probability distributions of the structure seminvariants, assuming as known the magnitudes $|E|$ in their neighborhoods; in the present case the 19 magnitudes $|E|$ in the second neighborhood of T. This enormous undertaking has been carried out in only a few cases so far, and a great deal of additional work in this direction remains to be done.

HH acknowledges partial support from the National Science Foundation, Grant CHE79—11282.

References

BUSETTA, B. (1976). *Acta Cryst.* **A32**, 139–143.

DUAX, W. & HAUPTMAN, H. (1972*a*). *J. Am. Chem. Soc.* **94**, 5467–5471.

DUAX, W. & HAUPTMAN, H. (1972*b*). *Acta Cryst.* **B28**, 2912–2916.

FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* **A33**, 694–696.

GIACOVAZZO, C. (1977). *Acta Cryst.* **A33**, 933–944.

HAUPTMAN, H. (1978). *Acta Cryst.* **A34**, 525–528.

HAUPTMAN, H. & DUAX, W. (1972). *Acta Cryst.* **A28**, 393–395.

HAUPTMAN, H. & FORTIER, S. (1977). *Acta Cryst.* **A33**, 697–701.

HAUPTMAN, H. & GREEN, E. (1978). *Acta Cryst.* **A34**, 224–229.

HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* **A35**, 371–381.

OLTHOF, G. J. (1979). Unpublished.

OVERBEEK, O. (1980). Thesis, Amsterdam.

PUTTEN, N. VAN DER (1979). Unpublished.

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.

SMITH, G. D., DUAX, W. L., LANGS, D. A., DETITTA, G. T., EDMONDS, J. W., ROHRER, D. C. & WEEKS, C. M. (1975). *J. Am. Chem. Soc.* **97**, 7242–7247.

Acta Cryst. (1980). **A36**, 897–903

Application of the Three-Phase Seminvariants to Phase Determination in $P2_1$

BY NIEK VAN DER PUTTEN AND HENK SCHENK

Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, 1018 *WS Amsterdam, The Netherlands*

AND HERBERT HAUPTMAN

Medical Foundation of Buffalo Inc., 73 *High Street, Buffalo, New York* 14203, *USA*

(*Received* 5 *November* 1979; *accepted* 28 *April* 1980)

Abstract

Formulas for estimating the values (0 or π) of the enantiomorph-insensitive two- and three-phase structure seminvariants in $P2_1$ and for identifying the enantiomorph-sensitive ones (*i.e.* those having the values $\pm\pi/2$) have recently been secured. These procedures are here integrated in order to identify still more reliably those seminvariants having the values 0, π or $\pm\pi/2$. The results are employed in two ways to strengthen direct-methods procedures. The first makes active use of the three-phase structure seminvariants alone and the second relies heavily on the three- and four-phase structure invariants but employs the seminvariant enantiomorph-specific figure of merit CRISEM to select the best of a number of possible solutions. The heavy dependence on the enantiomorph-sensitive seminvariants facilitates the selection and maintenance of the enantiomorph, often a source of difficulty in this space group.

0567-7394/80/060897-07$01.00

1. Introduction

Recently a number of probability distributions for two- and three-phase structure seminvariants were derived on the basis of their quartet, quintet and sextet extensions (Hauptman & Green, 1978; Hauptman & Potter, 1979; van der Putten, Schenk & Hauptman, 1980, respectively). Examples were given which showed that the reliability of the estimates of the seminvariants is good and, particularly in the case of the enantiomorph-sensitive seminvariants, applications to phase determining procedures in space group $P2_1$ were indicated as well. The first object of this paper is to show that integrating the different approaches leads to still more reliable estimates. Then, secondly, the estimates are employed in two direct-methods procedures: *ab initio* phase determination using the three-phase seminvariants alone and phase determination *via* the three-phase and four-phase invariants, employing the seminvariant enantiomorph-specific figure of merit CRISEM. The procedures give encouraging results and open up new possibilities for phase determination in difficult cases.

