

2. BASIC CRYSTALLOGRAPHY

of the unit cell: the bc plane or (100), the ac plane or (010) and the ab plane or (001).

For the bc plane or (100): indices $h = 1$, $k = 0$ and $l = 0$; $\mathbf{S}(100)$ is normal to this plane and has a length $1/d(100)$. Vector $\mathbf{S}(100)$ will be called \mathbf{a}^* .

For the ac plane or (010): indices $h = 0$, $k = 1$ and $l = 0$; $\mathbf{S}(010)$ is normal to this plane and has a length $1/d(010)$. Vector $\mathbf{S}(010)$ will be called \mathbf{b}^* .

For the ab plane or (001): indices $h = 0$, $k = 0$ and $l = 1$; $\mathbf{S}(001)$ is normal to this plane and has a length $1/d(001)$. Vector $\mathbf{S}(001)$ will be called \mathbf{c}^* .

From the definition of \mathbf{a}^* , \mathbf{b}^* and \mathbf{c}^* and the Laue conditions [equation (2.1.4.7)], the following properties of the vectors \mathbf{a}^* , \mathbf{b}^* and \mathbf{c}^* can be derived:

$$\mathbf{a}^* \cdot \mathbf{a} = \mathbf{a} \cdot \mathbf{a}^* = \mathbf{a} \cdot \mathbf{S}(100) = h = 1.$$

Similarly

$$\mathbf{b}^* \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{S}(010) = k = 1,$$

and

$$\mathbf{c}^* \cdot \mathbf{c} = \mathbf{c} \cdot \mathbf{S}(001) = l = 1.$$

However, $\mathbf{a}^* \cdot \mathbf{b} = 0$ and $\mathbf{a}^* \cdot \mathbf{c} = 0$ because \mathbf{a}^* is perpendicular to the (100) plane, which contains the b and c axes. Correspondingly, $\mathbf{b}^* \cdot \mathbf{a} = \mathbf{b}^* \cdot \mathbf{c} = 0$ and $\mathbf{c}^* \cdot \mathbf{a} = \mathbf{c}^* \cdot \mathbf{b} = 0$.

Proposition. The endpoints of the vectors $\mathbf{S}(hkl)$ form the points of a lattice constructed with the unit vectors \mathbf{a}^* , \mathbf{b}^* and \mathbf{c}^* .

Proof. Vector \mathbf{S} can be split into its coordinates along the three directions \mathbf{a}^* , \mathbf{b}^* and \mathbf{c}^* :

$$\mathbf{S} = X \cdot \mathbf{a}^* + Y \cdot \mathbf{b}^* + Z \cdot \mathbf{c}^*. \quad (2.1.5.1)$$

Our proposition is true if X , Y and Z are whole numbers and indeed they are. Multiply equation (2.1.5.1) on the left and right side by \mathbf{a} .

$$\begin{array}{cccc} \mathbf{a} \cdot \mathbf{S} & = & X \cdot \mathbf{a} \cdot \mathbf{a}^* & + & Y \cdot \mathbf{a} \cdot \mathbf{b}^* & + & Z \cdot \mathbf{a} \cdot \mathbf{c}^* \\ \vdots & & \vdots & & \vdots & & \vdots \\ = h & = & X \cdot 1 & = & 0 & = & 0. \end{array}$$

The conclusion is that $X = h$, $Y = k$ and $Z = l$, and, therefore,

$$\mathbf{S} = h \cdot \mathbf{a}^* + k \cdot \mathbf{b}^* + l \cdot \mathbf{c}^*.$$

The diffraction by a crystal [equation (2.1.4.6)] is only different from zero if the Laue conditions [equation (2.1.4.7)] are satisfied. All vectors $\mathbf{S}(hkl)$ are vectors in reciprocal space ending in reciprocal-lattice points and not in between. Each vector $\mathbf{S}(hkl)$ is normal to the set of planes (hkl) in real space and has a length $1/d(hkl)$ (Fig. 2.1.5.1).

