

2.3. PATTERSON AND MOLECULAR REPLACEMENT TECHNIQUES

orientation by a single set of three rotation angles that determine the rotation matrix $[\mathbf{E}]$.

Assume $[\mathbf{I}_n]$ ($n = 1, \dots, N$) is the collection of noncrystallographic symmetry point-group rotation matrices in the standard orientation. Then the operation $[\mathbf{E}]$ will bring the noncrystallographic symmetry point group to a new orientation and the noncrystallographic symmetry rotation matrices in this new orientation, $[\rho_n]$, are given by (Tong & Rossmann, 1990)

$$[\rho_n] = [\mathbf{E}][\mathbf{I}_n][\mathbf{E}]^{-1}. \quad (2.3.6.8)$$

For each rotation $[\mathbf{E}]$, the ordinary self-rotation-function value (R_n) for each of the noncrystallographic symmetry rotation matrices in the new orientation ($[\rho_n]$) is calculated. The locked self-rotation-function value (R_L) for this rotation is defined as the average of the ordinary rotation-function values over the noncrystallographic symmetry elements

$$R_L([\mathbf{E}]) = \frac{1}{N-1} \sum_{n=2}^N R_n,$$

where the summation starts from 2 as it is assumed that $[\mathbf{I}_1]$ is the identity matrix.

The locked self rotation function simplifies the task of interpreting the self rotation function for the orientation of a noncrystallographic symmetry assembly. Instead of searching for $N - 1$ peaks in the ordinary self rotation function, a single peak is sought in the locked self rotation function. It must be emphasized that this rotation ($[\mathbf{E}]$) in the locked self rotation function is most often a general rotation. The locked self rotation function also reduces the noise in the rotation-function calculation by a factor of $(N - 1)^{1/2}$ due to the averaging of the ordinary rotation-function values (Tong & Rossmann, 1990).

The symmetry of the locked self rotation function is generally rather complex and an analytical solution is often impossible (Tong & Rossmann, 1990). It depends not only on the crystallographic symmetry and the noncrystallographic symmetry, but also on the definition of the standard orientation of the noncrystallographic symmetry. For example, if the standard orientation is defined such that the twofold axes are parallel to the Cartesian coordinate axes for the 222 point group, a 90° rotation around the X , Y or Z axis, or a 120° rotation around the 111 direction, does not cause a net change to the standard orientation. Such rotations will appear as symmetry in the locked self rotation function (Tong & Rossmann, 1997). In practice, the locked self rotation function can be calculated rather quickly, especially if the fast rotation function is used. A large region of rotation space can be explored in the calculation of the locked rotation function and the solutions can then be clustered based on the resulting orientation of the noncrystallographic symmetry. For example, two rotations $[\mathbf{E}_1]$ and $[\mathbf{E}_2]$ that produce the same set of noncrystallographic symmetry matrices based on (2.3.6.8) are likely to be related by the symmetry of the locked self rotation function.

A locked cross rotation function can also be defined to determine the orientation, $[\mathbf{F}]$, of the known monomer structure relative to the noncrystallographic symmetry of the molecular assembly (Navaza *et al.*, 1998; Tong, 2001a; Tong & Rossmann, 1990, 1997). With the knowledge of $[\mathbf{F}]$ and the orientation of the noncrystallographic symmetry in the crystal $[\mathbf{E}]$, which can be determined from the locked self rotation function, the orientation of all the monomers in the crystal cell is given by

$$[\rho_n] = [\mathbf{E}][\mathbf{I}_n][\mathbf{F}].$$

Therefore, $[\rho_n]$ represents the rotational relationship between the monomer search model and the monomers of the assembly in the crystal. An ordinary cross-rotation-function value R_n can be calculated for each of the rotations $[\rho_n]$, and the locked cross-rotation-function value is defined as the average

$$R_L = (1/N) \sum_n R_n.$$

Like the locked self rotation function, the locked cross rotation function can determine the orientation of all the monomers of the noncrystallographic symmetry assembly with a single rotation.

