

Unique or Essentially Unique Results from One-Wavelength Anomalous Dispersion Data

BY J. KARLE

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC, USA 20375-5000

(Received 1 November 1984; accepted 26 February 1985)

Abstract

An experiment that is sensitive to anomalous dispersion effects will produce at one wavelength independent intensity information at a reciprocal-lattice point and its negative. These pairs of intensities are known as Bijvoet pairs. The usual analysis of the implications of Bijvoet pairs leads to the conclusion that they generate a twofold ambiguity in the evaluation of certain phase differences. In this paper, it is shown that additional information contained in the Bijvoet pairs, and not normally used in the analysis leading to the implication of twofold ambiguity, can be used to obtain unique or essentially unique values for the phase differences of interest with potentially useful accuracy. The accuracy, of course, depends upon the accuracy of the data, but a test example has shown considerable insensitivity to such errors. The analysis presented here is based on an exact algebraic analysis of the intensity equations associated with the anomalous dispersion technique. Although the theory is quite general, applying exactly to any number or type of anomalously scattering atoms at any number of wavelengths, the application here concerns the case of one type or one predominant type of anomalously scattering atoms in a one-wavelength experiment. It is noted that in the two equations associated with the Bijvoet pairs there are three unknown quantities. It is shown, however, that the two intensity data provide enough information to evaluate the three unknown quantities to good approximation in an essentially unique fashion, which, in addition, can be effected in a least-squares calculation. The phase information of interest that is obtained concerns the values of phase differences, $\varphi_{1,b}^n - \varphi_{2,b}^n$, between phases associated with the structure of nonanomalously scattering atoms and those associated with the structure of the anomalously scattering atoms, respectively, with all atoms scattering as if there were no anomalous dispersion.

Introduction

An exact algebraic analysis of multiple-wavelength anomalous dispersion data resulted in an essentially linear system of simultaneous equations (Karle, 1980). The unknown quantities are intensities and phase differences that would be obtained if there were

no anomalous dispersion. The equations are exact for any number or types of anomalously scattering atoms. Variation of the values of the observed intensities is expressed in the simultaneous equations by means of factors that modify the unknown quantities and can be evaluated from known values for the real and imaginary corrections to the atomic scattering factors. Isomorphous replacement information is readily incorporated into the system of equations.

An evident application of the simultaneous equations would be to measure intensity data at several wavelengths so that the number of independent data would at least equal the number of unknown quantities. Analyses should benefit in increased accuracy from an excess of independent data over unknown quantities. In this paper, a particularly simple but potentially useful case is considered, namely, a one-wavelength experiment applied to the case in which the substance of interest contains one type or one predominant type of anomalously scattering atom. An algebraic analysis of the case of one type of anomalous scatterer has recently been presented by Woolfson (1984) in terms of two sets of data, one with anomalous scattering and one without. The discussion here is restricted to only one set of data with anomalous effects.

In a one-wavelength experiment involving anomalous dispersion, values are obtained for independent intensities associated with a reciprocal vector and its negative, Bijvoet pairs. Each Bijvoet pair gives rise to two independent equations. Each of the equations contains as unknown quantities two intensities concerned with the nonanomalous scattering, and a phase difference, *i.e.* three unknown quantities. It would appear that two independent equations containing three unknown quantities would be of little use. However, the two unknown intensities bear an approximate relationship to the two measured intensities, which can be made use of. As a consequence, it will be seen that a system of equations can be generated that can be solved by a least-squares technique to yield values for the unknown intensities and the unknown phase differences in a unique or essentially unique fashion. The term 'essentially unique' implies the existence of a procedure that distinguishes between two alternative values, even when they are fairly close. In the latter case, a variety

of selections would approximate the correct answer well. Information concerning the structure of the anomalous scatterers is not required. The potential of the mathematical system to be described here may be compared to the usual analysis of a one-wavelength anomalous dispersion experiment in noncentric systems, which would, in general, imply the existence of a twofold ambiguity, even when the structure of the anomalous scatterers is known.

In the past, unique phase information has been extracted from one-wavelength anomalous dispersion experiments. A useful way for resolving the ambiguity that was normally obtained was suggested by Peerdeman & Bijvoet (1956) and by Ramachandran & Raman (1956). It was to choose, between the two alternatives found in their analyses, that phase that was closest to the phase that could be computed from known positions for the anomalously scattering atoms. This gave a correct choice in a majority of instances. This method was used by Dale, Hodgkin & Venkatesan (1963) in the study of the structure of an aquo cyanide of the natural vitamin B₁₂ nucleus containing cobalt. More recently, use was made in a probabilistic fashion by Hendrickson & Teeter (1981) of the known positions of anomalously scattering sulfur atoms in the investigation of the structure of the macromolecule, crambin, to facilitate the resolution of phase ambiguities. In all these applications, the known structure of the anomalously scattering atoms was used. The analysis presented here for one-wavelength anomalous dispersion data is distinguished from the earlier ones in that it gives essentially unique phase information without the use of structural information concerning the anomalously scattering atoms. In fact, only information concerning the chemical nature of these atoms is required.

Although this paper does not concern triplet phase invariants it is noteworthy that triplet phase invariants can be evaluated unambiguously from one-wavelength anomalous dispersion data by probabilistic means (Hauptman, 1982; Giacovazzo, 1983) and by the application of certain rules (Karle, 1984b).

