Fourier transform of $t(\mathbf{r})$ is $T(\mathbf{R})$ and, using Cartesian coordinates, $T(X, Y, Z) = T(X, Y)T(Z)$. This is a simplification in the computation of the diffraction intensity $I(\xi, l)$. One can write, after omitting the weight factors, the Fourier transform $T(X, Y)$ as

$$T(X, Y) = \text{sinc } \pi\lambda X \text{ sinc } \pi y_0 Y, \quad (29)$$

and $T(l)$, $Z = l/c$, is given by

$$T(l) = \text{sinc } \pi z_0 l/c. \quad (30)$$

The Fourier transform $T(X, Y)$ in cylindrical coordinates ξ, Φ is

$$T(\xi, \Phi) = \text{sinc } \pi\lambda\xi \cos \Phi \text{ sinc } \pi y_0 \xi \sin \Phi. \quad (31)$$

The averaged transform $\mathbb{U}_m(\xi)$ for the pillar of uniform density can be computed using (31) and (26).

Discussion

A new formula for the diffracted intensity $I(\xi, l)$ for a helical array of subunits has been derived. This formula is exact. It may have computational advantages in fiber diffraction but further study is required. The present formulation allows the examination of the effects of helical disorders. The treatment of disorder in the case of the elementary helix was described in paper I.

The electron density of each subunit is defined relative to its own origin independent of the helical parameters. This straightforward application of the Fourier transform of the subunit may prove to be useful in molecular model-building studies of biological helical structures and in a variety of disorder problems relating to helical structure. A long-term aim of this work is eventually to study the dynamics of biological structures.

I thank Professor G. F. Elliott for discussions and his helpful comments on the manuscript.

References

CHAMPENEY, D. (1973). In *Fourier Transforms and their Physical Applications*. London, New York: Academic Press.
COCHRAN, W., CRICK, F. H. C. & VAND, V. (1952). *Acta Cryst.* **5**, 581-586.
FRANKLIN, R. E. & KLUG, A. (1955). *Acta Cryst.* **8**, 777-780.
HASELGROVE, J. C. (1980). *J. Muscle Res. Cell Motil.* **1**, 177-191.
JEFFREYS, H. & JEFFREYS, B. (1962). In *Methods of Mathematical Physics*. Cambridge Univ. Press.
WORTHINGTON, C. R. (1961). *J. Mol. Biol.* **3**, 618-633.
WORTHINGTON, C. R. & ELLIOTT, G. F. (1989). *Acta Cryst.* A**45**, 645-654.

Acta Cryst. (1990). A**46**, 783-792

The Locked Rotation Function

BY LIANG TONG AND MICHAEL G. ROSSMANN

Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, *USA*

(*Received* 13 *February* 1990; *accepted* 9 *May* 1990)

Abstract

It frequently occurs that a biological assembly in a crystallographic asymmetric unit has more than one noncrystallographic symmetry operator. For instance, a tetramer might have the point group 222 or a spherical virus will have the point group 532. A self-rotation function searches for the direction and angle of rotation of the individual noncrystallographic symmetry operations, while a cross-rotation function searches for the relationship of a structure in one unit cell with similar structures in another cell. The power of the rotation function can be greatly enhanced by searching for all noncrystallographic symmetry operators simultaneously. The procedure described previously [Rossmann, Ford, Watson & Banaszak (1972). *J. Mol. Biol.* **64**, 237-249] has been generalized. The increased power of this 'locked' rotation function permits a good determination of the orientation of an icosahedral virus in the presence of less than 1% of the possible diffraction data to 7 Å resolution. In addition, the peak-to-noise ratio is substantially improved.

Introduction

The rotation function (Rossmann & Blow, 1962; Hoppe, 1957) determines the direction and angle of rotation of noncrystallographic symmetry operators in a crystal lattice. The latter is any operator that is valid within a local volume (as opposed to infinite volume) of the crystal lattice. In many cases, biological assemblies contain point groups of fairly high symmetry. When crystals are available, the rotation function can be used to determine the point group

0108-7673/90/100783-10$03.00

of the assembly. For instance, glyceraldehyde 3-phosphate dehydrogenase (GAPDH) was shown to be a tetramer with 222 symmetry (Rossmann, Ford, Watson & Banaszak, 1972). More recently, the rotation function has been used routinely to find the orientation of icosahedral viruses (point group 532) in their crystal lattice, as well as for many other applications.

If the point group of an assembly is known or can be assumed, then it is often possible to determine the orientation of the particle when the rotation-function peaks are otherwise hidden in noise. One example was in the analysis of GAPDH where the rotation function had numerous peaks in the $\kappa = 180°$ plane (a search for twofold axes). However, there was only one combination of three twofold peaks which were orthogonal to each other within less than 1° (Rossmann *et al.*, 1972). Alternatively, the requirement for there to be rotation-function peaks with given angular relationships to each other, as in the case of GAPDH, makes it possible to increase the precision of the determination of the orientation of the whole symmetrical object. This process has been used to advantage in the determination of the orientation of icosahedral viruses [*e.g.* for human rhinovirus 14 (HRV14; Arnold, Erickson, Fout, Frankenberger, Hecht, Luo, Rossmann & Rueckert, 1984) and poliovirus type 1 (Hogle, Chow & Filman, 1985)] where there are many twofold, threefold and fivefold axes at precisely known angles to each other. This constraint on the rotation function has been called a 'locked' rotation function (Rossmann *et al.*, 1972).

Previously, the search for the existence of simultaneous peaks in the rotation function was done either clumsily (Rossmann *et al.*, 1972) or in special cases only. In the case of HRV14, the virus sat on a crystallographic threefold axis. Hence, it was only necessary to rotate the icosahedron about the body diagonal of the cubic cell and sample the rotation function at the end of the various icosahedral symmetry operators. The sum of the rotation-function values was taken as a measure of the locked function. In this way it was possible to obtain an orientation that was later found to be accurate to 0·05° using 5 Å data. However, a general locked rotation function has never been considered.

