Acta Cryst. (1989). A**45**, 309–314

Simultaneous Search for Symmetry-Related Molecules in Cross-Rotation Functions

BY T. O. YEATES

Department of Molecular Biology, Research Institute of Scripps Clinic, La Jolla, CA 92037, *USA*

(*Received* 12 *January* 1988; *accepted* 8 *November* 1988)

Abstract

In a typical cross-rotation function, the Patterson function of a single search molecule is compared with an observed Patterson function, which contains a set of symmetry-related intramolecular vector sets. In principle, it is better to search for the symmetry-related molecules simultaneously, and Nordman [Am. Crystallogr. Assoc. (1986), Hamilton, Ontario, *Program Abstr.* p. 36] has reported success with an algorithm of this type. In this paper, the differences between the ordinary search and a simultaneous search are investigated, and it is shown that the combined presence of crystallographic symmetry and approximate symmetry of a search model may lead to significant bias in conventional rotation functions. The nature and magnitude of this symmetry bias are discussed. An efficient algorithm is derived for generating a modified unbiased cross-rotation function map from conventional rotation functions. Two examples are described that demonstrate improvement in the quality of the rotation function maps and the ability to obtain physically meaningful correlation coefficients.

1. Introduction

In the molecular replacement method, one begins with a search model which resembles the structure of the crystallized molecule, and attempts to determine its orientation and position in the unit cell. If the attempt is successful, approximate phases for the observed structure-factor amplitudes can be obtained. The first step is to determine the correct orientation of the model in the unit cell by analysis of rotation functions. Several methods for calculating rotation functions, both in Patterson space (Nordman & Nakatsu, 1963) and reciprocal space (Rossmann & Blow, 1962; Huber, 1969; Crowther, 1972) have been described.

Orientational information can be obtained, in the absence of phase information, from a Patterson function corresponding to the observed intensities. The correct orientation of the model is determined by identifying orientations for which the Patterson function corresponding to the observed intensities resembles a Patterson function (or an interatomic vector set) corresponding to the oriented search model. For most space groups, there are multiple copies of the molecule of interest in each unit cell, in orientations related by crystallographic symmetry. Each molecule contributes a set of intramolecular vectors (self vectors) to the observed Patterson function, which consequently resembles a superposition of symmetry-related intramolecular vector sets. Now, consider a model which is in the same orientation as one of the molecules in the unit cell. Since this orientation corresponds to a solution of the problem, it is expected that the oriented model Patterson function will resemble the observed Patterson function. However, the contributions from the symmetry-related molecules to the observed Patterson may adversely affect the similarity between the observed and model Patterson functions. Fortunately, the systematic error that arises from the presence of symmetry-related molecules can be removed. In what follows, we derive the details of the desired correction, including an efficient method for applying it, and discuss the nature of the modification and the cases for which it is expected to be significant.

2. Derivation

Consider a cross-rotation function between the observed Patterson function P_o and the Patterson function of the search model P_m. Let the correlation between P_o and P_m be C, which is a function of the orientation of the model. The orientation is expressed in terms of three Eulerian angles, α, β and γ, and $R(\alpha, \beta, \gamma)$ is the rotation matrix associated with the angles. The correlation function C is given by

$$\begin{aligned} C(\alpha, \beta, \gamma) = &\langle P_o(\mathbf{x}) | P_m[R(\alpha, \beta, \gamma) \cdot \mathbf{x}]\rangle \\ &\times \langle P_o(\mathbf{x}) | P_o(\mathbf{x})\rangle^{-1/2} \\ &\times \langle P_m[R(\alpha, \beta, \gamma) \cdot \mathbf{x}]| \\ &\times P_m[R(\alpha, \beta, \gamma) \cdot \mathbf{x}]\rangle^{-1/2} \end{aligned} \quad (1)$$

where the brackets indicate the integral over space of the product of the two enclosed functions. The domain for the integration is usually taken to be a sphere centered on the origin. The terms in the denominator of (1) are normalization constants. Since the denominator of (1) is independent of α, β and γ, we can write

$$C(\alpha, \beta, \gamma) \propto \langle P_o(\mathbf{x}) | P_m[R(\alpha, \beta, \gamma) \cdot \mathbf{x}]\rangle. \quad (2)$$

0108-7673/89/050309-06$03.00

The right-hand side of (2) is the value determined in ordinary summation-type rotation functions [Rossmann & Blow (1962) and the 'fast rotation function' of Crowther (1972)].

