PRESS, W. H., FLANNERY, B. P., TEUKOLSKY, S. A. & VETTERLING, W. T. (1988). *Numerical Recipes in C: the Art of Scientific Computing.* Cambridge Univ. Press.
STUBBS, G. J. (1974). *Acta Cryst.* A30, 639-645.
TIBBITTS, T. T., CASPAR, D. L. D., PHILLIPS, W. C. & GOODENOUGH, D. A. (1988). *Biophys. J.* 53, 634*a*.
TIBBITTS, T. T., CASPAR, D. L. D., PHILLIPS, W. C. & GOODENOUGH, D. A. (1990). *Biophys. J.* 57, 1025-1036.

Acta Cryst. (1993). A49, 545-557

Frequency Statistical Method for Evaluating Cosine Invariants of Three-Phase Relationships

BY DAVID A. LANGS

Medical Foundation of Buffalo, 73 High Street, Buffalo, NY 14203, USA

(*Received* 8 *June* 1992; *accepted* 26 *October* 1992)

Abstract

A new variation on the established procedure to evaluate three-phase structure invariants through quadrupole relationships is described. This method differs from earlier algebraic formulations in that the cosine-invariant estimates are based on a conditional observed frequency distribution of $|E|$ magnitudes for the quadrupole, rather than on the values of the magnitudes themselves. Successful applications of this method to a number of structures that ranged in size from 84 to 317 independent non-hydrogen light atoms are given.

Introduction

The three-phase crystallographic structure invariants play a central role in the determination of crystal structures by direct-phasing methods. Tangent-formula methods for small-molecule determinations have traditionally relied on the 0 (modulo 2π) probability estimate for these 'triples' (Karle & Hauptman, 1956). Efforts to extend these techniques to larger structures have required more precise estimates to be obtained for these phase invariants, though use of algebraic formulae (Karle & Hauptman, 1957; Vaughan, 1958; Hauptman, 1964; Hauptman, Fisher, Hancock & Norton, 1969; Karle, 1970; Duax, Weeks & Hauptman, 1972; Hauptman & Duax, 1972), determinantal joint probability distributions (Tsoucaris, 1970; Messager & Tsoucaris, 1972; Giacovazzo, 1976, 1977*a*; Karle, 1979, 1980) or probabilistic formulae, as applied to isomorphous-replacement or anomalous-dispersion data (Hauptman, 1982; Giacovazzo, 1983; Fortier, Moore & Frazer, 1985) and to the extended neighborhoods or phasing shells of data that define higher-order relationships into which these triples have been suitably embedded (Hauptman, 1975; Giacovazzo, 1977*b*; Karle, 1982). This report describes a new method to estimate three-phase invariants based on examination of the frequency distribution of $|E|$ magnitudes that complete a family of conditionally constructed quadrupoles that are common to the evaluated triple.

Background

One of the earliest strategies in direct-methods research was the development of formulae to evaluate crystal-structure phase invariants and semi-invariants as a means to determine crystal structures. This work was initiated about the same time that the rules for origin and enantiomorph specification and phase-extension techniques were being developed. Formulae to estimate the single-phase structure invariants had an immediate application; they provided a means to reduce the number of algebraic symbols that had to be permuted and tested for a selected starting group of phases. But algebraic formulae that were developed for the determination of the cosine values of the three-phase structure invariants, for example, for $P1$ symmetry (Karle & Hauptman, 1957),

$$\begin{aligned}|E_{\mathbf{h}}E_{-\mathbf{k}}E_{\mathbf{k}-\mathbf{h}}|\cos(\varphi_{\mathbf{h}}-\varphi_{\mathbf{k}}+\varphi_{\mathbf{k}-\mathbf{h}})&\\ \simeq N^{-1/2}(|E_{\mathbf{h}}|^2+|E_{\mathbf{k}}|^2+|E_{\mathbf{k}-\mathbf{h}}|^2-2)&\\ +\tfrac{1}{2}N^{3/2}\langle(|E_{\mathbf{l}}|^2-1)(|E_{\mathbf{l}-\mathbf{k}}|^2-1)(|E_{\mathbf{l}-\mathbf{h}}|^2-1)\rangle_{\mathbf{l}},&\quad(1)\end{aligned}$$

however, did not have an immediate impact on phasing practices. Firstly, these formulae were computationally demanding; the average of a product of $|E|^2-1$ magnitudes had to be computed over a range of $\mathbf{l}$ that sampled the whole of reciprocal space and, to ensure that the whole of reciprocal space was explored, the diffraction vectors of the triple, $\mathbf{h}$, $-\mathbf{k}$ and $\mathbf{k}-\mathbf{h}$, had to be cyclicly permuted within the framework of the formula. Secondly, numerical tests indicated that (1) tended to produce unreliable cosine

0108-7673/93/030545-13$06.00

estimates as a consequence of the unavoidable overlap of vectors in the Patterson function and attempts were made to modify these calculations to improve the results.

One variant of (1) that was developed to mitigate the Patterson-function overlap problem, subsequently referred to as the TPROD formulation (Hauptman, Fisher, Hancock & Norton, 1969), was

$$|E_h E_{-k} E_{k-h}| \cos(\varphi_h - \varphi_k + \varphi_{k-h}) \simeq R_3 + K\Psi, \quad (2)$$

$$R_3 = \tfrac{1}{4}N^{-1/2}[\tfrac{3}{2}(|E_h E_k|^2 + |E_k E_{k-h}|^2 + |E_h E_{k-h}|^2) + |E_h|^2 + |E_k|^2 + |E_{k-h}|^2 - \tfrac{7}{2}],$$

$$\Psi = \langle(|E_1|^{1/2} - \xi)(|E_{1-k}|^{1/2} - \xi)(|E_{1-h}|^{1/2} - \xi)\rangle_1,$$

where $\xi = \langle E_h|^{1/2}\rangle_h$ and K is a scale factor that can be adjusted as a function of the magnitude of $|E_h E_k E_{k-h}|$ in an effort to fit the calculated distribution of cosines to the theoretical distribution. The MDKS formula (Fisher, Hancock & Hauptman, 1970) is an alternate variant of (1) that utilizes conditional averages,

$$|E_h E_{-k} E_{k-h}| \cos(\varphi_h - \varphi_k + \varphi_{k-h}) \simeq M(D - KS), \quad (3)$$

where

$$D = \langle |E_{1-k}|^2 - 1 \,||E_1|, |E_{1-h}| \geq t\rangle_1,$$

$$S = \langle |E_{1-k}|^2 - 1 \,||E_1| \geq t\rangle_1 + \langle |E_{1-k}|^2 - 1 \,||E_{1-h}| \geq t\rangle_1,$$

M and K are scaling constants to fit the distribution of calculated cosine values and t is a conditional threshold value placed on certain of the E magnitudes that must be satisfied before the associated terms can be included in the averages. The exact value chosen for t is somewhat arbitrary, but good results are usually obtained when t is in the range 1.0 to 1.5.

These and related algebraic formulae will only provide reliable cosine estimates for triples whose A values ($A = 2|E_h E_{-k} E_{k-h}|/N^{1/2}$, where N is the effective number of equal atoms in the primitive unit cell) exceed some limit, say 0.5 for $\sum_1$ triples, 1.5 for zonal phase-restricted $\sum_2$ triples and 2.5 for unrestricted $\sum_2$ triples composed of non-centrosymmetric phases. Below these thresholds, the estimates tend not to be reliable. The estimates tend to be more accurate the larger the A value and the greater the number of terms that are included in computing the Ψ, D and S averages. The sensitivity of the TPROD and related cosine estimates can often be improved by the imposition of constraints on particular magnitudes ($|E_1|$, $|E_{1-h}|$, $|E_{1-k}|$) that are used to compute the averages, for example it may be required that one or two of these terms exceed some threshold, say 1.5, or lie in a particular range, say between 1.5 and 2.5 (Karle, 1970; Hauptman, Fisher & Weeks, 1971). Karle (1970) has further cautioned that one should adjust these amplitude constraints to determine the point at which the cosine estimates become stable and do not fluctuate unpredictably as a consequence of variation of the number of quadrupole contributors to the formulae.

Thus far these algebraic and probabilistic formulae have been found to be useful for the solution of problematic structures containing as many as 50 to 80 independent light atoms in the unit cell and efforts to solve larger structures by these methods have not been universally successful (Gilmore & Hauptman, 1985; Gilmore, 1991). This work reports a method to evaluate phase invariants that may improve the reliability of the estimates and extend the usefulness of these formula beyond 80 atoms' complexity.

The frequency distribution

A general characteristic of all earlier algebraic formulae is that the cosine estimates were an explicit function of the values of the $|E|$ magnitudes. Yet it is commonly known that certain properties of the structure, *e.g.* whether or not it is centrosymmetric, are as readily revealed by the shape of the distribution of E values as by statistics based on powers of the $|E|$ magnitudes. For example, although the average expected $|E|^2$ values of centric and non-centric structures are both equal to 1.0 and cannot be used to distinguish whether a center exists, the percentages of E values that fall into certain $|E|$-magnitude ranges, *e.g.* those that are less than 0.1 or exceed a value of 1.0, 2.0 or 3.0, can often indicate the correct choice (Howells, Phillips & Rogers, 1950). In this regard, the information given by the relative fraction of the number of E values in the tails of the distribution curve (an intensive property) is more sensitive to whether the structure is centric or acentric than the average $|E|^2$ (an extensive property) computed from all the data. It will be shown in the following paragraphs how the information present in the high $|E|$-magnitude tail of the frequency-distribution curve can be used to provide an alternate qualitative measure of the conditional expectation values $\langle |E_{1-k}|^2 - 1 \,||E_1|, \ldots, \geq t\rangle_1$, as they appear in the D and S terms of the MDKS formula.

When the computed value of $\langle |E_{1-k}|^2 - 1 \,||E|_1, \ldots, \geq t\rangle_1$ is greater than that given by the unrestricted average, $\langle |E_1|^2 - 1\rangle_1$, provided by the normal distribution of the full set of $|E|$ values, the distribution of $|E|$ values must be perturbed to include a higher percentage of large values in the conditional average and *vice versa*.

The actual nature of these perturbations, however, appears to be intermediate between two extremes: the distribution may be (*a*) uniformly perturbed, *i.e.* retain the shape of the normal distribution of $|E|$ values but either be shifted toward larger or smaller magnitudes, or (*b*) exhibit a non-uniform or skew perturbation such that the population of $|E|$ values

in the tails of the distribution may experience the greatest changes. In either case, the fraction of $|E|$ values in the tails of the distribution will be affected, but more so in case (b) than in case (a).

It also follows that, on occasion, triples that have similar computed values for their conditional averages, $\langle |E_{1-k}|^2-1 \,||E|_1, \ldots, \geq t\rangle_1$, may have disparate frequency distributions. Examples of phasing applications have been selected to demonstrate that this frequency criterion can be useful for cosine-invariant evaluation and structure determination for a number of difficult-to-solve structures. It remains to be shown to what extent this new procedure may have mitigated the deleterious effects that Patterson-function overlap has on the derived cosine estimates.