2. Recapitulation of the theoretical background

In space group $P2_1$ the linear combination of two phases

$$\varphi_{12} = \varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2} \quad (2.1)$$

is a structure seminvariant if and only if

$$(h_1 + h_2, 0, l_1 + l_2) \equiv 0 \text{ (mod (2,0,2))}. \quad (2.2)$$

By embedding φ_{12} in the quartet

$$\varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2} + \varphi_{\bar{h}_1k\bar{l}_1} + \varphi_{\bar{h}_2\bar{k}\bar{l}_2}, \quad (2.3)$$

one obtains an extension of the seminvariant φ_{12}. The space-group-dependent relationships among the phases enable one to write (2.3) as $2\varphi_{12}$. Hauptman & Green (1978) have shown that the conditional probability distribution $P_{1|5}$, of φ_{12}, given the E magnitudes of the two main terms, (h_1kl_1) and $(h_2\bar{k}l_2)$, and of the three cross terms, $(02k0)$, $(h_1 + h_2, 0, l_1 + l_2)$ and $(h_1 - h_2, 0, l_1 - l_2)$, is

$$P_{1|5}, = \frac{1}{K_{1|5}} \exp \{ - [(4\sigma_3^2 - \sigma_2\sigma_4)/\sigma_2^3] R_1^2 R_2^2 \cos 2\varphi_{12} \}$$
$$\times I_0 \{ (2^{3/2}\sigma_3 R_0/\sigma_2^{3/2})[R_1^4 + R_2^4 + 2R_1^2R_2^2 \cos 2\varphi_{12}]^{1/2} \}$$
$$\times \cosh \{ (2\sigma_3/\sigma_2^{3/2}) R_1 R_2 R_{12} \cos \varphi_{12} \}$$
$$\times \cosh \{ (2\sigma_3/\sigma_2^{3/2}) R_1 R_2 R_{1\bar{2}} \cos \varphi_{12} \} \quad (2.4)$$

in which $K_{1|5}$ is a suitable normalization factor,

$$R_1 = |E_{h_1kl_1}|, \quad R_2 = |E_{h_2\bar{k}l_2}|, \quad R_0 = |E_{02k0}|,$$
$$R_{12} = |E_{h_1+h_2,0,l_1+l_2}|, \quad R_{1\bar{2}} = |E_{h_1-h_2,0,l_1-l_2}|$$

and

$$\sigma_n = \sum_{j=1}^{N} f_j^n \quad (2.5)$$

where f_j is the zero-angle atomic scattering factor of the atom labelled j. $P_{1|5}$ shows that if R_1, R_2, R_0, R_{12} and $R_{1\bar{2}}$ are all large, then

$$\varphi_{12} \simeq 0 \text{ or } \pi; \quad (2.6)$$

but if R_1, R_2, are large and R_0, R_{12} and $R_{1\bar{2}}$ are all small, then

$$2\varphi_{12} \simeq \pi \quad \text{and} \quad \varphi_{12} \simeq \pm\pi/2. \quad (2.7)$$

In the latter case the structure seminvariant φ_{12} is enantiomorph sensitive. In contrast to the estimate (2.6) of φ_{12}, which is ambiguous, the estimate $+\pi/2$ of (2.7) corresponds to one enantiomorph and the estimate $-\pi/2$ to the other enantiomorph. The reliability of $\varphi_{12} \simeq \pi/2$ in the range $0 \leq \varphi_{12} \leq \pi/2$ is measured by the variance of the mode, defined by

$$\text{VAR} = \int_0^{\pi/2} P_{1|5}(|\varphi_{12}| - |\varphi_{12\text{mode}}|)^2 \, d\varphi_{12}. \quad (2.8)$$

In space group $P2_1$ the linear combination of three phases

$$T = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3} \quad (2.9)$$

is a structure seminvariant if and only if

$$(h_1 + h_2 + h_3, k_1 + k_2 + k_3, l_1 + l_2 + l_3) \equiv 0 \text{ (mod(2,0,2))}, \quad (2.10)$$

i.e if and only if $h_1 + h_2 + h_3$ and $l_1 + l_2 + l_3$ are both even and $k_1 + k_2 + k_3 = 0$. By embedding T and its symmetry-related variants in suitable quintets (extensions of T), one obtains a simple relation between T and the values of the quintets. In this way the discriminant Δ_5 of T is defined in terms of the discriminants of the quintet extensions, and extreme values of Δ_5 are correlated with extreme values of T in the sense that $T \simeq 0$ or π according as $\Delta_5 \gg 0$ or $\Delta_5 \ll 0$, respectively (for specific procedure and expressions, see Hauptman & Potter, 1979).

By embedding T in the sextet extension S defined by

$$S = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3} + \varphi_{\bar{h}_1k_1\bar{l}_1} + \varphi_{\bar{h}_2k_2\bar{l}_2} + \varphi_{\bar{h}_3k_3\bar{l}_3}, \quad (2.11)$$

one obtains, in view of (2.10) and the space-group-dependent relationships among the phases,

$$S = 2T. \quad (2.12)$$

Because the most extreme values of the discriminant Δ_6 of the sextet are correlated with S in the sense that $S \simeq 0$ or π according as $\Delta_6 \gg 0$ or $\Delta_6 \ll 0$, respectively, it follows from (2.12) that if

$$\Delta_6 \gg 0 \text{ then } T \simeq 0 \text{ or } \pi \quad (2.13)$$

and if

$$\Delta_6 \ll 0 \text{ then } T \simeq \pm\pi/2. \quad (2.14)$$

[Δ_6 is given by equation (2.9) in van der Putten, Schenk & Hauptman, 1980]. As with the quartet extension, the estimate (2.13) of T is ambiguous and the estimate (2.14) of T is enantiomorph sensitive.

