Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Multiwavelength anomalous solvent contrast (MASC): derivation of envelope structure-factor amplitudes and comparison with model values

M. Ramin,[a]*† W. Shepard,[a] R. Fourme[a] and R. Kahn[b]

[a]LURE (CNRS, CEA, MENRT), Université Paris-Sud, 91405 Orsay CEDEX, France, and [b]IBS J.-P. Ebel, 41 Avenue des Martyrs, 38027 Grenoble CEDEX, France

† Current address: Université d'Evry-Val-d'Essonne, Boulevard François Mitterand, 91025 Evry CEDEX, France.

Correspondence e-mail: ramin@lure.u-psud.fr

Received 22 January 1998
Accepted 30 April 1998

A previous article [Fourme *et al.* (1995). *J. Synchrotron Rad.* **2**, 36–48] presented the theoretical foundations of MASC, a new contrast-variation method using multiwavelength anomalous scattering, and reported the first experimental results. New experiments have been conducted both at the ESRF (Grenoble, France) and at LURE-DCI (Orsay, France), using cryocooled crystals of three proteins of known structures and very different molecular weights. Amplitudes of $\{\Gamma_T(\mathbf{h})\}$, the 'normal' structure factors of the anomalously scattering part of the crystal including the solvent zone and the ordered anomalous scattering sites (if any), have been extracted from multiwavelength data. In the very low resolution range ($d \geq 20$ Å), the agreement between experimental $\{|\Gamma_T(\mathbf{h})|\}$ and model values calculated from the bulk solvent is all the more satisfactory since the molecular weight of the protein is high. For spacings between 10 and 20 Å, the agreement between experimental $\{|\Gamma_T(\mathbf{h})|\}$ and model values is also satisfactory if one takes into account ordered anomalous scatterer sites. Such sites have been found in the three cases.

1. Introduction

The knowledge of the envelope of a biological macromolecule is a crucial step in the determination of its three-dimensional structure. The relevant information is essentially provided by low-resolution reflections, for which the completeness is generally poor in standard biocrystallography experiments. Contrast-variation methods in neutron or X-ray crystallography represent an exception to this situation. Experiments with chemical contrast variation, using X-rays (Dumas, 1988; Carter *et al.*, 1990) or isotopical contrast variation, using neutrons (Moras *et al.*, 1983; Roth *et al.*, 1984; Bentley *et al.*, 1984; Podjarny *et al.*, 1987; Timmins *et al.*, 1992; Pebay-Peroula *et al.*, 1995) have been conducted. With X-rays, the main limitation in chemical contrast variation comes from the fact that several samples are used with different solvent contents that may induce structural changes in the macromolecule leading to non-isomorphism.

The idea of combining multiwavelength anomalous scattering and contrast-variation techniques was introduced in a previous article (Fourme *et al.*, 1995), presenting the theoretical principles of a new method called MASC (an acronym suggested by W. Shepard for multiwavelength anomalous solvent contrast). Some experimental results obtained with this new technique were also presented. MASC is a physical contrast-variation method in which the change in electronic density is obtained by tuning the X-ray wavelength near one of the absorption edges of an anomalous scattering species

randomly dispersed inside solvent channels of the macromolecular single crystal (see also Fourme *et al.*, 1996). In MASC, the anomalous signals, quantified by Bijvoet and dispersive ratios are large at low resolution (d spacings $\geq$ 8–10 Å) and decrease rapidly with increasing resolution. These signals are optimized for a concentration of anomalous scatterers such that the electronic density in the solvent is equal to the mean electronic density inside the protein (the contrast-matching point).

