

## Conclusion

Complications appear in many-body searches when the first molecule or molecules are not correctly placed. In our experience, orientations seem to be detectable if the search model represents more than 1/35 of the unit cell content, but more robust agreement factors are needed, especially for TFs. Statistics are lacking with respect to requirements for the quality of the search model; they would allow a correct assessment of the full capabilities of the method.

## [34] Rotation Function Calculations with GLRF Program

By LIANG TONG and MICHAEL G. ROSSMANN

### Introduction

Calculation of rotation functions represents the first step in a structure determination by the molecular replacement method.<sup>1-3</sup> Rotation functions can be used to determine the orientation of a noncrystallographic axis in a crystal or to determine the orientation of a search model (an identical structure in another crystal form or a homologous structure) relative to the molecules in a crystal. In the second step of the molecular replacement method, the rotational parameters can be used to determine the position of the noncrystallographic symmetry in the crystal,<sup>4,5</sup> or to determine the position of the search model in the crystal.<sup>5-7</sup> The information on the orientation and position of a search model leads to an initial solution for the crystal structure. Finally, the orientation and position of the noncrystallographic symmetry axes can be used to determine an initial phasing set as well as to improve and extend the phase information.<sup>3</sup>

Given two identical or homologous molecules, *A* and *B*, the rotation [ $\rho$ ] that brings molecule *A* into the same orientation as molecule *B* will also bring the self (intramolecular) vectors within molecules *A* and *B* into

<sup>1</sup> M. G. Rossmann and D. M. Blow, *Acta Cryst.* **15**, 24 (1962).

<sup>2</sup> M. G. Rossmann, ed., "The Molecular Replacement Method." Gordon and Breach, New York, 1972.

<sup>3</sup> M. G. Rossmann, *Acta Cryst.* **A46**, 73 (1990).

<sup>4</sup> D. M. Blow, M. G. Rossmann, and B. A. Jeffery, *J. Mol. Biol.* **8**, 65 (1964).

<sup>5</sup> L. Tong, *J. Appl. Cryst.* **26**, 748 (1993).

<sup>6</sup> R. A. Crowther and D. M. Blow, *Acta Cryst.* **23**, 544 (1967).

<sup>7</sup> Y. Harada, A. Lifchitz, J. Berthou, and P. Jolles, *Acta Cryst.* **A37**, 398 (1981).

overlap. Hence, an ordinary rotation function<sup>1</sup> can be defined as the overlap of one Patterson function ( $P(u)$ ) and the rotated version of the other,

$$\mathfrak{R}([C]) = \int_U P_A(u) P_B([C]u) du \quad (1)$$

where in real space

$$x_B = [C]x_A + d \quad (2)$$

with  $d$  being the translation vector, and where

$$[C] = [\alpha_B][\rho][\beta_A] \quad (3)$$

$[\alpha_B]$  is the deorthogonalization matrix,<sup>1</sup> which converts angstrom coordinates relative to a Cartesian system to fractional coordinates relative to the crystal unit cell, and  $[\beta_A]$  is the orthogonalization matrix. The rotation  $[\rho]$  is usually represented as a set of Eulerian ( $\theta_1, \theta_2, \theta_3$ ) or polar ( $\varphi, \psi, \kappa$ ) angles.<sup>1</sup> The integration volume ( $U$ ) is generally a sphere or spherical shell centered at the Patterson origin. The radius ( $R$ ) of this sphere is chosen to exclude most of the intermolecular vectors in the crystal.

Rotation functions can be classified by how Eq. (1) is evaluated (e.g., as slow<sup>1</sup> or fast,<sup>8</sup> or as reciprocal or Patterson space<sup>9</sup> rotation functions); by whether the two molecules reside in the same or different crystals (self- or cross-rotation functions); and by how the presence of noncrystallographic symmetry is handled (ordinary or locked<sup>10,11</sup> rotation functions).

Rotation functions are generally evaluated on a search grid in Eulerian or polar angles covering either the region of interest or the entire (unique) region of the rotational space. Grid points that correspond to the correct rotation angles are generally expected to give high values for Eq. (1). Rotation function values at other grid points provide an estimate for the background noise. Several computer program packages are available for the calculation of rotation functions. These include PROTEIN,<sup>12</sup> CCP4,<sup>13</sup> MERLOT,<sup>14</sup> X-Plor,<sup>15</sup> and AMoRe.<sup>16</sup> These programs support only some of the rotation functions described previously.

<sup>8</sup> R. A. Crowther, in "The Molecular Replacement Method" (M. G. Rossmann, ed.), pp. 174–178. Gordon and Breach, New York, 1972.

<sup>9</sup> R. Huber, *Acta Cryst.* **19**, 353 (1965).

<sup>10</sup> M. G. Rossmann, G. C. Ford, H. C. Watson, and L. J. Banaszak, *J. Mol. Biol.* **64**, 237 (1972).

<sup>11</sup> L. Tong and M. G. Rossmann, *Acta Cryst.* **A46**, 783 (1990).

<sup>12</sup> W. Steigemann, Ph.D. thesis, Technische Universität, München, Germany (1974).

<sup>13</sup> "CCP4, the Cooperative Computing Project in Crystallography." SERC Daresbury Laboratory, Warrington, U.K., 1986.

<sup>14</sup> P. M. D. Fitzgerald, *J. Appl. Cryst.* **21**, 273 (1988).

