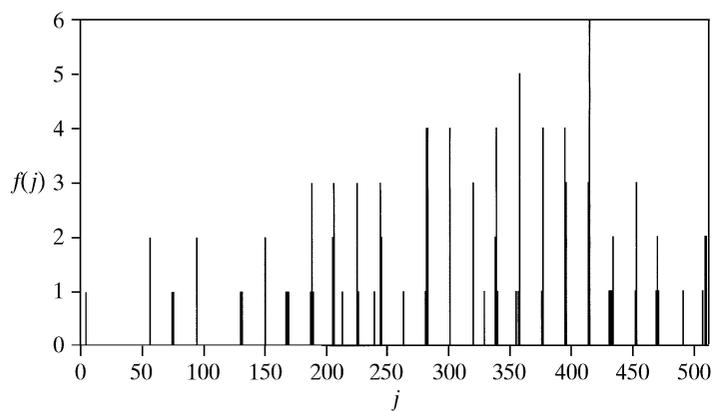11. DATA PROCESSING



Fig. 11.1.3.1. Frequency distribution of the projected reciprocal-lattice vectors for a suitably chosen direction of a diffraction pattern from a fibritin crystal (Tao *et al.*, 1997). Reproduced with permission from Steller *et al.* (1997). Copyright (1997) International Union of Crystallography.

Let $\mathbf{t}$ represent a dimensionless unit vector of a chosen direction. Then, the projection $p$ of a reciprocal-lattice point $\mathbf{x}$ onto the chosen vector $\mathbf{t}$ is given by

$$p = \mathbf{x} \cdot \mathbf{t}. \qquad (11.1.3.1)$$

To apply a discrete FFT algorithm, all such projections of the reciprocal-lattice points onto the chosen direction $\mathbf{t}$ are sampled in small increments of $p$. For the given direction, the values of the projections are in a range between the endpoints $p_{min}$ and $p_{max}$. If the maximum real cell dimension is assumed to be $a_{max}$, then the maximum number of reciprocal-lattice planes between the observed limits of $p$ is $(p_{max} - p_{min})/(1/a_{max})$. Hence, the number of useful grid points along the direction $\mathbf{t}$ should be

$$m = (p_{max} - p_{min})na_{max}, \qquad (11.1.3.2)$$

where $n$ represents the number of grid points between successive reciprocal-lattice planes and is normally set to 5. Then, the frequency $f(p)$ in the range $p < \mathbf{x} \cdot \mathbf{t} < p + \Delta p$ can be given as $f(p)\Delta p = f(j)$, where $j$ is the closest integer to $(p - p_{min})/\Delta p$ and $\Delta p = na_{max}$. Thus, the discrete Fourier transform of this frequency distribution will be given by the summation

$$F(k) = \sum_{j=0}^{m} f(j) \exp(2\pi i k j). \qquad (11.1.3.3)$$

The transform is then calculated using a fast Fourier algorithm for all integer values between 0 and $m/2$ (Fig. 11.1.3.2). The Fourier coefficients that best represent the periodicity of the frequency distribution will be large. The largest coefficient will occur at $k = 0$ and correspond to the number of vectors used in establishing the frequency distribution. The next set of large coefficients will correspond to the periodicity that represents every reciprocal-lattice plane. The ratio of this maximum to $F(0)$ will be a measure of the tightness of the frequency distribution around each lattice plane. Subsequent maxima will be due to periodicities spanning every second, third *etc.* frequency maximum and will thus be progressively smaller (Fig. 11.1.3.2). The largest $F(k)$ (when $k = l$), other than $F(0)$, will, therefore, correspond to an interval of $d^*$ between reciprocal-lattice planes in the direction of $\mathbf{t}$ where $d^* = l/(na_{max})$.