3. Introduction to the calculations

As described above the value of the three-phase structure seminvariant T in space group $P2_1$ is estimated *via* its quintet and sextet extensions and the value of the two-phase seminvariant φ_{12} is estimated *via* the quartet extension. It should be noted that the different probability expressions give numerical values for the same quantity and this implies that they can fruitfully be used together in order to arrive at more reliable estimates. In particular, it is possible to get two indications for the case that $T \simeq 0$, $T \simeq \pi$ and $T \simeq \pm\pi/2$.

The conditions for the estimation of $T \simeq 0$ or π and $T \simeq \pm\pi/2$ are summarized in Table 1. For three-phase seminvariants consisting of three phase-restricted reflections, only the associated quintet discriminants are calculated. In this case $T = 0$ or π, and so the associated sextet discriminant will not be calculated. T estimates of 0 or π correspond to $\Delta_5 \gg 0$ or $\Delta_5 \ll 0$, respectively. For three-phase seminvariants consisting of three general reflections the $T = 0$ estimates are expected to have large values of both the quintet and sextet discriminants ($\Delta_5 \gg 0$ and $\Delta_6 \gg 0$). If $\Delta_5 \ll 0$ and $\Delta_6 \gg 0$ then the most probable value of T is π; if $\Delta_6 \ll 0$ and $\Delta_5 \simeq 0$ then $T \simeq \pm\pi/2$. Finally, for seminvariants T consisting of two general reflections and one phase-restricted reflection, the associated sextet

$$S = 2T = \varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2} + \varphi_{h_30l_3} + \varphi_{\bar{h}_1k\bar{l}_1} + \varphi_{\bar{h}_2\bar{k}\bar{l}_2} + \varphi_{\bar{h}_30\bar{l}_3} \quad (3.1)$$

reduces to the quartet

$$2\varphi_{12} = 2T = \varphi_{h_1kl_1} + \varphi_{\bar{h}_1k\bar{l}_1} + \varphi_{h_2\bar{k}l_2} + \varphi_{\bar{h}_2\bar{k}\bar{l}_2} \quad (3.2)$$

because $\varphi_{h_30l_3} = 0$ or π. $T \simeq 0$ or π estimates are expected to correspond to $\Delta_5 \gg 0$ and $2\varphi_{12} \simeq 0$ or $\Delta_5 \ll 0$ and $2\varphi_{12} \simeq 0$, respectively. In case $\varphi_{12} \simeq \pm\pi/2$, VAR, which is given by (2.8), is small and $\Delta_5 \simeq 0$, then the most probable value of T is $\pm\pi/2$.

In order to test the respective estimation procedures they have been applied to five known structures in space group $P2_1$.

(1) Diethylmalonic acid (DIEMAL) (van der Putten, 1979), $C_7H_{12}O_4$, $Z = 4$, $N = 44$.

(2) Tribenzamide (TRIBEN) (Olthof, 1979), $C_{21}H_{15}NO_3$, $Z = 2$, $N = 50$.

(3) Aldosterone monohydrate (ALDO) (Duax & Hauptman, 1972), $C_{21}H_{28}O_5.H_2O$, $Z = 2$, $N = 54$.

(4) Valinomycin (VALI) (Smith, Duax, Langs, DeTitta, Edmonds, Rohrer & Weeks, 1975), $C_{54}H_{90}N_6O_{18}$, $Z = 2$, $N = 156$.

(5) 3-[(4-Bromophenyl)methylene]-6,6-dimethylbicyclo[3,1,1]heptan-2-one (SR20) (Germain, 1979), $C_{16}H_{17}BrO$, $Z = 8$, $N = 144$.

For DIEMAL, TRIBEN and ALDO, the 200 strongest reflections and, for VALI and SR20, the 500 strongest reflections were used to generate 5000–12 000 three-phase seminvariants having the largest values of

$$E_3 = \frac{\sigma_3}{\sigma_2^{3/2}} |E_{h_1k_1l_1} E_{h_2k_2l_2} E_{h_3k_3l_3}|. \quad (3.3)$$

Then the respective discriminants or phase-sum estimates of the quintet and sextet (or quartet) extensions are calculated. For the quintet discriminant calculation all possible extensions are screened and the most extreme value (either negative or positive) is used to define Δ_5 of the seminvariant. Where the reflection $H = 0$ appears as a cross term, its value is given by

$$E_0 = \sigma_2^{3/2}/\sigma_3, \quad (3.4)$$

which in the equal-atom case is identical to $E_0 = N^{1/2}$. The sextet extension is unique, as described above, and provides a straightforward calculation of Δ_6.