The reciprocal-lattice concept is most useful in constructing the directions of diffraction. The procedure is as follows:

Step 1: Draw the vector \mathbf{s}_o indicating the direction of the incident beam from a point M to the origin, O , of the reciprocal lattice. As in Section 2.1.4.2, the length of \mathbf{s}_o and thus the distance MO is $1/\lambda$ (Fig. 2.1.5.2).

Step 2: Construct a sphere with radius $1/\lambda$ and centre M . The sphere is called the Ewald sphere. The scattering object is thought to be placed at M .

Step 3: Move a reciprocal-lattice point P to the surface of the sphere. Reflection occurs with $\mathbf{s} = \mathbf{MP}$ as the reflected beam, but only if the reciprocal-lattice point P is on the surface of the sphere, because only then does $\mathbf{S}(hkl) = \mathbf{s} - \mathbf{s}_o$ (Section 2.1.4.2).

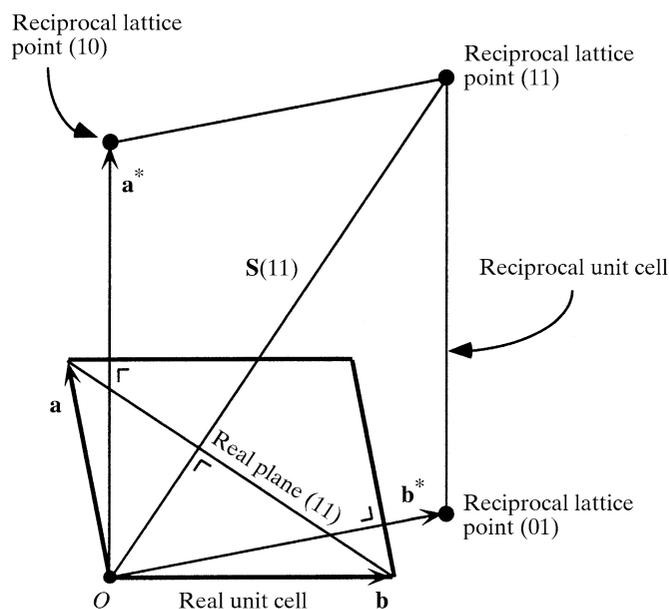


Figure 2.1.5.1

A two-dimensional real unit cell is drawn together with its reciprocal unit cell. The reciprocal-lattice points are the endpoints of the vectors $\mathbf{S}(hkl)$ [in three dimensions $\mathbf{S}(hkl)$]; for instance, vector $\mathbf{S}(11)$ starts at O and ends at reciprocal-lattice point (11). Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

Noncrystalline objects scatter differently. Their scattered waves are not restricted to reciprocal-lattice points passing through the Ewald sphere. They scatter in all directions.

2.1.6. Mosaicity and integrated reflection intensity

Crystals hardly ever have a perfect arrangement of their molecules, and crystals of macromolecules are certainly not perfect. Their crystal lattices show defects, which can sometimes be observed with an atomic force microscope or by interferometry. A schematic but useful way of looking at non-perfect crystals is through mosaicity; the crystal consists of a large number of tiny blocks. Each block is regarded as a perfect crystal, but the blocks are slightly misaligned with respect to each other. Scattering from different blocks is incoherent. Mosaicity causes a spread in the diffracted beams; when combined with the divergence of the beam from the X-ray source, this is called the effective mosaic spread. For the same crystal, effective mosaicity is smaller in a synchrotron beam with its lower divergence than in the laboratory. Protein crystals usually show a mosaic spread of 0.25–0.5°. Mosaic spread increases due to distortion of the lattice; this can happen as a result of flash freezing or radiation damage, for instance.

