tion, especially given the greater possibilities proffered by modern electron microscopes which operate at higher accelerating voltages and have well designed goniometer stages for sample orientation. Although the determination of crystal structures already solved by X-ray methods may seem pointless, one should be reminded that many interesting materials cannot be crystallized sufficiently for X-ray data collection. Hence this 'benchmark' comparison with a known structure emphasizes the utility of electron crystallography for yielding reasonably accurate structural geometries.

References

BRISSE, F. (1989). J. Electron Microsc. Tech. 11, 272-279.
COREY, R. B. (1938). J. Am. Chem. Soc. 60, 1598-1604.
COWLEY, J. M. (1981). Diffraction Physics, 2nd ed. Amsterdam: North-Holland.
COWLEY, J. M. & MOODIE, A. F. (1959). Acta Cryst. 12, 360-367.
COWLEY, J. M., REES, A. L. G. & SPINK, J. A. (1951). Proc. Phys. Soc. London Sect. A, 64, 609-619.
DEGEILH, R. & MARSH, R. E. (1959). Acta Cryst. 12, 1007-1014.
DORSET, D. L. (1976). Acta Cryst. A32, 207-215.
DOYLE, P. A. & TURNER, P. S. (1968). Acta Cryst. A24, 390-397.
HAUPTMAN, H. A. (1972). Crystal Structure Determination, the Role of the Cosine Seminvariants. New York: Plenum.
HAUPTMAN, H. & GREEN, E. A. (1976). Acta Cryst. A32, 45-49.
HEIDENREICH, R. D. (1964). Fundamentals of Transmission Electron Microscopy, pp. 202-203. New York: Interscience.
JAP, B. K., DOWNING, K. H. & WALIAN, P. J. (1990). J. Struct. Biol. 103, 57-63.
KARLE, I. L., DRAGONETTE, K. S. & BRENNER, S. A. (1965). Acta Cryst. 19, 713-716.
PEREZ, S. & CHANZY, H. (1989). J. Electron Microsc. Tech. 11, 280-285.
RAMACHANDRAN, G. N. & SRINIVASAN, R. (1970). Fourier Methods in Crystallography, pp. 62-66. New York: Interscience.
RIGAMONTI, R. (1936). Gazz. Chim. Ital. 66, 174-182.
SASS, H. J., BÜLDT, G., BECKMANN, E., ZEMLIN, F., VAN HEEL, M., ZEITLER, E., ROSENBUSCH, J. P., DORSET, D. L. & MASSALSKI, A. M. (1989). J. Mol. Biol. 209, 171-175.
SPENCE, J. C. H. (1981). Experimental High-Resolution Electron Microscopy. Oxford: Clarendon Press.
TURNER, P. S. & COWLEY, J. M. (1969). Acta Cryst. A25, 475-481.
VAINSHTEIN, B. K. (1955). Zh. Fiz. Khim. 29, 327-344.
VAINSHTEIN, B. K. (1964) Structure Analysis by Electron Diffraction, pp. 124, 402. Oxford: Pergamon Press.
WILSON, A. J. C. (1942). Nature (London), 150, 151-152.

The Direct Determination of Phase Invariants Provided by Diffraction Data Measured at Two Different Temperatures

BY DAVID A. LANGS

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

In memory of David Harker

'He had bought a large map representing the sea,
    Without the least vestige of land:
And the crew were much pleased when they found it to be
    A map they could all understand.

"What's the good of Mercator's North Poles and Equators,
    Tropics, Zones, and Meridian Lines?"
So the Bellman would cry: and the crew would reply
    "They are merely conventional signs!"'

(Louis Carrol, 'The Hunting of the Snark')

Abstract

A procedure is described for the determination of the crystal structure phase invariants of a compound based on diffraction data measured at two different temperatures. This temperature difference replacement (TDR) technique is shown to provide phase-invariant information from experimentally measured X-ray diffraction data for two different test structures. Although the new method does not appear to be as powerful as single-derivative isomorphous replacement (SIR) phasing, it does appear to be capable of reliably determining a limited number of negative as well as positive phase-restricted invariants for structures containing as many as 300 non-H atoms in the asymmetric unit.

Introduction

Common solutions to the crystallographic phase problem are often based on trigonometric diffraction magnitude relationships among derivative groups of data measured for the same or Friedel-related reflecting planes. These derivative groups can include a native and various isomorphous and anomalous-scattering crystal specimens as well as structures for which a partial molecular-replacement solution may be determined. The determination of non-centrosymmetric phases requires, in principle, a minimum of a native and two derivative data sets for which the underlying derivative substructure can be determined (Bokhoven, Schoone & Bijvoet, 1951; Harker, 1956). In practice the phases of certain reflections may be

unresolved by the data that are available and may require additional derivative measurements for their ultimate determination. Certain compounds may be crystallized for which it is impossible to introduce anomalous scatterers or heavy atoms in an isomorphous manner, and for which the size of the structural complexity may exceed the radius of phase determination offered by direct-methods procedures. We propose that the diffraction patterns provided by a single crystal measured at two different temperatures may be sufficiently perturbed by the non-uniform distribution of atomic temperature factors of the structure to be useful for phase determination.