2.3.7. Translation functions

2.3.7.1. Introduction

The problem of determining the position of a noncrystallographic symmetry element in space, or the position of a molecule of known orientation in a unit cell, has been reviewed by Rossmann (1972), Colman *et al.* (1976), Karle (1976), Argos & Rossmann (1980), Harada *et al.* (1981) and Beurskens (1981). All methods depend on the prior knowledge of the object's orientation implied by the rotation matrix $[\mathbf{C}]$. The various translation functions, T , derived below, can only be computed given this information.

The general translation function can be defined as

$$T(\mathbf{S}_x, \mathbf{S}_{x'}) = \int_U \rho_1(\mathbf{x}) \cdot \rho_2(\mathbf{x}') \, d\mathbf{x},$$

where T is a six-variable function given by each of the three components that define \mathbf{S}_x and $\mathbf{S}_{x'}$. Here \mathbf{S}_x and $\mathbf{S}_{x'}$ are equivalent reference positions of the objects, whose densities are $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x}')$. The translation function searches for the optimal overlap of the two objects after they have been similarly oriented. Following the same procedure used for the rotation-function derivation, Fourier summations are substituted for $\rho_1(\mathbf{x})$ and $\rho_2(\mathbf{x}')$. It can then be shown that

$$T(\mathbf{S}_x, \mathbf{S}_{x'}) = \int_U \left\{ \frac{1}{V_h} \sum_{\mathbf{h}} |\mathbf{F}_h| \exp[i(\alpha_h - 2\pi\mathbf{h} \cdot \mathbf{x})] \right\} \times \left\{ \frac{1}{V_p} \sum_{\mathbf{p}} |\mathbf{F}_p| \exp[i(\alpha_p - 2\pi\mathbf{p} \cdot \mathbf{x}')] \right\} d\mathbf{x}.$$

Using the substitution $\mathbf{x}' = [\mathbf{C}]\mathbf{x} + \mathbf{d}$ and simplifying leads to

$$T(\mathbf{S}_x, \mathbf{S}_{x'}) = \frac{1}{V_h V_p} \sum_{\mathbf{h}} \sum_{\mathbf{p}} |\mathbf{F}_h| |\mathbf{F}_p| \times \exp[i(\alpha_h + \alpha_p - 2\pi\mathbf{p} \cdot \mathbf{d})] \times \int_U \exp\{-2\pi i(\mathbf{h} + [\mathbf{C}]^T \mathbf{p}) \cdot \mathbf{x}\} \, d\mathbf{x}.$$

The integral is the diffraction function $G_{\mathbf{h}\mathbf{p}}$ (2.3.6.4). If the integration is taken over the volume U , centred at \mathbf{S}_x and $\mathbf{S}_{x'}$, it follows that

$$T(\mathbf{S}_x, \mathbf{S}_{x'}) = \frac{2}{V_h V_p} \sum_{\mathbf{h}} \sum_{\mathbf{p}} |\mathbf{F}_h| |\mathbf{F}_p| G_{\mathbf{h}\mathbf{p}} \times \cos[\alpha_h + \alpha_p - 2\pi(\mathbf{h} \cdot \mathbf{S}_x + \mathbf{p} \cdot \mathbf{S}_{x'})]. \quad (2.3.7.1)$$

