WLODAWER, A., BORKAKOTI, N., MOSS, D. S. & HOWLIN, B. (1986). *Acta Cryst.* B**42**, 379-387.
WLODAWER, A., SVENSSON, L. A., SJÖLIN, L. & GILLILAND, G. L. (1988). *Biochemistry*, **27**, 2705-2717.
YU, H., KARPLUS, M. & HENDRICKSON, W. A. (1985). *Acta Cryst.* B**41**, 191-201.
YU, N.-T., LIU, C. S. & O'SHEA, D. C. (1972). *J. Mol. Biol.* **70**, 117-132.

Acta Cryst. (1989). A**45**, 861-870

On the Calculation of Small-Angle Diffraction Patterns from Distorted Lattices

BY H. DE GRAAF

Department of Physiology, University of Amsterdam, Meibergdreef 15, 1105 *AZ Amsterdam, The Netherlands*

(*Received* 26 *August* 1988; *accepted* 8 *August* 1989)

Abstract

A method is presented for the calculation of the diffraction properties of distorted lattices. An isotropic correlation field is used and the influence of finite crystal size is taken into account. An approximate model is described which is based on a fit of the correlation function to a sum of Gaussians. This method, which is applicable to more complex structures than simple lattices, simplifies the calculation of complicated convolution integrals. The method is illustrated with the calculation of the diffraction properties of 2D hexagonal lattices and exponential correlation. Various aspects, such as variations in peak intensity, line width and line shifts, are discussed.

Introduction

Many models have been developed which deal with the statistical properties of disordered lattices. However, calculation of the diffraction properties of such lattices remains a difficult task (Welberry, 1985). One type of disorder which is particularly important in macromolecular crystals is deformation of the regular structure. The earliest concept which has been widely used to describe the diffraction properties of such lattices is that of the paracrystal, developed by Hosemann and co-workers (Hosemann & Bagchi, 1962). In the paracrystal model one considers the distance between neighbours as the basic stochastic variable. Fluctuations in the length and orientation of unit-cell vectors lead to displacement disorder. The advantage of this model, in its simplest form, is that expressions for diffraction properties can be obtained relatively easily. The main disadvantage is that, owing to the random-walk nature of the disturbances, the variance of the length of vectors between successively distant neighbours increases without bound. This implies that the crystal dimensions depend on the statistical properties. In later developments corrections have been made for this behaviour (Hosemann, 1975) but these do not allow a simple calculation of diffraction properties.

An alternative view regards the lattice points as the basic stochastic variables and treats the distorted lattice by allowing the lattice points to deviate from a perfectly regular arrangement. If one chooses a large single-site standard deviation as well as a sufficiently large neighbour correlation the ideal paracrystal and the distorted regular lattice can be shown to be equivalent in one dimension (Welberry, Miller & Carroll, 1980). Generalizing this approach to higher dimensions, Welberry and co-workers developed so-called Gaussian 'growth disorder models' (Welberry, Miller & Carroll, 1980; Welberry & Carroll, 1982). These models do not have the drawback of the ideal paracrystal and lead to manageable formulae for diffraction patterns. Also, examples of lattices, which can be used in optical analogue diffraction experiments, can be easily constructed. The starting point in these models is that the perturbations of the lattice points can be represented by a Gaussian probability distribution, with the additional assumption that, on a square lattice, the diagonal correlation coefficient can be expressed as a product of the primary axial coefficients. This assumption introduces an unavoidable anisotropy in the correlation field and the corresponding diffraction pattern. This feature is also present in the ideal paracrystal model.

Dropping this assumption leads to a general Gaussian model (Welberry & Carroll, 1983) which is much more realistic. However, samples of lattices can only be generated by Monte Carlo methods and expressions for derived diffraction patterns are complicated and hardly amenable to calculation, so that one still has to use optical diffraction analogues to study practical situations.

A third type of approach has been proposed by Thakur *et al.*, who explicitly introduced the interaction potential, felt by the scattering units, into the line-shape expression (Thakur, Tripathy & Lando, 1985). Anharmonic terms in this potential are responsible for line-broadening effects. Although this idea

0108-7673/89/120861-10$03.00

is attractive since temperature dependence of the diffraction line shape can be taken into account in a natural way, the method as applied to several polymer crystals is still essentially a 1D treatment (Thakur, Tripathy, Burkhart & Lando, 1985). A general treatment for higher dimensions would be much more involved.

In the following sections we shall outline a diffraction model which assumes a general Gaussian distribution for the lattice displacements and an isotropic interaction between lattice sites. The method results in relatively simple expressions for the diffracted intensity and can be used with relatively complex disorder models. This work has grown out of the study of equatorial small-angle X-ray diffraction by skeletal muscle fibres. These systems consist of hexagonally ordered protein structures, in 2D, with a variable degree of displacement disorder, depending on the physiological circumstances [see, for instance, Yu, Steven, Naylor, Gamble & Podolsky (1985)]. Therefore we shall use 2D hexagonal lattices to illustrate the method, although the theory is easily extended to higher dimensions and other lattices.

The model

Two types of hexagonal lattices are treated in this section. A simple hexagonal case is used to define the basic variables and concepts, including an approximate model. Subsequently, this model is applied to a more general case, representing skeletal muscle tissue, with more than one scattering object in the unit cell and different scattering factors.

1. *A simple hexagonal lattice*

For simplicity, here we shall consider point objects so that scattering factors can be left out of consideration. The scattered intensity of an assembly of N objects is then given by

$$I(\mathbf{s}) = \left| \sum_j \exp(i\mathbf{r}_j \cdot \mathbf{s}) \right|^2. \quad (1)$$

The object j is situated at position $\mathbf{r}_j$. The scattering vector $\mathbf{s}$ is defined as the difference of the wave vectors of the incident and scattered waves. The position vector $\mathbf{r}_j$ can be represented by a lattice translation vector $\mathbf{t}$ and an associated continuous Gaussian random vector with zero mean $\mathbf{d}_t$,

$$\mathbf{r}_j = \mathbf{t} + \mathbf{d}_t.$$

Interpreting (1) in terms of a lattice average one obtains the standard result

$$I(\mathbf{s}) = N \sum_t \exp(i\mathbf{s} \cdot \mathbf{t}) \langle \exp[i\mathbf{s} \cdot (\mathbf{d}_t - \mathbf{d}_0)] \rangle. \quad (2)$$

Since the perturbations $\mathbf{d}_t$ are Gaussian variables, the exponent also has a Gaussian distribution, which leads to

$$I(\mathbf{s}) = N \sum_t \exp(i\mathbf{s} \cdot \mathbf{t}) \exp\{-\tfrac{1}{2}\langle[\mathbf{s} \cdot (\mathbf{d}_t - \mathbf{d}_0)]^2\rangle\}. \quad (3)$$