Theory

We consider the case of a structure composed of atoms that scatter nonanomalously and atoms, all of the same type, that scatter anomalously. The appropriate equations, which were obtained from the aforementioned analysis (Karle, 1980), may be written

$$\begin{aligned} |F_{\lambda\mathbf{h}}|^2 = & |F_{1,\mathbf{h}}^n|^2 + \alpha_{\mathbf{h}}|F_{2,\mathbf{h}}^n|^2 \\ & + \beta_{\mathbf{h}}|F_{1,\mathbf{h}}^n||F_{2,\mathbf{h}}^n| \cos(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n) \\ & + \gamma_{\mathbf{h}}|F_{1,\mathbf{h}}^n||F_{2,\mathbf{h}}^n| \sin(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n) \end{aligned} \quad (1)$$

and

$$\begin{aligned} |F_{\lambda\mathbf{h}}|^2 = & |F_{1,\mathbf{h}}^n|^2 + \alpha_{\mathbf{h}}|F_{2,\mathbf{h}}^n|^2 \\ & + \beta_{\mathbf{h}}|F_{1,\mathbf{h}}^n||F_{2,\mathbf{h}}^n| \cos(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n) \\ & - \gamma_{\mathbf{h}}|F_{1,\mathbf{h}}^n||F_{2,\mathbf{h}}^n| \sin(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n), \end{aligned} \quad (2)$$

where

$$\alpha_{\mathbf{h}} = 1 + (f_{\lambda 2,\mathbf{h}}^a/f_{2,\mathbf{h}}^n)[(f_{\lambda 2,\mathbf{h}}^a/f_{2,\mathbf{h}}^n) + 2 \cos \delta_{\lambda 2,\mathbf{h}}] \quad (3)$$

$$\beta_{\mathbf{h}} = 2[1 + (f_{\lambda 2,\mathbf{h}}^a/f_{2,\mathbf{h}}^n) \cos \delta_{\lambda 2,\mathbf{h}}] \quad (4)$$

$$\gamma_{\mathbf{h}} = 2(f_{\lambda 2,\mathbf{h}}^a/f_{2,\mathbf{h}}^n) \sin \delta_{\lambda 2,\mathbf{h}} \quad (5)$$

$|F_{\lambda\mathbf{h}}|$ is a known structure-factor magnitude whose value is obtained from a measurement of the intensity at a particular wavelength, λ , for a given reciprocal vector, \mathbf{h} , $|F_{1,\mathbf{h}}^n|$ is the magnitude of the corresponding structure factor for the nonanomalously scattering atoms, $|F_{2,\mathbf{h}}^n|$ is the magnitude of the corresponding structure factor for the anomalously scattering atoms as if there were no anomalous scattering and $\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n$ is the difference between the phases associated with $|F_{1,\mathbf{h}}^n|$ and $|F_{2,\mathbf{h}}^n|$, respectively. Evidently, the subscript 1 refers to the non-anomalously scattering atoms and the subscript 2 refers to the anomalously scattering ones. Further definitions of the quantities in (3)–(5) follow from the definition of the atomic scattering factor for an atom of type q ,

$$f_{\lambda q,\mathbf{h}} = f_{q,\mathbf{h}}^n + f'_{\lambda q,\mathbf{h}} + if''_{\lambda q,\mathbf{h}}, \quad (6)$$

where $f_{q,\mathbf{h}}^n$ is the normal atomic scattering factors and $f'_{\lambda q,\mathbf{h}}$ and $f''_{\lambda q,\mathbf{h}}$ are the real and imaginary corrections. Equations (3)–(5) contain the following for $q = 2$:

$$f'_{\lambda q,\mathbf{h}} = (f_{\lambda q,\mathbf{h}}^a{}^2 + f_{\lambda q,\mathbf{h}}^n{}^2)^{1/2} \quad (7)$$

$$\delta_{\lambda q,\mathbf{h}} = \tan^{-1}(f''_{\lambda q,\mathbf{h}}/f'_{\lambda q,\mathbf{h}}). \quad (8)$$

The quantities f' and f'' are normally treated as independent of \mathbf{h} and are tabulated so that appropriate values for $\alpha_{\mathbf{h}}$, $\beta_{\mathbf{h}}$ and $\gamma_{\mathbf{h}}$ may be readily computed.

The two equations representing the Bijvoet pairs are (1) and (2). It is seen that these equations involve three unknown quantities, $|F_{1,\mathbf{h}}^n|$, $|F_{2,\mathbf{h}}^n|$ and $\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n$. This apparent imbalance between the number of independent data and number of unknown quantities can be overcome to yield unique or essentially unique values for $\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n$ without any knowledge of the heavy-atom structure. This is accomplished by recognizing that $|F_{1,\mathbf{h}}^n|$ and $|F_{2,\mathbf{h}}^n|$ are approximately definable in terms of the measured Bijvoet pairs and that the values obtained for $\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n$ from (1) and (2) are relatively insensitive to errors in the values for $|F_{1,\mathbf{h}}^n|$ and $|F_{2,\mathbf{h}}^n|$.

In this section three systems of defining equations are presented that can be treated in a least-squares fashion. System I is a general system, but can be replaced for convenience with system II or III, if desired. In fact, system II has been used with the

test calculations that are presented. The defining equations are treated in the usual least-squares fashion, namely, the sum of the squares of the errors in the defining equations are minimized.

Unlike most least-squares calculations, test calculations have shown that it is possible to proceed with systems I-III even though the number of independent equations does not exceed the number of unknown quantities. This is no doubt facilitated by the fact that in the systems all the defining equations are linear except for one that is quadratic, giving a broad range of convergence. Suitable starting values for the unknown quantities are thus fairly readily obtained.

In the next part, approximate statistical formulas are given for the evaluation of $|F_{1,h}^n|$ and $|F_{2,h}^n|$. The philosophy of their use is to recognize their approximate nature and consider a range of starting values for $|F_{1,h}^n|$ and $|F_{2,h}^n|$ based on the initial evaluations. The effect of the variation of starting values on the resulting values of the desired phase differences from the least-squares calculation is then observed for consistency. In some instances in the test calculations, it was found that no variation on the initial evaluation of $|F_{2,h}^n|$ was required for the succeeding least-squares analysis. Each time new starting values are employed, a new least-squares calculation is performed. Results obtained in good agreement with each other, independently of the broad range of starting values, are a measure of the reliability of the calculation. The ultimate test is, of course, the agreement of the results with the correct values.

A geometric analysis of the interrelationships among the mathematical quantities that arise is given in the Appendix.

