

Laboratory, Oxford; the author wishes to thank Dr I. G. Wood and Dr A. M. Glazer for their help with these experiments. Thanks are also due to Dr Glazer for a critical reading of the manuscript and for useful comments. Comments of the Editor on the problem of defining ferroelastics are also acknowledged with thanks.

#### References

- ABRAHAMS, S. C. (1971). *Mater. Res. Bull.* **6**, 881–890.  
 AIZU, K. (1969). *J. Phys. Soc. Jpn*, **27**, 1171–1178.  
 AIZU, K. (1970a). *Phys. Rev. B*, **2**, 754–772.  
 AIZU, K. (1970b). *J. Phys. Soc. Jpn*, **28**, 706–716.  
 ARVINDAKSHAN, C. (1958). *Z. Kristallogr.* **111**, 35–45.  
 BATS, J. W. (1978). *Acta Cryst.* **B34**, 1679–1681.  
 COWLEY, R. A. (1976). *Phys. Rev. B*, **13**, 4877–4885.  
 CRACKNELL, A. P. (1974). *Adv. Phys.* **23**, 673–866.  
 DESHPANDE, V. V., MIRZA, E. B. & PHADNIS, A. B. (1979).  
 Personal communication.  
 JANOVEC, V., DVOŘÁK, V. & PETZELT, J. (1975). *Czech. J. Phys.* **B25**, 1362–1396.  
 LONAPPAN, M. A. (1955). *Proc. Phys. Soc. London Sect. B*, **68**, 75–80.  
 MADAN, H. G. (1886). *Nature (London)*, **34**, 66–67.  
 PISTORIUS, C. F. W. T. (1970). *J. Chem. Phys.* **52**, 1009–1010.  
 RAMACHANDRAN, G. N. & LONAPPAN, M. A. (1957). *Acta Cryst.* **10**, 281–287.  
 RAMAN, C. V. & KRISHNAMURTI, D. (1953). *Proc. Indian Acad. Sci. Sect. A*, **36**, 315–334.  
 RAYLEIGH, LORD J. W. S. (1889). *Nature (London)*, **40**, 227–228.  
 SAPIRIEL, J. (1975). *Phys. Rev. B*, **12**, 5128–5140.  
 WADHAWAN, V. K. (1978a). *Mater. Res. Bull.* **13**, 1–8.  
 WADHAWAN, V. K. (1978b). *Curr. Sci.* **47**, 534–535.  
 WOOD, I. G. & GLAZER, A. M. (1980). *J. Appl. Cryst.* **13**, 217–223.  
 WYCKOFF, R. W. G. (1964). *Crystal Structures*, Vol. 2, pp. 383–385. New York: Interscience Publishers.  
 ZACHARIASEN, W. H. (1929). *Z. Kristallogr.* **71**, 501–516.

*Acta Cryst.* (1980). **A36**, 856–864

## The Use of Anomalous Scattering Effects to Phase Diffraction Patterns from Macromolecules

By JAMES C. PHILLIPS\* and KEITH O. HODGSON†

*Department of Chemistry, Stanford University, Stanford, California 94305, USA*

(Received 5 October 1979; accepted 25 March 1980)

### Abstract

Recent experiments using synchrotron radiation to measure anomalous scattering terms at wavelengths very close to *L*-shell elemental absorption edges [Phillips, Templeton, Templeton, & Hodgson (1978). *Science*, **201**, 257–259] have demonstrated that effects much larger than those expected on the basis of simple atomic scattering calculations are observed. These observations of the large wavelength dependence of the magnitudes of anomalous scattering terms has prompted our reexamination of how such effects can be used to phase single-crystal diffraction patterns. This paper describes a methodology for using information on the magnitude of anomalous scattering effects to plan a multiple-wavelength phasing experiment on a macromolecule-containing crystal. Different data collection strategies, such as measuring data at many wavelengths less accurately or at a few wavelengths

accurately, are compared. The method uses principles taken from standard MIR phasing techniques and also can be used to process multiple-wavelength data to obtain the phases. A numerical example of phasing, using the anomalous scattering curves determined experimentally for cesium, is presented and the results discussed in terms of the applicability of multiple-wavelength phasing for protein crystallographic studies. A survey of a variety of heavy metals suggests that large changes in anomalous scattering near the *L* absorption edges are quite general and will be of importance in multiple-wavelength phasing of single-crystal protein diffraction patterns.

### Introduction

It has long been apparent that anomalous scattering can be used to solve the 'phase problem' in crystallography. The experiment is best performed using a synchrotron radiation source, as the largest anomalous scattering effects are obtained only when the wave-

\* Present address: EMBL c/o DESY, Notkestrasse 85, 2000 Hamburg 52, Federal Republic of Germany.

† To whom correspondence should be addressed.

length used in the diffraction experiment is near (within about 0.25%) the elemental absorption edge. Here, we discuss in detail the anomalous-scattering multiple-wavelength phasing method for proteins with a specific example using experimental anomalous dispersion terms around the  $L_3$  edge of cesium.

The normal practice for solving the phase problem in protein-structure determination utilizes multiple isomorphous replacement (Green, Ingram & Perutz, 1954). The method uses changes in scattering produced by isomorphous metal-atom substitutions and requires the native protein crystals plus at least two isomorphous derivatives. In practice, a phase probability function is calculated and a Fourier transform with weighted coefficients is used (Blow & Crick, 1959). The use of many derivatives generally improves the quality of the map.

