Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

A MAD experiment performed at the white line of the iridium L_{III} absorption edge in lysozyme

Gwyndaf Evans[a,b]*† and Keith S. Wilson[a]‡

[a]European Molecular Biology Laboratory (EMBL) Hamburg, c/o DESY, Notkestrasse 85, 22603 Hamburg, Germany, and [b]Department of Physics, University of Warwick, Coventry CV4 7AL, England

† Present address: MRC Laboratory of Molecular Biology, MRC Centre, Hills Road, Cambridge CB2 2QH, England.
‡ Present address: Protein Structure Group, Department of Chemistry, University of York, Heslington, York YO1 5DD, England.

Correspondence e-mail: gwyndaf@anl.gov

Received 18 August 1997
Accepted 21 May 1998

A multiple-wavelength anomalous diffraction (MAD) experiment was performed on an iridium derivative of hen egg-white lysozyme. Diffraction data were measured at five wavelengths on the X31 EMBL beamline on the DORIS III storage ring and at two further wavelengths on the X11 beamline. Four iridium-binding sites were located from the dispersive anomalous differences between two wavelengths at the rising and falling inflection points of the Ir L_{III}-edge white line using direct methods. All other attempts to determine the heavy-atom positions failed. The results demonstrate an experimental method whereby systematic error in MAD data due to sample absorption can be reduced where a white line is present in the absorption spectrum of a heavy atom.

1. Introduction

The method of multiple-wavelength anomalous diffraction has been a subject of interest ever since synchrotron radiation first made it a practical method of obtaining phase information from diffracted intensities. In the late 1970s and early 1980s test studies were made by Phillips *et al.* (1977), who performed a seven-wavelength experiment around the Fe *K* absorption edge in rubredoxin, and Hendrickson & Teeter (1981), who solved the structure of the small protein crambin using the anomalous scattering of the S atoms from its three disulfide bridges. Around this time, ideas for multiple-wavelength anomolous diffraction (MAD) phasing were presented by Phillips & Hodgson (1980*a*), who adopted an approach analogous to that used in MIR phasing, and Karle (1980) and Hendrickson (1985), who developed an algebraic method for phase determination from multiwavelength data.

It was evident very early on that significant problems arise when attempting MAD because of the very small (a few percent) anomalous intensity differences which need to be measured. In order to improve the anomalous signal-to-noise ratio, statistical and systematic error within the diffraction data must either be minimized during the experiment or, in the case of systematic error, carefully removed afterwards during data analysis. In conjunction with this philosophy, the anomalous signal should be maximized by careful selection of the appropriate wavelengths for an experiment. A considerable number of 'optimized' single-wavelength experiments have been performed to provide additional phase information where a native and a single derivative form of a protein are available (Phillips & Hodgson, 1980*b*; Baker *et al.*, 1990). The wavelength has typically been set to be just short of the heavy atom's L_{III} absorption edge in the case of third transition-

series metals or the K edge for first transition-series metals, where f'' approaches its maximum.

The local variation in the anomalous scattering factors can be as much as 2–3 times greater than the theoretically calculated values (Cromer & Libermann, 1970) when a white line is present at an absorption edge. A study of the absorption edges of heavy atoms suitable for biological studies was made by Lye *et al.* (1980). In particular, large white-line features were observed at the L_{II} and L_{III} edges of the lanthanides and of three metals of the third transition series, Re, Os and Ir.

The limitations of the multiple-wavelength method were tested computationally by Phillips (1978) using anomalous scattering factors at wavelengths around the Cs L_{III} edge. It was observed that maximum contrast in the anomalous scattering factors could be obtained by utilizing the white-line feature of the Cs L_{III} edge and measuring data at three wavelengths corresponding to the rising inflection point of the edge (f' minimum), the white line (f'' maximum) and either the falling inflection point or a remote short-wavelength point (f' maximum).

This method and a phasing approach analagous to MIR were used in the solution of *Opsanus tau* parvalbumin by MAD (Kahn *et al.*, 1985). Two Ca atoms were substituted by Tb which has a large white line at its L_{III} edge (f'' = 19.9 e). Data were measured at three wavelengths corresponding to the white-line maximum and the rising and falling inflection points of the Tb white line.

More recent examples of multiple-wavelength experiments using heavy atoms with white lines have been performed (Weis *et al.*, 1991; Shapiro *et al.*, 1995; McDonald *et al.*, 1995). Here, the lanthanides Ho and Yb were used for the structure determination, together with programs using the algebraic approach to MAD phasing.

More recently, data measurement and analysis strategies for MAD have been investigated by Peterson *et al.* (1996) using crystals of a brominated oligonucleotide.

We describe here a white-line MAD experiment which has been performed on an Ir derivative of lysozyme, the purpose of which was (*a*) to test the apparatus used for X-ray wavelength stability (the results of which have been published elsewhere) and (*b*) to attempt to draw conclusions about the best possible choice of and number of wavelengths which should be used in a MAD experiment.

2. Experimental apparatus

Two of the EMBL X-ray beamlines, X31 and X11, at DORIS III in Hamburg were used for the experiment. The X31 line consists of a single channel-cut Si crystal monochromator which is positioned just upstream of a prealigned toroidal segmented quartz mirror which provides 1:1 demagnification. The wavelength is easily tuned on this line over 1–2 keV without any readjustment of the optics. On entering the experimental hutch, the X-rays first encounter a wavelength-stabilization apparatus and then pass through two sets of vertical and horizontal slits, each preceded by a small ionization chamber and separated by 150 mm. The X-rays are then incident on the sample and diffraction is recorded by an area detector. At the time of these experiments, X31 was still equipped with EMBL's 'RED' imaging-plate detector (the prototype of the MAR imaging-plate system). The apparatus in the experimental hutch is mounted on a fully adjustable optical bench which can be automatically aligned into the X-ray beam with the aid of the two ionization chambers. The band pass of this beamline was, at the time, dominated by the source size of the DORIS ring and a band pass of $\Delta E/E = 3.5 \times 10^{-4}$ was attainable using an Si(111) cut monochromator. Therefore, at the time of these studies, an Si(311) cut crystal was installed which produced a band pass of $\Delta E/E = 1.8 \times 10^{-4}$ with about 30% of the original flux on the sample crystal. Measurements of the band pass were made using an Si crystal analyser. The (573) reflection was measured in back-reflection at a Bragg angle of $\sim 70^\circ$. The analyser crystal is, in fact, an integral part of the stabilization apparatus. This apparatus and the wavelength-stabilization technique have previously been described in detail (Evans & Pettifer, 1996), but are briefly summarized here. The X-ray beam size of X31 is typically much larger than any protein-crystal samples and the residual beam is used to sample the X-ray wavelength. A single piece of silicon is used to back-diffract these residual X-rays into a pair of detectors which are positioned above and below the X-ray beam. The crystal is oriented such that hkl and $hk\bar{l}$ pairs are simultaneously excited at the desired wavelength for an experiment. By monitoring the intensities of the excited pair, the experiment is made sensitive to any fluctuations of the incident wavelength and it is possible to apply a small change to the monochromator angle and thus correct the shift. Using this technique, wavelength stability to typically better than 4×10^{-5} is achievable.

