

### 33. On the structure of a biological crystal determined by its Patterson function

*D. B. Litvin*

Crystallographers have been using X-rays to investigate the structure of biologically important macromolecules for over forty years.<sup>1</sup> One type of these is the so-called spheroidal or globular macromolecules, such as myoglobin and hemoglobin. Most globular macromolecules can be crystallized. In forming a crystal these macromolecules are not to any large extent distorted, there are in general only a few molecules in the unit cell of the crystal, and the identity of each molecule is preserved. The term 'biological crystal' in the title refers to such a crystal.

The object of investigating the structure of such biological crystals is to determine the structure of the macromolecules. To determine the structure of the crystal one attempts to calculate the electron density  $\rho(\tau)$  of the crystal *via* the Fourier expansion

$$\rho(\tau) = (1/V) \sum_k F(k) \exp(-2\pi i k \cdot \tau)$$

where the Fourier coefficients  $F(k) = |F(k)| \exp(i\alpha(k))$  are called the complex structure factors,  $\alpha(k)$  the phase of the structure factor, and  $V$  is the volume of the unit cell of the crystal.

To investigate the structure of these biological macromolecules using only the magnitudes of the structure factors, the so-called 'molecular replacement' method has been developed.<sup>2</sup> This method consists of three steps.

- 1) The rotation problem: determine the point group of the molecules in the crystal, and their orientation with respect to the translational symmetry vectors of the crystal.
- 2) The translation problem: determine the vectors between the molecular positions in the crystal. These vectors are called translation vectors.
- 3) The phase problem: using the information obtained in the first two steps, determine the phases of the structure factors.

The solutions of the rotation and translation problems are based on the analysis of a function which can be calculated from the magnitude of the structure factors  $F(k)$ . This is the Patterson function introduced in 1934;

$$P(r) = (1/V) \sum_k |F(k)|^2 \exp(2\pi i k \cdot r).$$

The function is the self-convolution (self-correlation) of the electron density of the crystal:

$$P(r) = \int \rho(r') \rho(r' + r) dr'.$$

Because in the biological crystal the macromolecules preserve their identity, the electron density of the crystal can be written as a sum of the electron densities of the molecules, and the Patterson function can be written as

$$P(r) = \sum_{j,k,t} P_{jkt}(r)$$

where  $j$  and  $k$  index the molecules in the unit cell of the crystal,  $t$  the translations of the crystal, and

$$P_{jkt}(r) = \int \rho_{j0}(r') \rho_{kt}(r' + r) dr'$$

the convolution of the electron density of the  $j$ th molecule in the unit cell with the  $k$ th molecule in the  $t$ th unit cell. The convolution function  $P_{jkt}(r)$ , in general, is called a cross-Patterson function, and in the special case where  $j = k$  and  $t = 0$ ,  $P_{jjo}(r)$  is called a self-Patterson. There are two problems for consideration:

### 1. The rotation problem

Here one wants to determine the point group and orientation of the molecules in a biological crystal. The crystals which we consider are assumed to be made up of only one kind of biological macromolecule.

To determine this, one looks for relations between the point group and orientation of the molecules in the crystal and properties of the Patterson function. Consider the electron density  $\rho_{j0}(\tau)$  of the  $j$ th molecule in the unit cell. This density is localized because of the finite dimension of the molecule, about the center of mass  $\tau_j$  of the  $j$ th molecule. The symmetry point group of the molecule is the set of all proper rotation matrices,  $P$ , such that

$$\rho_{j0}(\tau_j + Pu) = \rho_{j0}(\tau_j + u)$$

(because the biological macromolecules are made up of 'left-handed' amino acids, the symmetry point group of the molecule consists only of proper rotation matrices). We choose a coordinate system in the crystal; then the group of matrices  $P$ , which is the symmetry point group of the  $j$ th molecule, is defined with respect to the coordinate system. The orientation of the molecule is the orientation of the rotation axes of the rotations which are represented by these matrices. Another molecule in the crystal has as its symmetry point a group of matrices  $P'$  also defined with respect to this coordinate system. The two groups of matrices are in general different but equivalent, *i.e.*, they both belong to the same class of point groups and are denoted by the same symbol in, *e.g.*, international notation.