<sup>15</sup> A. T. Brünger, "X-Plor Manual." Yale University Press, New Haven, 1992.

<sup>16</sup> J. Navaza, *Acta Cryst.* **A50**, 157 (1994).

The program GLRF<sup>5,11</sup> supports the calculation of most of the different types of rotation functions. The program also supports several different conventions for Eulerian and polar angles as well as several different ways of orthogonalizing the crystal unit cell. The input commands to the program are keyword-based and free-formatted. A few examples of input files to GLRF are shown in the Appendix.

### Slow Rotation Function

If both Patterson functions are expressed as their Fourier transforms, Eq. (1) becomes<sup>1</sup>

$$\Re([C]) = \sum_h \sum_p F_h^2 F_p^2 G_{hp} \quad (4)$$

where

$$G_{hp} = \frac{3(\sin 2\pi HR - 2\pi HR \cos 2\pi HR)}{(2\pi HR)^3} \quad (5)$$

is the  $G$  function,<sup>1</sup> which represents the Fourier transform of a sphere of radius  $R$ .  $H$  is the length of the reciprocal space vector  $h + p[C]$ . The  $G$  function assumes the maximum value of 1 when  $h + p[C] = 0$ , but it quickly approaches and then oscillates around zero as  $h + p[C]$  deviates from zero<sup>1</sup> (Fig. 1).

The evaluation of Eq. (4) is generally time-consuming, hence the name slow rotation function. Several techniques have been developed to speed up the evaluation of the slow rotation function. The rotation function, being a fourth power function [Eq. (4)], tends to be dominated by the strong

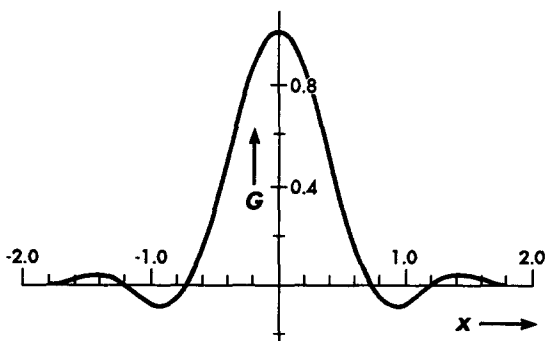


FIG. 1. Plot of the  $G$  function for values of  $HR$  in the range  $-2.0$  to  $2.0$ . Note that the  $G$  function values do not exceed  $0.1$  when  $|HR| \geq 0.72$ . [Reprinted from Rossmann and Blow.<sup>1</sup> Copyright by the International Union of Crystallography.]

reflections. Therefore, only the strong reflections of crystal *B* are needed in the calculation.<sup>17</sup> These reflections, which form the large terms in Eq. (4), are selected based on the criterion

$$I_p \geq \Delta \times \langle I_p \rangle \quad (6)$$

where  $\Delta$  is the cutoff value and the mean intensity  $\langle I_p \rangle$  is calculated for shells of equal reciprocal space volume. A cutoff value of 1.5 to 2.0, which usually selects about 10 to 20% of the observed reflections as large terms, is usually found to be adequate for calculations with protein crystal data.

Given the large terms (*p*), the summation over the reflections (*h*) of crystal *A* is limited to those terms that are within an interpolation box centered on the reciprocal lattice point closest to  $-p[C]$ , as the value of the *G* function is negligible for the other terms.<sup>1</sup> The size of this interpolation box is generally  $3 \times 3 \times 3$ . When the integration volume *U* is very small as compared to the volume of the crystal unit cell, the size of the interpolation box should be increased. However, the computation will be severely slowed down, as it is proportional to the volume of the interpolation box.

*G* function values are obtained by table lookup. The table can be set up as a function of *H* or  $H^2$ . The fastest approach, however, is to set up the table on a fine mesh<sup>5,17</sup> for all (nonintegral) values of  $h + p[C]$  in the interpolation box.

For ordinary self- and cross-rotation functions, the rotation function values vary relatively slowly as a function of the rotation angles. If a reasonably fine search grid is used relative to the resolution of the reflection data and the radius of the integration sphere (generally grid intervals of  $3-4^\circ$ ), it is unlikely for the neighbors of a grid point to have a high rotation function value if the grid point itself has a low value. The slow rotation function calculations can therefore be accelerated by ignoring such neighboring grid points.<sup>5</sup>

### Fast Rotation Function

For the fast rotation function,<sup>8</sup> the Patterson functions are expressed in terms of spherical polar coordinates (*r*,  $\theta$ ,  $\phi$ ). The angular portion ( $\theta$ ,  $\phi$ ) of the Patterson function is expanded in terms of the spherical harmonics functions. The radial portion is expanded in terms of the spherical Bessel functions. Rotating the Patterson function will only affect the angular portion of the expansion. Utilizing special properties of the spherical harmonics functions, the rotation function [Eq. (1)] becomes a Fourier transform in

<sup>17</sup> P. Tollin and M. G. Rossmann, *Acta Cryst.* **21**, 872 (1966).

