Acta Cryst. (1995). D51, 39–47

Structure Determination of OppA at 2.3 Å Resolution Using Multiple-Wavelength Anomalous Dispersion Methods

BY IAN D. GLOVER, RICHARD C. DENNY* AND N. DONATUS NGUTI

Department of Physics, Keele University, Keele, Staffordshire ST5 5BG, England

SEAN M. MCSWEENEY, STEVEN H. KINDER AND ANDREW W. THOMPSON†

SERC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, England

AND ELEANOR J. DODSON, ANTHONY J. WILKINSON AND JEREMY R. H. TAME

Department of Chemistry, The University of York, Heslington, York YO1 5DD, England

(*Received* 10 *March* 1994; *accepted* 16 *June* 1994)

Abstract

OppA is a 58.8 kDa bacterial transport protein involved in the transport of peptides across the cytoplasmic membrane of Gram-negative bacteria. It binds peptides from two to five residues in length but with little sequence specificity. OppA from *Salmonella typhimurium* has been cloned and expressed in *E. coli* and the protein cocrystallized with uranyl acetate, producing two distinct crystal forms with different uranium sites. Multiple-wavelength data collected about the uranium L_{III} edge have been collected at the Daresbury Synchrotron Radiation Source (SRS) to a nominal resolution limit of 2.3 Å. Maximum-likelihood phasing methods have been used in phase determination from the multiple-wavelength data giving a readily interpretable electron-density map, without any density modification. The electron-density map, calculated at 2.3 Å resolution shows OppA to be a bilobal, principally β-stranded, three-domain protein. The tri-lysine ligand molecule can be clearly seen in the peptide-binding site between the two lobes.

Introduction

Bacterial cells utilize a variety of nutrients with which they come into contact and have evolved a variety of transport systems to facilitate their uptake. In periplasmic transport systems a binding protein delivers the molecule to be transported to a membrane channel (Ames, 1986). These periplasmic binding proteins are soluble and diffuse freely in the periplasmic space. They bind ligands with high affinity, generally with a K_D of 0.1–1.0 μM. The structures of several binding proteins have been solved, with and without bound ligand, and are reviewed by Quiocho (1991). All the structures solved to date are bilobal two-domain proteins, the ligand being bound in a deep cleft between the domains. Upon binding the ligands become completely buried as the two domains of the binding protein close over them in a mechanism that has been likened to a Venus fly trap. Some binding proteins are highly specific for particular ligands but others will bind a limited variety of ligands; sulfate and phosphate-binding proteins will bind each other's substrate only with very low affinity, but Leu-Ile-Val-binding protein will bind to several different amino acids. Peptides are an important nutrient for bacteria and two periplasmic transport systems are dedicated to them, the oligopeptide and dipeptide permeases Opp and Dpp. There are three peptide transport systems, two of which are well characterized and are periplasmic binding protein dependent. These system have unusually large binding proteins, and the oligopeptide-binding protein, OppA is the largest binding protein known. The dipeptide-binding protein, DppA, is highly selective for dipeptides but OppA will bind peptides up to five residues in length. Both proteins show very little ligand-sequence specificity.

OppA from *Salmonella typhimurium* has been cloned and overexpressed (Higgins, Hardie, Jamieson & Powell, 1983; Hiles & Higgins, 1986) and the protein crystallized (Tolley, Derewenda, Hyde, Higgins & Wilkinson, 1988). The $P2_12_12_1$ form reported by Tolley *et al.* (1988) grown in the absence of uranium proved fragile and difficult to reproduce. In the presence of uranium, however, thicker more stable $P2_12_12$ crystals were produced which diffracted to beyond 2.1 Å resolution (Tame *et al.*, 1994).

* Present address: SERC Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, England.

† Present address: European Synchrotron Radiation Facility, BP 220, F38043 Grenoble, France.

Anomalous Patterson maps showed strong peaks indicating a single uranium site and were designated OppA-U. These crystals prove very sensitive to any heavy-atom soaking experiments in the presence or absence of uranium. Both crystal forms were grown using protein containing the co-purified peptides. More recently ligand-free protein has been prepared and mixed with tri-lysine to form a homogeneous protein–peptide complex. This complex crystallizes in space group $P2_12_12_1$ with similar but slightly larger unit-cell dimensions, and is designated OppA-L. Despite being physically more robust these crystals also proved to be sensitive to even the mildest chemical modification.

The presence of well ordered U and the difficulties in producing other heavy-atom derivatives made this an ideal case for multi-wavelength anomalous dispersion phasing methods using synchrotron radiation. Multiple-wavelength methods may be applied to any anomalous scatterer including metals in metalloproteins or selenium in selenoproteins [see for example, Guss *et al.* (1988); Hendrickson *et al.* (1989)]. The *L* edges of lanthanides or actinides such as uranium give considerably larger anomalous signals than the more commonly used *K* edges of lighter atoms and allow the extension of the method to larger structures.

The size of the protein and the low beam intensity at the short wavelength of the uranium L_{III} edge at 0.72 Å do, however, place additional demands on the experimental method. These include the requirement for data of very high quality and optimization of the anomalous signal by selection of wavelengths giving maximum anomalous contributions (Fig. 1). Phasing was attempted using both the *MADSYS* program suite (Hendrickson, Pahler, Smith, Phizackerley & Merritt, 1988; Hendrickson, 1991) and the standard *CCP*4 phasing program *MLPHARE*. It was found that the level of noise in the data limited the success of classical multiple-wavelength anomalous dispersion (MAD) methods. Maximum-likelihood methods which treat the multiple-wavelength data in a analogous way to the traditional multiple isomorphous replacement method, gave a readily interpretable electron-density map demonstrating the advantages of maximum-likelihood methods in overcoming signal noise.

