in all cases except *A* and *D* modifications have been extensive to the point of total remachining.

	(*A*) Syntex $P\bar{1}$	(*B*) GE XRD-490	(*C*) Datex–GE hybrid	(*D*) Hilger–Watts*	(*E*) Buerger–Supper†	(*F*) S. Samson‡ hybrid
Source						
Target size (mm)	10 × 1	12·5 × 0·8	12·5 × 0·8	10 × 1	8 × 0·4	12·5 × 0·8
Voltage (kV)	40	45	45	46	33	45
Current (mA)	20	18·5	15	16	20	18
Radiation	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Cu $K\alpha$	Mo $K\alpha$
Geometry						
Take-off-angle	6°	–	3°	3°	3°	–
Monochromator type	Graphite	Graphite	None	None	Graphite	Graphite
Monochromator mounting (ϱ)§	90°	90°	–	–	–	0°
Monochromator dispersion	0·3°	0·3°	–	–	–	0·3–0·4°
Source to crystal						
Distance from source to:						
1st aperture(cm) (size in mm)	5·5 (2·0)	4·5 (1·0)	–	–	8·0 (2·0)	4·0 (1·0)
Monochromator crystal	6·8	5·0	None	None	9·0	5·5
Aperture (size)	19·8 (1·0–1·5)	8·0 (1·0)	4·5 (1·5)	~4·0 (1·0)	11·5 (0·6)	12·0 (1·5)
Aperture (size)	21·3 (1·5–2·0)	14·5 (1·8)	14·3 (1·8)	~9·5 (1·0)	15·9 (0·6)	20·8 (1·8)
Crystal	27·3	15·5	15·0	23·5	19·4	22·5
Crystal to counter ‖						
Distances from crystal to:						
Aperture (size)	32·5 (2·5)	4·0 (1·0–1·5)	9·2 (1·0)	17·0 (3·0)	6·0 (2·2)	8·0 (2·11)
Aperture (size)	38·5 (1·0–2·0)	12·0 (2·0–2·5)	–	22·5 (3·0)	11·8 (1·5–2·0)	11·4 (5·04)
Counter	40·5	16·0	22·2	25·0	13·5	21·6
Noise count rate with X-rays						
off (per 100 s)	35	55	25	38	30	13

* Standard four-circle machine with no monochromator.

† This machine was essentially redesigned by, and rebuilt under the direction of, R. M. Stroud, and uses Datex automation for operation in a φ-step scan mode (two-circle Weissenberg geometry).

‡ Four-circle diffractometer assembled locally under the direction of S. Samson. Electronics and Alloys goniostat, Varian computer, Datex resoverdynes, and Ortec counter chain.

§ ϱ is the angle between the normals to the planes of incidence at the monochromator.

‖ Counter type: all diffractometers use scintillation counter. Hermetically sealed Tl-drifted NaI crystals are individually selected and checked regularly for decay and damage.

be referred to by a letter (*A–F*) which in turn corresponds to one arrangement in Table 1. Each of the figures refers to a specific case, although the discussion considers results obtained from many different experiments. It should be noted that these different arrangements include diffractometers both with and without monochromators, with both of the generally used monochromator settings, with copper and molybdenum $K\alpha$ radiation, and with or without helium-filled pathways to the counter.

Glass capillaries were 0·3–1·0 mm in diameter with walls approximately 0·01 mm thick. Crystals of trypsin, trypsinogen, cytochrome *c* and their derivatives with dimensions ranging from 0·1 × 0·1 × 0·35 mm to 0·4 × 0·4 × 0·7 mm were used in the crystal–capillary experiments. Trypsin crystals grow with acicular habit, cytochrome crystals are rectangular prisms, and trypsinogen crystals grow as approximately equidimensional trigonal bipyramids. Thus crystals of the most commonly encountered morphological forms have been used.

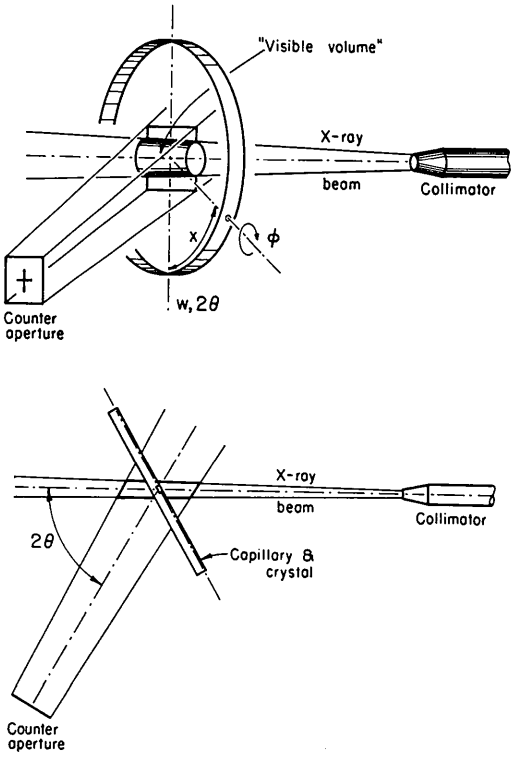

Fig. 1. The usual diffractometer data-collection arrangement. The 'visible volume' is the volume of air irradiated by the direct beam that is seen by the receiving-counter aperture. This volume changes only with 2θ, and governs the 2θ dependence of the air-scattered part of the background. Portions of the 'visible volume' are at times hidden from the X-ray beam and at times from the counter by the interposition of the capillary and crystal. The overall background will be affected by changes in the capillary–crystal setting angles, as well as the amount of capillary glass in the direct beam.