In Section 2.1.4.5, it was stated that the amplitude of the wave scattered by a crystal is proportional to the structure-factor amplitude $|F|$ and that its intensity is proportional to $|F|^2$. Of course, other factors also determine the intensity of the scattered beam, such as the wavelength, the intensity of the incident beam, the volume of the crystal *etc.* The intensity integrated over the entire region of the diffraction spot hkl is

$$I_{\text{int}}(hkl) = \frac{\lambda^3}{\omega V^2} \left(\frac{e^2}{mc^2} \right)^2 V_{\text{cr}} I_o LPT |F(hkl)|^2. \quad (2.1.6.1)$$

In equation (2.1.6.1), we recognize $I_o(e^2/mc^2)^2$ as part of the Thomson scattering for one electron, $I_{\text{el}} = I_o(e^2/mc^2)^2 \sin^2 \varphi$ [equations (2.1.4.1a) and (2.1.4.1b)] per unit solid angle. V_{cr} is the

2.1. INTRODUCTION TO BASIC CRYSTALLOGRAPHY

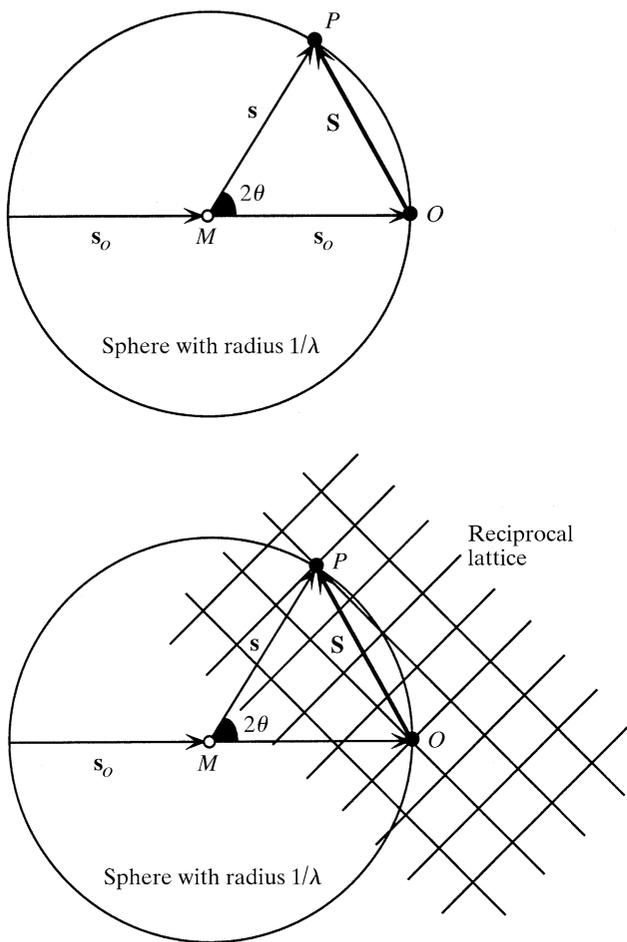


Figure 2.1.5.2

The circle is, in fact, a sphere with radius $1/\lambda$. s_o indicates the direction of the incident beam and has a length $1/\lambda$. The diffracted beam is indicated by vector s , which also has a length $1/\lambda$. Only reciprocal-lattice points on the surface of the sphere are in a reflecting position. Reproduced with permission from Dreth (1999). Copyright (1999) Springer-Verlag.

volume of the crystal and V is the volume of the unit cell. It is clear that the scattered intensity is proportional to the volume of the crystal. The term $1/V^2$ can be explained as follows. In a mosaic block, all unit cells scatter in phase. For a given volume of the individual blocks, the number of unit cells in a mosaic block, as well as the scattering amplitude, is proportional to $1/V$. The scattered intensity is then proportional to $1/V^2$. Because of the finite reflection width, scattering occurs not only for the reciprocal-lattice point when it is on the Ewald sphere, but also for a small volume around it. Since the sphere has radius $1/\lambda$, the solid angle for scattering, and thus the intensity, is proportional to $1/(1/\lambda)^2 = \lambda^2$.