2. RECIPROCAL SPACE IN CRYSTAL-STRUCTURE DETERMINATION

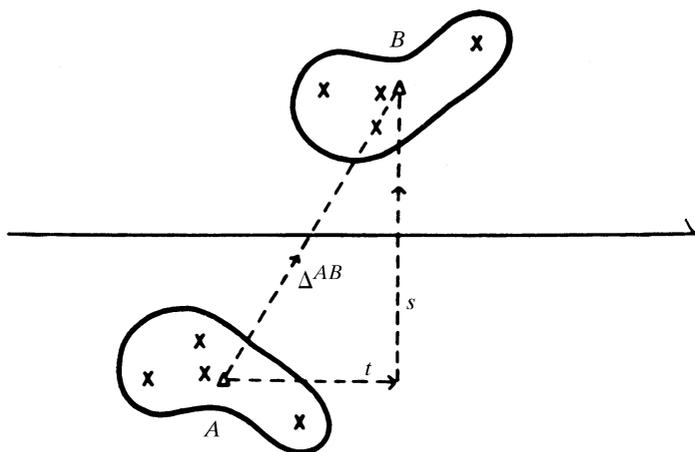


Fig. 2.3.7.1. Crosses represent atoms in a two-dimensional model structure. The triangles are the points chosen as approximate centres of molecules *A* and *B*. Δ^{AB} has components *t* and *s* parallel and perpendicular, respectively, to the screw rotation axis. [Reprinted from Rossmann *et al.* (1964).]

2.3.7.2. Position of a noncrystallographic element relating two unknown structures

The function (2.3.7.1) is quite general. For instance, the rotation function corresponds to a comparison of Patterson functions P_1 and P_2 at their origins. That is, the coefficients are F^2 , phases are zero and $\mathbf{S}_x = \mathbf{S}_{x'} = 0$. However, the determination of the translation between two objects requires the comparison of cross-vectors away from the origin.

Consider, for instance, the determination of the precise translation vector parallel to a rotation axis between two identical molecules of unknown structure. For simplicity, let the noncrystallographic axis be a dyad (Fig. 2.3.7.1). Fig. 2.3.7.2 shows the corresponding Patterson of the hypothetical point-atom structure. Opposite sets of cross-Patterson vectors in Fig. 2.3.7.2 are related by a twofold rotation and a translation equal to twice the precise vector in the original structure. A suitable translation function would then compare a Patterson at \mathbf{S} with the rotated Patterson at $-\mathbf{S}$. Hence, substituting $\mathbf{S}_x = \mathbf{S}$ and $\mathbf{S}_{x'} = -\mathbf{S}$ in (2.3.7.1),

$$T(\mathbf{S}) = \frac{2}{V^2} \sum_{\mathbf{h}} \sum_{\mathbf{p}} |\mathbf{F}_{\mathbf{h}}|^2 |\mathbf{F}_{\mathbf{p}}|^2 G_{\mathbf{h}\mathbf{p}} \cos[2\pi(\mathbf{h} - \mathbf{p}) \cdot \mathbf{S}]. \quad (2.3.7.2)$$

The opposite cross-vectors can be superimposed only if an evenfold rotation between the unknown molecules exists. The translation function (2.3.7.2) is thus applicable only in this special situation. There is no published translation method to determine the interrelation of two unknown structures in a crystallographic asymmetric unit or in two different crystal forms. However, another special situation exists if a molecular evenfold axis is parallel to a crystallographic evenfold axis. In this case, the position of the noncrystallographic symmetry element can be easily determined from the large peak in the corresponding Harker section of the Patterson.

In general, it is difficult or impossible to determine the positions of noncrystallographic axes (or their intersection at a molecular centre). However, the position of heavy atoms in isomorphous derivatives, which usually obey the noncrystallographic symmetry, can often determine this information.

2.3.7.3. Position of a known molecular structure in an unknown unit cell

The most common type of translation function occurs when looking for the position of a known molecular structure in an

