

The authors thank Dr C. H. Stam and Professor B. O. Loopstra for valuable discussions. One of the authors (GJO) is indebted to the Netherlands Organization for the Advancement of Pure Research (ZWO) for financial support.

### References

- AGARWAL, S. K., RASTOGI, R. P., VAN KONINGSVELD, H., GOUBITZ, K. & OLTHOF, G. J. (1980). *Tetrahedron*, **36**, 1435–1438.
- BUSETTA, B. (1976). *Acta Cryst.* **A32**, 139–143.
- COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.
- FISHER, J., HANCOCK, H. & HAUPTMAN, H. (1970). Report No. 7132, Naval Research Laboratory, Washington, DC, USA.
- GERMAIN, G. & WOOLFSON, M. M. (1968). *Acta Cryst.* **B24**, 91–96.
- HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* **B25**, 811–814.
- KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264–269.
- KRIEGER, W. & SCHENK, H. (1973). *Acta Cryst.* **A29**, 720–723.
- KROON, J., SPEK, A. L. & KRABBENDAM, H. (1977). *Acta Cryst.* **A33**, 382–385.
- MAULIK, P. R., VENKATASUBRAMANIAN, K., OLTHOF, G. J. & SCHENK, H. (1981). In preparation.
- OLTHOF, G. J., SINT, L. & SCHENK, H. (1979). *Acta Cryst.* **A35**, 941–946.
- OVERBEEK, A. R. & SCHENK, H. (1978). In *Computing in Crystallography*, edited by H. SCHENK, R. OLTHOF-HAZEKAMP, H. VAN KONINGSVELD & G. BASSI, pp. 108–112. Delft Univ. Press.
- PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* **A35**, 381–387.
- PUTTEN, N. VAN DER, SCHENK, H. & HAUPTMAN, H. (1980). *Acta Cryst.* **A36**, 891–897, 897–903.
- RYCHLEWSKA, U., BRATEK, M. D. & WIEWIÓROWSKI, M. (1978). *Acta Cryst.* **A34**, S106.
- SCHENK, H. (1971). *Acta Cryst.* **B27**, 2040–2042.
- SCHENK, H. (1973). *Acta Cryst.* **A29**, 503–507.
- SINT, L. & SCHENK, H. (1975). *Acta Cryst.* **A31**, S22.

*Acta Cryst.* (1981). **A37**, 689–691

## Phase Extension and Refinement for Small Structures

BY GER J. OLTHOF AND HENK SCHENK

*Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, Amsterdam, The Netherlands*

(Received 30 August 1980; accepted 3 March 1981)

### Abstract

A procedure is described to extend and refine phases starting from small starting sets. Tests with a number of real crystal structures show that the enantiomorph-maintaining refinement procedure [Olthof, Sint & Schenk (1979). *Acta Cryst.* **A35**, 941–946] can be successful in these cases as well, provided that a few modifications are introduced.

### Introduction

For quite a few structures, particularly in polar space groups and space groups without translational symmetry, it is difficult or even impossible to define and maintain the enantiomorph in the course of a direct phase determination (e.g. Schenk, 1972; Lessinger, 1976; Busetta, 1976; Woolfson, 1977). In these cases the definition of the enantiomorph can be achieved with enantiomorph-specific figures of merit, such as those based on quartets and quintets (van der Putten & Schenk, 1979) or three-phase seminvariants (van der

Putten, Schenk & Hauptman, 1980). We have shown (Olthof & Schenk, 1981) that in a symbolic addition procedure this problem can also be solved by means of the figure of merit DELCRI, which selects a small set of approximately correct phases. DELCRI is based on the relationship

$$\varphi_3 = -\varphi_H + \varphi_K + \varphi_{H-K}, \quad (1)$$

in which  $\varphi_3$  is approximated by  $s\Delta_3$ :

$$\varphi_H = \varphi_K + \varphi_{H-K} - s\Delta_3. \quad (2)$$

The  $\Delta_3$  values are empirical estimates, and the signs  $s$  are determined in the procedure.

If, however, the subsequent numerical phase extension and refinement is carried out by means of the tangent formula

$$\tan(\varphi_H) = \frac{\sum_K E_3 \sin(\varphi_K + \varphi_{H-K})}{\sum_K E_3 \cos(\varphi_K + \varphi_{H-K})} \quad (3)$$

with  $E_3 = N^{-1/2} |E_H E_K E_{H-K}|$ , in many cases the result is a set of centrosymmetric phases, in spite of the enantiomorph-specific starting set (e.g. Schenk, 1972; Lessinger, 1976). For large structures with a moderately large set of already known phases an enantiomorph-maintaining extension and refinement procedure has been developed, based on a modified tangent formula:

$$\tan(\varphi_H) = \frac{\sum_K E_3 \sin(\varphi_K + \varphi_{H-K} - s\Delta_3)}{\sum_K E_3 \cos(\varphi_K + \varphi_{H-K} - s\Delta_3)} \quad (4)$$