In order to arrive at the desired estimates $T \simeq 0$, $T \simeq \pi$ and $T \simeq \pm\pi/2$, the Δ_5, Δ_6 and φ_{12} estimates are then ordered in three different ways:

(1) in accordance with their Δ_5 values (Δ_6 is printed additionally);

(2) in accordance with their Δ_6 values (Δ_5 is printed additionally);

(3) in accordance with the variance of the φ_{12} estimates in those cases where $T = \varphi_{12}$ is predicted to be $\pm\pi/2$ (Δ_5 is printed additionally).

Table 1. *The conditions for the estimates $T \simeq 0$, $T \simeq \pi$ and $T \simeq \pm\pi/2$ in the cases that the seminvariant consists of:* (1) *three general reflections*, (2) *two general and one restricted reflection and* (3) *three restricted reflections, respectively*

In order to make the estimates, condition (*a*) must be fulfilled whilst condition (*b*) has a more supplementary character.

	Three general	2 + 1	Three restricted
$T \simeq 0$	(*a*) $\Delta_5 \gg 0$ (*b*) $\Delta_6 \gg 0$	(*a*) $\Delta_5 \gg 0$ (*b*) $2\varphi_{12} \simeq 0$	$\Delta_5 \gg 0$
$T \simeq \pi$	(*a*) $\Delta_5 \ll 0$ (*b*) $\Delta_6 \gg 0$	(*a*) $\Delta_5 \ll 0$ (*b*) $2\varphi_{12} \simeq 0$	$\Delta_5 \ll 0$
$T \simeq \pm\pi/2$	(*a*) $\Delta_6 \ll 0$ (*b*) $\Delta_5 \simeq 0$	(*a*) $\varphi_{12} \simeq \pm\pi/2$ and VAR small (*b*) $\Delta_5 \simeq 0$	Does not exist

4. Test results: estimation of the enantiomorph-insensitive seminvariants

Hauptman & Potter (1979) showed the reliability of the $T \simeq 0$ or π estimates for ALDO and VALI. In Table 2 a comparison is made for DIEMAL, TRIBEN and SR20 between the mean deviation of the triplet invariant $\varphi_3 = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{-\mathbf{h}-\mathbf{k}}$ from 0 and the mean deviation of the three-phase seminvariant T from its 0 or π estimate. The estimates have been calculated on the basis of the most important condition $\Delta_5 \ll 0$ or $\Delta_5 \gg 0$; the supplementary conditions following from Δ_6 and φ_{12} are not taken into account because, from a first calculation, they appeared not to lead to improve-

Table 2. *Average error for DIEMAL, TRIBEN and SR20 of the* $\sum_2$ *relations and of the* $T \simeq 0$ *and* $T \simeq \pi$ *estimates*

The $\sum_2$ relations are ordered in accordance with the E_3 values (4.1) and the seminvariants in accordance with their $|\Delta_5|$ values. The average errors $\langle$DEV ($\varphi_3 = 0$)$\rangle$ and $\langle$DEV ($T = 0/\pi$)$\rangle$ are given in millicycles for the number of $\sum_2$ relations (NR) and the number of seminvariants (NR) with largest E_3 and $|\Delta_5|$ values respectively, NR having the different values shown.

	DIEMAL		TRIBEN		SR20	
NR	$\langle$DEV ($\varphi_3=0$)$\rangle$	$\langle$DEV ($T=0/\pi$)$\rangle$	$\langle$DEV ($\varphi_3=0$)$\rangle$	$\langle$DEV ($T=0/\pi$)$\rangle$	$\langle$DEV ($\varphi_3=0$)$\rangle$	$\langle$DEV ($T=0/\pi$)$\rangle$
10	27	30	60	64	21	18
50	31	33	64	67	21	21
100	36	34	71	78	20	23
300	42	39	79	82	26	23
500	49	46	84	87	31	23
700	55	50	89	92	37	26
900	61	52	96	94	40	28
1200					45	34
1500					47	36
2000					52	48

ments. This has been done for the triplet invariants arranged in decreasing order of

$$E_3 = (\sigma_3/\sigma_2^{3/2})|E_{\mathbf{h}} E_{\mathbf{k}} E_{\mathbf{h+k}}| \quad (4.1)$$

and for the three-phase seminvariants arranged in decreasing order of $|\Delta_5|$. The lower limit for $|\Delta_5|$ is 30·0 for DIEMAL and SR20, and 20·0 for TRIBEN. The deviations are given in millicycles (1000 mc $= 2\pi$ rad). The number of seminvariants in this table for which one of the symmetry-related seminvariants happens to be an invariant is approximately 50, 75 and 15% for DIEMAL, TRIBEN and SR20, respectively.