Experimental

OppA was crystallized using the hanging-drop vapour-diffusion method from 15% PEG 4000/ 50 m*M* sodium acetate, pH 5.5 in the presence of 1 m*M* uranyl acetate and a slight molar excess of tri-lysine. Two crystal forms appear under these conditions (Tame *et al.*, 1994). OppA-L crystals, space group $P2_12_12_1$ with unit-cell dimensions a = 110.42, b = 77.04, c = 71.38 Å, grow typically to a size of 0.15 × 0.15 × 0.15 mm. Data were collected from five larger crystals of dimensions 0.3 × 0.2 × 0.2 mm. This was the crystal form used in this experiment. The other crystal form, OppA-U, has space group $P2_12_12$ with cell dimensions a = 1065, b = 74.5, c = 70.0 Å. The OppA-L crystal form contains up to eight U atoms per asymmetric unit, see below, whereas the OppA-U form contains only one well occupied site.

All data collection was carried out on station 9.5 at the Daresbury SRS (Thompson *et al.*, 1992) using radiation from a superconducting wiggler in the

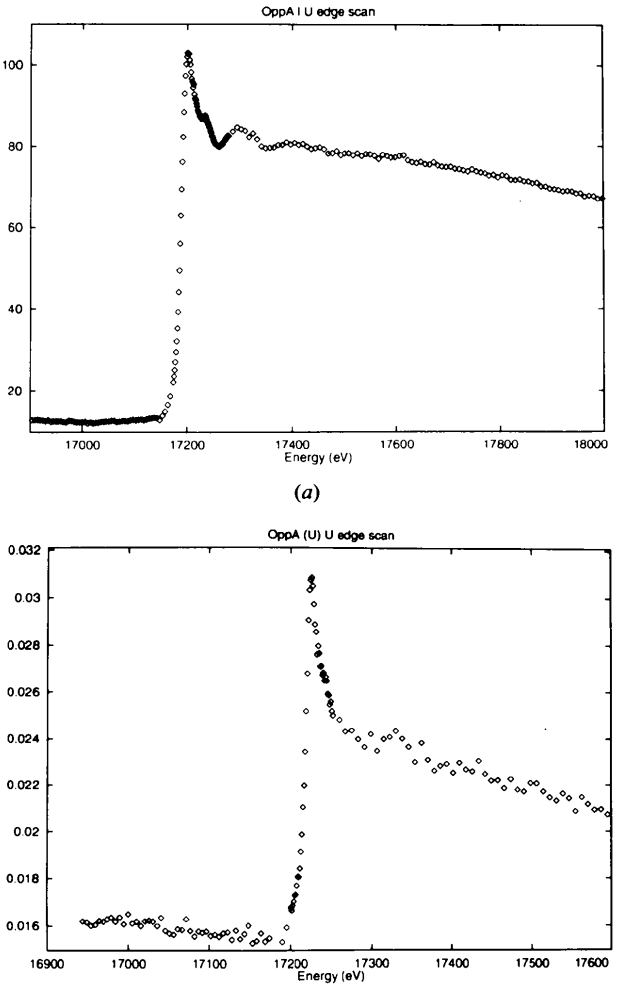


Fig. 1. (*a*) Fluorescence XANES spectrum from a single crystal of OppA-L (dimensions 0.3 × 0.2 × 0.2 mm) recorded using a single-wire proportional counter *in situ* on the diffractometer prior to data collection. Data were collected at nominal incident X-ray energies of 17.224 (λ_1), 17.243 (λ_2), 18.049 (λ_3) and 13.777 KeV (λ_4). (*b*) Fluoresence XANES spectrum recorded as for (*a*) from a crystal of the single uranium site crystal form, OppA-U, showing significantly enhanced white-line features.

storage ring operated at 2 GeV with circulating currents of approximately 250–150 mA. Monochromatic X-rays were produced by a channel-cut $Si^{[111]}$ monochromator with cooling on the first crystal downstream from a Pt-coated toroidal mirror providing point focused radiation at the sample. Initial calibration of X-ray energy was carried out with reference to the points of inflection in absorption spectra of Zn foil and uranium acetate powder samples. The energy resolution of the monochromator was controlled using slits giving a calculated energy bandpass of 8.7×10^{-4} ($\delta E/E$). The uranium absorption-edge position and profile was characterized from the X-ray absorption near-edge structure (XANES) spectra from a single crystal of OppA recorded using a single-wire proportional counter mounted on the goniostat. The XANES spectra were recorded immediately prior to data collection and at the start of each machine fill and the monochromator positions were updated accordingly.

XANES spectra recorded from crystals of OppA-U and OppA-L are shown in Fig. 1. The edge position shows a shift of 5 eV from uranium acetate demonstrating the importance of recording XANES spectra from the protein crystal itself rather than relying on model compounds for wavelength selection. Prominent white-line features were evident, allowing enhanced f'' components to be exploited in data collection. The X-ray energies at which data were collected were determined directly from the XANES spectra recorded immediately prior to data collection. λ_1 corresponds to the point of inflection of the edge (f' minimum), λ_2 to the white-line position (f'' maximum) and λ_3 chosen to give a calculated f' value half that at λ_1, allowing the collection of data with large dispersive differences whilst avoiding significant beam position movements due to large changes in monochromator angle. A fourth data set was subsequently collected at a remote wavelength of 0.9 Å, a wavelength at which the incident beam flux is greater and gives significant f'' and f' contributions. The XANES spectra may be transformed to yield experimental values of f' and f'', taking parameters such as energy bandpass and the presence of white-line features into account (Table 1) and used in wavelength selection. No evidence of absorption edge anisotropy was seen in the XANES spectra from OppA-L crystals, probably due to the presence of several U sites in the crystal. The spectra recorded from OppA-L and OppA-U do show significant differences, for OppA-U the white-line feature is significantly more pronounced, even though the total signal is reduced.

