Acta Cryst. (1981). A37, 642–644

On The Application of Phase Relationships to Complex Structures. XVIII. *RANTAN*–Random *MULTAN*

BY YAO JIA-XING*

Department of Physics, University of York, Heslington, York YO1 5DD, England

(*Received* 20 *October* 1980; *accepted* 19 *February* 1981)

Abstract

To overcome the disadvantages of a small starting set in *MULTAN* a technique is described whereby a large number of phases are given random values with low weights and then refined by a weighted tangent formula. The effectiveness of the process depends quite critically on assigning suitable weights to phase estimates and on a fairly tight control of the refinement. Examples of the application of the method are given for a number of structures, many of which are difficult to solve by direct methods.

Introduction

Multisolution direct methods, *e.g. MULTAN* and *SHELX*, are very successful in the automatic solution of crystal structures. These methods are in a state of steady development and *MULTAN*-80 (Main, Fiske, Hull, Lessinger, Germain, Declercq & Woolfson, 1980), the latest available version, not only includes a number of new features but contains very sophisticated decision-making routines (Main, 1978), so that it is normally run on default values for the various program parameters.

An important reason for the occasional failure of multisolution methods is the small size of the starting set. The latest versions of *MULTAN* and *SHELX* use the magic-integer concept to extend the size of the starting set but even so, in the early stages of the phase-development process, the validity of comparatively few triple-phase relationships control the success, or otherwise, of the whole exercise. It was to overcome this difficulty that the programs *MAGIC* and *YZARC* (Declercq, Germain & Woolfson, 1979) were developed. In these one treats simultaneously from 40 to 100 reflexions and the hundreds of relationships which link them, so there is less dependence on the validity of a few particular relationships.

The *YZARC* method (Baggio, Woolfson, Declercq & Germain, 1978) involves the assignment of random phase values to about 100 reflexions and the refinement of these values by the least-squares solution of a set of linear equations derived from the phase relationships. The proposal here is similar except in two important respects. Firstly, it is possible simultaneously to work with all the reflexions whose phases are needed and all the relationships which link them from the very beginning. Thus one may assign random phases to, say, 300 reflexions although, initially, these have low weights associated with them. Secondly, the refinement process is carried out either by the weighted tangent formula (WTF) (Main, Hull, Lessinger, Germain, Declercq & Woolfson, 1978), or by the statistically-weighted tangent formula (SWTF) (Hull & Irwin, 1978), both of which are available from *MULTAN*-80.

The weighting and refinement schemes

In the present method, called *RANTAN*, weights are allocated to the initial phase allocations as follows: origin-fixing reflexion, weight = 1·00; enantiomorph-fixing reflexion with special value, weight = 0·99; enantiomorph-fixing reflexion with general value, weight = 0·85; random phase, weight = 0·25. Any other 'known' phase, for example a $\sum_1$-determined phase, is included with weight $2P - 1$, where P is the probability of the indicated special value. This is the same as the weight used in *MULTAN*-80.

This system of weights, and especially the weight for a random phase, has been determined empirically; the method is usually not particularly sensitive to the weights used as long as they have approximately the values suggested here. However, later we shall describe an application where a different weight assignment was essential for success.

In the *RANTAN* method the initial random phase and its weight of 0·25 are not changed until a phase estimate is obtained with a new weight greater than 0·25. Then, and only then, is the phase allowed to vary and to follow its refinement path.

Experience has shown that sometimes one or other of the two tangent-formula procedures, WTF and

* Permanent address: Institute of Biophysics, Academia Sinica, Peking, People's Republic of China.

0567-7394/81/050642-03$01.00

SWTF, is effective and the other is not. Where there is an enantiomorph-defining problem then the SWTF is to be preferred. Normally WTF is used, because it is somewhat faster, but if it is not successful then another run is made with the SWTF.

Once the refinement is complete then the usual *MULTAN* figures of metit are computed and another trial set of random phases may then be allocated and refined in the same way. For all the structures so far examined a hundred trial sets of phases or less have been developed.

Some applications to trial structures

In Table 1 are shown some of the results of applying *RANTAN* to solve crystal structures. These twelve examples have been selected to illustrate the range of space groups to which *RANTAN* has been successfully applied. Their names are given in coded form but they represent a wide variety of space groups, structural complexity and difficulty as far as direct methods are concerned. Every one of these structures was solved by *RANTAN*, with all or most of the structure shown in the first *E* map. Some comments about some of these structures are now given.

MUNICH-1: dibenzo[*i,l*]pentacyclo[6.2.2.$0^{2,6}$.$0^{2,7}$.$0^{3,7}$]dodeca-9,11-diene (Szeimies-Seebach, Harnisch, Szeimies, Van Meerssche, Germain & Declercq, 1978).