The locked self-rotation function: rotation of a point group into an unknown cell

Let us consider a point group defined with respect to a set of orthogonal axes. This we shall call the 'standard' orientation. For an icosahedron it might be defined as in Fig. 1. Let $[I_i]$ be a rotation about the ith symmetry element ($i = 1, 2, \ldots, n$) of the point group relative to this set of axes. The operations $[I_i]$ leave the point group unchanged. Our task is to rotate the point group (*e.g.* Fig. 1) into the 'unknown' cell

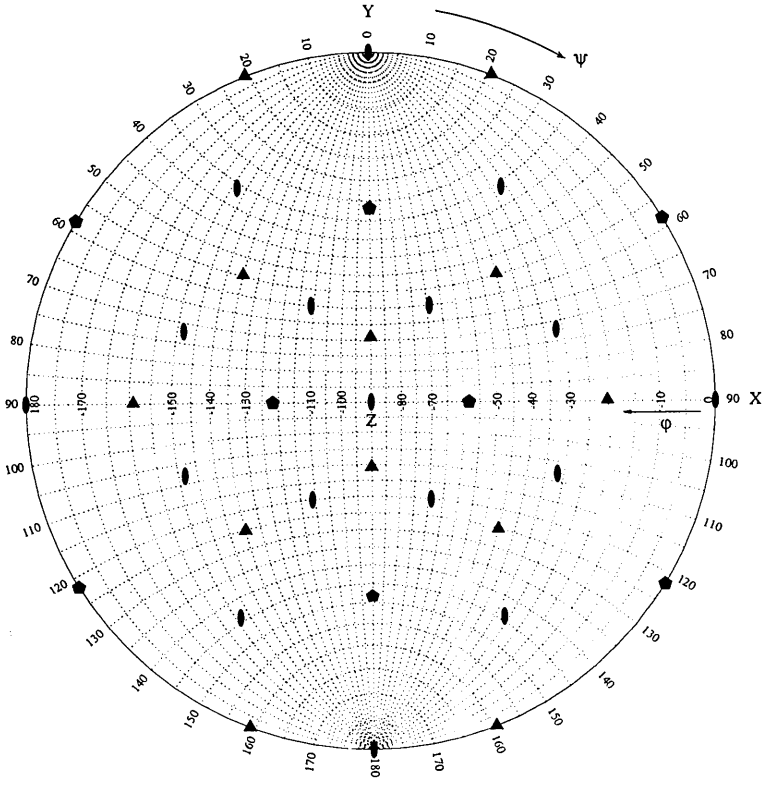

Fig. 1. Icosahedral 532 point group in the standard orientation relative to the orthogonal axes *X*, *Y*, *Z* plotted as a stereographic projection. The polar coordinates ψ, φ are defined as given by Rossmann & Blow (1962).

and to sample the rotation function for each of the n elements $[I_i]$. Rotation of the point group into the unknown cell can be defined, in general, by three angles (*e.g.* polar or Eulerian angles) unless one or more of the point-group symmetry elements is made to be coincident with crystallographic symmetry operators. Let $[E]$ be the rotation matrix which performs such a rotation.

Let a position vector within the point-group system be defined by $\mathbf{U}_1$. If $\mathbf{U}_i$ is the position of the vector after being operated on by $[I_i]$, then

$$\mathbf{U}_i = [I_i]\mathbf{U}_1. \tag{1}$$

Now if a position vector in the orthogonalized system of the crystal lattice is represented by $\mathbf{X}$, then by definition

$$\mathbf{X}_1 = [E]\mathbf{U}_1 \tag{2}$$

or similarly

$$\mathbf{X}_i = [E]\mathbf{U}_i$$

since the rotation $[E]$ takes all equivalent points of the point-group system $\mathbf{U}$ to points that are also equivalent in the crystal system. Hence, from (1),

$$\mathbf{X}_i = [E][I_i]\mathbf{U}_1$$

and, from (2),

$$= [E][I_i][E^{-1}]\mathbf{X}_1. \tag{3}$$

However,

$$\mathbf{X}_i = [\rho_i]\mathbf{X}_1, \tag{4}$$

where $[\rho_i]$ is a rotation which takes $\mathbf{X}_1$ to $\mathbf{X}_i$ in the unknown crystal lattice. By comparison of (3) and (4) we see that

$$[\rho_i] = [E][I_i][E^{-1}]. \tag{5}$$

The self-rotation function tests the degree of the superposition of the Patterson of the unknown structure on itself after a rotation by $[\rho]$. This requires the computation of the matrix $[C]$ which is (following Rossmann & Blow, 1962)

$$[C] = [\alpha][\rho][\beta]. \tag{6}$$

Here $[\alpha]$ deorthogonalizes and converts to fractional coordinates and $[\beta]$ orthogonalizes from the fractional coordinate system in the crystal ($[\alpha]^{-1} = [\beta]$ in this case). Thus, from (5) and (6),

$$[C_i] = [\alpha][E][I_i][E^{-1}][\beta]. \tag{7}$$

Here $[C_i]$ would be the rotation matrix in the unknown cell necessary to test the ith symmetry element of the point group after it has been rotated into the unknown cell by the matrix $[E]$. If R_i is the self-rotation-function value corresponding to the rotation $[C_i]$, then the locked rotation-function value

Table 1. *Examples of types of symmetry operators required for different point groups*

An icosahedron with 532 symmetry

Axes	Number	Number of operations/axis	Total number of operators
Fivefold	6	4	24
Threefold	10	2	20
Twofold	15	1	15
Total	31		59

Tetramer with 222 symmetry

Axes	Number	Number of operations/axis	Total number of operators
Twofold	3	1	3
Total	3		3

Note: Each operator is represented by a matrix $[I_i]$ in the text.