(Sint & Schenk, 1975; Olthof, Sint & Schenk, 1979). The signs  $s$  in (4) are taken so as to minimize

$$|-\varphi_H + \varphi_K + \varphi_{H-K} - s\Delta_3|. \quad (5)$$

If the values for both phases and  $\Delta_3$  are reasonable approximations to the correct ones the signs  $s$  will in general be correct. This implies that the enantiomorph is maintained. The two crucial points in this type of refinement are (a) the availability of the moderately large group of phased reflections which is used to estimate the  $\Delta_3$  values and (b) the number of terms contributing to the sums in (4), which must be at least five because otherwise the reliability of the  $\varphi_H$  values is too low. Busetta (1976) has shown that for an antibiotic ionophore ( $K^+C_{48}H_{86}O_{14}^-$ ,  $Z = 2$ ,  $P2_1$ ) a similar treatment was successful. It may be noted furthermore that the refinement (3) can also be improved by introducing better weighting schemes (Hull & Irwin, 1978; Giacovazzo, 1979).

In this paper it will be shown that for phasing small structures by direct methods the above procedure based on the modified tangent formula (4) can be generalized to yield reliable phases starting from a small DELCRI starting set. This procedure has been tested in five structure determinations in  $P2_1$  and  $Pca2_1$ , the results of which are summarized in the final paragraph.

### The procedure

The starting point of the extension and refinement procedure is a small set of reflections, the numerical phases of which have been determined with the figure of merit DELCRI in a symbolic addition procedure. Symbols and origin-defining reflections are used only, and thus in general this starting set consists of five to ten reflections. This set does not meet the requirements for a successful application of the extension and refinement procedure (Olthof, Sint & Schenk, 1979). Therefore in the first stage the starting set is used to calculate new phases applying the tangent formula (3). These are accepted only subject to rather strict criteria. This set of phases is not refined, since centrosymmetric phases would be obtained by use of (3), whereas for the

refinement with (4) the number of terms is too small. This extension process is repeated until the number of known phases becomes large enough to calculate a reliable  $\Delta_3$ - $E_3$  correlation. Usually this requires about 50 reflections. Then, to arrive at the  $\Delta_3$ - $E_3$  correlation, the triplet phase sums (1) are calculated. From this  $\delta(E_3)$  is obtained:

$$\delta(E_3) = \langle |\varphi_3| \rangle_{E_3}. \quad (6)$$

Next, from a plot of  $\delta$  against  $E_3$ , the  $\Delta_3$  function is obtained as the curve which fits best. Extensive details can be found in Olthof, Sint & Schenk (1979). This function is used in the adapted refinement (4).

The subsequent stages of the modified extension and refinement procedure consist of a number of extensions by means of (3), applying strict acceptance criteria, each followed by a refinement with (4). In each refinement cycle the value of

$$\text{CONS} = \sum_H \sum_K E_3 |-\varphi_H + \varphi_K + \varphi_{H-K} - s\Delta_3| \quad (7)$$

is calculated and when CONS does not improve further a new  $\Delta_3$ - $E_3$  correlation is calculated before starting the next stage. Generally the total number of stages is five.

### Test results

The phase extension and refinement procedure has been applied to five structures:

- (1) Diethylmalonic acid (DIEMAL) (van der Putten, unpublished),  $C_7H_{12}O_4$ ,  $Z = 4$ ,  $N = 44$ ,  $P2_1$ .
- (2) *N*-cyanomethylangustifoline (ANGUST) (Rychlewska, Bratek & Wiewiórowski, 1978),  $C_{16}H_{23}N_3O$ ,  $Z = 2$ ,  $N = 40$ ,  $P2_1$ .
- (3) 3-Chloro-1,3,4-triphenyl-2-azetidinone (AZET) (Colens, Declercq, Germain, Putzeys & Van Meerseche, 1974),  $C_{16}H_{23}ClNO$ ,  $Z = 8$ ,  $N = 192$ ,  $Pca2_1$ .
- (4) A naphthoquinone (INDIAN) (Agarwal, Rastogi, van Koningsveld, Goubitz & Olthof, 1980),  $C_{24}H_{26}O_4$ ,  $Z = 2$ ,  $N = 56$ ,  $P2_1$ .
- (5) Andrographolide (ANDRO) (Maulik, Venkatasubramanian, Olthof & Schenk, 1981),  $C_{20}H_{30}O_5$ ,  $Z = 2$ ,  $N = 50$ ,  $P2_1$ .

In the five cases 200, 250, 500, 300 and 300 reflections (referred to as NREF), respectively, have been used to carry out the direct-method procedure. Fourier maps calculated with the correct phases of these reflections showed that in the cases of DIEMAL and AZET one and three atoms, respectively, are not imaged and thus it is impossible for any direct method to solve these structures completely with NREF reflections.

