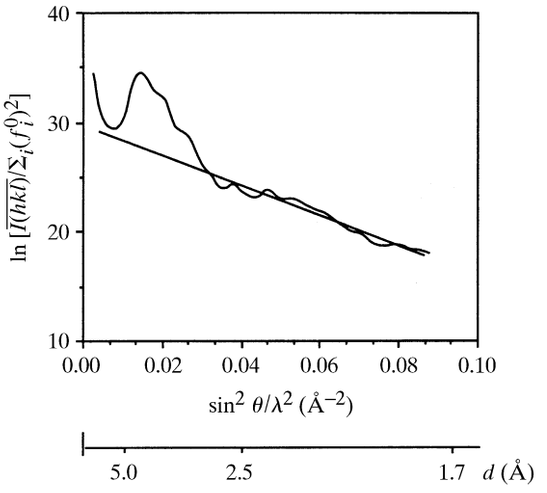

Fig. 2.1.4.10. The Wilson plot for phospholipase A_2 with data to 1.7 Å resolution. Only beyond 3 Å resolution is it possible to fit the curve to a straight line. Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

$$\overline{I(\mathbf{S})} = K\overline{I(\text{abs.},\mathbf{S})} = K\exp(-2B\sin^2\theta/\lambda^2)\sum_i (f_i^o)^2. \quad (2.1.4.12)$$

To determine K and B, equation (2.1.4.11) is written in the form

$$\ln[\overline{I(\mathbf{S})}/\sum_i (f_i^o)^2] = \ln K - 2B\sin^2\theta/\lambda^2. \quad (2.1.4.13)$$

Because f_i^o depends on $\sin\theta/\lambda$, average intensities, $\overline{I(\mathbf{S})}$, are calculated for shells of narrow $\sin\theta/\lambda$ ranges. $\ln[\overline{I(\mathbf{S})}/\sum_i (f_i^o)^2]$ is plotted against $\sin^2\theta/\lambda^2$. The result should be a straight line with slope $-2B$, intersecting the vertical axis at $\ln K$ (Fig. 2.1.4.10).

For proteins, the Wilson plot gives rather poor results because the assumption in deriving equation (2.1.4.11) that the angles, $[2\pi(\mathbf{r}_i - \mathbf{r}_j)\cdot\mathbf{S}]$, are evenly distributed over the range 0–2π for $i \neq j$ is not quite valid, especially not in the $\sin\theta/\lambda$ ranges at low resolution.

As discussed above, the average value of the structure factors, $F(\mathbf{S})$, decreases with the scattering angle because of two effects:

(1) the decrease in the atomic scattering factor f;

(2) the temperature factor.

This decrease is disturbing for statistical studies of structure-factor amplitudes. It is then an advantage to eliminate these effects by working with normalized structure factors, $E(\mathbf{S})$, defined by

$$E(\mathbf{S}) = F(\mathbf{S})\Big/\left(\sum_j f_j^2\right)^{1/2}$$

$$= F(\mathbf{S})\exp\left(B\sin^2\theta/\lambda^2\right)\Big/\left[\sum_j (f_j^o)^2\right]^{1/2}. \quad (2.1.4.14)$$

The application of equation (2.1.4.14) to $\overline{|E(\mathbf{S})|^2}$ gives

$$\overline{|E(\mathbf{S})|^2} = \overline{|F(\mathbf{S})|^2}\Big/\sum_j f_j^2 = \overline{|F(\mathbf{S})|^2}\Big/\overline{|F(\mathbf{S})|^2} = 1. \quad (2.1.4.15)$$

The average value, $\overline{|E(\mathbf{S})|^2}$, is equal to 1. The advantage of working with normalized structure factors is that the scaling is not important, because if equation (2.1.4.14) is written as

$$E(\mathbf{S}) = \frac{F(\mathbf{S})}{(\overline{|F(\mathbf{S})|^2})^{1/2}},$$

a scale factor affects numerator and denominator equally.

In practice, the normalized structure factors are derived from the observed data as follows:

$$E(\mathbf{S}) = F(\mathbf{S})\exp\left(B\sin^2\theta/\lambda^2\right)\Big/\left(\varepsilon|F(\mathbf{S})|^2\right)^{1/2}, \quad (2.1.4.16)$$

where ε is a correction factor for space-group symmetry. For general reflections it is 1, but it is greater than 1 for reflections having $\mathbf{h}$ parallel to a symmetry element. This can be understood as follows. For example, if m atoms are related by this symmetry element, $\mathbf{r}_j\cdot\mathbf{S}$ (with j from 1 to m) is the same in their contribution to the structure factor

$$F(\mathbf{h}) = \sum_{j=1}^{m} f_j\exp(2\pi i\mathbf{r}_j\cdot\mathbf{S}).$$

They act as one atom with scattering factor $m\times f$ rather than as m different atoms, each with scattering factor f. According to equation (2.1.4.11), this increases $F(\mathbf{h})$ by a factor $m^{1/2}$ on average. To make the F values of all reflections statistically comparable, $F(\mathbf{h})$ must be divided by $m^{1/2}$. For a detailed discussion, see *IT* B (2001), Chapter 2.1, by A. J. C. Wilson and U. Shmueli.