R_L will be given by

$$R_L = \left(\sum_{i=1}^{n} R_i\right) \Big/ n. \tag{8}$$

The locked rotation function can be explored in terms of the three angles that define the rotation matrix $[E]$. This function, much like the ordinary rotation function, will have considerable symmetry depending on the symmetry of the unknown Patterson and the assumed point group of the local symmetry elements. The noise of the locked rotation function can be expected to be $n^{-1/2}$ of the noise of the ordinary rotation function in which the position of each of the n point-group symmetry elements is separately derived. The $[I_i]$ $(i = 1, 2, \ldots, n)$ symmetry elements of the assumed point group can be determined from the minimal number of symmetry elements necessary to generate the different symmetry operators in the standard setting (Table 1).

The locked cross-rotation function

A cross-rotation function can be used to compare the orientation of a molecule in one crystal cell with that of a similar molecule in another crystal cell. This procedure is useful if the structure of a 'known' molecule is to be utilized to solve the structure of a related molecule in an unknown cell. If, however, the unknown cell has more than one molecule per asymmetric unit, then the cross-rotation function should show a fit to each of the orientations in the unknown cell. Very often a self-rotation function has already established the relationship between the various molecules in the unknown cell. This then determines the relationship between the acceptable peaks in the cross-rotation function. Only those peaks can be accepted which are consistent with both the self-rotation function of the unknown cell and the cross-rotation function between the known molecule and the unknown crystal structure. This permits the

computation of a 'locked cross-rotation function' by sampling simultaneously only those points in the cross-rotation function that are consistent with the self-rotation function of the unknown cell.

Let **U** be a position vector in the cell containing the known molecular structure and let **X** be a position vector in the unknown cell containing a number of copies of the molecule per crystallographic asymmetric unit. Then

$$\mathbf{X}_1 = [E]\mathbf{U} \tag{9}$$

where $[E]$ is a rotation matrix that relates the molecule in the known cell with the first molecule in the unknown cell. Now if $[\beta_U]$ is an orthogonalization matrix in the known cell, $[\alpha_X]$ is a deorthogonalization matrix in the unknown cell and lower-case characters denote fractional coordinate vectors, then

$$\mathbf{x}_1 = [\alpha_X][E][\beta_U]\mathbf{u}. \tag{10}$$

Within the unknown cell there exist $[I_i]$ $(i = 1, 2, \ldots, n)$ relationships between the various molecular copies in the asymmetric unit such that

$$\mathbf{X}_i = [I_i]\mathbf{X}_1$$

or

$$\mathbf{x}_i = [\alpha_X][I_i][\beta_X]\mathbf{x}_1. \tag{11}$$

From combination of (10) and (11) it follows that

$$\mathbf{x}_i = [\alpha_X][I_i][\beta_X][\alpha_X][E][\beta_U]\mathbf{u}$$

or

$$\mathbf{x}_i = [\alpha_X][I_i][E][\beta_U]\mathbf{u}. \tag{12}$$

Thus, the locked cross-rotation function R_L can be defined as before in (8) by using

$$[C_i] = [\alpha_X][I_i][E][\beta_U] \tag{13}$$

to evaluate the i rotation-function values R_i and then summing

$$R_L = \left(\sum_{i=1}^{n} R_i\right) \Big/ n.$$

Symmetry of the locked rotation function

Since the order of evaluation of the n values of the ordinary rotation function R_i that go into the computation of the locked rotation function is irrelevant, the locked rotation function has higher symmetry than the ordinary rotation function. An intuitive consideration shows that there are n symmetry-related asymmetric units in a plot of the locked rotation function within the asymmetric unit of the ordinary rotation function. The time of calculation of the locked rotation function will be n times that of the ordinary rotation function at each grid point. However, as the symmetry is n times as high, the time of calculation of the locked rotation function should be expected to be the same as the ordinary function provided the limits of the asymmetric unit of the locked rotation function can be determined in advance. Furthermore, the computation of the locked rotation function can be speeded up considerably because fewer 'large terms' are required for the calculation (see next section).

Consider the rotation $[E_A]$ that is applied to the ith symmetry operator $[I_i]$ of the standard point group. Then from (5)

$$[\rho] = [E_A][I_i][E_A^{-1}]. \tag{14}$$

Similarly, consider the rotation $[E_B]$ that produces the same set of local symmetry operators by rotating the jth symmetry operator $[I_j]$ into the same position previously occupied by $[I_i]$. Then

$$[\rho] = [E_B][I_j][E_B^{-1}]. \tag{15}$$

Comparison of (14) and (15) shows that

$$[E_B][I_j][E_B^{-1}] = [E_A][I_i][E_A^{-1}]. \tag{16}$$

Thus, the locked rotation function will have equivalent values for angles θ_A (corresponding to the matrix $[E_A]$) and θ_B (corresponding to the matrix $[E_B]$) which satisfy (16) for all combinations of i and j. The above derivation has omitted the possibility of superimposing the standard point group onto crystallographically related orientations in the unknown cell. If the mth and nth crystallographic symmetry operations are given by $[T_m]$ and $[T_n]$, then (16) expands to

$$[E_B][I_j][E_B^{-1}] = \{[\alpha^{-1}][T_n][\alpha]\}\{[E_A][I_i][E_A^{-1}]\} \times \{[\beta][T_m][\beta^{-1}]\}. \tag{17}$$

The symmetry of the locked cross-rotation function can be similarly derived. It can be shown that in this case

$$[I_j][E_B] = \{[\alpha^{-1}][T_n][\alpha]\}\{[I_i][E_A]\} \times \{[\beta][T_m][\beta^{-1}]\}. \tag{18}$$

Given a set of angles θ_A, (18) can be solved analytically for the symmetry-related angles θ_B. Equation (17) can be solved by a search through a predefined region of θ_B because an analytical solution is difficult to obtain. However, this is time consuming and, hence, a geometric approach is more useful. Below is shown an example for an icosahedral point group.

