points, using a multiplicative group M, that produces a matrix of a particularly nice form. Further, it restricts to an asymmetric unit for $P3$ on the space reducing the size of the matrix by $1/9$. Furthermore, the resulting matrix has the same tensor skew-circulant form as the big matrix and so can be used to compute the transform using the same efficient algorithms as the original transform.

As we have suggested earlier, the technical details in the resulting algorithm will be the subject of another paper where the construction of the asymmetric unit and all the blocks will be shown.

Summary

The finite Fourier transform on N data points is simply the evaluation of an $N \times N$ matrix times an N-vector to produce an N-vector result. The straightforward method of matrix multiplication requires a number of operations proportional to N^2. In terms of computer programming, the time required to compute the result, say $T_S(N)$, would quadruple if the input size were doubled. We can express this by writing $T_S(N) = CN^2$ where C is a constant depending, amongst many things, primarily on the machine on which the program is run and the coding of the program. Typical values of C are in the range of 40 μs for the VAX-11/785 to 400 ns for the CRAY X-MP. So even for moderate-size problems, say $N = 100$ to 10 000, the range of observed times is significant; measured in seconds to hours.

The remarkable aspect of the Fourier transform is that there exist 'fast' or 'efficient' methods which do the same evaluation in a time proportional to $N \log N$, or $T_F(N) = KN \log N$ where K is another constant, approximately the same as C. This means that using 'fast' methods reduces time costs for real problems by several orders of magnitude.

For problems in crystallography, the finite Fourier transform is run many times for the same space group on the same number of points. This problem has led to efforts to use the symmetry of the data to reduce N by the order of the group to save time and space in calculating the result. For groups built from $P2$ Ten Eyck (1973) was able to achieve both. Bantz & Zwick (1974) were able to use symmetry to reduce memory requirements for nearly all space groups.

The advantage of the approach presented in this paper is that we can use symmetry to reduce the data and still use an efficient $N \log N$ evaluation method. Although much work needs to be done to develop algorithms for all the space groups, the general method presented here shows that such algorithms exist. In principle, each group and each grid size leads to a different program. However, the methods presented in this paper enable us to generate these programs automatically. Moreover, by the very nature of the algorithms developed they can be naturally partitioned to calculate structures larger than available high-speed memory.

References

BANTZ, D. A. & ZWICK, M. (1974). *Acta Cryst.* A30, 257–260.
International Tables for Crystallography (1983). Vol. A, pp. 94, 480. Dordrecht: Reidel. (Present distributor Kluwer Academic Publishers, Dordrecht.)
RADER, C. M. (1968). *Proc. IEEE*, 5, 1107–1108.
TEN EYCK, L. F. (1973). *Acta Cryst.* A29, 183–191.
WINOGRAD, S. (1978). *Math. Comput.* 32, 175–199.

Increasing the Size of Search Fragments for use in Patterson Method Calculations – the Partial-Fragment Rotation Function

BY C. C. WILSON

Neutron Division, Rutherford Appleton Laboratory, Chilton, Didcot, Oxon OX11 0QX, England

Abstract

A method is described of expanding a molecular fragment for use in Patterson search procedures by the rotation of part of a model about a bond direction with respect to a fixed fragment, allowing the removal of an important degree of freedom in the model. The function has been incorporated into a computer program and it has been found possible to orient very small partial fragments in this way. The consequent expansion of a search model should assist in structural solution.

Introduction

The basis of Patterson search techniques lies in the provision of a reasonable model fragment for comparison with the observed data. In the reciprocal-

space calculations performed in the methods of Tollin & Cochran (1964); Tollin (1976); Wilson & Tollin (1985), the rotation function is calculated using the expression

$$\sigma(\theta_1, \theta_2, \theta_3)$$

$$= \sum_h |F_h^s|^2 \left[\sum_i (\cos 2\pi h.r_i)^2 + \sum_i (\sin 2\pi h.r_i)^2 \right] \quad (1)$$

where the $|F_h^s|^2$ are the sharpened squared structure factors, the r_i are the coordinates of the model, $(\theta_1, \theta_2, \theta_3)$ are the Eulerian angles of the rotation (Rossmann & Blow, 1962) and the summation is normally over the N_{ref} largest $|F_h^s|^2$, with $N_{ref} \sim 300$ in most cases. This function is closely related to the sum over all reciprocal-lattice points of the product of the observed $|F_h|^2$ with the model molecular transform. The first term in (1) can be regarded as the Fourier transform of the observed Patterson function, $P(r)$, $FT[P(r)_{obs}] = |F|^2$, and the second term as the Fourier transform of the calculated Patterson {FT $[P(r)_{calc}] = FT[\rho(r) * \rho(-r)] = \sum_j (f_j \cos 2\pi h.r_j)^2 + \sum_j (f_j \sin 2\pi h.r_j)^2$}, excluding the atomic scattering factors, where $*$ represents convolution. The summation (1) can therefore be regarded as a Patterson search, but without explicit evaluation of the Patterson function at any point, instead carrying out all calculations in reciprocal space.

