electron density in the unit cell:

$$F(hkl) = V \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho(xyz) \exp[2\pi i(hx + ky + lz)]\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z. \tag{2.1.7.5}$$

Because $F(hkl)$ is a vector in the Argand diagram with an amplitude $|F(hkl)|$ and a phase angle $\alpha(hkl)$,

$$F(hkl) = |F(hkl)| \exp[i\alpha(hkl)]$$

and

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l |F(hkl)| \exp[-2\pi i(hx + ky + lz) + i\alpha(hkl)]. \tag{2.1.7.6}$$

By applying equation (2.1.7.6), the electron-density distribution in the unit cell can be calculated, provided values of $|F(hkl)|$ and $\alpha(hkl)$ are known. From equation (2.1.6.1), it is clear that $|F(hkl)|$ can be derived, on a relative scale, from $I_{\rm int}(hkl)$ after a correction for the background and absorption, and after application of the Lorentz and polarization factor:

$$|F(hkl)| = \left[\frac{I_{\rm int}(hkl)}{LPT}\right]^{1/2}. \tag{2.1.7.7}$$

Contrary to the situation with crystals of small compounds, it is not easy to find the phase angles $\alpha(hkl)$ for crystals of macromolecules by direct methods, although these methods are in a state of development (see Part 16). Indirect methods to determine the protein phase angles are:

(1) isomorphous replacement (see Part 12);

(2) molecular replacement (see Part 13);

(3) multiple-wavelength anomalous dispersion (MAD) (see Part 14).

From equation (2.1.7.5), it is clear that the reflections $hkl$ and $\overline{hkl}$ have the same value for their structure-factor amplitudes, $|F(hkl)| = |F(\overline{hkl})|$, and for their intensities, $I(hkl) = I(\overline{hkl})$, but have opposite values for their phase angles, $\alpha(hkl) = -\alpha(\overline{hkl})$, assuming that anomalous dispersion can be neglected. Consequently, equation (2.1.7.6) reduces to

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l |F(hkl)| \cos[2\pi(hx + ky + lz) - \alpha(hkl)] \tag{2.1.7.8}$$

or

$$\rho(xyz) = F(000)/V + (2/V) \sum_h{}' \sum_k{}' \sum_l{}' |F(hkl)| \times \cos[2\pi(hx + ky + lz) - \alpha(hkl)]. \tag{2.1.7.9}$$

$\sum'$ denotes that $F(000)$ is excluded from the summation and that only the reflections $hkl$, and not $\overline{hkl}$, are considered.

The two reflections, $hkl$ and $\overline{hkl}$, are called Friedel or Bijvoet pairs.

If anomalous dispersion cannot be neglected, the two members of a Friedel pair have different values for their structure-factor amplitudes, and their phase angles no longer have opposite values. This is caused by the $f''$ contribution to the anomalous scattering (Fig. 2.1.7.1). Macromolecular crystals show anomalous dispersion if the structure contains, besides the light atoms, one or more heavier atoms. These can be present in the native structure or are introduced in the isomorphous replacement technique or in MAD analysis.