Diffraction data were recorded using a MAR Research image-plate detector system with a crystal-to-detector distance of 175 mm giving a nominal resolution limit of 2.3 Å. Data-collection protocols aimed at reducing the effect of systematic error were adopted, collecting all observations contributing to an individual phase determination as close together in time as possible. Crystals were freshly mounted, with b^* parallel to the rotation axis, and carefully aligned to enable the recording Bijvoet mates on the same image. Data were collected in single φ intervals at each of three wavelengths (0.7199, 0.7191 and 0.687 Å) before proceeding to the next φ interval. Oscillation ranges between 1.2 and 2° were employed, depending on crystal orientation and the need to avoid overlapping reflections, using exposure times between 75 and 200 seconds per degree. Two crystals were used to record an overlapping total of 107° at each of the three wavelengths followed by the collection of a complete data set at 0.9 Å wavelength from a third crystal. All data were collected at room temperature over a period of 5 days.

Data analysis

The data were processed using the *MOSFLM* suite of programs, modified to take account of the extended dynamic range of image-plate data. Lorentz and polarization corrections were carried out using the program *ABSCALE* (*CCP*4, SERC Daresbury Laboratory, 1979). No absorption corrections were applied to the data due to the short wavelengths employed. Scaling and merging of the profile-fitted integrated intensities was performed using the Fox & Holmes algorithm (1966) as implemented in the *CCP*4 programs *ROTAVATA* and *AGROVATA*.

The data processing is summarized in Table 1, showing the merging statistics for the data along with their calculated f' and f'' values. The average multiplicity of the data is 5.3. The merging R factors show that the data is of reasonable quality but the expected anomalous signal is of the same order or less than the reproducibility of the data (Table 2). The agreement between reflections recorded at different wavelengths is, however, better than may be expected from R_{merge} values given in Table 1, illustrating the advantage of the data-collection protocols used. (See R factor between data sets in Table 2.)

When using MAD data sets, where the signal is very low compared to the standard deviation of the observations, it is necessary to take particular care to minimize error in differences between the structure factors measured at different wavelengths (dispersive differences) and between $F(+)$ and $F(-)$, the anomalous differences.

The usual approach is to try to minimize random errors by careful and repeated measurements, and to try to guarantee that each set of observations to be used for phasing has the same systematic error. This

Table 1. *Summary of OppA-L data-processing statistics in space group $P2_12_12_1$ and for comparison in space group P1*

The anomalous scattering coefficients, f' and f'', derived from the transformed XANES spectrum for λ_1 and λ_2, values for other remote wavelengths calculated from theoretical cross sections (Cromer, 1983).

Data set	λ (Å)	Resolution (Å)	R_{merge}* (%)	R_{ano}† (%)	R_{merge} (no anom) (%)	%Complete	Multiplicity	N_{uni}	f'	f''
(*a*) Space group $P2_12_12_1$										
λ_1	0.7199	2.3	9.3	7.0	11.3	97.8	5.7	26625	−20.9	8.92
λ_2	0.7191	2.3	9.6	7.0	11.6	97.6	5.7	26544	−15.49	11.41
λ_3	0.687	2.2	9.7	6.8	11.5	96.6	5.0	32013	−7.74	9.16
λ_4	0.9000	2.3	10.8	5.2	11.5	96.2	7.68	26111	−5.50	5.84
(*b*) Space group *P*1										
λ_1	0.7199	2.3	9.6	9.3	10.3	75.8	2.0	78001	−20.9	8.92
λ_2	0.7191	2.3	10.2	9.1	10.6	75.1	2.0	77280	−15.49	11.41
λ_3	0.6870	2.2	10.9	8.9	10.4	82.7	1.5	104122	−7.74	9.16
λ_4	0.9000	2.3	9.9	7.9	10.4	77.2	2.2	79476	−5.50	5.84

* $R_{merge} = \sum_{hkl}|I - \langle I\rangle|/\sum_{hkl}\langle I\rangle$.
† $R_{ano} = \sum_{hkl}|I(+) - I(-)|/\sum_{hkl}\langle I\rangle$.

may be achieved firstly by careful design of the experiment; the same reflections should be measured from the same crystal, in the same orientation, and nearly at the same time. In addition applying local scaling can help to remove systematic errors. This requires that the data sets are not merged before local scales are calculated and that the merging appropriate to the space group is carried out after phasing. The data sets used here had high multiplicity which meant that the error in the average structure factor, $\langle F\rangle$, should be reduced by the square root of (multiplicity − 1).

To test whether it was better in this case to merge symmetry-equivalent data or not, we prepared two data sets, one without merging symmetry equivalents, the '*P*1' data set and the other with merging. Data-processing statistics for the $P2_12_12_1$ and the *P*1 data sets are given in Table 1.

The statistics for the observed and calculated dispersive and anomalous differences are given in Table 2. The mean observed anomalous differences for the $P2_12_12_1$ data sets are 50% larger than the calculated values and the mean observed dispersive differences between the $P2_12_12_1$ data sets at different wavelengths are about three times as large as the calculated values. For the *P*1 data set, the anomalous differences are about twice the expected value and the mean observed dispersive differences about four times as large as the calculated values. The improved agreement for the $P2_12_12_1$ data demonstrates the accuracy with which the anomalous differences have been measured. The approach of merging first is clearly more appropriate in this case. This approach may not always be best, particularly in the case of large systematic errors, such as that caused by absorption. Absorption is not a systematic source of error at the X-ray wavelength used in this experiment.