A second source of phase information is derived from the phenomenon of anomalous scattering. Absorption by heavy atoms leads to two extra terms in their contribution to the X-ray diffraction pattern, the real ( $f'$ ) and imaginary ( $f''$ ) components of the anomalous scattering. The real part has a phase equal to that of the normal scattering and the imaginary part contributes  $90^\circ$  out of phase. Consequently, in noncentrosymmetric crystals, Friedel's law is not obeyed. Thus, measurement of Friedel pairs can resolve the phase ambiguity when only a single isomorphous derivative is available. This effect is often used to improve the accuracy of phase information (Blow, 1958; Blow & Rossmann, 1961). Use of Friedel pairs in combination with isomorphous derivative information in the location of heavy atoms in protein crystals has been discussed by Kartha & Parthasarathy (1965) and Matthews (1966*a*). The way in which Friedel pairs are currently used in phase determination in protein crystallography was originally formulated by North (1965) and Matthews (1966*b*).

It has long been realized that anomalous scattering effects alone can be used to determine completely the phases of the diffraction pattern. A detailed explanation of the principles has been given by Raman (1959). As we have noted earlier (Phillips, Templeton, Templeton & Hodgson, 1978; Templeton, Templeton, Phillips & Hodgson, 1980), as the wavelength is varied through an absorption edge of the heavy atom, the real and imaginary parts of the anomalous scattering vary drastically. Changes in the real part are mathematically analogous to isomorphous-derivative changes. Thus, measurements using several different wavelengths can also be used to obtain the phase.

There has been a theoretical estimate of how well anomalous scattering from sulfur and copper can be used to solve the structure of organic molecules, using Mo and W  $L$ -emission radiation (Herzenberg & Lau, 1967). It was concluded that a structure containing 1000 light atoms could be solved.

A phase determination of the diffraction pattern from crystals of an iron-containing protein was made using data obtained with Ni and Co  $K\alpha$  radiations, wavelengths which occur on either side of the Fe  $K$  edge (Hoppe & Jakubowski, 1975). The phases obtained solely from the use of anomalous scattering differed by an average of  $50^\circ$  from those obtained by multiple isomorphous replacement. In a study using synchrotron radiation to provide variable wavelengths, an accuracy in the phase determination of  $60^\circ$  was obtained for another iron-containing protein, rubredoxin, despite the poor quality of the film data (Phillips, Wlodawer, Goodfellow, Watenpaugh, Sieker, Jensen & Hodgson, 1977).

Anomalous scattering effects have also been used in neutron diffraction. Singh & Ramaseshan (1968) have discussed the method which is analogous to that used for X-rays. Schoenborn (1975) obtained phases from a cadmium derivative of myoglobin using two wavelengths of neutrons. Koetzle & Hamilton (1975) used neutrons at three wavelengths to solve the phase problem for Na [Sm EDTA.  $8H_2O$ ].

An X-ray diffraction facility at a synchrotron radiation source has recently become available (Phillips, Cerino & Hodgson, 1979). This intense source can be used to collect single-crystal diffraction data at incident beam wavelengths between 3 and 0.5 Å.

It is the purpose of this paper to give first the mathematical basis of multiple-wavelength phasing in a form analogous to that used for the multiple isomorphous replacement method. Then, with the recently acquired knowledge of anomalous scattering effects at  $L$  edges (Phillips, Templeton, Templeton & Hodgson, 1978; Templeton, Templeton, Phillips & Hodgson, 1980), the question of how well a diffraction pattern might be phased using these phenomena will be examined. Specifically there will be a discussion of what and how many wavelengths should be used, whether Friedel pairs need be measured and what kind of accuracy can be expected for a given ratio of light to heavy atoms. The treatment given is valid for the case where there is one type of anomalous scatterer in the unit cell. Arndt (1978) has also discussed multiple-wavelength phasing using synchrotron radiation but the emphasis was on data collection rates rather than on what data to collect.

### The effect of anomalous scattering on the diffraction pattern

Fig. 1 shows the contributions to a Bragg reflection from various scatterers within the unit cell where the addition of each contribution of different amplitude and phase is represented as a sum of vectors on the complex plane. The total scattering vector ( $F$ ) is given by:

$$F = F_p + f_H + f' + f'' \quad (1)$$

where  $F_p$  = sum of scattering from the light atoms;  $f_H$  = normal scattering from the heavy atom;  $f'$ ,  $f''$  = real and imaginary parts of the anomalous scattering of the heavy atom, which are respectively  $180^\circ$  and  $90^\circ$  out of phase with  $F_H$ .

The  $|F|$  is the only measurable quantity. When the wavelength of the incident X-ray beam is varied through an absorption edge of the heavy atom,  $f'_H$  and  $f''_H$  will vary in magnitude while  $F_p$  and  $f_H$  stay essentially constant. Fig. 2 shows the summation of vectors for three different wavelengths. The Friedel-related reflection ( $\bar{F}$ ) is also shown for the three wavelengths, inverted through the real axis. The total scattering vectors are omitted for clarity. It is clear that  $F$  is altered slightly in both amplitude and phase as  $f'$  and  $f''$  vary.

#### Criteria governing the choice of wavelengths

A measurement of the amplitude of any three of the resultant  $F$ 's is sufficient in principle to obtain the phase of the Bragg reflection unambiguously. A graphical solution, the Harker (1956) construction, for the case of a measurement of a reflection at three wavelengths is shown in Fig. 3. Circles are drawn with radii equal to the measured  $|F|$  and centered on the appropriate heavy atom plus anomalous scattering vector. The common intersection of all three circles determines the correct phase. Any two measurements give two possible phases where the two circles intersect. Each pair of intersections is symmetric about the line joining the two heavy atom plus anomalous scattering vectors.