The ν-STAT formulae to evaluate triples

It should be noted that the D term in the MDKS formula (3) corresponds to the average value of $|E|^2-1$ for any one of the three magnitudes ($|E_1|$, $|E_{1-k}|$, $|E_{h-1}|$) that form a quadrupole with a triple, on the condition that the two other magnitudes are large and greater than some threshold t_1. One of the conditions for the TPROD and MDKS cosine estimates to be large and positive may be formulated from the single-variate magnitude distribution of the unrestricted third term of the doubly conditioned quadrupole relationship. If two of these three quadrupole terms are required to be large, *e.g.* $t_1 \geq 1.75$, the tendency for the D and Ψ terms to be large or small can be sensed by the perturbation of the normal frequency distribution of $|E|$ amplitudes of the third unrestricted term. Thus, if

$$\nu^+ = \text{quads}(|E_{1-k}| \geq t_1 \,||E_1|, |E_{h-1}| \geq t_1) \times [\text{quads}(|E_{1-k}| = \text{obs.} \,||E_1|, |E_{h-1}| \geq t_1)]^{-1}, \quad (4)$$

where quads is the number of quadrupoles satisfying the condition in parentheses and $t_1 \simeq 1.75$, is larger than the random experimentally determined frequency, $\nu(\text{random}) = (\text{number of } |E| \text{ values} \geq t_1)/(\text{total number of } |E| \text{ values})$, this may be taken as an indirect indication that $\varphi_h - \varphi_k + \varphi_{k-h} \simeq 0 \pmod{2\pi}$, regardless of the statistic computed from all the data as it may affect the TPROD or MDKS formulae. But if, by comparison,

$$\nu^- = \text{quads}(|E_{1-k}| \geq t_1 \,||E_1| \leq t_2, |E_{h-1}| \geq t_1) \times [\text{quads}(|E_{1-k}| = \text{obs.} \,||E_1| \leq t_2, |E_{h-1}| \geq t_1)]^{-1}, \quad (5)$$

where $t_1 \simeq 1.75$ and $t_2 \simeq 0.20$, is also by chance larger than $\nu(\text{random})$, it may be inferred that $\varphi_h - \varphi_k + \varphi_{k-h} \simeq \pi \pmod{2\pi}$ since the Ψ term in the TPROD formula would tend to be maximally negative. Here $|E_{1-k}| = \text{obs.}$ implies that the remaining term of the quadrupole has been measured so that the frequency for which it exceeds the threshold t_1 can be validly computed. The thresholds t_1 and t_2 are selected so that the number of $|E|$ magnitudes that are greater than t_1 is approximately equal to the number of those less than t_2. For non-centric structures, the expected probability that an $|E|$ magnitude exceeds some threshold value $|E_t|$ is given by $\nu(|E| \geq |E_t|) = \exp(-|E_t|^2)$, which, for the thresholds $t_1 = 1.75$ and $t_2 = 0.22$, select ~4.5% of the weakest and strongest data. For centrosymmetric data, comparable values are $t_1 = 2.00$ and $t_2 = 0.11$, as determined from the normal probability distribution. In general, t_1 and t_2 should be adjusted to ensure that there are at least 300 quadrupole contributors in the denominators of (4) and (5) to obtain frequency statistics with reasonably small variances. Equations (4) and (5) will be referred to as the ν-STAT formulae and ν^+ and ν^- as the cosine-invariant estimators for this procedure. If there are on average much fewer than 300 quadrupole terms in the evaluation of these estimators, they will not provide reliable cosine indications.

An important observation regarding cosine-invariant estimates of large structures, whether they be from TPROD, MDKS or ν-STAT methods, is that in general the strongest positive cosine indications are correctly indicated but only about 25% of the strongest negative indications are correct. The source of this difficulty is fairly easy to understand. Let us say that in a list of the top 100 triples for some centric structure we may typically have 90 triples cosines that are +1.0 and 10 that are −1.0. If only 80% of the cosine-invariant estimates are decisively correct, we will have 72 of the 90 positive triples and 8 of the 10 negative triples correctly indicated, but at the cost of getting 18 of the positive triples and 2 of the negative triples incorrectly indicated. Thus, by relying on the most positive cosine estimates, we can significantly reduce the fraction of aberrant triples from 10% (10/100) to 2.7% (2/74). However, even though the negative estimator is capable of correctly identifying 8 of the 10 negative triples, only 31% (8/26) of the strong negative indications will actually be negative as there are 18 false indications from positive triples in this list. Similar observations are seen to hold for non-centric structures. Thus, it can be very risky to use negative cosine estimates in an active phasing role to determine complex crystal structures.

Experience has shown that one of the better phasing strategies to employ under these circumstances is to use limited symbolic addition involving the strongest positive indications to produce a large starting set of reflections that can be expressed in terms of a small number of permuted symbols. The negative estimates are seldom used actively in this manner, but may be used to confirm a $\sum_1$ phase assignment for pairs of $\sum_1$ triples that indicate opposing signs. That is, if one of the two $\sum_1$ estimates is strongly

indicated to be positive, the other must exhibit a strong negative estimate to be consistent. Similar constructions to unmask and actively utilize aberrant $\sum_2$ triples' phase information exist in the form of inconsistent quadrupoles (Viterbo & Woolfson, 1973) and triples (Han & Langs, 1988).

Applications

The ν-STAT estimators ν^+ and ν^- were computed for lists of the strongest three-phase structure invariants for a number of complex test structures. These structures included ILED, $C_{60}H_{102}N_6O_{18}$, $P2_12_12_1$, $N=84$ (Pletnev, Galitskii, Ivanov, Smith, Weeks & Duax, 1980); ACE, $C_{60}H_{102}N_6O_{18}.2C_3H_6O$, $P3_2$, $N=92$ (Pletnev, Mikhailova, Ivanov, Langs, Grochulski & Duax, 1991); ENKA, $C_{72}H_{90}N_6O_{18}$, $P1$, $N=96$ (Smith & Kristenansky, 1991); HEXIL, $C_{80}H_{136}N_8O_{24}.14H_2O$, $P2_12_12_1$, $N=126$ (Pletnev, Ivanov, Langs, Strong & Duax, 1992) and GRAMA, $C_{198}H_{280}N_{40}O_{34}.15C_2H_5OH$, $P2_12_12_1$, $N=317$ (Langs, 1988).

A variety of different kinds of three-phase structure invariants were examined. These included $\sum_1$ triples, both general and zonal phase-restricted $\sum_2$ triples and a small group of inconsistent $\sum_2$ triples that were generated for the $P3_2$ ACE structure. The minimum A values chosen for the various kinds of invariants were generally in the range that was found to be useful for the TPROD and MDKS estimates, *viz* 0.25 for $\sum_1$ triples, 1.25 for zonal restricted $\sum_2$ triples and 2.5 for unrestricted $\sum_2$ triples. In most cases, the value chosen for t_1 was 1.75 and t_2 was adjusted so that the ratio of weakest to strongest $|E|$ magnitudes was approximately 1.2:1. The only exception was the ENKA structure, where t_1 had to be reduced to 1.65 to increase the average number of quadrupole contributors in the demoninators of (4) and (5) toward the target value of 300 or more.

The results are presented in Tables 1 to 5 for the five known test structures; the aberrant phase invariants are indicated by an asterisk (*). Similar to what had been previously noted for the TPROD and MDKS estimates, the reliability of the individual ν-STAT estimates were better as the A values of the triples became larger and, moreover, it appeared that this reliability could be extended to structures of greater than 100 atoms' complexity. Table 6 lists a group of 33 zonal reflections for the GRAMA structure that can be abstracted from the phase-invariant list of Table 5 and expressed in terms of five symbols. These 33 phases together with 9 additional unrestricted phases, expressed as magic integers, were actually used to solve this structure (Langs, 1988), which contains more than 300 independent atoms in the unit cell. It sould be cautioned that results such as these are sensitive to the accuracy with which the data have been measured and scaled, as well as to the degree to which the scaled $|E|$ values model the ideal point-atom structures upon which probabilistic direct methods have been derived (Langs, 1993).

Discussion of results

The phase-invariant analyses presented in Tables 1 to 5 survey a range of structural determinations of increasing complexity. The ILED data set represents a difficult but well behaved problem in that this structure can be determined from numerous different starting sets by phase extension or by random-phasing methods (Yao, 1981) and on average requires from 250 to 500 phasing trials to find a solution. The number of phasing trials, however, can be significantly reduced if one can take advantage of the phase-invariant information from $\sum_1$ and $\sum_2$ relationships in Table 1. Of the 46 largest $\sum_1$ indications in Table 1(a), 14 are indicated to be aberrant (*) and of these 14 triples only one, involving reflection 47, is observed to have $\nu^+ \leq \nu^-$, as is indicated (†). Of the 46 zonal restricted $\sum_2$ triples with A values greater than 1.20 listed in Table 1(b), there are five aberrant (*) entries and each of these has been detected with ν^+ being less than ν^-. How can this information be effectively used?

For instance, the first $\sum_1$ triple in Table 1(a) is a strong positive indication ($\nu^+=0.072$, $\nu^-=0.036$) that the phase of the 166th reflection, 008, is 180°. There are two contradicting phase indications for reflection 166 in the table (triples nos. 10 and 46) and for both of these triples $\nu^+ \leq \nu^-$, indicating that the phase indication of 0° is incorrect or aberrant, thus confirming that the true phase is 180°, as indicated by triple number 1. Inspection of Table 1(a) readily reveals that reflections 7, 9, 10, 47, 74, 166 and 424 are 0, π, 0, 0, π, π and π, respectively. Given that there are two other strong and conflicting sign indications for reflection 47 in the table (numbers 30 and 38) and the fact that the seventh $\sum_2$ triple in Table 1(b) strongly indicates that $\varphi_7-\varphi_{10}+\varphi_{47}=0 \pmod{2\pi}$, the potential error involving phase 47 can be avoided. After selection of the ten or so strongest $\sum_2$ cosine indications from Table 1(b), it is a fairly trivial matter to derive a starting phase set of 20 or more reflections that can be expressed in terms of only two symbols. In this way, this structure can be solved in considerably fewer than the 250 to 500 trials required if one does not employ this additional phasing information.