Let us define the standard icosahedron with a fivefold axis along Z (Fig. 2). This is convenient in defining Eulerian angles that determine $[E]$. The first rotation in Eulerian space is θ_1 about Z. A rotation of 72° will superimpose the icosahedron onto itself. Hence, the locked rotation function will repeat itself every 72° in θ_1. The second rotation (θ_2) is about the direction of the new X axis. This rotation will tilt the

fivefold axis away from the Z axis. However, there are already fivefold axes that are tilted by 63·4° from the Z axis, which suggests that the rotation by θ_2 need not exceed 63·4°. Since it is anticipated that there are 60 asymmetric units in the locked rotation function, it follows that the final angle θ_3 need only be explored from 0 to 90° if the three Eulerian rotations are not correlated.

The computer program

The program has been written in standard Fortran and is available for distribution upon request. The input commands to the program are all free formatted and keyword based. The program has been tested extensively on a Cyber 205 and a VAX/VMS 8550.

The program has been written to perform general rotation searches, permitting the calculation of ordinary rotation functions. In the case of locked rotation searches, minimal information is needed to specify the point-group symmetry of the particle. For example, to designate the 532 point-group symmetry of icosahedral viruses, only a fivefold and a non-perpendicular twofold axis (or a three- and a twofold axis *etc.*) need to be specified. The program can then generate all the 60 symmetry operations of the icosahedron.

Many different conventions exist in the literature as to the orthogonalization of the crystallographic unit cell, and for the definition of Eulerian and polar angles (Rossmann & Blow, 1962; Machin, 1985). The program has been coded to be able to support all of these conventions. The user of the program can choose among the supported conventions or insert new ones.

The 'slow' rotation function (Rossmann & Blow, 1962; Tollin & Rossmann, 1966) has been employed in the program due to the need to sample the rotation function at precise sets of angles when evaluating the locked rotation function. In order to speed up the calculation, the program keeps only the strongest reflections in the data set that is being rotated (Tollin & Rossmann, 1966). These reflections, also known as 'large terms', are selected based on the criterion that $I > (\text{cut-off} \times I_{avg})$, where the average intensity (I_{avg}) has been calculated in shells of resolution. A cut-off value which selects about 5% of the possible data in each shell is a useful guide for calculating ordinary rotation functions.

Testing the locked self-rotation function

The locked self-rotation function has been extensively tested with diffraction data from four icosahedral virus crystals (Table 2) - φX174 (Willingmann, Krishnaswamy, McKenna, Smith, Olson, Rossmann, Stow & Incardona, 1990; and unpublished results), tetragonal canine parvovirus (CPV; Luo, Tsao, Rossmann,

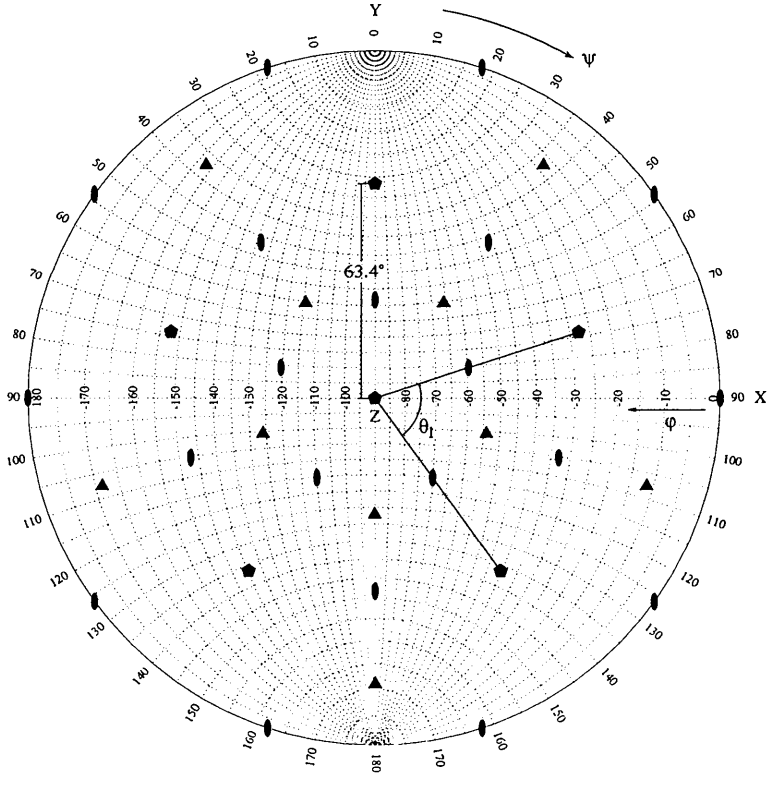

Fig. 2. Icosahedral 532 point group with a fivefold axis along Z. Rotation about Z by the Eulerian angle θ_1 through 72° will superimpose the figure onto itself. A subsequent rotation of $\theta_2 = 63{\cdot}4°$ will also superimpose the figure onto itself. Thus, limits of the asymmetric unit in the locked rotation function are $0 < \theta_1 < 72°$, $0 < \theta_2 < 63{\cdot}4°$.