The X11 line consists of a bent Ge(111) horizontally focusing single-reflection monochromator followed by a spherical vertically focusing mirror. Demagnification of about 4:1 is attained on this line. The experimental hutch is essentially the same as on X31, apart from an additional arm which rotates the whole optical bench about the monochromator vertical axis to facilitate wavelength changes. (The wavelength, however, is very rarely changed on this line.) The area detector used was the 180 mm diameter MAR imaging-plate scanner, which boasted a factor of four greater dynamic range over the 'RED' scanner. The geometry of the beamline makes it near to impossible to use the wavelength-stabilization device; however, the band pass was measured using the (375) reflection of an Si analyser crystal and found to be 3.0×10^{-3}, which clearly demonstrates that no stabilization is necessary, since any features in an absorption spectrum which were considerably less than the band pass would be completely smeared out. This beamline is clearly unsuitable for MAD, but was used here to obtain better quality data for use in the phasing procedures.

3. Experimental methods

Tetragonal crystals of hen egg-white lysozyme were prepared in the conventional manner (Blundell & Johnson, 1976).

Derivatives of the iridium complex were prepared by soaking the crystal in a 5 m*M* solution of Na_3IrCl_6 in 0.85 *M* NaCl at pH 4.7 (Blake, 1968). After a few days the crystals began to turn a pale yellow–brown colour which gradually turned to a dark brown after 2–3 weeks of soaking. All data were measured from a single crystal having dimensions 0.5 × 0.5 × 0.25 mm which had been soaked for two weeks. Iridium was chosen because it was known to bind to lysozyme (as established by Blake) and it also possessed a large white-line feature (determined from preliminary fluorescence experiments).

Fluorescence spectra were measured directly from the protein crystal on the X31 beamline using a photomultiplier tube coupled to a fast plastic scintillator. Selection of the X-ray wavelengths for data collection was made on the basis of f' and f'' anomalous scattering factors which were determined directly from the fluorescence spectra using the program *CHOOCH* (Evans, 1994) which is based on the Kramers–Kronig transformation algorithm described by Hoyt *et al.* (1984). Five wavelengths were chosen for the X31 narrow band-pass experiment. All but the remote wavelengths are shown in Fig. 1. Two further wavelengths were measured on the broad band-pass ($\Delta E/E \simeq 3 \times 10^{-3}$) higher intensity beamline X11. Wavelength selection on this line is much more difficult owing to its single-reflection monochromator geometry and was estimated from the diffraction pattern of lysozyme, assuming accurate measurement of the sample-to-detector distance. Fig. 2 illustrates how the choice of wavelengths on X31 is optimal for producing the largest differences in f' and f''.

In this experiment, no attempt was made to align the crystal in such a way that the two members of a Bijvoet pair would be excited simultaneously. This allowed truly complete data sets to be recorded without the need for realignment of the crystal. In addition, each data set was measured fully at one wavelength before proceeding to the next because, at the time of the experiment, the wavelength-stabilization software had not been developed to the stage where wavelengths could be rapidly interchanged. Reflections were recorded up to a minimum *d* spacing of 2.5 Å using the constant-dose method, *i.e.* the crystal sample received a similar overall X-ray dose for each exposure, as measured by an ionization chamber positioned just before the sample. For the data sets at λ_2, λ_3 and λ_4, where the anomalous scattering factors are very sensitive to wavelength changes, the wavelength-stabilization technique was used to produce a monochromatic beam stable to within 2.7×10^{-5} ($\Delta\lambda/\lambda$). The λ_1 and λ_5 sets were measured without using the apparatus, since the small slopes of f' and f'' at these

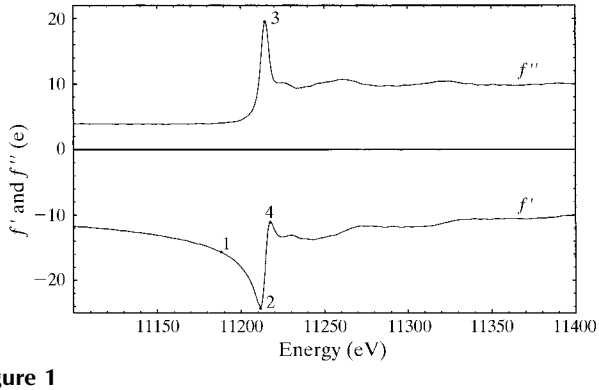

Figure 1
Plot of f' and f'' against energy around the iridium L_{III} absorption edge calculated from X-ray fluorescence measurements on an iridium derivative of crystalline lysozyme. The noise in the experimental fluorescence data was at the level of about 1%. The estimated error in the calculation of the f'' values is ±0.3 e which may be introduced as a result of the spline-fitting routine in the above edge region. Data sets 1, 2, 3 and 4 were measured at the X-ray energies labelled in the diagram.