Background

The failure of direct methods to solve many structures is often exacerbated by the poor diffracting quality of the crystals which are available. Concerns that the high-angle data may not be of sufficient accuracy or that the total number of data may be insufficient relative to the number of parameters that have to be determined may not be the most serious factors contributing to the inability of these methods to produce a solution. An important limitation of a weakly diffracting data set is that the atoms of the structure may not exhibit the same proportionate scattering power as a function of the Bragg angle θ. Whereas all atoms of a particular chemical type, say carbon, may scatter equally well in the lowest-resolution shell of the data, this balance may not be maintained for the higher-resolution shells due to the varying degrees of thermal vibration of these atoms in the structure. The proportionate scattering power will increase for those atoms with the lower temperature factors and decrease for those with the higher temperature-factor values. This variation is contrary to the point-atom assumption upon which the direct-methods phase-invariant probability estimates have been derived.

Crystal structures are composed of atoms that are constrained in molecular associations which allow varying mean-square vibrational displacements. Atoms defining the terminal groups on the surface of an organic molecule may be expected to vibrate more freely than those in the interior of the structure. When crystals of such molecular compounds are cooled, say to liquid-nitrogen temperatures, the thermal factors of the atoms of the structure decrease, but by a proportional, rather than a constant, amount. Atoms on the periphery of the molecule usually exhibit the greatest reduction in thermal motion and those near the center the least. The thermal vibrations on the surface of the molecule remain larger than those near the center, but the differences between the two extremes will be reduced. It has long been recognized that a low-temperature diffraction experiment can significantly extend the limit of resolution of the data by freezing out thermal motion and disorder and

markedly reduce the amount of crystal decomposition during the data measurement, but an added advantage may be that these low-temperature data will more nearly approximate the point-atom assumption required for direct phasing methods. This paper will demonstrate the validity of this conjecture by comparing the ambient and liquid-nitrogen values of the three-phase invariants of a structure for which the room-temperature diffraction data are quite limited. Secondly, it will be shown that the differences in magnitudes between the normalized structure factors of two such data sets can in principle provide an experimentally derived estimate of the values of these phase invariants.

Structure-factor relationships

The differences in magnitude between normalized structure factors measured at two widely separated temperatures can be analyzed in the same manner as the amplitude differences noted between derivative data sets as a consequence of isomorphous substitution or anomalous scattering. Given that the relevant temperature-dependent quasi-normalized E values of a diffraction data set may be defined as

$$E_\mathbf{h} = \sum f_j \exp\left(2\pi i \mathbf{h} \cdot \mathbf{r}_j\right)$$
$$\times \exp\left(-B_j s^2\right)/\left[\sum f_j^2 \exp\left(-2B_j s^2\right)\right]^{1/2}, \quad (1)$$

the differences between corresponding E values at two different temperatures, T_1 and T_2 ($T_1 \ll T_2$), may be considered to be a result of the scattering power of a real non-negative difference structure

$$\Delta E_\mathbf{h} = E_\mathbf{h}(T_1) - E_\mathbf{h}(T_2)$$
$$\simeq \sum f_j \exp\left(2\pi i \mathbf{h} \cdot \mathbf{r}_j\right)$$
$$\times \{\exp\left[-\Delta B_j(T_1)s^2\right] - \exp\left[-\Delta B_j(T_2)s^2\right]\} \quad (2)$$

where each data set has been normalized in (1) with respect to a minimum average B value, $\langle B_j(T_i)_{\min}\rangle$, to ensure that all the differences

$$\Delta B_j(T_1) = B_j(T_1) - \langle B_j(T_1)_{\min}\rangle \quad (2a)$$

are greater than zero. As a consequence of cooling the structure it is expected that all $\Delta B_j(T_1) < \Delta B_j(T_2)$ and that the quantity

$$f_j\{\exp\left[-\Delta B_j(T_1)s^2\right] - \exp\left[-\Delta B_j(T_2)s^2\right]\} \quad (2b)$$

will be greater than or equal to zero, as required for a real non-negative difference structure. The strongest contributors to (2) are those atoms that have the greatest reduction in thermal motion as a consequence of cooling the structure, thus maximizing (2b), while those having temperature factors in the vicinity of $\langle B_j(T_i)_{\min}\rangle$ will tend to cancel in (2b) on the assumption that the position vectors $\mathbf{r}_j$ remain isomorphous at the two temperatures.