The D and S terms from the MDKS calculation and their computed cosine estimate are presented as the three rightmost columns of Table 1. An inspection of the table discloses no strong correlation among the D and S terms and the ν^+ and ν^- estimators, even though both are derived from the same sample population of quadrupole relationships. Triples can be observed that have similar computed D or S values

Table 1. *List of triples for the test structure ILED ($P2_12_12_1$, $N = 84$ atoms), ranked on decreasing magnitude of A*

(*a*) $\sum_1$ triples $N1$, $N2$ and $N3$ are the serial numbers of the three phases that define the triple invariant, $\Phi = \varphi(N1) + \varphi(N2) + \varphi(N3) + T$, where $\varphi(I) = -\varphi(-I)$ and T is a phase shift due to symmetry operations on the constituent reflections to relate them to their standard parent forms. Of the 46 $\sum_1$ triples listed, there are 14 aberrant phase relationships, indicated by * to the right of column T. Only one of these 14 aberrant triples is noted to have $\nu^+ \geq \nu^-$ and is indicated by †. The ν-STAT formulae used 276 E values with $t_1 \geq 1.75$, 400 with $t_2 \leq 0.30$ and $\nu(\text{random}) = 0.055$. The D and S terms from the MDKS calculation and their computed cosine estimate are appended for comparison in the three rightmost columns. Although there are numerous entries in this and subsequent tables for which $\nu^+ < \nu^-$, which infers a negative cosine estimate, only a small fraction of these, say 25%, will actually be observed to be negative for reasons that are stated in the paragraph following equation (5).

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T	D	S	$\cos\Phi_c$
0	3	4	0	−3	4	0	0	−8	1	−1	−166	0.072	0.036	1.470	π	0.675	0.228	1.10
1	7	30	1	7	−30	−2	−14	0	13	−13	−9	0.019	0.052	1.036	π	0.107	0.181	0.17
1	7	27	1	7	−27	−2	−14	0	42	−42	−9	0.019	0.058	0.736	0*	0.194	0.132	0.32
1	7	6	1	7	−6	−2	−14	0	57	−57	−9	0.000	0.012	0.655	π	0.019	0.104	0.03
1	7	26	1	7	−26	−2	−14	0	75	−75	−9	0.070	0.026	0.585	π	0.510	0.070	0.83
1	7	0	1	7	0	−2	−14	0	92	−92	−9	0.000	0.056	0.538	π	−0.417	0.065	−0.68
9	4	3	−9	4	3	0	−8	−6	24	−24	−272	0.019	0.046	0.538	π*	−0.044	0.027	−0.07
5	0	6	5	0	6	−10	0	−12	15	−15	−452	0.000	0.037	0.527	0	−0.476	−0.056	−0.78
2	1	37	2	1	−37	−4	−2	0	49	−49	−74	0.041	0.028	0.512	π	0.305	0.042	0.50
0	2	4	0	−2	4	0	0	−8	40	−40	−166	0.039	0.045	0.502	0*	0.377	0.266	0.62
0	6	30	0	6	−30	0	−12	0	23	−23	−424	0.037	0.051	0.493	0*	0.298	0.162	0.49
2	7	27	2	7	−27	−4	−14	0	118	−118	−10	0.043	0.069	0.475	π*	0.269	0.181	0.44
1	1	35	1	1	−35	−2	−2	0	72	−72	−47	0.000	0.032	0.467	0	−0.093	0.041	−0.15
3	14	2	3	−14	2	−6	0	−4	43	−43	−195	0.023	0.040	0.462	0	0.213	0.087	0.35
1	13	1	1	−13	1	−2	0	−2	35	−35	−322	0.056	0.000	0.455	0	0.420	0.149	0.69
1	7	10	1	7	−10	−2	−14	0	145	−145	−9	0.072	0.035	0.454	π	0.606	0.109	0.99
1	1	0	1	1	0	−2	−2	0	94	−94	−47	0.055	0.032	0.430	π*†	0.332	0.264	0.54
1	7	32	1	7	−32	−2	−14	0	182	−182	−9	0.000	0.055	0.400	π	0.040	0.161	0.07
2	7	5	1	7	−5	−4	−14	0	165	−165	−10	0.000	0.043	0.398	π*	0.261	0.070	−0.43
0	1	20	0	−1	20	0	0	−40	63	−63	−201	0.000	0.041	0.396	π*	0.431	0.059	−0.70
2	7	29	2	7	−29	−4	−14	0	197	−197	−10	0.000	0.045	0.375	π*	−0.004	0.106	0.00
0	6	31	0	6	−31	0	−12	0	51	−51	−424	0.082	0.076	0.363	π	0.483	0.084	0.79
4	2	0	4	2	0	−8	−4	0	74	−74	−234	0.029	0.060	0.361	0	0.106	0.023	0.17
4	9	6	4	−9	6	−8	0	−12	150	−150	−44	0.018	0.024	0.357	π	−0.046	0.003	−0.08
2	7	26	2	7	−26	−4	−14	0	224	−224	−10	0.048	0.043	0.352	0	0.441	0.068	0.72
2	7	8	2	7	−8	−4	−14	0	222	−222	−10	0.070	0.052	0.352	0	0.534	0.213	0.87
2	14	10	2	−14	10	−4	0	−20	95	−95	−173	0.000	0.025	0.347	0	−0.219	−0.013	−0.36
3	6	11	3	6	−11	−6	−12	0	251	−251	−7	0.050	0.014	0.343	0	0.520	0.069	0.85
9	6	13	−9	6	13	0	−12	−26	136	−136	−108	0.000	0.010	0.331	π*	−0.157	0.001	−0.26
1	1	27	1	1	−27	−2	−2	0	176	−176	−47	0.066	0.050	0.324	0	0.554	0.078	0.91
4	2	16	4	2	−16	−8	−4	0	100	−100	−234	0.000	0.036	0.324	0	−0.146	0.008	−0.24
1	3	4	1	3	−4	−2	−6	0	107	−107	−219	0.076	0.052	0.315	π	0.505	0.039	0.83
1	7	3	1	7	−3	−2	−14	0	291	−291	−9	0.027	0.071	0.314	0*	−0.004	0.079	0.00
2	7	28	2	7	−28	−4	−14	0	286	−286	−10	0.035	0.028	0.309	0	0.316	0.159	0.52
2	7	32	2	7	−32	−4	−14	0	279	−279	−10	0.024	0.040	0.309	0	0.139	0.158	0.23
8	0	20	−8	0	20	0	0	−40	126	−126	−201	0.000	0.065	0.307	0	−0.267	0.172	−0.44
5	2	8	−5	2	8	0	−4	−16	210	−210	−46	0.013	0.042	0.306	π	−0.099	0.007	−0.16
1	1	17	1	1	−17	−2	−2	0	228	−228	−47	0.050	0.016	0.291	0	0.406	0.115	0.66
9	3	1	−9	3	1	0	−6	−2	101	−101	−422	0.000	0.070	0.281	0	−0.304	0.018	−0.50
3	15	14	3	−15	14	−6	0	−28	167	−167	−151	0.000	0.035	0.274	π	−0.084	−0.046	−0.14
1	14	11	1	−14	11	−2	0	−22	256	−256	−55	0.047	0.017	0.266	π	0.237	0.031	0.39
4	0	20	−4	0	20	0	0	−40	173	−173	−201	0.038	0.064	0.256	0	0.042	0.050	0.07
0	6	13	0	6	13	0	−12	−26	221	−221	−108	0.000	0.054	0.255	0	−0.270	0.000	−0.44
5	3	8	5	−3	8	−10	0	−16	131	−131	−478	0.000	0.026	0.252	π*	−0.137	−0.098	−0.22
3	6	10	3	6	−10	−6	−12	0	394	−394	−7	0.032	0.040	0.252	π*	0.188	0.079	0.31
6	0	4	−6	0	4	0	0	−8	195	−195	−166	0.025	0.044	0.251	0*	−0.038	0.019	−0.06

(*b*) The top 46 $\sum_2$ zonal restricted triples. There are 5 aberrant triples, marked *, that have $\nu^+ \leq \nu^-$. Of the 21 triples that have $\nu^+ \geq \nu^-$, all are correctly indicated as being reliable.

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T	D	S	$\cos\Phi_c$
0	7	30	0	7	−27	0	−14	−3	2	16	−17	0.053	0.037	2.782	0	0.348	0.183	0.91
0	3	4	0	−6	30	0	3	−34	1	−23	79	0.039	0.027	2.361	π	0.323	0.130	0.85
0	3	4	0	−4	16	0	1	−20	1	−46	63	0.038	0.036	2.214	π	0.209	0.152	0.55
0	3	4	0	8	−7	0	−11	3	1	33	−99	0.042	0.053	2.170	π	0.170	0.066	0.45
0	3	4	0	−7	27	0	4	−31	1	−16	163	0.031	0.014	2.157	π	0.373	0.166	0.98
0	3	4	0	6	−30	0	−9	26	1	23	−158	0.103	0.036	2.120	π	0.810	0.123	2.13
6	12	0	−4	−14	0	−2	2	0	7	−10	47	0.078	0.029	2.098	0	0.792	0.143	2.08
0	7	30	0	8	−36	0	−15	6	2	50	−64	0.024	0.000	2.039	π	0.206	0.081	0.54
6	12	0	−2	−14	0	−4	2	0	7	−9	74	0.054	0.028	2.014	0	0.550	0.108	1.45
0	3	4	0	7	−14	0	−10	10	1	11	−367	0.015	0.043	1.912	π	0.038	0.099	0.10
0	7	30	0	2	−4	0	−9	−26	2	40	−158	0.052	0.036	1.865	0	0.553	0.216	1.45
0	3	4	0	−7	14	0	4	−18	1	−11	455	0.018	0.046	1.819	π	0.122	0.109	0.32
0	7	30	0	4	−16	0	−11	−14	2	46	−213	0.016	0.006	1.719	0	0.002	0.070	0.00

Table 1(*b*) (*cont.*)