All the subsequent phase calculations require that the anomalous-scattering sites first be identified to a reasonable degree of accuracy from Patterson or direct methods. Using data collected at multiple wavelengths a variety of coefficients may be used in the location of anomalous scatterers. Identification of heavy-atom sites was carried out using Pattersons based on the anomalous differences $[(F+) - (F-)]$, dipersive differences $(F\lambda_i - F\lambda_j)$ and combined $\sum(\Delta F\lambda)^2 + \sum(\Delta F+/-)^2$ Patterson syntheses. The Harker sections in the anomalous and dispersive difference Patterson maps clearly indicated a number of uranium sites but were complicated by the fact that the most significant features were cross vectors between sites with common *x*, *y* or *z* coordinates.

Several sites were previously identified from difference Pattersons calculated from data sets collected at York. The multi-wavelength data gave very high quality anomalous Pattersons, reflecting their overall quality. The quality of the anomalous difference Pattersons for data collected at λ_1 and λ_2 was especially high, and the *SHELX*90 (Sheldrick, 1990) Patterson search procedure identified seven sites. The cross and self vectors between the eight uranium sites in the final, refined, model are plotted on two of the sections in Fig. 2.

Results

Phasing

Attempts at phasing the diffraction data were carried out using two approaches, the approach of Hendrickson (1985, 1991) as adapted from the formalism of Karle (1980) in the program *MADLSQ* and associated software, and the *CCP*4 program *MLPHARE* which exploits maximum-likelihood methods in phase calculation (Otwinowski, 1991; Bricogne, 1991).

Table 2. *The agreement between equivalent reflections recorded at different wavelengths in the OppA-L data sets*

The agreement is shown in space groups $P2_12_12_1$ and $P1$. The poorer agreement between λ_4 and the other data sets reflects the fact that this data set was collected with different crystals.

Data set	λ_1	λ_2	λ_3	λ_4
(*a*) Space group $P2_12_12_1$				
λ_1				
Mean dispersive difference	—	13.1	20.3	39.3
Mean anomalous difference	34.4	34.3	31.6	25.3
R factor between data sets (%)	—	3.5	5.4	10.5
No. of reflections	25642	25553	25512	25086
Calculated dispersive difference*	0.0	4.3	9.7	19.7
Calculated anomalous difference†	23.7	23.9	19.9	12.4
λ_2				
Mean dispersive difference		—	18.0	37.8
Mean anomalous difference		34.3	31.6	25.3
R factor between data sets (%)		—	4.8	10.1
No. of reflections		25561	25468	25086
Calculated dispersive difference*		4.3	9.7	19.7
Calculated anomalous difference†		23.9	19.9	12.4
λ_3				
Mean dispersive difference			0	29.7
Mean anomalous difference			31.6	25.3
R factor between data sets (%)				9.5
No. of reflections			25981	25052
Calculated dispersive difference*			9.7	19.7
Calculated anomalous difference†			19.9	12.4
λ_4				
Mean dispersive difference				
Mean anomalous difference				25.3
R factor between data sets (%)				
No. of reflections				25113
Calculated dispersive difference*				19.7
Calculated anomalous difference†				12.4
(*b*) Space group $P1$				
λ_1				
Mean dispersive difference		20.0	29.3	50.7
Mean anomalous difference	46.2	45.8	44.4	41.8
R factor between data sets (%)	—	5.3	7.8	13.7
No. of reflections	77995	76891	71080	57549
λ_2				
Mean dispersive difference			27.7	50.3
Mean anomalous difference		45.8	44.4	41.8
R factor between data sets (%)			7.4	13.6
No. of reflections		77279	70671	56979
λ_3				
Mean dispersive difference				45.0
Mean anomalous difference			43.8	43.8
R factor between data sets (%)			—	12.1
No. of reflections			100035	70420
λ_4				
Mean dispersive difference				
Mean anomalous difference				41.3
R factor between data sets (%)				
No. of reflections				79471

* $F_H'(\lambda_i) - F_H'(\lambda_j)$.
† $F_H''(\lambda_i)$.

The first stage in the MAD phase determination following the method of Hendrickson is the correct scaling of data collected at different wavelengths, as carried out by the program *WVLSCL*. Small differences in intensity due to anomalous scattering must be preserved whilst placing the data on the correct relative scale. The pattern of average Bijvoet and dispersive differences calculated from the scaled data follows that expected from a comparison of theoretical values (Table 3). These diffraction ratios are, however, systematically overestimated and values calculated from the centric data, expected to be free of anomalous effects, reflect the level of noise in the data. These values are particularly high in the high-resolution shells where noise constitutes approximately half of the anomalous signal in the data.

The *WVLSCL*-scaled data were then used in the calculation of the normal scattering component of the anomalously scattering atoms, ${}_0F_A$, using the program *MADLSQ*. This procedure was applied to sets including all data, data from separate crystals and three and four wavlength data. In all cases the calculated ${}_0F_A$ values showed systematically high r.m.s. values compared to the total normal scattering and a large number of outliers in the intensity distributions. Patterson maps calculated with these derived ${}_0F_A$ values were noisy and not interpretable in terms of a consistent set of anomalous scattering sites despite a series of stringent rejection criteria based on absolute magnitude, reliability (σ and Q) and resolution range. The inability to determine the anomalous scattering partial structure using these approaches precluded any further phase calculation using this method. Anomalous-scattering factors refined during this procedure, however, showed good agreement with those calculated from the transformed XANES spectra.