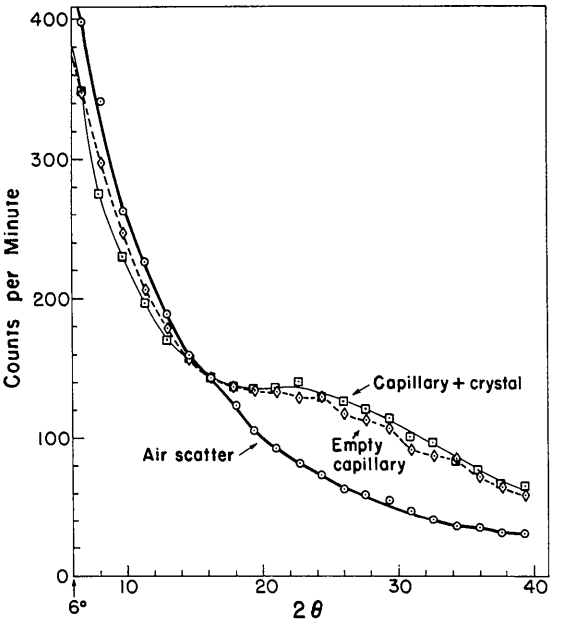

Fig. 2. The background 2θ dependence for: (*a*) air scatter only (○, 5 min/point), (*b*) an empty capillary (◇, 3 min/point), and (*c*) for a trypsinogen crystal, mounted in the same 0·5 mm diameter capillary (□, 3 min/point). All measurements were made at $\chi = 90°$, using diffractometer *A* as described in the text.

Backgrounds were measured under the following conditions:

(1) without crystal or capillary in the beam;

(2) with a well centered capillary (mounted coaxially with φ) in the beam;

(3) with a capillary miscentered, but mounted parallel to the φ axis; and

(4) with a crystal in a capillary such that the crystal, but not the capillary, is centered in the beam (the standard data-collection condition).

Background counting times were usually 3 to 5 min per point for the empirical background curves, and 10 to 120 s for individual backgrounds used for comparison. Reflection intensities were measured in both the ω scan and the Wyckoff step-scan (Wyckoff *et al.*, 1967) modes. Measurements were made over the ranges $0 \le 2\theta \le 46°$; $0 \le \varphi \le 360°$; and $0 \le \chi \le 360°$. Attenuators were used to obtain data for small 2θ angles down to 0°.

Results

Fig. 1 shows the usual diffractometer data collection arrangement and illustrates the origins of the different variations in the background intensity.

Primary effects: 2θ *dependence*

Fig. 2 shows the overall background as a function of 2θ: (*a*) with no crystal, no capillary; (*b*) for a well centered capillary; and (*c*) for a capillary containing a trypsinogen crystal (dimensions approximately 0·3 ×

0·25×0·35 mm). The intrinsic absorption of a glass capillary was measured and found to be generally less than 25%. The absorptive component due to the glass which affects the curves of Fig. 2 must consequently be less than about 25%. It is much less where 'visible volume' dimensions (see Fig. 1) are greater than the diameter of the capillary – which is often the case. It is therefore clear that scattering by air is a major contributing factor to the background. In the presence of a capillary the background is reduced by absorption by the glass capillary. The background intensity increases at higher 2θ angles where scatter from the glass predominates over its absorption. Variations in the orientation of the capillary lead to φ- and χ-dependent secondary perturbations in the background.

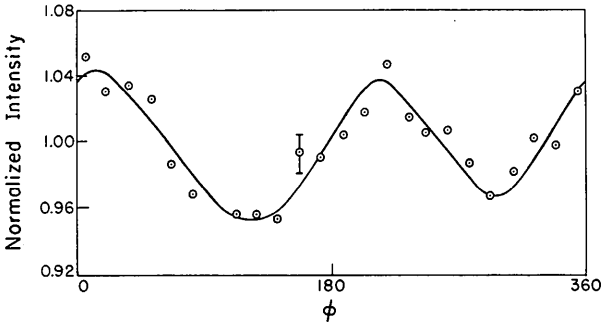

Fig. 3. The background φ dependence observed at $2\theta=22{\cdot}6°$ and $\chi=90°$ for a DIP-trypsin crystal. The average intensity is normalized to 1·0 (Diffractometer *A*).

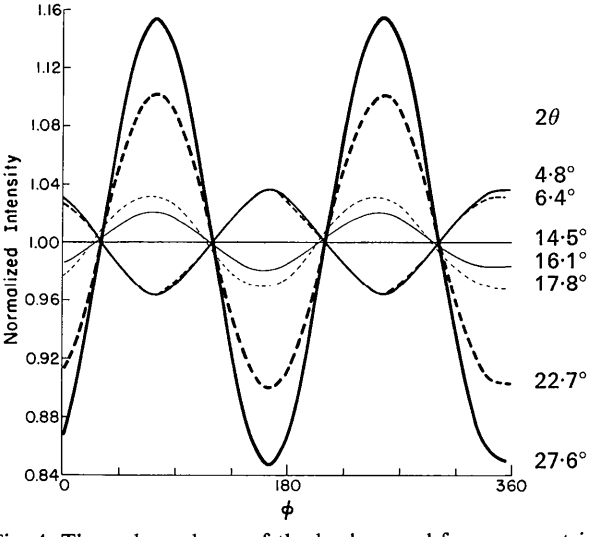

Fig. 4. The φ dependence of the background for an eccentrically mounted empty capillary at various 2θ values ($\chi=90°$ in all cases). While the amplitude and sign of the variations change with 2θ, the phase is constant. The curves were measured at $2\theta=27{\cdot}6°$, 22·7°, 17·8°, 16·1°, 14·5°, 6·4°, and 4·8°. The traces, except for $2\theta=14{\cdot}5°$, represent the best least-squares fit of 24 points measured at 15° intervals to the function: Normalized intensity $=1{\cdot}0+a\cos 2(\varphi-\varphi_0)$. The average intensity at each 2θ is normalized to 1·0. The curve for $2\theta=14{\cdot}5°$ is the best straight line through the data at this 2θ value. Measurements were made on diffractometer *A*.