$\theta_1$  and  $\theta_3$  of Eulerian angles, whereas the coefficients of this transform are dependent on  $\theta_2$ . The rotation function is then evaluated by the fast Fourier transform technique for each section of  $\theta_2$ . An improved implementation of the fast rotation function has been presented<sup>18</sup> and is available in the program AMoRe.<sup>16</sup> The fast rotation function can also be calculated in terms of polar angles,<sup>19</sup> as is available in the program MERLOT.<sup>14</sup>

The large-term approach developed for the slow rotation function can also be used for the fast rotation function. Test calculations have shown, however, that a much smaller cutoff value ( $\Delta$ ) should be used (roughly between 0.25 and 0.5) to obtain good fast rotation function results. More large-term reflections might be required for the fast than for the slow rotation function because of the poor convergence properties of the spherical Bessel expansions.<sup>18</sup> The large-term approach is better than that used in several other programs, where an absolute cutoff value on the structure factor amplitudes is applied. This could lead to under-representation of the high resolution reflection data in the calculation.

### Ordinary Self-Rotation Function

When the two molecules *A* and *B* exist in the same crystal, an ordinary self-rotation function can be used to determine the orientation and the angle of rotation of the noncrystallographic symmetry axis relating the two molecules. Polar angles are generally preferred for representing the rotation [ $\rho$ ] in ordinary self-rotation functions. For example, if the molecules are expected to be related by a twofold axis,  $\kappa$  can be fixed at  $180^\circ$  and the rotation search can be limited to the  $\varphi$  and  $\psi$  angles. Calculations in polar angles are performed with the slow rotation function in the program GLRF.

If the two molecules are related by an improper rotation, a full three-dimensional search may be necessary. In such cases, it is generally more advantageous to represent the rotation [ $\rho$ ] as a set of Eulerian angles. This simplifies the definition of the unique region of the rotational space<sup>1,20,21</sup> and makes it easier to take advantage of the fast rotation function. Once the rotation [ $\rho$ ] is known, it can be converted into a set of polar angles. This conversion can be done automatically in the GLRF program.

The ordinary self-rotation function will always achieve its maximum value when the rotation [ $\rho$ ] is the identity matrix. In the program GLRF, the ordinary self-rotation function is scaled such that this maximum value

<sup>18</sup> J. Navaza, *Acta Cryst.* **A43**, 645 (1987).

<sup>19</sup> N. Tanaka, *Acta Cryst.* **A33**, 191 (1977).

<sup>20</sup> P. Tollin, P. Main, and M. G. Rossmann, *Acta Cryst.* **20**, 404 (1966).

<sup>21</sup> S. N. Rao, J.-H. Jih, and J. A. Hartsuck, *Acta Cryst.* **A36**, 878 (1980).

is always 1000. This brings the ordinary self-rotation function onto an "absolute" scale and makes it easier to compare the results from different calculations. The rotations corresponding to the crystallographic symmetry operators should all produce rotation function values of 1000.

### Ordinary Cross-Rotation Function

If two molecules are in separate crystals, the ordinary cross-rotation function can be used to determine their relative orientation. In the case where an identical or homologous structure is used as the model to solve another structure, the known model can be placed in an artificial cell with dimensions about twice that of the largest dimension of the model to avoid the overlap of the self-Pattersons around neighboring unit cell origins. Eulerian angles are more convenient for the ordinary cross-rotation function, as the rotation function symmetry in the Eulerian angles can be easily determined in all but the cubic systems.<sup>20,21</sup> The angle limits that define the unique regions of Eulerian rotational space for noncubic Laue group symmetries of crystals *A* and *B* have been tabulated<sup>21</sup> and these limits can be automatically assigned in the GLRF program. The angle limits for the unique regions for the cubic Laue groups ( $m\bar{3}$  or  $m\bar{3}m$ ) can be specified based on the corresponding lower orthorhombic (*mmm*) or tetragonal (*4/mmm*) Laue groups. The rotation function will then contain crystallographically related regions.

The fast rotation function is particularly useful for the complete evaluation of the ordinary cross-rotation function due to its speed. With modern computers and the improvements in speed performance as mentioned previously, the evaluation of the slow rotation function is no longer extraordinarily expensive in time. The slow rotation function employs different approximations as compared to the fast rotation function. Therefore, in cases where results from the fast rotation function prove unsatisfactory, the slow rotation function may turn out to be more advantageous.

The grid interval in the rotation angles for the initial searches should be set as coarse as possible to reduce the total number of grid points that need to be calculated. The interval depends on the high resolution limit (*D*) of the reflection data and the radius of the integration sphere and is roughly given by  $2 \sin^{-1}(D/2a)$ , where *a* is the average unit cell length. A grid interval of 3 to 4° is generally sufficient for calculations with protein crystal data up to about 3 Å resolution. The grid interval along  $\theta_1$  and  $\theta_3$  for the fast rotation function is roughly 2.8° in the program GLRF. Once a peak is identified from the initial searches, fine searches should be performed around this peak, covering a smaller region and using a smaller grid interval (usually around 1°) to determine the angles more accurately.

The program GLRF can automatically perform these fine searches, with the slow rotation function, for the top peaks in the rotation function. Alternatively, Patterson correlation refinements can be used, as is available in the program X-Plor.<sup>15</sup> A different type of Patterson correlation refinement is possible as well (Tong, unpublished results, 1996).