2.1.5. Reciprocal space and the Ewald sphere

A most convenient tool in X-ray crystallography is the reciprocal lattice. Unlike real or direct space, reciprocal space is imaginary. The reciprocal lattice is a superior instrument for constructing the X-ray diffraction pattern, and it will be introduced in the following way. Remember that vector $\mathbf{S}(hkl)$ is perpendicular to a reflecting plane and has a length $|\mathbf{S}(hkl)| = 2\sin\theta/\lambda = 1/d(hkl)$ (Section 2.1.4.5). This will now be applied to the boundary planes 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 a^*, b^* and 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 $\mathbf{a}$.

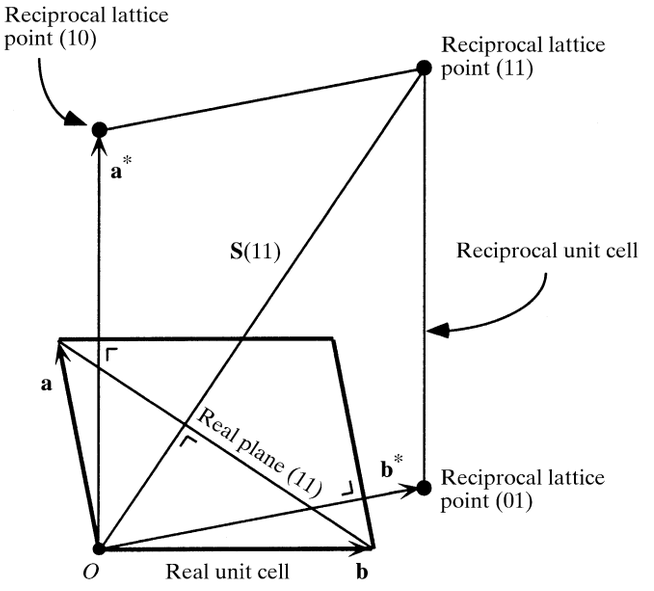

Fig. 2.1.5.1. A two-dimensional real unit cell is drawn together with its reciprocal unit cell. The reciprocal-lattice points are the endpoints of the vectors $\mathbf{S}(hk)$ [in three dimensions $\mathbf{S}(hkl)$]; for instance, vector $\mathbf{S}(11)$ starts at O and ends at reciprocal-lattice point (11). Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

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

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

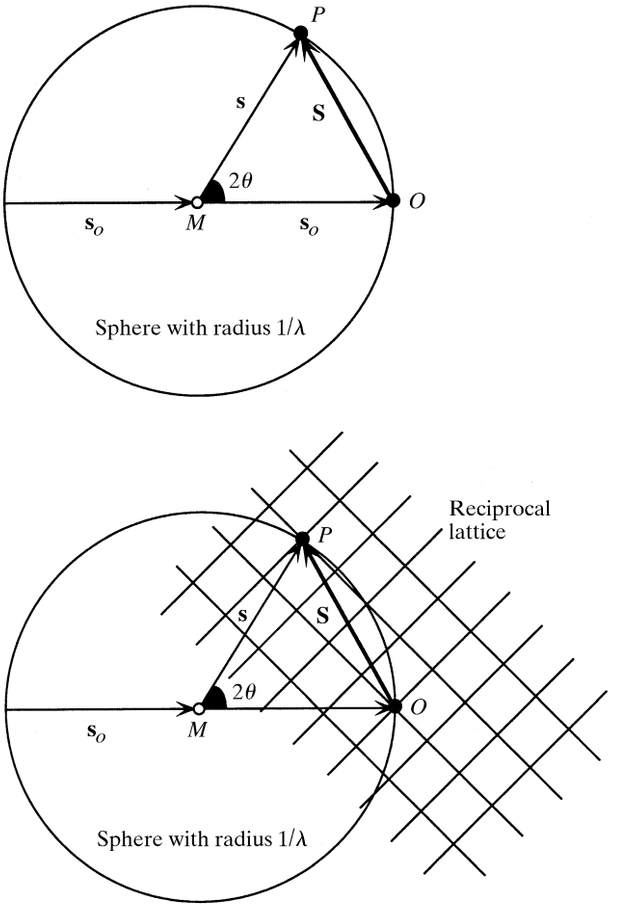

Fig. 2.1.5.2. The circle is, in fact, a sphere with radius $1/\lambda$. $\mathbf{s}_O$ indicates the direction of the incident beam and has a length $1/\lambda$. The diffracted beam is indicated by vector $\mathbf{s}$, which also has a length $1/\lambda$. Only reciprocal-lattice points on the surface of the sphere are in a reflecting position. Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

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

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

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

references