Estimates for $|F_{1,h}^n|$ and $|F_{2,h}^n|$

A statistical argument would suggest that $|F_{1,h}^n|$ could be estimated from

$$|F_{1,h}^n| \sim 0.5 W_{1,h} (|F_{\lambda h}| + |F_{\lambda \bar{h}}|), \quad (9)$$

where

$$W_{1,h} = \left\{ \frac{\sum_{j=1}^{N_{\text{non}}} f_{jh}^2}{\sum_{j=1}^{N_{\text{non}}} f_{jh}^2 + \sum_{j=1}^{N_{\text{ano}}} [(f_{jh}^n + f_j)^2 + f_j^2]} \right\}^{1/2} \quad (10)$$

and N_{non} and N_{ano} are the number of nonanomalously and anomalously scattering atoms, respectively, in the unit cell. Equations (9) and (10) may be compared with similar ones (Karle, 1984a) whose purpose it is to calculate values for the $|F_h^n|$, the structure-factor magnitudes for the total structure when all atoms scatter nonanomalously.

For $|F_{2,h}^n|$, it can be shown that

$$|F_{2,h}^n|^2 = S \{ \|F_{\lambda h}\| - |F_{\lambda \bar{h}}| / [2(f_{\lambda 2}^n / f_{2,h}^n)] \}^2, \quad (11)$$

where $f_{\lambda 2}^n$ is treated as independent of scattering angle and S is a scale factor that is equal to 1 when the angles $\varphi_{\lambda h}$ and $-\varphi_{\lambda \bar{h}}$ are equal. The latter circumstance gives a minimum value for $|F_{2,h}^n|^2$. An estimate of values for S can be based on test examples having the same atomic composition as the substance of interest.

There are two probability theories giving identical results, except for notation (Hauptman, 1982; Giacovazzo, 1983), which afford an alternative way to estimate $|F_{2,h}^n|^2$. A test calculation with exact data for cytochrome c550.PtCl₄²⁻ showed these theories to underestimate the values of $|F_{2,h}^n|^2$ by about 1-30% for the first 2000 differences, $\|F_{\lambda h}\| - |F_{\lambda \bar{h}}|$, listed with the largest first. The underestimation increased roughly as the differences decreased. This may occur because heavy-atom structures in macromolecules are quite simple and therefore are not comprised of a sufficient number of atoms to have the statistical properties predicted by standard probabilistic analyses. Some special studies of this matter have appeared recently (Shmueli, Weiss, Kiefer & Wilson, 1984; Karle, 1984c).

System I

Equations (1) and (2) can be considered as linear in four variables defined as

$$x_1 = |F_{1,h}^n|^2 \quad (12)$$

$$x_2 = |F_{2,h}^n|^2 \quad (13)$$

$$x_3 = |F_{1,h}^n| |F_{2,h}^n| \cos(\varphi_{1,h}^n - \varphi_{2,h}^n) \quad (14)$$

$$x_4 = |F_{1,h}^n| |F_{2,h}^n| \sin(\varphi_{1,h}^n - \varphi_{2,h}^n). \quad (15)$$

Rewriting (1) and (2) gives

$$|F_{\lambda h}|^2 = x_1 + \alpha_h x_2 + \beta_h x_3 + \gamma_h x_4 \quad (16)$$

and

$$|F_{\lambda \bar{h}}|^2 = x_1 + \alpha_h x_2 + \beta_h x_3 - \gamma_h x_4. \quad (17)$$

In addition, there exists a quadratic relationship among the variables

$$x_1 x_2 = x_3^2 + x_4^2. \quad (18)$$

If (11), which defines x_2 , is added to (16)-(18), there are four equations defining four unknown quantities. Three of the equations are linear in the variables and one is quadratic. This system is amenable to least-squares solution with one important caution. The predominance of (18) in the system must be tempered. This can be accomplished in many ways. Two ways that are used are to either divide (18) by x_2 or take the square root of both sides of (18). In the case of the former, the possibility exists that in some intermediate step in the least-squares process x_2 could become quite small and cause instability and, in the case of the latter, an intermediate excursion of $x_1 x_2$ into negative values would inhibit the taking of the

square root. Suitable computer programming can anticipate and overcome such eventualities.

Initial values for the unknown quantities can be obtained for x_1 from (9), for x_2 from (11), for x_3 from

$$x_3 = (|F_{\lambda b}|^2 + |F_{\lambda \bar{b}}|^2 - 2\alpha_b x_2 - 2x_1) / 2\beta_b \quad (19)$$

once initial values are obtained for x_1 and x_2 and for x_4 from

$$x_4 = (|F_{\lambda b}|^2 - |F_{\lambda \bar{b}}|^2) / 2\gamma_b, \quad (20)$$

where (19) and (20) are obtained by adding and subtracting (16) and (17), respectively.

In the case that the only information available is values for $|F_{\lambda b}|^2$ and $|F_{\lambda \bar{b}}|^2$, a suitable system of equations for a least-squares procedure involves the use of (11) and (16)–(18). Equations (19) and (20) can replace (16) and (17) in the least-squares calculation. In test examples, it has been found that convergence is enhanced by combining all the equations, giving for the defining equations (11) and (16)–(20). These six equations are called system I. In practice, it has been found useful to consider several initial values for x_1 and carry through the least-squares calculation several times. In that case, it is possible to dispense with the calculation of $W_{1,b}$ in (9) and simply base the initial values on

$$|F_{1,b}^n| \sim 0.5\epsilon (|F_{\lambda b}| + |F_{\lambda \bar{b}}|), \quad (21)$$

where ϵ can assume, for example, the values 0.7, 1.0 and 1.3. It is most likely, for the simple substitution of heavy atoms in a macromolecule, that the range of values for $|F_{1,b}^n|$, made accessible in the least-squares calculation by the latter initial values, would include the correct value for $|F_{1,b}^n|$.