The lattice average in the exponent is evaluated in terms of correlation coefficients. In a general Gaussian model the probability of a given arrangement of the scattering points is given by a multivariate Gaussian distribution, which involves the interaction between lattice sites. The nature of this interaction determines the form of these correlation coefficients in a complicated way. Isotropic correlation can be introduced by assuming an isotropic interaction between lattice sites of the Heisenberg type (Kittel, 1976). This interaction does not necessarily have to be restricted to nearest neighbours. Although not explicitly mentioned, such an interaction was used by Welberry & Carroll (1983). For an interaction of this type all cross-correlation between the x and y components of the distortion vectors vanishes and the correlation between x components is equal to the correlation between y components. Also, the single-site standard deviations in the x and y direction are equal. Thus one obtains

$$\langle[\mathbf{s} \cdot (\mathbf{d}_t - \mathbf{d}_0)]^2\rangle = 2s^2\sigma^2(1 - \rho_t). \quad (4)$$

The single-site standard deviation σ and the correlation coefficient ρ_t are given by

$$\sigma^2 = \langle d_{0x}^2 \rangle \text{ and } \rho_t = \langle d_{tx} d_{0x} \rangle / \sigma^2, \quad -1 < \rho_t < 1.$$

In the isotropic case the correlation coefficient ρ_t depends on the length of the lattice vector $\mathbf{t}$ and not on the direction. Substitution of (4) into (3) and expansion of the exponential with the lattice average yields

$$I(\mathbf{s}) = N \exp(-s^2\sigma^2) \times \sum_n [(s^2\sigma^2)^n / n!] \sum_t \exp(i\mathbf{s} \cdot \mathbf{t}) \rho_t^n. \quad (5)$$

At this point one can consider two extreme cases. In the first case the correlation between lattice sites is negligibly small, *i.e.* $\rho_t = 0$ for $\mathbf{t} \neq 0$. Then, for $\mathbf{t} \neq 0$, the term with $n = 0$ is the only non-vanishing term and one obtains for (5)

$$I(\mathbf{s}) = N\left[1 + \exp(-s^2\sigma^2) \sum_{t \neq 0} \exp(i\mathbf{s} \cdot \mathbf{t})\right], \quad (6)$$

which includes the familiar Bragg intensity corrected for uncorrelated disorder by the Debye–Waller factor and a diffuse background. In the second case (the opposite extreme, $\rho_t = 1$) the ρ_t^n become independent of n and one has a perfect lattice.

In the general case the higher-order terms, which give rise to additional diffuse scattering, cannot be neglected. In order to deal with these terms we replace the correlation coefficients ρ_t by a continuous correlation function $\rho(\mathbf{r})$, defined such that $\rho(\mathbf{t}) = \rho_t$ and we

replace the Fourier sum by an integral,

$$\sum_{\mathbf{t}} \exp(i\mathbf{s}\cdot\mathbf{t})\rho_{\mathbf{t}}^{n} = V^{-1}\int d\mathbf{r} \exp(i\mathbf{s}\cdot\mathbf{r})\Big\{\rho(\mathbf{r})^{n}\sum_{\mathbf{t}}\delta(\mathbf{r}-\mathbf{t}) \times [B(\mathbf{r}) * B(-\mathbf{r})]\Big\}, \quad (7)$$

where the autocorrelation of the shape function $B(\mathbf{r})$ is introduced in order to define the integration over all space (Cowley, 1984). The * stands for convolution and V is the crystal area. The summation over $\mathbf{t}$ on the right-hand side of (7) is taken over an infinite lattice. Applying the convolution theorem one obtains

$$\sum_{\mathbf{t}} \exp(i\mathbf{s}\cdot\mathbf{t})\rho_{\mathbf{t}}^{n} = \Omega^{-1}R_n(\mathbf{s}) * C(\mathbf{s}), \quad (8)$$

with

$$C(\mathbf{s}) = V^{-1}\sum_{\mathbf{g}}\delta(\mathbf{s}-\mathbf{g}) * |B(\mathbf{s})|^2, \quad (9)$$

$$R_n(\mathbf{s}) * = (4\pi^2)^{-n}[R(\mathbf{s}) *]^n, \quad (10)$$

$$R(\mathbf{s}) = \int d\mathbf{r} \exp(i\mathbf{s}\cdot\mathbf{r})\rho(\mathbf{r}), \quad (11)$$

$$B(\mathbf{s}) = \int d\mathbf{r} \exp(i\mathbf{s}\cdot\mathbf{r})B(\mathbf{r}). \quad (12)$$

In (8) Ω represents the unit-cell area. The summation over n in (5) leads to a rapid increase in the number of convolutions that have to be taken into account in (10). Moreover, each convolution is a two-centre integral, which is in general difficult to evaluate. It is clear that (5) is hardly accessible to analytical treatment for a general form of crystal shape and $\rho(\mathbf{r})$.

For a finite single crystal with sharply defined boundary, $B(\mathbf{s})$ is an oscillating function which determines the width of the diffraction lines in the two extremes. Owing to instrumental resolution or the use of polycrystalline samples these fringes are normally not observed. This justifies the replacement of this function by a smoothed version $\bar{B}(\mathbf{s})$ for which we take a Gaussian, giving

$$|\bar{B}(\mathbf{s})|^2 = \bar{V}^2 \exp[-\tfrac{1}{2}(s^2/\sigma_R^2)]. \quad (13)$$

Whereas (12) yields $B(0) = V$, the smoothed version yields an effective area $\bar{V}$, which can be related to an effective number of lattice points $\bar{N}$ and effective crystal radius R_{eff} by

$$\bar{V} = \bar{N}\Omega = \pi R_{\text{eff}}^2,$$

so that σ_R is defined by

$$2R_{\text{eff}}^2\sigma_R^2 = 1.$$

Now we redefine (5) as

$$\bar{I}(\mathbf{s}) = (\bar{N}/\Omega)\exp(-s^2\sigma^2) \times \sum_{n}[(s^2\sigma^2)^n/n!][R_n(\mathbf{s}) * C(\mathbf{s})]. \quad (14)$$

One can think of (14) as representing diffraction from an ensemble of finite lattices with a variation in size and shape, but equal orientation of crystallographic axes.

It should be emphasized that for large values of b the effect of the finite crystal size is important, since the line widths of the diffuse peaks are then comparable to or smaller than the width of $|B(\mathbf{s})|^2$. If one used an infinite-lattice approximation, then $|B(\mathbf{s})|^2$ would behave as a δ function, with negligible width, which instead of (14) gives for (5)

$$I(\mathbf{s}) = (N/\Omega)\exp(-s^2\sigma^2) \times \sum_{n}[(s^2\sigma^2)^n/n!]\sum_{\mathbf{g}}R_n(\mathbf{s}-\mathbf{g}).$$

At a reciprocal-lattice point this function increases without bound for increasing correlation length. For the paracrystalline limit in macromolecular crystals and in particular for biological structures this is an undesirable feature since in these cases the crystal size often cannot be taken as infinite. In our expressions the intensity at the reciprocal-lattice points approximates that of a finite perfect crystal, for b going to infinity.