5. Test results: enantiomorph-sensitive seminvariants

The estimation of the 2 + 1 enantiomorph-sensitive three-phase seminvariant $T = \varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2} + \varphi_{h_30l_3}$ was carried out by the estimation of the enantiomorph-sensitive two-phase seminvariant $\varphi_{12} = \varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2}$ *via* its quartet extension. Of course one $\varphi_{12} \simeq \pm\pi/2$ estimate will result in more than one $T \simeq \pm\pi/2$ estimate in this way at the same time. In most cases many reflections with indices h_30l_3, for which $h_1 + h_2 + h_3 \equiv 0 \pmod 2$ and $l_1 + l_2 + l_3 \equiv 0 \pmod 2$ is valid, are present among the reflections for which the three-phase seminvariants are generated. In Table 3 the test results of the $\varphi_{12} \simeq \pm\pi/2$ estimates calculated by means of the $P_{1|5}$ expression are shown for DIEMAL, TRIBEN, ALDO, VALI and SR20. The secondary condition was that $\langle|\Delta_5|\rangle_{h_30l_3} < 15{\cdot}0$ where $\langle|\Delta_5|\rangle_{h_30l_3}$ is the average of the maximal quintet discriminant values $|\Delta_5|$, associated with all three-phase seminvariants, constructed from the pair $\varphi_{12} = \varphi_{h_1kl_1} + \varphi_{h_2\bar{k}l_2}$ and different $\varphi_{h_30l_3}$'s. From the table it can be seen that small variances correspond to fairly reliable estimates. In general $\langle$VAR$\rangle$ values smaller than 400 correspond to useful seminvariant estimates. Without giving details it can be noted that the use of the secondary condition for $|\Delta_5|$ improved the results appreciably.

The test results of the $T \simeq \pm\pi/2$ estimates *via* their sextet extensions are given for the five test structures in Table 4. Apart from the condition that the sextet discriminant value Δ_6 had to be negative, the maximal discriminant value $|\Delta_5|$ of the associated quintet was also required to be smaller than 15·0. The results on the basis of the sextet discriminant alone are presented elsewhere (van der Putten, Schenk & Hauptman, 1980). From a comparison of these results it is clear that the use of the secondary condition for Δ_5 improves the results.

6. Concluding remarks concerning the test results

The probabilistic theory of the three-phase seminvariant T in $P2_1$ using the extension concept has been shown to lead to a fairly powerful technique for obtaining reliable estimates of T. In particular, the estimates $T \simeq 0$, or π, or $\pm\pi/2$ with Δ_5 and Δ_6 respectively may well give the additional enantiomorph-insensitive and -sensitive information needed for strengthening the phase-determination process in difficult cases. Also it is important to notice that the most reliable estimates are for the relations among the strongest reflections.

Van der Putten, Schenk & Hauptman (1980) have shown that the formalism of the estimate $T \simeq \pm\pi/2$ *via* the sextet extension can be applied to all three-phase sums (variants)

$$V = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3} \quad (6.1)$$

provided only that $k_1 + k_2 + k_3 = 0$. They also showed that the power of the enantiomorph-sensitive relations is increased if these variants V are generated and estimated as well.

Table 3. *Average error and variance* (2.8) *of the* $\varphi_{12} \simeq \pm\pi/2$ *estimates for the five test structures*

The seminvariants are arranged in increasing order of the variance. The average errors ⟨DEV⟩ (in millicycles) and the average variance times 0·01 (⟨VAR⟩) are given for the number of seminvariants (NR) with smallest variance, NR having the different values shown.

	DIEMAL		TRIBEN		ALDO		VALI		SR20	
NR	⟨DEV⟩	⟨VAR⟩	⟨DEV⟩	⟨VAR⟩	⟨DEV⟩	⟨VAR⟩	⟨DEV⟩	⟨VAR⟩	⟨DEV⟩	⟨VAR⟩
1	3	213	240	407	146	496	29	301	13	35
5	34	240	167	520	136	684	147	343	15	44
10	24	269	156	547			121	375	16	49
15	45	294	131	566			87	392	40	56
20	56	309	117	579			80	409	58	64
25	75	323	119	591			95	423	80	75

Table 4. *Average error of the* $T \simeq \pm\pi/2$ *estimates in groups of seminvariants* (*NR in each group*) *for the five test structures*

The seminvariants are arranged in increasing order of their Δ_6 values and the value given in the column labeled Δ_6 is the largest (*i.e.* least negative) one in the group. The average errors ⟨DEV⟩ are given for the numbers of seminvariants (NR) with smallest Δ_6, NR having the different values shown.