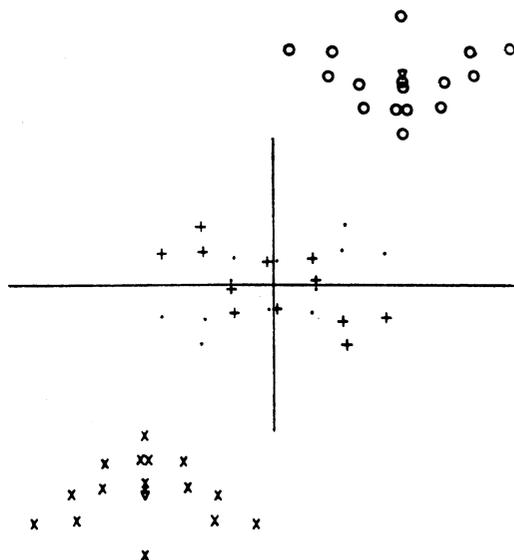


Fig. 2.3.7.2. Vectors arising from the structure in Fig. 2.3.7.1. The self-vectors of molecules *A* and *B* are represented by + and ; the cross-vectors from molecules *A* to *B* and *B* to *A* by x and o. Triangles mark the position of $+\Delta^{AB}$ and $-\Delta^{AB}$. [Reprinted from Rossmann *et al.* (1964).]

unknown crystal. For instance, if the structure of an enzyme has previously been determined by the isomorphous replacement method, then the structure of the same enzyme from another species can often be solved by molecular replacement [*e.g.* Grau *et al.* (1981)]. However, there are some severe pitfalls when, for instance, there are gross conformational changes [*e.g.* Moras *et al.* (1980)]. This type of translation function could also be useful in the interpolation of *E* maps produced by direct methods. Here there may often be confusion as a consequence of a number of molecular images related by translations (Karle, 1976; Beurskens, 1981; Egert & Sheldrick, 1985).

Tollin's (1966) *Q* function and Crowther & Blow's (1967) translation function are essentially identical (Tollin, 1969) and depend on a prior knowledge of the search molecule as well as its orientation in the unknown cell. The derivation given here, however, is somewhat more general and follows the derivation of Argos & Rossmann (1980), and should be compared with the method of Harada *et al.* (1981).

If the known molecular structure is correctly oriented into a cell (\mathbf{p}) of an unknown structure and placed at \mathbf{S} with respect to a defined origin, then a suitable translation function is

$$T(\mathbf{S}) = \sum_{\mathbf{p}} |\mathbf{F}_{\mathbf{p}, \text{obs}}|^2 |\mathbf{F}_{\mathbf{p}}(\mathbf{S})|^2. \quad (2.3.7.3)$$

This definition is preferable to one based on an *R*-factor calculation as it is more amenable to computation and is independent of a relative scale factor.

The structure factor $\mathbf{F}_{\mathbf{p}}(\mathbf{S})$ can be calculated by modifying expression (2.3.8.9) (see below). That is,

$$\mathbf{F}_{\mathbf{p}}(\mathbf{S}) = \frac{U}{V_{\mathbf{h}}} \sum_{n=1}^N \exp(2\pi i \mathbf{p} \cdot \mathbf{S}_n) \left[\sum_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} G_{\mathbf{h}\mathbf{p}_n} \exp(-2\pi i \mathbf{h} \cdot \mathbf{S}) \right],$$

where $V_{\mathbf{h}}$ is the volume of cell (\mathbf{h}) and \mathbf{S}_n is the position, in the *n*th crystallographic asymmetric unit, of cell (\mathbf{p}) corresponding to \mathbf{S} in known cell (\mathbf{h}). Let

$$A_{\mathbf{p}, n} \exp(i\gamma_n) = \sum_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} G_{\mathbf{h}\mathbf{p}_n} \exp(-2\pi i \mathbf{h} \cdot \mathbf{S}),$$

2.3. PATTERSON AND MOLECULAR REPLACEMENT TECHNIQUES

which are the coefficients of the molecular transform for the known molecule placed into the n th asymmetric unit of the \mathbf{p} cell. Thus

$$\mathbf{F}_{\mathbf{p}}(\mathbf{S}) = \frac{U}{V_{\mathbf{h}}} \sum_{n=1}^N A_{\mathbf{p},n} \exp[i(\gamma + 2\pi\mathbf{p} \cdot \mathbf{S}_n)]$$