System II

The least-squares procedure can be varied in several ways. For example, an alternative procedure could be based on the use of (11) to provide a value for $|F_{2,b}^n|$ that is held constant throughout the calculation. Treating $|F_{2,b}^n|$ as a known constant can give rise to a somewhat different set of equations that replace (16)–(20). We define new variables

$$x_5 = |F_{1,b}^n| \cos(\varphi_{1,b}^n - \varphi_{2,b}^n) \quad (22)$$

$$x_6 = |F_{1,b}^n| \sin(\varphi_{1,b}^n - \varphi_{2,b}^n) \quad (23)$$

and have the following new set of defining equations for the least-squares system:

$$|F_{\lambda b}|^2 - x_1 - \alpha_b |F_{2,b}^n|^2 - \beta_b |F_{2,b}^n| x_5 - \gamma_b |F_{2,b}^n| x_6 = 0 \quad (24)$$

$$|F_{\lambda \bar{b}}|^2 - x_1 - \alpha_b |F_{2,b}^n|^2 - \beta_b |F_{2,b}^n| x_5 + \gamma_b |F_{2,b}^n| x_6 = 0 \quad (25)$$

$$x_1 - x_5^2 - x_6^2 = 0 \quad (26)$$

$$|F_{\lambda b}|^2 + |F_{\lambda \bar{b}}|^2 - 2x_1 - 2\alpha_b |F_{2,b}^n|^2 - 2\beta_b |F_{2,b}^n| x_5 = 0 \quad (27)$$

$$[(|F_{\lambda b}|^2 - |F_{\lambda \bar{b}}|^2) / (2\gamma_b |F_{2,b}^n|)] - x_6 = 0, \quad (28)$$

where the value for $|F_{2,b}^n|$ is obtained from (11). Initial

values for x_1 can be obtained from (21), x_5 from (27) and x_6 from (28). The five equations (24)–(28) are called least-squares system II.

The case of isomorphous replacement (system III)

In the case that isomorphous replacement data are also available, measured values of the intensities for the native macromolecule provide values for the $|F_{1,b}^n|$. It is then possible to solve for $|F_{2,b}^n|$ and $\varphi_{1,b}^n - \varphi_{2,b}^n$. In this case, approximate information for $|F_{2,b}^n|$ from (11) could distort the results and this equation is therefore either omitted from the least-squares system I (11, 16–20) or included with low weight. Values obtained for the $|F_{2,b}^n|$ can be used to determine the heavy-atom structure and therefore, ultimately, values for the $\varphi_{2,b}^n$. Having values for the $\varphi_{1,b}^n - \varphi_{2,b}^n$ and, separately, values for $\varphi_{2,b}^n$ leads to values for the desired $\varphi_{1,b}^n$.

An alternative set of equations to system I for the least-squares calculation when $|F_{1,b}^n|$ is known can be considered. We define new variables

$$x_7 = |F_{2,b}^n| \cos(\varphi_{1,b}^n - \varphi_{2,b}^n) \quad (29)$$

$$x_8 = |F_{2,b}^n| \sin(\varphi_{1,b}^n - \varphi_{2,b}^n) \quad (30)$$

and have the following new set of defining equations for the least-squares system:

$$|F_{\lambda b}|^2 - |F_{1,b}^n|^2 - \alpha_b x_2 - \beta_b |F_{1,b}^n| x_7 - \gamma_b |F_{1,b}^n| x_8 = 0 \quad (31)$$

$$|F_{\lambda \bar{b}}|^2 - |F_{1,b}^n|^2 - \alpha_b x_2 - \beta_b |F_{1,b}^n| x_7 + \gamma_b |F_{1,b}^n| x_8 = 0 \quad (32)$$

$$x_2 - x_7^2 - x_8^2 = 0 \quad (33)$$

$$|F_{\lambda b}|^2 + |F_{\lambda \bar{b}}|^2 - 2|F_{1,b}^n|^2 - 2\alpha_b x_2 - 2\beta_b |F_{1,b}^n| x_7 = 0 \quad (34)$$

$$[(|F_{\lambda b}|^2 - |F_{\lambda \bar{b}}|^2) / (2\gamma_b |F_{1,b}^n|)] - x_8 = 0, \quad (35)$$

where the value of $|F_{1,b}^n|$ is obtained from experiment. Initial values for x_2 can be obtained from (11), x_7 from (34) and x_8 from (35). The five equations (31)–(35) are called least-squares system III. Equation (11) may be added to system III with a low weight. The reason for including it with a low weight is that it is possible, because of experimental error, to have inconsistent values among the three quantities $|F_{\lambda b}|$, $|F_{\lambda \bar{b}}|$ and $|F_{1,b}^n|$. If these values are not greatly inconsistent, the least squares may be brought to reasonable convergence by including (11) with a low weight. The weight should be a compromise between dominance of the result by (11), as obtained from too large a weight, and insufficient influence to affect the convergence, as obtained from too small a weight. In one application to a problem in macromolecular structure determination, a weight of 0.01 for (11) was found to be suitable.

Test calculations

Test calculations were performed on exact data and also on data into which errors were introduced. The

Table 1. Calculation of least-squares system II for samples of data from cytochrome c550.PtCl₄²⁻ in which only the Pt atoms scatter anomalously

The samples are based on \mathbf{h} associated with an ordered sequence of $\|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$ in which the largest one is first. The effect of errors in the data was tested by multiplying $\|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$ by the error factor listed in the second column and effecting the error, in this case, by readjusting the value of $|F_{\lambda\bar{\mathbf{h}}}|$. The actual distribution of the errors between $|F_{\lambda\mathbf{h}}|$ and $|F_{\lambda\bar{\mathbf{h}}}|$ does not materially affect the results. The average magnitudes of error for $|F_{1,\mathbf{h}}^n|$, $|F_{2,\mathbf{h}}^n|$ and $(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n)$ are seen in columns 3, 4 and 5, respectively. The total number of independent data is 3250 at 2.5 Å resolution and the radiation is Cu K α .