As an example of the application of (14) we shall take a Gaussian and an exponential form for the correlation function. For the Gaussian form of the correlation function we have

$$\rho(\mathbf{r}) = \exp[-\tfrac{1}{2}(r^2/b^2)], \quad (15)$$

where b is a correlation length. The 2D Fourier transform (10) is

$$R_n(\mathbf{s}) = (2\pi\sigma_n^2)^{-1}\exp[-\tfrac{1}{2}(s^2/\sigma_n^2)], \quad \sigma_n = (n^{1/2})/b. \quad (16)$$

The convolution in (8) can be written as

$$R_n(\mathbf{s}) * C(\mathbf{s}) = V^{-1}\sum_{\mathbf{g}}\int d\mathbf{s}'[R_n(\mathbf{s}'-\mathbf{s})|\bar{B}(\mathbf{s}'-\mathbf{g})|^2], \quad (17)$$

which contains integrals over two Gaussian functions centred on different locations in reciprocal space. In this case the integration is easily done, giving again a Gaussian with variance equal to the sum of the variances of the original distributions (Shavitt, 1963). Applying this result to (17) and substituting in (14) we obtain

$$\bar{I}(\mathbf{s}) = \bar{N}^2\exp(-s^2\sigma^2)\sum_{n}[(s^2\sigma^2)^n/n!] \times \sum_{\mathbf{g}}\frac{(b\sigma_R)^2}{(b\sigma_R)^2+n}\exp\left[-\frac{1}{2}\frac{(\mathbf{s}-\mathbf{g})^2}{\sigma_R^2+n/b^2}\right]. \quad (18)$$

For the exponential form we have

$$\rho(\mathbf{r}) = \exp(-r/b) \quad (19)$$

and for the Fourier transform (10)

$$R_n(\mathbf{s}) = (2\pi)^{-1}\left\{\frac{(n/b)^2}{[(n/b)^2+s^2]^3}\right\}^{1/2}. \quad (20)$$

Table 1. *Parameters obtained by least-squares fitting of* (21) *with three and four Gaussians using the constraint* (23)

	Three Gaussians		Four Gaussians	
i	c_i	γ_i	c_i	γ_i
1	0·006332	6·40452	0·0009781	11·8006
2	0·17406	1·59286	0·031879	2·95908
3	0·78266	0·61772	0·30658	1·14461
4	—	—	0·64794	0·53144

The evaluation of the convolution in (17) with (20) is much more difficult. To overcome this difficulty one could change $|B(\mathbf{s})|^2$ to the Fourier transform of an exponential, corresponding to an approximation of a flat cylindrical shape function. We shall not do this, since such a function is not practical in more complicated situations, as will be discussed later. When a Gaussian shape function is used one can approximate the integral by fitting (20) to a sum of Gaussians. If

$$[1/(1+x^2)^3]^{1/2} \simeq \sum_i c_i \exp[-\tfrac{1}{2}(x^2/\gamma_i^2)] \quad (21)$$

then

$$\left\{\frac{(n/b)^2}{[(n/b)^2+(\mathbf{s}'-\mathbf{s})^2]^3}\right\}^{1/2} \simeq (b/n)^2 \sum_i c_i \exp\left[-\frac{1}{2}\frac{(\mathbf{s}'-\mathbf{s})^2}{(n\gamma_i/b)^2}\right]. \quad (22)$$

We have obtained c_i's and γ_i's by a least-squares fit of (21) with three and four Gaussians. These parameters are given in Table 1. We have kept

$$\sum_i c_i\gamma_i^2 = 1 \quad (23)$$

as a constraint in the fit, otherwise the perfect-lattice limit would not be correctly represented in (24), (29) and (39). Substitution of (22) with (20) in (14) and use of the two-centre convolution property of Gaussians gives for the intensity

$$\bar{I}(\mathbf{s}) = \bar{N}^2 \exp(-s^2\sigma^2) \sum_n [(s^2\sigma^2)^n/n!] \times \sum_i c_i \sum_{\mathbf{g}} \frac{(b\sigma_R)^2}{(b\sigma_R/\gamma_i)^2+n^2} \times \exp\left[-\frac{1}{2}\frac{(\mathbf{s}-\mathbf{g})^2}{\sigma_R^2+(n\gamma_i/b)^2}\right]. \quad (24)$$

The same technique can be used to obtain a formula for a more general monotonically decreasing circularly symmetric type of correlation function. However, if the multiple convolutions in (10) cannot be evaluated then the Gaussian approximation should be applied to $R(\mathbf{s})$ instead of $R_n(\mathbf{s})$. For a large amplitude of the lattice deformation expressed by the single-site standard deviation one has to include many orders of n. For the higher-order terms the convolutions in (10) result in strong error propagation. This requires an accurate approximation of $R(\mathbf{s})$, especially in the tail of the function. We shall illustrate this method with the exponential correlation function. Substitution of (22) with $n=1$ in (10) gives

$$R_n(\mathbf{s}) * = (4\pi^2)^{-n} \times \left\{2\pi b^2 \sum_i c_i \exp[-\tfrac{1}{2}s^2/(\gamma_i/b)^2] *\right\}^n. \quad (25)$$

For three Gaussians we use a twofold binomial expansion, which, together with the two-centre convolution property, yields

$$\begin{aligned} R_n(\mathbf{s}) * &= [b^2/(2\pi)]^n \sum_{p_1,p_2} \binom{n}{p_1}\binom{p_1}{p_2} \\ &\quad \times \{c_1 \exp[-\tfrac{1}{2}s^2/(\gamma_1/b)^2] *\}^{n-p_1} \\ &\quad \times \{c_2 \exp[-\tfrac{1}{2}s^2/(\gamma_2/b)^2] *\}^{p_1-p_2} \\ &\quad \times \{c_3 \exp[-\tfrac{1}{2}s^2/(\gamma_3/b)^2] *\}^{p_2} \\ &= \sum_{p_1,p_2} \binom{n}{p_1}\binom{p_1}{p_2} \Gamma (2\pi Z^2)^{-1} \\ &\quad \times \exp[-\tfrac{1}{2}(s^2/Z^2)] *, \end{aligned} \quad (26)$$

with

$$Z^2 = [(n-p_1)\gamma_1^2 + (p_1-p_2)\gamma_2^2 + p_2\gamma_3^2]/b^2 \quad (27)$$

and

$$\Gamma = (c_1\gamma_1^2)^{n-p_1}(c_2\gamma_2^2)^{p_1-p_2}(c_3\gamma_3^2)^{p_2}. \quad (28)$$

Insertion of (26) into (14) yields

$$\bar{I}(\mathbf{s}) = \bar{N}^2 \exp(-s^2\sigma^2) \times \sum_{n,p_1,p_2} \frac{(s^2\sigma^2)^n}{(n-p_1)!\,(p_1-p_2)!\,p_2!} \times \Gamma \sum_{\mathbf{g}} \frac{\sigma_R^2}{Z^2+\sigma_R^2} \exp\left[-\frac{1}{2}\frac{(\mathbf{s}-\mathbf{g})^2}{Z^2+\sigma_R^2}\right]. \quad (29)$$

When more Gaussians are included in the fit one has to increase the number of binomial expansions, which is a straightforward procedure.