or

$$\mathbf{F}_{\mathbf{p}}(\mathbf{S}) = \frac{U}{V_{\mathbf{h}}} \sum_{n=1}^N A_{\mathbf{p},n} \exp[i(\gamma_n + 2\pi\mathbf{p}_n \cdot \mathbf{S})],$$

where $\mathbf{p}_n = [\mathbf{C}_n^T]\mathbf{p}$ and $\mathbf{S} = \mathbf{S}_1$. Hence

$$|\mathbf{F}_{\mathbf{p}}(\mathbf{S})|^2 = \left(\frac{U}{V_{\mathbf{h}}}\right)^2 \sum_n \sum_m (A_{\mathbf{p},n} A_{\mathbf{p},m} \times \exp\{i[2\pi(\mathbf{p}_n - \mathbf{p}_m) \cdot \mathbf{S} + (\gamma_n - \gamma_m)]\}),$$

and then from (2.3.7.3)

$$T(\mathbf{S}) = \left(\frac{U}{V_{\mathbf{h}}}\right)^2 \sum_{\mathbf{p}} \sum_n \sum_m (|\mathbf{F}_{\mathbf{p},\text{obs}}|^2 A_{\mathbf{p},n} A_{\mathbf{p},m} \times \exp\{i[2\pi(\mathbf{p}_n - \mathbf{p}_m) \cdot \mathbf{S} + (\gamma_n - \gamma_m)]\}), \quad (2.3.7.4)$$

which is a Fourier summation with known coefficients $\{|\mathbf{F}_{\mathbf{p},\text{obs}}|^2 A_{\mathbf{p},n} A_{\mathbf{p},m} \times \exp[i(\gamma_n - \gamma_m)]\}$ such that $T(\mathbf{S})$ will be a maximum at the correct molecular position.

Terms with $n = m$ in expression (2.3.7.4) can be omitted as they are independent of \mathbf{S} and only contribute a constant to the value of $T(\mathbf{S})$. For terms with $n \neq m$, the indices take on special values. For instance, if the \mathbf{p} cell is monoclinic with its unique axis parallel to \mathbf{b} such that $\mathbf{p}_1 = (p, q, r)$ and $\mathbf{p}_2 = (\bar{p}, q, \bar{r})$, then $\mathbf{p}_1 - \mathbf{p}_2$ would be $(2p, 0, 2r)$. Hence, $T(\mathbf{S})$ would be a two-dimensional function consistent with the physical requirement that the translation component, parallel to the twofold monoclinic axis, is arbitrary.

Crowther & Blow (1967) show that if $\mathbf{F}_{\mathbf{M}}$ are the structure factors of a known molecule correctly oriented within the cell of the unknown structure at an arbitrary molecular origin, then (altering the notation very slightly from above)

$$T(\mathbf{S}) = \sum_{\mathbf{p}} |\mathbf{F}_{\text{obs}}(\mathbf{p})|^2 \mathbf{F}_{\mathbf{M}}(\mathbf{p}) \mathbf{F}_{\mathbf{M}}^*(\mathbf{p}[\mathbf{C}]) \exp(-2\pi i \mathbf{p} \cdot \mathbf{S}),$$

where $[\mathbf{C}]$ is a crystallographic symmetry operator relative to which the molecular origin is to be determined. This is of the same form as (2.3.7.4) but concerns the special case where the \mathbf{h} cell, into which the known molecule was placed, has the same dimensions as the \mathbf{p} cell.

The translation function as defined by (2.3.7.4) is on an arbitrary scale, which makes it difficult to compare results from different calculations. Translation functions can also be defined based on the crystallographic R factor or a correlation coefficient (CC). In particular, CCs based on reflection intensities can be evaluated by Fourier methods (Navaza & Vernoslova, 1995), although it is still computationally more expensive than the evaluation of (2.3.7.4). Alternatively, the translation function can be calculated first with (2.3.7.4), and then the R factor and CC can be calculated for the resulting top solutions.