This structure, with 40 atoms in the asymmetric unit and space group *C*2, can only be solved by *MULTAN* with great difficulty. The best set from *RANTAN*, which shows 38 of the atoms, has figures of merit: ABSFOM 1·0093; PSIZERO 1·262; RESID 19·93; CFOM 2·8503.

RR: 3,3-dimethyl-4,5,9,10,11,12-hexa(carboxymethyl)-tetracyclo(7.2.1.$0^{2,4}$.$0^{2,8}$)dodeca-5,7,10-triene (Declercq, Germain & Henke, 1973).

The *CONVERGENCE* program in *MULTAN* chose three reflexions, 981, 10,5,3 and 133, for fixing the origin and enantiomorph. The general reflection 133 must be allocated two values of phase, 45 and 315°, according to the *MULTAN* origin-fixing scheme, and this was done in *RANTAN*. Thus pairs of trials (1, 2), (3, 4), *etc.* began with the same phases, including random phases, but with changes in the phase value of 133. For each of the pair of trials (21, 22) *RANTAN* gets the same solution, in other words, the refinement of the phases sometimes depends on the pattern of random phases rather than the fixing of one phase.

APAPA: adenylyl-(3′,5′)-adenylyl-(3′,5′)-adenosine hexahydrate (Suck, Manor & Saenger, 1976).

This again is a structure which cannot be solved by a straightforward application of *MULTAN*. *RANTAN* clearly showed in an *E* map the two fragments containing two phosphorus atoms and 39 other atoms. All the other non-hydrogen atoms were found by a weighted Fourier synthesis.

AZET: 3-chloro-1,3,4-triphenyl-2-azetidinone (Colens, Declercq, Germain, Putzeys & Van Meerssche, 1974).

This is a difficult structure to solve because of the enantiomorph problem, where a false centre of symmetry is found in the *E* maps, which makes their interpretation almost impossible. The first application of *RANTAN* with the SWTF did not solve the structure. It was decided to increase the weight for a random phase to 0·45. The rationale for this is that, clearly, the fixing of a few phases at explicit values was not fixing the enantiomorph too well but if, by chance, the pattern of random phases favoured one enantio-

Table 1. *Some of the crystal structures solved by RANTAN*

Z number of molecules in the unit cell; NOA number of atoms in the asymmetric unit; NOT number of trials; NSS number of correct solutions; NAE number of atoms from the first *E* map; NOR number of reference.

Name	Formula	Space group	*Z*	NOA	NOT	NSS	NAE	NOR
SEVIL	$C_{10}H_{13}NO_3$	$P\bar{1}$	2	14	64	1	14	(1)
TUR-10	$C_{15}H_{24}O_2$	$P6_322$	12	17	64	1	17	(2)
PYRBEN	$C_{22}H_{14}O_2$	$P4_1$	4	24	64	1	24	(3)
INOS	$C_6H_{12}O_6.H_2O$	$P2_1/n$	8	26	64	3	26	(4)
MUNICH-4	$C_{27}H_{22}O$	Cc	4	28	100	1	23	(5)
TUR-11	$C_{15}H_{24}O_4$	$P2_1$	4	38	32	1	37	(6)
RR	$C_{26}H_{28}O_{12}$	$Pn2_1a$	4	38	64	6	36	(7)
MUNICH-1	$C_{20}H_{16}$	$C2$	8	40	100	1	38	(8)
AZET	$C_{21}H_{16}ClNO$	$Pca2_1$	8	48	100	1	30	(9)
ERGO	$C_{28}H_{44}O$	$P2_12_12_1$	8	58	48	1	45	(10)
APAPA	$C_{30}H_{35}N_{15}O_{16}P_2.6H_2O$	$P4_12_12$	8	69	64	1	41	(11)
SCHWZ2	$C_{46}H_{70}O_{27}$	$P1$	1	73	100	1	50	(12)

References: (1) Conde, Lopez Castro, Marquez, Declercq & Germain (1979); (2) unpublished; (3) Bernstein *et al.* (1975); (4) unpublished; (5) Szeimies-Seebach, Szeimies, Van Meerssche, Germain & Declercq (1979); (6) Brackman *et al.* (1981); (7) Declercq *et al.* (1973); (8) Szeimies-Seebach *et al.* (1978); (9) Colens *et al.* (1974); (10) Hull, Leban, Main, White & Woolfson (1976); (11) Suck *et al.* (1976); (12) Schweizer (1980).

morph then one could encourage the development of this solution by increasing the weights of the random phases. Experiments have shown that the optimum value of weight for the initial random phases varies from structure to structure. The value 0·25 seems to work for most structures but the device of varying this value is available if *RANTAN* is unsuccessful.

Of the other structures in Table 1, other than those described above, ERGO and SCHWZ2 are particularly difficult to solve by previous direct methods.