The algebraic approach developed initially for MAD (Hendrickson, 1985) can also be applied to MASC (Fourme *et al.*, 1995). The basic equations are,

$$\begin{aligned} |{}^{\lambda}F(\pm\mathbf{h})|^2 = {} & |{}^{\circ}F(\mathbf{h})|^2 + a(\lambda)|\Gamma(\mathbf{h})|^2 \\ & + b(\lambda)|{}^{\circ}F(\mathbf{h})\parallel\Gamma(\mathbf{h})|\cos(\Delta\varphi) \\ & \pm c(\lambda)|{}^{\circ}F(\mathbf{h})\parallel\Gamma(\mathbf{h})|\sin(\Delta\varphi), \end{aligned} \quad (1)$$

where ${}^{\lambda}F(\pm\mathbf{h})$ is the total (wavelength-dependent) structure factor and ${}^{\circ}F(\mathbf{h})$ is the total (wavelength-independent) normal structure factor, and $\Gamma(\mathbf{h}) = -{}^{0}\rho_{\rm sA}G(\mathbf{h})\exp(-Bs^2/4)$ is the normal component of the structure factor of the disordered anomalous scatterers in the bulk solvent. ${}^{0}\rho_{\rm sA}$ is the normal electronic density, assumed to be uniform, of anomalous scatterers in the solvent domain, and the effect of the disorder in this phase is described by a Debye–Waller-like factor B. $G(\mathbf{h})$ is the structure-factor value at $\mathbf{h}$ of the indicator function $\chi(\mathbf{r})$ (Bricogne, 1974) which takes the value of one inside the ordered phase of the crystal (macromolecule and ordered water molecules) and zero in the disordered phase (solvent), thus delineating the envelope of the macromolecule. $\Delta\varphi$ is the difference between the phases of ${}^{\circ}F(\mathbf{h})$ and $\Gamma(\mathbf{h})$.

The scattering factor of the anomalous scattering species is expressed as

$${}^{\lambda}f = {}^{\circ}f + {}^{\lambda}f' + 1{}^{\lambda}f'',$$

and a, b and c encode the whole wavelength-dependence with

$$a(\lambda) = ({}^{\lambda}f'^2 + {}^{\lambda}f''^2)/{}^{\circ}f^2,$$
$$b(\lambda) = 2{}^{\lambda}f'/{}^{\circ}f \text{ and } c(\lambda) = 2{}^{\lambda}f''/{}^{\circ}f.$$

Since equation (1) has the same form as MAD equations, the program *MADLSQ* (Hendrickson, 1985) can be readily used to obtain $|{}^{\circ}F(\mathbf{h})|$, $|\Gamma(\mathbf{h})|$ and $\Delta\varphi$ values.

The first and relatively straightforward output of MASC, as shown in this article, is the derivation of the $\{\Gamma(\mathbf{h})\}$ amplitudes. The determination of the macromolecular envelope from these amplitudes is a problem which is currently under study, and for which promising results have been obtained in a test case (R. Kahn *et al.*, in preparation). This determination is probably the most useful output of any MASC study. Further, the $\{\Gamma(\mathbf{h})\}$ phases can be calculated from the partial structure (with the correct enantiomorph) of anomalous scatterers; the derivation of the $\{{}^{\circ}F(\mathbf{h})\}$ phases is then straightforward.

If anomalous scatterers have in addition ordered sites (*i.e.* the mixed MASC/MAD case), the expression of $\Gamma_T(\mathbf{h})$, the total normal structure factor of anomalous scatterers, is

$$\Gamma_T(\mathbf{h}) = \Gamma(\mathbf{h}) + {}^{\circ}F_A(\mathbf{h}), \quad (2)$$

where ${}^{\circ}F_A(\mathbf{h})$ is the normal structure factor of ordered anomalous scatterers. In the mixed MASC/MAD case, Bragg reflections can be separated into two parts: the low-resolution data ($d \geq$ 8–10 Å) where MASC effects are dominant if the concentration of anomalous scatterers is not too low (*cf.* the case of HEWL below), and from which the molecular envelope can be determined as described previously; and the medium and high-resolution data ($d \leq$ 8 Å), where MAD effects are dominant and can be also exploited to obtain phase information.