Secondary effects: φ *dependence*

When it is necessary to collect data with a crystal mounted in a capillary whose *diameter* is greater than the radius of the X-ray beam, there can be a significant systematic variation of the background radiation with the diffractometer φ setting. Because there is little difference between the angular distribution of the background of an empty capillary and that of a capillary containing a crystal and its mother liquor, the φ dependence must be almost wholly due to the eccentricity of the capillary about the φ axis.

Fig. 3 shows typical φ dependence of background intensity for a capillary containing a protein crystal, and Fig. 4 shows this dependence for an eccentrically mounted empty capillary at different 2θ values. A well centered capillary shows no φ dependence, while with eccentrically mounted capillaries the background variation is an approximately sinusoidal function of 2φ as the capillary rotates in and out of the beam. At low 2θ values the background, primarily due to air scatter, is modulated by absorption of the eccentric capillary; therefore, the background is greatest when the capillary blocks the direct beam least ($\varphi=0$ and 180° in Fig. 5). At higher 2θ angles the capillary scatters more radiation than it absorbs, and backgrounds are highest at $\varphi=90$ and 270° in Fig. 5. At intermediate 2θ values, near the crossing point of Fig. 2, the absorption and scattering of the capillary are nearly equal and the backgrounds are essentially φ-independent. The extent of the background variation with φ will depend on the eccentricity and diameter of the capillary and the value of 2θ. The curves in Fig. 6 depict the background as a function of 2θ measured at two values of φ; one when the low-angle background is a minimum [curve $A_{\varphi_0,\chi_{90}}(2\theta)$] and the other when the low-angle background is a maximum [curve $B_{\varphi_1,\chi_{90}}(2\theta)$]. The magnitude of the differences between the two curves is the amplitude of the φ-dependent variation. There may be circumstances under which these curves would not cross, although we have not observed this. In such a case the amplitude of the φ-dependent variation with 2θ would not change sign.

For a DIP-trypsin crystal in a 0·6 mm capillary and a 1 mm beam, the background at 23° in 2θ varied 11% in intensity with rotation about φ. The average reflection intensity in the shell $22<2\theta<26°$ for this crystal was seven times the background level. For a reflection with this 'average' intensity, using interpolated backgrounds uncorrected for the φ dependence would produce a systematic error up to 1·6% in its net intensity. Weaker reflections, which are still important for heavy-atom refinement and difference maps, are subject to substantially larger errors: a reflection with raw intensity twice background, *although at least* 10 σ (*standard deviations*) *above background*, would be in error by as much as 11%. These errors are systematic and must be corrected for, if accurate phases are to be calculated.

The following scheme was tested for generating interpolated backgrounds corrected for φ dependence:

(1) Measure the φ dependence at a low 2θ value ($\sim 6°$) with a φ scan at $\chi = 90°$.

(2) Measure the 2θ dependence, $A_{\varphi_0,\chi_{90}}(2\theta)$, at the φ value (φ_0) for which the background is a minimum on the φ scan.

(3) Measure the 2θ dependence, $B_{\varphi_1,\chi_{90}}(2\theta)$, at the φ value ($\varphi_1 = \varphi_0 + 90°$) for which the background is a maximum on the φ scan.

(4) Tabulate the difference, $D_\varphi(2\theta) = B_{\varphi_1,\chi_{90}}(2\theta) - A_{\varphi_0,\chi_{90}}(2\theta)$. [Normally, it is convenient to interpolate the values between the observed backgrounds on the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $B_{\varphi_1,\chi_{90}}(2\theta)$ curves at intervals in $\sin(\theta)/\lambda$.]

(5) Calculate the backgrounds, $\mathrm{BG}(2\theta,\varphi)$. For background corrections (just as for absorption corrections) the φ setting angle, φ_s must be modified to account for geometrical contributions to rotation with respect to the incident beam from χ and ω.*

$$\mathrm{BG}(2\theta,\varphi_s) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\varphi(2\theta)\sin^2(\varphi_c - \varphi_0) \quad \text{(I)}$$

where

$$\varphi_c = \varphi_s - \tan^{-1}(\cos\chi\tan\omega)\,. \quad \text{(II)}$$

For 1500 reflections observed over a wide range of 2θ, φ, and χ for a DIP-trypsin crystal ($0{\cdot}25\times 0{\cdot}3\times 0{\cdot}4$ mm) mounted in a 1 mm capillary on diffractometer *A*, the agreement between observed and interpolated backgrounds typically improved by 40% when φ-dependent anisotropic, rather than isotropic, interpolation was used. This scheme assumes that the φ dependence is sinusoidal† and that the amplitude of the variation is the difference $[D_\varphi(2\theta)]$ between the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $B_{\varphi_1,\chi_{90}}(2\theta)$ curves of Fig. 6. If possible, crystals should be mounted only in capillaries whose diameters are smaller than the radius of the X-ray beam. Under these conditions there is no observable φ dependence.

