SCHERINGER, C. (1977). *Acta Cryst.* A**44**, 588-592.
SCHERINGER, C. & REITZ, H. (1976). *Acta Cryst.* A**32**, 271-273.
SFUN/LIBRARY (1989). Version 2.1. *Fortran Subroutines for Evaluating Special Functions.* Houston, Texas: IMSL.
SIEGEL, J., GUTIÉRREZ, A., SCHWEIZER, W. B., ERMER, O. & MISLOW, K. (1986). *J. Am. Chem. Soc.* **108**, 1569-1575.
STENKAMP, R. E. & JENSEN, L. H. (1984). *Acta Cryst.* A**40**, 251-254.
STEVENS, E. D., RYS, J. & COPPENS, P. (1977). *Acta Cryst.* A**33**, 333-338.
STEWART, R. F. (1968). *Acta Cryst.* A**24**, 497-505.
WILSON, A. J. C. (1976). *Acta Cryst.* A**32**, 781-783.
WILSON, A. J. C. (1979). *Acta Cryst.* A**35**, 122-130.
ZYGMUND, A. (1988). *Trigonometric Series*, 2nd ed., Vols. I & II Combined, Vol. II, pp. 248-250. Cambridge Univ. Press.

Acta Cryst. (1992). A**48**, 172-174

An Efficient Molecular-Replacement Translation Function Based on the Evaluation of Direct-Methods Phase Invariants

BY DAVID A. LANGS

Medical Foundation of Buffalo, Inc., Buffalo, New York 14203, *USA*

(*Received* 13 *May* 1991; *accepted* 12 *August* 1991)

Abstract

Traditional molecular-replacement translation functions are based on direct- or reciprocal-space correlations between the observed diffraction amplitudes and the calculated amplitudes and phases of the symmetry-related molecular transforms of the search fragment as a function of the displacement vector. An alternative method that has been described is based on evaluating a list of phase invariants as a function of the position of the search model in the unit cell and seeking those regions which satisfy the expectation value of these invariants as predicted by probability theory. As originally formulated, this procedure required the iterative computation of the phases and the evaluation of the list of invariants as the search model was stepped over the grid points defining the asymmetric portion of the unit cell. A new computational procedure is described whereby the values of the invariants are expressed solely as a function of the displacement vector **r** as a Fourier series that can be evaluated by a standard fast Fourier transform (FFT) without having to compute and insert the values of the phases based on the search model at each grid point.

Introduction

Translation functions are computational algorithms by which the true unit-cell location of a correctly oriented molecule or fragment, whose absolute position is unknown, may be determined. Various techniques to determine translation solutions exist and include vector search methods (Nordman & Schilling, 1970), electron-density convolutions (Rossmann & Blow, 1962), Patterson correlation functions (Vand & Pepinsky, 1956) and computing the crystallographic residual on a grid encompassing the searched space (Booth, 1945; Bhuiya & Stanley, 1964). An up-to-date and thorough review of the literature by Beurskens and co-workers is highly recommended to those who wish to become more familiar with these established methods (Beurskens, Gould, Bruins Slot & Bosman, 1987).

An alternative method that has been proposed and tested involves the evaluation of direct-methods phase invariants as a function of the position of the fragment in the unit cell (Fortier & Langs, 1979). These grid search procedures produced encouraging results for a number of different phase-invariant types, including both negative and positive quartets (Hauptman, 1974) as well as the three-phase seminvariants (Hauptman & Green, 1978). A figure of merit is computed at each grid point, based on an E-weighted sum of the cosine values of a particular invariant type, and the translation solution is expected to produce either a positive maximum or a negative minimum, depending on whether the group of invariants is expected to be positive or negative. For example, NQEST (DeTitta, Edmonds, Langs & Hauptman, 1975) is a figure of merit that is based on the expected negative value of quartets, $E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}}E_{\mathbf{m}}$, for which the magnitudes of the four main terms $E_{\mathbf{h}}$, $E_{\mathbf{k}}$, $E_{\mathbf{l}}$ and $E_{\mathbf{m}}$ are large, and the magnitudes of the three cross terms $E_{\mathbf{h}+\mathbf{k}}$, $E_{\mathbf{h}+\mathbf{l}}$ and $E_{\mathbf{h}+\mathbf{m}}$ are small, *e.g.* less than 0.70, where $\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{m}=\mathbf{0}$, $B=2|E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}}E_{\mathbf{m}}|/N$ and N is the number of equivalent atoms in the primitive unit cell.