Two preliminary MASC experiments at room temperature, reported in Fourme *et al.* (1995), were carried out on two proteins, HEWL and P64k. HEWL (hen egg-white lysozyme) is a small protein (molecular weight 14.3 kDa) crystallized in a tetragonal form (space group $P4_32_12$, $a = b = 79.24$, $c = 37.86$ Å at room temperature, one molecule per asymmetric unit) with a known structure (Blake *et al.*, 1965). P64k is a protein (molecular weight 64 kDa) extracted from the outer membrane of *Neisseria meningitidis*. Crystals of P64k were grown in ammonium sulfate and the structure was solved at LURE (Li de La Sierra *et al.*, 1997). These crystals (space group $P4_32_12$, $a = b = 140.1$, $c = 79.8$ Å at room temperature, one molecule per asymmetric unit) lack the lipoyl domain and contain only the lipoamide dehydrogenase domain. The name P64k will refer to this domain throughout the rest of this article. A pioneering experiment on tetragonal HEWL grown from a solution of ytterbium chloride was performed at LURE-DCI using the D23 setup and a spherical drift MWPC (multiwire proportional chamber); several data sets were collected at wavelengths around the Yb $L_{\rm III}$ absorption edge. A preliminary MASC experiment on P64k, soaked in ammonium selenate, was later carried out at the ESRF (Grenoble, France) on the TROIKA beamline at wavelengths around the Se K edge. In both experiments, the crystals deteriorated rapidly under the high flux of synchrotron radiation. Further MASC experiments, which are reported in this article, were thus conducted at low temperatures on HEWL, P64k and, in addition, on xylose isomerase (XI) crystals. XI is extracted from *Actinoplanes missouriensis*, and crystallizes as a tetramer (total molecule weight 173.2 kDa) in the asymmetric unit (space group $P3_221$, $a = b = 143.45$, $c = 231.50$ Å at room temperature). The crystal structure of XI has been determined by Rey *et al.* (1988).

In this paper, we shall describe methods used for the extraction of $\{|\Gamma_T(\mathbf{h})|\}$ values from multiwavelength data sets collected on crystals of these three monomeric or oligomeric proteins with different molecular weights in the range 14.3–173.2 kDa, using Yb or Se as anomalous scatterers dispersed in the solvent. The known three-dimensional structures of these proteins were used as models. Model and experimental amplitudes are in satisfactory agreement in the resolution range where they are the largest. The existence of ordered anomalous scatterers at the protein surface has been proved for the three cases. The use of other anomalous scatterers (*e.g.* Rb, Br *etc.*) for MASC experiments with the same proteins is currently under way. These results are not presented here.

Table 1
Data collection and processing for P64k in 3.5 *M* $(NH_4)_2SeO_4$.

Protein	P64k		
Synchrotron beamline	TROIKA (ESRF)		
Anomalous scatterer	Se		
Concentration of anomalous scatterer (*M*)	3.5		
Temperature (K)	124		
Space group	$P4_32_12$		
Unit-cell parameters (Å)	140.62, 140.62, 77.02		
$\Delta V/V$ unit cell (%)	−2.8		
Solvent fraction (%)	60.5		
Crystal-to-detector distance (mm)	370		
Wavelengths (Å)	0.9919	0.9795	0.9793
	Remote low energy	$\|f'\|$ max	f'' max
Exposure time per frame (s)	30	45	60
Angular increment per frame (°)	2.0	2.0	2.0
Total rotation range (°)	100.0	100.0	100.0
Measured reflections	44222	45453	45310
Unique reflections	6008	6107	6106
Resolution limits (Å)	100.9–4.19	100.0–4.18	100.0–4.18
Data completeness (%)	98.2	99.9	99.9
Overall $R_{\rm merge}$ (%)	5.5	5.1	5.6
$\langle I\rangle/\langle\sigma(I)\rangle$	5.2	9.4	8.0
$N_{\rm obs}/N_{\rm theoretical}$ ($d \geq 20$ Å)	73/73		
Completeness for $d \geq 20$ Å (%)	100		