Now, if there are n crystallographically related molecules in the unit cell, the observed Patterson function will consist of the superposition (or summation) of n symmetry-related intramolecular vector sets (self vectors), as well as intermolecular vector sets (cross vectors) between symmetry-related molecules. In most rotation functions, the observed Patterson function is correlated with a Patterson function of a single search molecule. As previously mentioned, when the correlation function is calculated in this way, the $(n-1)$ symmetry-related self-vector sets in the observed Patterson contribute to the background noise of the rotation function. Nordman (1986) has pointed out that it is better to evaluate the correlation between the observed Patterson and the superposition of symmetry-related self-vector sets of the model. Therefore, an improved correlation function, N, can be written as

$$N(\alpha,\beta,\gamma)=\left\langle P_o(\mathbf{x})\left|\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right.\right\rangle \times\langle P_o(\mathbf{x})|P_o(\mathbf{x})\rangle^{-1/2} \times\left\langle\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right| \times\left.\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right\rangle^{-1/2} \tag{3}$$

where the S_i are the proper symmetry operations of the crystal point group. It is important to note that the rotations S_i and R do not commute, and that the symmetry operations S_i are applied after the rotation R. While the algorithm of Rossmann & Blow (1962) could be modified to evaluate (3) directly, this approach would be computationally intensive (at least n times as long as the ordinary cross-rotation function evaluation). In addition, the 'fast rotation function' algorithm (Crowther, 1972) cannot be modified to evaluate the expression in (3). Fortunately, (3) simplifies to a form which may be evaluated from two ordinary rotation functions of the type in (2). This simplification allows rapid evaluation, using fast rotation functions for cases where efficiency and speed are required.

The 'symmetry-corrected' correlation, $N(\alpha,\beta,\gamma)$, can be evaluated from a cross-rotation function and a self-rotation function, as follows. For simplicity, we temporarily substitute D for the denominator of (3).

$$N(\alpha,\beta,\gamma)=\frac{\langle P_o(\mathbf{x})|(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\})\rangle}{D}. \tag{4}$$

Rearrangement gives

$$N(\alpha,\beta,\gamma)=\frac{\sum_{i=1}^{n}\langle P_o(\mathbf{x})|P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\rangle}{D}. \tag{5}$$

Since the product integral in the numerator of (5) is not affected by rotating the P_o and P_m functions simultaneously,

$$N(\alpha,\beta,\gamma)=\frac{\sum_{i=1}^{n}\langle P_o(S_i^{-1}\,.\,\mathbf{x})|P_m[R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\rangle}{D}. \tag{6}$$

Since S_i (and S_i^{-1}) are symmetry operations of the function P_o,

$$P_o(S_i^{-1}\,.\,\mathbf{x})=P_o(\mathbf{x})$$

and

$$N(\alpha,\beta,\gamma)=n\langle P_o(\mathbf{x})|P_m[R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\rangle/D \propto C(\alpha,\beta,\gamma)/D. \tag{7}$$

Therefore, the numerator of (3) corresponds to the usual cross-rotation function $C(\alpha,\beta,\gamma)$.

Now we show how the denominator D can be determined from the self-rotation function of the model Patterson. Recall that

$$D(\alpha,\beta,\gamma)=\langle P_o(\mathbf{x})|P_o(\mathbf{x})\rangle^{1/2} \times\left\langle\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right| \times\left.\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right\rangle^{1/2}. \tag{8}$$

Since $\langle P_o(\mathbf{x})|P_o(\mathbf{x})\rangle^{1/2}$ is independent of α, β and γ,

$$D(\alpha,\beta,\gamma)\propto\left\langle\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right| \times\left.\left(\sum_{i=1}^{n}\{P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\}\right)\right\rangle^{1/2}. \tag{9}$$