It is obvious that in order to improve the reliability of the rotation function $\sigma(\theta_1, \theta_2, \theta_3)$, one must make both terms in the summation as reliable as possible. The $|F_h^s|^2$ are made reliable by good experimental and sensible sharpening procedures and the model transform is made most reliable by inclusion of as much correct stereochemical information as possible. This argument holds also for subsequent stages in the Patterson method calculations such as the Q (translation) functions (2) (Tollin & Cochran, 1964; Tollin, 1966),

$$Q(R_0) = \sum_h |F_h^s|^2 \sum_{j,j'=1}^{n} \cos 2\pi h.[r_j + R_0 - T(r_{j'} + R_0)] \quad (2)$$

where $T(r)$ is a symmetry transformation, in which again the provision of the best model for the second term would improve the reliability of the calculation.

The purpose of this paper is to indicate how the second of these requirements can be met more satisfactorily and provide better models for rotation and translation searches using Patterson methods.

The partial-fragment rotation function

In many crystal structures the situation can arise whereby two (or more) groups whose stereochemistries are reasonably well known are linked *via* an undetermined degree of freedom, for example through a bond, the torsion angle of the two groups

relative to one another about that bond being unknown. Two typical examples of this are in the biphenyl structures and in nucleoside structures where a planar base is connected to a puckered sugar group through the glycosidic N–Cl′ bond, but many more such structures exist.

If the orientation of one of these groups can be found, then if it were possible to eliminate the degree of freedom about the linking bond, the model for use in Patterson search calculations could then be extended to encompass both groups.

Essentially this amounts to writing the rotation function (1) in the form

$$\sigma(\theta_1, \theta_2, \theta_3, \theta_p)$$

$$= \sum_h |F_h^s|^2 \left[\sum_{i=1}^{n_1} (\cos 2\pi h.r_i)^2 + \sum_{i=1}^{n_1} (\sin 2\pi h.r_i)^2 \right]$$

$$+ \sum_h |F_h^s|^2 \left[\sum_{j=n_1}^{n_2} (\cos 2\pi h.r_j)^2 \right.$$

$$\left. + \sum_{j=n_1}^{n_2} (\sin 2\pi h.r_j)^2 \right] \quad (3)$$

where the n_1 atoms i $[r_i = r_i(\theta_1, \theta_2, \theta_3)]$ are those in the group whose orientation is known and fixed, and the $(n_2 - n_1)$ atoms j $[r_j = r_j(\theta_1, \theta_2, \theta_3, \theta_p)]$ are those in the partial fragment whose orientation with respect to the n_1 fixed atoms is unknown.

Thus the $(n_2 - n_1)$ atoms j have an additional partial rotation angle (representing the extra degree of freedom) θ_p about the relevant bond direction which is applied before imposition of the standard Eulerian rotation $(\theta_1, \theta_2, \theta_3)$ leading to calculation of the rotation function. The purpose of the partial-fragment rotation function (PFRF) would then be to find the value of θ_p by which the $(n_2 - n_1)$ atoms j should be rotated to give a maximum value in the modified rotation function (3).

While this procedure is very straightforward there is a possible problem related to the fact that even in the standard rotation function (1) the significant features pertaining to orientational information sit on a very high background value, and this situation is exacerbated by the inclusion of the $(n_2 - n_1)$ extra atoms in the search model. It is obviously of some interest to discover for how few additional atoms the PFRF will produce significant and correct structural information.

The PFRF calculation (3) was incorporated into the Patterson methods program *PATMET* (Wilson & Tollin, 1986), being implemented *via* extra coordinate transformations sitting on top of the existing rotation-function calculations. Clashes between the rotating and the fixed fragment can eliminate possible conformational regions from consideration and this is incorporated in the code. The results of several tests of the procedure are outlined below.

Location of the orientation of a planar group is standard procedure in *PATMET* calculations (Tollin & Cochran, 1964; Wilson & Tollin, 1985), and a facility is provided whereby a published set of coordinates can be supplied which are then transformed and used in the search procedure. The code can deal with any combination of oblique or orthogonal coordinate systems in the structure and the model. Published structural information can also be used for the partial fragment on which the θ_p operation is to be performed.