The density map of a properly scaled difference structure should have strong positive peaks at the positions of those atoms which experience the greatest decrease in B_j upon cooling, and peaks of vanishingly small positive density at the sites of atoms which have temperature factors in the vicinity of $B_j(T_i)_{min}$ at both temperatures. The map should be non-negative at all atomic sites if the difference coefficients, ΔE_h, that created it are to be considered in a probabilistic direct-methods sense. It is important to realize that this non-negative criterion could be jeopardized if one does not reliably estimate $\langle B_j(T_i)_{min}\rangle$ for each of the two data sets and the E values are not scaled properly. It remains to be demonstrated whether the mean $\langle B_j(T_i)\rangle$ normally determined in scaling procedures can be used in place of $\langle B_j(T_i)_{min}\rangle$, on the assumption that the differences between $\langle B_j(T_i)\rangle$ and $\langle B_j(T_i)_{min}\rangle$ at both temperatures may tend to be similar, and thus not affect the relative scales of $E_h(T_1)$ and $E_h(T_2)$ as they appear in (2).

The experimental magnitudes $|E_h(T_1)|$ and $|E_h(T_2)|$ provide us with a means of estimating ΔE_h, which we cannot measure directly. For centric phase-restricted magnitudes we have

$$\Delta E_h = [|E_h(T_1)| - |E_h(T_2)|] \cos \varphi_h(T_2), \qquad (3)$$

on the assumption that $\cos \varphi_h(T_1)$ will tend to equal $\cos \varphi_h(T_2)$ with phase probabilities (see Woolfson, 1956) of

$$P\{\cos [\varphi_h(T_1) - \varphi_h(T_2)]\} = \tfrac{1}{2} + \tfrac{1}{2} \tanh (X), \qquad (4)$$

$$P(\Delta E_h) = \tfrac{1}{2} + \tfrac{1}{2} \tanh [\Delta E_h X / \sigma^2(\Delta E_h)] \qquad (5)$$

where $X = |E_h(T_1)E_h(T_2)|/\varepsilon$, ε is the estimated variance of ΔE_h from (2), φ_h is the phase of E_h for a particular temperature or difference structure and $\sigma^2(\Delta E_h)$ is the estimated variance in the difference $|E_h(T_1)| - |E_h(T_2)|$ estimated from propagation of experimental error calculations.

For non-centric reflections we have a similar development for the expected value of the phase difference (Cochran, 1955; Sim, 1960)

$$\varepsilon\{\cos [\varphi_h(T_1) - \varphi_h(T_2)]\} = I_1(2X)/I_0(2X) \qquad (6)$$

from which the expectation magnitude of the difference structure amplitudes

$$\varepsilon(|\Delta E_h|) = (|E_h(T_1)|^2 + |E_h(T_2)|^2 - 2|E_h(T_1)E_h(T_2)|$$
$$\times \varepsilon\{\cos [\varphi_h(T_1) - \varphi_h(T_2)]\})^{1/2} \qquad (7)$$

may be obtained.

Phase-invariant relationships

In most applications the temperature difference replacement (TDR) substructure described by (2) may be far too complex to determine by Patterson or direct phasing methods to be of any use to attempt

to phase the entire native structure by conventional isomorphous-replacement methods. In lieu of this possibility, it was decided to investigate whether the phase-invariant relationships linking the native and TDR substructures might provide useful information to disclose phasing inconsistencies before these equations are applied.

If one considers the example for a three-phase structure invariant, $\Phi = \varphi_h + \varphi_k + \varphi_l$, on the condition that $h + k + l = 0$, it follows from (2) that since $\Delta E_h = E_h(T_1) - E_h(T_2)$

$$|\Delta E_h \Delta E_k \Delta E_l| \cos \Phi(\Delta E_h)$$

$$= |E_h(T_1)E_k(T_1)E_l(T_1)| \cos \Phi_{111}$$

$$- |E_h(T_2)E_k(T_2)E_l(T_2)| \cos \Phi_{222}$$

$$+ |E_h(T_2)E_k(T_2)E_l(T_1)| \cos \Phi_{221}$$

$$- |E_h(T_1)E_k(T_1)E_l(T_2)| \cos \Phi_{112}$$

$$+ |E_h(T_1)E_k(T_2)E_l(T_2)| \cos \Phi_{122}$$

$$- |E_h(T_2)E_k(T_1)E_l(T_1)| \cos \Phi_{211}$$

$$+ |E_h(T_2)E_k(T_1)E_l(T_2)| \cos \Phi_{212}$$

$$- |E_h(T_1)E_k(T_2)E_l(T_1)| \cos \Phi_{121} \qquad (8)$$

where the phase-invariant values of $\cos \Phi(\Delta E_h)$ and the various $\cos \Phi_{ijk}$ are *a priori* unknown. Given that probabilistic estimates of the relative signs of the ΔE_{hi} are known from (3) and that $\varphi_{hi}(T_1) \simeq \varphi_{hi}(T_2)$ for those data for which $|E_{hi}(T_1)E_{hi}(T_2)|$ is large as indicated in (4) and (6), it follows that the various $\cos \Phi_{ijk}$ on the right-hand side of (8) are all approximately equal such that the expression may be further simplified to