MLPHARE

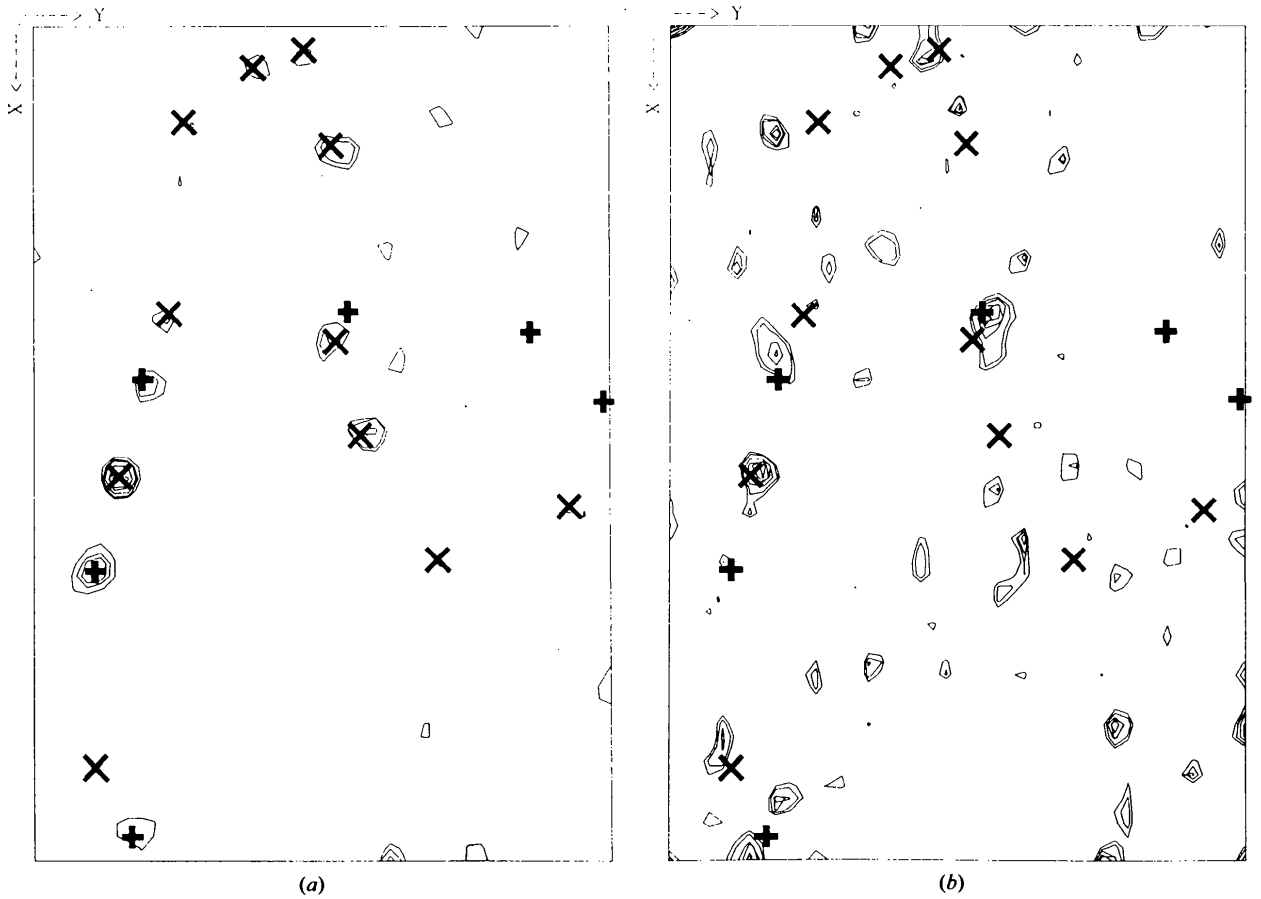
The *CCP*4 programs *CAD* and *SCALEIT* were used to combine data and place them on the same relative scale. *SCALEIT* also estimates the largest acceptable anomalous and isomorphous intensity differences. Larger differences were rejected as outliers due to experimental error. These rejections were important due to the sensitivity of Patterson methods and *SHELX*90 to erroneous large intensity terms.

The use of *MLPHARE* in phasing from multi-wavelength data is also discussed briefly by Ramakrishnan, Finch, Gravaziano, Lee & Sweet (1993) who solved the structure of seleno-methionyl histone H5. This method treats the data sets collected at different wavelengths as conventional isomorphous heavy-atom derivatives, one data set being chosen as the 'native' data set (with anomalous scattering). Dispersive differences between data sets collected at different wavelengths give rise to

'isomorphous' differences which are refined as apparent real occupancies of the anomalous scatterers. For convenience the data set with the maximum $|f''|$ was used as the native data set to maintain a consistent positive dispersive difference between all data sets. In other cases it is preferable to use the most complete data set as the native and adjust the sign of the dispersive differences appropriately.

Initial estimates of the positions of the anomalous scatterers were refined in iterative cycles of phase refinement. The coordinates (x, y and z), temperature factor, real (dispersive) and anomalous (absorptive) occupancies for each site may each be refined or fixed. Phase probability weighted by the maximum likelihood is integrated at set intervals round the phase circle heavy-atom contributions estimated for the variation in the structure factors measured at each wavelength. It is more robust than previous methods of phasing and very effective in screening out the signal from the background noise in the measurements. It refines incorrect heavy-atom or anomalous scatterer sites to very low occupancies and is less likely to become trapped in false minima. Real occupancies and coordinates may be refined first using centric data and the anomalous occupancies and coordinates can subsequently be refined using all data. The real occupancies of the anomalous scatterer in the native, reference data set are fixed at zero throughout this procedure.