H			−*K*			*K*−*H*			*N*1	*N*2	*N*3	ν^+	ν^-	*A*	*T*	*D*	*S*	cos Φ_c
5	0	6	−7	0	16	2	0	−22	15	−21	−55	0.018	0.015	1.718	π	−0.032	−0.033	−0.08
0	7	30	0	−8	7	0	1	−37	2	−33	285	0.031	0.049	1.709	π	0.152	0.128	0.40
0	7	27	0	6	−30	0	−13	3	16	23	−52	0.043	0.027	1.706	0	0.359	0.173	0.94
0	3	4	0	−7	−25	0	4	21	1	−54	296	0.046	0.024	1.673	0	0.334	0.091	0.88
0	3	4	0	11	−3	0	−14	−1	1	99	−181	0.005	0.020	1.652	0	−0.068	0.070	−0.18
0	3	4	0	3	9	0	−6	−13	1	81	−221	0.021	0.038	1.645	0	0.179	0.104	0.47
6	0	3	−8	0	19	2	0	−22	18	−31	−55	0.000	0.025	1.635	0	−0.096	0.009	−0.25
6	0	1	−4	0	−32	−2	0	31	5	−68	73	0.009	0.027	1.610	π^*	0.072	0.034	0.19
0	3	4	0	13	−3	0	−16	−1	1	52	−358	0.015	0.035	1.596	0	0.073	0.138	0.19
2	14	0	−3	−15	0	1	1	0	9	−41	94	0.054	0.052	1.584	0	0.581	0.188	1.53
0	7	27	0	−8	−7	0	1	−20	16	−33	63	0.042	0.051	1.579	π	0.266	0.112	0.70
0	3	4	0	−2	−22	0	−1	18	1	−80	−139	0.016	0.032	1.544	π^*	0.153	0.134	0.40
4	14	0	−3	−15	0	−1	1	0	10	−41	94	0.089	0.045	1.530	π	0.896	0.232	2.36
0	7	30	0	−2	4	0	−5	−34	2	−40	−481	0.036	0.048	1.488	0	0.537	0.250	1.41
0	7	30	0	−3	−9	0	−4	−21	2	−81	−296	0.014	0.041	1.472	0	−0.094	0.049	−0.25
0	5	25	0	7	−32	0	−12	7	20	27	−120	0.038	0.024	1.461	0	0.352	0.087	0.93
0	3	4	0	2	27	0	−5	−31	1	153	−257	0.025	0.029	1.458	0*	0.015	0.072	0.04
0	5	25	0	−2	4	0	−3	−29	20	−40	−88	0.032	0.025	1.449	0	0.240	0.136	0.63
0	3	4	0	5	−10	0	−8	6	1	164	−272	0.050	0.027	1.412	π	0.486	0.115	1.28
0	14	3	0	−2	4	0	−12	−7	17	−40	−120	0.024	0.057	1.398	0	0.153	0.149	0.40
4	0	37	−5	0	−35	1	0	−2	8	−14	−392	0.018	0.029	1.393	π	0.120	0.185	0.32
0	7	14	0	−2	−4	0	−5	−10	11	−40	−164	0.014	0.044	1.380	0	0.172	0.194	0.45
5	0	6	−7	0	−16	2	0	10	15	−21	230	0.017	0.033	1.371	0	−0.105	−0.062	−0.28
0	3	4	0	−1	−37	0	−2	33	1	−285	−320	0.015	0.019	1.249	π	−0.077	0.059	−0.20
0	3	4	0	−9	−2	0	6	−2	1	−199	422	0.017	0.031	1.242	0	0.070	0.100	0.18
0	4	16	0	−7	−25	0	3	9	46	−54	81	0.045	0.057	1.241	0*	0.117	0.076	0.31
0	7	27	0	−2	4	0	−5	−31	16	−40	−257	0.044	0.031	1.240	0	0.494	0.197	1.30
2	2	0	−2	0	−22	0	−2	22	47	−55	−80	0.040	0.025	1.236	0	0.281	0.047	0.74
6	0	1	−8	0	9	2	0	−10	5	−132	−230	0.000	0.032	1.234	0	−0.239	−0.018	−0.63
6	0	1	2	0	−22	−8	0	21	5	55	423	0.061	0.045	1.222	π	0.515	0.032	1.35
6	0	1	−10	0	−10	4	0	9	5	−172	178	0.000	0.018	1.205	0	−0.042	−0.076	−0.11
0	7	30	0	−8	−12	0	1	−18	2	−255	319	0.029	0.047	1.203	π^*	0.145	0.074	0.38
0	7	14	0	−4	−16	0	−3	2	11	−46	−299	0.029	0.033	1.202	π	0.276	0.087	0.73

for their conditional averages but have dissimilar ν^+ or ν^- values computed from the distribution of the same $|E|$ magnitudes. No other comparisons are made here between the ν-STAT formulae and similar cosine-invariant estimation procedures for triples other than to say that there do not appear to have been any successful *ab initio* applications to any structure containing more than 80 non-hydrogen light atoms by any of the earlier methods. Of the five structural examples cited in this paper, all contain more than 80 atoms and, apart from the ILED example, each required the ν-STAT analysis to facilitate the structure-determination process.

The remaining four structural examples represent problems that cannot easily be determined by standard direct-phasing procedures. The 92-atom $P3_2$ structure of ACE, for example, will not readily yield to phase-permutation methods and, although a solution can be found by random-phasing methods, thousands of trials are generally required. Table 2 provides an answer as to why this particular structure is so troublesome. Of the thousands of triples used to solve this structure, a small percentage are inconsistent (Han & Langs, 1988). The two strongest $\sum_2$ triple indications for this structure, with *A* values of 3.31, represent an inconsistent pair (see Table 2). Reflection numbers 1, 10 and 11 can form two non-identical triples, one for which $\varphi_1+\varphi_{10}+\varphi_{11}+120°$ is assumed to equal zero (mod 2π) and the other for which $\varphi_1+\varphi_{10}+\varphi_{11}+240°$ is assumed to be zero. Clearly, it is impossible for both of these triples to be satisfied. In fact, from the solved structure it is known that the first invariant ($\cos\Phi=0.84$) sums to 32°, while the second ($\cos\Phi=-0.88$) sums to 152°. If one does nothing to these data, the two indications will average and a 92° phase error will be propagated at the onset of the phase determination. The ν-STAT cosine estimates, however, correctly indicate that the first triple is the more reliable of the two and with the second triple discarded the phase error may be reduced to 32°. Thus, by employment of a trivial amount of phase-invariant editing, the structure can be solved in a manageable number of trials.

The third example to be discussed is the *P*1 structure of ENKA. The problem with this structure is that, if a basis set is selected from among the strongest *E* values, the phases rapidly refine to over-consistency to produce 'U-atom' solutions. Four different phase-permutation trials, each exploring 8190 phase sets, all failed to produce a single solution as all NQEST figures of merit (De Titta, Edmonds, Langs & Hauptman, 1975) turned positive and most approached a value of +1.0 within three or four refinement cycles. Random-phasing methods experienced the same fate, even when special weights (Hull & Irwin, 1978) were employed to mitigate such ill conditioning. In 2000 random-phase sets, no solutions were produced that had a stable negative NQEST value. Table 3 lists the top 98 $\sum_2$ triples for this structure that have *A* values greater than 2.5. In addition to the 10 aberrant triples

Table 2. *Partial listing of $\sum_2$ triples for ACE ($P3_2$, N=92 atoms)*

Most of the pairs of triples in the partial listing shown below are either consistent or inconsistent, depending on the value of T. The strongest two triples in the entire $P3_2$ data set are marked ‡; however, it is physically impossible for them both to be correct. The value of the first invariant exactly equals that of the second in the list minus 120°. Note that the particular pairs of triples marked § are not symmetry related and their values are totally independent of one another. The calculation used 212 E values with $t_1 \geq 1.75$, 300 with $t_2 \leq 0.25$ and ν(random)=0.061.

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T(°)	cos Φ
1	3	2	−3	−13	1	2	10	−3	1	11	10	0.037	0.032	3.31	120	0.84‡
1	3	2	2	−13	−3	−3	10	1	1	11	10	0.030	0.054	3.31	240	−0.88‡
2	3	2	−4	−3	2	2	0	−4	3	125	4	0.072	0.039	2.24	0	0.97
2	3	2	2	−3	−4	−4	0	2	3	125	4	0.077	2.24	0.032	0	0.97
2	0	2	2	−1	−2	−4	1	0	4	20	133	0.080	0.042	1.91	240	0.98§
2	0	2	−2	1	2	0	−1	−4	4	−20	−133	0.087	0.050	1.91	240	0.87§
1	3	2	2	−2	−3	−3	−1	1	1	134	107	0.044	0.057	1.71	120	−0.49
1	3	2	−3	−2	1	2	−1	−3	1	134	107	0.028	0.055	1.71	240	−0.51
1	−13	2	2	−11	−3	−3	24	1	11	30	173	0.043	0.031	1.58	240	0.89
1	−13	2	−3	−11	1	2	24	−3	11	30	173	0.038	0.056	1.58	120	−0.05
0	−20	6	−3	20	0	3	0	−6	28	48	155	0.054	0.050	1.41	240	0.43§
0	−20	6	3	20	−3	−3	0	−3	28	48	−155	0.055	0.046	1.41	120	0.74§
1	10	2	2	−11	−3	−3	1	1	10	30	291	0.044	0.040	1.40	0	−0.98
1	10	2	−3	−11	1	2	1	−3	10	30	291	0.043	0.044	1.40	0	−0.98
1	0	1	1	−23	−2	−2	23	1	31	70	158	0.61	0.051	1.29	120	0.01
1	0	1	−2	−23	1	1	23	−2	31	70	158	0.098	0.043	1.29	240	0.86
0	−1	2	1	1	−1	−1	0	−1	20	356	−31	0.100	0.044	1.28	240	0.98§
0	−1	2	−1	1	0	1	0	−2	20	356	31	0.103	0.040	1.28	120	0.97§

Table 3. *List of $\sum_2$ triples for structure ENKA ($P1$, N=96 atoms)*

Asterisks (*) mark ten aberrant triples (cos $\Phi \leq 0.5$) for which ν^- exceeds ν^+ by more than 0.01 units. A † in the rightmost column identifies triples for which ν^+ exceeds ν^- by 0.03 and may confidently be accepted as reliable. There were 434 E values with $t_1 \geq 1.65$, 600 E values with $t_2 \leq 0.36$ and ν (random)=0.071.