The GLRF program also supports the calculation of a "prealigned" cross-rotation function.<sup>10,22</sup> Sometimes it is known that a direction in crystal *A* is aligned with a direction in crystal *B* (for example, the alignment of the noncrystallographic twofold axis of a dimer in two different crystal forms). In such cases, a prerotation can be used to align the two directions first. Subsequent rotation function calculations are then carried out in polar angles, varying only the angle  $\kappa$ . It should be cautioned that the head-to-head and head-to-tail alignments of the two directions will in general not be equivalent, and both alignments should therefore be tried.

### Locked Self-Rotation Function

Many macromolecular crystals contain assemblies that have a simple point group symmetry. For example, a tetramer may have the point group 222 or an icosahedral virus will have the point group 532. The ordinary self-rotation function can be used to determine the orientation of each noncrystallographic symmetry axis individually. However, if the symmetry of the point group can be assumed beforehand, the orientations of all *N* of the symmetry elements of the point group can be determined simultaneously, with the locked self-rotation function.<sup>10,11</sup> Instead of searching for *N* orientations separately, the locked self-rotation function searches for one general rotation [*E*] which will bring the noncrystallographic symmetry point group in a standard orientation (see below) to the orientation of the macromolecular assembly in the crystal. Moreover, the averaging over the *N* individual rotation function values leads to a reduction, by a factor of about  $N^{1/2}$ , in the background noise level.<sup>11</sup> The locked rotation function should therefore be more powerful than the ordinary rotation function.

A standard orientation is defined for the noncrystallographic symmetry point group. For example, a 222 point group can be defined with its three twofold axes along the *X*, *Y*, and *Z* axes of a Cartesian coordinate system. If [*I<sub>n</sub>*](*n* = 1, . . . , *N*) represents the set of rotation matrices for the symmetry operators of the point group in the standard orientation (where [*I*<sub>1</sub>] is the identity matrix), the rotation matrices [ $\rho_n$ ] of the point group after a rotation [*E*] has been applied to the standard orientation are given by

$$[\rho_n] = [E][I_n][E]^{-1} \quad (7)$$

<sup>22</sup> L. Tong, H.-K. Choi, W. Minor, and M. G. Rossmann, *Acta Cryst.* **A48**, 430 (1992).

An ordinary self-rotation function value ( $\mathfrak{R}_n$ ) can be calculated corresponding to each of the rotation matrices in the new orientation. The locked self-rotation function value ( $\mathfrak{R}_L$ ) is defined as the average of the individual values

$$\mathfrak{R}_L = \frac{1}{N-1} \sum_{n=2}^N \mathfrak{R}_n \quad (8)$$

where  $\mathfrak{R}_1$  is omitted as  $[\rho_1]$  is the identity matrix.

Locked self-rotation function calculations are more readily performed in Eulerian angles except in special cases involving a noncrystallographic symmetry axis aligned with a direction in the crystal, where polar angles may be more appropriate. The unique region of the rotational space for a locked self-rotation function depends on the symmetry of the crystal, the symmetry of the point group, and the definition of the standard orientation. For each point group, there exists a set of rotational transformation matrices ( $[H_k]$ ,  $k = 1, \dots, K$ ) that will leave the orientations of the point group symmetry elements unchanged. For example, the set of transformation matrices for point group 23 will have the symmetry 432 (Table I). The set of rotation matrices that is related to a rotation  $[E]$  by the symmetry of the locked self-rotation function is then given by  $[\beta][T_m][\alpha][E][H_k]$ , where  $[T_m]$  is the rotational component of the  $m$ th crystallographic symmetry operator. Therefore, the unique region of the rotational space for a locked self-rotation function can be much smaller than that for the corresponding ordinary self-rotation function (Table I).

For point groups 222, 4, 422, 3, 32, 6, 622, 23, and 432, the standard orientation can be defined such that the orientations of the point group symmetry elements are identical to those at the origin in the corresponding

TABLE I  
SYMMETRY OF THE POINT GROUP  
TRANSFORMATION MATRICES

Symmetry of point group	Symmetry of transformation matrices $[H]^a$
222	432
422	822 <sup>b</sup>
32	622
622	622
23	432
432	432

<sup>a</sup> See text for definition.

<sup>b</sup> The twofold axes are separated by 22.5°.



space group (for example, space group *P23* for the point group 23). The standard orientation for other point groups can be defined such that the highest-order rotation symmetry axis is along the Cartesian *Z* axis in order to simplify the definition of the unique region of the rotational space.<sup>11</sup> The program GLRF provides several different ways for defining the standard orientation of the noncrystallographic symmetry point group. The program requires only the minimum number of the point group symmetry elements, being able to generate from them, by pairwise multiplication, the entire point group. For example, one way of defining the standard orientation of point group 532 would be to provide the orientation of a fivefold and a nonorthogonal twofold axis.

The locked self-rotation function can be calculated with the slow rotation function in the program GLRF. If the rotation function value for any of the noncrystallographic symmetry elements is below a predefined cutoff at a search grid point, the program will skip to the next grid point. This will prevent the program from spending time on rotation angles that are unlikely to produce a large value for the locked self-rotation function. Alternatively, the fast rotation function can be used to calculate an ordinary self-rotation function covering the entire Eulerian space. The program GLRF can then evaluate the locked self-rotation function by linear interpolation among the computed values in the ordinary self-rotation function, resulting in a significant speedup in the calculation as compared to the evaluation by the slow rotation function. The discussion here also applies to the calculation of the locked cross-rotation functions (see as follows).