	DIEMAL		TRIBEN		ALDO		VALI		SR20	
NR	⟨DEV⟩	Δ_6	⟨DEV⟩	Δ_6	⟨DEV⟩	Δ_6	⟨DEV⟩	Δ_6	⟨DEV⟩	Δ_6
1	13	−31·1	14	−2·9	48	−4·6	13	−13·4	249	−149·4
5	34	−24·6	49	−2·2	47	−2·9	36	−4·2	133	−95·7
10	23	−18·7	90	−1·6	88	−2·1	60	−3·9	133	−52·3
15	25	−16·3	101	−1·5	101	−1·9	69	−3·1	126	−45·4
20	27	−14·0	113	−1·4	98	−1·8	61	−2·7	112	−39·4
25	30	−12·1	108	−1·3	86	−1·6	73	−2·6	118	−28·4
30	28	−11·1	103	−1·2	95	−1·4	74	−2·0	112	−26·3
35	29	−9·9	106	−1·2	95	−1·2	70	−2·0	119	−22·2
40	35	−9·0	114	−1·1	101	−1·2	78	−1·8	115	−17·6
45	41	−8·6	118	−1·1	103	−1·1	85	−1·5	115	−16·9
50	40	−7·8	118	−1·0	105	−1·1	88	−1·4	119	−16·2

Finally, the estimate $\varphi_{12} = \pm\pi/2$ by means of the P_{115} formula (Table 3) may also prove to be useful, but the variances must be considered carefully. Further, the test results with SR20 (Tables 3 and 4) show that structures with heavy atoms may cause trouble if one relies on the Δ_6 values or variances of the $\varphi_{12} \simeq \pm\pi/2$ estimates on an absolute scale.

7. Application of three-phase seminvariants in phasing procedures

The application of enantiomorph-insensitive and -sensitive three-phase seminvariants together are particularly useful in phasing procedures in space group $P2_1$. On the one hand $P2_1$ is one of the most frequently occurring space groups, and on the other hand parts of the phase determination in $P2_1$ may give rise to more serious problems than in other space groups. The problems lie mainly in the difficulty of enantiomorph definition, in the drift of the phases in the phase extension and refinement process towards zero, and in the preference of figures of merit for centrosymmetric solutions (see *e.g.* Schenk, 1972; Hull & Irwin, 1978; van der Putten & Schenk, 1979; Busetta, 1976). The three-phase seminvariants were implemented in the symbolic addition program *SIMPEL* (Overbeek, 1980) in two different ways.

(i) *PROC* I. The complete phasing procedure, including the evaluation of the possible solutions by means of figures of merit, is carried out with three-phase seminvariants alone. The procedure is similar to the *SIMPEL* procedure, in which the enantiomorph specification is introduced in the FOM calculation.

(ii) *PROC* II. The phasing procedure is carried out with three- and four-phase invariants. However, the evaluation and sorting of the possible solutions are executed with the seminvariant criterion CRISEM.

8. The basic *SIMPEL* concept with three-phase seminvariants alone

The basic steps in the phase-development scheme, using three-phase seminvariants alone, are

(*a*) program *SEMCAL*: generates the three-phase seminvariants;

(*b*) program *DISCAL*: estimates the three-phase seminvariants;

(*c*) program *SORCEM*: produces a redundant seminvariant file;

(*d*) program *STASEM*: generates a starting set;

(*e*) program *SYMBAD*: assigns new symbolic phases using the enantiomorph-insensitive seminvariants alone;

(*f*) program *QCRIT*: calculates Q criteria (Schenk, 1971*a*) based on the seminvariant relations from *SYMBAD* for all permutations of test values given by the user. The sets of symbol values with the lowest criterion values can be used for *QREF*. The enantiomorph is defined automatically, because the *QCRIT* values have one by one the same value;

(*g*) program *QREF*: refines, by least squares, the rough values from *QCRIT* (Schenk, 1971*b*). The symbol-value combination with the lowest figure of merit should be the most probable one;

(*h*) program *TANREF*: does a weighted tangent refinement with the ($T \simeq 0$ or π) seminvariants, starting with the symbol values of the best *QREF* solution.