## 11.1.4. Exploring all possible directions to find a good set of basis vectors

The polar coordinates $\psi$, $\varphi$ will be used to define the direction $\mathbf{t}$, where $\psi$ defines the angle between the X-ray beam and the chosen direction $\mathbf{t}$. The Fourier analysis is performed for each direction $\mathbf{t}$ in the range $0 < \psi \leq \pi/2, 0 < \varphi \leq 2\pi$. A suitable angular increment in $\psi$ was determined empirically to be about 0.03 rad (1.7°). For each value of $\psi$, the increment in $\varphi$ is taken to be the closest integral value to $(2\pi \sin \psi)/0.03$. This procedure results in $\sim 7300$ separate, roughly equally spaced, directions.

For each direction $\mathbf{t}$, the distribution of the corresponding $F(k)$ coefficients is surveyed to locate the largest local maximum at $k = l$. The $\psi$ and $\varphi$ values associated with the 30 largest maxima are selected for refinement by a local search procedure to obtain an accuracy of $10^{-4}$ rad ($\sim 0.006°$). If the initial angular increment (0.03 rad) used for the hemisphere search was reduced, then it would not be necessary to refine quite as many local maxima. However, to increase the efficiency of the search procedure, the ratio of angular increments to the number of refined positions was chosen to minimize the total computing time. The $F(l)$ values of the refined positions are then sorted by size. Directions are chosen from these vectors to give a linearly independent set of three basis vectors of a primitive real-space unit cell. These are then converted to the basis vectors of the reciprocal cell. The components of the three reciprocal-cell axes along the three camera axes are the nine components of the crystal orientation matrix $[A]$ (11.1.2.2).

The final step in the selection of the best $[A]$ matrix is to choose various nonlinear combinations of the refined vectors that have the biggest $F(l)$ values. That set of three vectors which gives the best indexing results is then chosen to represent the crystal orientation matrix $[A]$. A useful criterion is to determine the nonintegral Miller indices $h'$ from (11.1.2.4) using the $[A]$ matrix and the known reciprocal-lattice vectors $\mathbf{x}$. Any reflection for which any component $|\mathbf{h} - \mathbf{h}'|$ is bigger than, say, 0.2 is rejected. The best $[A]$ matrix is chosen as the one with the least number of rejections. In most cases, the best combination corresponds to taking the three largest $F(l)$ values.

The program goes on to determine a reduced cell from the cell obtained by the above indexing procedure (Kim, 1989). The reduced cell is then analysed in terms of the 44 lattice characters (Burzlaff *et al.*, 1992; Kabsch, 1993) in order to evaluate the most likely Bravais lattice and crystal system.
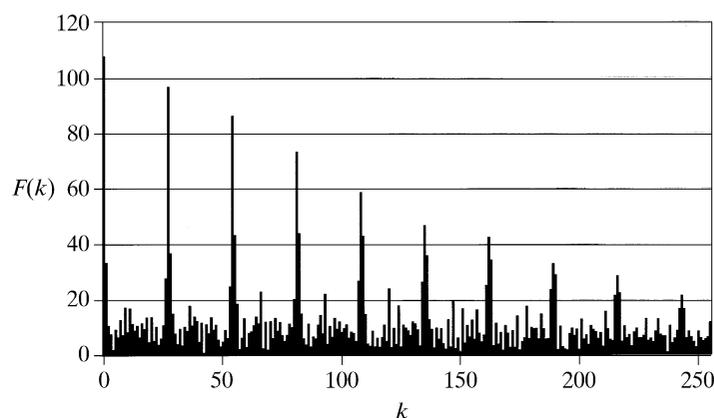


Fig. 11.1.3.2. Fourier analysis of the distribution shown in Fig. 11.1.3.1. The first maximum, other than $F(0)$, is at $k = 27$, corresponding to $(1/d^*) = 41.9$ Å and a value of $F(27) = 97.0$. Reproduced with permission from Steller *et al.* (1997). Copyright (1997) International Union of Crystallography.

210

**references**