The locked self-rotation function has proven to be useful for virus crystallography.<sup>23-25</sup> It simplifies the task of determining the orientation of the icosahedral particle, and it reduces the calculation time and the amount of manual intervention. Because of the greater power of the locked self-rotation function compared to the ordinary self-rotation function, a large-term cutoff value ( $\Delta$ ) of 3 to 5, selecting only 1 to 5% of the observed reflection data, is generally sufficient for the locked self-rotation function of icosahedral virus crystal data. A higher large-term cutoff value (in the range of 2 to 3) than is possible with an ordinary rotation function should also be sufficient for locked self-rotation function calculations when the noncrystallographic point group symmetry is lower. If the fast rotation function is used in the calculation, however, the large-term cutoff value

<sup>23</sup> M. A. Oliveira, R. Zhao, W.-M. Lee, M. J. Kremer, I. Minor, R. R. Rueckert, G. D. Diana, D. C. Pevear, F. J. Dutko, M. A. McKinlay, and M. G. Rossmann, *Structure* **1**, 51 (1993).

<sup>24</sup> M. Agbandje, R. McKenna, M. G. Rossmann, M. L. Strassheim, and C. R. Parrish, *Proteins* **16**, 155 (1993).

<sup>25</sup> A. Zlotnick, B. R. McKinney, S. Munshi, J. Bibler, M. G. Rossmann, and J. E. Johnson, *Acta Cryst.* **D49**, 580 (1993).

should still be 0.25 to 0.5. The locked self-rotation function was used to confirm the orientation of the noncrystallographic symmetry elements in crystals of bacterioferritin, where the point group symmetry of the assembly is 432.<sup>26</sup>

The locked self-rotation function should also be applicable to cases where the symmetry of the macromolecular assembly does not obey that of a point group (improper rotational symmetry). Such applications will be limited, however, as it is then difficult to define the standard orientation. Moreover, as the noncrystallographic symmetry is no longer a closed point group, the unique region of the rotational space will be much larger.

### Locked Cross-Rotation Function

The presence of noncrystallographic symmetry can also be utilized in cross-rotation function calculations, when only the monomer of the macromolecular assembly is used as the search model. For a rotation  $[F]$  that brings the monomeric search model into the same orientation as one of the molecules of the assembly, there will be a set of other rotations  $[\rho_n]$  ( $n = 1, \dots, N$ ) that will bring the model into the same orientations as the other molecules of the assembly. Assuming that  $[J_n]$  represents the set of noncrystallographic symmetry rotation matrices for the assembly in the crystal, which can be determined from either ordinary or locked self-rotation functions, the set of rotations is given by

$$[\rho_n] = [J_n][F] \quad (9)$$

Therefore, a locked cross-rotation function value<sup>11</sup> can be defined as the average of the individual ordinary cross-rotation function values,

$$\mathfrak{R}_L = \frac{1}{N} \sum_n \mathfrak{R}_n \quad (10)$$

The locked cross-rotation function, like the locked self-rotation function, is best calculated in Eulerian angles. The unique region of rotational space for a locked cross-rotation function as defined above depends not only on the symmetry of the macromolecular assembly, but also on the actual orientation of the assembly in the crystal, making it rather difficult to define the angle limits.

The following formulation of the locked cross-rotation function is preferred as it can simplify the definition of the unique region of rotational space. Assume that  $[I_n]$  is the set of noncrystallographic symmetry rotation

<sup>26</sup> F. Frolow, A. J. Kalb (Gilboa), and J. Yariv, *Acta Cryst.* **D49**, 597 (1993).

matrices in the standard orientation and that  $[E]$  is the rotation that brings the standard orientation to that of the assembly in the crystal. For a rotation  $[F]$  that brings the monomeric search model into the same orientation as one of the molecules of the assembly in the standard orientation, the set of rotations relating the search model to the assembly in the crystal is given by

$$[\rho_n] = [E][I_n][F] \quad (11)$$

With this definition, the unique region of the rotational space for the locked cross-rotation function depends on the symmetry of the point group and the definition of the standard orientation and therefore can be determined more easily. As discussed for the locked self-rotation function, the standard orientation for simple point groups can be defined such that the rotation axes are oriented to be coincident with those of the corresponding space groups. The angle limits for the unique region can then be obtained from the table,<sup>21</sup> using  $P\bar{1}$  as the Laue group for the rotated Patterson map and the Laue group corresponding to the point group for the stationary Patterson map. It should be noted that the unique region is independent of the space group symmetry of the crystal, as the set of rotation matrices  $[J_n]$  or the rotation  $[E]$  applies only to the macromolecular assembly in one crystallographic asymmetric unit. Considering all the crystallographically related assemblies at the same time will eliminate this independence, but it will not reduce the total number of calculations that is necessary. A locked cross-rotation function may therefore require the coverage of a larger rotational space than the corresponding ordinary cross-rotation function.

### Presentation of Rotation Functions

The program GLRF can perform a peak search to identify those angles that give rise to high rotation function values. It lists two sets of angles for each peak. The first set corresponds to the actual angles that are used in the rotation function calculation. The nature of the second set of angles can be defined by the user. For example, if a three-dimensional ordinary self-rotation function calculation is carried out in Eulerian angles, the second set of angles in the peak listing can be specified to be polar angles so that the orientation and the angle of rotation of the noncrystallographic symmetry axis can be identified.