2. *A generalization*

In this section we apply the approximate model to the actomyosin lattice of skeletal muscle tissue which has two types of major protein structures in the unit cell. Objects of type m (myosin) are taken at the lattice points and objects of type a (actin) at the trigonal positions, as indicated in Fig. 1. The proteins are represented by Gaussian electron distributions with different widths and amplitudes. Scattering factors are taken into account. For the disorder we make the following assumptions. The deformation of the unit cell is determined by the displacements of the lattice points as in the simple hexagonal case. The

positions of the actin units are given by the centre of gravity of the three neighbour objects of type m. This assumption is a reasonable first-order approximation if the distortion is of an elastic nature. Additional uncorrelated deviations from these positions can be incorporated into Debye-Waller factors. Referring to Fig. 1 we now have

$$\mathbf{r}_t = \mathbf{t} + \mathbf{d}_t$$
$$\mathbf{r}_{1t} = \mathbf{t} + \mathbf{u}_1 + \mathbf{d}_{1t}$$
$$\mathbf{r}_{2t} = \mathbf{t} + \mathbf{u}_2 + \mathbf{d}_{2t}$$

with

$$\mathbf{d}_{1t} = (\mathbf{d}_t + \mathbf{d}_{t+a} + \mathbf{d}_{t+b})/3$$
$$\mathbf{d}_{2t} = (\mathbf{d}_t + \mathbf{d}_{t-a} + \mathbf{d}_{t-b})/3,$$

which instead of (1) gives for the scattered intensity

$$I(\mathbf{s}) = \left| \sum_t \exp(i\mathbf{s}\cdot\mathbf{t}) \{ f_m \exp(i\mathbf{s}\cdot\mathbf{d}_t) + f_a \exp[i\mathbf{s}\cdot(\mathbf{u}_1 + \mathbf{d}_{1t})] + f_a \exp[i\mathbf{s}\cdot(\mathbf{u}_2 + \mathbf{d}_{2t})] \} \right|^2, \quad (30)$$

with scattering factors given by

$$f_m = g_m \exp(-\tfrac{1}{2} b_m^2 s^2)$$
$$f_a = g_a \exp(-\tfrac{1}{2} b_a^2 s^2),$$

where b_m and b_a are the standard deviations of the electron distributions of type m and a respectively. Normalization of mass per unit cell leads to the assumption

$$g_m + 2g_a = 1.$$

Following the steps leading to (5) we obtain

$$I(\mathbf{s}) = N \sum_t \exp(i\mathbf{s}\cdot\mathbf{t}) (f_m'^2 \exp[s^2\sigma^2\rho_0(\mathbf{t})] + 2f_m' f_a' \{\exp(i\mathbf{s}\cdot\mathbf{u}_1) \exp[s^2\sigma^2\rho_1(\mathbf{t})/3] + \text{c.c.}\} + f_a'^2 \{2 \exp[s^2\sigma^2\rho_2(\mathbf{t})/9] + \exp(2i\mathbf{s}\cdot\mathbf{u}_1) \exp[s^2\sigma^2\rho_3(\mathbf{t})/9] + \text{c.c.}\}) \quad (31)$$

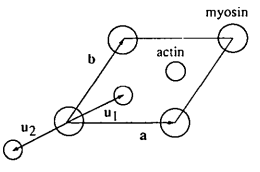

Fig. 1. Unit cell of the actomyosin lattice in skeletal muscle tissue. **a** and **b** denote primitive lattice vectors. $\mathbf{u}_1$ and $\mathbf{u}_2$ locate the actin units in the basis. Myosin is taken at the origin.

with

$$f_m' = \exp(-s^2\sigma^2/2) f_m$$
$$f_a' = \exp\{-s^2\sigma^2[1 + 2\rho(\mathbf{a})]/6\} f_a$$

and

$$\rho_0(\mathbf{t}) = \rho(\mathbf{t}),$$
$$\rho_1(\mathbf{t}) = \rho(\mathbf{t}) + \rho(\mathbf{t}+\mathbf{a}) + \rho(\mathbf{t}+\mathbf{b}),$$
$$\rho_2(\mathbf{t}) = 3\rho(\mathbf{t}) + \rho(\mathbf{t}-\mathbf{a}) + \rho(\mathbf{t}+\mathbf{a}) + \rho(\mathbf{t}-\mathbf{b}) + \rho(\mathbf{t}+\mathbf{b}) + \rho(\mathbf{t}+\mathbf{a}-\mathbf{b}) + \rho(\mathbf{t}-\mathbf{a}+\mathbf{b}), \quad (32)$$
$$\rho_3(\mathbf{t}) = \rho(\mathbf{t}) + \rho(\mathbf{t}+2\mathbf{a}) + \rho(\mathbf{t}+2\mathbf{b}) + 2\rho(\mathbf{t}+\mathbf{a}) + 2\rho(\mathbf{t}+\mathbf{b}) + 2\rho(\mathbf{t}+\mathbf{a}+\mathbf{b}).$$

Expanding the exponentials involving the correlation functions and using (7) we obtain

$$\bar{I}(\mathbf{s}) = (\bar{N}/\Omega) \sum_n [(s^2\sigma^2)^n/n!] (f_m'^2 A_0^n(\mathbf{s}) + f_m' f_a' [2 \exp(i\mathbf{s}\cdot\mathbf{u}_1) A_1^n(\mathbf{s}) + \text{c.c.}] + f_a'^2 \{2A_2^n(\mathbf{s}) + [\exp(2i\mathbf{s}\cdot\mathbf{u}_1) A_3^n(\mathbf{s}) + \text{c.c.}]\}), \quad (33)$$

with

$$A_0^n(\mathbf{s}) = (4\pi^2)^{-n} [R(\mathbf{s}) *]^n C(\mathbf{s})$$
$$A_1^n(\mathbf{s}) = (4\pi^2)^{-n} [\varphi(\mathbf{s}) R(\mathbf{s}) *]^n C(\mathbf{s})$$
$$A_2^n(\mathbf{s}) = (4\pi^2)^{-n} [|\varphi(\mathbf{s})|^2 R(\mathbf{s}) *]^n C(\mathbf{s}) \quad (34)$$
$$A_3^n(\mathbf{s}) = (4\pi^2)^{-n} [\varphi(\mathbf{s})^2 R(\mathbf{s}) *]^n C(\mathbf{s})$$

and

$$\varphi(\mathbf{s}) = [1 + \exp(-i\mathbf{s}\cdot\mathbf{a}) + \exp(-i\mathbf{s}\cdot\mathbf{b})]/3. \quad (35)$$