Table 2. *Results of locked self-rotation function test cases*

Virus crystal*	φX174		Tetragonal CPV		Mengo virus		HRV14	
a (Å)	306·0		257·5		441·4		445·1	
b (Å)	361·1		257·5		427·3		445·1	
c (Å)	299·7		807·0		421·9		445·1	
β (°)	92·91							
Space group	$P2_1$		$P4_12_12$		$P2_12_12_1$		$P2_13$	
	Initial	Final	Initial	Final	Initial	Final	Initial	Final
Resolution (Å)	30–15	10–7	30–16	12–8	30–15	10–7**	30–15	10–7
Number of observed reflections	2971	21 001	2093	13 300	6980	54 832	4191	29 548
Number of possible reflections	9000	66 000	3000	20 000	11 000	80 000	4500	30 000
Cut off†	5·0	2·0	4·5	2·8	6·0	3·5	4·0	3·5
Number of large terms	12	2452	11	486	34	1362	33	729
% observed reflections	0·40	11·68	0·53	3·65	0·49	2·48	0·79	2·47
Search interval (°)	8·0	0·1	5·0	0·1	8·0	0·1	8·0	0·1
Search grid	12×12×12	3×3×3	10×10×10	1×3×3	12×12×12	3×3×3	4×4×4	1×1×9
Peak position‡ θ_1	82·00	82·15	0·00	0·00	0·00	0·00**	0·00	−45·00
Peak position‡ θ_2	92·00	92·35	25·00	25·80	88·00	90·00**	0·00	54·74
Peak position‡ θ_3	82·00	81·65	15·00	16·30	40·00	39·85**	0·00	113·95
Peak height§	70	79	50	73	65	190	125	123
Peak/σ¶	8.3	11	4·8	10	5·2	27	5·2	6
Next peak height	44	—	43	—	53	—	56	—
Next peak/σ	5·0	—	4·0	—	4·1	—	1·3	—
Cyber 205 CPU (min)	11	42	22	11	50	49	8	31

* Radius of integration was chosen as 150 Å in each case. Diameter of particles is about 260 Å for φX174 and CPV, and 300 Å for Mengo virus and HRV14.

† Cut off implies that reflections greater than (cut off × I_{avg}) were used as large terms.

‡ Eulerian angles are defined as in Rossmann & Blow (1962). Polar angles (φ, ψ, κ) (Rossmann & Blow, 1962) are given in the case of HRV14 where the particle is situated on a crystallographic threefold axis. The angles are applied to the standard orientation as defined in Fig. 1.

§ Peak heights are normalized so that the origin peak is 1000.

¶ For the initial solution, peak/σ shows the peak height above mean height divided by the r.m.s. deviation from the mean. For the final solution, peak/σ shows the ratio of the peak height over the estimated average background.

** On increasing the resolution of the data to 5 Å, the ordinary rotation-function peaks split into two parts because an icosahedral twofold axis is about 1° away from being parallel to the crystal b axis (Luo *et al.*, 1989).

Basak & Compans, 1988; and unpublished results), Mengo virus (Luo, Vriend, Kamer, Minor, Arnold, Rossmann, Boege, Scraba, Duke & Palmenberg, 1987; Luo, Vriend, Kamer & Rossmann, 1989) and HRV14 (Rossmann, Arnold, Erickson, Frankenberger, Griffith, Hecht, Johnson, Kamer, Luo, Mosser, Rueckert, Sherry & Vriend, 1985; Arnold, Vriend, Luo, Griffith, Kamer, Erickson, Johnson & Rossmann, 1987). There is one icosahedral virus particle per crystallographic asymmetric unit for φX174 and Mengo virus crystals. In HRV14 the particle is sitting on a crystallographic threefold axis, thus allowing only one degree of rotational freedom, which is the rotation about the cell body diagonal. Similarly, in CPV the particle is situated on a twofold axis.

An initial rough search using a coarse grid and relatively few large terms at low resolution yielded an approximate position in each case. Subsequent searches with progressively finer grids, more large terms and higher-resolution reflection data gave the precise orientation of the virus particle in the unit cell. These results were compared with those from the ordinary rotation function for φX174 and CPV, where no structure is yet available, or with the results of the atomic resolution structure for Mengo virus and HRV14. The φX174 results will be described in greater detail here (Table 2).

The φX174 crystals have space group $P2_1$ with $a = 306{\cdot}0$, $b = 361{\cdot}1$, $c = 299{\cdot}7$ Å and $\beta = 92{\cdot}9°$. The reflection data set available at the time the tests were conducted contained only 30% of the total possible reflections to 7 Å resolution, although the diffraction patterns could be observed to at least 2·8 Å resolution. All functions were scaled so that the ordinary rotation function had a value of 1000 at its origin, corresponding to the superposition of the Patterson onto itself without rotation.

The initial coarse search was conducted for data between 30 and 15 Å resolution with 2971 observed reflections. From these, 12 terms were selected that were larger than five times the mean intensity of all the observed reflections in the resolution shell. The successful selection of such a small number of terms may depend critically on their uniform distribution in reciprocal space. The initial search was carried out in Eulerian space (defined as in Rossmann & Blow, 1962) between 0 and 88° in all three angles, using 8° intervals, and took 11 CPU min on a Cyber 205. The highest peak in the map was located at 24, 8, 72° with a height of 70 and was 8·3σ above the background. The next lower peak had a height of 44 and was 5·0σ above background. Subsequent searches were conducted about the symmetry-related position at 82, 92, 82°. [The peaks at 24, 8, 72° and 82, 92, 82° are related by (17).] Reflection data between 15 and 9 and

Table 3. *Reduction of noise level in the locked rotation function for φX174*

Large term selection			Locked rotation function				Ordinary rotation function				
Cut-off value	Number of terms	% of observed data	Peak height	bgd*	σ(bgd)	S/N†	⟨Peak height⟩	⟨bgd⟩	⟨σ(bgd)⟩	⟨S/N⟩	Noise reduction‡
3·0	772	3·7	78·9	6·5	1·4	51·7	79·5	6·8	9·1	8·0	4·9
5·0	83	0·4	75·0	5·8	1·4	49·4	76·0	6·0	13·5	5·2	9·6
6·5	17	0·08	59·9	5·5	2·0	27·2	60·8	7·1	22·6	2·4	11·3

* bgd gives the mean function height away from the peak position.
† S/N means signal-to-noise ratio which is given by [(peak height − bgd)/σ(bgd)].
‡ Noise reduction of the locked rotation function is given by [σ(bgd) for the ordinary rotation function/σ(bgd) for the locked rotation function].

then 10 and 7 Å resolution were used for the finer searches. At 10–7 Å resolution, a cut-off value of 2·0 resulted in 2452 reflections (11·7% of the observed data) being saved as large terms. Successive search intervals of 2·0, 0·5, 0·2 and 0·1° were used around the position 82, 92, 82°. The final search, with a 0·1° interval, localized the peak to 82·15, 92·35, 81·65°.