$$|\Delta E_h \Delta E_k \Delta E_l| \cos \Phi(\Delta E)$$

$$\simeq [|E_h(T_1)E_k(T_1)E_l(T_1)|$$

$$- |E_h(T_2)E_k(T_2)E_l(T_2)|$$

$$+ |E_h(T_2)E_k(T_2)E_l(T_1)|$$

$$- |E_h(T_1)E_k(T_1)E_l(T_2)|$$

$$+ |E_h(T_1)E_k(T_2)E_l(T_2)|$$

$$- |E_h(T_2)E_k(T_1)E_l(T_1)|$$

$$+ |E_h(T_2)E_k(T_1)E_l(T_2)|$$

$$- |E_h(T_1)E_k(T_2)E_l(T_1)|] \cos \Phi$$

$$\simeq \Delta A \cos \Phi, \qquad (9)$$

where $\cos \Phi$ represents the average weighted value of the various $\cos \Phi_{ijk}$ and the sign of the quantity represented by ΔA can indicate whether $\cos \Phi(\Delta E)$ and $\cos \Phi$ have similar or opposite signs. For non-centrosymmetric data it must be remembered that the observed differences in magnitudes between $E_h(T_1)$ and $E_h(T_2)$ will always be an underestimate for $|\Delta E_h|$

in those cases that $\varphi_h(T_1)$ and $\varphi_h(T_2)$ are not precisely equal, such that the exact value of $|\Delta E_h \Delta E_k \Delta E_l|$ is not known. For centrosymmetric structures, however, these $|\Delta E_h|$ values can be reliably determined, either in the probabilistic sense according to (3)–(5) or by experimental verification to determine whether the intensity of a particular reflection extinguishes to zero at some intermediate temperature to indicate a phase change of 180° (Boyes-Watson & Perutz, 1943). Given that these $|\Delta E_h|$ may be determined for centrosymmetric structures, it may be shown that

$$|\Delta E_h \Delta E_k \Delta E_l| \cos \Phi(\Delta E) = \Delta E_h \Delta E_k \Delta E_l \cos \Phi, \quad (10)$$

where the sign of $\Delta E_h \Delta E_k \Delta E_l$ is a direct indication of whether the two cosine values are equal or opposite in sign, irrespective of the individual magnitudes of the various $E_{hi}(T_j)$. If the magnitude of the structure product $|\Delta E_h \Delta E_k \Delta E_l|$ of the TDR substructure is large, the probability that $\cos \Phi(\Delta E)$ is positive increases, such that the sign of $\Delta E_h \Delta E_k \Delta E_l$ tends to be equal to the sign of $\cos \Phi$.

These results are similar to those obtained by Karle (1983) from the observed magnitudes of an isomorphous-replacement set of data. The interpretation of the sign of the product of the magnitude differences for TDR data, however, is less stringent than that associated with SIR data. In the SIR analysis it is assumed that the heavy-atom replacement structure is restricted to a very few sites in the cell and the structure invariants of the heavy-atom substructure tend very strongly toward a $\cos \Phi(\Delta E)$ of $+1$ if $|\Delta E_h \Delta E_k \Delta E_l|$ is only moderately large, such that the sign of $\Delta E_h \Delta E_k \Delta E_l$ will with high probability be the sign of $\cos \Phi$. The TDR substructure, however, is not associated with a discrete number of specific atomic sites, but rather reflects a sampling over all atomic sites, with varying non-negative weights as indicated in (2b). The TDR substructure may often be only marginally less complex than the native structure itself, such that $\cos \Phi(\Delta E)$ will tend to $+1$ with a lower probability, except for those instances where $|\Delta E_h \Delta E_k \Delta E_l|$ is unusually large. The sign of $\Delta E_h \Delta E_k \Delta E_l$, nonetheless, does reveal the relative signs of $\cos \Phi(\Delta E)$ and $\cos \Phi$, regardless of the magnitude of $|\Delta E_h \Delta E_k \Delta E_l|$ and the true sign of $\cos \Phi(\Delta E)$.

Computational analysis

Data were obtained from two crystal structures that had been measured at two widely separated temperatures. The first structure was the tripeptide Pyr-Phe-Pro-t-CH$_3$ (Smith, 1990): $C_{20}H_{25}N_7O_3$, $P2_1$, $a = 13\cdot397$, $b = 6\cdot171$, $c = 13\cdot331$ Å, $\beta = 118\cdot44°$, $Z = 2$ (120 K cell data). Data were recorded at 120 and 298 K with Mo $K\alpha$ radiation from the same crystal on the same instrument (Nicolet $P3$). The second structure that was examined was gramicidin A (Langs,

Table 1. $P2_1$ tripeptide at 120 and 298 K

Data are sorted into equi-proportionate shells based on $(\sin \Theta)/\lambda$, where there are NUM reflections in each shell and SMAX is the maximum $(\sin \Theta)/\lambda$ value (Å^{-1}) for the shell.