The Fourier transforms of the translated correlation functions in (32) give rise to complex exponential factors. These phase factors are incorporated in φ. For narrow diffraction lines (compared with the reciprocal-lattice spacing) they represent slowly varying waves with a value of unity at the reciprocal-lattice points. In this case they could be set to one. For broader lines, however, this approximation does not conserve the hexagonal symmetry of the scattered intensity, so here they have to be taken into account, making the evaluation of the multiple convolutions more difficult. One meets two-centre integrals which are Fourier transforms of two-centre distributions. These types of integrals have been studied in the context of scattering from bonded atoms (McWeeny, 1952, 1953). The results show that the evaluation of (34) for exponential distributions would be extremely tedious. For the Gaussian case, however, the

integration is much simpler since one has

$$\begin{aligned}&\exp(i\mathbf{s}\cdot\mathbf{a}_1-\tfrac{1}{2}s^2/\sigma_1^2)*\exp(i\mathbf{s}\cdot\mathbf{a}_2-\tfrac{1}{2}s^2/\sigma_2^2)\\&=2\pi[\sigma_1^2\sigma_2^2/(\sigma_1^2+\sigma_2^2)]\\&\quad\times\exp[-\tfrac{1}{2}(\mathbf{a}_1-\mathbf{a}_2)^2/(\sigma_1^{-2}+\sigma_2^{-2})]\\&\quad\times\exp[i\mathbf{s}\cdot(\sigma_1^2\mathbf{a}_1+\sigma_2^2\mathbf{a}_2)/(\sigma_1^2+\sigma_2^2)\\&\quad-\tfrac{1}{2}s^2/(\sigma_1^2+\sigma_2^2)].\end{aligned}\tag{36}$$

Here, $\mathbf{a}_1$ and $\mathbf{a}_2$ are arbitrary real-space vectors. In (34) the difference between these vectors is never larger than one lattice constant. For sufficiently large correlation lengths this implies that the exponential amplitude factor on the right-hand side of (36) can be taken equal to unity. In this case the right-hand side has the same form as each of the operands on the left. Hence, an n-fold convolution also results in this form, yielding

$$\begin{aligned}&\exp(i\mathbf{s}\cdot\mathbf{a}_1-\tfrac{1}{2}s^2/\sigma_1^2)*\ldots*\exp(i\mathbf{s}\cdot\mathbf{a}_n-\tfrac{1}{2}s^2/\sigma_n^2)\\&=(2\pi)^{n-1}[(\sigma_1^2\ldots\sigma_n^2)/(\sigma_1^2+\ldots+\sigma_n^2)]\\&\quad\times\exp[i\mathbf{s}\cdot(\sigma_1^2\mathbf{a}_1+\ldots+\sigma_n^2\mathbf{a}_n)/(\sigma_1^2+\ldots+\sigma_n^2)\\&\quad-\tfrac{1}{2}s^2/(\sigma_1^2+\ldots+\sigma_n^2)].\end{aligned}\tag{37}$$

This property makes complex disorder models tractable, since now multiple convolutions do not lead to increasingly difficult expressions. Analogous to (26) one obtains for (34)

$$A_i^n(\mathbf{s})=\sum_{p_1,p_2}\binom{n}{p_1}\binom{p_1}{p_2}\psi_i\Gamma\frac{1}{2\pi Z^2}\times\exp(-\tfrac{1}{2}s^2/Z^2)*C(\mathbf{s}),\tag{38}$$

for $i=0, 1, 2, 3$, with

$$\psi_0=1,\qquad \psi_1=\Phi(\mathbf{s}),$$
$$\psi_2=|\Phi(\mathbf{s})|^2,\qquad \psi_3=\Phi(\mathbf{s})^2$$

and

$$\begin{aligned}\Phi(\mathbf{s})&=\varphi_{1,Z}(\mathbf{s})^{n-p_1}\varphi_{2,Z}(\mathbf{s})^{p_1-p_2}\varphi_{3,Z}(\mathbf{s})^{p_2}\\\varphi_{j,Z}(\mathbf{s})&=[1+\exp(-i\mathbf{s}\cdot\mathbf{a}\sigma_j^2/Z^2)\\&\quad+\exp(-i\mathbf{s}\cdot\mathbf{b}\sigma_j^2/Z^2)]/3,\end{aligned}$$

where $\sigma_j=\gamma_j/b$. In order to obtain this result one has to use additional binomial expansions in intermediate steps and liberally use the convolution approximation (37). The validity of this approach will be discussed further in the next section.

Finally, after evaluating the convolution with $C(\mathbf{s})$ in (38) we can write the intensity (33) as

$$\bar{I}(\mathbf{s})=\bar{N}^2\sum_{n,p_1,p_2}\frac{(s^2\sigma^2)^n}{(n-p_1)!\,(p_1-p_2)!\,p_2!}\Gamma\frac{\sigma_R^2}{Z^2+\sigma_R^2}\times\sum_{\mathbf{g}}F_{\mathbf{g}}(\mathbf{s})^2\exp[-\tfrac{1}{2}(\mathbf{s}-\mathbf{g})^2/(Z^2+\sigma_R^2)],\tag{39}$$

with

$$\begin{aligned}F_{\mathbf{g}}(\mathbf{s})&=f_m'+f_a'[\exp(i\mathbf{s}\cdot\mathbf{u}_1)\bar{\Phi}(\mathbf{s}-\mathbf{g})+\text{c.c.}]\\\bar{\Phi}(\mathbf{s})&=\bar{\varphi}_{1,Z}(\mathbf{s})^{n-p_1}\bar{\varphi}_{2,Z}(\mathbf{s})^{p_1-p_2}\bar{\varphi}_{3,Z}(\mathbf{s})^{p_2}\\\bar{\varphi}_{j,Z}(\mathbf{s})&=\{1+\exp[-i\mathbf{s}\cdot\mathbf{a}\sigma_j^2/(Z^2+\sigma_R^2)]\\&\quad+\exp[-i\mathbf{s}\cdot\mathbf{b}\sigma_j^2/(Z^2+\sigma_R^2)]\}/3.\end{aligned}$$

Equation (39) has the correct hexagonal symmetry (if the crystal is not too small) and is valid for correlation lengths substantially larger than a lattice constant.

Discussion

The general character of the expressions for the scattered intensity is similar to that of the expressions derived by Welberry & Carroll (1983) for their general Gaussian model. The amplitude of the disturbances in the lattice is related to the single-site standard deviation σ. For small values of σ the Bragg scattering dominates the spectrum for all values of the correlation length, *i.e.* the $n=0$ term is larger than the contribution of the higher-order terms. This is the situation usually found in well ordered atomic crystals. For larger values of σ, diffuse scattering becomes more important and can exceed the Bragg intensity. The character of the diffuse terms is determined by σ as well as the correlation length b. If we take σ so large that Bragg reflection is undetectable due to the small Debye–Waller factor and at the same time b so large that the diffuse scattering results in pronounced peaks, then one can speak of a paracrystal-like situation. In this situation the low-spatial-frequency components dominate the power spectrum of the lattice disturbances and the diffuse pattern reflects the local ordering. In this case many orders of n have to be taken into account, in for example (5), since it is seen in the intensity expressions that larger values of σ lead to a slower convergence of the series expansion. The same effect is observed for increasing distance in reciprocal space since the length of the scattering vector plays the same role as σ in the series expansion. The series expansion represents an increasing function of the scattering vector s, which is, however, progressively more attenuated by the Debye–Waller factor for larger distances in reciprocal space. In these circumstances, the general appearance of the diffraction pattern of a simple lattice of δ functions is that of a pronounced first-order peak with a few weaker and broader next-higher orders.