A special, unitary scattering factor was added to the form factor list, with f' and $f'' = 1.00$ electron. This meant that the real and anomalous occupancies were proportional to the number of electrons constituting $(F'\lambda_i - F'\lambda_j)$ and $(F''\lambda_i - F''\lambda_i)$. The results are shown in Table 4, in which the occupancies have been corrected to correspond to structure factors which are on an approximately absolute scale. In the course of the refinement there is no means of distinguishing the sign of either the real or anomalous origin. A knowledge of the wavelengths used allows



(*a*) (*b*)

Fig. 2. Harker sections, ($w = 1/2$), of the Patterson maps calculated with coefficients (*a*) anomalous differences recorded at a wavelength of 0.719 Å, corresponding to the white-line feature observed in XANES spectrum of OppA-L and (*b*) dispersive differences between data sets λ_1 and λ_2. The Harker vectors corresponding to self vectors (+) and cross vectors (×) resulting from the Patterson search procedure of *SHELX*90 are superimposed.

Table 3. *Anomalous diffraction ratios (Hendrickson et al., 1985)*

(*a*) Bijovet difference ratios [r.m.s. $(\Delta F+/-)$/r.m.s. $(|F|)$] are given as diagonal elements; ratios calculated with centric data only are given in parentheses. Dispersive difference ratios [r.m.s. $(F\lambda_i - F\lambda_j)$/r.m.s. $(|F|)$] are the off-diagonal terms. The results are presented for all data, in the 200–4.5 Å resolution shell and 2.95–2.5 Å shell and shown for calculated, expected values on the basis of the f' and f'' components at each wavelength given in Table 1.

Observed ratios, all data.

λ(Å)	0.900	0.7199	0.7191	0.687
0.900	0.067 (0.042)	0.117	0.115	0.130
0.7177		0.110 (0.055)	0.061	0.093
0.7199			0.113	0.076 (0.057)
0.687				0.106 (0.059)

Observed ratios, resolution shell 200–4.5 Å

λ (Å)	0.900	0.7911	0.7191	0.687
0.900	0.062 (0.033)	0.104	0.102	0.103
0.7199		0.090	0.052 (0.046)	0.064
0.7191			0.095 (0.48)	0.060
0.0687				0.085 (0.047)

Observed ratios, resolution shell 2.9–2.5 Å

λ (Å)	0.900	0.7199	0.7191	0.687
0.900	0.083 (0.058)	0.134	0.132	0.143
0.7199		0.126 (0.075)	0.085	0.130
0.7191			0.131	0.087 (0.078)
0.687				0.122 (0.085)

(*b*) Theoretical values of difference and dispersive diffraction ratios calculated for a single uranium site and eight uranium sites, the latter numbers given in parentheses.

Theoretical values, 1 site and (8 sites)

λ (Å)	0.900	0.7199	0.7191	0.687
0.900	0.019 (0.053)	0.025 (0.069)	0.016 (0.045)	0.004 (0.010)
0.7199		0.28 (0.080)	0.009 (0.024)	0.021 (0.059)
0.7191			0.036 (0.102)	0.012 (0.035)
0.687				0.029 (0.082)

Table 4. *The refined real (R) and anomalous (A) occupancies and temperature factors (B values) of the eight uranium sites identified using Patterson and Patterson search methods*

The occupancies were refined using unitary form factors and reflect the number of electrons constituting the dispersive differences between data collected at different wavelengths and contributing to the anomalous differences using data on an approximately absolute scale.

Data set	λ_1	λ_2	λ_3	λ_4
Site 1				
Occupancy (*R*)	0.000	2.19	4.35	9.53
Occupancy (*A*)	7.89	7.99	7.00	4.51
B value (Å^3)	19.08	19.85	23.4	22.6
Site 2				
Occupancy (*R*)	0.00	1.83	4.86	8.96
Occupancy (*A*)	7.84	7.83	6.96	4.65
B value (Å^2)	23.9	24.4	26.4	27.2
Site 3				
Occupancy (*R*)	0.00	1.72	3.85	9.81
Occupancy (*A*)	6.48	6.53	5.94	3.66
B value (Å^2)	25.5	26.0	28.5	27.9
Site 4				
Occupancy (*R*)	0.00	1.28	4.58	8.72
Occupancy (*A*)	6.96	7.29	6.61	4.12
B value (Å^2)	20.8	23.1	25.5	25.8
Site 5				
Occupancy (*R*)	0.00	1.50	3.90	7.01
Occupancy (*A*)	7.30	7.39	6.69	4.15
B value (Å$_2$)	21.7	22.7	26.0	24.7
Site 6				
Occupancy (*R*)	0.00	1.62	3.83	7.69
Occupancy (*A*)	4.39	4.60	3.86	2.53
B value (Å^2)	25.3	28.3	26.3	28.7
Site 7				
Occupancy (*R*)	0.00	1.34	2.29	3.72
Occupancy (*A*)	2.97	3.10	2.73	−0.25
B value (Å^2)	20.8	23.1	23.5	30.5
Site 8				
Occupancy (*R*)	0.00	1.41	2.99	7.24
Occupancy (*A*)	3.81	3.81	3.49	2.02
B value (Å^2)	26.1	25.9	31.8	29.2

the correct choice of the sign of the real occupancies but it is still necessary to calculate two electron-density maps, on both hands of the anomalous occupancy and determine which is correct.

The electron-density map calculated at 2.3 Å using the *MLPHARE* phases was of high quality and readily interpretable in terms of the known sequence of OppA. The protein main- and side-chain atoms could be readily traced and the tri-lysine substrate identified and located. The correlation of the *MLPHARE*-phased electron-density map with F_{calc} maps calculated from the model refined against the λ_1 data (*R* factor = 19.74% for all data in the 10–2.3 Å resolution shell) is extremely high, 0.6067, indicating the very high quality of the initial *MLPHARE*-phased map as shown in Fig. 3.