In addition, the program GLRF can produce contour plots of the rotation function values. Calculations in Eulerian angles are plotted in a Cartesian system with  $\theta_1$ ,  $\theta_2$ , and  $\theta_3$  along the three coordinate axes. Calcula-

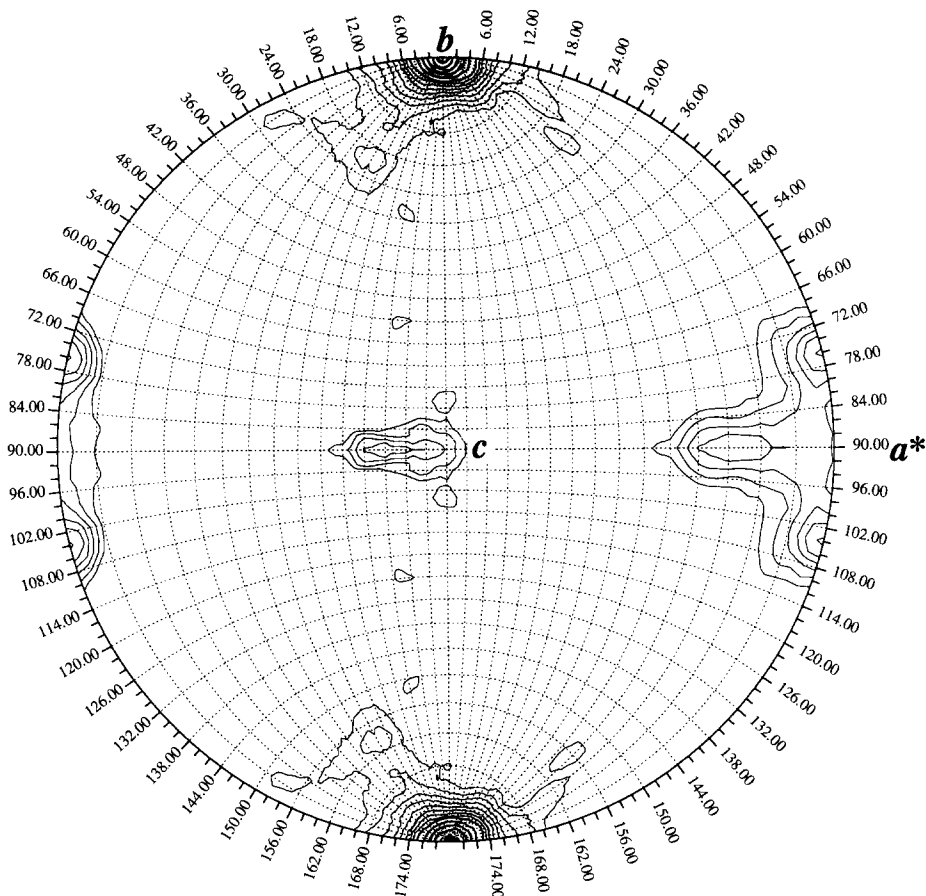


FIG. 2. A stereographic projection of the  $\kappa = 180^\circ$  section for the self-rotation function of Sindbis virus capsid protein type 3 crystal data. A total of 3863 reflections between 10.0 and 3.5 Å resolution were used in the calculation of Patterson map A. The large-term cutoff was 1.5, selecting 839 reflections in the calculation of Patterson map B. The radius of integration was 20 Å. [Reprinted from Tong *et al.*<sup>22</sup> Copyright by the International Union of Crystallography.]

tions in polar angles are plotted as stereographic projections for sections of constant  $\kappa$  (Fig. 2). While the peak listing provides information only for the heights of the peaks, the contour plots will also show the shapes of the peaks, which often can be more informative. For example, sharp and well-defined contours may suggest a high degree of similarity between the molecules, and contours with an elongated shape may indicate the presence of two closely positioned peaks.

### Electron Density Correlation Functions

The rotation function is based on the correlation of one Patterson function with the rotated version of another and can be calculated with the knowledge of the structure factor amplitudes. When phase information is available for the reflections, a function can be similarly defined based on the correlation of the electron density functions of two identical or homologous molecules<sup>1,4,5,27</sup>:

$$\mathfrak{S}([C], s_A, s_B) = \sum_h \sum_p \mathbf{F}_h \mathbf{F}_p G_{hp} e^{-2\pi i h s_A} e^{-2\pi i p s_B} \quad (12)$$

where  $[C]$  is the rotational relationship between the two molecules, and  $s_A$  and  $s_B$  are the centers of the spheres within which the electron density functions are being compared. The function  $\mathfrak{S}$  calculates the correlation within spherical volumes and hence avoids the definition of molecular masks.

Equation (12) can be used in three ways, and these applications are all supported by the program GLRF. First, given the rotational relationship  $[C]$  and a position (for example, the center) in molecule  $A$  ( $s_A$ ), the corresponding position in molecule  $B$  ( $s_B$ ) can be determined by a Fourier transformation with  $p$  as indices. Second, once initial estimates for  $[C]$ ,  $s_A$ , and  $s_B$  are available, Eq. (12) can be used to optimize these values. An initial estimate for  $[C]$  is usually obtained from rotation function calculations. An initial estimate for  $s_A$  can usually be obtained by examining the electron density map and selecting a position that is roughly at the center of the molecular envelope. The initial estimate for  $s_B$  can then be obtained from Eq. (12) by Fourier transform. Subsequently, only values for  $[C]$  and  $s_B$  (or  $s_A$ ) need to be changed in the optimization.