General comments

RANTAN shares characteristics with both the *YZARC* and *MULTAN* procedures but, on balance, it has some advantages over both these other methods. Although its performance varies from structure to structure it seems to be at least as powerful as *MULTAN* and is probably more powerful. The disadvantage it possesses, compared with *MULTAN*, is that it takes somewhat more computing time (see Table 2). If the early-figure-of-merit (EFOM) idea is used, as in *MULTAN*-80, so that unlikely sets of phases can be weeded out before they are fully refined, then a great saving of time can be made. By applying EFOM after the sixth cycle of refinement the time per trial for ERGO was reduced from 15·140 to 8·376 s.

No direct comparison with *YZARC* has been made as far as power is concerned although, since they share the large starting-set characteristic, they are probably similarly effective. However, here, *RANTAN* has a considerable advantage in terms of computer time, being more than ten times faster if systems of similar size are being processed. That comparison must be qualified by the fact that *YZARC* is usually employed with a smaller system and also that the original *YZARC* least-squares refinement process is being replaced by a gradient refinement method which is three times faster.

RANTAN can be implemented by simple modifications of the *MULTAN*-80 system and a distributable version of it is being prepared. It is also hoped, in due course, to have available a version which can be deployed in a minicomputer.

Table 2. *A comparison of computer time of MULTAN, RANTAN and YZARC*

N is the number of phases; M is the number of relationships; T is the number of trials.

		MULTAN	*RANTAN*	*YZARC*
ERGO*	N, M, T	332, 6141, 112	332, 6141, 112	66, 382, 50
	Time per trial (s)	13·140	15·140	7
AZET†	N, M, T	200, 3132, 40	200, 3132, 40	200, 3201, 27
	Time per trial (s)	4·635	8·367	100

* Using the WTF in *MULTAN* and *RANTAN*.
† Using the SWTF in *MULTAN* and *RANTAN*.

I should like to express my gratitude to Professor M. M. Woolfson for helpful discussions and encouragement with this project. I would like to thank Dr P. Main and Mr S. J. Fiske for their help with *MULTAN*-80. My thanks are also due to the Royal Society of London and the Academia Sinica for their support of my stay in the United Kingdom.

The help and resources of the University of York Computer Service are also gratefully acknowledged.

References

BAGGIO, R., WOOLFSON, M. M., DECLERCQ, J. P. & GERMAIN, G. (1978). *Acta Cryst.* A**34**, 883–892.

BERNSTEIN, J., REGEV, H., HERBSTEIN, F. H., MAIN, P., RIZVI, S. H., SASVARI, K. & TURCSANYI, B. (1976). *Proc. R. Soc. London Ser. A*, **347**, 419–434.

BRACKMAN, J. C., DALOZE, D., DUPONT, A., TURSCH, B., DECLERCQ, J. P., GERMAIN, G. & VAN MEERSSCHE, M. (1981). Submitted to *Tetrahedron*.

COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.

CONDE, A., LOPEZ CASTRO, A., MARQUEZ, R., DECLERCQ, J. P. & GERMAIN, G. (1979). *Acta Cryst.* B**35**, 2228–2229.

DECLERCQ, J. P., GERMAIN, G. & HENKE, H. (1973). *Cryst. Struct. Commun.* **2**, 405–409.

DECLERCQ, J. P., GERMAIN, G. & WOOLFSON, M. M. (1979). *Acta Cryst.* A**35**, 622–626.

HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* A**34**, 863–870.

HULL, S. E., LEBAN, I., MAIN, P., WHITE, P. S. & WOOLFSON, M. M. (1976). *Acta Cryst.* B**32**, 2374–2381.

MAIN, P. (1978). *Computing in Crystallography*, edited by H. SCHENK, R. OLTHOF-HAZEKAMP, H. VAN KONINGSVELD & G. C. BASSI, pp. 93–107. Delft Univ. Press.

MAIN, P., FISKE, S. J., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1980). *MULTAN*-80, *A system of computer programs for the automatic solution of crystal structures from X-ray diffraction data*. Univ. of York.

MAIN, P., HULL, S. E., LESSINGER, L., GERMAIN, G., DECLERCQ, J. P. & WOOLFSON, M. M. (1978). *MULTAN, A system of computer programs for the automatic solution of crystal structure from X-ray diffraction data*. Univ. of York.

SCHWEIZER, B. (1980). Private communication.

SUCK, D., MANOR, P. C. & SAENGER, W. (1976). *Acta Cryst.* B**32**, 1727–1737.

SZEIMIES-SEEBACH, U., HARNISCH, J., SZEIMIES, G., VAN MEERSSCHE, M., GERMAIN, G. & DECLERCQ, J. P. (1978). *Angew. Chem. Int. Ed. Engl.* **17**, 848–850.

SZEIMIES-SEEBACH, U., SZEIMIES, G., VAN MEERSSCHE, M., GERMAIN, G. & DECLERCQ, J. P. (1979). *Nouv. J. Chim.* pp. 357–358.