| | | 120 K | 298 K | | 120 K/298 K |
SMAX	NUM	$\langle|E|^2\rangle$	$\langle|E|^2\rangle$	$\langle|\Delta E|\rangle$	$\langle|E|\rangle/\langle|E|\rangle$
0·303	271	0·871	0·871	0·042	1·000
0·381	264	0·762	0·724	0·058	1·026
0·436	244	1·160	1·092	0·080	1·031
0·480	250	1·288	1·242	0·100	1·018
0·517	260	1·140	1·139	0·100	1·000
0·549	244	0·895	0·853	0·112	1·025
0·578	246	0·911	0·837	0·112	1·043
0·605	245	0·953	0·889	0·129	1·035
0·629	231	0·960	0·955	0·151	1·003
0·647	185	1·016	1·014	0·168	1·001

1988), a large polypeptide which crystallized from ethanol: $C_{228}H_{370}N_{40}O_{49}$, $P2_12_12_1$, $a = 31\cdot595$, $b = 32\cdot369$, $c = 24\cdot219$ Å, $Z = 4$ (120 K cell data). The gramicidin A data were recorded with Cu $K\alpha$ radiation from two separate crystals on two separate instruments (Nicolet $P3$, 120 K; Nonius CAD-4, 277 K). The 277 K data were obtained from a crystal that was sealed in a capillary with a trace of mother liquor to prevent solvent loss and decomposition, the 120 K crystal specimen was coated with a thin film of silicone grease and was flash frozen in the cold-temperature stream.

The data sets were processed to determine the absolute scale and overall anisotropic temperature parameters (Levy, Thiessen & Brown, 1970; Sheriff & Hendrickson, 1987; Blessing & Langs, 1988) and these values were used to derive the normalized structure amplitudes which were employed in this analysis. The average $\langle|E|^2\rangle$ statistic for the two data sets of each compound were compared as a function of increasing $(\sin \Theta)/\lambda$. A small corrective scale and isotropic thermal adjustment was applied to the room-temperature data sets to improve the fit with the low-temperature data. The average values of the differences in magnitude between the high- and low-temperature data sets were tabulated as a function of $(\sin \Theta)/\lambda$. The overall average $\langle|\Delta E|\rangle$ for the tripeptide and gramicidin structures were 0·10 and 0·28 respectively, and the local averages of $\langle|\Delta E|\rangle$ were observed to increase as a function of increasing $(\sin \Theta)/\lambda$, as expected. The $(\sin \Theta)/\lambda$ distributions of the adjusted structure amplitudes for the tripeptide and gramicidin structures are presented in Tables 1 and 2, respectively.

The 500 strongest $|E|$ values of the tripeptide and the 800 strongest $|E|$ values of the gramicidin data sets were employed to generate all triples which had A values greater than $1\cdot0$. The two lists included 142 phase-restricted triples for the tripeptide and 131 phase-restricted triples for the gramicidin structure, that is triples formed solely by zonal phase-restricted reflections. The values of $\Delta E_h \Delta E_k \Delta E_l$ were computed for each of these lists of triples, the triples sorted on

Table 2. *Gramicidin A data at* 120 *and* 277 K

The data are sorted as described in Table 1.

SMAX	NUM	120 K $\langle|E|^2\rangle$	277 K $\langle|E|^2\rangle$	$\langle|\Delta E|\rangle$	120 K/277 K $\langle|E|\rangle/\langle|E|\rangle$
0·188	807	1·445	1·490	0·147	0·985
0·236	765	0·959	0·956	0·141	1·001
0·271	739	1·034	1·010	0·167	1·012
0·298	707	0·705	0·674	0·169	1·023
0·321	686	0·646	0·589	0·162	1·048
0·341	648	0·714	0·661	0·184	1·040
0·359	629	0·653	0·595	0·190	1·048
0·375	586	0·662	0·636	0·203	1·021
0·390	549	0·728	0·663	0·209	1·047
0·404	554	0·704	0·605	0·238	1·079
0·417	503	0·756	0·687	0·236	1·049
0·430	472	1·029	0·951	0·274	1·040
0·441	488	0·982	0·900	0·281	1·045
0·452	433	0·964	0·955	0·294	1·005
0·463	455	1·005	0·980	0·321	1·013

the decreasing value of $\Delta E_h \Delta E_k \Delta E_l$ and the signs of $\Delta E_h \Delta E_k \Delta E_l$ and the values of cos Φ from the refined structures were compared.