The orientation of the icosahedral particle in φX174 was previously determined using the ordinary rotation function. This involved the determination of the location of the 31 rotation axes (Table 1) and their superposition onto an ideal icosahedron. Similar reflection data were used in both the ordinary and the locked rotation-function searches. The two results differ by a rotation of less than 0·05° for the standard icosahedron. The ordinary rotation-function results showed a r.m.s. deviation of 0·15° between the observed axial orientation and an ideal icosahedron.

The locked rotation-function search is a simplified process involving only the location of a single peak as compared to the location of each of the 31 axes in the ordinary rotation function. Moreover, the locked rotation-function map has less background variation as compared to the ordinary rotation function (Table 3).

The results for orthorhombic Mengo virus, whose structure has been determined (Luo *et al.*, 1987, 1989), also show good correspondence. Fig. 3 shows the orientation of one of the four particles in the Mengo unit cell as output directly from the locked-rotation-function program, using 10–7 Å resolution data. One of the icosahedral twofold axes is only 1° off parallel

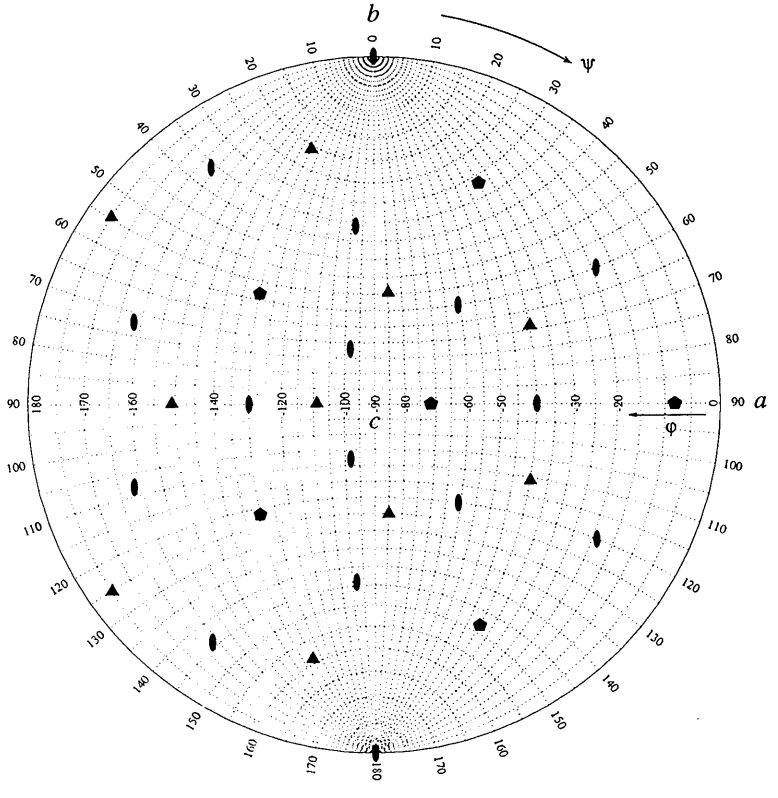

Fig. 3. Stereographic projection of the orientation of one Mengo virus icosahedral particle in the $P2_12_12_1$ crystal cell. This plot is obtained directly from the locked rotation-function program.

from the crystallographic *b* axis (Luo *et al.*, 1989). This tilt can only be recognized by 5 Å resolution or better data.

Parameter dependence of the locked self-rotation function

The computation of an ordinary rotation function or a locked rotation function is dependent on the choice of a variety of parameters, some of which are determined by the nature of the problem and others which can be selected to arbitrate between speed of calculation and quality of the result. Some of the parameters that determine a rotation function are: (1) symmetry of the Laue group associated with the two Pattersons that are being compared; (2) symmetry of the point group of the molecule in the asymmetric unit; (3) percentage of data available for each of the two crystal forms; (4) percentage of data selected as large terms from the available observations; (5) resolution limits of the selected data; and (6) radius of integration.

Usually a higher-symmetry space group reduces the quality of a rotation function because the self-vectors around the origin are a sum of those derived from the different particles in the various asymmetric units of the cell. The locked rotation function helps to alleviate this problem by checking the value of the rotation function at various points simultaneously. Table 2 shows that, in general, taking into account the percentage of data available, the peak-to-background ratio is best for lower-symmetry space groups.

Tests to determine the effect of changing the percentage of selected large terms used in computing the locked rotation function were performed with the φX174 data. While these tests were being conducted, there was available only 30% of the theoretically possible data. Presumably, as the percentage of observed data increases, so the number of large terms could be reduced even further. The width and height above background of the current peak were used as criteria for these tests.

A series of cut-off values, between 2·0 and 7·0, was used to select the large terms. A θ_1 scan was then made in 1° intervals, for each cut-off value, across the correct peak of the locked rotation function (Fig. 4). The peak profile looked similar, independent of whether 11·7% (2452 reflections) or only 0·13% (28 reflections) of the observed data were used as large terms. When even fewer reflections were used (a cut-off of 6·5 or 7·0 yielding 16 or 7 reflections, respectively), there was only a slight decrease in the peak height and peak width relative to the original peak. Similar results were obtained in scans of θ_2 and θ_3 across the same peak. It can, therefore, be assumed that using only 0·5% of the observed data as large terms in the presence of 60-fold redundancy and a monoclinic space group is, in general, sufficient to locate the correct locked rotation-function peak.