9. The seminvariant figure of merit: CRISEM

On the basis of the extreme estimates 0 and π and of the enantiomorph-sensitive estimates $\pm\pi/2$ of the three-phase seminvariants, the figure of merit CRISEM can be formulated as follows:

$$\mathrm{CRISEM} = \sum_i W_{5_i} |\varphi_{\mathbf{h}_i} + \varphi_{\mathbf{k}_i} + \varphi_{\mathbf{l}_i} - T_i(0/\pi)| + \sum_j W_{6_j} \big|\, |\varphi_{\mathbf{h}_j} + \varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j}| (\mathrm{mod}\ \pi) - T_j(\pi/2) \big| \quad (9.1)$$

in which $\varphi_{\mathbf{h}_i}, \varphi_{\mathbf{k}_i}, \varphi_{\mathbf{l}_i}$ and $\varphi_{\mathbf{h}_j}, \varphi_{\mathbf{k}_j}, \varphi_{\mathbf{l}_j}$ are determined in a direct phase procedure, $W_{5_i} = E_3^*$ with E_3^* defined in § 8 and $W_{6_j} = 0{\cdot}5^* |\Delta_6|^{1/2}$. $T_i(0/\pi) = 0$ or π and $T_j(\pi/2) = \pm\pi/2$ are estimated in accordance with the Δ_5 values and Δ_6 values respectively. The weight W_{6_j} has been chosen semi-empirically to be in accord with the observed distribution of errors in the two kinds of contributors to (9.1) in certain test structures.

It must be noted that CRISEM consists of an enantiomorph-insensitive first part and an enantiomorph-sensitive second part. Further, the last part discriminates only in the range $0 \leq |\varphi_{\mathbf{h}_j} + \varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j}| < \pi$. The CRISEM figure of merit can be used in both multisolution tangent refinement and symbolic addition procedures. In the first procedure for each solution a CRISEM value can be calculated. To bring these into a correct relative scale it is necessary to use, instead of (9.1), the related criterion

$$\mathrm{CRISEM} = \sum_i W_{5_i} |\varphi_{\mathbf{h}_i} + \varphi_{\mathbf{k}_i} + \varphi_{\mathbf{l}_i} - T_i(0/\pi)| \Big/ \sum_i W_{5_i} + \sum_j W_{6_j} \big|\, |\varphi_{\mathbf{h}_j} + \varphi_{\mathbf{k}_j} + \varphi_{\mathbf{l}_j}| (\mathrm{mod}\ \pi) - T_j(\pi/2) \big| \Big/ \sum_j W_{6_j}. \quad (9.2)$$

In the symbolic addition procedures the phrases $\varphi_{\mathbf{h}_i}$ *etc.* are expressed in terms of the symbols X_n and thus (9.2) can be rewritten as:

$$\mathrm{CRISEM}\ (X_1, X_2, \ldots, X_n) = \sum_i W_{5_i} |\sum_n A_{in} X_n - T_i(0/\pi)| + \sum_j W_{6_j} \big|\, |\sum_n A_{jn} X_n| (\mathrm{mod}\ \pi) - T_j(\pi/2) \big|. \quad (9.3)$$

Then for sets of numerical values of X_n the CRISEM figure of merit can be evaluated. Starting from a set of parameter values in (9.3) it is possible to use an iterative least-squares procedure to refine the X_n values. The function to be minimized is

$$R(X_n) = \sum_i W_{5_i} |\sum_n A_{in} X_n - T_i(0/\pi)|^2 + \sum_j W_{6_j} \big| \sum_n |A_{jn} X_n| (\mathrm{mod}\ \pi) - T_j(\pi/2) \big|^2. \quad (9.4)$$

10. Applications

The *SIMPEL* procedure using enantiomorph-insensitive seminvariants only (*PROC* I) and the normal *SIMPEL* procedure using triplets and quartets together with the figure of merit CRISEM (*PROC* II) have been applied to three known structures, DIEMAL, ALDO and TRIBEN. These structures were chosen because one (DIEMAL) belongs to the group of the fairly easily solvable structures and the other two belong to the group of fairly difficult or difficult structures. For all structures about 6000 three-phase seminvariants were generated for the 200 reflections with the strongest E values. The limit for the E_3 value was 1·5, 1·5 and 1·0 for DIEMAL, ALDO and TRIBEN respectively. The results of the phase determination with *PROC* I and *PROC* II are summarized in Table 5. This table is headed in the following way.