It follows that if  $P$  is the symmetry point group of the  $j$ th molecule, then the self-convolution of the electron density of the  $j$ th molecule, *i.e.*, the  $j$ th self-Patterson function

$$P_{jjo}(\tau) = \int \rho_{j0}(\tau')\rho_{j0}(\tau' + \tau)d\tau$$

is invariant under all rotations of  $P$ , *i.e.*,  $P_{jjo}(\tau) = P_{jjo}(P\tau)$ .  $P_{jjo}(\tau)$  is also invariant under inversion, and consequently  $P \times \bar{1}$  is an invariance point group of  $P_{jjo}(\tau)$ . We shall assume that  $P \times \bar{1}$  is the symmetry point group of  $P_{jjo}(\tau)$ .

As this self-Patterson is localized in a volume about the origin of the Patterson function, Rossmann and Blow,<sup>3</sup> in order to determine from the Patterson function the point group and orientation of the molecules, introduced (in 1962) the rotation function  $R(A)$ :

$$R(A) = \int_U P(\tau)P(A\tau)d\tau$$

where  $A$  is a proper rotation, and the integration is over a volume about the origin of the Patterson function. This is an overlap integral of a volume about the origin of the Patterson function with a rotation image of the same volume. Relative maxima of this rotation function, as a function of  $A$ , are called peaks of the rotation function. Obviously if  $P_{jj0}(r)$  is invariant under a rotation  $P$  then there will be a peak in the rotation function at  $A = P$ . Consequently by determining the peaks of this rotation function, one can obtain information on this point group of the molecules in the crystal. A general method is available to determine systematically the point group symmetry of the molecules from information provided by the rotation function.<sup>4</sup> To drive this method, one needs to apply only very elementary group theoretical arguments.

Consider a biological crystal consisting of identical molecules generated by a space group  $G$  from a single molecule at position  $r_1$ . Let  $T$  denote the translational subgroup of  $G$ , and  $(R_j | \tau_j) j = 1, \dots, n$ , the coset representatives of  $T$  in  $G$ . We will consider the case where  $r_1$  is a general position, *i.e.*, the  $n$  vectors  $r_j = (R_j | \tau_j)r_1$  are distinct. We then have  $n$  molecules in the unit cell, and the electron density of the molecule at  $r_j$  is related to the electron density of the molecule at  $r_1$  by

$$\rho_{j0}(r_j + u) = \rho_{10}(r_1 + R_j u)$$

where  $r_j = (R_j | \tau_j)r_1$ . That is, we have  $n$  identical molecules in different orientations in the unit cell, and their mutual orientation is determined by the rotations of the space group of the crystal. It also follows that the self-Pattersons  $P_{jj0}(r)$ ,  $j = 1, \dots, n$ , which are all localized about the origin of the Patterson, are identical, in different orientation, and their mutual orientations are also determined by the rotations of the space group of the crystal.

All peaks of the rotation function correspond to rotations which

- 1) Leave a self-Patterson  $P_{jj0}$ , for some  $j$ , invariant; or
- 2) Rotate a self-Patterson  $P_{jj0}$  into the orientation of a self-Patterson  $P_{kk0}$ , where  $j \neq k$ .

Let  $P$  denote the symmetry point group of the molecule at  $r_1$ :

- 1) The group of rotations  $\{R(jj)\} = \{R_j P R_j^{-1}\}$  is the symmetry point group of the self-Patterson  $P_{jjo}$ ; and
- 2) The set of rotations  $\{R(jk)\} = \{R_k P R_j^{-1}\}$  is the set of all rotations which rotates  $P_{jjo}$  into the orientation of  $P_{kk0}(r)$ .  
Therefore, all peaks of the rotation function  $R(A)$  correspond to all the rotations contained in the set of rotations:<sup>4</sup>

$$\{\{R(jk)\} \mid j, k = 1, \dots, n\}. \quad (1)$$

One now has a systematic method to analyze the data obtained from the rotation function.