However, in equation (2.1.6.1), the scattered intensity is proportional to λ^3 . The extra λ dependence is related to the time t it takes for the reciprocal-lattice 'point' to pass through the surface of the Ewald sphere. With an angular speed of rotation ω , a reciprocal-lattice point at a distance $1/d$ from the origin of the reciprocal lattice moves with a linear speed $v = (1/d)\omega$ if the rotation axis is normal to the plane containing the incident and reflected beam. For the actual passage through the surface of the Ewald sphere, the component perpendicular to the surface is needed: $v_{\perp} = (1/d)\omega \cos \theta = \omega \sin 2\theta/\lambda$. Therefore, the time t required to pass through the surface is proportional to $(1/\omega)(\lambda/\sin 2\theta)$. This introduces the extra λ term in equation (2.1.6.1) as well as the ω dependence and a $1/\sin 2\theta$ term. The latter represents the Lorentz factor L . It is a geometric correction

factor for the hkl reflections; here it is $1/\sin 2\theta$, but it is different for other data-collection geometries.

The factor P in equation (2.1.6.1) is the polarization factor. For the polarized incident beam used in deriving equation (2.1.4.1a), $P = \sin^2 \varphi$, where φ is the angle between the polarization direction of the beam and the scattering direction. It is easy to verify that $\varphi = 90^\circ - 2\theta$, where θ is the reflecting angle (Fig. 2.1.4.9). P depends on the degree of polarization of the incident beam. For a completely unpolarized beam, $P = (1 + \cos^2 2\theta)/2$.

In equation (2.1.6.1), T is the transmission factor: $T = 1 - A$, where A is the absorption factor. When X-rays travel through matter, they suffer absorption. The overall absorption follows Beer's law:

$$I = I_o \exp(-\mu d),$$

where I_o is the intensity of the incident beam, d is the path length in the material and μ is the total linear absorption coefficient. μ can be obtained as the sum of the atomic mass absorption coefficients of the elements $(\mu_m)_i$:

$$\mu = \rho \sum_i g_i (\mu_m)_i,$$

where ρ is the density of the absorbing material and g_i is the mass fraction of element i .

Atomic mass absorption coefficients $(\mu_m)_i$ for the elements are listed in Tables 4.2.4.3 (and 4.2.4.1) of *ITC* (2004) as a function of a large number of wavelengths. The absorption is wavelength-dependent and is generally much stronger for longer wavelengths. This is the result of several processes. For the X-ray wavelengths applied in crystallography, the processes are scattering and photoelectric absorption. Moreover, at the reflection position, the intensity may be reduced by extinction.

Scattering is the result of a collision between the X-ray photons and the electrons. One can distinguish two kinds of scattering: Compton scattering and Rayleigh scattering. In Compton scattering, the photons lose part of their energy in the collision process (inelastic scattering), resulting in scattered photons with a lower energy and a longer wavelength. Compton scattering contributes to the background in an X-ray diffraction experiment. In Rayleigh scattering, the photons are elastically scattered, do not lose energy, and leave the material with their wavelength unchanged. In a crystal, they interfere with each other and give rise to the Bragg reflections. Between the Bragg reflections, there is no loss of energy due to elastic scattering and the incident beam is hardly reduced. In the Bragg positions, if the reduction in intensity of the incident beam due to elastic scattering can still be neglected, the crystal is considered an ideal mosaic. For non-ideal mosaic crystals, the beam intensity is reduced by *extinction*:

- (1) The blocks are too large, and multiple reflection occurs within a block. At each reflection process, the phase angle shifts $\pi/2$ (Section 2.1.4.3.2). After two reflections, the beam travels in the same direction as the incident beam but with a phase difference of π , and this reduces the intensity.
- (2) The angular spread of the blocks is too small. The incident beam is partly reflected by blocks close to the surface and the resulting beam is the incident beam for the lower-lying blocks that are also in reflecting position.

Extinction is not a serious problem in protein X-ray crystallography.