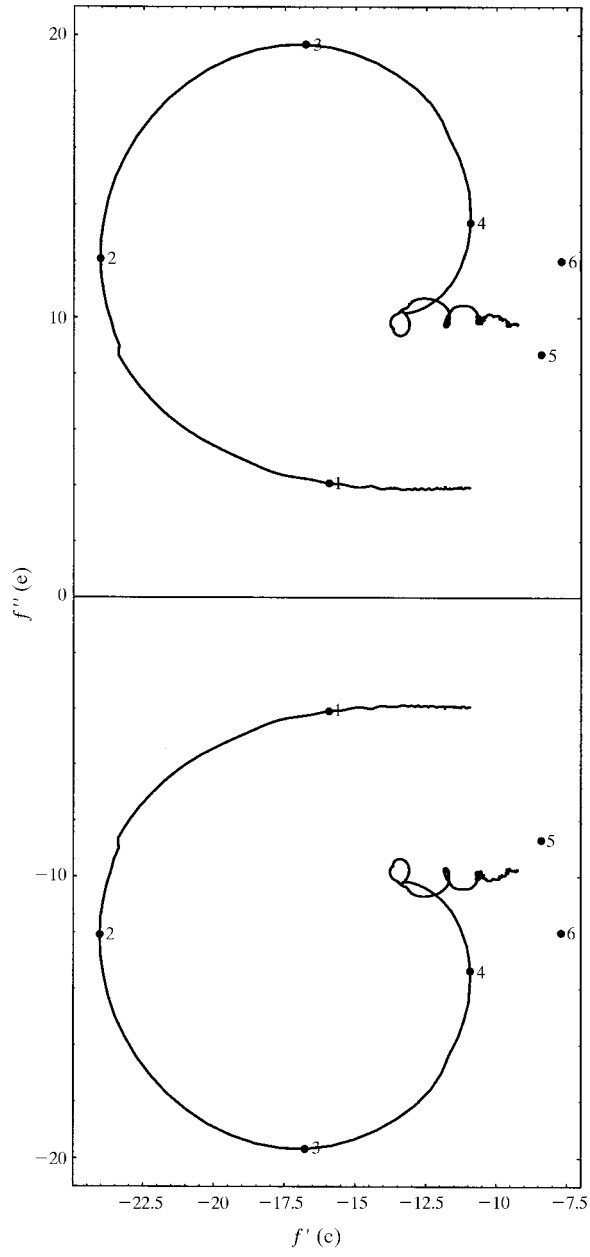

Figure 2
Plot of f' *versus* f'' for the experimentally determined anomalous-scattering factors. The positions corresponding to the wavelengths used to measure the seven data sets are labelled. This demonstrates the rationale where wavelengths are chosen which maximize the anomalous contrast in both the f' and f'' scattering factors of iridium.

Table 1
Diffraction data from the iridium derivative of lysozyme.

Data were collected at five X-ray energies at and around the iridium L_{III} absorption edge using the X31 beamline (λ_1–λ_5) and at two more energies using the X11 beamline (λ_6–λ_7). The table lists these energies along with the values of the experimentally determined anomalous scattering factors for iridium.

Data set	Description	Wavelength (Å)	Energy (eV)	R_{merge} (%)	Redundancy (%)	Complete (%)	N_{unique}	f'' (e)	f' (e)
λ_1	Below edge	1.10800	11190.0	4.5	4.51	99.4	7776	4.1	−16.1
λ_2	Rising edge	1.10587 (5)	11211.5 (5)	6.0	4.18	99.6	7791	12.1	−24.0
λ_3	White-line peak	1.10555 (5)	11214.7 (5)	6.9	4.37	99.9	7886	19.7	−16.8
λ_4	Falling edge	1.10527 (5)	11217.6 (5)	6.3	4.36	99.5	7776	13.4	−10.9
λ_5	Above edge	1.00800	12300.0	6.8	4.53	98.5	7337	8.7†	−8.4†
λ_6	Above L_I edge	0.92011	13475.0	3.5	4.32	99.5	8020	12.0†	−7.7†
λ_7	Just below edge	1.10701	11200.0	4.8	4.28	99.2	7664	ND‡	ND‡

† Values calculated using the program *CROSSEC* (Cromer & Libermann, 1970). ‡ No direct measurement of f' and f'' was possible for data set 7, and values from *CROSSEC* are not applicable for the region around an absorption edge.

Table 2
The results of scaling the multi-wavelength data together, assuming data set 4 to be native.

Shown are the percentage r.m.s. Bijvoet and dispersive differences, the mean differences as tabulated in the *SCALEIT* output and the corresponding imaginary and real scattering-factor magnitudes. The observed differences for λ_5 data are larger that expected from the magnitude of f'' at this wavelength. The R_{merge} value for this data set was also slightly higher than expected. The probable cause of this was crystal miscentring which was resolved after measuring this data set. (λ_5 was actually the first wavelength measured.) The relatively lower observed differences for λ_6 and λ_7 are representative of the lower statistical error in these data sets.

Data set	R.m.s. Bijvoet difference (%)	R.m.s. dispersive difference (%)	Mean Δ_{ano}	Mean Δ_{iso}	f'' (e)	$\Delta f'$ (e)
λ_1	4.06	3.98	7.0	6.7	4.1	7.9
λ_2	5.99	0.00	10.6	0.0	12.1	0.0
λ_3	8.07	4.25	15.1	6.9	19.7	7.2
λ_4	6.25	4.68	11.2	8.0	13.4	13.1
λ_5	6.21	5.42	10.8	9.5	8.7†	15.6
λ_6	4.08	4.88	7.5	8.7	12.0†	16.3
λ_7	3.80	4.34	6.2	7.3	ND	ND

† Values calculated using the program *CROSSEC* (Cromer & Libermann, 1970).

wavelengths made the experiment insensitive to the typical wavelength shifts of $\sim 5 \times 10^{-4}$.

The λ_6 to λ_7 sets were measured on the X11 beamline. The wavelength of X11 is typically set to 0.9201 Å which corresponds to the peak of the L_I absorption edge of Ir. This was therefore chosen as the λ_6 wavelength. The wavelength was then changed to correspond to the f' minimum of the Ir L_{III} edge. The data collection lasted 8 h, 2 h of which were occupied by the change of wavelength.

4. Data analysis and partial structure determination.

Integrated intensities were determined from the raw diffraction data and scaled using *DENZO* and *SCALEPACK* (Otwinowski, 1993). The results of the preliminary data analysis are given in Table 1. The value of R_{merge} assumes Freidel pairs to be equivalent and therefore gives an indication of the relative anomalous signal present in the data. The data sets are almost fully complete to 2.5 Å. This is a result of the high symmetry of the crystal and the offset in crystal orientation.

Scaling together of the data sets at different wavelengths was performed with *SCALEIT* (Collaborative Computational Project, Number 4, 1994). In dealing with multiwavelength data where there are no native data, one of the wavelengths is selected as a pseudo-native set. In principle any wavelength may be chosen, but the wavelength where f' is at its most negative is convenient, since this is where the scattering power of the heavy atom and, therefore, of the whole structure is at its minimum. Thus, a parallelism is maintained between the multiwavelength and the MIR case. Table 2 shows the results of the inter-wavelength scaling using the λ_2 data as the 'native'.

Bijvoet difference Patterson maps calculated using λ_3 data appeared to be very noisy, contained no significant peaks and proved impossible to interpret, as did dispersive difference Patterson maps using $(\lambda_4 - \lambda_2)$ and $(\lambda_4 - \lambda_5)$ data. Direct methods were therefore employed to solve the heavy-atom structure. *MULTAN* (Main *et al.*, 1980) has already been successfully used to solve heavy-atom structures using isomorphous differences between native and derivative data (Navia & Sigler, 1974; Wilson, 1978) and using Bijvoet differences (Mukherjee *et al.*, 1989).