If the two molecules are present in the same crystal, the projection of  $s_B - s_A$  onto the direction of the rotation axis  $[C]$  will give the translation element ( $d_{\parallel}$ ) along this axis. The average of  $s_A$  and  $s_B$  is the position of the noncrystallographic symmetry axis in the unit cell. If  $s$  is assumed to lie on this rotation axis, then  $s = [C]s + d_{\parallel}$ . By substitution into Eq. (12),

$$\mathfrak{S}([C], s, d_{\parallel}) = \sum_h \sum_p \mathbf{F}_h \mathbf{F}_p G_{hp} e^{-2\pi i p d_{\parallel}} e^{-2\pi i (h+p)s} \quad (13)$$

Therefore, the position ( $s$ ) of the noncrystallographic symmetry axis can be determined by a Fourier transformation with  $h + p$  as indices. The transform should produce a long, sausage-shaped feature that corresponds to the position of the noncrystallographic symmetry axis in the unit cell.  $[C]$  can be obtained from rotation function calculations, and  $d_{\parallel}$  is zero for

<sup>27</sup> M. G. Rossmann, D. M. Blow, M. M. Harding, and E. Collier, *Acta Cryst.* **17**, 338 (1964).

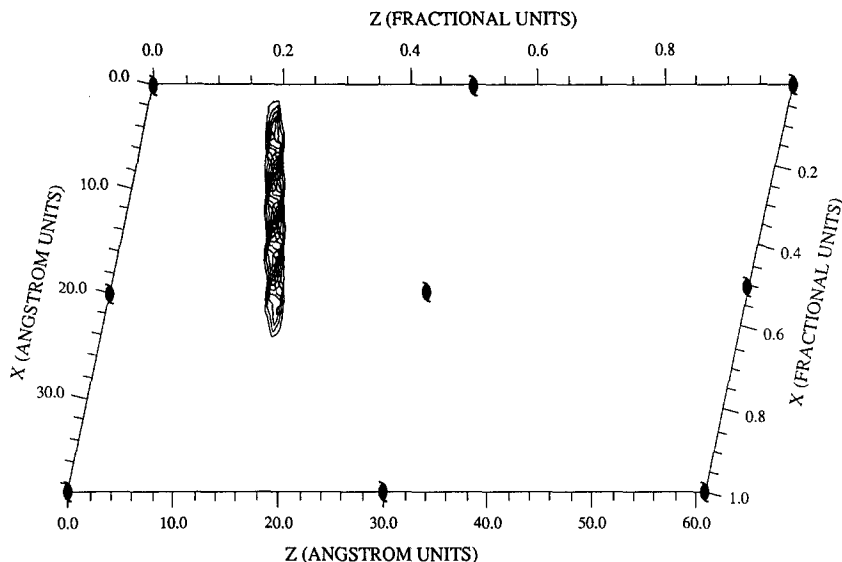


FIG. 3. Electron density correlation function [Eq. (13)] showing the position of the noncrystallographic twofold axis in Sindbis virus capsid protein type 3 crystal. Reflection data were selected between 10.0 and 5.0 Å resolution. Solvent-flattened, MIR phases were used in the calculation. The large-term cutoff was 1.5 and the radius of integration was 20 Å. Sections 9 through 13 (0.1 through 0.15 in fractional units) in  $y$  are shown superimposed in the plot. The first contour level was drawn at  $4\sigma$  and subsequent contours were drawn at intervals of  $\sigma$ , where  $\sigma$  is the rms deviation from the mean of the function. [Reprinted from Tong *et al.*<sup>22</sup> Copyright by the International Union of Crystallography.]

symmetrical aggregates with closed point groups. Equation (13) can then be used once initial phase information is available to locate the position of the noncrystallographic symmetry axis. For a noncrystallographic twofold axis,  $d_{\parallel}$  can be determined by a special Patterson correlation function,<sup>27</sup> which is also supported by the GLRF program.

Accurate values for  $[C]$  and the center ( $s$ ) of the noncrystallographic symmetry axis are prerequisites before initiating molecular replacement averaging for phase improvement and phase extension. The approach presented here avoids the tedious task of defining the molecular masks and represents a quick and simple way of obtaining accurate parameters for the noncrystallographic symmetry elements. Equations (12) and (13) have been used successfully in the structure determinations of  $\alpha$ -chymotrypsin,<sup>4</sup> Sindbis virus capsid protein,<sup>21</sup> and thiolase.<sup>28</sup> In these cases, the initial

<sup>28</sup> M. Mathieu, J. Zeelen, R. A. Pauptit, R. Erdmann, W.-H. Kunau, and R. K. Wierenga, *Structure* **2**, 797 (1994).

electron density map was of rather poor quality and the noncrystallographic twofold axis in the electron density could not be identified from visual inspections of the map. Calculations with Eq. (13), however, clearly showed the positions of the noncrystallographic two-fold axis (Fig. 3). A position was then identified to be roughly at the center of one of the monomers by examining the electron density map with the results from Eq. (13). The corresponding position in the other monomer was determined by the Fourier transform in Eq. (12). After the rotational and positional parameters were optimized with Eq. (12), the average of the two positions then gave the accurate location for the noncrystallographic twofold axis in the unit cell.