MLPHARE phases derived from MAD data are ideal for phase refinement using density-modification procedures. The well phased reflections are distributed over all resolution shells and poorly phased reflections are assigned sensible, low weights. In this case the *MLPHARE* phases were subjected to phase refinement using density-modification procedures as

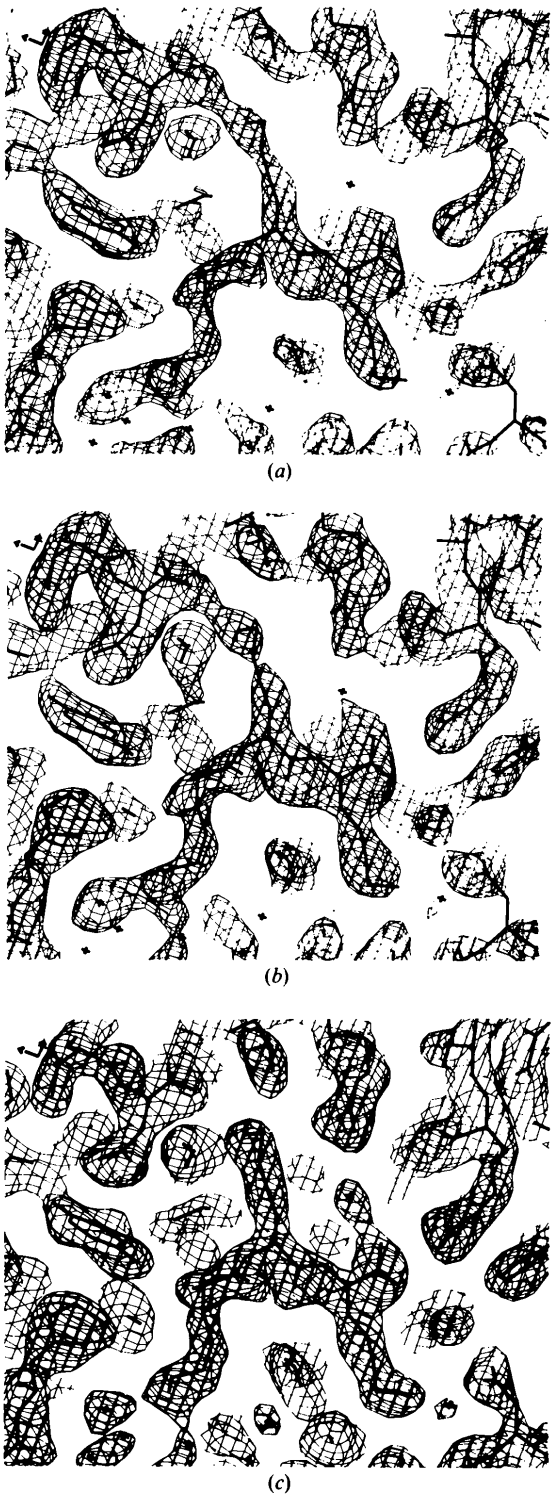


Fig. 3. A portion of the electron-density map calculated at 2.3 Å resolution in the region of the bound tri-lysine ligand with the coordinates resulting from the 2.3 Å resolution refinement of the structure against the λ_1 data. The maps were calculated in (*a*) with the phases resulting from *MLPHARE*, (*b*) maximum likelihood phases after density modification and (*c*) $2F_o - F_c$ map at the current state of refinement ($R = 19.74\%$).

implemented in the program *SQUASH*. The results of the density modification are summarized in Fig. 4. The resulting correlation between the *MLPHARE* and modified maps was 0.8912, and the correlation between the modified and F_{calc} maps calculated from the refined model rose to 0.7038. Details of the refined stucture have been published separately (Tame *et al.*, 1994).*

Discussion

Multiple-wavelength anomalous dispersion has long been established as a useful method of phase determination in macromolecular crystallography. The

* Atomic coordinates and structure factors have been deposited with the Protein Data Bank, Brookhaven National Laboratory (Reference: 1OLB, R1OLBSF). Free copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England (Reference: LI0177).

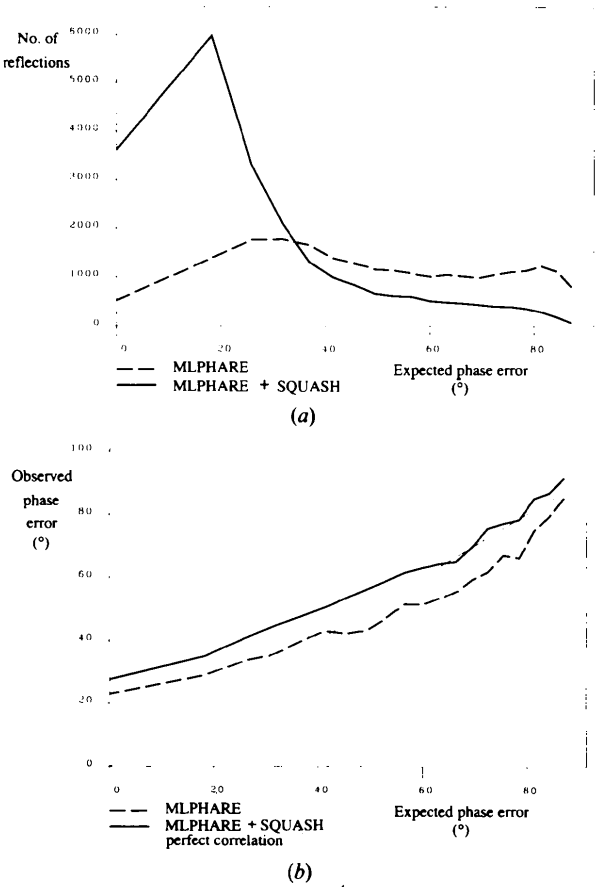


Fig. 4. Plots of (*a*) the expected phase error as a function of number of phased reflections and (*b*) observed compared to expected phase error for structure determination of OppA using *MLPHARE* and *MLPHARE* combined with density modification using *SQUASH*.