Initial attempts with *MULTAN* using the λ_3 Bijvoet differences were performed by employing rejection criteria similar to those of Mukherjee *et al.*, *i.e.* Bijvoet differences were rejected if $|\Delta_{Bij}| > 3.5 \times$ r.m.s. ($|\Delta_{Bij}|$) or if $F < 6\sigma(F)$ for either of the contributing structure factors. The $6\sigma(F)$ cutoff was chosen to maintain a high accuracy of data for input but still allow a sufficient number of phase relationships to be used by *MULTAN*. The rejection cutoff of $3.5 \times$ r.m.s. ($|\Delta_{Bij}|$) was, in fact, the limit suggested by the output of *SCALEIT*. A number of attempts were made using different resolution ranges starting with all the data between 20 and 2.5 Å and then lowering the high-resolution cutoff in steps of 0.5 Å until only data between 20 and 6 Å were used. In addition the 5–2.5 Å shell of data was also used in one attempt. The phase

Table 3
The fractional coordinates of the five peaks in the *E* map calculated from the best *MULTAN* solution as judged by the figures of merit.

Shown alongside are the corresponding refined atomic positions of the iridium atoms. The four highest peaks in the *E* map corresponded to the four most occupied iridium sites. There was, however, no correlation between the peak heights in the *E* map and the final refined occupancies of the Ir atoms.

MULTAN E-map peaks				Refined Ir positions			Refined occupancies	
x	*y*	*z*	Height	*x*	*y*	*z*	Relative	Absolute
0.620	0.614	0.398	1675	0.615	0.612	0.368	0.42	0.24
0.388	0.586	0.874	961	0.393	0.584	0.879	0.48	0.27
0.190	0.566	0.544	680	0.186	0.586	0.548	1.00	0.56
0.362	0.618	0.627	847	0.382	0.645	0.635	0.30	0.17
0.193	0.571	0.748	517	No corresponding site				

sets obtained typically produced a number of peaks in the *E* map at special positions and were therefore rejected.

This procedure was then repeated using the dispersive differences between sets λ_2 and λ_4. The f' difference between these two wavelengths is 13.1 e and the f'' values at the two points are 12.1 and 13.4 e, respectively. In all other respects the data sets are essentially identical owing to the very small wavelength difference between them of 5×10^{-4} Å (~6.1 eV). Using a similar argument to that given by Mukherjee *et al.*, it can be assumed that for the largest values of $^{2,4}\Delta_{\rm dis}$ where $\cos(\varphi_P - \varphi_H) \simeq 1$, the dispersive differences are a good estimate of the length of the vector difference between the two structure factors at these wavelengths, *i.e.* $^{2,4}\Delta_{\rm dis} \simeq |^2\mathbf{F}_H'| - |^4\mathbf{F}_H'|$. Dispersive differences were rejected if $|^{2,4}\Delta_{\rm dis}| > 3.5 \times$ r.m.s $|^{2,4}\Delta_{\rm dis}|$ or if $F < 6\sigma(F)$ for either of the contributing structure factors. On this basis 699 and four terms, respectively, were rejected from the data and 3655 differences were used. After normalization of the structure factors to give *E* values, only the 200 strongest and 100 weakest reflections were input into *MULTAN*. A total of 1538 phase relationships were initially determined and, of these, 1067 were accepted after convergence. The 100 weakest reflections took part in a total of 1360 phase relationships. A total of 16 phase sets were output. The best solutions were indicated by the three figures of merit listed by *MULTAN*. The first, ABSFOM, is an indication of a phase set's deviation from random, where a value of 0 is random and 1 indicates a true solution. The second, RESID, shows how the triplet invariants deviate from their statistically predicted behaviour, where lower values generally indicate the best solution. Finally, PSIZERO tests whether the phase relationships which include the 100 weakest reflections are consistent with the phase set produced using the 200 strongest reflections. Lower values indicate better agreement. In addition, a combined FOM is calculated as a weighted average of ABSFOM, RESID and PSIZERO where the relative weights are 0.6, 1.2 and 1.2, respectively.

The best combined FOM (2.993) was observed for phase set 5 which also had the lowest RESID (26.74) and the highest ABSFOM (0.623). The PSIZERO value of 2.719 was the third lowest out of the 16 solutions. The *E* map calculated using this phase set displayed five large peaks, none of which were at special positions. These heavy-atom positions were compared with those found in anomalous difference Fourier maps calculated with refined phases of tetragonal lysozyme. In fact, eight sites were found in the difference Fourier map, three of which were very weak. Table 3 shows the fractional coordinates of the heavy-atom positions determined by *MULTAN* and the corresponding positions after positional refinement. The lowest of the five peaks in the *E* map had no corresponding site in the difference Fourier map. The large number of iridium sites found compared to the single site determined by Blake may reflect the long soaking time used for the derivative crystals. Also, the large number of sites of varying occupancy could explain the difficulties experienced in interpreting the Patterson maps.

5. Results

Phase determination was performed using *MLPHARE* (Otwinowski, 1991; Bricogne, 1991; Collaborative Computational Project, Number 4, 1994). Again, data set 2 was chosen as the 'native' set but was then included again as a 'derivative' such that its Bijvoet differences contributed to the phase calculation but its dispersive difference contribution was zero. Heavy-atom refinement and phase calculation were performed several times while including different data sets into the procedure. Data to 2.75 Å was used for heavy-atom refinement and phase calculation. Using each phase set, real-space correlation coefficients (RSCC's) were calculated on a residue-by-residue basis using *OVERLAP MAP* (Collaborative Computational Project, Number 4, 1994) and observed phase discrepancies were calculated against the refined structure of the derivative (crystallographic *R* factor = 14.6%). A total of eight Ir sites were used in the phase calculation. Density modification was then performed on the experimentally determined phases using *DM* (Cowtan, 1994). The different phasing procedures are described below and the results are summarized in Table 4. Corresponding electron-density maps are shown in Fig. 3.

(*a*) (λ_1, λ_2, λ_3, λ_4, λ_5). Phasing was performed using all five data sets measured on the X31 beamline. This solution gave the highest observed figures of merit (FOMs) from *MLPHARE* of 0.548 for acentric and 0.582 for centric reflections. The RSCC of the electron-density map for all atoms was 0.554 before density modification and 0.610 after.

(*b*) (λ_2, λ_3, λ_4, λ_5). The low remote wavelength was excluded. The FOMs dropped with respect to solution (*a*) but the experimental RSCCs are essentially the same. However, after density modification the final RSCCs are higher than those for solution 1.