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	cos Φ
11	2	0	−1	2	−2	−10	−4	2	2	−24	−28	0.039	0.049	3.77	0.65
9	6	−1	−12	11	−1	3	−17	2	1	−36	47	0.053	0.033	3.68	0.89
0	16	−3	10	−12	3	−10	−4	0	10	11	−53	0.97	0.008	3.60	0.99‡
0	16	−3	−7	−2	6	7	−14	−3	10	−12	41	0.032	0.043	3.52	0.91
1	1	−2	−11	−6	4	10	5	−2	5	−25	48	0.069	0.023	3.43	1.00†
11	2	0	1	−17	2	−12	15	−2	2	33	−57	0.077	0.058	3.41	0.65
9	−9	4	−10	12	−3	1	−3	−1	7	−11	100	0.009	0.051	3.33	0.66
4	−11	−5	1	2	−1	−5	9	6	8	9	−142	0.064	0.063	3.27	0.81
1	1	−2	−12	9	−1	11	−10	3	5	−22	69	0.044	0.060	3.26	0.46*
8	4	−5	−2	−16	0	−6	12	5	4	−23	−70	0.105	0.057	3.25	0.96†
9	6	−1	−1	2	−1	−8	−8	3	1	−24	−164	0.056	0.039	3.19	0.96
4	−11	−5	−1	2	1	−3	13	4	8	−9	−176	0.060	0.047	3.19	1.00†
8	4	−5	−12	9	−1	4	−13	6	4	−22	80	0.074	0.039	3.19	0.96†
1	1	−2	1	−14	−1	−2	13	3	5	31	−64	0.045	0.077	3.17	0.97
11	2	0	0	−16	3	−11	14	−3	2	−10	−281	0.049	0.067	3.12	0.02*
1	1	−2	3	9	4	−4	−10	−2	5	38	−63	0.036	0.033	3.12	1.00
4	−11	−5	−7	4	0	3	7	5	8	−40	55	0.044	0.069	3.12	0.84
4	−1	7	1	2	−1	−5	−1	−6	3	9	−269	0.065	0.053	3.09	0.85
8	0	5	1	2	−1	−9	−2	−4	6	9	−229	0.050	0.081	3.08	0.87
8	4	−5	−5	−5	6	−3	1	−1	4	−16	−124	0.089	0.093	3.08	0.83
8	4	−5	−7	−2	6	−1	−2	−1	4	−12	−150	0.083	0.048	3.07	0.78†
9	6	−1	−1	−1	2	−8	−5	−1	1	−5	−523	0.035	0.033	3.07	0.99
8	0	5	−1	−2	1	−7	2	−6	6	−9	−246	0.092	0.071	3.06	0.69
8	−6	−6	−12	9	−1	4	−3	7	18	−22	43	0.079	0.066	3.06	0.38
8	4	−5	−1	−2	1	−7	−2	4	4	−9	−275	0.137	0.035	3.04	1.00†
9	−11	4	−12	17	−1	3	−6	−3	15	−17	59	0.018	0.031	3.03	0.90
4	−11	−5	2	16	0	−6	−5	5	8	23	−92	0.079	0.050	3.03	0.87
11	2	0	−1	1	−2	−10	−3	2	2	−58	−84	0.065	0.051	3.00	0.99
8	4	−5	−4	11	5	−4	−15	0	4	−8	−321	0.082	0.068	3.00	0.82
4	−1	7	4	−11	−5	−8	12	−2	3	8	−340	0.066	0.038	2.99	0.36
9	6	−1	−9	11	−4	0	−17	5	1	−15	−361	0.032	0.032	2.95	0.97
9	−9	4	−13	6	3	4	3	−7	7	−27	105	0.054	0.066	2.94	0.12*
1	1	−2	−1	14	1	0	−15	1	5	−31	−126	0.076	0.035	2.91	1.00†
4	−11	−5	−7	−2	6	3	13	−1	8	−12	183	0.033	0.068	2.91	0.85
1	2	−1	−12	8	−2	11	−10	3	9	−50	69	0.000	0.072	2.91	0.77
9	−11	4	−6	18	1	−3	−7	−5	15	−30	−55	0.031	0.025	2.91	0.94
1	2	−1	1	−2	2	−2	0	−1	9	24	−122	0.036	0.058	2.90	0.94
9	6	−1	−8	−6	−1	−1	0	2	1	−39	−219	0.071	0.051	2.90	0.98
9	6	−1	−10	−3	2	1	−3	−1	1	−84	100	0.044	0.065	2.88	0.19*
9	6	−1	−1	−9	6	−8	3	−5	1	−79	−110	0.092	0.061	2.88	1.00†
4	−1	7	0	−3	−6	−4	4	−1	3	−62	−74	0.035	0.029	2.87	0.37
10	−12	3	−7	14	3	−3	−2	−6	11	−41	−71	0.079	0.045	2.86	0.78†

Table 3 (*cont.*)

	H			$-K$		$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	cos Φ
1	2	−1	−6	14	−3	5	−16	4	9	−21	146	0.100	0.046	2.86	0.99†
1	1	−2	−3	17	−2	2	−18	4	5	−47	113	0.062	0.040	2.83	0.96
7	2	−6	−11	1	0	4	−3	6	12	−37	60	0.031	0.041	2.83	0.72
1	2	−1	10	−15	2	−11	13	−1	9	51	−86	0.072	0.028	2.82	0.95†
1	2	−1	4	−3	7	−5	1	−6	9	43	−101	0.114	0.058	2.80	0.98†
4	−1	7	−1	−2	1	−3	3	−8	3	−9	−469	0.046	0.041	2.78	0.90
0	16	−3	10	−13	1	−10	−3	2	10	49	−84	0.032	0.063	2.78	−0.71*
8	4	−5	−9	−5	4	1	1	1	4	−26	191	0.067	0.060	2.77	0.99
1	2	−1	−2	−16	0	1	14	1	9	−23	189	0.078	0.064	2.76	0.86
10	−12	3	0	16	−1	−10	−4	−2	11	13	−221	0.083	0.045	2.74	0.90†
11	2	0	−10	−2	−2	−1	0	2	2	−45	−219	0.054	0.070	2.72	−0.03*
9	5	−4	−10	−4	2	1	−1	2	26	−28	58	0.031	0.024	2.71	0.78
9	−9	4	−1	−2	1	−8	11	−5	7	−9	−467	0.028	0.062	2.70	0.99
0	16	−1	−2	−16	0	2	0	1	13	−23	122	0.073	0.053	2.69	0.70
9	6	−1	1	−14	−1	−10	8	2	1	31	−403	0.063	0.029	2.69	0.84†
11	2	0	−11	13	−1	0	−15	1	2	−86	−126	0.082	0.039	2.68	0.98†
1	2	−1	−10	15	−2	9	−17	3	9	−51	127	0.042	0.051	2.67	0.58
8	4	−5	−1	2	−2	−7	−6	7	4	−24	−267	0.046	0.082	2.67	0.17*
8	0	5	−8	−6	−1	0	6	−4	6	−39	179	0.029	0.049	2.66	0.81
7	2	−6	−10	−4	0	3	2	6	12	−53	71	0.013	0.019	2.66	1.00
9	6	−1	−3	−14	2	−6	8	−1	1	−118	−143	0.023	0.049	2.65	0.99
1	2	−1	−11	3	1	10	−5	0	9	−65	93	0.030	0.040	2.65	0.70
9	6	−1	−9	−5	4	0	−1	−3	1	−26	−442	0.061	0.035	0.265	0.99
9	6	−1	−12	6	0	3	−12	1	1	−54	323	0.024	0.053	2.64	1.00
1	−2	2	9	5	−4	−10	−3	2	24	26	−84	0.075	0.012	2.62	0.96†
8	4	−5	−4	−10	−2	−4	6	7	4	−63	−129	0.052	0.073	2.62	0.96
1	2	−1	−10	−2	−2	9	0	3	9	−45	170	0.079	0.039	2.61	0.92†
1	−2	2	−7	14	3	6	−12	−5	24	−41	70	0.042	0.011	2.60	1.00
6	−18	−1	−7	19	−1	1	−1	2	30	−46	58	0.061	0.018	2.59	0.90†
8	4	−5	0	−16	1	−8	12	4	4	−13	−427	0.133	0.068	2.58	0.98†
9	−9	4	−2	8	0	−7	1	−4	7	−32	−254	0.012	0.038	2.58	0.88
9	6	−1	−9	−7	3	0	1	−2	1	−106	193	0.040	0.060	2.57	−0.10*
11	2	0	−9	−7	3	−2	5	−3	2	−106	−137	0.013	0.032	2.57	0.93
0	16	−3	1	−14	−1	−1	−2	4	10	31	−209	0.036	0.018	2.57	0.94
9	6	−1	−3	−2	−6	−6	−4	7	1	−71	−277	0.056	0.063	2.57	0.92
1	−2	2	13	−6	−3	−14	8	1	24	27	−95	0.012	0.037	2.56	0.07*
11	2	0	−4	3	−6	−7	−5	6	2	−60	−232	0.058	0.027	2.56	0.98†
5	5	−6	8	−14	3	−13	9	3	16	44	−89	0.148	0.056	2.55	0.96†
4	−1	7	−3	3	−6	−1	−2	−1	3	−73	−150	0.041	0.075	2.55	1.00
12	−17	1	−10	−4	2	−2	21	−3	17	−28	−120	0.000	0.045	2.55	0.63
1	2	−1	−3	15	−4	2	−17	5	9	−14	393	0.060	0.030	2.55	0.74†
9	−9	4	3	−6	−3	−12	15	−1	7	59	−172	0.018	0.031	2.54	0.99
1	1	−2	−10	−4	2	9	3	0	5	−28	314	0.048	0.034	2.54	1.00
0	16	−1	1	−2	2	−1	−14	−1	13	24	−189	0.083	0.062	2.53	1.00
8	4	−5	3	−15	4	−11	11	1	4	14	−446	0.064	0.044	2.53	1.00
9	6	−1	1	−7	3	−10	1	−2	1	154	−157	0.036	0.035	2.52	0.74
8	0	5	−3	−7	−5	−5	7	0	6	−55	−207	0.064	0.057	2.52	0.93
1	−14	−1	−2	5	7	1	9	−6	31	−34	79	0.063	0.067	2.52	0.93
0	16	−1	3	−17	2	−3	1	−1	13	47	−124	0.040	0.064	2.51	0.12*
6	−18	−1	−2	5	7	−4	13	−6	30	−34	−80	0.018	0.059	2.51	0.92
1	2	−1	6	−12	−5	−7	10	6	9	70	−141	0.066	0.070	2.50	0.97
4	−11	−5	−4	4	−1	0	7	6	8	−74	130	0.024	0.053	2.50	0.84
3	−15	4	0	−3	−6	−3	18	2	14	−62	−78	0.020	0.044	2.50	0.99
8	4	−5	−3	−13	−1	−5	9	6	4	−83	−142	0.058	0.056	2.50	0.98
11	2	0	−11	1	0	0	−3	0	2	−37	−416	0.049	0.045	2.50	0.88
1	1	−2	−11	1	2	10	−2	0	5	−77	160	0.038	0.038	2.50	0.88

(*) that are correctly identified, $\nu^+ \leq \nu^-$ and cos $\Phi \leq$ 0.5, 23 strong positive cosine indications are noted (†) for which ν^+ exceeds ν^- by 0.03 units. A basis set of 29 reflections was extracted from these 23 triples that could be expressed in terms of three origin definers (phases 5, 8 and 9), 12 magic-integer symbols selecting the enantiomorph (phases 1, 2, 4, 10, 11, 12, 13, 21, 22, 23, 25 and 31) and 14 symbolic relationships defining 14 other phases. Standard tangent-formula refinement of 8190 phase sets [$N = 2$ magic-integer scheme (Main, 1977)] produced four solutions with a stable negative NQEST value of −0.20 from which more than 90 atoms of the structure were discernible. Six other partial solutions were also obtained for which NQEST was stable in the vicinity of ∼−0.05 and from which significant chemically recognizable segments of the structure could be identified for fragment recycling. None of the earlier trials produced any solutions that were as readily interpretable as the worst of this latter group.

The fourth structure, HEXIL, is included as an embarrassing example of shaken confidence in the active use of phase-invariant information. At least a dozen or more attempts were made to solve this structure with the methods described above; a number of trials used basis sets of up to 45 reflections and

Table 4. *List of triples for HEXIL* ($P2_12_12_1$, $N=126$)

(*a*) The top 25 $\sum_1$ triples. Of 15 aberrant (*) triples indicated, 12 have $v^+ \leq v^-$, only 3 are incorrectly indicated (†). 430 E values with $t_1 \geq 1.75$, 600 with $t_2 \leq 0.35$, ν(random)=0.067.