In all cases two sets of starting phases were used, one obtained by application of the enantiomorph-specific figure of merit DELCRI in the interactive symbolic addition program *SIMPEL* (Overbeek & Schenk, 1978), and the other consisting of the correct phases. In both sets the same reflections are involved. The results of the application of the modified extension and refinement procedure are summarized in Table 1, together with the results of the application of the normal tangent procedure for comparison. The latter refinements are carried out with the guide lines of the *MULTAN* system (Main, 1980) with the exception of the Hull & Irwin (1978) weighting scheme.

In Table 1 the number of atoms (NAT) among the NP strongest peaks in the respective Fourier maps are given as well as the percentages of correct atoms, calculated with respect to the complete structure.

From Table 1 it can be concluded that refinements employing the tangent formula (3) are less satisfactory than those with the modified tangent procedure.

If the modified extension and refinement procedure is used satisfactory results are obtained in almost all cases. In the case of ANDRO starting with the correct phases the procedure was not able to reach the  $k = \text{odd}$  reflections, which led to a pseudo-translation of 0.5 with respect to the  $b$  axis in the final Fourier map. Apparently the choice of only one  $k = \text{odd}$  reflection in the starting set led to this result, which can be remedied easily.

From the averaged percentages it can be concluded that a DELCRI figure of merit starting set extended by means of our modified procedure reveals about 90% of the structure, whereas the tangent procedure leads to only  $\frac{2}{3}$  of the structure. Although one out of eight atoms still cannot be traced, it is obvious that by the present procedure structures are solved with a greatly enhanced chance of success.

The authors thank Dr C. H. Stam and Professor B. O. Loopstra for valuable discussions. One of the authors (GJO) is indebted to the Netherlands Organisation for the Advancement of Pure Research (ZWO) for financial support.

Table 1. *Numbers of atoms and percentages of correct information in four E maps for five different structures*

In column 2 the procedure for finding the phases is stated [present means the use of the present extension and refinement procedure; tangent means the use of the tangent formula (3) only] and in column 3 the origin of the starting set. The final three columns give the number of highest peaks in the Fourier map (NP), the number of atoms recognizable among these peaks (NAT), and the percentage of the structure (NPER) which is represented by the NAT atoms.

	Procedure	Starting set	NP	NAT	NPER
DIEMAL	Present	DELCRI	29	19	86
	Present	Correct	29	19	86
	Tangent	DELCRI	29	19	86
	Tangent	Correct	29	20	91
ANGUST	Present	DELCRI	27	17	85
	Present	Correct	27	20	100
	Tangent	DELCRI	27	13	65
	Tangent	Correct	27	20	100
AZET	Present	DELCRI	64	36	75
	Present	Correct	64	36	75
	Tangent	DELCRI	64	30	62
	Tangent	Correct	64	28	58
INDIAN	Present	DELCRI	37	25	89
	Present	Correct	37	27	96
	Tangent	DELCRI	37	16	57
	Tangent	Correct	37	27	96
ANDRO	Present	DELCRI	33	25	100
	Present	Correct	33	11	44
	Tangent	DELCRI	33	13	52
	Tangent	Correct	33	7	28
Averaged percentages	Present	DELCRI			87
	Present	Correct			82
	Tangent	DELCRI			65
	Tangent	Correct			75

## References

- AGARWAL, S. K., RASTOGI, R. P., VAN KONINGSVELD, H., GOUBITZ, K. & OLTHOF, G. J. (1980). *Tetrahedron*, **36**, 1435–1438.
- BUSETTA, B. (1976). *Acta Cryst.* **A32**, 139–143.
- COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.
- GIACOVAZZO, C. (1979). *Acta Cryst.* **A35**, 757–764.
- HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* **A34**, 863–870.
- LESSINGER, L. (1976). *Acta Cryst.* **A32**, 538–550.
- MAIN, P. (1980). In *Computing in Crystallography*, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN. Bangalore: Indian Academy of Sciences.
- MAULIK, P. R., VENKATASUBRAMANIAN, K., OLTHOF, G. J. & SCHENK, H. (1981). In preparation.
- OLTHOF, G. J. & SCHENK, H. (1981). *Acta Cryst.* **A37**, 684–689.
- OLTHOF, G. J., SINT, L. & SCHENK, H. (1979). *Acta Cryst.* **A35**, 941–946.
- OVERBEEK, A. R. & SCHENK, H. (1978). In *Computing in Crystallography*, edited by H. SCHENK, R. OLTHOF-HAZEKAMP, H. VAN KONINGSVELD & G. BASSI, pp. 108–112. Delft Univ. Press.
- PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* **A35**, 381–387.
- PUTTEN, N. VAN DER, SCHENK, H. & HAUPTMAN, H. (1980). *Acta Cryst.* **A36**, 897–903.
- RYCHLEWSKA, U., BRATEK, M. D. & WIEWIÓROWSKI, M. (1978). *Acta Cryst.* **A34**, S106.
- SCHENK, H. (1972). *Acta Cryst.* **A28**, 412–422.
- SINT, L. & SCHENK, H. (1975). *Acta Cryst.* **A31**, S22.
- WOOLFSON, M. M. (1977). *Acta Cryst.* **A33**, 219–225.