The intensity expressions (18), (24), (29) and (39) have been investigated on the computer. Most computations were done on a Sun4/280. The systems which have been studied consisted of a hexagonal lattice with 900 lattice points, corresponding to $R_{\text{eff}}=15{\cdot}7$ lattice constants. For a hexagonal lattice with lattice constant of 460 Å the diameter of this crystal is

roughly 1μm. These values are representative for the myosin filament lattice found in a myofibril of skeletal muscle tissue (Yu, Steven, Naylor, Gamble & Podolsky, 1985). At first sight, increasing the number of Gaussians in the fit (21) leads to a polynomial increase of CPU time in (29) and (39). However, the summations in these expressions contain many terms which are small enough to be neglected. A suitable criterion to skip the execution of a loop over p_i can be found by considering the effect of the variation of p_i on Z. If Z does not change significantly for the range of allowed p_i values, then Z can be taken independent of p_i, in which case the summation over p_i can be eliminated. Careful tuning of this approximation brings down the CPU time for the computation of (29) and (39) to that of the order necessary for (24). For $\sigma = 0{\cdot}3$ and four Gaussians in the fit only 1% accuracy is lost.

In order to investigate the effect of variation of the disorder parameters we have evaluated (24) for various values of b and σ. The number of relevant terms in the expansion increases rapidly with σ and s. For $\sigma = 0{\cdot}3$ lattice constant and s 3·5 times the value for the first-order Bragg reflection, approximately 80 terms should be included to obtain an accuracy of 2%. For these values all variables could be kept within the range of standard double-precision reals. We observed that a slight deviation from unity for the correlation coefficient already gives a substantial intensity decrease for the higher-order reflections. In the case of the simple hexagonal lattice, for a correlation length of 500 lattice constants (*i.e.* much larger than the crystal size), the amplitude of the 11 reflection is 75% of that of the 10 reflection. Line broadening increases for lower b. These effects are stronger for higher values of σ, for equal b. This is a familiar result since it is known that it is the product of σ and b, which is related to the standard deviation of the length of the cell edges in the paracrystal model, which determines the line width. For correlation lengths lower than 1 lattice constant diffuse scattering vanishes and a small first-order Bragg peak remains. For $\sigma = 0{\cdot}3$ the magnitude of this peak is only 1% of that of the perfect lattice intensity. For higher values of σ this peak will be negligible.

In Fig. 2, we compare the intensity calculated with (18), (24) and (29). For $b = 50$, scans are made along three directions in reciprocal space, passing through the 10, 11 and 21 reflections. The results are summed. It is seen that the patterns obtained with (18) (Fig. 2*a*) and (24) (Fig. 2*d*) differ considerably. The exponential correlation gives much more line broadening than the Gaussian correlation function. By approximation of the exponential correlation function with three Gaussians one obtains the pattern in Fig. 2(*b*), which is only qualitatively similar to Fig. 2(*d*). Clearly three Gaussians are not enough to give a good representation of the multiple convolutions in (29). The result of a fit with four Gaussians, shown in Fig. 2(*c*), does give a fair representation of (24). The average intensity difference at the maximum of the peaks between (24) and (29) is of the order of 10%. The agreement is worse for higher angles because here the number of convolutions in (10) is higher and the result is therefore more affected by error propagation.

In Fig. 3 detailed plots of the scattered intensity in the low-angle region of reciprocal space are shown for correlation lengths with strong line broadening. The figure shows contour levels of the intensity in the irreducible low-angle part of reciprocal space and the corresponding powder pattern, obtained by sampling at angles of 1°. Comparison of corresponding panels in Figs. 3(*a*) and (*b*) shows considerable difference in the low-intensity levels between the peaks. Notice that the contour levels are drawn on an exponential scale. Obviously, the Gaussian approximation cannot accurately represent the intensity in this region for exponential correlation, even with four terms in the fit. The higher intensity near the peaks, however, is very similar in both figures. This is reflected in the good correspondence of the powder patterns.

Fig. 3(*c*) shows results obtained with (39). A mass ratio myosin/actin of 2 was chosen, giving $g_m = 0{\cdot}5$

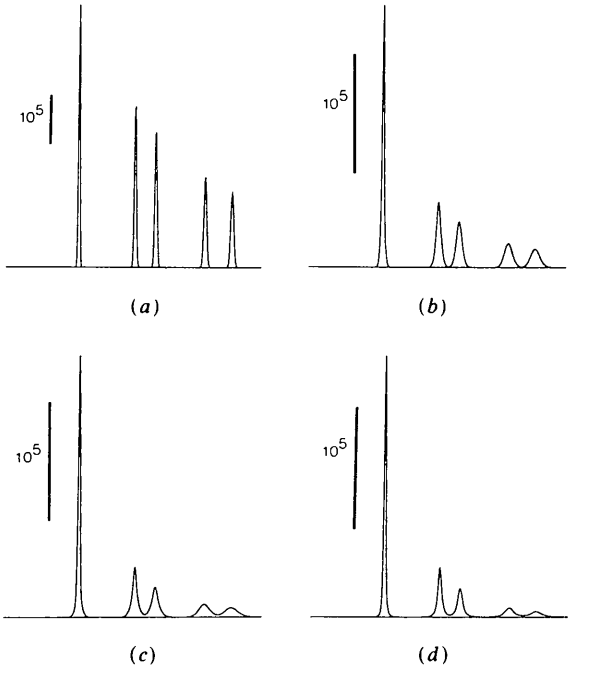

Fig. 2. Summed intensity profiles along three directions in the reciprocal space of a distorted simple hexagonal lattice of δ functions. The scans pass through the 10, 11 and 21 reflections. $\sigma = 0{\cdot}3$ and $b = 50$ (in lattice constants). Vertical bars indicate the intensity scale. (*a*) was obtained with equation (18). In (*b*) and (*c*) equation (29) was used with three and four Gaussians respectively. (*d*) was calculated with equation (24).

and $g_a = 0{\cdot}25$. These values reflect the corresponding protein composition in muscle. The width of the electron distribution is taken as 0·1 lattice constant for both species (*i.e.* $b_m = b_a = 0{\cdot}1$). The perfect lattice intensities for the 10, 11, 20, 21 and 30 reflections are obtained as 1·0, 5·58, 0·21, 0·04 and 0·24 relative to the 10 peak. It is seen that for small correlation lengths the 11 peak is severely broadened while its amplitude is greatly reduced. We recall that in deriving (39) the exponential amplitude factors in (36) have been neglected. Since in the intermediate steps a cumulation of these factors occurs, the validity of (37) remains to be discussed. It is not difficult to obtain the equivalent of (37) with the amplitude factors included. This would, however, lead to additional and much more complicated summations in (38), due