H			$-K$			$K-H$			$N1$	$N2$	$N3$	v^+	v^-	A	T
4	0	2	4	0	2	−8	0	−4	54	54	−4	0.030	0.051	0.886	0
4	0	9	−4	0	9	0	0	−18	9	−9	−532	0.040	0.089	0.754	0*
5	0	4	5	0	4	−10	0	−8	119	119	−7	0.023	0.043	0.633	0*
2	0	21	2	0	21	−4	0	−42	71	71	−67	0.016	0.061	0.544	0*
4	4	4	4	−4	4	−8	0	−8	74	−74	−177	0.022	0.040	0.460	0
7	2	11	−7	2	11	0	−4	−22	88	−88	−142	0.007	0.052	0.451	π*
9	0	14	−9	0	14	0	0	−28	48	−48	−387	0.036	0.060	0.435	π
4	2	2	4	−2	2	−8	0	−4	364	−364	−4	0.033	0.048	0.388	0
0	2	11	0	2	11	0	−4	−22	131	131	−142	0.046	0.029	0.383	0
1	4	24	1	4	−24	−2	−8	0	53	−53	−616	0.041	0.040	0.374	π
7	0	12	−7	0	12	0	0	−24	91	−91	−369	0.028	0.034	0.365	π
0	4	38	0	4	−38	0	−8	0	136	136	−301	0.073	0.035	0.325	0*†
0	4	22	0	4	−22	0	−8	0	142	142	−301	0.046	0.064	0.321	0*
3	1	18	3	−1	18	−6	0	−36	313	−313	−39	0.040	0.041	0.302	π*
1	2	21	−1	2	21	0	−4	−42	121	−121	−580	0.045	0.025	0.298	π*†
5	11	4	5	−11	4	−10	0	−8	464	−464	−7	0.000	0.039	0.296	π
4	1	16	4	−1	16	−8	0	−32	196	−196	−186	0.60	0.72	0.291	π*
4	0	20	−4	0	20	0	0	−40	276	−276	−86	0.005	0.052	0.291	0
4	1	5	−4	1	5	0	−2	−10	375	−375	−31	0.051	0.043	0.283	π*†
7	0	20	−7	0	20	0	0	−40	290	−290	−86	0.011	0.044	0.279	π*
2	6	21	2	−6	21	−4	0	−42	324	−324	−67	0.027	0.039	0.277	π*
0	4	24	0	4	−24	0	−8	0	188	188	−301	0.033	0.066	0.274	0*
1	2	11	−1	2	11	0	−4	−22	264	−264	−142	0.035	0.047	0.273	π*
5	0	14	−5	0	14	0	0	−28	182	−182	−387	0.053	0.026	0.265	π
0	11	9	0	−11	9	0	0	−18	172	−172	−532	0.000	0.036	0.255	0*

(*b*) The top 77 zonal restricted $\sum_2$ triples. Of 15 aberrant (*) triples in this list, all are flagged with v^+ less than v^-. The reflection with an incorrect phase is indicated by ‡ (see text).

H			$-K$			$K-H$			$N1$	$N2$	$N3$	v^+	v^-	A	T
8	0	4	−4	0	−13	−4	0	9	4	−5	9	0.057	0.049	3.962	π
9	0	10	−3	0	−46	−6	0	36	1	−33	39	0.048	0.056	3.377	0
2	0	45	4	0	−9	−6	0	−36	2	9	−39	0.054	0.073	3.303	0
9	0	10	−10	0	−8	1	0	−2	1	−7	−323	0.067	0.036	2.969	π
9	0	10	−7	0	−31	−2	0	21	1	−65	71	0.072	0.077	2.780	π
2	0	45	5	0	−11	−7	0	−34	2	30	−51	0.029	0.059	2.671	π
9	0	10	1	0	−27	−10	0	17	1	12	339	0.088	0.041	2.652	0
9	0	10	−4	0	−2	−5	0	−8	1	−54	−116	0.044	0.069	2.639	0
8	0	4	−1	0	27	−7	0	−31	4	−12	−65	0.070	0.031	2.564	π
4	0	13	−9	0	−2	5	0	−11	5	−26	−30	0.061	0.061	2.525	0
4	0	19	−1	0	27	−3	0	−46	10	−12	−33	0.041	0.048	2.442	π
4	0	9	−9	0	2	5	0	−11	9	−26	−30	0.056	0.059	2.409	π
2	0	45	−8	0	−4	6	0	−41	2	−4	−489	0.044	0.036	2.397	π
4	0	13	1	0	−47	−5	0	34	5	20	46	0.034	0.078	2.389	0*
2	0	45	−9	0	−14	7	0	−31	2	−48	−65	0.078	0.053	2.338	0
0	3	41	0	−9	3	0	7	−44	8	−13	59	0.064	0.036	2.337	π
4	0	9	3	0	25	−7	0	−34	9	14	−51	0.020	0.064	2.336	0
4	0	13	−9	0	4	5	0	−17	5	−22	−64	0.039	0.042	2.273	π*
4	0	13	3	0	−25	−7	0	12	5	14	91	0.033	0.044	2.273	0
9	0	10	−8	0	13	−1	0	−23	1	−89	−213	0.019	0.053	2.196	0
1	0	27	3	0	−25	−4	0	−2	12	14	−54	0.034	0.047	2.128	π
9	0	10	−9	0	14	0	0	−24	1	−48	−369	0.015	0.068	2.126	π
2	0	45	−4	0	−9	2	0	−36	2	−9	−500	0.028	0.063	2.110	0
2	0	45	−4	0	−19	2	0	−26	2	−10	−423	0.046	0.061	2.105	0*
2	0	45	−1	0	−47	−1	0	2	2	−20	323	0.065	0.082	2.080	π
4	0	13	4	0	19	−8	0	−32	5	10	−186	0.066	0.069	2.079	0
9	0	10	−1	0	−42	−8	0	32	1	−153	186	0.086	0.064	2.068	0
9	0	10	−7	0	34	−2	0	−44	1	−51	−446	0.063	0.048	2.035	π
4	0	19	−9	0	−2	5	0	−17	10	−26	−64	0.034	0.067	2.020	0
4	0	19	1	0	−27	−5	0	8	10	12	116	0.022	0.035	2.015	0*
4	0	13	1	0	−27	−5	0	14	5	12	182	0.056	0.058	2.011	0
1	0	47	5	0	−11	−6	0	−36	20	30	−39	0.062	0.060	1.976	π
10	0	8	−3	0	25	−7	0	−33	7	−14	−179	0.005	0.027	1.968	π
2	0	45	3	0	−49	−5	0	4	2	114	119	0.070	0.053	1.935	0
4	0	13	5	0	11	−9	0	−24	5	30	−148	0.030	0.061	1.920	0
2	0	45	−7	0	−31	5	0	−14	2	−65	−182	0.025	0.048	1.913	π
2	0	45	7	0	−34	−9	0	−11	2	51	−235	0.053	0.059	1.882	π*
9	0	10	−5	0	8	−4	0	−18	1	−116	−403	0.060	0.064	1.876	π
9	0	10	−8	0	−8	−1	0	−2	1	−177	−323	0.061	0.052	1.825	0
4	0	13	3	0	−46	−7	0	33	5	33	179	0.034	0.039	1.810	π
8	0	4	−9	0	−2	1	0	−2	4	−26	−323	0.073	0.041	1.810	π
0	9	3	0	−2	10	0	−7	−13	13	−31	−102	0.048	0.038	1.786	0

Table 4(*b*) (*cont.*)

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T
4	0	9	−9	0	−4	5	0	−5	9	−22	−220	0.048	0.063	1.763	0
10	0	8	−7	0	34	−3	0	−42	7	−51	−113	0.028	0.060	1.763	π
10	0	8	−2	0	−21	−8	0	13	7	−71	89	0.014	0.038	1.753	π
4	0	9	−4	0	19	0	0	−28	9	−10	−387	0.033	0.050	1.714	0*
4	0	19	5	0	−34	−9	0	15	10	46	107	0.033	0.058	1.695	π*
4	0	9	4	0	−19	−8	0	10	9	10	409	0.051	0.051	1.695	0
4	0	13	−4	−5	0	0	5	−13	5	−16	411	0.047	0.052	1.670	0*‡
4	0	9	1	0	−47	−5	0	38	9	20	322	0.045	0.066	1.665	0
4	0	13	−5	0	−11	1	0	−2	5	−30	−323	0.082	0.062	1.662	π
3	0	25	1	0	−47	−4	0	22	14	20	211	0.024	0.052	1.650	π
2	0	45	−2	0	−21	0	0	−24	2	−71	−369	0.024	0.036	1.649	0
1	0	27	−9	0	−14	8	0	−13	12	−48	−89	0.045	0.066	1.648	π
2	0	45	−5	0	−17	3	0	−28	2	−64	−431	0.058	0.027	1.614	π
1	0	27	−10	0	−5	9	0	−22	12	−23	−239	0.031	0.042	1.595	π
5	2	0	−5	0	−11	0	−2	11	29	−30	−131	0.053	0.059	1.588	π*
4	0	9	−5	0	−11	1	0	2	9	−30	323	0.067	0.067	1.586	0
1	0	47	7	0	−34	−8	0	−13	20	51	−89	0.071	0.069	1.577	π
3	0	25	−10	0	−5	7	0	−20	14	−23	−290	0.006	0.027	1.539	π*
8	0	4	2	0	−21	−10	0	17	4	71	339	0.039	0.037	1.532	π
3	0	25	2	0	−21	−5	0	−4	14	71	−119	0.027	0.021	1.528	0
4	0	9	−9	0	−14	5	0	5	9	−48	220	0.033	0.056	1.527	0*
2	0	45	3	0	−50	−5	0	5	2	201	220	0.031	0.052	1.523	π
1	0	47	−4	0	2	3	0	−49	20	−54	−114	0.056	0.081	1.511	0
4	0	19	−9	0	−15	5	0	−4	10	−107	−119	0.038	0.049	1.497	π
10	0	5	−3	0	−17	−7	0	12	23	−75	91	0.009	0.031	1.488	π
4	0	19	−1	0	−47	−3	0	28	10	−20	431	0.053	0.045	1.484	π
8	0	4	−5	0	−34	−3	0	30	4	−46	475	0.025	0.067	1.483	π
4	0	9	2	0	21	−6	0	−30	9	71	−207	0.027	0.042	1.482	0
2	0	45	−8	0	−32	6	0	−13	2	−186	−289	0.026	0.056	1.476	π*
10	0	5	−3	0	17	−7	0	−22	23	−75	−95	0.009	0.026	1.475	π
4	0	19	−9	0	−14	5	0	−5	10	−48	−220	0.032	0.053	1.463	0
3	0	25	6	0	−36	−9	0	11	14	39	235	0.039	0.046	1.449	0*
9	0	2	−4	0	2	−5	0	−4	26	−54	−119	0.020	0.075	1.449	0*
8	0	4	−3	0	−42	−5	0	38	4	−113	322	0.021	0.044	1.447	π*
0	2	10	0	−2	−50	0	0	40	31	−72	86	0.065	0.033	1.444	0