Sample	Error factor	Average % error $ F_{1,\mathbf{h}}^n $	Average % error $ F_{2,\mathbf{h}}^n $	Average error (rad) $\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n$
1-100	1.50	21	49	0.21
901-1000	1.50	21	50	0.35
1601-1700	1.50	27	51	0.48
1-100	1.25	14	25	0.16
901-1000	1.25	20	30	0.34
1601-1700	1.25	26	37	0.47
1-100	1.00	13	7	0.17
901-1000	1.00	19	17	0.37
1601-1700	1.00	27	28	0.49
1-100	0.75	19	25	0.29
901-1000	0.75	20	27	0.47
1601-1700	0.75	30	30	0.56
1-100	0.50	24	50	0.48
901-1000	0.50	22	51	0.59
1601-1700	0.50	33	53	0.65

data were computed at 2.5 Å resolution for Cu K α radiation from the coordinates for cytochrome c550.PtCl₄²⁻ from *Paracoccus denitrificans* (Timkovich & Dickerson, 1976). The structure factors were computed in two ways. One calculation introduced anomalous effects from the Pt atom alone and the second included anomalous effects from the Pt, Fe, S and Cl atoms. The first calculation models the case when there would be only one type of anomalous scatterer. This calculation not only represents an important experimental case, but also provides a basis of comparison for determining the effect on the errors of including all four types of anomalous scatterers in the data while treating the data as if the Pt atoms were the one predominant type of anomalous scatterer.

Calculations based on system II are presented in Table 1. They concern only the data that contain anomalous effects from the Pt atoms alone. In system II, $|F_{2,\mathbf{h}}^n|$ is estimated from a statistical analysis and the least-squares system is solved for the values of $|F_{1,\mathbf{h}}^n|$ and $(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n)$. The value of $|F_{2,\mathbf{h}}^n|$ was determined from (11) where the value of S was obtained from a statistical calculation based on an arbitrary structure with the same chemical composition as cytochrome c550.PtCl₄²⁻. The factor S assumes that value for which the average value of $|F_{2,\mathbf{h}}^n|$ in some sample of reflections, as computed from (11), is equal to the

average value of $|F_{2,\mathbf{h}}^n|$ for the same type of sample, as obtained from the arbitrary structure.

The samples of reflections for which the statistical estimates of S were made coincide with those listed in the first column of Table 1. The numerical sequence refers to the sets of reflections for which the calculations of system II were made and are based on the sequence of values of $\|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$, with the largest first. If a value for $|F_{\lambda\mathbf{h}}|$ is very small, it may be detrimental to the calculations. In this case, no lower limit was used, but it is worth considering in future calculations. The starting values for $|F_{1,\mathbf{h}}^n|$ were the three values described for (21). The three calculations converged either to the same three values for $|F_{1,\mathbf{h}}^n|$ and $(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n)$ or to two the same and one different. The values accepted were either three the same or two the same. A rejection criterion could be based on the magnitude of the discrepancy when two results differ from the third. This was not done here.

Errors were introduced by taking the correct values for $|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$, multiplying them by the factors listed in column 2 of Table 1 and readjusting the values of the $|F_{\lambda\bar{\mathbf{h}}}|$ to be consistent with the resulting smaller or larger differences. An even distribution of the errors among both the $|F_{\lambda\mathbf{h}}|$ and $|F_{\lambda\bar{\mathbf{h}}}|$ did not alter the results significantly. The introduction of rather large errors into the differences, $|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$, had a remarkably small effect on the average errors for $|F_{1,\mathbf{h}}^n|$ and $(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n)$.

The other type of error that could be considered is an error in the scaling of $|F_{\lambda\mathbf{h}}|$ and $|F_{\lambda\bar{\mathbf{h}}}|$. Because of the nature of the defining equations, this type of error affects only the accuracy of values of $|F_{1,\mathbf{h}}^n|$ and $|F_{2,\mathbf{h}}^n|$ but not those for the desired phase differences $(\varphi_{1,\mathbf{h}}^n - \varphi_{2,\mathbf{h}}^n)$. Because scaling error does not affect the values of the phase differences, it was not investigated further.

Test calculations based on data from cytochrome c550.PtCl₄²⁻ in which Pt, Fe, Cl and S atoms were the anomalous scatterers are shown in Table 2. In the analysis of the data by use of system II, the Pt atoms were treated as the sole predominant type of anomalous scatterer. The nature of the computations was the same as those that contributed to Table 1 with the exception that it was often necessary to modify the value of $|F_{2,\mathbf{h}}^n|$, as obtained from (11), in order to bring the system to convergence. This was done by modifying the starting value of $|F_{2,\mathbf{h}}^n|$ by multiplying by the factors, 1.9, 1.6, 1.3, 1.0, 0.7, in succession until convergence was obtained. The results shown in Table 2 are seen to be comparable to those in Table 1 with only a modest increase in the average errors.

It is seen from Table 1 that only the largest differences, $\|F_{\lambda\mathbf{h}}\| - |F_{\lambda\bar{\mathbf{h}}}|$, give reliable values for the $|F_{2,\mathbf{h}}^n|$ and then only if the errors in these differences are small. Table 2 shows that when the four types of anomalous scatterers are considered, none of the

Table 2. Calculation of least-squares system II for samples of data from cytochrome c550.PtCl₄²⁻ in which Pt, Fe, Cl and S atoms scatter anomalously

The samples are based on h associated with an ordered sequence of $\|F_{\lambda h}\| - |F_{\lambda \bar{h}}|$ in which the largest one is first. The effect of errors in the data was tested by multiplying $\|F_{\lambda h}\| - |F_{\lambda \bar{h}}|$ by the error factor listed in the second column and effecting the error, in this case, by readjusting the value $|F_{\lambda h}|$. The actual distribution of errors between $|F_{\lambda h}|$ and $|F_{\lambda \bar{h}}|$ does not materially affect the results. The average magnitudes of error for $|F_{1,h}^n|$, $|F_{2,h}^n|$ and $(\varphi_{1,h}^n - \varphi_{2,h}^n)$ are seen in columns 3, 4 and 5, respectively. The total number of independent data is 3250 at 2.5 Å resolution and the radiation is Cu K α .