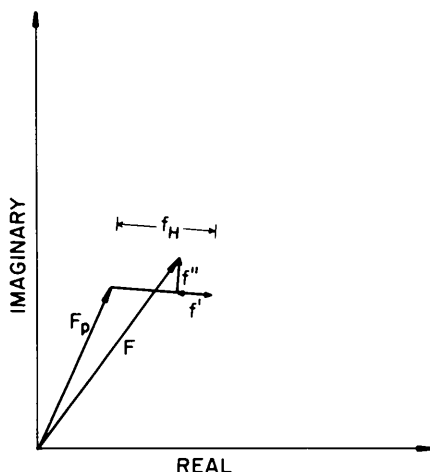


Fig. 1. The amplitude and phase of a Bragg reflection from a protein crystal containing a heavy atom is represented as a vector in the complex plane ( $F$ ).  $F$  is the vector sum of the protein contribution ( $F_p$ ), the contribution from the normal scattering of the heavy atom ( $f_H$ ) and that from the real ( $f'$ ) and imaginary ( $f''$ ) parts of the heavy-atom anomalous scatter.  $f'$  and  $f''$  are respectively  $180^\circ$  and  $90^\circ$  out of phase with  $f_H$ . The anomalous scattering vectors for the heavy-atom contribution ( $f'$  and  $f''$ ) are drawn for the general non-centrosymmetric case in this and subsequent figures.

Fig. 4 is the solution for the case of the measurement of Friedel pairs at one wavelength and one of the pair at a second wavelength. Measurement of the reflection at two wavelengths plus the Friedel-related reflection at a third wavelength would also lead to an unambiguous value for the phase.

In deciding what measurements to make, it is clear that the largest possible changes in  $F$  should be induced so that they can be accurately measured. Also, the ambiguity should be broken by the best choice of measurements such that the line of symmetry of any

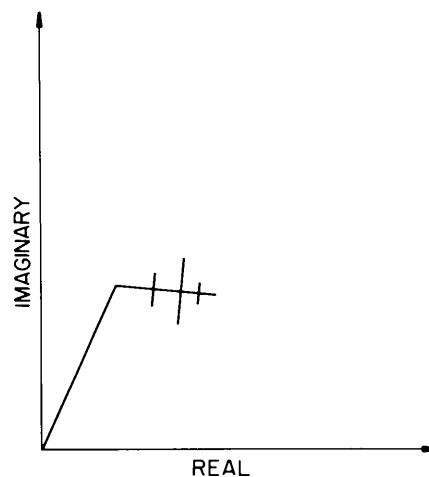


Fig. 2. The contributions to a Bragg reflection at three wavelengths through an absorption edge.  $f'$  and  $f''$  vary rapidly with wavelength but  $F_p$  and  $F_H$  are essentially constant. Also shown is the Friedel-related reflection, reflected through the real axis.  $f''$  adds differently to the Friedel-pair reflections and thus their amplitudes are different. Total vectors have been omitted for clarity.

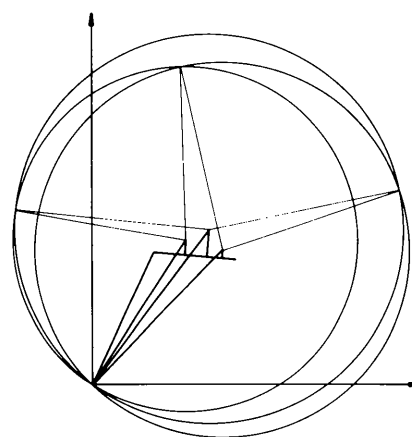


Fig. 3. Harker construction showing how the measurement of a Bragg reflection at three wavelengths can lead to an unambiguous solution for the phase. Any pair of measurements gives two possible solutions. The third measurement resolves this ambiguity.

pair is in a different direction to that of the other pairs. Hoppe & Jakubowski (1975) suggested that measurements be made such that two of the lines of symmetry be as close as possible to right angles with each other. This is correct if only two of the three possible pairings of three measurements are used in phase determination, but is modified if more use of the measurements is made.

In the multiple isomorphous replacement method phases are obtained by calculating a phase probability function,  $P(\Phi)$ , according to the formula

$$P(\Phi) = \exp [-x^2(\Phi)/2\sigma^2], \quad (2)$$

where  $x(\Phi)$  (the 'lack of closure') is the difference between the derivative amplitude as measured and as calculated from the native amplitude and the heavy-atom parameters assuming a phase  $\Phi$ .  $\sigma$  is the estimated average error in the measurements for that particular derivative. The total probability function is the product of the functions for each derivative. The phase used to calculate electron density maps is then the weighted mean of this probability function (Blow & Crick, 1959).

When this method is used with normal MIR data, there is generally no attempt to relate the measurements between two derivatives as there can be worse errors due to imperfect isomorphism between two derivatives than between a native and a derivative. Also, the introduction of a heavy atom into the crystal often causes problems. Crystals may crack, diffract less strongly, and suffer more greatly from radiation damage. Thus the error in a measurement between two derivatives can be greater than that between a native and a derivative. A third factor is the error due to the imperfect estimation of the heavy-atom parameters which are again compounded if two derivative measurements are compared.