required the examination of as many as 64 000 phase combinations. The structure was eventually solved in a random-phasing trial with a basis set of 18 substantialized phases (Woolfson, 1954). Of the 2560 phase sets examined, the one with the best figures of merit (NQEST = −0.097, $R = 0.257$) produced an E map from which a peptide-like fragment of ten atoms could be discerned. Fragment recycling produced five solutions in ten random-phasing trials (Yao, 1983); a total of 100 atoms were located in the next E map. Of the ten atoms selected from the original peptide test fragment, nine corresponded to real atomic positions and one was chosen in error. A re-examination of the previous cosine-invariant-analysis-based trials revealed that one trial had 36 of 37 zonal basis-set phases correct for one of the 256 permutation combinations allowed by the eight symbols involved. The lone phase error affected reflection no. 411, which was imprudently phased through the 49th triple indicated by ‡ in Table 4(*b*). When all 37 phases were tested with their correct values, the tangent-formula refinement still failed to produce a solution! An inspection of the basis-set reflection indices revealed that only five phases had non-zero k values, so the six largest unrestricted E values were added to the basis set to increase the number of phases from 5 to 11. A tangent-formula calculation was performed for the 252 magic-integer permutations for these 6 additional phases while the 37 zonal phases were held fixed at their known values. Still no solution was obtained. When the 37 zonal phases were introduced as a known basis set into the random-phasing process, a solution appeared on average once in only every 100 random trials. Thus, if the ten-atom fragment had not been recognized, it might have been necessary to sort through 256×100 random phase sets to obtain a clear-cut solution by these methods.

The last example, GRAMA, rebuilds our confidence in these methods. Table 5 lists 100 $\sum_1$ and $\sum_2$ triples analyzed for this structure. Of the 19 $\sum_1$ triples, 7 are aberrant (*) and correctly flagged with $\nu^+ \leq \nu^-$ and, of the remaining 81 $\sum_2$ triples, 14 are aberrant and only the two marked (†) have been incorrectly flagged. The first of these, $\varphi_2 + \varphi_9 - \varphi_{442}$, is a marginal positive indication ($\nu^+ = 0.053$, $\nu^- = 0.052$) that was not trusted in the original analysis of this structure. The second of these, $\varphi_2 + \varphi_{16} - \varphi_{998}$, only affected the 998th phase, which did not have a strong influence on the direction of the phasing. Table 6 summarizes a group of 33 zonal reflections that can be expressed in terms of four origin and enantiomorph definers, four $\sum_1$ phases, five symbols (a, b, c, d, $\alpha/2$) and eight symbolic relationships. Of this group of phases, only φ_{998} was later determined to be in error. A tangent-formula refinement based on these 33 zonal and 9 other unrestricted phases

Table 5. *List of triples for the structure GRAMA* ($P2_12_12_1$, $N=317$ *atoms*)

(*a*) $\sum_1$ triples. Aberrant triples (*) are indicated at the far right of the table. Incorrectly indicated triples are marked †. The calculation used 1102 E values with $t_1 \geq 1.75$, 1500 with $t_2 \leq 0.32$ and ν(random)=0.051.

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T
0	27	2	0	−27	2	0	0	−4	1	−1	−348	0.066	0.039	1.96	π
2	0	1	−2	0	1	0	0	−2	3	−3	−9	0.096	0.063	1.91	0
16	0	1	−16	0	1	0	0	−2	4	−4	−9	0.054	0.047	1.83	0
2	27	3	2	−27	3	−4	0	−6	8	−8	−2	0.032	0.043	1.69	0
3	6	2	3	−6	2	−6	0	−4	5	−5	−19	0.033	0.038	1.41	0*
0	33	1	0	−33	1	0	0	−2	12	−12	−9	0.059	0.033	1.18	0
0	25	1	0	−25	1	0	0	−2	15	−15	−9	0.057	0.055	1.14	0
8	0	1	−8	0	1	0	0	−2	36	−36	−9	0.054	0.040	0.77	0
17	0	1	−17	0	1	0	0	−2	39	−39	−9	0.024	0.042	0.76	π*
0	0	2	0	0	2	0	0	−4	9	−9	−348	0.056	0.078	0.74	0*
2	0	5	−2	0	5	0	0	−10	16	−16	−122	0.052	0.042	0.72	0
2	35	3	2	−35	3	−4	0	−6	77	−77	−2	0.067	0.031	0.69	0
0	35	2	0	−35	2	0	0	−4	11	−11	−348	0.036	0.038	0.69	π
2	14	3	2	−14	3	−4	0	−6	108	−108	−2	0.014	0.034	0.61	π*
0	8	2	0	−8	2	0	0	−4	21	−21	−348	0.033	0.049	0.57	0*
0	8	2	0	8	−2	0	−16	0	21	−21	−350	0.033	0.038	0.57	0
15	0	2	−15	0	2	0	0	−4	24	−24	−348	0.033	0.048	0.56	π
3	14	2	3	−14	2	−6	0	−4	62	−62	−19	0.020	0.033	0.55	0*
5	0	5	−5	0	5	0	0	−10	34	−34	−122	0.031	0.040	0.52	π*

(*b*) Zonal restricted $\sum_2$ triples

H			$-K$			$K-H$			$N1$	$N2$	$N3$	ν^+	ν^-	A	T
4	0	6	−2	0	−1	−2	0	−5	2	−3	−16	0.062	0.047	3.78	0
0	27	2	0	0	2	0	−27	−4	1	9	−35	0.055	0.047	3.53	0
0	27	2	0	8	2	0	−35	−4	1	21	−40	0.063	0.031	3.05	0
4	0	6	16	0	−1	−20	0	−5	2	4	−72	0.036	0.038	2.79	0
0	27	2	0	−35	−2	0	8	0	1	−11	161	0.055	0.028	2.70	0
2	0	1	16	0	−1	−18	0	0	3	4	−88	0.028	0.040	2.69	0
4	0	6	−2	0	1	−2	0	−7	2	−3	−178	0.056	0.044	2.51	0
2	0	1	−6	0	−4	4	0	3	3	−19	32	0.035	0.052	2.47	0
2	0	1	15	0	−2	−17	0	1	3	24	39	0.038	0.042	2.32	π
0	33	1	0	−25	1	0	−8	−2	12	−15	−21	0.030	0.037	2.24	0
0	27	2	0	−26	13	0	−1	−15	1	−27	−366	0.035	0.041	2.19	π
0	27	2	0	−13	−6	0	−14	4	1	−48	−124	0.029	0.031	2.18	0
0	0	2	0	35	2	0	−35	−4	9	11	−40	0.043	0.042	2.09	0
16	0	1	0	−25	−1	−16	25	0	4	−15	93	0.029	0.041	2.08	π
2	0	1	0	−33	−1	−2	33	0	3	−12	135	0.049	0.026	2.06	π
0	27	2	0	−33	−1	0	6	−1	1	−12	1045	0.048	0.038	2.03	π
0	27	2	0	−25	−1	0	−2	−1	1	−15	−1028	0.035	0.032	2.01	0
4	0	6	−6	0	−4	2	0	−2	2	−19	−110	0.036	0.050	2.00	0
2	0	1	0	−25	−1	−2	25	0	3	−15	159	0.043	0.042	1.97	π
0	35	2	0	−8	2	0	−27	−4	11	−21	−35	0.041	0.034	1.91	0
18	0	12	−15	0	2	−3	0	−14	7	−24	−52	0.030	0.037	1.91	π
2	0	1	4	0	−3	−6	0	2	3	32	84	0.035	0.045	1.90	0
4	0	6	0	0	2	−4	0	−8	2	9	−442	0.053	0.052	1.85	0*†
2	0	5	6	0	−4	−8	0	−1	16	19	−36	0.037	0.038	1.85	0*
0	27	2	0	−26	−13	0	−1	11	1	−27	−1055	0.065	0.035	1.84	0
0	27	2	0	−27	4	0	0	−6	1	−35	−620	0.051	0.027	1.83	π
3	6	0	−1	27	0	−2	−33	0	6	14	−135	0.037	0.029	1.83	π
2	0	1	−8	0	1	6	0	−2	3	−36	−84	0.054	0.040	1.83	0
2	0	1	−4	0	−3	2	0	2	3	−32	110	0.043	0.054	1.81	0*
2	0	1	15	0	2	−17	0	−3	3	24	−210	0.035	0.046	1.81	0
2	0	1	−2	0	−5	0	0	4	3	−16	348	0.067	0.052	1.78	0
2	0	1	−5	0	9	3	0	−10	3	−31	−139	0.038	0.037	1.77	0
4	0	6	−5	0	9	1	0	15	2	−31	−149	0.044	0.045	1.76	π
0	0	2	−6	0	−4	6	0	2	9	−19	84	0.038	0.047	1.73	0
0	5	4	0	8	2	0	−13	−6	17	21	−48	0.017	0.032	1.73	0*
1	27	0	5	6	0	−6	−33	0	14	22	−54	0.038	0.036	1.71	0
16	0	1	−5	0	9	−11	0	−10	4	−31	−170	0.035	0.37	1.69	π
18	0	12	−15	0	−2	−3	0	−10	7	−24	−139	0.035	0.018	1.64	0
0	27	2	0	0	−10	0	−27	8	1	−122	−379	0.046	0.039	1.63	π
2	0	1	−2	0	5	0	0	−6	3	−16	−620	0.073	0.040	1.63	0
0	0	2	2	0	5	−2	0	−7	9	16	−178	0.056	0.044	1.62	0
2	0	5	−5	0	9	3	0	−14	16	−31	−52	0.054	0.026	1.60	0
2	0	1	−2	0	9	0	0	−10	3	−56	−122	0.062	0.042	1.59	0
2	0	1	−3	0	14	1	0	−15	3	−52	−149	0.030	0.040	1.58	π
0	27	2	0	−13	−5	0	−14	3	1	−68	−805	0.028	0.032	1.58	π
0	33	1	0	−25	−1	0	−8	0	12	−15	−161	0.045	0.036	1.57	0
0	0	2	0	−8	−2	0	8	0	9	−21	161	0.038	0.042	1.55	0
0	27	2	0	−27	−6	0	0	4	1	−216	348	0.048	0.040	1.54	0
2	0	1	3	0	14	−5	0	−15	3	52	−191	0.040	0.058	1.54	0
4	0	6	2	0	5	−6	0	−11	2	16	−998	0.048	0.038	1.51	0*†

Table 5(*b*) (*cont.*)