Rearrangement gives

$$D(\alpha,\beta,\gamma)\propto\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\langle P_m[S_i\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]| \times P_m[S_j\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\rangle\right)^{1/2} \tag{10}$$

and

$$D(\alpha,\beta,\gamma)\propto\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\langle P_m(\mathbf{x})|P_m[R^{-1}(\alpha,\beta,\gamma) \,.\,S_i^{-1}\,.\,S_j\,.\,R(\alpha,\beta,\gamma)\,.\,\mathbf{x}]\rangle\right)^{1/2}. \tag{11}$$

The bracketed value in (11) corresponds to a value of the self-rotation function of the search model,

Table 1. *Range of the symmetry-correction term for cross-rotation functions*

Results for three different models in two different Laue symmetries are given. The results are based upon calculated intensity data from 10 to 2·5 Å, and a search radius of 14 Å. The structures of myoglobin (Kendrew *et al.*, 1960; Watson, 1969) and cyctochrome c (Carter *et al.*, 1985) were taken from the Brookhaven Protein Databank (Bernstein *et al.*, 1977). See text for definitions.

Model	Number of amino acids	Laue symmetry (n)	Range of the symmetry-correction term, D
Polyalanine α-helix	20	*mmm* (4)	0·81–1·99
Polyalanine α-helix	20	4/*mmm* (8)	0·69–2·81
Myoglobin helix no. 7 (residues 100–118)	19	*mmm* (4)	0·89–1·56
Myoglobin helix no. 7 (residues 100–118)	19	4/*mmm* (8)	0·79–1·96
Cytochrome c_5	153	*mmm* (4)	0·93–1·29
Cytochrome c_5	153	4/*mmm* (8)	0·86–1·47

sampled at an orientation determined by $R(\alpha, \beta, \gamma)$, S_i and S_j. Letting Q be the self-rotation function of the model, we obtain

$$Q(\alpha', \beta', \gamma') = \langle P_m(\mathbf{x}) | P_m[R(\alpha', \beta', \gamma') \cdot \mathbf{x}] \rangle. \quad (12)$$

Equations (7), (11) and (12) give

$$N(\alpha, \beta, \gamma) \propto \frac{C(\alpha, \beta, \gamma)}{[\sum_{i=1}^{n} \sum_{j=1}^{n} Q(\alpha'_{ij}, \beta'_{ij}, \gamma'_{ij})]^{1/2}} \quad (13)$$

where α'_{ij}, β'_{ij}, and γ'_{ij} are defined by

$$R(\alpha'_{ij}, \beta'_{ij}, \gamma'_{ij}) = R^{-1}(\alpha, \beta, \gamma) \cdot S_i^{-1} \cdot S_j \cdot R(\alpha, \beta, \gamma). \quad (14)$$

Since $(S_i^{-1} \cdot S_j)$ must also be a symmetry element, S_k, of the function P_o, the denominator of (13) can be simplified to give

$$\left(\sum_{i=1}^{n} \sum_{j=1}^{n} Q(\alpha'_{ij}, \beta'_{ij}, \gamma'_{ij}) \right)^{1/2} = \left(n \sum_{k=1}^{n} Q(\alpha'_k, \beta'_k, \gamma'_k) \right)^{1/2}. \quad (15)$$

From (13), the final form for $N(\alpha, \beta, \gamma)$ is

$$N(\alpha, \beta, \gamma) \propto \frac{C(\alpha, \beta, \gamma)}{[\sum_{k=1}^{n} Q(\alpha'_k, \beta'_k, \gamma'_k)]^{1/2}} \quad (16)$$

where α'_k, β'_k and γ'_k are related to α, β and γ by

$$R(\alpha'_k, \beta'_k, \gamma'_k) = R^{-1}(\alpha, \beta, \gamma) \cdot S_k \cdot R(\alpha, \beta, \gamma). \quad (17)$$

Equations (16) and (17) show how N can be evaluated from the functions C and Q. The function C is a cross-rotation function, and Q can be computed from a single self-rotation function of the search model. In practice, α', β' and γ' may not correspond to grid points in the discretely sampled self-rotation function. In such cases, interpolation or approximation to the nearest grid point may be required.