It was determined that the agreement between $\Delta E_h \Delta E_k \Delta E_l$ and cos Φ could be improved if all three $|\Delta E|$ values were required to be greater than some minimum value in an attempt to resolve these differences from the random and systematic errors in both sets of data. A minimum threshold value for $|\Delta E|$ that was one third of the overall average $\langle|\Delta E|\rangle$ was found to be quite effective, that is 0·03 for the tripeptide and 0·10 for the gramicidin data. In addition, since the $\langle|\Delta E|\rangle$ between the two temperature sets is expected to approach zero in the low $(\sin\Theta)/\lambda$ limit of the data, it is best to omit significantly large ΔE contributors from the analysis if they occur below some minimal $(\sin\Theta)/\lambda$ value. For the tripeptide it was determined that a $(\sin\Theta)/\lambda$ cutoff of 0·35 Å^{-1} appeared to be effective, at which point from Table 1 it can be seen that the local $\langle|\Delta E|\rangle$ is only about 0·05 and corresponds to a 2·5% measurable difference for an E value of 2·5. The average $\langle|\Delta E|\rangle$ distribution for the gramicidin structure is approximately three times larger than that noted for the tripeptide (Table 2). This reflects both the ordering of the ethanol solvent structure as the crystal is cooled and the large systematic errors in scaling the two data sets together as a consequence of neglecting the differences in absorption due to the different crystal sizes and the added scattering of the capillary in the room-temperature experiment. It was noted that the $(\sin\Theta)/\lambda$ threshold for the gramicidin data could be lowered to 0·05 Å^{-1}, at which point this criterion appeared to have its greatest influence in discriminating against a large group of triples, involving a small number of low-angle reflections, which were observed to have cos Θ in disagreement with $\Delta E_h \Delta E_k \Delta E_l$.

The effect of applying the $|\Delta E|$ threshold and minimum $(\sin\Theta)/\lambda$ cutoff in analyzing the two lists of triples dramatically reduces the number of invariants which are to be evaluated. The tripeptide analysis produced a final list of 17 triples and the gramicidin analysis provided 24 triples as is noted in Tables 3 and 4 respectively. The tripeptide produced nine positive $\Delta E_h \Delta E_k \Delta E_l$ indications of which eight were correct and five of the six negative triples in the list were also identified by a negative $\Delta E_h \Delta E_k \Delta E_l$ value. The sign errors indicated by an * in Table 3 are generally associated with the smaller $|\Delta E_h \Delta E_k \Delta E_l|$ magnitudes, as is to be expected since cos $\Phi(\Delta E)$ of the TDR substructure is required to be $-1·0$. Similar observations may be made with regard to the results from the gramicidin analysis in Table 4. It is possible that these latter results would be improved if the data were recorded from the same crystal on the same instrument. The method appeared to work well when the data sets were initially normalized using an anisotropic thermal model and subsequently readjusted to ensure that the $\langle|E_h|^2\rangle$ averages in $(\sin\Theta)/\lambda$ shells were approximately equal as indicated in Tables 1 and 2. These tests indicate that the $\langle B_j(T_i)\rangle$ obtained by this procedure can be used in place of $\langle B_j(T_i)_{\min}\rangle$, probably for the reason that was provided earlier.

Although the new method does not appear to be as powerful as SIR phasing, it does appear to be capable of reliably determining a limited number of negative as well as positive phase-restricted invariants for structures containing as many as 300 non-H atoms in the asymmetric unit. The fact that it can reliably identify negative triple invariants may offer a powerful phasing strategy for identifying additional negative triples through quadrupole constructions (Viterbo & Woolfson, 1973) and cosine-invariant evaluation procedures (Karle & Hauptman, 1958; Hauptman, Fisher, Hancock & Norton, 1969; Hauptman, 1970).

Our calculations indicate that these algebraic formulae also tend to resolve the positive and negative cos Φ values for noncentrosymmetric three-phase invariants. An analysis of 6000 unrestricted triples of the tripeptide structure $[|\Delta E|_{\min} = 0·1, (\sin\Theta)/\lambda_{\min} = 0·25$ Å$^{-1}, |\Delta E_h \Delta E_k \Delta E_l|_{\min} = 0·02]$ produced a list of 47 positive and 28 negative $\Delta E_h \Delta E_k \Delta E_l$ cosine indications. It was noted that 46 of the 47 positive estimated triples had positive cos Φ values, but only 4 of the 28 negative estimates actually had negative cos Φ values in spite of the fact that one third of these values actually exceeded 60°. These results are quite similar to what is obtained by other cosine-invariant estimation procedures, in that (a) the positive estimates can identify a small group of exceedingly reliable invariants from which to initiate the phasing procedure, but that (b) the list of negative estimated invariants is seldom reliable enough to employ actively as 180° phase-shifted values and the best strategy may be simply to delete them from the phasing process.

Table 3. *Zonal restricted triples analysis of the tripeptide structure*

The triples are ranked on the descending value of $\Delta E_h \Delta E_k \Delta E_l$. NSER gives the position of the invariant in the original A-value-sorted list of 142 triples. The Miller indices of the three E values in each triple and the value of the cosine invariant computed from the phases of the refined low-temperature structure are indicated.