- 1) From the rotation function calculate all rotations which correspond to all peaks.
- 2) Determine the point groups such that the set of distinct rotations in equation (1) is identical with the set of rotations corresponding to peaks of the rotation function. If  $P$  is such a point group, the conjugate point groups  $R_j P R_j^{-1}$ ,  $j = 1, \dots, n$ , where  $R_j$  is a rotation of the space group of the crystal, are also such point groups. If there is only one set of such conjugate subgroups  $R_j P R_j^{-1}$ ,  $j = 1, \dots, n$ , we say the  $P$  is the symmetry point group of the molecules. The orientation of the rotation axes of these conjugate subgroups, with respect to the translational vectors of the crystal, determine the orientation of the molecules in the crystal. If there is more than one set of such conjugate subgroups, the solution of the rotation function problem is not uniquely determined by this method. However, in practice it is impracticable to search for all peaks of the rotation function, *i.e.*, to calculate  $R(A)$  for all possible rotations  $A$ . One must then use an alternative method:
  - 1) From the rotation function calculate the rotations corresponding to some of the peaks.
  - 2) Determine the point groups such that the set of distinct rotations of equation (1) includes all those rotations found in step 1 from the rotation function.
  - 3) Determine if there are peaks of the rotation function corresponding to the additional rotations of equation (1).

**Example: satellite tobacco necrosis virus (STNV).**

A few years ago there was a debate as to whether the protein coat of this 'spherical' virus was of cubic  $O(432)$  or icosahedral  $I(532)$  point symmetry. A rotation function study was then made of a crystal containing two STNV molecules in the unit cell.<sup>5</sup> The crystal was monoclinic of space group symmetry  $C_2^3 (C2)$ , and the orientations of the two STNV molecules were related by a rotation of  $180^\circ$ . It was felt that a rotation function study of this crystal would easily determine the point group since there are 4-fold rotations contained in the cubic point group and none in the icosahedral, and 5-fold rotations in the icosahedral and none in the cubic.

A set of strong peaks was found with corresponding rotations which were exactly those proper rotations of a cube, and this was interpreted as meaning that the STNV molecules were of cubic symmetry. There were peaks corresponding to 5-fold rotations, which are characteristic of icosahedral symmetry, but these peaks were much lower than the cubic peaks.

This interpretation was immediately challenged and it was shown that all the peaks of the rotation function corresponding to 5-fold rotations could be interpreted as two molecules of icosahedral symmetry in two different orientations related by the rotation of  $180^\circ$  of the space group of the crystal.<sup>6</sup> There is general agreement that this is the correct interpretation and that STNV molecules do have icosahedral symmetry. However in this reinterpretation the stronger cubic peaks were explained away in an argument that approximated this monoclinic crystal as being cubic!

All these peaks can be explained using the above formalism and taking the point group of the STNV molecules as being icosahedral:<sup>4</sup> if one calculates the set of rotations in equation (1) taking  $j, k = 1, 2$ ,  $R_2$  the rotation of  $180^\circ$  of the space group, and  $P$  as the icosahedral symmetry point group of one of the molecules, one finds 240 rotations. These include the icosahedral rotations of both molecules, and a set of cubic rotations, exactly that set of cubic rotations determined from the rotation function! Each of these cubic rotations either leaves both molecules invariant or interchanges the two orientations, ex-

plaining the high corresponding peaks, since all other rotations either leave only one molecule invariant, or rotate one molecule into the orientation of the other. One finds also that this set, equation (1), of rotations contains additional rotations which have not yet been determined. However, even without determining peaks of the rotation function corresponding to these rotations, it does seem that the symmetry point group of the STNV molecules is icosahedral.

## 2. The translation function

Information on the point group and orientation of the molecules in a biological crystal is found, using the rotation function, by considering that volume of the Patterson about the origin of the Patterson function. Information on the translation vectors between molecules is found using a similar method, but considering other parts of the Patterson. One uses a so-called translation function  $T(x, A)$  introduced by Rossmann, Blow, Harding, and Collier,<sup>7</sup>

$$T(x, A) = \int_U P(x+r)P(x+Ar)dr.$$

The translation function, like the rotation function, is an overlap integral of a volume  $U$  of the Patterson function with a rotated image of the same volume, but unlike the rotation function, the center of the volume is now a variable, and not restricted to be at the origin of the Patterson function.