(*c*) (λ_2, λ_3, λ_4, λ_6). Substituting λ_6 for λ_5 resulted in a significant increase in the acentric FOMs and an improvement in both the experimental and *DM* maps.

Table 4
Result of phasing with several combinations of data measured at different wavelengths.

Mean phase deviations and correlation coefficient are shown for the maps produced by phases from *MLPHARE* and the phases after running *DM*. Also shown are the mean figures of merit for the *MLPHARE* phases. Phasing was performed using all of the bound iridium sites found in the anomalous difference Fourier map.

	Data set combination	FOM			ΔΦ (°)	RSCC (experimental)			ΔΦ (°)	RSCC (*DM*)		
		Acentric	Centric	All	(exp)	All	Main	Side	(*DM*)	All	Main	Side
(*a*)	$\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5$	0.548	0.582	0.562	51.7	0.554	0.674	0.532	50.4	0.610	0.705	0.591
(*b*)	$\lambda_2, \lambda_3, \lambda_4, \lambda_5$	0.528	0.553	0.534	51.2	0.558	0.679	0.538	48.8	0.636	0.732	0.614
(*c*)	$\lambda_2, \lambda_3, \lambda_4, \lambda_6$	0.591	0.572	0.586	50.7	0.574	0.698	0.555	48.0	0.666	0.754	0.641
(*d*)	$\lambda_2, \lambda_3, \lambda_4$	0.442	0.431	0.439	51.2	0.529	0.655	0.500	50.8	0.590	0.688	0.561
(*e*)	$\lambda_2, \lambda_3, \lambda_5$	0.436	0.446	0.434	51.2	0.535	0.652	0.510	50.0	0.607	0.691	0.582
(*f*)	$\lambda_2, \lambda_3, \lambda_6$	0.530	0.493	0.521	50.5	0.561	0.687	0.544	47.1	0.671	0.754	0.653
(*g*)	$\lambda_2, \lambda_4, \lambda_6$	0.495	0.513	0.499	48.6	0.577	0.691	0.555	45.3	0.683	0.768	0.662
(*h*)	λ_2, λ_4	0.318	0.324	0.320	51.2	0.501	0.621	0.465	54.4	0.533	0.617	0.506
(*i*)	λ_6, λ_7	0.292	0.246	0.281	52.6	0.466	0.586	0.443	52.4	0.550	0.623	0.529
(*j*)	λ_2, λ_6	0.403	0.392	0.400	48.7	0.543	0.666	0.522	46.3	0.668	0.746	0.644

(*d*) (λ_2, λ_3, λ_4). Phases were then calculated using only the three wavelengths measured around the white line within 6 eV of each other. There was a noticeable drop in the FOMs compared to the five- and four-wavelength solutions to 0.442 (centric) and 0.431 (acentric). The RSCCs for both experimental and *DM* maps also decreased slightly.

(*e*) (λ_2, λ_3, λ_5). The falling-edge inflection-point data set (λ_4) was then replaced by the short-wavelength data set (λ_5). An improvement is observed in the centric FOM compared to (*d*) due to the increased contrast in f' (15.6 between λ_2 and λ_5 compared with 13.1 between λ_2 and λ_4). The centric FOMs are lower, presumably owing to the corresponding decrease in f'' contribution from λ_5 with respect to λ_4. The RSCCs after *DM* are comparable to the five-wavelength solution (*a*).

(*f*) (λ_2, λ_3, λ_6). The remote-wavelength data set (λ_5) is replaced by the remote data set from X11 (λ_6). An significant improvement in the FOMs and RSCCs is observed. The improvements arise from the factor of two lower statistical error in the remote data set λ_6 compared with λ_5 and the larger f'' (12.0 e) of the λ_6 data. The $\Delta f'$ values for λ_6 and λ_5 are similar.

(*g*) (λ_2, λ_4, λ_6). The inflection-point data sets λ_2 and λ_4 were used plus the remote X11 data set (λ_6). Removal of the peak data set (λ_3) has lowered the acentric FOM. However, the centric FOM has increased. The RSCCs for both the experimental and *DM* maps are the highest of all the solutions.

(*h*) (λ_2, λ_4). Phases were calculated using only the rising and falling inflection-point data sets (λ_2 and λ_4) measured on X31. The FOMs are much lower than those obtained using three, four or five wavelengths. The RSCCs after density modification are the lowest of all the solutions.

(*i*) (λ_6, λ_7). Phases were calculated with only the X11 data (λ_6 and λ_7). The FOMs are very poor because of the very small dispersive differences between the two data sets. The experimental RSCCs are also the lowest of all; however, in this case *DM* has a significant effect on the RSCC, although the resulting electron density still shows poor connectivity.

(*j*) (λ_2, λ_6). Finally, the rising inflection-point data (λ_2) and the X11 remote data set (λ_6) were used. This is clearly the better of the two-wavelength solutions and although the FOMs are poor compared to solutions (*a*)–(*g*), the RSCCs before and after density modification are comparable. After running *DM* the RSCCs are bettered only by solutions (*f*) and (*g*).

In all the cases described above, the overall phase discrepancy ($\Delta\Phi$) was lowered by the density-modification procedure. The best improvements were observed for the four-wavelength solutions (*b*) and (*c*) and the three-wavelength solutions (*f*) and (*g*), where decreases of $\sim 3^\circ$ were observed.

From Table 4 we can make a number of observations. Firstly, when comparing the three-wavelength solutions, (*d*), (*e*), (*f*) and (*g*), with the four- and five-wavelength solutions, (*a*), (*b*) and (*c*), we see that the improvement in the phasing achieved by including λ_6 into the phasing [solutions (*f*) and (*g*)] instead of λ_3 and λ_5 from solutions (*d*) and (*e*), respectively, is more significant that the improvement seen by including one or two extra wavelengths. The FOMs, however, would suggest otherwise. Even so, when comparing cases (*b*) and (*d*), addition of an extra wavelength has proved valuable, since additional phase information has been added owing to the larger $\Delta f'$ of λ_5 over λ_4.

When comparing solutions (*c*) and (*g*), we observe that addition of λ_3 into the phasing reduces the RSCCs. This is very probably because of the greater effect of crystal absorption at the white-line peak introducing a systematic error into the λ_3 data. A similar but lesser effect is also seen when comparing (*c*) and (*f*), where λ_4 is included. Moreover, when (*f*) and (*g*) are compared, we observe an improvement in the RSCCs when the λ_3 data set is replaced with the λ_4 set. We have, in this case, sacrificed contrast in f'' in order to reduce systematic error contributions from absorption and improve the phasing power from the dispersive signal by ensuring that at least two of the data sets have identical absorption errors.