Sample	Error factor	Average % error $ F_{1,h}^n $	Average % error $ F_{2,h}^n $	Average error (rad) $\varphi_{1,h}^n - \varphi_{2,h}^n$
1-100	1.50	30	85	0.29 _p
901-1000	1.50	31	79	0.40
1601-1700	1.50	27	75	0.67
1-100	1.25	24	56	0.25
901-1000	1.25	30	56	0.38
1601-1700	1.25	27	55	0.69
1-100	1.00	22	31	0.28
901-1000	1.00	30	38	0.41
1601-1700	1.00	27	43	0.72
1-100	0.75	27	23	0.41
901-1000	0.75	33	30	0.47
1601-1700	0.75	28	42	0.75
1-100	0.50	31	39	0.67
901-1000	0.50	35	45	0.59
1601-1700	0.50	29	52	0.80

estimates of $|F_{2,h}^n|$ is reliably obtained. A two-wavelength experiment, which would contribute additional defining equations, or accurate information concerning the values of $|F_{1,h}^n|$, as may be obtained from an isomorphous replacement experiment, would afford an opportunity to obtain more accurate values for the $|F_{2,h}^n|$ and the other unknown quantities as well. Reliable data for the $|F_{2,h}^n|$ would be valuable for the determination of the heavy-atom structure. Knowledge of the latter gives values for the $\varphi_{2,h}^n$. Once these are known, the desired values for the $\varphi_{1,h}^n$ may be obtained from the known values of the differences, $\varphi_{1,h}^n - \varphi_{2,h}^n$. The values for the $\varphi_{1,h}^n$ then permit the immediate calculation of a Fourier map of the electron distribution.

Summary remarks

Least-squares techniques for treating data obtained from one-wavelength anomalous dispersion experiments have been presented, which led, in test calculations, to unique evaluations of the phase differences, $\varphi_{1,h}^n - \varphi_{2,h}^n$, within acceptable ranges of accuracy, even when the errors in the differences $\|F_{\lambda h}\| - |F_{\lambda \bar{h}}|$ were rather large. The uniqueness of the results derives from inherent information in the measured intensities concerning the normal structure factors for the non-anomalous and anomalously scattering atoms. The calculations involved are very simple and can be carried out at a relatively high rate. It has been pointed

out elsewhere (Karle, 1984c) that the phase differences ($\varphi_{1,h}^n - \varphi_{2,h}^n$) can be formed into triplet phase invariants for the nonanomalously scattering atoms for use in phase determination if difficulties should arise in the determination of the heavy-atom structure.

In order to make the calculations described in this paper, the only information required concerning the anomalously scattering atoms is the chemical nature of the predominant type of scatterer. The calculations should have fairly broad application in practice. The presence of more than one predominant anomalous scatterer of widely differing atomic number may require the use of multiple-wavelength experiments and the associated theory (Karle, 1980).

Use of multiple-wavelength data has the potential of enhancing the accuracy of the analyses in any case. The point of the presentation here, however, has been to show how much information is derivable from a one-wavelength experiment and the potential utility of such an experiment.

I wish to thank Mr Stephen Brenner for writing the appropriate programs and making the computations reported here.

This work was supported in part by USPHS grant GM30902.

APPENDIX

Geometric analysis

In this part a geometric construction is discussed that shows the relationships between the quantities that enter into the three systems of defining equations that are solved in a least-squares fashion to obtain values for the phase differences $\varphi_{1,h}^n - \varphi_{2,h}^n$. The construction not only provides a graphical understanding of the quantities involved and how they are related, but it also affords a clear insight into the basis for obtaining unique or essentially unique values for the phase differences in a one-wavelength anomalous dispersion experiment when there is one type or one predominant type of anomalous scatterer present.

The equations used in the construction are

$$F_{\lambda h} = F_h^n + F_{\lambda h}^a \quad (36)$$

$$F_{\lambda \bar{h}}^* = F_{\bar{h}}^{n*} + F_{\lambda \bar{h}}^{a*} \quad (37)$$

$$F_h^n = F_{\bar{h}}^{n*} \quad (38)$$

$$F_{\lambda \bar{h}}^* = F_h^n + F_{\lambda \bar{h}}^{a*} \quad (39)$$

$$F_h^n = F_{1,h}^n + F_{2,h}^n \quad (40)$$

$$F_{\lambda h} = F_{1,h}^n + F_{2,h}^n + F_{\lambda h}^a \quad (41)$$

$$F_{\lambda \bar{h}}^* = F_{1,h}^n + F_{2,h}^n + F_{\lambda \bar{h}}^{a*} \quad (42)$$

$$F_{\lambda h}^a = \exp(i\delta_{\lambda 2})(f_{\lambda 2}^n/f_{2,h}^n)F_{2,h}^n \quad (43)$$

$$F_{\lambda \bar{h}}^{a*} = \exp(-i\delta_{\lambda 2})(f_{\lambda 2}^n/f_{2,h}^n)F_{2,h}^n \quad (44)$$

$$F_{\lambda h} = F_{\lambda \bar{h}}^* + 2i(f_{\lambda 2}^n/f_{2,h}^n)F_{2,h}^n \quad (45)$$

Equation (36) defines the structure factor for a structure that scatters anomalously, $F_{\lambda h}$, in terms of a structure factor appropriate to the same structure when all atoms scatter nonanomalously, F_h^n , plus a structure factor that represents the total contribution from the anomalous scattering, both real and imaginary, $F_{\lambda h}^a$. Except for the quantity F_h^n , all quantities appearing in (36)–(45) are defined in the body of the paper. Equation (36) is the source equation from which (1) and (2) were derived (Karle, 1980). Equation (37) follows from (36) and is written as the complex conjugate because it is convenient to use it in that form in the construction. Since (38) is valid when atoms do not scatter anomalously, (39) follows. Equation (40) is a statement that F_h^n is composed of the sum of the structure factors for the atoms that scatter nonanomalously, $F_{1,h}^n$, and the structure factors for the atoms that scatter anomalously when the anomalous part of the scattering is omitted, $F_{2,h}^n$. Equations (41) and (42) follow from (36)–(40). Equations (43) and (44) are valid when there is one type of anomalous scatterer. They follow from relations that arise in the derivation of (1) and (2) (Karle, 1980). Equation (45) is immediately derivable from (36), (39), (43) and (44).