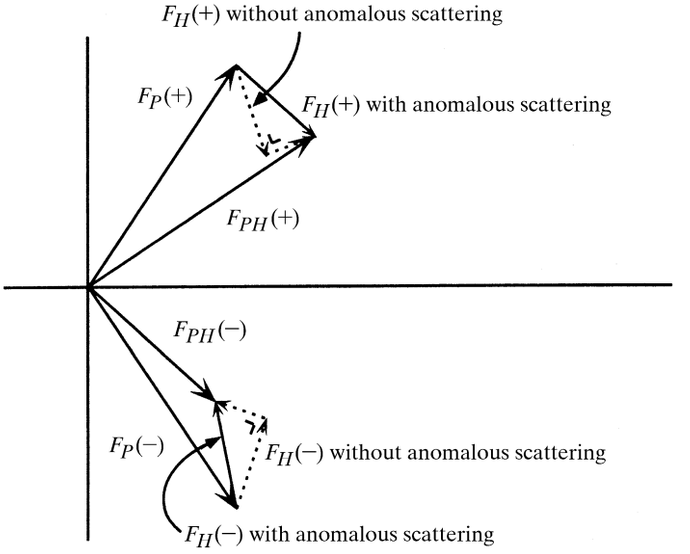


Fig. 2.1.7.1. An Argand diagram for the structure factors of the two members of a Friedel pair. (+) represents $hkl$ and (−) represents $\overline{hkl}$. $F_P$ is the contribution to the structure factor by the non-anomalously scattering protein atoms and $F_H$ is that for the anomalously scattering atoms. $F_H$ consists of a real part with an imaginary part perpendicular to it. The real parts are mirror images with respect to the horizontal axis. The imaginary parts are rotated counterclockwise with respect to the real parts (Section 2.1.4.4). The result is that the total structure factors, $F_{PH}(+)$ and $F_{PH}(-)$, have different amplitudes and phase angles. Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

## 2.1.8. Symmetry in the diffraction pattern

In the previous section, it was noted that $I(hkl) = I(\overline{hkl})$ if anomalous scattering can be neglected. In this case, the effect is that the diffraction pattern has a centre of symmetry. This is also true for the reciprocal lattice if the reciprocal-lattice points $(hkl)$ are weighted with their $I(hkl)$ values. If the crystal structure has symmetry elements, they are also found in the diffraction pattern and in the weighted reciprocal lattice. Macromolecular crystals of biological origin are enantiomorphic and the symmetry operators in the crystal are restricted to rotation axes and screw axes. It is evident that a rotation of the real lattice will cause the same rotation of the reciprocal lattice. If this rotation is the result of a symmetry operation around an axis, the crystal structure looks exactly the same as before the rotation, and the same must be true for the weighted reciprocal lattice. However, screw axes in the crystal lattice reduce to normal (non-screw) rotation axes in the weighted reciprocal lattice, as has been shown by Waser (1955). We follow his arguments, but must first introduce matrix notation for convenience.

If $\mathbf{r}$ is a position vector and $\mathbf{h}$ a vector in reciprocal space, the scalar product

$$\mathbf{h} \cdot \mathbf{r} = (h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*) \cdot (\mathbf{a}x + \mathbf{b}y + \mathbf{c}z) = hx + ky + lz,$$

or in matrix notation,

$$(hkl) \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{h}^T \mathbf{r},$$

where $(hkl) = \mathbf{h}^T$ is a row vector and $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{r}$ is a column vector. $\mathbf{h}^T$ is the transpose of column vector $\mathbf{h}$ (rows and columns are interchanged). In this notation, the structure factor is given by

$$F(\mathbf{h}) = \int_{\text{cell}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{h}^T \cdot \mathbf{r})\ \mathrm{d}v_{\text{real}}. \quad (2.1.8.1)$$

The symmetry operation of a screw axis is a combination of a rotation and a translation. The rotation can be represented by the matrix $\mathbf{R}$ and the translation by the vector $\mathbf{t}$. Because of the screw-axis symmetry, $\rho(\mathbf{R}\cdot\mathbf{r}+\mathbf{t}) = \rho(\mathbf{r})$.

$F(\mathbf{h})$ can also be expressed as

$$\begin{aligned} F(\mathbf{h}) &= \int_{\text{cell}} \rho(\mathbf{R}\cdot\mathbf{r}+\mathbf{t}) \exp[2\pi i \mathbf{h}^T \cdot (\mathbf{R}\cdot\mathbf{r}+\mathbf{t})]\ \mathrm{d}v_{\text{real}} \\ &= \exp(2\pi i \mathbf{h}^T \cdot \mathbf{t}) \int_{\text{cell}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{h}^T \cdot \mathbf{R} \cdot \mathbf{r})\ \mathrm{d}v_{\text{real}}. \end{aligned} \quad (2.1.8.2)$$

Because $\mathbf{h}^T \cdot \mathbf{R} = (\mathbf{R}^T \cdot \mathbf{h})^T$, where $\mathbf{R}^T$ is the transpose of the matrix $\mathbf{R}$, equation (2.1.8.2) can be written as

$$F(\mathbf{h}) = \exp(2\pi i \mathbf{h}^T \cdot \mathbf{t}) \int_{\text{cell}} \rho(\mathbf{r}) \exp[2\pi i (\mathbf{R}^T \cdot \mathbf{h})^T \cdot \mathbf{r}]\ \mathrm{d}v_{\text{real}}. \quad (2.1.8.3)$$