* Any additional correction for a 2θ component in φ_c (due to rotation with respect to the counter) is much less significant in general where the visible volume (Fig. 1) is larger than the irradiated portion of the capillary. The correction for χ and ω deals with the source of scattering directly, while the 2θ correction deals with only a small component of the scattered radiation. In any case an additional 2θ correction would be asymmetric and complex and is usually unneccessary.

† We have always observed a certain asymmetry in the φ scans which we presume to be due to the difference in position of the eccentrically mounted capillary with respect to the counter at φ and $\varphi + 180°$. In order to account for this asymmetry, we have calculated backgrounds using the expression:

$$\mathrm{BG}(2\theta,\varphi) = A_{\varphi_0,\chi_{90}}(2\theta) + D_\varphi(2\theta)\left[\frac{F(\varphi_c) - F(\varphi_0)}{F(\varphi_1) - F(\varphi_0)}\right]$$

where, for example,

$F(\varphi) = a(x-x_0)^2 + b(y-y_0)^2 + c(x-x_0)(y-y_0)$; $x = \sin\varphi$, $y = \cos\varphi$, and x_0, y_0, a, b, and c are refinable parameters. Alternatively, one could use an interpolated φ curve for $F(\varphi)$. However, we have obtained the best results using the sine function in equation (I).

χ dependence

The background χ dependence is analogous to the φ dependence, but a rigorous definition of the problem is more difficult as a result of more complex geometry. However, the χ dependence may be understood qualitatively. As χ varies from 0 to 90°, the amount of glass in the equatorial plane increases approximately as the secant of χ. This, in turn, increases the absorption of air-scattered radiation. The exact nature of this angular background variation with χ depends strongly on the experimental conditions.‡

‡ We have observed that each diffractometer has a characteristic χ–2θ dependence, as shown in Figs. 7 and 8. Subtle differences in the geometry of the diffractometers (*e. g.*, collimator-to-crystal distance) are apparently responsible for the difference in these patterns. The correction scheme outlined below has proven effective for data collected on different diffractometers with different χ–2θ patterns.

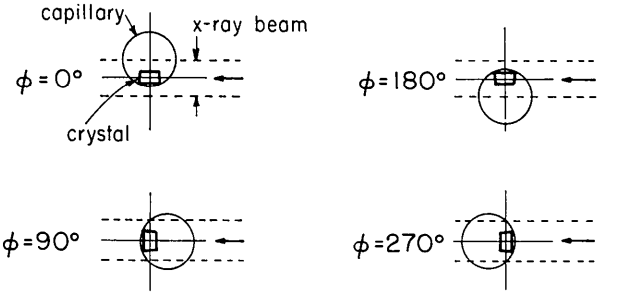

Fig. 5. View parallel to the φ axis of a capillary and crystal, showing the amount of direct-beam radiation they intercept at various φ settings. The capillary diameter is greater than the radius of the X-ray beam.

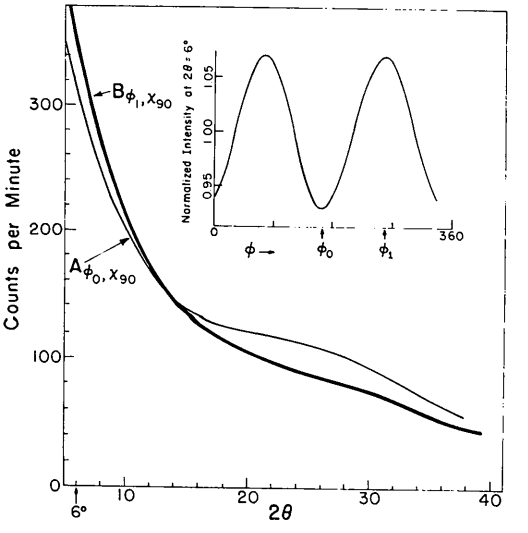

Fig. 6. Curves $A_{\varphi_0,\chi_{90}}$ and $B_{\varphi_1,\chi_{90}}$ the background 2θ dependence observed at two φ settings, 90° apart, for a capillary-mounted DIP-trypsin crystal. Inset, a φ scan for this crystal taken at $2\theta = 6°$, showing the positions of the φ_0 and φ_1 settings (Diffractometer *A*).

Figs. 7 and 8 show the 2θ dependence of the background radiation for DIP-trypsin and cytochrome c crystals at $\chi=0°$ [$C_{\varphi_0,\chi_0}(2\theta)$] and $\chi=90°$ [$A_{\varphi_0,\chi_{90}}(2\theta)$], and the variation of the χ dependence for several values of 2θ. Just as with the φ dependence, the amplitude of the χ variation depends upon the difference between the backgrounds at the extremes, $\chi=0$ and $\chi=90°$. Depending upon the experimental conditions, it is possible to observe a crossover of the $A_{\varphi_0,\chi_{90}}(2\theta)$ and $C_{\varphi_0,\chi_0}(2\theta)$ curves analogous to that for the φ curves. In such cases, the high-angle background increases as χ goes from 0 to 90°, and the shape of the curve can vary from that observed at low 2θ (Fig. 8).