Tests of the function

Since the work outlined here was prompted by the particular problem of orienting puckered ribose and deoxyribose sugar groups relative to planar bases in nucleosides, a selection of these structures was used to test the PFRF.

Table 1 summarizes the results of the tests and Fig. 1 shows a typical plot of the partial-fragment rotation function for the example of 2',3',5'-tri-*O*-acetylguanosine (TAG) using 2',3',5'-tri-*O*-acetyladenosine (TAA) as model. Inset in Fig. 1 are pictures of the molecules showing the rotation being performed in this calculation. The peak at ~128° is close to the expected solution, which for these two structures should give a value equal to the difference in the χ (glycosyl) torsion angles, which is 134°.

The first group of calculations was used to verify the potential applicability of the procedure; these were tests using as a model parts of the correct structure of each molecule. All results from these tests were fairly hopeful, although the problems inherent in rotation-function calculations using small fragments can be seen in the fact that even in these 'idealized' cases the solutions are not always unambiguous. The quantities p_1^2 and p^2, representing the scattering power of the partial fragment with respect to the asymmetric unit and the unit-cell contents respectively, give some indication of the relative sizes of the fragments whose orientations are being determined.

In the second group more realistic tests of the procedure were performed, using published related structures or in one case an ideal-geometry-generated structure as models. In all cases to date the correct θ_p value has been indicated, mostly as the best solution in the PFRF, or failing this as one of the next few highest.

As can be seen from Table 1, the models being oriented in the PFRF calculations tends to be rather small and so these results are certainly impressive. Indeed in the extreme case where just a C–O fragment was being oriented relative to a $C_9N_5O_2$ fixed group within a $C_{16}N_5O_7$ structure, the high ranking of the correctly oriented solution is quite remarkable.
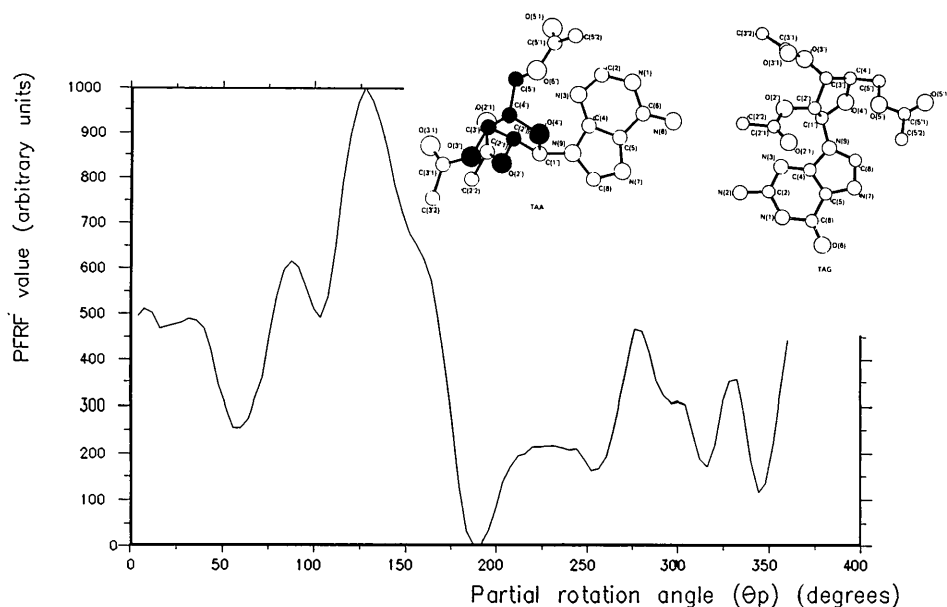

Fig. 1. Plot of the partial-fragment rotation function for TAG using TAA as model, showing the peak at 128°. Inset left: view of the molecular structure of TAA. Inset right: view of the molecular structure of TAG. The partial rotation was performed about the N(9)–C(1') bond, in this case using the blocked atoms [C(2'), O(2'), C(3'), O(3'), C(4'), C(5'), O(4')] as partial fragment. The χ angle in these structures is defined as the C(4)–N(9)–C(1')–O(4') torsion angle (see text).