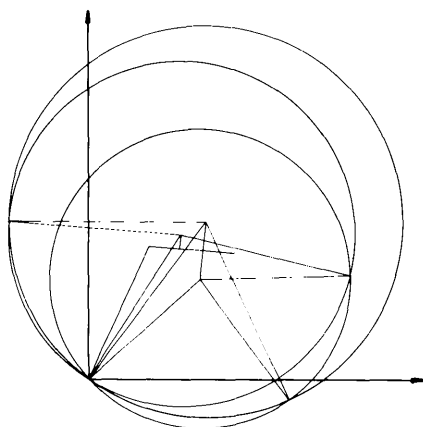


Fig. 4. As for Fig. 3, but now a measurement of Friedel-related reflections plus a measurement of one of the pair at a second wavelength is assumed. Again an unambiguous phase determination is possible.

In a phase determination by multiple-wavelength measurements, the above considerations are greatly modified. There are no isomorphism errors. Measurements would be made on the same crystal, which would contain the heavy atom. Therefore, all of the data will be of similar quality (except that Friedel-related reflections differ by worse absorption errors) and errors in the estimation of heavy-atom parameters would affect the correlation of any pair of measurements in exactly the same way. In each phase analysis, the total phase probability function can thus be taken as the product of those between all possible pairs of the measurements assumed. If data from several measurements are correlated in this way, it is possible that the correct choice of the minimum three measurements is not that of Hoppe & Jakubowski (1975). Rather than choose measurements such that two of the symmetry lines are at right angles, it is conceivable that the three symmetry lines are  $60^\circ$  apart from each other.

Fig. 5 shows some values of the anomalous scattering terms for cesium at the Cs  $L_{III}$  edge (Templeton, Templeton, Phillips & Hodgson, 1980) plotted with  $f'$  as abscissa and  $\pm f''$  as ordinate to show what vector differences between measurements at different wavelengths and Friedel pairs are possible with a small change of wavelength. Table 1 lists the values used (the point  $f' = -13, f'' = 13$  is an interpolation). It is quite possible to obtain large vector changes closely perpendicular to each other. Particularly note that a measurement of a Friedel pair

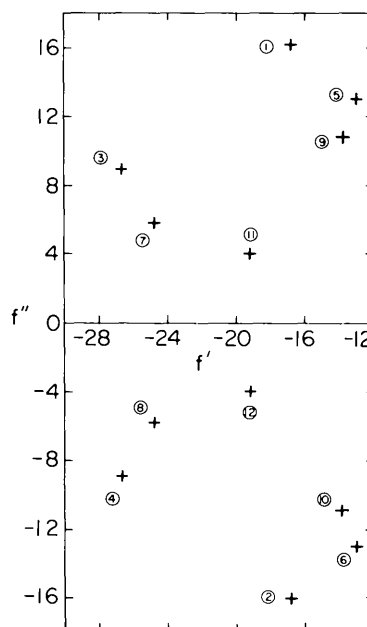


Fig. 5. Plot of  $f'$  versus  $\pm f''$  for cesium through the  $L_{III}$  edge. This plot shows the vector changes to the total  $F$  which it is possible to make by altering the wavelength or measuring Friedel related reflections. The points are referred to by their numbers in the text.

Table 1. *Anomalous scattering terms used in the estimates of the accuracy of multiple-wavelength phasing*

$\lambda$ (Å)	$\lambda/\lambda_{III}$	$f'$	$f''$	Comments
2.4793	1.0021	-19.2	4.0	measured
2.4751	1.0004	-24.8	5.8	measured
2.4737	0.9999	-26.7	8.9	measured
2.4722	0.9993	-16.8	16.1	measured
2.4710	0.9988	-13	13	interpolation
2.4703	0.9985	-13.9	10.8	measured

(points 1 and 2 on Fig. 5) coupled with measurements at two other wavelengths (points 3 and 9 or 5) is close to the supposed ideal case with the largest variation in one component of the anomalous scattering associated with little change in the other. Moreover, the wavelengths involved are within 0.28% of each other. The approximation that other scattering factors are not changing with wavelength will be valid in such a small range. Points 1, 4 and 6 are closer to the 60° condition while still being quite widely apart.

#### *Consideration of the number of wavelengths to be measured*

As described above, each measurement makes possible the calculation of more phase probability functions, the  $J$ th measurement gives  $(J - 1)$  more. In some senses this is like the effect of having several isomorphous derivatives. However, as data from more wavelengths are considered, the magnitudes of the intensity changes will decline. Also, there is little possibility of phasing reflections for which the structure factor of the anomalous scatterer is low, whereas this becomes increasingly less true as more derivatives with different sites are used in the MIR method. Collecting more data also uses more synchrotron radiation beam time, at present a precious commodity. There is also the question of the availability of a sufficient quantity of crystals.

#### **A method for assessing various data collection strategies**

When the anomalous scattering terms are known [either from diffraction experiments (Phillips, Templeton, Templeton & Hodgson, 1978) or absorption measurements (Templeton, Templeton, Phillips & Hodgson, 1980)], as is the case with cesium, it is possible to consider the above questions in a quantitative way. To do this, a Fortran computer program was written which generates artificial data at various wavelengths for an arbitrary phase and analyses them assuming different combinations of measurements.

The program steps are as follows.

(1) A random number generator is used to obtain an arbitrary phase from 0 to 360°.

(2) A structure factor for the heavy atom is obtained from the amplitude of the sum of  $N$  randomly oriented unit vectors [angles obtained by the same procedure as in step (1)].  $N$  is the product of the number of heavy-atom sites and the number of symmetry-related sites.