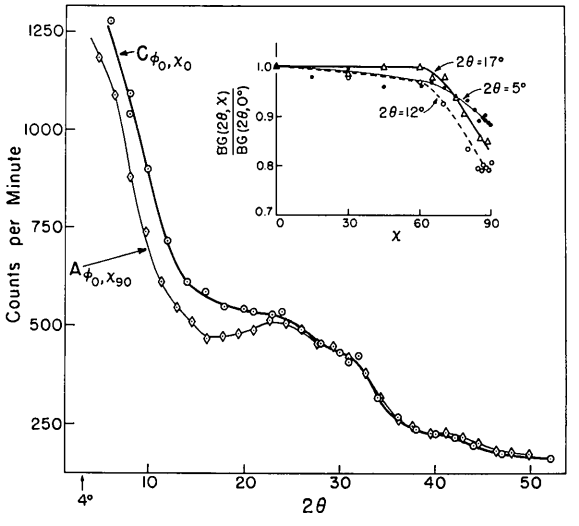

Fig. 7. Curves $A_{\varphi_0,\chi_{90}}$ and C_{φ_0,χ_0}: the background 2θ dependence observed at $\chi=90$ and $\chi=0°$ on diffractometer A for a DIP-trypsin crystal mounted in a 1 mm capillary. Each data point was counted for 4 min. Inset: the χ dependence observed at several 2θ values. Each point was measured for 5 min.

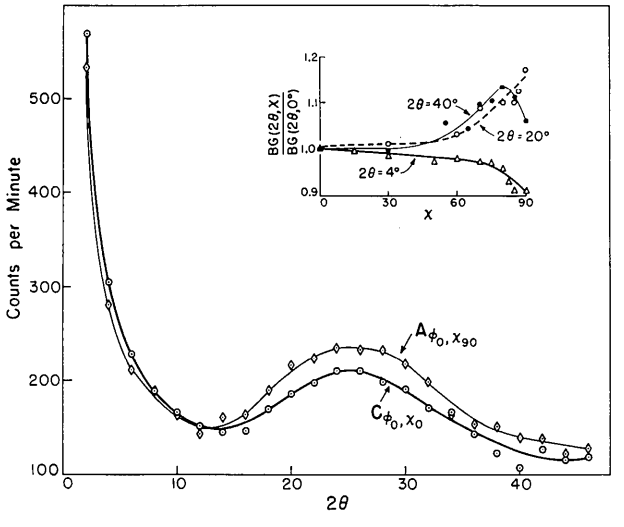

Fig. 8. Curves $A_{\varphi_0,\chi_{90}}$ and C_{φ_0,χ_0}: the background 2θ dependence observed at $\chi=90$ and $\chi=0°$ for a cytochrome c crystal using diffractometer B. Inset: background χ dependence at several 2θ values. All points were measured for 4 min.

The χ-dependent background variation can represent a significant fraction of the total background. We have routinely observed 15% differences in backgrounds between $\chi=0$ and $\chi=90°$, and differences as high as 25% are not uncommon. The following interpolation scheme, analogous to that used for φ, was used to correct for the χ dependence:

(1) Measure the 2θ dependence, $C_{\varphi_0,\chi_0}(2\theta)$, at $\chi=0°$.

(2) Measure the 2θ dependence, $A_{\varphi_0,\chi_{90}}(2\theta)$, at $\chi=90°$.

(3) Tabulate the differences, $D_\chi(2\theta)=C_{\varphi_0,\chi_0}(2\theta)-A_{\varphi_0,\chi_{90}}(2\theta)$.

(4) Measure the χ dependence at a $2\theta=2\theta_m$ for which the difference, $D_\chi(2\theta)$, is large. If possible, $2\theta_m$ should be representative of the 2θ's in the data set.

(5) Calculate the backgrounds from:

$$\mathrm{BG}(2\theta,\chi)=A_{\varphi_0,\chi_{90}}(2\theta)+D_\chi(2\theta)\left[\frac{F(\chi)-F(90°)}{F(0°)-F(90°)}\right]_{2\theta_m} \quad \text{(III)}$$

where $F(\chi)$ is either an empirical function representing the χ dependence or an interpolated χ curve at $2\theta_m$. Our best results have been obtained with:

$$F(\chi)=a\exp\{-b/(\cos\chi+c)\} \quad \text{(IV)}$$

where the coefficients a,b,c are determined by a least-squares fit. If the background shown in Fig. 7, for example, is not corrected for χ dependence, *i.e.*, if isotropic interpolated backgrounds (depending only on 2θ) are used, systematic errors of up to 30% of the background can result. χ-dependent (anisotropic) interpolation can be improved upon by subdividing the data into smaller 2θ ranges and applying appropriate χ curves to the data in each range.

When both φ and χ dependences are present, backgrounds may be approximated by combining the φ and χ corrections:

$$\mathrm{BG}(2\theta,\varphi,\chi)=A_{\varphi_0,\chi_{90}}(2\theta)+D_\varphi(2\theta)\sin^2(\varphi_c-\varphi_0)+D_\chi(2\theta)\left[\frac{F(\chi)-F(90°)}{F(0°)-F(90°)}\right]_{2\theta_m}. \quad \text{(V)}$$

Non-linear least squares

A different approach to the estimation of backgrounds is the use of non-linear least squares to approximate backgrounds as some function of 2θ, φ, and χ. This function is determined by fitting backgrounds measured for short times in the vicinity of each reflection. The reliability of such a procedure depends on the choice of a well behaved function with a relatively small number of parameters. Its accuracy also depends on the number of data and time spent on measuring the individual backgrounds.