$$\mathrm{NQEST}=\sum_{\mathbf{h},\mathbf{k},\mathbf{l}} B\cos(\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{\mathbf{l}}+\varphi_{\mathbf{m}})\Big/\sum_{\mathbf{h},\mathbf{k},\mathbf{l}} B. \qquad (1)$$

NQEST is evaluated from the phases computed from the search model at each point in the grid search and

0108-7673/92/020172-03$03.00

should produce a large negative minimum in the vicinity of the true location-vector solution. Similar expressions for the triple invariants, three-phase seminvariants and quartets having a positive expected value, as a consequence of the three cross terms being large, *e.g.* greater than 1.5, can be formulated. It will be shown that these phase-invariant cosines, which are a function of the phases $\varphi_{\mathbf{h}}$, may be completely expressed as a function of the displacement vector $\mathbf{r}$, independent of the values of the phases computed at each grid point, such that these translation-function figures of merit can be efficiently calculated as a FFT.

Analysis

Phased normalized structure factors, $E_{\mathbf{h}}$, are readily expressed in terms of the transforms of oriented molecular fragments, $E_{\mathbf{h}_j p}$, and displacement vectors, $\mathbf{r}_p$, which relate the arbitrary coordinates, $\mathbf{r}_{jp}$, of the transform to the crystallographic coordinates, $\mathbf{r}_j$, of the correctly positioned group by means of $\mathbf{r}_j = \mathbf{r}_p + \mathbf{r}_{jp}$. Thus, for a crystal structure possessing n equivalent space-group positions and m independent correctly oriented molecules in the asymmetric unit

$$E_{\mathbf{h}} = \sum_{j}^{n} \sum_{p}^{m} E_{\mathbf{h}_j p} \exp 2\pi i(\mathbf{h}_j \cdot \mathbf{r}_p + \mathbf{h} \cdot \mathbf{t}_j), \qquad (2)$$

where $\mathbf{h}_j = \mathbf{R}_j^t \cdot \mathbf{h}$, where $\mathbf{R}_j$ is the matrix operator, and $\mathbf{t}_j$ the translation component of the jth equivalent space-group position. Similar relationships exist for properly scaled $F_{\mathbf{h}}$ data, but in the context of evaluating direct-methods phase invariants, the $E_{\mathbf{h}}$ amplitudes are to be preferred. It can be shown that the three-phase structure products

$$E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}} = \sum_{j,j',j''}^{n} \sum_{p,q,u}^{m} E_{\mathbf{h}_j p} E_{\mathbf{k}_{j'} q} E_{\mathbf{l}_{j''} u} \times \exp 2\pi i(\mathbf{h}_j \cdot \mathbf{r}_p + \mathbf{k}_{j'} \cdot \mathbf{r}_q + \mathbf{l}_{j''} \cdot \mathbf{r}_u), \qquad (3)$$

where $\mathbf{h}+\mathbf{k}+\mathbf{l}=\mathbf{0}$, can be calculated, provided one has reliable estimates for the various phase components $\mathbf{h}_j \cdot \mathbf{r}_p$, which are generally unknown. Here, for the sake of a cleaner notation, the phase due to the translation components of the symmetry operators indicated in (2) are incorporated into the phase of the individual transforms. Various methods for estimating these phase invariants (3) utilizing correctly oriented molecular fragments for which the position vectors $\mathbf{r}_p$ are not explicitly known have been published elsewhere (Kroon & Krabbendam, 1970; Thiessen & Busing, 1974; Main, 1976; Langs, 1985). A similar expansion for four-phase quartet invariants is shown to be