(3) The heavy-atom anomalous scattering vectors for various wavelengths are added to a constant protein vector using the above obtained phase structure factor. The relative scattering amplitudes of the protein and the anomalous scattering atom are set using the following formula derived by Crick & Magdoff (1956).

$$2\Delta F/F = \Delta I/I = [2N_H/N_L]^{1/2} f_A/f_L. \quad (3)$$

$\Delta F/F$ ,  $\Delta I/I$  = relative amplitude and intensity change due to the isomorphous addition of the anomalously scattering electrons;  $N_H$  = the number of heavy atoms in the unit cell;  $N_L$  = the number of light atoms approximated by the molecular weight divided by 14.7;  $f_A$  = anomalous scattering factors;  $f_L$  = average normal scattering factor for the protein (taken to be seven electrons, an average of C, N and O).

(4) Gaussian distributed noise with a r.m.s. error of  $\sigma$  is added to the resultant  $F$ 's to obtain the artificial data.

(5) The phase is recovered from the data assuming various different measurements have been made. The lack of closure [ $x(\varphi)_{12}$ ] between two measured amplitudes  $M_1$  and  $M_2$  at wavelengths  $i$  and  $j$  is calculated as follows. From (1),

$$\mathbf{M}_1 = \mathbf{F}_p + \mathbf{f}_H + \mathbf{f}'_i \pm \mathbf{f}''_i, \quad (4)$$

where the sign of  $\mathbf{f}''_i$  is set according to which of the Friedel pair  $\mathbf{M}_1$  corresponds to. Similarly,

$$\mathbf{M}_2 = \mathbf{F}_p + \mathbf{f}_H + \mathbf{f}'_j \pm \mathbf{f}''_j. \quad (5)$$

Therefore,

$$\mathbf{M}_1 = \mathbf{M}_2 + \mathbf{f}'_i - \mathbf{f}'_j \pm \mathbf{f}''_i \mp \mathbf{f}''_j. \quad (6)$$

By squaring (6) the lack of closure due to error is found to be

$$\begin{aligned} x(\varphi)_{12} = & M_1 - [M_2^2 + (f'_i - f'_j)^2 + (\pm f''_i \mp f''_j)^2 \\ & + 2M_2 \cos \varphi (f'_i - f'_j) \\ & + 2M_2 \sin \varphi (\pm f''_i \mp f''_j)]^{1/2}, \end{aligned} \quad (7)$$

where  $\varphi$  is the phase difference between the heavy-atom vector and  $\mathbf{M}_2$ . Phase probability functions for all possible pairings of the measurements are assumed and they are multiplied together to obtain the final phase.

(6) The differences between the correct phase and those obtained by the different phasing methods are accumulated.

(7) The procedure is repeated many times then average phase differences and other information are tabulated.

The estimates of the average phase error expected, as obtained by the above calculation, will be lower than that to be expected in an actual experiment as several effects have not been taken into account. Specifically, the heavy-atom parameters and the anomalous scattering terms are assumed to be known precisely, which would not be true for a real experiment. The formula of Crick & Magdoff (1956) used to set the scale between the heavy atom and protein contributions assumes average  $F$ 's based on Wilson statistics, which, it is known, protein crystals do not follow closely. The structure factor of the protein was not varied, although varying this would probably not affect the results appreciably, and what is required is an estimate of the ability to phase average reflections rather than very weak or very strong ones. Also, the weakest and strongest measurements are measured the least and most accurately respectively, so keeping a constant protein vector with a constant error should be a fairly good approximation. However, the estimates presented should be taken only as upper limits on the phasing power of the anomalous-scattering multiple-wavelength method. However, they can certainly be used as guides in the planning of an experiment, indicating the relative advantages of one data collection strategy over another.

### An examination of results of the numerical methods using Cs anomalous scattering for phasing

The effects of making different types and numbers of measurements were investigated using the above described computer program as a function of the molecular weight of the protein and the error in the measurements. Molecular weights of 12 000, 25 000, 50 000 and 100 000 per anomalously scattering Cs atom were used. The range of  $\sigma$  was 0.5–3.0% of  $F$ , corresponding to a 1–6% error in the intensity measurement. All results given are for averages over 1000 reflections and for four heavy-atom positions within the unit cell (e.g.  $P2_12_12_1$  symmetry). In tests with 500 reflections, no substantial differences in the results were found so 1000 reflections are sufficient for good averaging of all effects to occur. The mean phase errors calculated by the program are given in Table 2 for all of the ten cases described below.

In case I, the minimum of three measurements (points 1, 2 and 3 of Fig. 5) is assumed. The points are chosen to maximize the Friedel-pair difference and then to obtain the largest possible difference of  $f'$  at the second wavelength. In case II, the same measurements plus the Friedel pair at the second wavelength (point 4) are assumed to be measured. In case III again the minimum of three measurements (points 3, 5 and 6) is assumed. There is a trade-off of a smaller  $f''$  for a larger difference in  $f'$  between this and case I. Case IV

Table 2. Estimated mean phase error for various phasing methods, assumed error in intensities and protein molecular weight