This point is also well illustrated by comparing solutions (*d*) and (*g*), where substituting λ_3 for λ_6 produces a significant improvement in the phasing. It is interesting to note that the Cullis *R* factors for the Bijvoet differences are essentially identical over the whole resolution range for λ_3 and λ_6. This arises from a factor of two improvement in the statistics for λ_6 compared with λ_3, with a f'' signal reduced by about a half. The

Cullis R factor for the dispersive differences are, however, significantly better for the λ_6 data set.

It is evident that even a well measured two-wavelength solution can still provide good phasing information as can be seen from solution (*j*). This supports the observations of Peterson *et al.* (1996), who successfully obtained a fully interpretable electron-density map of a brominated oligonucleotide using only two wavelengths. The poor experimental RSCCs obtained from solution (*i*) are attributed to poor wavelength calibration of the λ_7 data set combined with the broader band pass, which smeared out the sharp features of the white line and modified the value of f' at its minimum from -24 to -17 e. This was estimated by convoluting the experimentally determined anomalous-scattering-factor curves in Fig. 1 with a Gaussian having a FWHM of 33 eV ($\Delta E/E = 3 \times 10^{-3}$) which approximates the effect of the broad band pass. The electron-density maps from solutions (*h*) and (*i*) were untraceable due to poor connectivity of the maps.

The trend in the FOMs is more correlated with the number of wavelengths used in the phasing than with the individual quality of the solution obtained, although within a group of solutions calculated with the same number of wavelengths, their behaviour appears to correlate with the choice of individual wavelengths. This is presumably a result of calculating a combined probability distribution function from individual functions which are assumed to be statistically independent but which are, in fact, very highly correlated. The FOMs are clearly not a good estimator of the quality of phasing.

The overall quality of the phasing for these data was not of a very high standard. This can be attributed to a number of factors. Firstly, the data was measured at room temperature. Although there was no significant indication of radiation damage, refinement of the structure showed a large number of disordered side chains. Since, in most cases, the Ir atoms were bound to external side chains, their thermal displacement parameters were in turn very large. In addition, the occupancies of the Ir sites varied very widely, suggesting that a number of the atoms may have been very weakly bound. These factors would greatly reduce the contribution of anomalous electrons at higher resolution. Indeed, above 3 Å the phasing power falls off rapidly, although the $R_{\rm merge}$ values for data in these resolution shells are only 2–3% worse that the overall $R_{\rm merge}$ values for each data set. Even under these circumstances the quality of the electron-density maps for the best solutions would have allowed interpretation of large portions of the structure. The data was of a sufficient quality, therefore, to allow conclusions to be drawn about the relative merits of the various strategies employed above.

6. Discussion

The results presented here suggest that when a significant white line exists at a heavy-atom absorption edge the dispersive anomalous differences may provide a more accurate set of measurements with which to determine heavy-atom positions than Bijvoet differences.

The use of Bijvoet differences in heavy-atom determination can suffer from the inherent problem that when a data set is measured at the f' maximum wavelength, this necessarily implies that the heavy-atom absorption contribution is also at a maximum. The absorption errors introduced into the data will therefore be at their largest.

However, a set of dispersive anomalous differences measured at the rising and falling inflection points of a white line (with the condition imposed that f'' is equal at both wavelengths) will be inherently more accurate for the following reasons. Dispersive differences are, in this case, intensity differences between two reflections having identical indices at two different wavelengths where

(i) the position of the reflections on the detector surface differ by $\sim$100 µm only, thus minimizing any errors which may be introduced by an imperfect non-uniformity of response correction for the detector,

(ii) the path lengths through the protein crystal are identical for both reflections and

(iii) both reflections are measured where the sample linear-absorption coefficient is the same.

This implies that these reflections will have identical systematic error contributions arising from crystal absorption and detector imperfections. The absorption corrections between reflections having differing indices will, of course, still be different and depend on the path length through the crystal traversed by the reflected beam. This approach appears to have had a significant effect on the phase calculation; the two data sets at the white-line inflection points (λ_4 and λ_5) and the third remote wavelength set (λ_6) provided the best phase set. Inclusion of an additional wavelength (λ_3) in fact lowers the RSCCs of the *DM* map and has no significant effect on the experimental map.

In all probability, careful alignment of the crystal for this experiment may well have reduced some of the systematic error in the f'' measurements due to absorption. The morphology of the crystal used in this experiment would have indeed been helpful in reducing these errors. Peterson *et al.* (1996) demonstrated this clearly by measuring data at the f'' peak of the Br K edge with an aligned crystal of a brominated oligonucleotide and with the same crystal misaligned. Using information from only two wavelengths, they found that inclusion of the misaligned-crystal data gave an effectively uninterpretable map, whereas inclusion of the aligned-crystal data provided clearly better maps. It is not surprising, however, that absorption would be a crucial factor in these experiments, as there are two Br atoms per 244 light atoms and the crystal has a plate-like morphology (dimensions $0.2 \times 0.1 \times 0.01$ mm).

A commonly discussed question when performing MAD experiments is the number of wavelengths that should be measured. It is generally accepted that at least three are required in order to obtain a map which could allow the structure to be interpreted, although two is the theoretical minimum. However, if experimental time permits, it is usually the case that a fourth wavelength is measured at another remote wavelength. We have presented some evidence to