The ordinary rotation function is much more sensitive than the locked rotation function to the percentage of reflections that are used as large terms (Table 3). The major difference between the ordinary and locked rotation functions lies in the background noise level, σ(bgd). Whereas the noise level in the locked rotation function was only slightly affected by reducing the number of large terms from 3·7 to 0·08% (Table 3), that in the ordinary rotation function more than doubled. Consistent with the earlier discussion, the amount of noise reduction was found to be roughly proportional to $n^{1/2}$, or 7·7 for 60-fold redundancy.

Three different ranges (10–7, 20–10 and 30–15 Å) were used to investigate the resolution dependence of the locked self-rotation function. Two separate cases were tested, one using about 4% of the observed data as large terms and the other using 1%. In both cases, the peak progressively widened as the resolution of the data was decreased (Fig. 5). At 30–15 Å

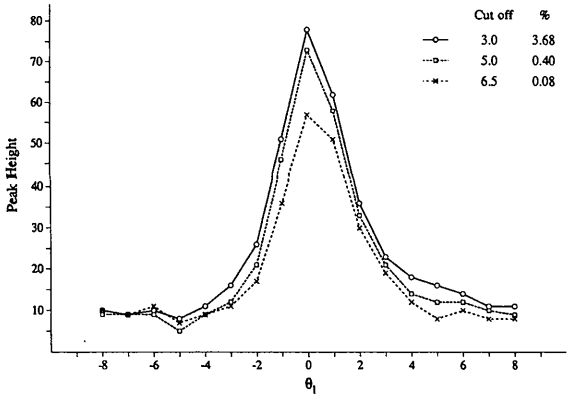

Fig. 4. Section through the correct locked rotation-function peak in θ_1 to assess the effect of the number of large terms included in the calculation. The resolution range used was 10–7 Å and the radius of integration was 150 Å for φX174 diffraction data.

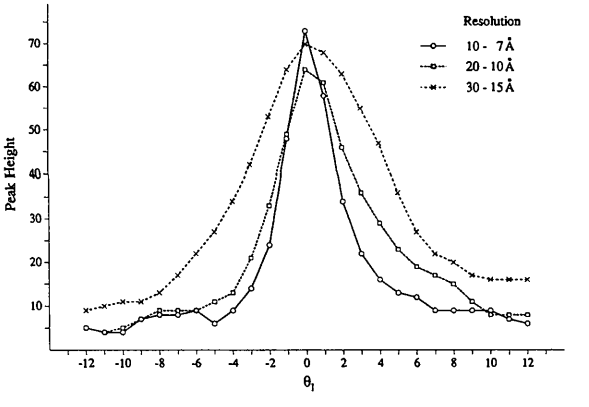

Fig. 5. Section through correct locked rotation-function peak in θ_1 to assess the effect of resolution range. The percentage of large terms was 1% of the observed φX174 diffraction data and the radius of integration was 150 Å.

Table 4. *Locked cross-rotation function for monellin*

	All atoms						Main-chain atoms only					
	Peak†	θ_1	θ_2	θ_3	Height	σ‡	Peak	θ_1	θ_2	θ_3	Height	σ‡
Ordinary	1	30	41	54	100	4·0*A	1	26	43	60	100	3·7*A
rotation	2	30	61	56	58	2·3	2	0	121	46	49	1·8*B
function	3	30	83	50	50	2·0*A	3	26	79	54	47	1·8*A
	4	30	13	58	49	2·0	4	4	107	52	46	1·7
	5	0	15	48	28	1·1*B	5	36	13	54	41	1·5
Locked	1	29	40	55	100	2·9*	1	27	44	59	100	2·6*
rotation	2	21	20	67	65	1·9	2	15	38	73	80	2·1
function A§							3	5	60	75	40	1·1
Locked	1	0	16	44	100	2·3*	1	−4	14	50	100	2·1*
rotation	2	18	16	48	86	2·0	2	−16	10	64	95	2·0
function B	3	16	24	28	82	1·9	3	4	−6	50	74	1·6
	4	20	12	68	71	1·6	4	−4	−4	56	74	1·6
	5	18	−6	66	50	1·2	5	−6	−6	60	74	1·6

The correct solutions are:

	Peak	θ_1	θ_2	θ_3	
Ordinary rotation function	1	27	42	55	A
	2	30	83	50	A
	3	0	14	48	B
	4	0	120	47	B
Locked rotation function A	1	27	42	55	A
Locked rotation function B	1	0	14	48	B

* The correct peaks are marked by an asterisk and are identified by the corresponding dimer A or B.
† Peak positions are given in descending order of height.
‡ σ designates the number of standard deviations above background.
§ Locked rotation functions A and B correspond to assuming the orientation of the twofold axis in dimer A or in dimer B.

resolution, the half-width of the peak was 8°, suggesting that a locked-rotation-function search can be initiated by using 30–15 Å reflection data for particles around 300 Å in diameter and using 5–8° as the search interval.

Testing the locked cross-rotation function with data from monellin

Crystals of monellin (a sweet protein) belong to space group $P2_1$ with $a = 39{\cdot}84$, $b = 87{\cdot}20$, $c = 72{\cdot}06$ Å and $\gamma = 107{\cdot}3°$. There are four molecules in the asymmetric unit consisting of two independent dimers (Ogata, Hatada, Tomlinson, Shin & Kim, 1987). The refined atomic structure for one of the four molecules was used as the search model. The model was rotated by Eulerian angles 125, 42, 153° and placed in a triclinic cell with $a = b = c = 90$ Å and $\alpha = \beta = \gamma = 90°$. Structure factors were calculated to 3·5 Å resolution using an overall temperature factor of 30 Å^2.

An ordinary cross-rotation function was calculated using 2° intervals and 8–3·5 Å resolution reflection data. The cut-off value was set at 2·0, giving 12·6% of the observed reflections as large terms (Table 4). Two different search models were used - one with all the atoms and the other with only the main-chain atoms. Although peaks corresponding to three out of the four molecules are among the top five peaks in this map, the peak heights are mostly indistinguishable from the background.