Mol. wt	$\sigma$ (%)	Case									
		I	II	III	IV	V	VI	VII	VIII	IX	X
120 000	1	6.3	5.2	6.5	5.2	4.4	5.6	4.2	4.1	4.0	10.7
	2	12.1	10.3	11.5	9.4	7.6	11.0	7.4	7.2	7.0	19.2
	3	18.8	14.5	16.0	13.4	11.0	15.0	10.6	10.8	10.4	28.4
	4	23.6	19.2	21.7	17.2	14.3	21.8	12.9	13.9	12.9	37.3
	5	29.5	25.0	27.0	21.8	18.4	26.5	16.9	16.9	15.9	43.2
	6	30.4	26.4	30.6	26.1	22.0	28.5	20.1	20.5	19.4	47.3
25 000	1	8.5	7.5	8.1	6.7	5.6	8.1	5.2	5.2	4.7	14.2
	2	15.6	12.8	15.0	11.9	9.8	14.5	9.5	8.6	8.1	26.6
	3	26.5	22.1	24.0	20.0	16.8	23.1	16.2	14.5	14.0	39.5
	4	32.7	29.4	30.3	25.9	22.4	29.8	20.5	20.3	18.8	48.1
	5	34.9	31.4	32.8	27.8	23.6	32.6	22.1	22.1	20.2	50.5
	6	42.5	38.1	41.2	35.7	30.7	39.4	29.4	27.7	26.1	59.3
50 000	1	12.2	10.3	9.7	8.3	7.2	9.8	6.5	6.8	6.2	19.8
	2	21.2	18.1	20.2	16.3	13.8	19.3	12.9	12.9	11.6	33.2
	3	30.4	26.5	29.2	24.8	20.2	28.5	19.4	17.9	16.9	46.0
	4	41.3	36.1	38.5	32.2	28.5	36.6	26.7	26.3	23.6	57.8
	5	43.4	38.6	43.7	38.6	33.0	40.9	30.9	30.6	29.0	60.2
	6	50.2	46.8	47.7	43.3	38.4	47.8	36.2	35.1	32.8	66.8
100 000	1	17.3	14.4	14.6	12.9	11.1	14.8	10.0	9.8	9.1	26.3
	2	30.9	26.4	26.9	23.3	20.5	26.7	18.7	19.4	17.8	45.1
	3	41.7	38.0	37.9	33.0	30.0	39.0	27.6	28.2	25.1	56.3
	4	47.5	43.8	48.1	41.1	35.9	47.8	33.5	33.8	30.9	63.6
	5	51.1	47.7	53.3	49.3	42.9	51.6	40.7	40.9	38.2	68.9
	6	56.6	55.3	58.3	52.9	47.9	58.3	45.1	45.5	42.4	72.0

again assumes that the Friedel-pair reflection at the second wavelength has also been measured (point 4). Case V assumes all six measurements of cases I–IV have been made (points 1, 2, 3, 4, 5 and 6). Case VI, involving measurements at points 2, 3 and 5, is a compromise between cases I and III.

In order to investigate whether adding measurements at other wavelengths would improve the phasing, cases VII–IX were tested. In case VII, points 7 and 8 are added to those of case V. In case VIII, points 9 and 10 are added to those of case V and case IX assumes all ten measurements at the five wavelengths have been made.

If single-counter diffractometry is the data collection method, it can be inefficient to measure Friedel pairs. Data collection is halted while driving between the pairs, which takes much longer than positioning on a reflection with one index higher and this must be done every few measurements if all phasing measurements of one Bragg reflection are to be made close together in time on the same crystal. Case X assumes measurements of all positive-index reflections only at three wavelengths (points 1, 3 and 11).

Judging from the table of mean phase errors (summarized in Table 2), of the four cases which assume only that the minimum of three measurements have been made, case VI appears to be slightly better than case III. Case I, utilizing the maximum Friedel-pair splitting, is less favorable than cases III and VI, where somewhat more  $f'$  changes occur. Case X is clearly unfavorable and would only be used in specialized circumstances, where the requirements for measuring Friedel pairs causes great inefficiencies in data collection, and there is a high ratio of heavy to light atoms.

It can again be seen from Table 2 that of the two cases which assume all measurements at two wavelengths have been made, case IV appears better than case II, just as case III was better than case I assuming three of the four measurements. When data from all three measurements are added, there is a further improvement in the phase determination. Adding a fourth or fifth wavelength again reduces the phase error. However, in general it would be more efficient to spend data collection time in improving accuracy of the best measurements rather than collecting data at many wavelengths.

A two or three wavelength case with the emphasis on accuracy of the data collection appears most optimal. However, there is a limit to the accuracy with which measurements can be made. If it turns out that there are no limitations other than instrument accuracy, *i.e.* a plentiful supply of samples and data collection time are available, measurements at many wavelengths will further improve the phase determination.

The above conclusions as to possible data collection methods apply only where the anomalous scattering

curves are similar to those for cesium. For each anomalous scatterer, a similar analysis should be performed in order to decide on the best data collection method, once the anomalous scattering curves are known.

### Some other considerations in the use of anomalous scattering

#### *Location of the heavy atom*

Before phases can be calculated, it is necessary to locate the heavy atom. This can be done by computing a Patterson map or by direct methods. Approximations to the heavy-atom  $F$ 's are required, and may be obtained from the multiple-wavelength data by the following method. From (7) one can show

$$\mathbf{M}_1 - \mathbf{M}_2 = (f'_i - f'_j) \cos \varphi + (\pm f''_i \pm f''_j) \sin \varphi, \quad (8)$$

where second-order terms have been neglected and it is assumed that there is no error in the measurements. This equation may be rewritten

$$\Delta = f_H(r_1 \cos \varphi + r_2 \sin \theta), \quad (9)$$

where  $r_1 = (f'_i - f'_j)/f_H$  and  $r_2 = (\pm f''_i \pm f''_j)/f_H$  are ratios of normal scattering to combinations of anomalous scattering terms. The ratios are known constants valid for all reflections if the anomalous scattering terms for the particular element used are known. When  $N$  measurements of a Bragg reflection with varying anomalous scattering components have been made then one has  $N(N - 1)/2$  equations of the form

$$\Delta_k = f_H(r_{k1} \cos \varphi + r_{k2} \sin \varphi), \quad (10)$$

$r_{k1}$  and  $r_{k2}$  being the appropriate ratios for the  $k$ th pair of measurements. This set of linear equations with  $f_H \cos \varphi$  and  $f_H \sin \varphi$  as unknowns is overdetermined if more than three measurements have been made and can be solved by the method of least squares to minimize the effect of errors in the  $\Delta_k$ 's. Then the Patterson coefficient can be obtained by use of the relationship  $\cos^2 \varphi + \sin^2 \varphi = 1$ .