3. Difference between the ordinary rotation function and the modified rotation function

The denominator of (16) is the symmetry correction term which relates the function N and the ordinary cross-rotation function C. This term is defined by the internal symmetry properties (Q) of the model and the symmetry (S) of the crystal, and is a function of orientation.

When $k = 1$, with S_1 the identity transformation, (12) and (17) show that the value of $Q(\alpha', \beta', \gamma')$ is given by the value of the self-rotation function of the model at the identity transformation. When no approximate symmetry exists in the model Patterson function, P_m, the other terms ($k \neq 1$) in the denominator of (16) will be small. In such a case, Q will be nearly independent of α', β' and γ', and N will be nearly proportional to C.

However, if P_m has some approximate internal symmetry, and can be oriented in such a way that it obeys a symmetry operation of P_o to a significant degree, then the value of C will be greater in magnitude than the value of N at that orientation. This result predicts that unusually large peaks (or valleys) may occur in ordinary rotation functions for orientations at which the model approximates the observed crystal symmetry.

The significance, or lack of significance, of the symmetry-correction term is dictated by the range of the denominator in (16). This range depends strongly on the particular problem, but values for a few examples are given in Table 1. The correction terms have been scaled to correspond to an origin peak of magnitude one in the self-rotation function of the model. For small models, and particularly for models with real or approximate symmetry, the upper limit of the correction term may approach the square root of n.

4. Examples

The presence of symmetry bias in cross-rotation functions is illustrated by the cross-rotation function between a polyalanine α-helix and melittin, a component of bee venom. Melittin is a 26-amino-acid polypeptide, containing two α-helices (Terwilliger & Eisenberg, 1982). Structure factors were calculated from an idealized polyalanine helix in space group $P1$ and from melittin in the space group in which it was crystallized ($C222_1$). Data between 10 and 2·5 Å were used, with a search radius of 14 Å. At this resolution, the helix resembles a radially symmetric

rod, and so the correct orientation of the helix about its long axis is difficult to determine from the cross-rotation function. In our calculation, this degree of freedom corresponds to the first Euler angle, α. Consequently, the angles β and γ specify the direction of a helical axis, but not the orientation about the axis. A section (of constant α) through a cross-rotation function between an α-helix and melittin is shown in Fig. 1. The section shows the true crystallographic symmetry, as well as approximate symmetry corresponding to orientations of incorrect polarity for the helix, and approximate radial symmetry. Two unrelated peaks, corresponding to the two helices in melittin,* are seen in the rotation function (Fig. 1). However, the peaks are significantly skewed and shifted from their expected positions to positions corresponding to orientations at which the helix axis would lie in a crystallographic plane. At such orientations, the ideal helix would approximately conform to one of the crystallographic twofold symmetry operations. The arguments of the previous sections have predicted bias towards these positions.

The symmetry-correction function for this problem was evaluated, and the corresponding section is shown in Fig. 2. Pronounced peaks appear in the directions of the crystallographic axes, connected by ridges in the crystallographic planes. Division of the ordinary cross-rotation map by the symmetry-correction map produces the map in Fig. 3. In this modified map, the peaks appear nearer to their expected positions, and the peaks become partially resolved from other peaks related by approximate symmetry.

A more general example, where there is no obvious symmetry in the search model, is provided by the cross-rotation function of tetragonal hen egg white lysozyme (Diamond, 1974), using the amino-terminal half of the protein (residues 1–60 of 129) as the search model. 'Observed' structure factors were calculated from the complete atomic model in space group $P4_32_12$, excluding the solvent molecules. Data between 10 and 2·5 Å were included and the Patterson search radius was 14 Å.

In an ordinary fast rotation function, the highest peak in the map is 4·79 standard deviations above the mean value and corresponds to the correct orientation for the search fragment. Taking into account the constants in the denominator of (1), this peak has a correlation coefficient of 0·214. In addition, five other distinct peaks exceed 2·5 standard deviations (see Table 2). A section through the asymmetric contents of the rotation function map is shown in Fig. 4(*a*). The correct peak and two other significant peaks are visible.

