2.1. INTRODUCTION TO BASIC CRYSTALLOGRAPHY

$$F(\mathbf{h}) = \int_{\text{cell}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{h}^T \cdot \mathbf{r}) \, dv_{\text{real}}. \qquad (2.1.8.1)$$

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

$$F(\mathbf{h}) = \int_{\text{cell}} \rho(\mathbf{R} \cdot \mathbf{r} + \mathbf{t}) \exp[2\pi i \mathbf{h}^T \cdot (\mathbf{R} \cdot \mathbf{r} + \mathbf{t})] \, dv_{\text{real}}$$

$$= \exp(2\pi i \mathbf{h}^T \cdot \mathbf{t}) \int_{\text{cell}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{h}^T \cdot \mathbf{R} \cdot \mathbf{r}) \, dv_{\text{real}}. \qquad (2.1.8.2)$$

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

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

$$(2.1.8.3)$$

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 $\mathbf{t} \neq 0$:

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

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

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

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 $\mathbf{b}/2$.

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

$$(2.1.8.6)$$

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

$$F(0k0) = [1 + \exp(\pi i k)] \int_{\text{half the cell}} \rho(\mathbf{r}) \exp(2\pi i ky) \, dv_{\text{real}}.$$

$$(2.1.8.7)$$

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

$$F(\mathbf{h}) = \{1 + \exp[\pi i(h + k)]\} \int_{\text{half the cell}} \rho(\mathbf{r}) \exp(2\pi i \mathbf{h}^T \cdot \mathbf{r}) \, dv_{\text{real}}.$$

$$(2.1.8.8)$$

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, $P(uvw) = P(\mathbf{u})$:

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

or, written as an exponential function,

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

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:

$$P(\mathbf{u}) = \int_{\mathbf{r}} \rho(\mathbf{r}) \rho(\mathbf{r} + \mathbf{u}) \, dv_{\text{real}}. \qquad (2.1.9.3)$$

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 ρ 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

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

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-

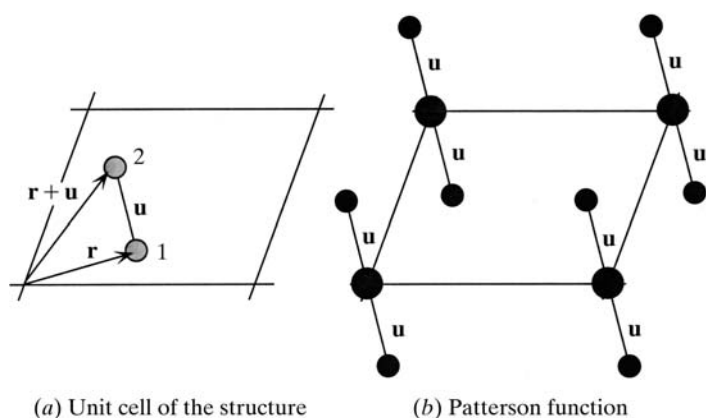*(a)* Unit cell of the structure   *(b)* Patterson function

Fig. 2.1.9.1. (*a*) A two-dimensional unit cell with two atoms. (*b*) The corresponding Patterson function. Reproduced with permission from Drenth (1999). Copyright (1999) Springer-Verlag.

tion, in which the heavy-atom positions are derived from a difference Patterson calculated with coefficients $(|F_{PH}| - |F_P|)^2$, where $|F_{PH}|$ is the structure-factor amplitude of the heavy-atom derivative and $|F_P|$ is that of the native protein (see Part 12). The vectors between the heavy atoms are the most prominent features in such a map.

The number of peaks in a Patterson map increases much faster than the number of atoms. For n atoms in the real unit cell, there are n^2 Patterson peaks, n of them superimposed at the origin, and $n \times (n - 1)$ elsewhere in the Patterson cell. Because the atomic electron densities cover a certain region and the width of a Patterson peak at $\mathbf{u}$ is roughly the sum of the widths of the atoms connected by $\mathbf{u}$, overlap of peaks is a real problem in the interpretation of a Patterson map. It can almost only be done for unit cells with a restricted number of atoms unless some extra information is available. For crystals of macromolecules, it is certainly impossible to derive the structure from an interpretation of the Patterson map.