#### *Data reduction considerations*

Data taken at multiple wavelengths would in practice be analysed in a similar fashion to MIR data. After heavy-atom sites are found, there would be a refinement of the phases by adjustment of the heavy-atom parameters as with MIR methods (Dickerson, Kendrew & Strandberg, 1961). Schoenborn (1975) did this during the anomalous scattering phasing of the neutron diffraction data from myoglobin. It is a simple matter to adapt existing computer programs to calculate the

above described phase probability functions. Estimates of the mean lack-of-closure error ( $\sigma$ ) can be obtained from comparison of the observed and calculated lack of closure on centric reflections, exactly as with MIR methods (Blow & Crick, 1959).

It should be noted that phase refinement of multiple-wavelength data may prove simpler than that of MIR data. For either method, each measurement contributes a constraining equation to the refinement procedure but in the MIR method each derivative also adds new variables (the coordinates, occupation and thermal motion of each site), while with multiple-wavelength data the number of variables is constant, since the heavy-atom sites remain fixed irrespective of the changes in its anomalous scattering, and many measurements are included. In high-symmetry space groups there are few independent reflections of low resolution. In such space groups, there can be problems of convergence of MIR data if all derivatives have several heavy-atom sites, as there are many variables and only a few measurements to fix the variables. Refinement of one derivative multiple wavelength data could prove more productive.

### Summary

The methodology described here can be used to govern the choice of data collection strategy for a multiple-wavelength phasing experiment and, furthermore, can be applied to obtain the phases from the data thus collected. The detailed knowledge of the wavelength dependence of the anomalous scattering effects is required for optimal use of the proposed methods. As we have suggested (Phillips, Templeton, Templeton & Hodgson, 1978; Templeton, Templeton, Phillips & Hodgson, 1980), this information can be obtained by an absorption spectroscopy measurement of the protein metal derivative being studied. It is also clear from the measurements discussed here, and a large number of other measurements at  $L$  edges made at storage-ring sources, that measurements to make use of the rapid changes in  $f'$  and  $f''$  with wavelength can be made using existing monochromators.

It is important to observe that the multiple-wavelength method requires only one heavy-atom-containing protein crystal. Unfortunately, the  $L$  edges of all the common biologically important metals have absorption edges in the soft X-ray region so preparation of one derivative is still required (however, it need not be isomorphous with the native crystal). It is, in principle, possible to use  $K$  absorption edges for these experiments as well but the smaller effects at  $K$  edges (approximately 1/3 of those seen at  $L_{III}$  edges) imposes more stringent requirements on the statistics of the data.

We have surveyed the high-resolution  $L$  X-ray absorption spectra of a variety of lanthanide and heavy transition-metal complexes. The  $L_{III}$  absorption edges for several elements are shown in Fig. 6. It is in fact now quite clear that large changes in anomalous scattering terms are generally observed near the  $L_{III}$  and  $L_{II}$  edges for many heavy atoms because of the strong resonance-like absorptions (so-called 'white lines') that occur at absorption threshold. Besides the strong absorption at threshold, the spectra in Fig. 6 also show weaker modulations to higher energies which result from EXAFS (Cramer & Hodgson, 1979). Neither the 'white lines' nor the modulation due to EXAFS (which will also be reflected in  $f''$  and  $f'$ ) are accounted for by atomic scattering theory. Many elements commonly used for making heavy-atom derivatives of proteins such as uranium, osmium, platinum and the lanthanides exhibit these 'white lines'.

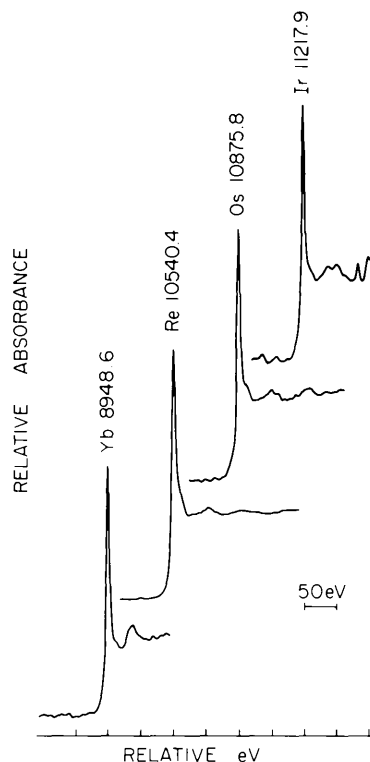


Fig. 6. The  $L_{III}$  absorption edges for Yb and several heavy elements. The peak heights are normalized to unity and the energy scale is relative (increasing energy left to right) to facilitate comparison. Absolute energy values are given for the maximum of each absorption peak. The complexes are  $\text{Yb}(\text{CdPm})_3$ ,  $\text{ReCl}_3$ ,  $\text{OsCl}_6(\text{Bu}_4\text{N})_2$  and  $\text{NaIrCl}_6$ . The strong resonance-like absorptions at threshold are classically referred to as 'white lines'. The rapid variance of absorption coefficient with wavelength results in very large changes of the anomalous scattering terms (see text). The weaker modulations seen above the absorption edge (in energy) are a result of EXAFS.