$$E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}} E_{\mathbf{m}} = \sum_{j,j',j'',j'''}^{n} \sum_{p}^{m} E_{\mathbf{h}_j p} E_{\mathbf{k}_{j'} p} E_{\mathbf{l}_{j''} p} E_{\mathbf{m}_{j'''} p} \times \exp 2\pi i[(\mathbf{h}_j + \mathbf{k}_{j'} + \mathbf{l}_{j''} + \mathbf{m}_{j'''}) \cdot \mathbf{r}_p], \qquad (4)$$

where for the sake of simplicity only one molecular displacement vector $\mathbf{r}_p$ is considered. Phase-invariant translation functions that will produce a solution maximum when $\mathbf{r} = \mathbf{r}_p$ may be obtained by formulating these structure products as a function of the displacement vector $\mathbf{r}$,

$$E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}(\mathbf{r}) = \sum_{j,j',j''}^{n} \sum_{p}^{m} E_{\mathbf{h}_j p} E_{\mathbf{k}_{j'} p} E_{\mathbf{l}_{j''} p} \times \exp 2\pi i\, [(\mathbf{h}_j + \mathbf{k}_{j'} + \mathbf{l}_{j''}) \cdot \mathbf{r}], \qquad (5)$$

and multiplying the computed value of $E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}(\mathbf{r})$ by the fixed structure product magnitude $|E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}|$ observed for the data.

$$\mathrm{MF}(\mathbf{r}) = \sum_{\mathbf{h},\mathbf{k}} S_{\mathbf{hkl}} |E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}| E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}(\mathbf{r}) \qquad (6)$$

where $S_{\mathbf{hkl}}$ is the expected sign of the invariant (+1 for triples and positive quartets and −1 for negative quartets). This function can be evaluated for the full range of values for $\mathbf{r}$ as a FFT, thus allowing all cosine-invariant values to be efficiently computed over the permitted subspace of the cell, rather than over the values of the generated sets of amplitudes $|E_{\mathbf{h}}(\mathrm{calc.})|$ and phases $\varphi_{\mathbf{h}}$ computed at each grid point, *i.e.*

$$\mathrm{MF}(\mathbf{r}) = \sum_{\mathbf{h},\mathbf{k}} S_{\mathbf{hkl}} |E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}|\, |E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{l}}(\mathrm{calc.})| \times \cos(\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}}), \qquad (7)$$

a computation that is extremely inefficient.

Results

Phase-invariant translation functions were evaluated using data that were measured for the orthorhombic $P2_12_12_1$ crystal structure of the dodecadepsipeptide isoleucinomycin (Pletnev, Galitskii, Smith, Weeks & Duax, 1980) : $C_{60}H_{102}N_6O_{18}$, $a = 11.516$, $b = 15.705$, $c = 39.310$ Å, $Z = 4$, $N = 84 \times 4$. The 700 largest E values were used to compile a list of 20 000 triples, 10 000 positive quartets (cross terms greater than 1.15) and 8000 negative quartets (cross terms less than 0.70). The phase-invariant functions for these three classes of invariants were computed by the FFT algorithm at a grid interval of ~0.25 Å for search fragments of various sizes (42, 21, 15 and 10 atom fractions of the 84 atom structure) and the peaks from the corresponding maps were ranked according to size and compared. In addition, it was decided to compute composite Fourier maps representing the weighted average of the coefficients used in the positive and negative quartet syntheses. The results are presented in Table 1.

An inspection of the results in Table 1 indicates that the phase-invariant translation function can produce useful information for search fragments that represent as little as 20% of the structure. The rank

Table 1. *Relative peak height and rank of phase-invariant translation-function solutions for fragments of varying sizes*

Phase-invariant translation-function analysis of isoleucinomycin using triples, positive quartets, negative quartets and a combination of both positive and negative quartets. Column one notes the number of atoms out of 84 that describe the relative size of the molecular fragment. The columns labeled rank give the position of the solution vector in the magnitude-sorted peak list and ratio gives the relative size of the solution vector as compared to the largest spurious peak.