## Appendix: Examples of Input Files to Program GLRF

### A. Ordinary self rotation function calculation.

```
!  
! Comments can be introduced anywhere in the input by using  
! the special character !. They can also be introduced with  
! the COMMENT command.  
!  
! Only the first four letters of each command is needed.  
!  
title Ordinary self RF calculation  
print srf.prt  
!  
! Conventions  
!  
polar xyk  
euler zxz  
orthogonalization axabz  
!  
! Crystal A  
!  
acell 38.8 79.7 60.8 90 102.2 90  
asymmetry p21  
aobs-file native.dat  
aformat (3i4, 2f8.2)  
acutoff 2 1 0  
apower 2  
nshell 8  
origin true           ! Patterson origin removal  
!  
! Crystal B  
! As a self RF is being calculated, the program will automatically  
! copy the information for crystal A into B  
!  
cutoff 1.5           ! Large term cutoff  
!  
! Search parameters  
!  
self true  
cross false  
fast false
```

```

resolution 10 3
radius 15
rcutoff 20 ! Cut-off value for removing neighbors
boxsize 3 3 3
gevaluation 2 ! G function from fine mesh table
sangle polar
slimits 1 0 180 2 ! Search limits in phi
slimits 2 0 180 2 ! Search limits in psi
slimits 3 180 180 0 ! kappa is held at 180
peak-search 3 40
pkfitting 5 1.5 ! automatic fine searches
oangle polar xzk
mapfile srf.map
!
! Contour plot
!
cntfile srf.ps ! A contour plot in stereographic
cntlevel 300 1000 50 ! projection will be produced
cntsection 1 1
!
stop

```

## B. Ordinary cross rotation function calculation.

```

!
title Ordinary cross RF calculation
print xrf.prt
!
! Conventions
!
polar xyk
euler zyz ! The fast RF is calculated with the
orthogonalization axabz ! ZYZ convention for Eulerian angles
!
! Crystal A
!
acell 38.8 79.7 60.8 90 102.2 90
asymmetry p21
aobs-file native.dat
aformat (3i4, 2f8.2)
acutoff 2 1 0
apower 2
nshell 8
origin true ! Patterson origin removal
!
! Crystal B
!
bcell 57.0 57.0 109.8 90 90 90
bsymmetry p43212
bobs-file fobs.dat
bformat (3i4, 2f8.2)
bcutoff 2 1 0
bpower 2
cutoff 0.25 ! Large term cutoff
!
! Search parameters
!

```



```

self false
cross true
fast true
resolution 10 3
radius 15
sangle euler
slimits 2 270 270 3      ! The program will automatically assign
peak-search 4 50        ! the angle limits
pkfitting 10 1.5        ! Automatic fine searches
oangle euler zxz
section 213              ! sectioning of the map file and the
mapfile xrf.map         ! contour plot
!
! Contour plot
!
cntfile xrf.ps           ! The CNTSection command is not given.
cntlevel 500 1000 50    ! Default is to plot 5 sections
!                        ! containing the top peaks.
stop

```

### C. Locked self rotation function calculation.

```

!
title Locked self RF calculation
print lrf.prt
!
! Standard orientation
!
polar xyk
locsymmetry 0 0 1 5 vector      ! 532 icosahedral symmetry
locsymmetry -90 58.28253 180 2 polar
!
! locsymmetry 1 0 0 2 vector      ! point group 23
! locsymmetry 1 1 1 3 vector
locexpand true
!
! Conventions
!
polar xzk
euler zxz
orthogonalization axabz
!
! Crystal A
!
acell 306 361.1 299.7 90 92.91 90
asymmetry p21
aobs-file native.dat
aformat (3i4, 2f8.2)
acutoff 2 1 0
apower 2
nshell 8
origin true                  ! Patterson origin removal
!
! Crystal B
!
cutoff 5.0                   ! Large term cutoff
!

```

```
! Search parameters
!
self true
cross false
fast false           ! Calculate by slow rotation function
resolution 30 15    ! Fast RF is possible as well
radius 150
rcutoff 10
boxsize 3 3 3
gevaluation 2
sangle euler
slimits 1 0 90 6    ! the angle limits
slimits 2 0 64 6
slimits 3 0 72 6
peak-search 4 50
oangle euler zyz
section 213         ! sectioning of the map file
mapfile lrf.map
!
stop
```

## Acknowledgments

We are grateful for the comments and suggestions of many of our colleagues. The initial work on GLRF was supported by grants to MGR by the NSF and NIH.

## [35] Phased Translation Function

By GRAHAM A. BENTLEY

### Introduction

Translation functions used in molecular replacement to position a correctly oriented model in the unit cell of an unknown crystal structure can be classified into three distinct types. The first category<sup>1,2</sup> relies on comparing the Patterson function of the structure under study with that calculated from an oriented model placed at some specified position in the unit cell of the unknown molecule. When the position of the model is coincident with the unknown molecule, the interatomic vectors between symmetry-related components in one Patterson will largely coincide with those of the other Patterson and, in the ideal case, the correlation between the two

<sup>1</sup> R. A. Crowther and D. M. Blow, *Acta Cryst.* **23**, 544 (1967).

<sup>2</sup> Y. Harada, A. Lifchitz, J. Berthou, and P. Jollés, *Acta Cryst.* **A37**, 398 (1981).