The locked rotation function used exactly the same reflection data. Two separate locked rotation functions were computed because the local noncrystallographic twofold axes are not related to each other in a simple manner. In both searches, using either all the atoms or main-chain atoms only, the highest peak in the maps corresponded to the correct solution and, in three of the four cases, there was a clear separation between the top peak and the second peak. Each search took between 2 and 4 CPU h on a VAX/VMS 8550.

Discussion

The locked rotation function is extremely useful in determining the orientation of icosahedral virus particles in their unit cell. The calculation time, as well as the amount of manual intervention, can be significantly reduced when compared to the ordinary rotation function. With the solution of the rotation problem, the translation problem can then be tackled using a homologous structure, an electron microscopy image or the isomorphous replacement method. The abundance of local symmetry can also be employed to improve and extend the initial phases by molecular averaging (Rossmann, 1990). For proteins, the noncrystallographic symmetry is, in general, of lower symmetry. Nevertheless, the locked cross-rotation function appears to be a powerful aid even in the presence of only a single local twofold axis. The calculation of the locked rotation function requires considerably fewer large terms than the ordinary

rotation function to attain interpretable results and, thus, provides a significant saving of computation time. Furthermore, an accurate solution can be obtained from a reasonably small number of strong reflections using relatively low-resolution data.

Dr S. Krishnaswamy carried out initial tests on a locked rotation function generated by point-group-specific modifications to the laboratory's standard program. These he showed to be highly successful on HRV14 and Mengo virus and provided the incentive for the study reported here. We are most grateful for and encouraged by his initial work. We thank Robert McKenna and Peter Willingmann for providing the φX174 diffraction data before publication, Walter Keller and Kathy Smith for the tetragonal CPV data and Sung-Hou Kim for permission to use the monellin data and coordinates. We also thank Helene Prongay and Sharon Wilder for help in the preparation of this manuscript. The work was supported by grants to MGR from the National Science Foundation and the National Institutes of Health.

References

ARNOLD, E., ERICKSON, J. W., FOUT, G. S., FRANKENBERGER, E. A., HECHT, H. J., LUO, M., ROSSMANN, M. G. & RUECKERT, R. R. (1984). *J. Mol. Biol.* **177**, 417-430.

ARNOLD, E., VRIEND, G., LUO, M., GRIFFITH, J. P., KAMER, G., ERICKSON, J. W., JOHNSON, J. E. & ROSSMANN, M. G. (1987). *Acta Cryst.* **A43**, 346-361.

HOGLE, J. M., CHOW, M. & FILMAN, D. J. (1985). *Science*, **229**, 1358-1365.

HOPPE, W. (1957). *Acta Cryst.* **10**, 750-751.

LUO, M., TSAO, J., ROSSMANN, M. G., BASAK, S. & COMPANS, R. W. (1988). *J. Mol. Biol.* **200**, 209-211.

LUO, M., VRIEND, G., KAMER, G., MINOR, I., ARNOLD, E., ROSSMANN, M. G., BOEGE, U., SCRABA, D. G., DUKE, G. M. & PALMENBERG, A. C. (1987). *Science*, **235**, 182-191.

LUO, M., VRIEND, G., KAMER, G. & ROSSMANN, M. G. (1989). *Acta Cryst.* B**45**, 85-92.

MACHIN, P. A. (1985). Editor. *Molecular Replacement. Proceedings of the Daresbury Study Weekend*, 15-16 *February* 1985. Warrington: SERC Daresbury Laboratory.

OGATA, C., HATADA, M., TOMLINSON, G., SHIN, W.-C. & KIM, S.-H. (1987). *Nature (London)*, **328**, 739-742.

ROSSMANN, M. G. (1990). *Acta Cryst.* **A46**, 73-82.

ROSSMANN, M. G., ARNOLD, E., ERICKSON, J. W., FRANKENBERGER, E. A., GRIFFITH, J. P., HECHT, H. J., JOHNSON, J. E., KAMER, G., LUO, M., MOSSER, A. G., RUECKERT, R. R., SHERRY, B. & VRIEND, G. (1985). *Nature (London)*, **317**, 145-153.

ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* **15**, 24-31.

ROSSMANN, M. G., FORD, G. C., WATSON, H. C. & BANASZAK, L. J. (1972). *J. Mol. Biol.* **64**, 237-249.

TOLLIN, P. & ROSSMANN, M. G. (1966). *Acta Cryst.* **21**, 872-876.

WILLINGMANN, P., KRISHNASWAMY, S., MCKENNA, R., SMITH, T. J., OLSON, N. H., ROSSMANN, M. G., STOW, P. L. & INCARDONA, N. L. (1990). *J. Mol. Biol.* **212**, 345-350.

Acta Cryst. (1990). A**46**, 792-798

Diffraction Theory for Diffuse Scattering by Correlated Microdomains in Materials with Several Atoms per Unit Cell

BY R. B. NEDER, F. FREY AND H. SCHULZ

Institut für Kristallographie und Mineralogie, Theresienstrasse 41, 8000 *München* 2, *Federal Republic of Germany*

(*Received* 3 *August* 1989; *accepted* 25 *May* 1990)

Abstract

A general diffraction theory is presented for the diffuse scattering by correlated microdomains within a disordered structure. The theory applies to crystals with several atoms per unit cell and several types of different microdomains. An analytical expression is given for an assumed distribution function of the microdomains within the disordered matrix of the host. Since the analytical Fourier transform of this distribution function is also given, very fast calculation of the diffuse intensity is possible.

Introduction

Quantitative interpretations of diffuse scattering are frequently done on the basis of the Warren short-range-order (SRO) parameters. In the case of simple alloy structures, good agreement of calculated and observed intensities is obtained. Hayakawa & Cohen (1975) presented a generalized solution for structures with several sublattices. Any description by SRO parameters, however, will not yield information on the actual distribution of the defects. Extended defects cannot be described. In particular, if the distri-

0108-7673/90/100792-07$03.00