The symmetry-correction map was generated as previously described. The correction term ranges from

* The crystallographic asymmetric unit contains a melittin dimer with a non-crystallographic twofold nearly parallel to a crystallographic twofold axis. Therefore, peaks related by the non-crystallographic axis are nearly superimposed and indistinguishable in the rotation function.

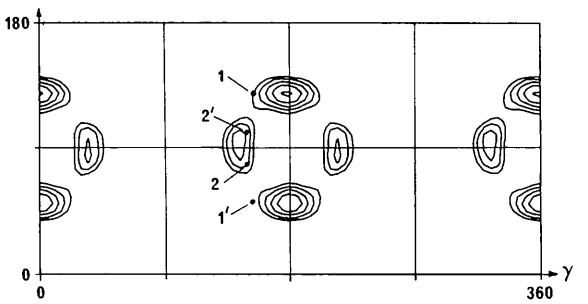


Fig. 1. Symmetry bias in the cross-rotation function between a polyalanine α-helix and melittin. The map is a section of constant α through the expected peak for the orientation of melittin helix no. 2. Due to approximate radial symmetry of the probe helix, the rotation function is nearly independent of α and peaks corresponding to helix no. 1 and to helices of wrong polarity are all visible in one section. The expected directions for the axes of helices nos. 1 and 2 are labeled. Orientations of incorrect polarity are labeled, and denoted by a prime ('). The remaining peaks are related by crystallographic symmetry. The correct peaks are not resolved from or distinguishable from peaks related by approximate symmetry and reverse polarity. The section is contoured at 1·2σ, 1·4σ, 1·6σ and 1·8σ.

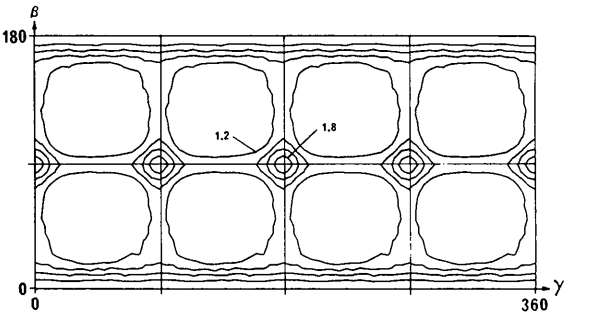


Fig. 2. Symmetry-correction map for the cross-rotation function between a polyalanine α-helix and melittin. Contours range from 1·2 to 1·8 in increments of 0·2.

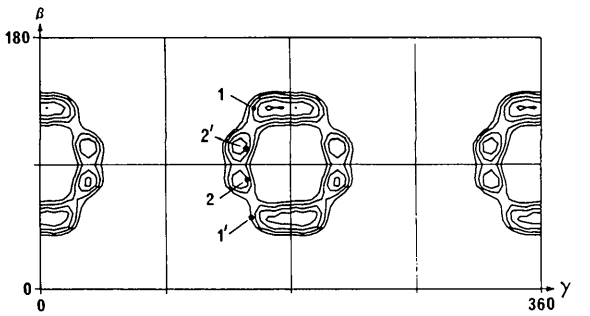


Fig. 3. Symmetry-corrected cross-rotation function between a polyalanine α-helix and melittin. The true helical directions for melittin helices nos. 1 and 2 are indicated. As before, the directions corresponding to incorrect polarity of the helices are denoted by a prime ('). The section is contoured at 1·2σ, 1·4σ, 1·6σ and 1·8σ, as in Fig. 1.