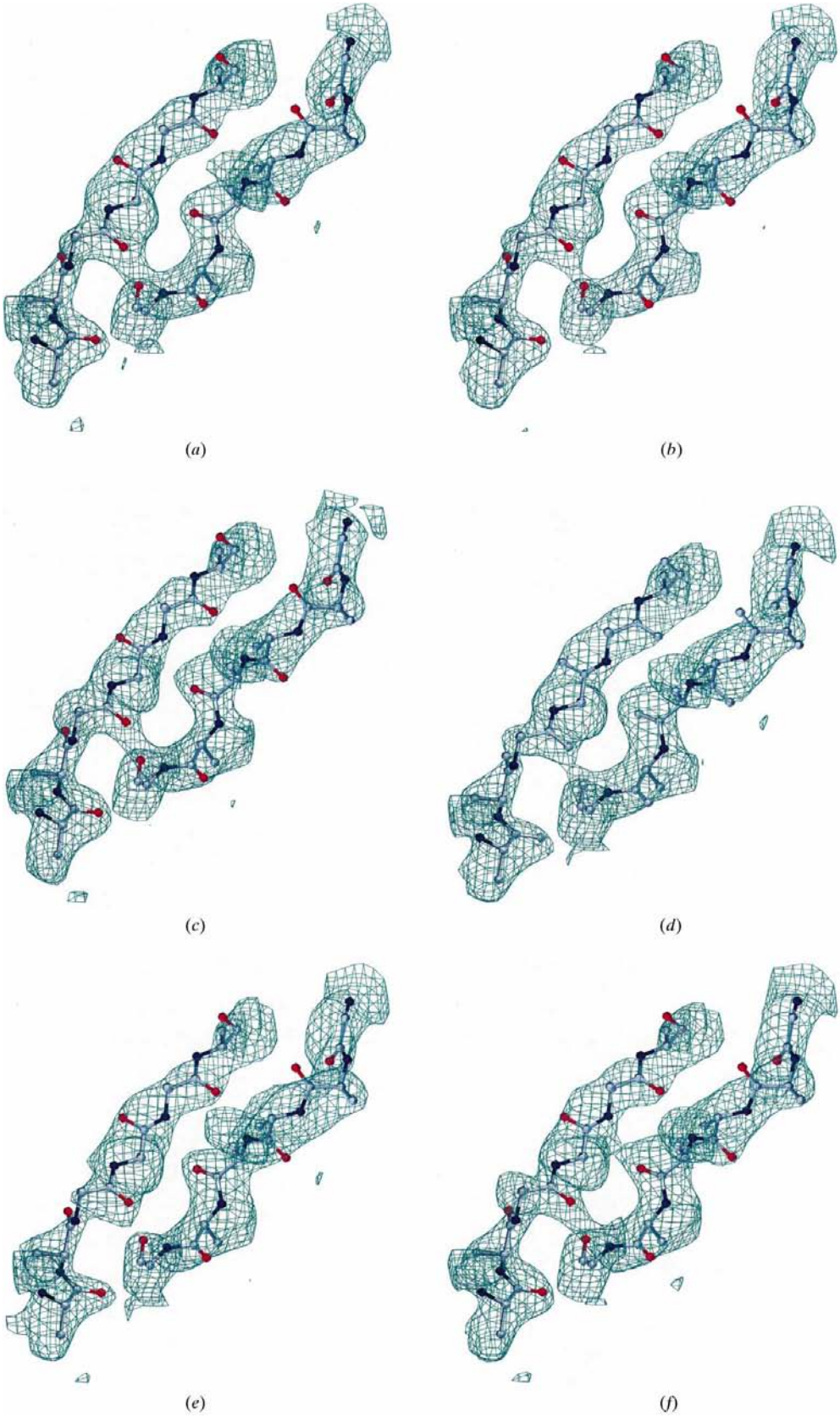

Figure 3
Experimental electron density for the ten phasing solutions (*a*) to (*j*) shown around a β-strand region extending from residues 40–45 and 49–54. The backbone is shown as polyalanine to demonstrate better the main-chain connectivity in this region of the electron density.

suggest that it may be advantageous to measure more accurate data at three wavelengths only, rather than measuring a greater number of wavelengths. This would imply repeating the measurements at one or more of the three wavelengths to reduce the statistical error in the data. The improvement obtained in our case also reflects a factor of four increase in the detector's dynamic range and the increased flux of the X11 beamline over X31. These factors meant that the overall experimental time on X11 was not significantly increased. Ordinarily, obtaining a factor of two improvement in the statistics of one of three data sets would mean doubling the overall time for the experiment. This is not always an option.

One could argue that measuring one or more of the wavelengths twice to improve the statistics of that measurement is essentially equivalent to measuring another data set of similar quality at a different wavelength where f' and f'' are similar and including it into the phase calculation. So why not measure another wavelength? Once three wavelengths have been chosen, one wavelength to maximize f'' and two more to produce a large dispersive signal, $\Delta f'$, more wavelengths contain very little additional phase information but do increase the number of parameters being refined during heavy-atom refinement and phasing procedure. This being the case, repeating the measurement of one wavelength, thus improving the accuracy of the data, and keeping the number of refined parameters lower should be a more desirable approach.

These results demonstrate the importance of maximizing the contrast in f'. On a phase diagram, the centres of phase circles representing two measurements at one wavelength are separated by $2\mathbf{F}''$ and measurements at different wavelengths are separated by $\mathbf{F}'$ (where $\mathbf{F}'$ and $\mathbf{F}''$ are the anomalous contributions to the heavy-atom structure factors). This implies that in order to approach the ideal situation of having the centres of phase circles at the vertices of an equilateral triangle (Phillips, 1978), we must place emphasis on obtaining the largest possible differences in f'.

Experiments of this type clearly require a beamline which is well behaved in terms of its X-ray energy stability. The rate of change in the anomalous scattering factors at the inflection