The situation can be improved through sharpening the Patterson peaks by simulating the atoms as point scatterers. This can be achieved by replacing the $|F(\mathbf{h})|^2$ values with modified intensities which, on average, do not decrease with $\sin\theta/\lambda$. For instance, suitable intensities for this purpose are the squared normalized structure-factor amplitudes $|E(\mathbf{h})|^2$ (Section 2.1.4.6), the average of which is 1 at all $\sin\theta/\lambda$. A disadvantage of sharpening to point peaks is the occurrence of diffraction ripples around the sharp peaks, induced by truncation of the Fourier series in equation (2.1.9.1). Therefore, modified intensities corresponding to less sharpened peaks are sometimes used [*IT* B (2001), Chapter 2.3, pp. 236–237]. Diffraction ripples that seriously disturb the Patterson map are generated by the high origin peak, and, particularly for sharpened maps, it is advisable to remove this peak. This implies

that $P(\mathbf{u} = 0) = 0$ [equation (2.1.9.1)]. It is easy to verify that this requires coefficients $[|F(\mathbf{h})|^2 - \langle|F(\mathbf{h})|^2\rangle]$ for the $|F(\mathbf{h})|^2$ map and $[|E(\mathbf{h})|^2 - 1]$ for the $|E(\mathbf{h})|^2$ map. Note that the term for $\mathbf{h} = 0$ is omitted and that the average of $|F(\mathbf{h})|^2$ must be taken for the appropriate $\sin\theta/\lambda$ region.

The symmetry in a Patterson map is related to the symmetry in the electron-density map, but it is not necessarily the same. For instance, screw axes in the real cell become non-screw axes in the Patterson cell, because all interatomic vectors start at the origin. It is possible, however, to distinguish between screw axes and non-screw axes by the concentration of peaks in the Patterson map. For instance, the consequence of a twofold symmetry axis along $\mathbf{b}$ is the presence of a large number of peaks in the ($u0w$) plane of the Patterson map. For a screw axis with translation $\frac{1}{2}$ along $\mathbf{b}$, the peaks lie in the ($u\frac{1}{2}w$) plane. Such planes are called Harker planes (Harker, 1936). Peaks in Harker planes usually form the start of the interpretation of a Patterson map. Harker lines result from mirror planes, which do not occur in macromolecular crystal structures of biological origin.

Despite the improvements that can be made to the Patterson function, for structures containing atoms of nearly equal weight its complete interpretation can only be achieved for a restricted number of atoms per cell unless some extra information is available. Nowadays, most structure determinations of small compounds are based on direct methods for phase determination. However, these may fail for structures showing strong regularity. In these cases, Patterson interpretation is used as an alternative tool, sometimes in combination with direct methods. It is interesting to see that the value of the Patterson function has shifted from the small-compound field to macromolecular crystallography, where it plays an extremely useful role:

(1) in the isomorphous replacement method, the positions of the very limited number of heavy atoms attached to the macromolecule can be derived from a difference Patterson map, as mentioned earlier in this section;

(2) anomalous scatterers can be located by calculating a Patterson map with coefficients $[|F_{PH}(\mathbf{h})| - |F_{PH}(-\mathbf{h})|]^2$, in which $|F_{PH}(\mathbf{h})|$ is the structure-factor amplitude of the protein containing the anomalous scatterer;

(3) molecular replacement is based on the property that the Patterson map is a map of vectors between atoms in the real structure, combined with the fact that such a vector map is (apart from a rotation) similar for two homologous structures: the unknown and a known model structure.

Acknowledgements

References

Burzlaff, H. & Zimmermann, H. (1995). *Bravais lattices and other classifications*. In *International tables for crystallography*, Vol. A. *Space-group symmetry*, edited by Th. Hahn, pp. 739–741. Dordrecht: Kluwer Academic Publishers.

Drenth, J. (1999). *Principles of protein X-ray crystallography*. New York: Springer-Verlag.

Haas, C. & Drenth, J. (1995). *The interaction energy between two protein molecules related to physical properties of their solution and their crystals and implications for crystal growth. J. Cryst. Growth*, **154**, 126–135.

Harker, D. (1936). *The application of the three-dimensional Patterson method and the crystal structures of proustite, Ag_3AsS_3, and pyrargyrite, Ag_3SbS_3. J. Chem. Phys.* **4**, 381–390.

Heitler, W. G. (1966). *The quantum theory of radiation*, 3rd ed. Oxford University Press.

Hönl, H. (1933). *Atomfaktor für Röntgenstrahlen als Problem der Dispersionstheorie (K-Schale). Ann. Phys.* **18**, 625–655.

International Tables for Crystallography (1995). Vol. A. *Space-group symmetry*, edited by Th. Hahn. Dordrecht: Kluwer Academic Publishers.

International Tables for Crystallography (1999). Vol. C. *Mathematical, physical and chemical tables*, edited by A. J. C. Wilson & E. Prince. Dordrecht: Kluwer Academic Publishers.

International Tables for Crystallography (2001). Vol. B. *Reciprocal space*, edited by U. Shmueli. Dordrecht: Kluwer Academic Publishers.