The construction shown in Fig. 1 is based on (36)–(45). Equations (1) and (2) were derived from these equations without approximation and their characteristics are properly represented by the characteristics of the diagram. We now proceed to construct the diagram in Fig. 1 in a stepwise fashion. We assume

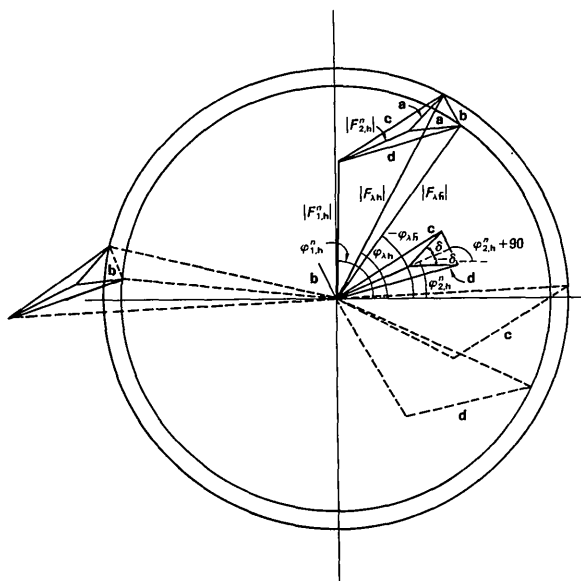


Fig. 1. A construction showing the relationships among the quantities occurring in the equations that are used to evaluate the phase differences $\varphi_{1,h}^n - \varphi_{2,h}^n$. Interpretation of the construction can be correlated with the experiences encountered in performing the test calculations described in the text. $\mathbf{a} = |F_{\lambda h}^a| = |F_{\lambda h}^n| = (f_{\lambda 2}^a/f_{2,h}^n)|F_{2,h}^n|$; $\mathbf{b} = 2(f_{\lambda 2}^n/f_{2,h}^n)|F_{2,h}^n|$; $\mathbf{c} = |F_{2,h}^n + F_{\lambda h}^a|$; $\mathbf{d} = |F_{2,h}^n + F_{\lambda h}^*|$.

that the magnitudes $|F_{\lambda h}|$, $|F_{\lambda h}^n|$ and $|F_{2,h}^n|$ are known. In actual practice, values for $|F_{\lambda h}|$ and $|F_{\lambda h}^n|$ are obtained from experiment and, as described in the paper, values for $|F_{2,h}^n|$ may be obtained approximately from values for $|F_{\lambda h}|$ and $|F_{\lambda h}^n|$ by use of (11). It is also assumed that the chemical identity of the anomalously scattering atoms is known. A value is selected for $\varphi_{2,h}^n$, the angle associated with $|F_{2,h}^n|$, to facilitate the construction of the diagram. The value of $\varphi_{2,h}^n$ is arbitrary and is not determined by the mathematics. This means that the resulting diagram is not determined in orientation. Rather, it can be arbitrarily rotated about an axis perpendicular to the plane of the diagram and placed at the origin.

On the basis of the assumed information, the triangle having solid lines with sides $|F_{\lambda h}|$, $|F_{\lambda h}^n|$ and \mathbf{b} can be drawn, representing (45). A second triangle ambiguously placed can also be drawn, as shown by the triangle with dashed lines and the side labeled \mathbf{b} . It can be obtained from the original triangle by rotating about the vector of length \mathbf{b} placed at the origin. We continue with consideration of the triangle with solid lines and proceed with the construction of (41). All information is available to compute $F_{2,h}^n + F_{\lambda h}^a$ on the right side of (41). It is labeled with a \mathbf{c} in Fig. 1. The head of this vector is attached to the head of the vector of magnitude $|F_{\lambda h}^n|$. Once this is done, it is possible to complete (41) by drawing the vector $F_{1,h}^n$ to close the triangle. Note that a value for $|F_{1,h}^n|$ has been determined here and that a value for the associated angle $\varphi_{1,h}^n$ has also been determined relative to the assumed value for $\varphi_{2,h}^n$. The vector of magnitude \mathbf{d} can now be drawn from the tip of $F_{1,h}^n$ to the tip of $F_{\lambda h}^*$ to form a triangle that represents (42). The vector of magnitude \mathbf{d} is entirely consistent with the construction that could be made from $F_{2,h}^n + F_{\lambda h}^*$, where the magnitude of $F_{\lambda h}^*$ is labeled with \mathbf{a} .

The triangle with sides of magnitude \mathbf{b} , \mathbf{c} and \mathbf{d} is also placed at the origin in order to help illustrate some angles involved in the construction and also to indicate that for a given $\varphi_{2,h}^n$ there is no other place that the construction shown in solid lines can occur. The only other position in which the triangle with sides of length $|F_{1,h}^n|$, \mathbf{c} and $|F_{\lambda h}|$ can occur is indicated by the triangle having dashed lines and a side of length \mathbf{c} . It can be placed by rotating the triangle with solid lines about the vector of length \mathbf{c} placed at the origin. A similar circumstance applies to the triangle with sides of length $|F_{1,h}^n|$, \mathbf{d} and $|F_{\lambda h}^n|$. Its ambiguous alternative is shown by the dashed triangle with a side of length \mathbf{d} . It is evident that the elements of the main construction do not combine elsewhere on the circle of rotation for a fixed value of $\varphi_{2,h}^n$.