NR OF $T \simeq \pm\pi/2$ USED: the number of enantiomorph-sensitive three-phase seminvariants used,

Table 5. *Results of the phase determination with PROC I and PROC II for DIEMAL, ALDO and TRIBEN*

	NR OF $T \simeq \pm\pi/2$ USED	NR OF $T \simeq 0, \pi$ USED	Small starting set NR OF ORIG. REFL.	NR OF SYMBOLS	NR OF EXAM. SOL.	NR IN *QREF* + ⟨error⟩	NR IN LSR AFTER CRISEM + ⟨error⟩
DIEMAL: *PROC* I		528	3	3 (3 + 0)	64	1 ⟨31 mc⟩	
DIEMAL: *PROC* II	20	528	3	4 (3 + 1)	128		1 ⟨26 mc⟩
ALDO: *PROC* I		387	3	4 (3 + 1)	128	2 ⟨76 mc⟩	
ALDO: *PROC* II	20	387	3	5 (2 + 3)	128		3 ⟨60 mc⟩
TRIBEN: *PROC* I		424	3	5 (5 + 0)	1024	1 ⟨24 mc⟩	
TRIBEN: *PROC* II	20	424	3	5 (5 + 0)	1024		6 ⟨90 mc⟩

of which the sums of the indices of the reflections satisfy (2.10).

NR OF $T \simeq 0/\pi$ USED: the number of enantiomorph-insensitive seminvariants used. The upper limits of $|\Delta_5|$ for the invariants and seminvariants were 80·0 and 30·0 (20·0 for TRIBEN) respectively.

NR OF ORIG. REFL.: the number of reflections for origin definition.

NR OF SYMBOLS: the number of reflections in the small starting set with symbolic phases. The numbers of restricted and general reflections are shown in parentheses. The starting set for TRIBEN had one more restricted reflection with a numerical value reliably obtained from $\sum_1$.

NR OF EXAM. SOL.: the number of solution sets examined in *QCRIT* (*PROC* I) or CRISEM (*PROC* II).

NR IN *QREF*: number of the most plausible solution sets in the refinement by *QREF*. Also the mean error from the correct symbol values is given in millicycles.

NR IN LSR AFTER CRISEM: number of the best plausible set of symbol values in the least-squares refinements criterion $R(X_n)$. Also the mean error from the correct symbol values is given in millicycles.

From the last two columns in Table 5 it can be seen that both procedures give encouraging results, especially since the last two examples concern problem structures. We have also tried a tangent refinement with the very small set of the enantiomorph-insensitive seminvariants alone for all solutions mentioned in Table 5. The calculated E maps revealed a large recognizable fragment of the structure in all cases. The smallest fragment consisted of 17 of the 25 atoms of TRIBEN within the 30 largest peaks in the E map calculated with the phases from the tangent refinement with seminvariants alone for solution NR6 in CRISEM. Other tests with two known and one unknown structure showed that the figure of merit CRISEM is fairly dependent on the number of good estimates for the enantiomorph-sensitive three-phase seminvariants; but they also showed the relation between the value of the sextet discriminant and the reliability of the estimation of $T \simeq \pm\pi/2$. Use of the $\pm\pi/2$ estimates of all variants (6.1) appears to enhance the reliability of CRISEM. In summary, CRISEM is expected to be a reliable figure of merit provided that there are a fairly large number of three-phase (semin)variants having Δ_6 values $< -2·0$.

HH acknowledges support for this research by the National Science Foundation (CHE79-11282). Our special thanks are extended to Steve Potter and Ok Overbeek for their help in using some of the computer programs.

References

BUSETTA, B. (1976). *Acta Cryst.* A**32**, 139–143.
DUAX, W. & HAUPTMAN, H. (1972). *J. Am. Chem. Soc.* **94**, 5467–5471.
GERMAIN, G. (1979). Unpublished.
HAUPTMAN, H. & GREEN, E. (1978). *Acta Cryst.* A**34**, 224–229.
HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* A**35**, 371–381.
HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* A**34**, 863–870.
OLTHOF, G. J. (1979). Unpublished.
OVERBEEK, O. (1980). Thesis, Amsterdam.
PUTTEN, N. VAN DER (1979). Unpublished.
PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* A**35**, 381–387.
PUTTEN, N. VAN DER, SCHENK, H. & HAUPTMAN, H. (1980). *Acta Cryst.* A**36**, 891–897.
SCHENK, H. (1971*a*). *Acta Cryst.* B**27**, 2037–2039.
SCHENK, H. (1971*b*). *Acta Cryst.* B**27**, 2040–2042.
SCHENK, H. (1972). *Acta Cryst.* A**28**, 412–422.
SMITH, G. D., DUAX, W. L., LANGS, D. A., DETITTA, G. T., EDMONDS, J. W., ROHRER, D. C. & WEEKS, C. M. (1975). *J. Am. Chem. Soc.* **97**, 7242–7247.