The dependence of background intensity on 2θ may be satisfactorily accounted for by a third-order function of 2θ:

$$\mathrm{BG}(2\theta)=a+bT+cT^2+dT^3, \qquad \text{(VI)}$$

where $T=2\theta$ and the lower case letters a,b,c, *etc.*, are parameters to be determined. Some improvement can be achieved when the data are separated into low, middle, and high-angle regions in 2θ. Other terms are added to BG(2θ) to account for the variation of BG (2θ) as functions of φ and χ.

In any region of reciprocal space (not necessarily over all 2θ ranges), it is sufficient to approximate the changing magnitude and sign of the φ correction with a third-order polynomial:

$$\Delta B_1(\varphi,2\theta)=\sin^2(\varphi-\varphi_0)\,(e+fT+gT^2+hT^3) \qquad \text{(VII)}$$

where φ_0 is a parameter corresponding to the φ value for minimum or maximum background. If the background does not have a φ dependence, $\Delta B_1(\varphi,2\theta)$ is zero.

Since the magnitude, curvature, and sometimes the sign of the χ correction are also 2θ-dependent, we have tried various combinations of terms of the form $T^i\chi^j$; $i=1$ to 4; $j=1$ to 4 to determine which terms are useful to best correct for the χ dependence. An expression of the form

$$\Delta B_2(\chi,2\theta)=iT\chi+jT\chi^2+kT\chi^3+lT^2\chi^2+mT^3\chi \qquad \text{(VIII)}$$

provides the best compromise between the number of refined parameters and the overall quality of fit.

The overall expression for least-squares refinement is:

$$\mathrm{BG}(2\theta,\varphi,\chi)=\mathrm{BG}(2\theta)+\Delta B_1(\varphi,2\theta)+\Delta B_2(\chi,2\theta)\,. \qquad \text{(IX)}$$

This function can be used in any 2θ range. However, BG(2θ) and $\Delta B_2(\chi,2\theta)$ can often be simplified for middle and high 2θ ranges with essentially no loss of accuracy: for middle or high 2θ, $\mathrm{BG}(2\theta)=a+bT$ suffices; for the middle range, $\Delta B_2(\chi,2\theta)=iT\chi+jT\chi^2+kT\chi^3+lT^2\chi^2$; and for the high range $\Delta B_2(\chi,2\theta)=iT\chi+jT\chi^2+kT\chi^3$.

Discussion

The anisotropic interpolation (AI) and non-linear least-squares (LS) techniques have been used routinely to compare and measure background intensities for crystals of several proteins on different diffractometers. Data for the comparison to be discussed first were taken from a DIP-trypsin crystal using diffractometer *A*. The data shown in Table 2 and Fig. 9 compare the isotropic interpolation (II), AI and LS methods of estimating background intensity with individual backgrounds, BG_{obs}, each measured for 40 s.

Fig. 9 shows the distribution of errors, $\Delta\mathrm{BG}/\sigma_{obs}$, where

$$\Delta\mathrm{BG}=(\mathrm{BG}_{obs}-\mathrm{BG}_{calc}) \text{ and } \sigma_{obs}=\sqrt{\mathrm{BG}_{obs}}\,,$$

for each of the three methods and for a perfectly normal distribution. For the II method the distribution of error is far from normal because of the systematic error introduced by ignoring the anisotropy in the background. The error distribution is skewed and distorted regardless of the setting angles used for observing the 2θ curve (see Fig. 10).

The error distributions for the AI and LS methods, on the other hand, are very nearly Gaussian,* indi-

Table 2. *Error analysis of background-approximation techniques*

Estimated backgrounds compared to backgrounds measured for 40 s per point.*

Method	$\langle\Delta\mathrm{BG}^2\rangle^{1/2}$†	s‡	R_{BG}%§	GOF‖	R_F%**
II	61·7	55·3	10·0	2·22	1·76
AI	41·2	30·8	6·0	1·38	0·88
LS	32·3	17·1	5·3	1·19	0·87

* Results are for 1956 reflections measured on diffractometer (*A*) as described in the text. The average observed background was 501 counts per min, and the rms σ (BG_{obs}) was 27·4 counts per min. The comparisons are all tabulated in counts per min.

† $\Delta\mathrm{BG}=\mathrm{BG}_{obs}-\mathrm{BG}_{calc}$.

‡ s is the error component due to errors arising from the II, AI or LS methods, $s^2=\langle\Delta\mathrm{BG}^2\rangle-\langle\sigma^2_{obs}\rangle$.

§ $R_{BG}=\sum|\Delta\mathrm{BG}|/\sum\mathrm{BG}_{obs}$.

‖ $\mathrm{GOF}=\left[\dfrac{\sum(\Delta\mathrm{BG}/\sigma_{obs})^2}{1956\text{ pts.}}\right]^{1/2}$, the 'goodness-of-fit'.

** $R_F=\sum|F_{BGobs}-F_{BGcalc}|/\sum|F_{BGobs}|$.

* We have shown this in two ways: least-squares fitting a Gaussian to the error frequency curves in Fig. 9, and by direct comparison to perfectly normal distributions by means of normal probability plots (Abrahams & Keve, 1968).