This translation function is non-zero when:

- 1) The volume  $U$  intersects with a cross-Patterson  $P_{jkt}(r)$  and the intersection is left invariant by the rotation  $A$  about  $x$ ;  
or
- 2) The volume  $U$  intersects two cross-Pattersons  $P_{jkt}(r)$  and  $P_{j'k't'}(r)$  and one is transformed into the other by the rotation  $A$  and  $X$ . The relative maxima of a translation function  $T(x, A)$  as a function of  $x$ , for constant  $A$ , are called the peaks of the translation function, and the positions of these peaks are related to the translation vectors between molecules in the crystal. The group theoretical arguments which enter

into determining the relationship between the peaks of the translation function and the translation vectors between molecules are similar to those used in predicting the peaks of the rotation function.

Necessary and sufficient conditions that the translation function  $T(x, A)$  has non-zero values associated with the transformation of  $P_{jkt}(\mathbf{r})$  into  $P_{j'k't'}(\mathbf{r})$  are:<sup>8</sup>

1) the rotation  $A$  is such that

$$P_{jkt}(\mathbf{r}_k - \mathbf{r}_j + \mathbf{t} + A^{-1}\mathbf{u}) = P_{j'k't'}(\mathbf{r}'_k - \mathbf{r}'_j + \mathbf{t}' + \mathbf{u});$$

2)  $A(\mathbf{r}_k - \mathbf{r}_j + \mathbf{t} - \mathbf{x}) = \mathbf{r}'_k - \mathbf{r}'_j + \mathbf{t}' - \mathbf{x}$ ;

3) the vector  $\frac{1}{2}\mathbf{y}$ , where  $-\mathbf{y} = \mathbf{r}_k - \mathbf{r}_j + \mathbf{t} - \mathbf{x}$  is within the volume  $U$ .

The first two conditions demand that  $P_{jkt}(\mathbf{r})$  is transformed into  $P_{j'k't'}(\mathbf{r})$  by a rotation about  $A$  about the point  $\mathbf{x}$ , and the third condition demands that the volume  $U$  of the Patterson function centered at  $\mathbf{x}$  intersects both  $P_{jkt}(\mathbf{r})$  and  $P_{j'k't'}(\mathbf{r})$ . Using group theoretical arguments like those used for the rotation function one can show that all rotations which satisfy the first condition are those denoted by  $\{A(jk, j'k')\}$

$$\{A(jk, j'k')\} = [\{R(jj')\} \cap \{R(kk')\}] + \bar{1}[\{R(jk')\} \cap \{R(kj')\}].$$

If this set of rotations is empty, the  $P_{jkt}(\mathbf{r})$  and  $P_{j'k't'}(\mathbf{r})$  are not congruent. For rotations  $A$  contained in  $\{A(jk, j'k')\}$ , the positions  $\mathbf{x}$  are calculated from the second condition.

## References

1. Bragg, L. and North, A. C. T. In *Progress in Physics, Biophysics, a reprint series*. W. A. Benjamin: New York (1969).
2. Rossmann, M. G. (ed.) *The Molecular Replacement Method*. Gordon and Breach: New York (1972).
3. Rossmann, M. G. and Blow, D. M. *Acta Cryst.* **15**, 24 (1962).
4. Litvin, D. B. *Acta Cryst.* **A31**, 407 (1975).
5. Akervall, K., Strandberg, B., Rossmann, M. G., Bengtson, V., Fridborg, K., Johannison, H., Kannan, K., Lovgren, S., Petef, G., Oberg, B., Eaker, D., Hjerten, S., Ryden, L. and Morning, I. *Cold Spring Harbor Symp. Quant. Biol.* **36**, 469 and 487 (1971).



6. Klug, A. *Cold Spring Harbor Symp. Quant. Biol.* **36**, 483 (1971).
7. Rossmann, M. G., Blow, D. M., Harding, M. M. and Collier, E. *Acta Cryst.* **17**, 338 (1964).
8. Litvin, D. B. *Acta Cryst.* **A33**, 62 (1977).

*An Offprint from*

# **Patterson and Pattersons**

**FIFTY YEARS OF THE  
PATTERSON FUNCTION**

*Proceedings of a Symposium held at  
the Institute for Cancer Research, the Fox Chase  
Cancer Center, Philadelphia, PA, USA,  
November 13-15 1984*

**EDITED BY**

**JENNY P. GLUSKER**

**BETTY K. PATTERSON**

**AND**

**MIRIAM ROSSI**

**INTERNATIONAL UNION OF CRYSTALLOGRAPHY  
OXFORD UNIVERSITY PRESS**

**1987**

**ISBN 0 19 855230 0**