2. Experimental

2.1. P64k and XI at the ESRF

A series of experiments was carried out in June 1995 on the TROIKA station at the ESRF for both XI and P64k crystals. An intense and highly monochromated beam was obtained by using a flat (333) silicon crystal combined with the undulator source (relative bandpass 7×10^{-5} at 1 Å). The detector was a MAR image-plate system. In order to shorten the dead time between images, a diameter of 18 cm out of 30 cm was effectively scanned after each X-ray exposure. In order to record Bragg reflections at very low resolution, a small beamstop was mounted a few millimetres from the detector entrance window and carefully centred. A helium cone was installed between the crystal and the detector to reduce absorption of X-rays by air, as well as the background due to X-ray scattering in the air path between the crystal and the beamstop. High-quality crystals of both XI and P64k were grown in ammonium sulfate. Selenate and sulfate ions are isostructural (they are both tetrahedral oxyanions), enabling ammonium sulfate in the mother liquor to be substituted with multimolar concentrations of $(NH_4)_2SeO_4$ in a K_2HPO_4 buffer *via* simple soaking for 4–5 h. Crystals were then transferred to solutions containing ammonium selenate at the desired concentration with 25% glycerol used as a cryoprotectant. Samples were held in nylon loops of sub-millimetre diameter (Teng, 1990) and kept under a stream of cold nitrogen gas at 124 K during the whole data collection. For each protein, two crystals were exposed to X-rays, the only difference between these crystals being the concentration of ammonium selenate in the solvent. In this manner, both chemical and anomalous contrast variation data sets were obtained. Wavelengths were selected after recording the Se *K* edge fluorescence spectra of a capillary filled with 0.1 *M* $(NH_4)_2SeO_4$. In the case of P64k samples, the first crystal was soaked in 3.5 *M* $(NH_4)_2SeO_4$ (a concentration ensuring approximate contrast matching) and diffraction data at three wavelengths (Table 1) around the Se *K* edge were recorded, while for the second one with 2.5 *M* $(NH_4)_2SeO_4$, only one data set was collected at a remote low-energy wavelength (0.9919 Å). In the case of XI, data sets were collected at four wavelengths in 2 *M* $(NH_4)_2SeO_4$ (Table 2) for the first crystal and at three wavelengths in 3 *M* $(NH_4)_2SeO_4$ for the second crystal. Each crystal was oriented in order to collect Bijvoet mates simultaneously on the same image. For P64k and XI, **a** and the bisector of *a* and *b* were set parallel to the spindle axis, respectively. Evidence of fluorescence appeared for the wavelengths near the absorption edge on the first XI crystal exposed to X-rays. As expected, the fluorescence level was maximum at the peak (f'' maximum) wavelength. The second XI crystal was rinsed for a few seconds in ammonium sulfate solution before flash-freezing, in order to remove the excess of anomalous scatterers in the mother liquor surrounding the crystal. This reduced the fluorescence background quite substantially. There are other ways to improve the signal-to-noise ratio of the diffraction signal for an image flooded by fluorescence. (i) Reducing the oscillation range $\Delta\Phi$ of each image. The $\Delta\Phi$ value used in our experiments could not be reduced below 0.5°, due to the fairly long readout time of the image-plate scanner. CCD-based systems would allow smaller oscillations. (ii) Increasing the sample-to-detector distance *D*, as the fluorescence intensity is emitted isotropically and thus varies as $1/D^2$. In view of their superior quality, we report here results for the P64k crystal soaked in 3.5 *M* $(NH_4)_2SeO_4$ and the XI crystal soaked in 2 *M* $(NH_4)_2SeO_4$.

2.2. HEWL at LURE

The second series of experiments was performed during November and December 1996 on the DW21b setup at LURE-DCI, installed on a wiggler line and designed for anomalous dispersion experiments in biocrystallography. The combination of a cylindrical rhodium-coated mirror and a double Si(111) crystal monochromator provides vertical and sagittal focusing and a relative bandpass of 4×10^{-4} at 1 Å. The detector is an image-plate system, diameter 180 mm, built by J. Hendrix and A. Lentfer at the EMBL Outstation, Hamburg. The standard beamstop was replaced by a small beamstop (diameter 2.5 mm) placed close to the detector entrance window and moveable along polar coordinates.

A preliminary MASC experiment with Yb and HEWL had been carried out in 1993 at the D23 station (Fourme *et al.*, 1995) using single crystals of HEWL crystallized from solutions of 0.3–0.5 *M* $YbCl_3$. This combination was chosen

Table 2
Data collection and processing for XI in 2 *M* $(NH_4)_2SeO_4$.