A correct solution should also produce satisfactory packing arrangements of the molecular models in the crystal. Packing

functions have been derived that estimate the amount of overlap among the models (Harada *et al.*, 1981; Hendrickson & Ward, 1976; Rabinovich & Shakked, 1984; Simpson *et al.*, 2001), and such considerations can frequently limit the search volume very considerably. Alternatively, a simple enumeration of the actual close contacts among different molecules in the crystal (for example, Ca–Ca distances less than 3 Å) has also been found to be an effective way of eliminating those solutions that produce unreasonable crystal packing (Jogl *et al.*, 2001; Tong, 1993). If conformational differences are expected between the search atomic model and the actual structure, care must be taken when applying this packing check.

2.3.7.4. Position of a noncrystallographic symmetry element in a poorly defined electron-density map

If an initial set of poor phases, for example from an SIR derivative, are available and the rotation function has given the orientation of a noncrystallographic rotation axis, it is possible to search the electron-density map systematically to determine the translation axis position. The translation function must, therefore, measure the quality of superposition of the poor electron-density map on itself. Hence $\mathbf{S}_x = \mathbf{S}_{x'} = \mathbf{S}$ and the function (2.3.7.1) now becomes

$$T(\mathbf{S}) = \frac{2}{V_{\mathbf{h}}^2} \sum_{\mathbf{h}} \sum_{\mathbf{p}} |\mathbf{F}_{\mathbf{h}}| |\mathbf{F}_{\mathbf{p}}| G_{\mathbf{hp}} \cos[\alpha_{\mathbf{h}} + \alpha_{\mathbf{p}} - 2\pi(\mathbf{h} + \mathbf{p}) \cdot \mathbf{S}].$$

This real-space translation function has been used successfully to determine the intermolecular dyad axis for α -chymotrypsin (Blow *et al.*, 1964) and to verify the position of immunoglobulin domains (Colman & Fehlhammer, 1976).

2.3.7.5. Locked translation function

In a translation search, an atomic model with a given orientation is moved systematically through the unit cell. In such a situation, the structure-factor equation takes on the special form (Harada *et al.*, 1981; Rae, 1977; Tong, 1993)

$$\mathbf{F}_{\mathbf{h}}^c = \sum_n \mathbf{F}_{\mathbf{h},n} \exp(2\pi i \mathbf{h}^T [\mathbf{T}_n] \mathbf{S}),$$

where \mathbf{S} is the translation vector and the summation goes over the crystallographic symmetry operators. $\mathbf{F}_{\mathbf{h},n}$ is the structure factor calculated based only on the n th symmetry-related molecule,

$$\mathbf{F}_{\mathbf{h},n} = \sum_j f_j \exp\{2\pi i \mathbf{h}^T ([\mathbf{T}_n] \mathbf{x}_j^0 + \mathbf{t}_n)\},$$

where \mathbf{x}_j^0 represents the atomic position of the model at the reference position and the summation goes over all the atoms.

Noting equation (2.3.7.3), the translation function is given by

$$T(\mathbf{S}) = \sum_{\mathbf{h}} \sum_n |\mathbf{F}_{\mathbf{h}}^c|^2 |\mathbf{F}_{\mathbf{h},n}|^2 + \sum_{\mathbf{h}} \sum_n \sum_{m \neq n} |\mathbf{F}_{\mathbf{h}}^c|^2 |\mathbf{F}_{\mathbf{h},n}| |\mathbf{F}_{\mathbf{h},m}| \exp\{-2\pi i \mathbf{h}^T ([\mathbf{T}_m] - [\mathbf{T}_n]) \mathbf{S}\}, \quad (2.3.7.5)$$

where the second term is the ordinary translation function, analogous to (2.3.7.4). The first term of (2.3.7.5) depends on the orientation of the model. Maximization of this term, or its correlation coefficient equivalent, is the basis behind the Patterson-correlation refinement (Brünger, 1990; Tong, 1996b) and the direct rotation function (DeLano & Brünger, 1995). It is

2. RECIPROCAL SPACE IN CRYSTAL-STRUCTURE DETERMINATION

also related to the intensity-based domain refinement (Yeates & Rini, 1990).