The question arises concerning the source of the alternative result often obtained in the calculations discussed in the paper, *i.e.* two different results were often obtained as the value of $|F_{1,h}^n|$ was varied through its three assigned starting values. The expla-

nation follows. On the basis of the information available, the triangle with sides \mathbf{b} , \mathbf{c} and \mathbf{d} could just as well have been attached to the side \mathbf{b} of the alternative dashed-line triangle. The magnitudes $|F_{\lambda\mathbf{b}}|$, $|F_{\lambda\bar{\mathbf{b}}}|$ and $|F_{2,\mathbf{b}}^n|$ would still be preserved as well as the assumed value for $\varphi_{2,\mathbf{b}}^n$. As seen in Fig. 1, such a triangle with sides of length \mathbf{b} , \mathbf{c} and \mathbf{d} has been attached to the dashed triangle with the common base of vectors of length \mathbf{c} and \mathbf{d} extending far to the left. This determines an alternative and quite different vector, $F_{1,\mathbf{b}}^n$. It is quite long, about 2.4 times longer than the initially determined $|F_{1,\mathbf{b}}^n|$, extending as a dashed line from the origin to the base of the vectors of length \mathbf{c} and \mathbf{d} . The associated angle, $\varphi_{1,\mathbf{b}}^n$, is also rotated somewhat more than 90° farther than the initially determined $\varphi_{1,\mathbf{b}}^n$. The magnitude of the initial $|F_{1,\mathbf{b}}^n|$ is about 0.63 of the average value of $|F_{\lambda\mathbf{b}}|$ and $|F_{\lambda\bar{\mathbf{b}}}|$ and that of the alternative is about 1.5 times larger.

It is evident now from Fig. 1 how two alternative sets of results arise and that the alternatives would be distinguishable by use of approximate knowledge of the value of $|F_{1,\mathbf{b}}^n|$. It is of interest to review the assumptions inherent in the diagram as they relate to practical circumstances. There are experimental errors in $|F_{\lambda\mathbf{b}}|$ and $|F_{\lambda\bar{\mathbf{b}}}|$ and $|F_{2,\mathbf{b}}^n|$ can be obtained only approximately from the latter two intensities. The arbitrariness of $\varphi_{2,\mathbf{b}}^n$ does not play a role in the calculations since the quantity evaluated is $\varphi_{1,\mathbf{b}}^n - \varphi_{2,\mathbf{b}}^n$, which is invariant to rotation of the diagram in Fig. 1 around the origin. The effects of the various uncertainties are illustrated in the test calculations. Because of the uncertainties, the variation of starting values for $|F_{1,\mathbf{b}}^n|$ and, on occasion, $|F_{2,\mathbf{b}}^n|$ was introduced into the calculations in order to explore the field of convergence. With the computer used for the test calculations,

6000 distinct least-squares computations were performed in one minute.

The diagram in Fig. 1 emphasizes the important practical significance of having additional information concerning $|F_{1,\mathbf{b}}^n|$, $|F_{2,\mathbf{b}}^n|$ and $\varphi_{2,\mathbf{b}}^n$. As noted, information concerning $|F_{1,\mathbf{b}}^n|$ is available from an isomorphous replacement experiment since $|F_{1,\mathbf{b}}^n|$ represents the magnitude of the structure factor for the native substance. If the structure of the anomalous scatterers is determined initially, values for the $|F_{2,\mathbf{b}}^n|$ are available to enhance the accuracy of the calculations and values for the $\varphi_{2,\mathbf{b}}^n$ are available for the evaluation of the $\varphi_{1,\mathbf{b}}^n$ from values of the $\varphi_{1,\mathbf{b}}^n - \varphi_{2,\mathbf{b}}^n$. The immediate calculation of the electron distribution of the structure of interest would follow.

References

- DALE, D., HODGKIN, D. C. & VENKATESAN, K. (1963). *Crystallography and Crystal Perfection*, edited by G. N. RAMACHANDRAN, pp. 237-242. New York, London: Academic Press.
- GIACOVANZO, C. (1983). *Acta Cryst.* **A39**, 585-592.
- HAUPTMAN, H. (1982). *Acta Cryst.* **A38**, 632-641.
- HENDRICKSON, W. A. & TEETER, M. (1981). *Nature (London)*, **290**, 107-113.
- KARLE, J. (1980). *Int. J. Quantum Chem.* **7**, 357-367.
- KARLE, J. (1984a). *Acta Cryst.* **A40**, 1-4.
- KARLE, J. (1984b). *Acta Cryst.* **A40**, 4-11.
- KARLE, J. (1984c). *Acta Cryst.* **A40**, 526-531.
- PEERDEMAN, A. F. & BIJVOET, J. M. (1956). *Acta Cryst.* **9**, 1012-1015.
- RAMACHANDRAN, G. N. & RAMAN, S. (1956). *Curr. Sci.* **25**, 348-351.
- SHMUELI, U., WEISS, G. H., KIEFER, J. E. & WILSON, A. J. C. (1984). *Acta Cryst.* **A40**, 651-660.
- TIMKOVICH, R. & DICKERSON, R. E. (1976). *J. Biol. Chem.* **251**, 4033-4046.
- WOOLFSON, M. M. (1984). *Acta Cryst.* **A40**, 32-34.

Acta Cryst. (1985). **A41**, 394-399

A Comparison of Weissenberg and Diffractometer Methods for the Measurement of Diffuse Scattering from Disordered Molecular Crystals

BY T. R. WELBERRY

Research School of Chemistry, Australian National University, GPO Box 4, Canberra City, ACT 2601, Australia

AND A. M. GLAZER

Clarendon Laboratory, Parks Road, Oxford OX1 3PU, England

(Received 29 May 1984; accepted 28 February 1985)

Abstract

Diffuse X-ray scattering from a disordered molecular crystal [1,4-dibromo-2,5-diethyl-3,6-dimethylbenzene (BEMB1)] has been measured by diffractometer and conventional Weissenberg-film techniques, and

a detailed comparison of the two sets of data has been made. An interexperimental agreement factor between the two sets of intensities was about 22%. Statistical tests on the data revealed that a substantial part of this discrepancy was mainly due to systematic differences attributable to alignment problems associ-