Protein	XI			
Synchrotron beamline	TROIKA (ESRF)			
Anomalous scatterer	Se			
Concentration of anomalous scatterer (*M*)	2			
Temperature (K)	124			
Space group	$P3_221$			
Unit-cell parameters (Å)	141.91, 141.91, 227.48			
$\Delta V/V$ unit cell (%)	−3.9			
Solvent fraction (%)	62.6			
Crystal-to-detector distance (mm)	370			
Wavelengths (Å)	0.9919	0.9795	0.9793	0.9791
	Remote low energy	$\|f'\|$ max	f'' max	Remote high energy
Exposure time per frame (s)	25	25	25	25
Angular increment per frame (°)	1.2	1.2	0.5	1.2
Total rotation range (°)	96.0	96.0	96.0	96.0
Measured reflections	110358	114299	114009	114302
Unique reflections	19684	20430	19429	20456
Resolution limits (Å)	105.4–4.18	105.4–4.12	105.4–4.13	105.4–4.12
Data completeness (%)	97.3	95.4	96.2	95.5
Overall R_{merge} (%)	3.7	7.0	19.4	10.6
$\langle I\rangle/\langle\sigma(I)\rangle$	17.3	11.8	3.9	6.9
$N_{obs}/N_{theoretical}$ ($d \geq 20$ Å)	217/224			
Completeness for $d \geq 20$ Å (%)	96.9			

because lanthanides feature a sharp white line at the L_{III} absorption edge which is, in the case of ytterbium, at a wavelength (1.3862 Å) most convenient for an instrument installed on a bending-magnet line. However, the diffraction data was collected at room temperature and the completeness was only 60%. The experiment was thus repeated on two crystals at 124 K. Both were grown in NaCl solutions, then soaked for a day in 0.5 *M* $YbCl_3$ solutions containing 10% glycerol and a CH_3CO_2Na buffer (pH = 4.6). They were briefly washed in 1.5 *M* NaCl and 10% glycerol solutions just before flash-freezing for reasons mentioned above. Data sets at four wavelengths for both crystals were collected at 124 K. The first crystal could not be oriented so that Bijvoet pairs appeared on the same image, due to the limited rotation ranges of the goniometer head. In contrast, the second crystal was oriented with **a** parallel to the spindle axis. Results are reported for this latter crystal.

3. Data processing

3.1. Integration

We present here the results concerning the data processing of three multiwavelength data sets, with one crystal per sample selected as indicated above. All data were integrated using the program *MOSFLM* (Leslie, 1992). Cell parameters were refined independently for each wavelength by the post-refinement technique (Winkler *et al.*, 1979; Rossmann *et al.*, 1979) implemented in *MOSFLM*. A systematic decrease of the unit-cell volume after cryocooling was observed (Tables 1, 2 and 3). Intensities were estimated using the profile-fitting option. In order to keep as many low-resolution reflections as possible, the acceptance limit of the background gradient to the average background was increased to take into account fairly large fluctuations of the background level near the beamstop shadow. Similarly, the profiles of some low-resolution reflections would deviate significantly from the standard estimated profile, but this could be overcome either by raising the tolerance levels or using the 'integrated' intensity option.