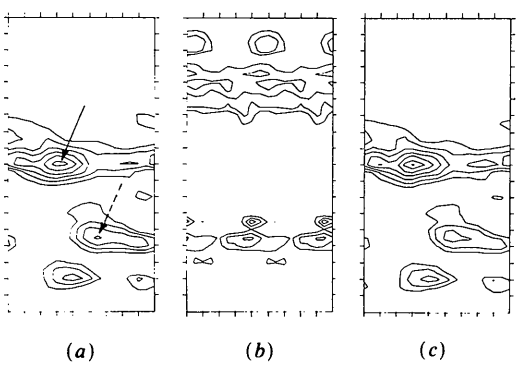
Table 2. *Peak statistics for the ordinary and modified cross-rotation functions*

Comparison of the peak statistics for the ordinary and the modified cross-rotation functions for tetragonal lysozyme, using half of the structure as the search model. See text. Peak no. 1 corresponds to the correct orientation of the model. Correlation coefficients are obtained from the rotation function values by evaluating the constant terms in the denominator of equation (1).

	Ordinary rotation function		Modified rotation function	
Peak no.	Height in standard deviations	Correlation coefficient [C in equation (1)]	Height in standard deviations	Correlation coefficient [N in equation (3)]
1	4·79	0·214	4·84	0·529
2	3·22	0·144	2·83	0·309
3	3·13	0·140	2·71	0·296
4	3·01	0·134	2·83	0·309
5	2·73	0·122	2·53	0·277
6	2·72	0·122	2·45	0·267

0·80 to 1·49. There is some coincidence between elevated regions in the correction map and some of the incorrect peaks in the cross-rotation function map. The corresponding section through this map is shown in Fig. 4(*b*).

A symmetry-corrected map is obtained by dividing the cross-rotation function map by the correction map. In this new map, the correct peak is slightly higher in terms of standard deviations (4·84 σ), and all five of the incorrect peaks are reduced in significance (see Table 2). All three of the peaks above three standard deviations in the ordinary rotation function map are reduced to below three standard deviations. No additional significant peaks are present in the new map. The corresponding section of the modified map is shown in Fig. 4(*c*), and a modest improvement in the quality of the map is apparent. More striking than the slight improvement in the quality of the map is the dramatic increase in the actual values of the correlation coefficient (Table 2). The value of the new correlation coefficient [N in (3)] for the correct peak is 0·529, which is reasonable for a search model that comprises half of the complete structure.*



(*a*) (*b*) (*c*)

Fig. 4. (*a*) Section of constant α through the cross-rotation function of tetragonal lysozyme, using half of the protein as a search model. β is horizontal and γ is vertical, with the axes marked in 10° increments. The correct peak is indicated by the solid arrow. The second highest noise peak in the map also lies in this section and is shown by the dashed arrow. (*b*) Corresponding section of the symmetry-correction map. (*c*) Corresponding section of the symmetry-corrected map. The correct peak is slightly higher than in the ordinary cross-rotation map and the highest noise peak in the section is slightly lower. See text. All of the maps are contoured at 1·5σ, 2·0σ, 2·5σ, ...

5. Concluding remarks

In an ordinary cross-rotation function, the observed Patterson function (or Patterson map) is compared with various orientations of a Patterson function (or interatomic vector set) of a single model molecule. The fact that the observed Patterson function may contain several superimposed symmetry-related intramolecular vector sets is usually ignored. Nordman (1986) has suggested the importance of accounting for these symmetry-related molecules in the cross-rotation function. The analysis presented in this paper validates this point, and elucidates the nature of the symmetry bias which occurs in an ordinary cross-rotation function. The bias results from interactions between the crystallographic symmetry and any real or approximate symmetry in the search model. The effect is most significant at orientations for which any approximate symmetry of the model resembles the crystallographic symmetry. The effect becomes increasingly significant with higher crystallographic symmetry, and more symmetrical search models. The analysis also demonstrates how the modified, or symmetry-corrected, cross-rotation function can be evaluated efficiently from two ordinary rotation functions, thereby allowing the use of 'fast rotation functions'.

Examples with test cases show that modest improvements in the quality of the cross-rotation function maps can be achieved. Furthermore, the correlation coefficient from the modified rotation function, N, is physically more meaningful than that

* A statistical argument suggests a value of $(60/129)^{1/2} = 0{\cdot}682$. The lower observed value of 0·529 can be attributed to the presence of intermolecular vectors in P_o.

of the ordinary rotation function, as all molecules present are accounted for. In fact, in the test case, the correlation coefficient is consistent with the fraction of the structure present in the search model. This symmetry-corrected rotation function may therefore provide a more objective measure of the reliability of possible rotation function solutions.