In the presence of noncrystallographic symmetry, the locked self rotation function can be used to define the orientation of the noncrystallographic symmetry point group in the crystal. If an atomic model is available for the monomer but not for the entire oligomer, the locked cross rotation function can be used to determine the orientation of this monomer in the oligomer. The locked translation function can then be used to determine the position of this monomer relative to the centre of the noncrystallographic symmetry point group (Tong, 1996*b*, 2001*a*), which will produce a model for the entire oligomer. The centre of this oligomer in the crystal can be defined by a simple translation search.

With the knowledge of the orientation of one monomer of the oligomer, the first term of (2.3.7.5) is dependent on the position of this monomer relative to the centre of the noncrystallographic symmetry oligomer (Tong, 1996*b*). The atomic positions of the entire noncrystallographic symmetry oligomer in the standard orientation are given by

$$\mathbf{X}_{n,j} = [\mathbf{I}_n](\mathbf{F}\mathbf{X}_j^0 + \mathbf{V}_0),$$

where \mathbf{X}_j^0 are the atomic positions of the monomer model, centred at (0, 0, 0); \mathbf{F} is the orientation of this model in the oligomer in the standard orientation; \mathbf{V}_0 is the position of this monomer relative to the centre of the oligomer; and $[\mathbf{I}_n]$ is the n th noncrystallographic symmetry rotation matrix in the standard orientation. The atomic positions of the noncrystallographic symmetry oligomer in the crystal unit cell, centred at the origin, are given by

$$\mathbf{x}_{n,j} = [\mathbf{a}][\mathbf{E}]\mathbf{X}_{n,j} = [\mathbf{a}][\mathbf{E}][\mathbf{I}_n](\mathbf{F}\mathbf{X}_j^0 + \mathbf{V}_0),$$

where $[\mathbf{E}]$ is the orientation of the noncrystallographic symmetry in the crystal unit cell and $[\mathbf{a}]$ is the deorthogonalization matrix.

By incorporating the calculated structure factors based on this noncrystallographic symmetry oligomer into the first term of (2.3.7.5), the locked translation function is given by

$$\begin{aligned} T_L(\mathbf{V}_0) &= \sum_{\mathbf{h}} |\mathbf{F}_{\mathbf{h}}^0|^2 |\mathbf{F}_{\mathbf{h}}|^2 \\ &= \sum_{\mathbf{h}} \sum_n \sum_{m \neq n} |\mathbf{F}_{\mathbf{h},n}^0|^2 \mathbf{F}_{\mathbf{h},n} \mathbf{F}_{\mathbf{h},m}^* \exp\{-2\pi i \mathbf{h}([\theta_m] - [\theta_n])\mathbf{V}_0\}, \end{aligned} \quad (2.3.7.6)$$

where $[\theta_n] = [\mathbf{a}][\mathbf{E}][\mathbf{I}_n]$ and $\mathbf{F}_{\mathbf{h},n} = \sum_j f_j \exp(2\pi i \mathbf{h}[\theta_n][\mathbf{F}]\mathbf{X}_j^0)$. A constant term $\sum_{\mathbf{h}} \sum_n |\mathbf{F}_{\mathbf{h},n}^0|^2 |\mathbf{F}_{\mathbf{h},n}|^2$ has been omitted from this equation.

Conceptually, the locked translation function is based on the overlap of intermolecular vectors within the noncrystallographic symmetry oligomer and the observed Patterson map (Tong, 1996*b*). The equation for the locked translation function, (2.3.7.6), bears remarkable resemblance to that for the ordinary Patterson-correlation translation function, (2.3.7.5), with the interchange of the crystallographic ($[\mathbf{T}_n]$) and noncrystallographic symmetry ($[\theta_n]$) parameters.