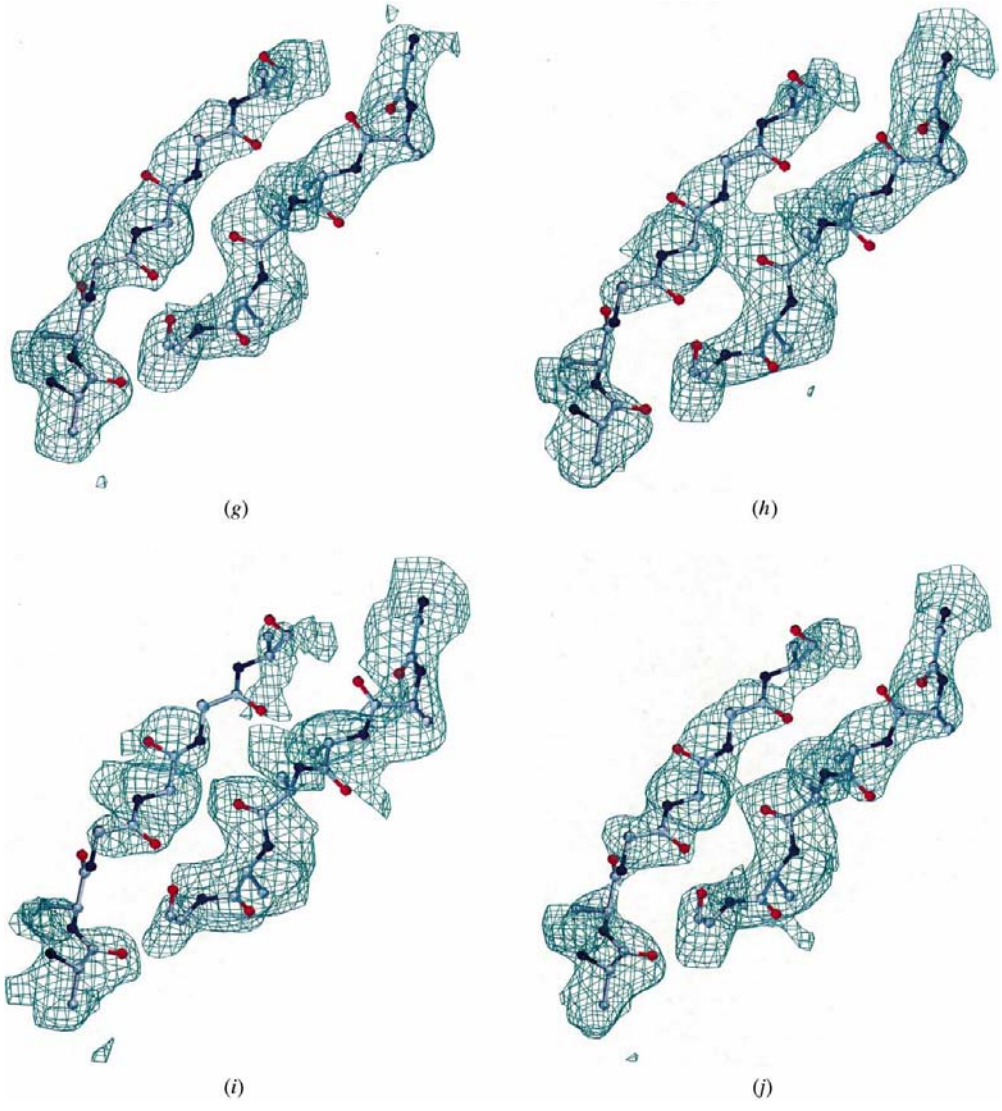

Figure 3 (continued)

points are $\sim$2.5 e (eV)$^{-1}$. A shift of 1 eV could therefore introduce as much as a 20% error into the measured anomalous components of the diffraction intensities. These errors, in turn, can map into gross errors in phase calculations. Wavelength stability can therefore make or break an experiment. The method of aligning the sample crystal and measuring Bijvoet pairs close together in time is clearly one way of reducing the effects of poor energy stability as well as minimizing absorption errors.

These results support the conclusions drawn by Phillips (1978), where he suggested that measuring fewer wavelengths with higher accuracy brought greater benefits than increasing the number of wavelengths used. However, decisions about data collection have to take into account all experimental factors, such as crystal lifetime, detector dynamic range, detector read-out time and beam-time availability, and any number of other unknowns which are an integral part of data collection at any synchrotron light source.

GE thanks the European Molecular Biology Laboratory for a pre-doctoral fellowship.

References

Baker, P., Farrants, G. W., Stillman, T. J., Britton, K. L., Helliwell, J. R. & Rice, D. W. (1990). *Acta Cryst.* A**46**, 721–725.

Blake, C. (1968). *Adv. Protein Chem.* **23**, 59–120.

Blundell, T. & Johnson, L. N. (1976). *Protein Crystallography*, p. 69. London: Academic Press.

Bricogne, G. (1991). *Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 60–68. Warrington: Daresbury Laboratory.

Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.

Cowtan, K. (1994). *Jnt CCP4 ESF-EACBM Newslett. Protein Crystallogr.* **31**, 34–38.

Cromer, D. & Libermann, D. (1970). *J. Chem. Phys.* **53**, 1891–1898.

Evans, G. (1994). PhD thesis, University of Warwick, England.

Evans, G. & Pettifer, R. F. (1996). *Rev. Sci. Instrum.* **67**(10), 3428–3433.

Hendrickson, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.

Hendrickson, W. A. & Teeter, M. M. (1981). *Nature (London)*, **290**, 107–113.

Hoyt, J., de Fontaine, D. & Warburton, W. (1984). *J. Appl. Cryst.* **17**, 344–351.

Kahn, R., Fourme, R., Bosshard, R., Chiadmi, M., Risler, J. L., Dideberg, O. & Wery, J. P. (1985). *FEBS Lett.* **179**(1), 133–137.

Karle, J. (1980). *Int. J. Quantum Chem. Quantum Biol. Symp.* **7**, 357–367.

Lye, R., Phillips, J. C., Kaplan, D., Doniach, S., & Hodgson, K. O. (1980). *Proc. Natl Acad. Sci. USA*, **77**, 5884–5888.

McDonald, N. Q., Panayotatos, N. & Hendrickson, W. A. (1995). *EMBO J.* **14**, 2689–2699.

Main, P., Fiske, S. J., Hul, S. E., Lessinger, G., Germain, G., Declercq, J. P. & Woolfson, M. M. (1980). *MULTAN80: A System of Computer Programs for the Automatic Solution of Crystal Structures from X-ray Diffraction Data.* Universities of York, England, and Louvain, Belgium.

Mukherjee, A., Helliwell, J. & Main, P. (1989). *Acta Cryst.* A**45**, 715–718.

Navia, M. & Sigler, P. (1974). *Acta Cryst.* A**30**, 706–712.

Otwinowski, Z. (1991). *Isomorphous Replacement and Anomalous Scattering*, edited by W. Wolf, P. R. Evans & A. G. W. Leslie, pp. 80−86. Warrington: Daresbury Laboratory.

Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend: Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.

Peterson, M. R., Harrop, S. J., McSweeney, S. M., Leonard, G. A., Thompson, A. W., Hunter, W. N. & Helliwell, J. R. (1996). *J. Synchrotron Rad.* **3**, 24–34.

Phillips, J. C. (1978). PhD thesis, Stanford University, CA, USA.

Phillips, J. C. & Hodgson, K. O. (1980*a*). *Acta Cryst.* A**36**, 856–864.

Phillips, J. C. & Hodgson, K. O. (1980*b*). *Synchrotron Radiation Research*, edited by H. Winick & S. Doniach, pp. 565−605. New York: Plenum Press.

Phillips, J. C., Wlodawer, A., Goodfellow, J. M., Watenpaugh, K. D., Sieker, L. C., Jensen, L. H. & Hodgson, K. O. (1977). *Acta Cryst.* A**33**, 445–455.

Shapiro, L., Fannon, A. M., Kwong, P. D., Thompson, A., Lehman, M. S., Grubel, G., Legrand, J.-F., Als-Nielsen, J., Colman, D. R. & Hendrickson, W. A. (1995). *Nature (London)*, **374**, 327–337.

Weis, W. I., Kahn, R., Fourme, R., Drickamer, K. & Hendrickson, W. A. (1991). *Science*, **254**, 1608–1615.

Wilson, K. S. (1978). *Acta Cryst.* B**34**, 1599–1608.