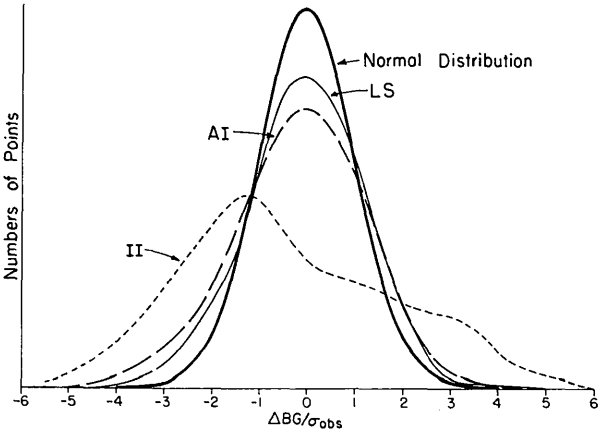

Fig. 9. The distribution of errors between individually observed (40 s) backgrounds and backgrounds calculated using the three techniques described in the text. The statistics plotted are the numbers of points within intervals of $\Delta\mathrm{BG}/\sigma_{obs}$, where $\Delta\mathrm{BG}=\mathrm{BG}_{obs}-\mathrm{BG}_{calc}$ and $\sigma_{obs}=\sqrt{\mathrm{BG}_{obs}}$. The total area under each curve is equivalent to 1956 data points. The curves are frequency distributions for the errors of the II, or isotropic interpolated backgrounds, the AI, or anisotropic interpolated backgrounds, and the LS, or least-squares backgrounds. The heavy curve represents an ideal Gaussian distribution of error for the same number of points. (DIP-trypsin crystal on diffractometer *A*).

cating that the systematic errors resulting from the II method have been largely eliminated. The slightly increased breadth of the AI and LS error distributions relative to the perfectly normal distribution is a consequence of the errors introduced by these methods. We can estimate the extent of these errors, s, using the approximation:

$$s^2 \simeq \langle \Delta BG^2 \rangle - \langle \sigma^2_{obs} \rangle \qquad (X)$$

where $\langle \Delta BG^2 \rangle$ is the mean squared ΔBG and $\langle \sigma^2_{obs} \rangle$ is the mean squared error in the individually observed backgrounds. For the data illustrated in Fig. 9 the estimated average error for backgrounds calculated using the AI technique, with empirical background curves observed for 4 min per point, is roughly equivalent to that which would be obtained by measuring each background for 32 s. For LS, s is equivalent to error expected for backgrounds measured for 91 s each. The value of s for AI depends upon the time spent measuring each point for the interpolated curves, while that for LS depends on the counting times for individually measured backgrounds. In this example, individual 40 s backgrounds were used to calculate the LS function. In another case results were essentially identical when 10 s backgrounds were used, and this method is clearly most useful when still shorter times are used.

We have compared the effects of the II, AI and LS methods on the structure factors by calculating the R values (R_F) between the data sets obtained using observed and calculated backgrounds (see Table 1). The R_F value for the II corrected data is approximately twice as large as that calculated for the AI and LS corrected data. Applying absorption corrections to the scan intensities *before* subtracting II backgrounds, as recommended by Hill & Banaszak, increases the R_F value to 1·95%. Clearly the AI and LS methods are substantial improvements over the II method.

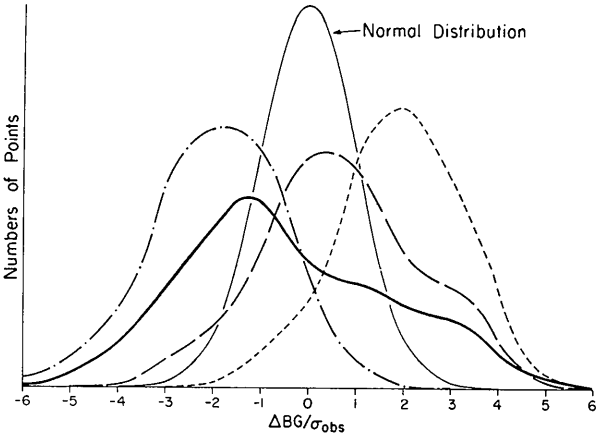

Fig. 10. The distribution of errors when different empirical 2θ curves are used in the II method. The frequency distributions are for backgrounds calculated using: a 2θ curve measured at $\varphi=\varphi_0$, $\chi=0°$ (— · —), a 2θ curve measured at $\varphi=\varphi_1$, $\chi=90°$ (----), a 2θ curve measured at $\varphi=\varphi_0$, $\chi=90°$ (——); and a 2θ curve calculated from the φ_1,χ_{90} and φ_0,χ_{90} curves, representing a 2θ curve which would be obtained if $\varphi \sim \frac{1}{2}(\varphi_0+\varphi_1)$ (— —). The light curve (——) represents an ideal Gaussian distribution of error. As in Fig. 9, the area under each curve represents 1956 data points distributed over the range $4·6 < 2\theta < 46°$.

Conclusion

The most significant conclusions to be drawn from these studies of background intensity are:

(1) The principal 2θ-dependent component of background intensity arises from X-rays scattered by the volume of air illuminated by the direct beam and 'seen' by the receiving-counter aperture. Modulation of this air scatter by capillary absorption and scatter gives rise to φ and χ-dependent variations in the background.

(2) The background intensity is essentially uncorrelated with the crystal absorption. Under certain experimental conditions the two phenomena may appear to be correlated; for example, where a flat crystal rests on the side of the capillary such that its long axis lies parallel to the φ axis of the diffractometer.

(3) As a consequence of (2), the background intensity should always be subtracted from the scan intensity before the absorption corrections are applied, contrary to the suggestion of Hill & Banaszak (1973).

(4) The anisotropy of background scatter as a function of 2θ, φ, and χ can be satisfactorily accounted for by using a simple interpolative procedure, or by fitting a suitable function to many backgrounds measured for short times near each reflection.

(5) The use of an isotropic background interpolation is usually unsatisfactory for accurate data reduction. This method can introduce systematic errors of up to 30% in the estimated background intensities. Errors of this sort lead to systematic errors in phases and increase the noise levels in difference maps.