The ideas presented here also have implications for symmetry effects in self-rotation functions. Our analysis suggests that signal amplification will occur at orientations that leave the point-group symmetry of the crystal invariant. For example, in space group $P3$, we expect anomalously high peaks at orientations corresponding to 180° rotations about axes perpendicular to the threefold crystallographic symmetry axis.

Discussions with D. C. Rees, B. T. Hsu and C. E. Nordman are appreciated. This is publication no. 5597-MB of the Research Institute of Scripps Clinic.

References

BERNSTEIN, F. C., KOETZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535-542.

CARTER, D. C., MELIS, K. A., O'DONNELL, S. E., BURGESS, B. K., FUREY, W. F. JR, WANG, B.-C. & STOUT, C. D. (1985). *J. Mol. Biol.* **184**, 279-295.

CROWTHER, R. A. (1972). *The Molecular Replacement Method*, edited by M. G. ROSSMANN, pp. 173-178. New York: Gordon & Breach.

DIAMOND, R. (1974). *J. Mol. Biol.* **82**, 371-391.

HUBER, R. (1969). *Crystallographic Computing*, edited by F. AHMED, pp. 96-102. Munksgaard: Copenhagen.

KENDREW, J. C., DICKERSON, R. E., STRANDBERG, B. E., HART, R. G., DAVIES, D. R., PHILLIPS, D. C. & SHORE, V. C. (1960). *Nature (London)*, **185**, 422-427.

NORDMAN, C. E. (1986). Am. Crystallogr. Assoc., Hamilton, Ontario. *Program Abstr.* p. 36.

NORDMAN, C. E. & NAKATSU, K. (1963). *J. Am. Chem. Soc.* **85**, 353-354.

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

TERWILLIGER, T. C. & EISENBERG, D. (1982). *J. Biol. Chem.* **257**, 6016-6022.

WATSON, H. C. (1969). *Prog. Stereochem.* **4**, 299-333.

Acta Cryst. (1989). A**45**, 314-319

Cauchy Distribution, Intensity Statistics and Phases of Reflections from Crystal Planes

BY G. B. MITRA

CSS Department, Indian Association for the Cultivation of Science, Jadavpur, Calcutta 700032, *India*

AND SABITA DAS

Department of Physics, Victoria Institution (College), 78B Acharya Prafulla Chandra Road, Calcutta 700009, *India*

(*Received* 1 *May* 1987; *accepted* 2 *November* 1988)

Abstract

Recently near-Gaussian distributions have been of much interest in the field of crystallographic statistics. In the present work, expressions for a truncated Cauchy distribution corresponding to acentric and centric cases have been derived. Expressions for P_+, the probability of sign relations for centric crystals, and for P_φ, the probability of the tangent relationship for acentric crystals, have been derived on the basis of the Cauchy distribution of structure factor components. Theoretical $N(Z)$ curves for centric and acentric Cauchy distributions have been compared with those for acentric, centric and bicentric Gaussian distributions. The $N(Z)$ curve for the Cauchy acentric distribution follows closely that for the Gaussian acentric up to $Z = 0{\cdot}5$. It then takes an upward turn and surpasses the Gaussian bicentric curve at high Z values. A similar trend is shown by the $N(Z)$ curve for the Cauchy centric distribution after being approximately intermediate between the Gaussian centric and bicentric cases up to $Z = 0{\cdot}5$. The results of P_+ and P_φ have been compared with some known crystal data and the agreement is quite satisfactory for the cases studied.

1. Introduction

The intensity statistics introduced by Wilson (1949, 1950) and its further extension to the phase problem by Cochran & Woolfson (1955) and by Cochran (1955) were based on the hypothesis that the structure-factor components obey the Gaussian probability distribution law. Bertaut (1955, 1960), Klug (1958), Mitra & Belgaumkar (1973), Shmueli (1979), Shmueli & Wilson (1981) and others have used near-Gaussian functions like the Gram-Charlier and Edgeworth

0108-7673/89/050314-06$03.00