Table 1. *Summary of tests of the partial-fragment rotation function*

Structure	Model	Fixed fragment	Rotating partial fragment	Partial fragment p_1^2	p^2	Solution (height/height of next highest peak)
TAG*	TAG	C_6N_5O	C_4O_3	0·25	0·12	1st (*998/525*)
$(C_{16}N_5O_8)$	TAG	C_6N_5O	C_3O	0·13	0·07	1st (*999/396*)
TAA†	TAA	C_6N_5	C_9O_7	0·59	0·15	1st (*999/883*)
$(C_{16}N_5O_7)$	TAA	C_6N_5	C_3O	0·14	0·03	3rd (*909/999*)
	TAA	C_6N_5	C_4O_3	0·26	0·06	4th (*792/999*)
	TAA	C_6N_5	CO	0·07	0·02	2nd (*918/999*)
MAT‡	MAT	$C_6N_2O_2$	C_3O	0·19	0·05	1st (*999/714*)
$(C_{12}N_2O_6)$						
TAU§	IDU¶	$C_5N_2O_2$	C_4O_2	0·23	0·06	2nd (*997/999*)
$(C_{15}N_2O_9)$	4TU**	$C_5N_2O_2$	C_4O_4	0·32	0·08	3rd (*704/998*)
TAG	TAA	C_6N_5	C_4O_4	0·29	0·14	1st (*999/694*)
$(C_{16}N_5O_8)$	TAA	C_6N_5	C_4O_3	0·25	0·12	1st (*999/616*)
TAA	COORD††	C_6N_5	C_3O	0·14	0·03	3rd (*711/998*)
$(C_{16}N_5O_7)$						
MAT	AT‡‡	$C_6N_2O_2$	C_4O_3	0·36	0·09	1st (*999/614*)
$(C_{12}N_2O_6)$	AT	$C_6N_2O_2$	C_3O	0·19	0·05	2nd (*973/999*)

* 2′,3′,5′-Tri-*O*-acetylguanosine: Wilson, Low & Tollin (1985).
† 2′,3′,5′,-Tri-*O*-acetyladenosine: Wilson, Tollin & Howie (1986).
‡ 3′-*O*-Acetylthymidine: Eccleston, Wilson & Howie (1988).
§ 2′,3′,5′-Tri-*O*-acetyluridine: Low & Wilson (1984).
¶ 5-Iodo-2′-deoxyuridine: Camerman & Trotter (1965).
** 4-Thiouridine: Saenger & Scheit (1970).
†† Model generated by *COORD* program: Bell, S. (1987). Private communication.
‡‡ 3′,5′-Di-*O*-acetylthymidine: Wilson, Low, Tollin & Wilson (1984).

Concluding remarks

From the results of these tests, the accuracy of the models obtained from the PFRF tends to be ≤10° in θ_p, which is quite adequate for provision of useful information to Patterson search procedures.

Obviously it is rather disappointing that the function does not produce exactly the correct θ_p orientation in every case as the highest value, but when one considers the size of the models being oriented the performance of the function must be regarded as reasonably impressive.

For the crystallographer who has recourse to Patterson methods for difficult structures for either structural solution or to give a model for *ab initio* provision of information to direct methods, every expansion of the search model is important. Even in the worst cases, when the PFRF merely reduces the possible values of a degree of freedom to just a few small regions, this information will significantly reduce the number of possible models to be tried, and hence should assist in structural solution. In the best cases, of course, the model available is immediately expanded and the probability of a solution being achieved increases concomitantly.

References

CAMERMAN, N. & TROTTER, J. (1965). *Acta Cryst.* 18, 203-211.
ECCLESTON, R. S., WILSON, C. C. & HOWIE, R. A. (1988). *Acta Cryst.* C44. In the press.
LOW, J. N. & WILSON, C. C. (1984). *Acta Cryst.* C40, 1030-1032.
ROSSMANN, M. G. & BLOW, D. M. (1962). *Acta Cryst.* 15, 24-31.
SAENGER, W. & SCHEIT, K. H. (1970). *J. Mol. Biol.* 50, 153-169.
TOLLIN, P. (1966). *Acta Cryst.* 21, 613-614.
TOLLIN, P. (1976). In *Crystallographic Computing Techniques*, pp. 212-221, edited by F. R. AHMED, K. HUML & B. SEDLACEK. Copenhagen: Munksgaard.
TOLLIN, P. & COCHRAN, W. (1964). *Acta Cryst.* 17, 1322-1324.
WILSON, C. C., LOW, J. N. & TOLLIN, P. (1985). *Acta Cryst.* C41, 1123-1125.
WILSON, C. C., LOW, J. N., TOLLIN, P. & WILSON, H. R. (1984). *Acta Cryst.* C40, 1712-1715.
WILSON, C. C. & TOLLIN, P. (1985). *PATMET. A Program for the Determination of Orientation and Position of a Known Fragment Within the Unit Cell.* Users Manual. Univ. of Dundee, Scotland.
WILSON, C. C. & TOLLIN, P. (1986). *J. Appl. Cryst.* 19, 411-412.
WILSON, C. C., TOLLIN, P. & HOWIE, R. A. (1986). *Acta Cryst.* C42, 697-700.