Absorption curves as a function of the X-ray wavelength show anomalies at absorption edges. At such an edge, electrons are

2. BASIC CRYSTALLOGRAPHY

ejected from the atom or are elevated to a higher-energy bound state, the photons disappear completely and the X-ray beam is strongly absorbed. This is called *photoelectric absorption*. At an absorption edge, the frequency of the X-ray beam ν is equal to the frequency ν_K , ν_L or ν_M corresponding to the energy of the K , L or M state. According to equation (2.1.4.4), anomalous scattering is maximal at an absorption edge.

2.1.7. Calculation of electron density

In equation (2.1.4.6), the wave $W_{\text{cr}}(\mathbf{S})$ scattered by the crystal is given as the sum of the atomic contributions, as in equation (2.1.4.5) for the scattering by a unit cell. In the derivation of equation (2.1.4.5), it is assumed that the atoms are spherically symmetric (Section 2.1.4.3) and that density changes due to chemical bonding are neglected. A more exact expression for the wave scattered by a crystal, in the absence of anomalous scattering, is

$$W_{\text{cr}}(\mathbf{S}) = \int_{\text{crystal}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}) \, dV_{\text{real}}. \quad (2.1.7.1)$$

The integration is over all electrons in the crystal. $\rho(\mathbf{r})$ is the electron-density distribution in each unit cell. The operation on the electron-density distribution in equation (2.1.7.1) is called Fourier transformation, and $W_{\text{cr}}(\mathbf{S})$ is the Fourier transform of $\rho(\mathbf{r})$. It can be shown that $\rho(\mathbf{r})$ is obtained by an inverse Fourier transformation:

$$\rho(\mathbf{r}) = \int_{\mathbf{S}} W_{\text{cr}}(\mathbf{S}) \exp(-2\pi i \mathbf{r} \cdot \mathbf{S}) \, dV_{\text{reciprocal}}. \quad (2.1.7.2)$$

In contrast to $\rho(\mathbf{r})$, $W_{\text{cr}}(\mathbf{S})$ is not a continuous function but, because of the Laue conditions, it is only different from zero at the reciprocal-lattice points \mathbf{h} ($= hkl$). In equation (2.1.4.6), $W_{\text{cr}}(\mathbf{S})$ is the product of the structure factor and three delta functions. The structure factor at the reciprocal-lattice points is $F(\mathbf{h})$, and the product of the three delta functions is $1/V$, the volume of one reciprocal unit cell. Therefore, $W_{\text{cr}}(\mathbf{S})$ in equation (2.1.7.2) can be replaced by $F(\mathbf{h})/V$, and equation (2.1.7.2) itself by

$$\rho(\mathbf{r}) = (1/V) \sum_{\mathbf{h}} F(\mathbf{h}) \exp(-2\pi i \mathbf{r} \cdot \mathbf{h}). \quad (2.1.7.3)$$

If x , y and z are fractional coordinates in the unit cell, $\mathbf{r} \cdot \mathbf{S} = (\mathbf{a} \cdot x + \mathbf{b} \cdot y + \mathbf{c} \cdot z) \cdot \mathbf{S} = \mathbf{a} \cdot \mathbf{S} \cdot x + \mathbf{b} \cdot \mathbf{S} \cdot y + \mathbf{c} \cdot \mathbf{S} \cdot z = hx + ky + lz$, and an alternative expression for the electron density is

$$\rho(xyz) = (1/V) \sum_h \sum_k \sum_l F(hkl) \exp[-2\pi i(hx + ky + lz)]. \quad (2.1.7.4)$$

Instead of expressing $F(\mathbf{S})$ as a summation over the atoms [equation (2.1.4.5)], it can be expressed as an integration over the electron density in the unit cell:

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

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

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

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 $I_{\text{int}}(hkl)$ after a correction for the background and absorption, and after application of the Lorentz and polarization factor:

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

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

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

or

$$\rho(xyz) = F(000)/V + (2/V) \sum'_h \sum'_k \sum'_l |F(hkl)| \times \cos[2\pi(hx + ky + lz) - \alpha(hkl)]. \quad (2.1.7.9)$$

\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.

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