	H			−*K*			*K*−*H*		*N*1	*N*2	*N*3	ν^+	ν^-	*A*	*T*
4	0	6	−6	0	−8	2	0	2	2	−95	110	0.028	0.040	1.50	0*
16	0	1	2	0	−5	−18	0	4	4	16	920	0.044	0.029	1.49	0
2	0	5	4	0	−3	−6	0	−2	16	32	−84	0.034	0.042	1.47	0*
0	0	2	0	13	6	0	−13	−8	9	48	−83	0.032	0.042	1.46	0
4	0	6	−15	0	−2	11	0	−4	2	−24	−894	0.023	0.039	1.45	π*
0	27	2	0	−26	−17	0	−1	15	1	−308	−366	0.051	0.032	1.45	0
2	0	1	−7	0	−2	5	0	1	3	−37	448	0.042	0.043	1.45	0*
16	0	1	2	0	−9	−18	0	8	4	56	227	0.013	0.034	1.44	0
2	0	5	4	0	3	−6	0	−8	16	32	−95	0.040	0.037	1.44	0
16	0	1	−6	0	4	−10	0	−5	4	−19	−927	0.024	0.030	1.44	0
18	0	12	−7	0	−2	−11	0	−10	7	−37	−170	0.024	0.032	1.42	0
6	0	4	−15	0	2	9	0	−6	19	−24	−136	0.012	0.040	1.42	0
2	0	5	−4	0	−3	2	0	−2	16	−32	−110	0.054	0.039	1.41	0
2	0	5	15	0	−2	−17	0	−3	16	24	−210	0.022	0.027	1.40	π
4	0	6	6	0	4	−10	0	−10	2	19	−1262	0.034	0.043	1.39	0*
0	0	2	0	27	4	0	−27	−6	9	35	−216	0.042	0.052	1.35	0*
2	0	1	2	0	−9	−4	0	8	3	56	442	0.060	0.039	1.34	0
0	0	2	17	0	1	−17	0	−3	9	39	−210	0.034	0.037	1.34	0
2	0	5	−5	0	5	3	0	−10	16	−34	−139	0.043	0.044	1.33	0
18	0	12	2	0	5	−20	0	−17	7	16	−911	0.037	0.031	1.32	0
0	27	2	0	−27	−8	0	0	6	1	−379	620	0.049	0.028	1.30	0
16	0	1	−2	0	9	−14	0	−10	4	−56	−493	0.033	0.041	1.30	0
4	0	6	7	0	−2	−11	0	−4	2	37	−894	0.022	0.047	1.29	π*
4	0	6	11	0	10	−15	0	−16	2	170	−194	0.017	0.026	1.29	0
16	0	1	−5	0	−5	−11	0	4	4	−34	894	0.027	0.035	1.29	π
16	0	1	−1	0	15	−15	0	−16	4	−149	−194	0.037	0.028	1.28	π
6	0	4	−4	0	3	−2	0	−7	19	−32	−178	0.028	0.036	1.27	0*
4	0	6	−3	0	14	−1	0	−20	2	−52	−707	0.026	0.039	1.27	π
18	0	12	−20	0	−5	2	0	−7	7	−72	−178	0.009	0.038	1.27	π*
0	0	2	−2	0	−9	2	0	7	9	−56	178	0.044	0.044	1.26	0
2	0	5	−20	0	−5	18	0	0	16	−72	88	0.025	0.034	1.25	0

Table 6. *Basis set of* 33 *zonal restricted phases obtained for GRAMA as determined from Table* 5

Five symbols (*a*, *b*, *c*, *d*, $\alpha/2$) and eight symbolic relationships ($\varphi_i = \varphi_j \pm \varphi_k$, coded as $i = j \pm k$) are employed to express this set. Reflection no. 998 is in error (*) and has a true phase value of 90°.

Serial no.	φ			Serial no.	φ	Serial no.	φ	Serial no.	φ
1	$\pi/2$	(0, 27, 2)	Origin and enantiomorph	16	$-\pi/2$	620	π	216	$-\pi/2$
3	$\pi/2$	(2, 0, 1)		35	$\pi/2$	14	$\pi/2$	998	$-\pi/2$*
6	$\pi/2$	(3, 6, 0)		21	*a*	379	$-\pi/2$	4	$\alpha/2$
12	$\pi/2$	(0, 33, 1)		40	=21+1	84	*b*	920	=4+16
				11	=21+35	36	=3−84	442	π
2	+	Σ_1 phases		161	=11−1	52	*c*	27	*d*
9	+			178	$-\pi/2$	31	=16−52	1055	=1−27
122	+			135	+	56	$\pi/2$		
348	π			1045	$=\pi$	15	=12−161		

produced a phase set (NQEST = −0.093, $R = 0.419$) and an E map from which numerous small peptide fragments could be discerned. A 60-atom helical model was fitted to a number of small fragments, the phases produced by the model were subjected to ten separate random-phasing trails and two improved solutions were obtained from which 109 atoms of the backbone of the helix could be traced. A total of 334 full and partial occupancy carbon, nitrogen and oxygen positions were eventually determined as the structural model was refined. It would have been virtually impossible to achieve this goal by traditional tangent-formula phasing methods had one not relied on the ν-STAT formulae described in this paper.

The development of the ν-STAT formulae and structural applications described here were supported by NIH grants GM32812 and GM46733.

References

De Titta, G. T., Edmonds, J. W., Langs, D. A. & Hauptman, H. (1975). *Acta Cryst.* A**31**, 472-479.

Duax, W. L., Weeks, C. M. & Hauptman, H. (1972). *Acta Cryst.* B**28**, 1857-1863.

Fisher, J., Hancock, J. & Hauptman, H. (1970). Report No. 7157. Naval Research Laboratory, Washington, DC, USA.

Fortier, S., Moore, N. J. & Frazer, M. E. (1985). *Acta Cryst.* A**41**, 571-577.

Giacovazzo, C. (1976). *Acta Cryst.* A**32**, 967-976.

Giacovazzo, C. (1977*a*). *Acta Cryst.* A**33**, 527-531.

Giacovazzo, C. (1977*b*). *Acta Cryst.* A**33**, 933-944.

Giacovazzo, C. (1983). *Acta Cryst.* A**39**, 585-592.

Gilmore, C. J. (1991). *Direct Methods of Solving Crystal Structures*, edited by H. Schenk, pp. 177-186. New York: Plenum.

Gilmore, C. J. & Hauptman, H. (1985). *Acta Cryst.*, A**41**, 457-462.

Han, F. & Langs, D. A. (1988). *Acta Cryst.* A**44**, 563-566.

Hauptman, H. (1964). *Acta Cryst.* **17**, 1421-1433.

Hauptman, H. (1975). *Acta Cryst.* A**31**, 680-687.

Hauptman, H. (1982). *Acta Cryst.* A**38**, 289-294, 632-641.

HAUPTMAN, H. & DUAX, W. L. (1972). *Acta Cryst.* A**28**, 393-395.
HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. (1969). *Acta Cryst.* B**25**, 811-814.
HAUPTMAN, H., FISHER, J. & WEEKS, C. M. (1971). *Acta Cryst.* B**27**, 1550-1561.
HOWELLS, E. R., PHILLIPS, D. C. & ROGERS, D. (1950). *Acta Cryst.* **3**, 210-214.
HULL, S. E. & IRWIN, M. J. (1978). *Acta Cryst.* A**34**, 863-870.
KARLE, J. (1970). *Acta Cryst.* B**26**, 1614-1617.
KARLE, J. (1979). *Proc. Natl Acad. Sci. USA*, **76**, 2089-2093.
KARLE, J. (1980). *Proc. Natl Acad. Sci. USA* **77**, 5-9.
KARLE, J. (1982). *Acta Cryst.* A**38**, 327-333.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635-651.
KARLE, J. & HAUPTMAN, H. (1957). *Acta Cryst.* **10**, 515-524.
LANGS, D. A. (1988). *Science*, **241**, 188-191.
LANGS, D. A. (1993). *Acta Cryst.* D**49**, 158-167.
MAIN, P. (1977). *Acta Cryst.* A**33**, 750-757.
MESSAGER, J. C. & TSOUCARIS, G. (1972). *Acta Cryst.* A**28**, 482-484.
PLETNEV, V. Z., GALITSKII, N. M., IVANOV, V. T., SMITH, G. D., WEEKS, C. M. & DUAX, W. L. (1980). *Biopolymers*, **19**, 1517-1534.
PLETNEV, V. Z., IVANOV, V. T., LANGS, D. A., STRONG, P. D. & DUAX, W. L. (1992). *Biopolymers*, **32**, 819-827.
PLETNEV, V. Z. MIKHAILOVA, I. YU., IVANOV, V. T., LANGS, D. A., GROCHULSKI, P. & DUAX, W. L. (1991). *Biopolymers*, **31**, 409-415.
SMITH, G. D. & KRISTENANSKY, J. (1991). Unpublished work.
TSOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492-499.
VAUGHAN, P. A. (1958). *Acta Cryst.* **11**, 111-115.
VITERBO, D. & WOOLFSON, M. M. (1973). *Acta Cryst.* A**29**, 205-208.
WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 65-67.
YAO, J. (1981). *Acta Cryst.* A**37**, 642-644.
YAO, J. (1983). *Acta Cryst.* A**39**, 35-37.

Acta Cryst. (1993). A**49**, 557-569

On Direct-Methods Phase Information from Differences Between Isomorphous Structure Factors

BY CHRISTOS E. KYRIAKIDIS, RENÉ PESCHAR AND HENK SCHENK

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

(*Received* 4 *June* 1992; *accepted* 2 *November* 1992)

Abstract

An efficient procedure is presented for the derivation of joint probability distributions of isomorphous data sets. The new technique is based on the use of the differences of isomorphous structure factors as random variables. It will be shown that the usual probabilistic techniques, applied to these random variables, finally result in the joint probability distribution of three single differences of isomorphous structure factors comprising three doublet and eight triplet phase sums. An advantage of the new technique is that the inherent correlation between the isomorphous data sets is removed if a probabilistic procedure is set up for the small difference itself. In this way, an enormous mathematical simplification is obtained while the final results are much better than those obtainable by previous probabilistic expressions. The final triplet distribution seems to be of sufficient quality to be used in a normal direct-methods procedure. In contrast to usual approaches, the heavy-atom substructure need not be solved first. The probabilistic expression will be explained in detail for one and three single differences. Applications for the cases of single anomalous scattering, two different wavelengths and single isomorphous replacement (excluding anomalous-scattering effects) for both real and randomly generated data show the strength of the method.

Abbreviations

c.f.	Characteristic function
j.p.d.	Joint probability distribution
c.p.d.	Conditional probability distribution
(p.)r.v.	(Primitive) random variable
s.f.	Structure factor
SD	Single difference
DM	Direct methods
DR	Diffraction ratio
SIR(N)AS	Single isomorphous replacement (neglecting) anomalous scattering
SAS	Single-wavelength anomalous scattering
2DW	Two different wavelengths

1. Introduction

The crystal structures of relatively small molecules with up to 100 independent atoms are readily determined from diffraction intensities by means of DM techniques relying on the mathematical application of a j.p.d. of complex-valued structure factors. DM estimate phases from the intensities and when these

0108-7673/93/030557-13$06.00