method places considerable demands on instrumentation, particularly monochromator systems and data quality. Whilst care should always be taken to minimize systematic error, noise is inevitable in experimental data. In the case of OppA the diffraction ratios calculated in *WVLSCL*, while reflecting the expected pattern of values, show that the anomalous contributions are overestimated and the noise level half that of the signal. This clearly hampers the performance of classical multiple-wavelength analysis, where global parameters are refined on a per-reflection basis. Here the analysis and calculations, though apparently successful do not allow the calculation of an unambiguous Patterson synthesis, location of the anomalous scatterers or subsequently the calculation of an electron-density map.

The maximum-likelihood method, however, treats heavy-atom parameters as global, and therefore seems less susceptible to the level of noise in the data. Heavy-atom refinement proceeded smoothly in *MLPHARE*, yielding phases accurate enough for an excellent electron-density map to be calculated directly. The method is also extremely simple and relatively rapid. Initial attempts at phasing with *MLPHARE* allowed a good clearly interpretable electron-density map to be calculated within 3 h of the identification of the anomalous-scattering sites. OppA is considerably larger than the majority of structures phased by MAD methods and it is, therefore, understandable that the noise level in the anomalous data is high. We suggest that the maximum-likelihood methods are more appropriate for the treatment of multiple-wavelength anomalous-dispersion data in such cases.

We would like to thank the Universities of Keele and York for their support and the staff at the Daresbury Laboratory for the provision of synchrotron radiation. This work was funded in part by and SERC project grant and the provision of a BAP small grant of synchrotron beam time. IDG and JT would also like to thank Dr Rose Todd for introducing them to each other.

References

AMES, G. F.-L. (1986). *Ann. Rev. Biochem.* **55**, 397–425.
BRICOGNE, G. (1991). *Isomorphous Replacement and Anomalous Scattering*, edited by W. WOLF, P. R. EVANS & A. G. W. LESLIE, pp. 60–68. Warrington: SERC Daresbury Laboratory.
BRÜNGER, A. T. (1988). *J. Mol. Biol.* **203**, 803–816.
CROMER, D. T. (1983). *J. Appl. Cryst.* **16**, 437.
FOX, G. C. & HOLMES, K. C. (1966). *Acta Cryst.* **20**, 886–891.
GUSS, J. M., MERRITT, E. A., PHIZACKERLEY, R. P., HEDMAN, B., MURATA, M., HODGSON, K. O. & FREEMAN, H. C. (1988). *Science*, **241**, 806–811.
HENDRICKSON, W. A. (1985). *Trans. Am. Crystallogr. Assoc.* **21**, 11–21.
HENDRICKSON, W. A., SMITH, J. L., PHIZACKERLEY, R. P. & MERRIT, E. A. (1988). *Proteins*, **4**, 77–78.
HENDRICKSON, W. A., PAHLER, A., SMITH, J. L., SATOW, Y., MERRITT, E. A. & PHIZACKERLEY, R. P. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190–2194.
HENDRICKSON, W. A. (1991). *Science*, **254**, 51–58.
HILES, I. D. & HIGGINS, C. F. (1986). *Eur. J. Biochem.* **158**, 561–567.
HIGGINS, C. F., HARDIE, M. M., JAMIESON, D. J. & POWELL, L. M. (1983). *J. Bacteriol.* **153**, 830–836.
JONES, T. A., ZOU, J.-Y., COWAN, S. W. & KJELDGAARD, M. (1991). *Acta Cryst.* A**47**, 110–119.
KARLE, J. (1980). *Int. J. Quantum Chem.* **7**, 357–367.
OTWINOWSKI, Z. (1991). *Isomorphous Replacement and Anomalous Scattering*, edited by W. WOLF, P. R. EVANS & A. G. W. LESLIE, pp. 80–86. Warrington: SERC Daresbury Laboratory.
QUIOCHO, F. A. (1991). *Curr. Op. Struct. Biol.* **1**, 922–933.
RAMAKRISHNAN, V., FINCH, J. T., GRAVAZIANO, C., LEE, P. L. & SWEET, R. M. (1993). *Nature (London)*, **362**, 219–223.
SERC Daresbury Laboratory (1979). *CCP4. A Suite of Programs for Protein Crystallography*. SERC Daresbury Laboratory, Daresbury, Warrington WA4 4AD, England.
SHELDRICK, G. M. (1990). *Acta Cryst.* A**46**, 467–473.
TAME, J. R. H., MURSHUDOV, G. N., DODSON, E. J., NEIL, T. K., DODSON, G. G., HIGGINS, C. F. & WILKINSON, A. J. (1994). *Science*, **264**, 1578–1581.
THOMPSON, A. W., HABASH, J., HARROP, S., HELLIWELL, J. R., NAVE, C., ATKINSON, P., HASNAIN, S. S., GLOVER, I. D., MOORE, P. R., HARRIS, N., KINDER, S. & BUFFEY, S. (1992). *Rev. Sci. Instrum.* **63**, 1062–1064.
TOLLEY, S. P., DEREWWNDA, Z., HYDE, S. C., HIGGINS, C. F. & WILKINSON, A. J. (1988). *J. Mol. Biol.* **204**, 493–494.