By definition, the integral in equation (2.1.8.3) is $F(\mathbf{R}^T \cdot \mathbf{h})$, and, therefore

$$F(\mathbf{h}) = \exp(2\pi i \mathbf{h}^T \cdot \mathbf{t}) F(\mathbf{R}^T \cdot \mathbf{h}).$$

Conclusion: The phase angles of the two structure factors are different for $\mathbf{t} \neq 0$:

$$\alpha(\mathbf{h}) = \alpha(\mathbf{R}^T \cdot \mathbf{h}) + 2\pi \mathbf{h}^T \cdot \mathbf{t}, \quad (2.1.8.4)$$

but the structure-factor amplitudes and, therefore, the intensities are always equal:

$$I(\mathbf{h}) = I(\mathbf{R}^T \cdot \mathbf{h}) \quad \text{or} \quad I[(\mathbf{R}^T)^{-1} \cdot \mathbf{h}] = I(\mathbf{h}). \quad (2.1.8.5)$$

The matrices $(\mathbf{R}^T)^{-1}$ in reciprocal space and $\mathbf{R}$ in direct space denote rotation over the same angle. Therefore, both an $n$-fold screw axis and an $n$-fold rotation axis in the crystal correspond to an $n$-fold axis in the weighted reciprocal lattice.

However, screw axes distinguish themselves from non-screw axes by extinction of some reflections along the line in reciprocal space corresponding to the screw-axis direction. This will be shown for a twofold screw axis along the monoclinic $b$ axis.

The electron density at $\mathbf{r}$, $\rho(\mathbf{r})$, is then equal to the electron density at $\mathbf{R}\cdot\mathbf{r}+\mathbf{t}$, where $\mathbf{R}\cdot\mathbf{r}$ is a rotation that leaves the value of the $y$ coordinate unchanged. $\mathbf{t}$ is equal to $\mathbf{b}/2$.

$$F(\mathbf{h}) = \int_{\text{half the cell}} \rho(\mathbf{r})\{\exp(2\pi i \mathbf{h}^T \cdot \mathbf{r}) + \exp[2\pi i \mathbf{h}^T (\mathbf{R}\cdot\mathbf{r}+\mathbf{t})]\}\ \mathrm{d}v_{\text{real}}. \quad (2.1.8.6)$$

For the $(0k0)$ reflections, ($\mathbf{h}$ along $\mathbf{b}^*$) is $\mathbf{h} = k\mathbf{b}^*$, giving

$$\mathbf{h}^T \cdot \mathbf{r} = \mathbf{h}^T \cdot \mathbf{R} \cdot \mathbf{r} = 0 + ky + 0 \text{ and } \mathbf{h}^T \cdot \mathbf{t} = k/2.$$

This simplifies equation (2.1.8.6) to

$$F(0k0) = [1 + \exp(\pi i k)] \int_{\text{half the cell}} \rho(\mathbf{r}) \exp(2\pi i k y)\ \mathrm{d}v_{\text{real}}. \quad (2.1.8.7)$$

If $k$ is odd, $F(0k0) = 0$, because $1 + \exp(\pi i k) = 0$.

This type of systematic absence, due to screw components in the symmetry elements, occurs along lines in reciprocal space. Other types of absence apply to all $hkl$ reflections. They result from the centring of the unit cell (Fig. 2.1.1.4). Suppose the unit cell is centred in the $ab$ plane ($C$ centring). Consequently, the electron density at $\mathbf{r}$ is equal to the electron density at $\mathbf{r}+\mathbf{t}$, with $\mathbf{t} = \mathbf{a}/2 + \mathbf{b}/2$ and $\mathbf{h}^T \cdot \mathbf{t} = h/2 + k/2$. The structure factor can then be written as

$$F(\mathbf{h}) = \{1 + \exp[\pi i (h+k)]\} \int_{\text{half the cell}} \rho(\mathbf{r}) \exp\left(2\pi i \mathbf{h}^T \cdot \mathbf{r}\right)\ \mathrm{d}v_{\text{real}}. \quad (2.1.8.8)$$

The conclusion is that when $(h+k)$ is odd, the structure factors are zero and no diffracted intensity is observed for those reflections.