3.2. Scaling

Scaling of the data was performed *via* the *CCP*4 (Collaborative Computational Project, Number 4, 1994) program *SCALA* (Evans, 1993) in two stages. The first step consisted of an internal scaling for each wavelength taking into account the deterioration of the sample and the intensity decay of the synchrotron radiation incident beam during data collection. In the second step, the internally scaled data set of the remote low-energy wavelength was merged as a function of the image number and used as a reference set to scale the other sets. This scale factor was varied according to the spindle axis rotation and the coordinates on the detector sensitive area divided into either nine (XI) or 16 zones (HEWL and P64k). This treatment corresponds to a 'pseudo'-local scaling function and minimizes absorption effects. No *B* factor was applied at this stage. The resulting scale factors were then applied to intensities *via* the *CCP*4 program *AGROVATA*. Data were merged with the same program. Extraction of structure-factor amplitudes followed, which were put on a approximate absolute scale (*CCP*4 program *TRUNCATE*). Anomalous differences for acentric reflections were calculated for Bijvoet pairs recorded on the same or adjacent images. Concerning data completeness (Table 3), seven out of 13 reflections with $d \geq 20$ Å were collected for HEWL, the six remaining reflections falling into the blind zone of the reciprocal space. For XI, the completeness is good for reflections with $d \geq 20$ Å as only seven out of 224 are lacking. These seven reflections belong to the blind zone of the reciprocal space. The P64k data set is complete (73 reflections with $d \geq 20$ Å). Experimental conditions and statistics are reported in Tables 1, 2 and 3 for P64k, XI and HEWL crystals, respectively. The internal agreement between symmetry-equivalent reflections is good except for the peak data set of XI, as a result of strong fluorescence in this latter case. As in a MAD experiment, the strength of anomalous signals in a MASC experiment can be quantitatively estimated through Bijvoet and dispersive ratios (Fourme *et al.*, 1995). Figs. 1(*a*), 1(*b*) and 1(*c*) show Bijvoet ratios $\langle\big||^{\lambda}F(+\mathbf{h})| - |^{\lambda}F(-\mathbf{h})|\big|\rangle/\langle|^{\lambda}F(\mathbf{h})|\rangle$, where $\langle\ \rangle$ denote the r.m.s. value and $|^{\lambda}F(\mathbf{h})| = [|^{\lambda}F(+\mathbf{h})| + |^{\lambda}F(-\mathbf{h})|]/2$, for P64k, XI and HEWL respectively, over the whole resolution range and for the various wavelengths. These ratios have been corrected for

Table 3
Data collection and processing for HEWL in 0.5 *M* $YbCl_3$.

Protein	HEWL			
Synchrotron beamline	DW21b (LURE-DCI)			
Anomalous scatterer	Yb			
Concentration of anomalous scatterer (*M*)	0.5			
Temperature (K)	124			
Space group	$P4_32_12$			
Unit-cell parameters (Å)	78.68, 78.68, 37.05			
$\Delta V/V$ unit cell (%)	−3.5			
Solvent fraction (%)	29.3			
Crystal-to-detector distance (mm)	250			
Wavelengths (Å)	1.3928	1.3859	1.3852	1.3776
	Remote low energy	$\|f'\|$ max	f'' max	Remote high energy
Exposure time per frame (s)	90	90	90	90
Angular increment per frame (°)	3.0	3.0	3.0	3.0
Total rotation range (°)	135.0	135.0	135.0	135.0
Measured reflections	10257	10427	10436	10617
Unique reflections	1092	1112	1111	1129
Resolution limits (Å)	33.5–3.95	33.5–3.93	33.5–3.92	33.5–3.91
Data completeness (%)	84.7	86.4	86.3	87.8
Overall $R_{\rm merge}$ (%)	2.8	2.5	2.7	2.3
$\langle I\rangle/\langle\sigma(I)\rangle$	18.3	20.9	19.6	23.3
$N_{\rm obs}/N_{\rm theoretical}$ ($d \geq 20$ Å)	7/13			
Completeness for $d \geq 20$ Å (%)	53.8			

noise by computing $R_{\rm merge}$ on intensities for centric reflections only, which gave an estimation of the noise for each resolution bin. Corrected Bijvoet ratios are large at very low resolution and decrease rapidly with increasing resolution up to about 10 Å. In the medium-resolution range (between 10 and 4 Å), these ratios remain significant except for the remote wavelength, which is an indication of the existence of ordered sites giving rise to MAD effects. As expected, Bijvoet ratios are highest at the peak wavelength corresponding to the maximum value of ${}^{\lambda}f''$.

