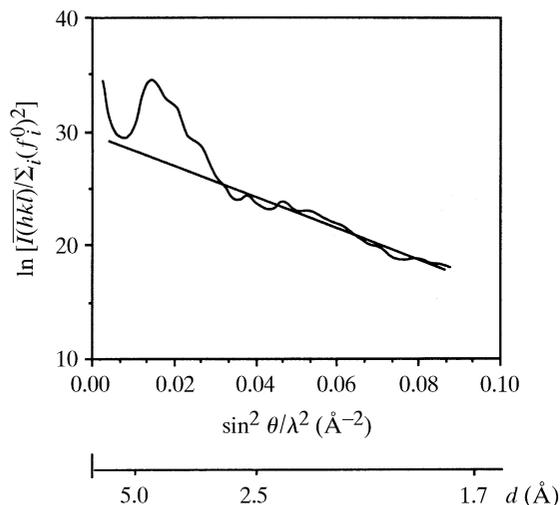


2.1. INTRODUCTION TO BASIC CRYSTALLOGRAPHY

**Figure 2.1.4.10**

The Wilson plot for phospholipase A₂ 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.

The value of $|F(\mathbf{S})|$ can be regarded as the effective number of electrons per unit cell scattering in the direction corresponding to \mathbf{S} . This is true if the values of $|F(\mathbf{S})|$ are on an absolute scale; this means that the unit of scattering is the scattering by one electron in a specific direction. The experimental values of $|F(\mathbf{S})|$ are normally on an arbitrary scale. The average value of the scattered intensity, $\overline{I(\text{abs.}, \mathbf{S})}$, on an absolute scale is $\overline{I(\text{abs.}, \mathbf{S})} = \overline{|F(\mathbf{S})|^2} = \sum_i f_i^2$, where f_i is the atomic scattering factor reduced by the temperature factor. This can be understood as follows:

$$\begin{aligned} I(\text{abs.}, \mathbf{S}) &= F(\mathbf{S}) \cdot F^*(\mathbf{S}) = |F(\mathbf{S})|^2 \\ &= \sum_i \sum_j f_i f_j \exp[2\pi i(\mathbf{r}_i - \mathbf{r}_j) \cdot \mathbf{S}]. \end{aligned} \quad (2.1.4.10)$$

For a large number of reflections, \mathbf{S} varies considerably, and assuming that the angles $[2\pi(\mathbf{r}_i - \mathbf{r}_j) \cdot \mathbf{S}]$ are evenly distributed over the range $0-2\pi$ for $i \neq j$, the average value for the terms with $i \neq j$ will be zero and only the terms with $i = j$ remain, giving

$$\overline{|F(\mathbf{S})|^2} = \overline{I(\text{abs.}, \mathbf{S})} = \sum_i f_i^2. \quad (2.1.4.11)$$

Because of the thermal vibrations

$$f_i^2 = \exp(-2B_i \sin^2 \theta / \lambda^2) (f_i^o)^2,$$

where i denotes a specific atom and f_i^o is the scattering factor for the atom i at rest.

It is sometimes necessary to transform the intensities and the structure factors from an arbitrary to an absolute scale. Wilson (1942) proposed a method for estimating the required scale factor K and, as an additional bonus, the thermal parameter B averaged over the atoms:

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

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 $0-2\pi$ 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

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

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

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

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

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

a scale factor affects numerator and denominator equally.

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

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

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* (2008), Chapter 2.1, by U. Shmueli and A. J. C. Wilson.

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