	Triples		Positive quartets		Negative quartets		Positive and negative quartets	
N/84	Rank	Ratio	Rank	Ratio	Rank	Ratio	Rank	Ratio
42	1	1.55	1	1.49	1	1.18	1	1.84
21	1	1.45	1	1.06	11	0.84	1	1.41
15	6	0.93	25	0.68	16	0.63	2	0.91
10	25	0.69	46	0.54	29	0.63	12	0.71

and relative size of the solution vector tends to decrease as the fragment becomes smaller, but the patterns of spurious peaks produced by the triples and negative quartets, or the positive and negative quartets, do not appear to be similar. As a consequence, the maps produced by these syntheses may be added, and the magnitude of the solution vector will be enhanced as the spurious peaks do not reinforce one another. This may be seen in Table 1 by inspecting the improvement in the peak rank and intensity ratio shown by the joint positive- and negative-quartet analysis as compared to the results produced by the positive and negative quartets separately.

The computational efficiency of the FFT algorithm as compared to the traditional methods of phase-invariant evaluation over the grid of the cell is impressive. The three-dimensional analysis for the structure which is presented in Table 1 required less than 10 min on a VAX 8600 computer system. The traditional method would require 2 or 3 d, a time saving on the order of 500-fold.

Research supported in part by NIH grant GM32812.

References

BEURSKENS, P. T., GOULD, R. O., BRUINS SLOT, H. J. & BOSMAN, W. P. (1987). *Z. Kristallogr.* **179**, 127–159.

BHUIYA, A. K. & STANLEY, E. (1964). *Acta Cryst.* **17**, 746–748.

BOOTH, A. D. (1945). *Philos. Mag.* **36**, 609–615.

DETITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* A**31**, 472–479.

FORTIER, S. & LANGS, D. A. (1979). Proc. Am. Crystallogr. Assoc. Meet., Boston, MA. Abstract PB31.

HAUPTMAN, H. (1974). *Acta Cryst.* A**30**, 472–476, 822–829.

HAUPTMAN, H. & GREEN, E. A. (1978). *Acta Cryst.* A**34**, 224–229.

KROON, J. & KRABBENDAM, H. (1970). *Acta Cryst.* B**26**, 312–314.

LANGS, D. A. (1985). *Acta Cryst.* A**41**, 583–586.

MAIN, P. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 97–105. Copenhagen: Munksgaard.

NORDMAN, C. E. & SCHILLING, J. W. (1970). *Crystallographic Computing*, edited by F. R. AHMED, S. R. HALL & C. P. HUBER, pp. 110–114. Copenhagen: Munksgaard.

PLETNEV, V. Z., GALITSKII, N. M., SMITH, G. D., WEEKS, C. M. & DUAX, W. L. (1980). *Biopolymers* **19**, 1517–1534.

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

THIESSEN, W. E. & BUSING, W. R. (1974). *Acta Cryst.* A**30**, 814–821.

VAND, V. & PEPINSKY, R. (1956). *Z. Kristallogr.* **108**, 1–14.

Acta Cryst. (1992). A**48**, 174–180

The Generalized Debye–Waller Factor for Atoms in a Twisted Environment. I. Sites at the Core of a Double Helix

BY W. PRANDL

Institut für Kristallographie, Universität Tübingen, Charlottenstrasse 33, 7400 *Tübingen, Germany*

AND F. DUNSTETTER

Laboratoire Léon Brillouin, CEA-CNRS, CEN Saclay, 91191 *Gif-sur-Yvette CEDEX, France*

(*Received* 4 *April* 1991, *accepted* 12 *August* 1991)

Dedicated to Professor H. Jagodzinski on his 75th birthday

Abstract

Lattice sites which have a proper or improper subgroup of 222 as the site group may exhibit an anharmonic twisted local potential. The generalized Debye–Waller factor for atoms occupying such sites is derived. If the twist axis is fixed by symmetry considerations the approximation used needs, in addition to three harmonic parameters, one anharmonicity parameter, essentially the local pitch. In the most general case with no symmetry restrictions three Eulerian angles are needed in addition to the pitch.

0108-7673/92/020174-07$03.00