Both the AI and LS methods substantially reduce the systematic errors inherent in the II method. As a result, the differences between observed backgrounds and estimated backgrounds closely follow a normal distribution. Both methods represent significant savings in time and are improvements over the II method. The overall accuracy of either method is limited by the extent and duration of the background sampling, and is generally commensurate with that obtained for individual backgrounds. When deciding between the two methods, one must choose between the greater savings in time provided by the AI method, or the somewhat better accuracy of the LS technique at the cost of increased crystal exposure and decay.

Our results identify the principal sources of background intensity. Consequently, there are several experimental steps which can be taken to reduce background intensity: (1) Capillaries should ideally be chosen to be smaller in diameter than the radius of the incident X-ray beam. This renders φ dependence insignificant, diminishes χ dependence of the background and reduces the high-angle background due to capillary

scatter. (2) Helium-filled tubes are often used to reduce absorption by air in the input or output beam directions. However, as the main problem is generally one of peak-to-background ratio, these measures only slightly affect this ratio, and in any case do not reduce background intensity significantly. It would seem highly advantageous to fill the 'visible volume' with helium, so affecting the background intensity directly. This is obviously difficult to do as it implies either a helium-filled enclosure over the entire diffractometer, or a helium-filled chamber mounted around the φ axis which surrounds the capillary and crystal completely. There are obvious mechanical difficulties in building such a device. First, it must be almost X-ray transparent over the angular ranges used. Second, it must be moderately well sealed to minimize leakage of helium if it encloses the φ drive shaft bearing. A mylar cylinder with solid supports mounted onto the top of the goniometer head would seem to be a good compromise allowing for rigid support of the top and bottom in the X-ray shadow.

(3) Any means of restricting the visible volume will reduce the background in almost direct proportion to the volume change. This can be achieved by placing the final restricting aperture and the scatter cap on the input collimator as close to the crystal as possible. Similarly, there should be a defining aperture as close to the crystal as possible in the crystal–counter pathway, and a second one close to the counter.

We thank Drs A. Kossiakoff and R. Swanson for generously supplying background data, Mr R. Almassy for helpful assistance, and Dr R. E. Marsh and Mr J. Greif for valuable discussions. We are grateful to Dr Sten Samson for allowing us to collect data on diffractometer *F*, a machine designed in the most part by himself, and for his supervision in the redesign of diffractometer *B*. We also recognize his constant attention in the redesign and improvements made to diffractometer *C*.

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* A**27**, 157–165.
DICKERSON, R. E., KOPKA, M. L., VARNUM, J. & WEINZIERL, J. E. (1967). *Acta Cryst.* **23**, 511–522.
HENDERSON, R. & MOFFATT, J. K. (1971). *Acta Cryst.* B**27**, 1414–1420.
HILL, E. J. & BANASZAK, L. J. (1973). *Acta Cryst.* B**29**, 372.
JENSEN, L. (1972). Personal communication.
KRIEGER, M., KAY, L. M. & STROUD, R. M. (1974). *J. Mol. Biol.* **83**, 209–230.
MATHEWS, F. S., LEVINE, M. & ARGOS, P. (1972). *J. Mol. Biol.* **64**, 449–464.
SALEMME, F. R., FREER, S. T., XUONG, Ng. H., ALDEN, R. & KRAUT, J. (1973). *J. Biol. Chem.* **248**, 3910–3921.
WATSON, H. C., SHOTTON, D. M., COX, J. M. & MUIRHEAD, H. (1970). *Nature, Lond.* **225**, 806–811.
WYCKOFF, H. W., DOSCHER, M., TSERNOGLOU, D., INAGAMI, T., JOHNSON, L. N., HARDMAN, K. D., ALLEWELL, N. M., KELLY, D. M. & RICHARDS, F. M. (1967). *J. Mol. Biol.* **27**, 563–578.
WYCKOFF, H. W., TSERNOGLOU, D., HANSON, A. W., KNOX, J. R., LEE, B. & RICHARDS, F. M. (1970). *J. Biol. Chem.* **245**, 305–328.

Acta Cryst. (1974). A**30**, 748

Highly Anisotropic Extinction

BY F. R. THORNLEY AND R. J. NELMES

Department of Physics, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, Scotland

(*Received* 7 *May* 1974; *accepted* 20 *May* 1974)

An investigation has been made of some large anisotropies observed in the X-ray scattering from a single crystal of Cr–Cl boracite, $Cr_3B_7O_{13}Cl$, in its room-temperature cubic phase. The model of Coppens & Hamilton [*Acta Cryst.* (1970), A**26**, 71–83] for anisotropic secondary extinction has been used to describe the results. Both the type I (domain misorientation) and the type II (domain shape) extinction of that model were found to be present. A new expression for the orientation dependence of type I extinction is introduced, which is believed to be more appropriate to the normal experimental situation. With this modification, the model was able to reproduce fairly well the observed changes in integrated intensity on rotation about the scattering vector. The components of the tensors describing the two types of extinction indicated much greater angular misorientation about the growth axis of the crystal than perpendicular to it, and a domain semi-axis smaller along the growth axis than perpendicular to it.

Introduction

During the collection of X-ray diffraction data for the refinement of the crystal structure of Cr–Cl boracite, $Cr_3B_7O_{13}Cl$, in its room-temperature cubic phase (Nelmes & Thornley, 1974), large differences between the integrated intensities of symmetry-related reflexions were observed. These differences were much too great