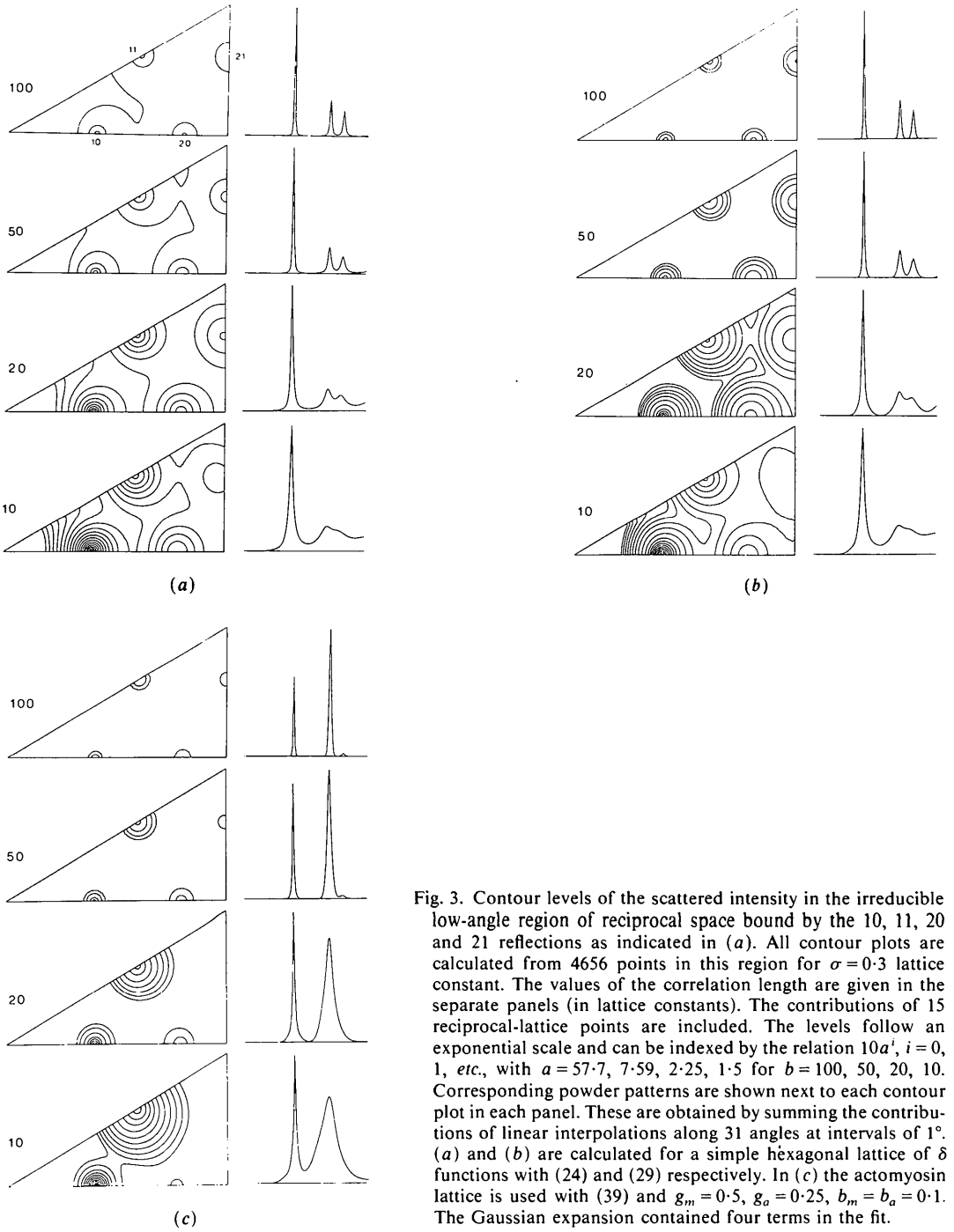

Fig. 3. Contour levels of the scattered intensity in the irreducible low-angle region of reciprocal space bound by the 10, 11, 20 and 21 reflections as indicated in (*a*). All contour plots are calculated from 4656 points in this region for $\sigma = 0{\cdot}3$ lattice constant. The values of the correlation length are given in the separate panels (in lattice constants). The contributions of 15 reciprocal-lattice points are included. The levels follow an exponential scale and can be indexed by the relation $10a^i$, $i = 0$, 1, *etc.*, with $a = 57{\cdot}7$, 7·59, 2·25, 1·5 for $b = 100$, 50, 20, 10. Corresponding powder patterns are shown next to each contour plot in each panel. These are obtained by summing the contributions of linear interpolations along 31 angles at intervals of 1°. (*a*) and (*b*) are calculated for a simple hexagonal lattice of δ functions with (24) and (29) respectively. In (*c*) the actomyosin lattice is used with (39) and $g_m = 0{\cdot}5$, $g_a = 0{\cdot}25$, $b_m = b_a = 0{\cdot}1$. The Gaussian expansion contained four terms in the fit.

to additional binomial expansions. For correlation lengths as low as 10 lattice constants the amplitude factors can reduce the intensity contribution of terms in these summations by more than 50%. It is possible to modify (39) to take this effect into account, in its worst case, in the following way. If in an additional summation all terms are included, but with their amplitude factors replaced by the smallest value which the amplitude factor can have for that summation, then (39) would include an exponential factor depending on the summation indices n and p_i only, with a value between zero and unity. With this factor included, the intensity of the 11 peak in Fig. 3(c) for $b=10$ drops 15%, while for $b=20$ this is 5%. The shape is virtually unaltered. Since this worst-case analysis heavily overestimates the intensity reduction, we can safely use (39) for values of the correlation length down to 10 lattice constants at the the expense of some overestimation of the intensity of the higher-order peaks for values of b near 10.

In Figs. 3(a), (b) and (c), for correlation lengths smaller than about 20 lattice constants, we observed line shifts. The maximum of the broad scattering peaks shifts in the direction of the central beam, as is clearly visible in the shift of the 21 position in Fig. 3(a). For $b=10$, the 11 peak in Fig. 3(c) deviates 2% from its Bragg position. These shifts are a result of the balance between the low-pass filtering effect by the Debye–Waller factor on broad components in the intensity expressions and high-pass filtering by the powers of n of the product $s\sigma$, which are rapidly increasing functions of s. The latter thus counteract the effect of the Debye–Waller factor. Multiplication of a broad symmetric function centred on a reciprocal-lattice point with a rapidly decreasing exponential function yields an asymmetric function with a shifted position of its maximum. Since this shift is different for different diffraction orders the position ratios as valid for Bragg reflections are affected.

A related point concerns the shape of the diffraction spots. Although 15 reciprocal-lattice points are taken into account in calculating Fig. 3, the intensity near the peaks mainly reflects the local contribution for all correlation lengths. The rounded appearance is a direct result of the explicit introduction of isotropic correlation, resulting in circularly symmetric Fourier transforms of the correlation function. In the case of severe line broadening the attenuating and amplifying effects of the mechanisms mentioned before can lead to asymmetric peaks which do not have to be perfectly circularly symmetric. These effects are visible in the contour levels in Figs. 3(a) and (b). In Fig. 3(c) the scattering factors give additional low-pass filtering. The powder pattern peaks are asymmetric for $b=10$ and $b=20$.