NSER	A	H			K			$-H-K$			$\cos \Phi$	$\Delta E_h \Delta E_k \Delta E_l$
14	2·45	16	0	−5	−3	0	−10	−13	0	15	1·0	$20·6 \times 10^{-3}$
116	1·11	9	0	−12	5	0	12	−14	0	0	1·0	16·8
94	1·26	9	0	−12	2	0	10	−11	0	2	1·0	12·0
22	2·14	16	0	−5	−4	0	−10	−12	0	15	1·0	10·2
95	1·24	11	0	−7	−2	0	−10	−9	0	17	1·0	7·1
133	1·03	6	0	−11	5	0	9	−11	0	2	1·0	5·8
41	1·71	8	0	5	4	0	−14	−12	0	9	1·0	5·7
70	1·39	10	0	−6	2	0	11	−12	0	−5	−1·0	3·8*
96	1·24	4	0	10	9	0	−10	−13	0	0	1·0	2·0
134	1·03	3	0	10	7	0	−9	−10	0	−1	1·0	−2·1*
49	1·62	9	0	−12	4	0	12	−13	0	0	1·0	−2·2*
132	1·03	1	0	−13	12	0	2	−13	0	11	−1·0	−2·5
78	1·31	2	0	−11	14	0	0	−16	0	11	−1·0	−3·1
99	1·23	5	0	11	7	0	−9	−12	0	−2	−1·0	−5·2
128	1·04	1	0	−13	11	0	−2	−12	0	15	−1·0	−5·9
17	2·28	9	0	−12	8	0	5	−17	0	7	1·0	−6·4*
90	1·27	3	0	13	9	0	−11	−12	0	−2	−1·0	−7·2

Table 4. *Zonal restricted triples analysis for gramicidin A*

Columns are labeled as indicated in Table 3. The tenth triple in the list is phase restricted to ±90° and thus cannot be determined by these methods.

NSER	A	H			K			$-H-K$			$\cos \Phi$	$\Delta E_h \Delta E_k \Delta E_l$
14	2·08	16	0	1	0	−25	−1	−16	25	0	1·0	$185·0 \times 10^{-3}$
65	1·41	2	0	5	15	0	−2	−17	0	−3	1·0	169·4
59	1·44	16	0	1	2	0	−9	−18	0	8	1·0	86·4
64	1·41	2	0	5	2	0	−2	−4	0	−3	1·0	67·1
124	1·06	15	0	2	−6	0	−8	−9	0	6	1·0	61·6
87	1·22	16	0	1	2	0	7	−18	0	−8	−1·0	60·5*
73	1·30	16	0	1	−2	0	9	−14	0	−10	1·0	59·4
111	1·10	8	0	1	0	0	10	−8	0	−11	1·0	57·9
60	1·44	2	0	5	4	0	3	−6	0	−8	1·0	51·0
47	1·56	4	0	6	0	−27	−6	−4	27	0	0·0	36·7
116	1·09	8	0	1	−2	0	−9	−6	0	8	−1·0	29·0*
70	1·33	2	0	5	3	0	−10	−5	0	5	1·0	26·4
4	2·79	4	0	6	16	0	−1	−20	0	−5	1·0	13·2
52	1·50	4	0	6	2	0	2	−6	0	−8	−1·0	13·1*
50	1·54	0	27	2	0	0	4	0	−27	−6	1·0	6·9
127	1·05	2	0	5	1	0	−15	−3	0	10	1·0	6·6
26	1·83	0	27	2	0	0	−6	0	−27	4	1·0	6·2
12	2·18	0	27	2	0	−13	−6	0	−14	4	1·0	−9·3*
40	1·63	0	27	2	0	0	−10	0	−27	8	1·0	−11·8*
105	1·13	0	27	2	0	−13	−14	0	−14	12	−1·0	−11·8
117	1·08	2	0	5	0	0	4	−2	0	−9	1·0	−34·3*
35	1·73	0	5	4	0	8	2	0	−13	−6	−1·0	−39·6
114	1·09	5	0	5	2	0	−7	−7	0	2	−1·0	−46·5
122	1·07	4	0	3	3	0	−5	−7	0	2	−1·0	−53·3

Probabilistic methods for estimating the general unrestricted three-phase invariant cosines for SIR data have been derived by Hauptman (1982). In this work the mixed moments of the SIR phase-invariant distributions have been defined as

$$\alpha_{nm} = \sum_j f_j^n g_j^m \qquad (11)$$

where f_j and g_j are isomorphously related atoms of the native and derivative structures. The TDR substructure is not stoichiometrically discrete, however, and the relative scattering powers of the various f_j and g_j will exhibit strong non-uniform, $\sin \Theta$ dependencies. It will thus be necessary to evaluate α_{nm} as

a function of $\sin \Theta$ and interpolate these moments as they apply to specific structure-factor amplitudes in the Hauptman formula. It remains to be demonstrated whether reliable Θ-dependent moment values may be obtained empirically from an expression such as

$$\alpha_{nm}(s) \simeq \langle |F_h^n G_h^m| \rangle_s. \qquad (12)$$

The four sets of diffraction data were recorded by Drs W. A. Pangborn and D. C. Swensen and processed utilizing computer software developed by R. H. Blessing. I thank Dr G. D. Smith for calling my attention to the tripeptide data and allowing their use in the initial phases of this study. The research

described here was supported in part by NIH grants HL32303 and GM32812.