2.3.7.6. Computer programs for rotation and translation function calculations

Several programs are currently in popular use for the calculation of rotation and translation functions. These include *AMoRe* (Navaza, 1994, 2001*a*), *BEAST* (Read, 2001*b*), *CCP4* (Collaborative Computational Project, Number 4, 1994), *CNS* (Brünger *et al.*, 1998), *COMO* (Jogl *et al.*, 2001), *EPMR*

(Kissinger *et al.*, 1999), *GLRF* (part of the *Replace* package) (Tong, 1993, 2001*a*; Tong & Rossmann, 1990, 1997), *Molrep* (Vagin & Teplyakov, 2000) and *Phaser* (Storoni *et al.*, 2004).

The correct placement of an atomic model in a crystal unit cell is generally a six-dimensional problem, with three degrees of rotational freedom and three degrees of translational freedom. Systematic examination of all six degrees of freedom at the same time is computationally expensive and cannot be used routinely (Fujinaga & Read, 1987; Rabinovich & Shakked, 1984; Sheriff *et al.*, 1999). On the other hand, directed sampling of the six degrees of freedom, driven by a stochastic or genetic algorithm (Chang & Lewis, 1997; Glykos & Kokkinidis, 2000; Kissinger *et al.*, 1999), has been successful in solving structures.

Traditionally, the calculations are divided into a rotational component (the rotation function) and a translational component (the translation function). Only a few rotation angles (for example the top few peaks of the rotation function) are manually passed to the translation function for examination (Fitzgerald, 1988). With the power of modern computers, it is now possible to perform limited six-dimensional searches, with the sampling of the rotational degrees of freedom guided by the rotation function. For example, the top peaks of the rotation function (Navaza, 1994) and their neighbours (Urzhumtsev & Podjarny, 1995) can be automatically examined by the translation function. A more general approach is to examine all rotation-function grid points with values greater than a certain threshold (Tong, 1996*a*). Such combined molecular replacement protocols have been found to be very powerful in solving new structures.

2.3.8. Molecular replacement

2.3.8.1. Using a known molecular fragment

The most straightforward application of the molecular replacement method occurs when the orientation and position of a known molecular fragment in an unknown cell have been previously determined. The simple procedure is to apply the rotation and translation operations to the known fragment. This will place it into one 'standard' asymmetric unit of the unknown cell. Then the crystal operators (assuming no further noncrystallographic operators are present in the unknown cell) are applied to generate the complete unit cell of the unknown structure. Structure factors can then be calculated from the rotated and translated known molecule into the unknown cell. The resultant model can be refined in numerous ways.

More generally, consider a molecule placed in any crystal cell (\mathbf{h}), within which coordinate positions shall be designated by \mathbf{x} . Let the corresponding structure factors be $\mathbf{F}_{\mathbf{h}}$. It is then possible to compute the structure factors $\mathbf{F}_{\mathbf{p}}$ for another cell (\mathbf{p}) into which the same molecule has been placed N times related by the crystallographic symmetry operators $[\mathbf{C}_1], \mathbf{d}_1; [\mathbf{C}_2], \mathbf{d}_2; \dots; [\mathbf{C}_N], \mathbf{d}_N$. Let the electron density at a point \mathbf{y}_1 in the first crystallographic asymmetric unit be spatially related to the point \mathbf{y}_n in the n th asymmetric unit of the \mathbf{p} crystal such that

$$\rho(\mathbf{y}_n) = \rho(\mathbf{y}_1), \quad (2.3.8.1)$$

where

$$\mathbf{y}_n = [\mathbf{C}_n]\mathbf{y}_1 + \mathbf{d}_n. \quad (2.3.8.2)$$

From the definition of a structure factor,

$$\mathbf{F}_{\mathbf{p}} = \sum_{n=1}^N \int_U \rho(\mathbf{y}_n) \exp(2\pi i \mathbf{p} \cdot \mathbf{y}_n) d\mathbf{y}_n, \quad (2.3.8.3)$$