In this context we should mention the line-shape-analysis method of Warren & Averbach (1950) and Warren (1955), from which intensity expressions similar to ours can be derived. The difference from our treatment lies in their explicit assumption that the peaks must not be too broad. In our formulae this means that $s\sigma$ would be replaced by $g\sigma$. From the previous discussion it is clear that, with this modification, only perfectly symmetric lines are obtained and no line shifts. This indicates that, for small correlation lengths, the present method gives more accurate results than a Warren & Averbach-type analysis.

Welberry & Carroll (1983) use the degree of local circular symmetry around diffuse peaks as an indication of the isotropy in the diffraction patterns obtained from optical diffraction analogues generated by a Monte Carlo method. They used a model on a square lattice with one axial and one diagonal nearest-neighbour interaction. It was found that with an axial nearest-neighbour correlation coefficient of 0·9 the most isotropic case gave a diagonal correlation coefficient of 0·88. In the non-isotropic growth disorder model the diagonal correlation is given by the product of the axial correlations, *i.e.* 0·81, while the axial correlation decays exponentially. When we calibrate our exponential correlation function (19) with a nearest-neighbour coefficient of 0·9 we obtain a diagonal coefficient of 0·86. For higher-order diagonal correlation values the exponential correlation function gives lower values than the Monte Carlo model but higher values than the growth-disorder models. The discrepancy suggests that more-distant-neighbour interactions are necessary to describe an isotropic exponential correlation field on a square lattice. Welberry & Carroll also pointed out that in 2D and 3D more complicated correlation functions are expected to result from nearest-neighbour-only interactions.

Concluding remarks

We have shown that relatively simple expressions for diffraction patterns can be obtained from a general Gaussian model of disorder by the explicit introduction of an isotropic correlation field. The expressions are valid for large single-site standard deviations and relatively small correlation lengths and for a finite crystal size. This is particularly relevant for paracrystal-like behaviour in biological systems, like muscle, where the size of the ordered region is often limited. Although in this paper we have only studied exponential and Gaussian correlation, the method is suitable for other forms of the correlation function in 2D or 3D. In this case, as in more complex distorted crystals, the evaluation of the necessary integrals will be more difficult so that it is useful to fit the correlation function or its Fourier transform to a number of Gaussians in order to obtain an approximate solution to the problem. This formalism can be sufficiently flexible

and accurate to serve as a convenient, although approximate, alternative for optical diffraction experiments or numerical simulations of more complicated systems with more than one scattering object in the unit cell and with form factors taken into account.

This study was supported by the Netherlands Organization for the Advancement of Pure Research (NWO).

References

COWLEY, J. M. (1984). *Difraction Physics.* Amsterdam: North-Holland.
HOSEMANN, R. (1975). *J. Polym. Sci.* **50**, 265-281.
HOSEMANN, R. & BAGCHI, S. N. (1962). *Direct Analysis of Diffraction by Matter.* Amsterdam: North-Holland.
KITTEL, C. (1976). *Introduction to Solid State Physics.* New York: John Wiley.
MCWEENY, R. (1952). *Acta Cryst.* **5**, 463-468.
MCWEENY, R. (1953). *Acta Cryst.* **6**, 631-637.
SHAVITT, I. (1963). *Methods Comput. Phys.* **2**, 1-45.
THAKUR, M., TRIPATHY, S. K., BURKHART, C. W. & LANDO, J. B. (1985). *Acta Cryst.* A**41**, 30-34.
THAKUR, M., TRIPATHY, S. K. & LANDO, J. B. (1985). *Acta Cryst.* A**41**, 26-30.
WARREN, B. E. (1955). *Acta Cryst.* **8**, 483-486.
WARREN, B. E. & AVERBACH, B. L. (1950). *J. Appl. Phys.* **21**, 595-599.
WELBERRY, T. R. (1985). *Rep. Prog. Phys.* **48**, 1543-1593.
WELBERRY, T. R. & CARROLL, C. E. (1982). *Acta Cryst.* A**38**, 761-772.
WELBERRY, T. R. & CARROLL, C. E. (1983). *Acta Cryst.* A**39**, 233-245.
WELBERRY, T. R., MILLER, G. H. & CARROLL, C. E. (1980). *Acta Cryst.* A**36**, 921-929.
YU, L. C., STEVEN, A. C., NAYLOR, G. R. S., GAMBLE, R. C. & PODOLSKY, R. J. (1985). *Biophys. J.* **47**, 311-321.

Acta Cryst. (1989). A**45**, 870-878

Quantitative Determination of Phases of X-ray Reflection from Three-Beam Diffraction. III. Experiments on Mosaic Crystals

BY SHIH-LIN CHANG, MAU-TZAI HUANG AND MAU-TSU TANG

Department of Physics, National Tsing Hua University, Hsinchu, Taiwan 30043

AND CHIH-HAO LEE

Synchrotron Radiation Research Center, Hsinchu, Taiwan 30043

(*Received* 18 *July* 1989; *accepted* 10 *August* 1989)

Abstract

The method proposed by Chang & Tang [*Acta Cryst.* (1988), A**44**, 1065-1072] for quantitative determination of phases of X-ray reflection from three-beam diffraction profiles is applied to an organic crystal of (2*R*, 3*R*)-3-acetoxy-5,7-dihydroxy-6-methylflavanone, with a mosaic spread of 0·08°. The expression for the kinematical diffraction intensities is modified according to the kinematical theory of X-ray diffraction in the multibeam regime. Three-beam *Umweganregung* and *Aufhellung* diffraction profiles are analyzed. With the help of the modified intensity expression, the crystal symmetry imposed by the space group and the three-beam diffraction geometry, four acentric phases and fourteen centric phases of structure-factor triplets are determined.

I. Introduction

A theoretical consideration of quantitative phase determination using three-beam multiple diffraction has recently been reported (Chang & Tang, 1988). Phase determination for diffractions from perfect crystals was also demonstrated subsequently (Tang & Chang, 1988). These two reports are hereafter referred to as papers I and II in the following discussions. In paper I, a general formalism was derived from the dynamical theory of X-ray diffraction. The first-order approximation in the iterative procedure for the fundamental equation of the wavefield was adopted for *Umweganregung* three-beam diffractions, where the multiple diffracted intensity is greater than the two-beam intensity background. Boundary conditions were considered for plate-like crystals. For *Aufhellung* three-beam diffractions, where the multiple diffraction intensity is lower than the two-beam background, a second-order approximation was proposed to deduce the expressions for diffraction intensities. In paper II, the experiments for plate-like perfect crystals of GaAs demonstrated the possibility of determining quantitatively the phases of acentric reflections.

In mosaic crystals, lattice perfection is not guaranteed. According to Zachariasen (1945), diffractions

0108-7673/89/120870-09$03.00