### 2.1.9. The Patterson function

In 1934, A. L. Patterson presented a method for locating the atomic positions in not too complicated molecules without knowledge of the phase angles (Patterson, 1934). The method involves the calculation of the Patterson function, $P(uvw) = P(\mathbf{u})$:

$$P(\mathbf{u}) = (1/V) \sum_{\mathbf{h}} |F(\mathbf{h})|^2 \cos(2\pi \mathbf{h} \cdot \mathbf{u}), \quad (2.1.9.1)$$

or, written as an exponential function,

$$P(\mathbf{u}) = (1/V) \sum_{\mathbf{h}} |F(\mathbf{h})|^2 \exp(2\pi \mathbf{h} \cdot \mathbf{u}). \quad (2.1.9.2)$$

Equations (2.1.9.1) and (2.1.9.2) give the same result, because in the definition of $P(\mathbf{u})$ anomalous dispersion is neglected, resulting in $|F(\mathbf{h})|^2 = |F(-\mathbf{h})|^2$. Comparison with equations (2.1.7.3) and (2.1.7.6) shows that the Patterson function $P(\mathbf{u})$ is a Fourier summation with coefficients $|F(\mathbf{h})|^2$ instead of $F(\mathbf{h}) = |F(\mathbf{h})| \exp[i\alpha(\mathbf{h})]$. The periodicity, and thus the unit cell, are the same for the electron density and the Patterson function. For the Patterson function, many authors prefer to use $\mathbf{u}$ rather than $\mathbf{r}$ as the position vector.

The fundamental advantage of Patterson's discovery is that, in contrast to the calculation of $\rho(\mathbf{r})$, no phase information is needed for calculating $P(\mathbf{u})$.

The Patterson map can be obtained directly after the intensities of the reflections have been measured and corrected. However, what kind of information does it provide? This can be understood from an alternative expression for the Patterson function:

$$P(\mathbf{u}) = \int_{\mathbf{r}} \rho(\mathbf{r}) \rho(\mathbf{r}+\mathbf{u})\ \mathrm{d}v_{\text{real}}. \quad (2.1.9.3)$$

Equation (2.1.9.3) leads to the same result as equation (2.1.9.1), as can be proved easily by substituting expression (2.1.7.3) for $\rho$ in the right-hand side of equation (2.1.9.3).

On the right-hand side of the equation, the electron density $\rho(\mathbf{r})$ at position $\mathbf{r}$ in the unit cell is multiplied by the electron density $\rho(\mathbf{r}+\mathbf{u})$ at position $\mathbf{r}+\mathbf{u}$; the integration is over all vectors $\mathbf{r}$ in the unit cell. The result of the integration is that the Patterson map will show peaks at the end of vectors $\mathbf{u}$ between atoms in the unit cell of the structure; all these Patterson vectors start at the origin of the Patterson cell. This can best be understood with a simple example. In Fig. 2.1.9.1, a two-dimensional unit cell is drawn containing only two atoms (1 and 2). To calculate the Patterson map, a vector $\mathbf{u}$ must be moved through this cell, and, according to equation (2.1.9.3), for every position and orientation of $\mathbf{u}$, the electron densities at the beginning and at the end of $\mathbf{u}$ must be multiplied. It is clear that this product will generally be zero unless the length and the orientation of $\mathbf{u}$ are such that it begins in atom 1 and ends in atom 2, or the other way around. If so, there is a peak in the Patterson map at the end of vector $\mathbf{u}$ and at the end of vector $-\mathbf{u}$, implying that the Patterson map is always centrosymmetric. The origin itself, where vector $\mathbf{u} = 0$, always has a high peak because

$$P(\mathbf{u}=0) = \int_{\mathbf{r}} \rho(\mathbf{r}) \rho(\mathbf{r})\ \mathrm{d}v_{\text{real}} = \sum_{\mathbf{h}} |F(\mathbf{h})|^2.$$

The origin peak is equal to the sum of the squared local electron densities. The height of each non-origin peak is proportional to the product of $\rho(\mathbf{r})$ and $\rho(\mathbf{r}+\mathbf{u})$. This is an important feature in the isomorphous replacement method for protein-structure determina-



references