Since  $f''$  is related to the absorption cross section by the optical theorem and  $f'$  can be derived from  $f''$  using the Kramers–Kronig dispersion relation, it is possible to estimate the magnitude of anomalous scattering terms from absorption data (Lye, Phillips, Kaplan, Doniach & Hodgson, 1980). Such an estimation for several lanthanides reveals values of  $f''$  as great as 28 electrons and  $f'$  values as large as  $-30$  electrons. The magnitudes of these effects at  $L_{III}$  edges for lanthanides are, for example, even larger than those observed for cesium as described herein. It is expected that these very large effects, which can only be effectively exploited with a synchrotron source, will be increasingly used in protein crystallography.

We wish to thank Professors David Templeton and Seb Doniach and Paul Phizackerley for numerous helpful discussions. This work was supported by the National Institutes of Health through grant CA 16748. Synchrotron radiation beam time was provided by the Stanford Synchrotron Radiation Laboratory which is a National facility supported by the NSF (grant DMR-77-2874) in cooperation with the Stanford Linear Accelerator Center and the US Department of Energy.

#### References

- ARNDT, U. W. (1978). *Nucl. Instrum. Methods*, **152**, 307–311.
- BLOW, D. M. (1958). *Proc. R. Soc. London Ser. A*, **247**, 302–336.
- BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–799.
- BLOW, D. M. & ROSSMANN, M. G. (1961). *Acta Cryst.* **14**, 1195–1202.
- CRAMER, S. P. & HODGSON, H. O. (1979). *Prog. Inorg. Chem.* **25**, 1–39.
- CRICK, F. H. C. & MAGDOFF, B. S. (1956). *Acta Cryst.* **9**, 901–908.
- DICKERSON, R. E., KENDREW, J. C. & STRANDBERG, B. (1961). *Acta Cryst.* **14**, 1188–1195.
- GREEN, D. W., INGRAM, V. M. & PERUTZ, M. F. (1954). *Proc. R. Soc. London*, **225**, 287–307.
- HARKER, D. (1956). *Acta Cryst.* **9**, 1–9.
- HERZENBERG, A. & LAU, H. S. M. (1967). *Acta Cryst.* **22**, 24–28.
- HOPPE, W. & JAKUBOWSKI, U. (1975). *Anomalous Scattering*, edited by S. RAMASESHAN & S. C. ABRAHAMS, pp. 437–461. Copenhagen: Munksgaard.
- KARTHA, G. & PARTHASARATHY, R. (1965). *Acta Cryst.* **18**, 745–749.
- KOETZLE, T. F. & HAMILTON, W. C. (1975). *Anomalous Scattering*, edited by S. RAMASESHAN & S. C. ABRAHAMS, pp. 489–502. Copenhagen: Munksgaard.
- LYE, R. C., PHILLIPS, J. C., KAPLAN, D., DONIACH, S. & HODGSON, K. O. (1980). Submitted to *Proc. Natl Acad. Sci. USA*.
- MATTHEWS, B. W. (1966a). *Acta Cryst.* **20**, 230–239.
- MATTHEWS, B. W. (1966b). *Acta Cryst.* **20**, 82–86.
- NORTH, A. C. T. (1965). *Acta Cryst.* **18**, 212–216.
- PHILLIPS, J. C., CERINO, J. A. & HODGSON, K. O. (1979). *J. Appl. Cryst.* **12**, 592–600.
- PHILLIPS, J. C., TEMPLETON, D. H., TEMPLETON, L. K. & HODGSON, K. O. (1978). *Science*, **201**, 257–259.
- PHILLIPS, J. C., WLODAWER, A., GOODFELLOW, J. M., WATENPAUGH, K. D., SIEKER, L. C., JENSEN, L. H. & HODGSON, K. O. (1977). *Acta Cryst.* **A33**, 445–455.
- RAMAN, S. (1959). *Proc. Indian Acad. Sci. Sect. A*, **50**, 95–107.
- SCHOENBORN, B. P. (1975). *Anomalous Scattering*, edited by S. RAMASESHAN & S. C. ABRAHAMS, pp. 407–416. Copenhagen: Munksgaard.
- SINGH, A. K. & RAMASESHAN, S. (1968). *Acta Cryst.* **B24**, 35–39.
- TEMPLETON, D. H., TEMPLETON, L. K., PHILLIPS, J. C. & HODGSON, K. O. (1980). *Acta Cryst.* **A36**, 436–442.

*Acta Cryst.* (1980) **A36**, 864–869

## Classification of Directions in Crystallographic Point Groups According to the Symmetry Principle

BY R. V. GALIULIN

*Institute of Crystallography of the USSR Academy of Sciences, 59 Leninsky Prospect, Moscow, USSR*

(Received 9 June 1979; accepted 8 April 1980)

### Abstract

It is shown that a complete geometrical classification of directions in crystallographic point groups may be constructed by means of partitioning directions according to connected totalities whose directions possess the same complete group of symmetry. In all

the 32 crystallographic point groups there exist 358 different connected regions of directions.

The set of all symmetry operations of a crystallographic point group which transfer the given direction into itself will be referred to as the *direction group* of