References

BLESSING, R. H. & LANGS, D. A. (1988). *Acta Cryst.* A44, 729-735.

BOKHOVEN, C., SCHOONE, J. C. & BIJVOET, J. M. (1951). *Acta Cryst.* 4, 275-280.

BOYES-WATSON, J. & PERUTZ, M. F. (1943). *Nature (London),* 151, 714-716.

COCHRAN, W. (1955). *Acta Cryst.* 8, 473-478.

HARKER, D. (1956). *Acta Cryst.* 9, 1-9.

HAUPTMAN, H. (1970). Am. Crystallogr. Assoc. Meet., New Orleans, Louisiana, USA. Abstract no. B8.

HAUPTMAN, H. (1982). *Acta Cryst.* A38, 289-294.

HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* B25, 811-814.

KARLE, J. (1983). *Acta Cryst.* A39, 800-815.

KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* 11, 264-269.

LANGS, D. A. (1988). *Science,* 241, 188-191.

LEVY, H. A., THIESSEN, W. E. & (in part) BROWN, G. M. (1970). Am. Crystallogr. Assoc. Meet., New Orleans, Louisiana, USA. Abstract no. B6.

SHERIFF, S. & HENDRICKSON, W. A. (1987). *Acta Cryst.* A43, 118-121.

SIM, G. A. (1960). *Acta Cryst.* 13, 511-512.

SMITH, G. D. (1990). Personal communication.

VITERBO, D. & WOOLFSON, M. M. (1973). *Acta Cryst.* A29, 205-208.

WOOLFSON, M. M. (1956). *Acta Cryst.* 9, 804-810.

An Analytical Packing Function Employing Fourier Transforms

BY MILTON T. STUBBS AND ROBERT HUBER

Max-Planck-Institut für Biochemie, D-8033 Martinsried bei München, Germany

Abstract

An analytical expression for molecular overlap as a function of position is presented; it can be calculated by means of Fourier transforms. Overlap functions between pairs of symmetry elements can be combined to give a crystallographic packing function. Multiplication of this packing function with the translation function increases the significance of the latter, as shown for a trigonal test example.

Introduction

Patterson-search (Hoppe, 1957a, b; Nordman & Nakatsu, 1963; Huber, 1965) or reciprocal-space (Rossmann & Blow, 1962; Crowther & Blow, 1967; Lattman & Love, 1970; see Rossmann, 1972) molecular-replacement techniques are used increasingly for macromolecular structure solution when a partial or similar structure is known. The method is carried out in two stages: determination of the model orientation in the new crystal (rotation function), followed by translation of the rotated model with respect to the new cell axes (translation function). It is not infrequent that a reasonable solution to the rotation function can be obtained with no corresponding translation vector. Use has been made of packing functions in order to discriminate peaks in the translation function on the grounds of reasonable crystallographic packing. A short summary of available packing algorithms has been given by Fitzgerald (1990); they operate as follows.

(i) Cohen & Suh (unpublished). The shape of the protein is approximated by a number of spheres, from which intersphere distances are calculated.

(ii) Bott & Sarma (1976). A criterion is defined for bad contacts; when the number of bad contacts exceeds a user-defined number, the translation vector **t** is abandoned.

(iii) Hendrickson & Ward (1976). The molecular shape is defined by a shape function $M(\mathbf{r})$, where

$$M(\mathbf{r}) = \begin{cases} 1 & \text{if } \mathbf{r} \text{ is intramolecular} \\ 0 & \text{otherwise.} \end{cases}$$

The packing function is then calculated using the relation

$$P(\mathbf{t}) = \frac{\int M(\mathbf{r}) \cup M([R]\mathbf{r} + \mathbf{t}) \, d^3\mathbf{r}}{\int M(\mathbf{r}) \, d^3\mathbf{r}}$$

where $[R]$ denotes a crystallographic rotation matrix and **t** the translation vector.

(iv) Harada, Lifchitz, Berthou & Jolles (1981). In investigating the use of a correlation coefficient to determine the translation function,

$$\mathscr{C}(\mathbf{t}) = \frac{\sum_{h}^{S} F_o^2(\mathbf{h}) \|F_c(\mathbf{h}, \mathbf{t})\|^2}{\left[\sum_{h}^{S} F_o^4(\mathbf{h}) \sum_{h}^{S} \|F_c(\mathbf{h}, \mathbf{t})\|^4 \right]^{1/2}},$$

approximations were made to allow utilization of FFT methods, resulting in a function

$$\mathscr{C}(\mathbf{t}) \simeq TO(\mathbf{t})/O(\mathbf{t})$$