3.3. Derivation of $\{|\Gamma_T(\mathbf{h})|\}$

As discussed in the introduction, the program *MADLSQ* (Hendrickson, 1985) can be used to obtain $|{}^{\circ}F(\mathbf{h})|$, $|\Gamma_T(\mathbf{h})|$ and $\Delta\varphi$ values from the set of $|{}^{\lambda}F(\pm\mathbf{h})|^2$. Refinement of ${}^{\lambda}f'$ and ${}^{\lambda}f''$ values is performed by minimizing a global residual corresponding to the sum over all reflections of local residuals. Initial values for ${}^{\lambda}f'$ and ${}^{\lambda}f''$ may be derived following the standard procedure (Hendrickson *et al.*, 1988) implemented in the *MADSYS* suite (Hendrickson, 1991). f'' values are derived from absorption or fluorescence spectra, with an absolute scaling based on theoretical values at wavelengths remote from the selected absorption edge. Then f' values are obtained from f'' values by numerical integration, using the Kramers–Kronig relation.

In the case of Yb in HEWL, a fluorescence spectrum of a 0.1 *M* $YbCl_3$ solution near the Yb $L_{\rm III}$ edge was recorded on the DW21b station (Fig. 2). Theoretical values for ${}^{\lambda}f'$ and ${}^{\lambda}f''$ (Sasaki, 1984) were chosen as initial values.

In the case of Se in P64k and XI, the fluorescence spectrum recorded at the ESRF of a 0.1 *M* ammonium selenate solution did not extend far enough from the absorption edge. Accordingly, an EXAFS spectrum of a 0.1 *M* selenate solution was later recorded at LURE-DCI on the D13 setup with an energy resolution close to that of the ESRF experiment (Fig. 3*a*). This spectrum features a sharp resonance. From this absorption spectrum, ${}^{\lambda}f'$ and ${}^{\lambda}f''$ experimental curves were derived (Fig. 3*b*). The largest values of f'' and $|f'|$ obtained are 9.4 and 11.8, respectively. MASC data were recorded at the ESRF using a Si(333) monochromator and an undulator source. The ESRF setup ensures a more parallel beam and a narrower bandpass than the bending magnet and the Si(311) monochromator of D13. Consequently, maximum values of $|f'|$ and f'' measured at LURE are probably somewhat underestimated.

The signal-to-noise ratio for the $|\Gamma_T(\mathbf{h})|$ values has been estimated by computing, for ten resolution bins with Δs^2 constant (where s is the resolution) and $d \geq 7$ Å, $\langle|\Gamma_T(\mathbf{h})|/\sigma[|\Gamma_T(\mathbf{h})|]\rangle$, with $\sigma[|\Gamma_T(\mathbf{h})|]$ the standard deviation of $|\Gamma_T(\mathbf{h})|$. Unrealistically large values for this ratio have been discarded by using a cutoff. For each protein, this ratio decreases *versus* resolution. In a similar way, $\langle\sigma[\Delta\varphi(\mathbf{h})]\rangle$ has been estimated for the same resolution bins for the three proteins. Figs. 4(*a*) and 4(*b*) represent $\langle|\Gamma_T(\mathbf{h})|/\sigma[|\Gamma_T(\mathbf{h})|]\rangle$ and $\langle\sigma[\Delta\varphi(\mathbf{h})]\rangle$, respectively, over the resolution range ∞–7 Å for the three proteins. The cutoff values for $\langle|\Gamma_T(\mathbf{h})|/\sigma[|\Gamma_T(\mathbf{h})|]\rangle$ are indicated. The highest values for $\langle|\Gamma_T(\mathbf{h})|/\sigma[|\Gamma_T(\mathbf{h})|]\rangle$ and the smallest ones for $\langle\sigma[\Delta\varphi(\mathbf{h})]\rangle$ on the whole resolution range are obtained for HEWL. Both standard deviations $\sigma[|\Gamma_T(\mathbf{h})|]$ and $\sigma[\Delta\varphi(\mathbf{h})]$ are computed in *MADLSQ* from the standard deviations $\sigma(p_i)$ (i = 1, 2, 3, 4) with $p_1 = |{}^{\circ}F(\mathbf{h})|^2$, $p_2 = |\Gamma_T(\mathbf{h})|^2$, $p_3 = |{}^{\circ}F(\mathbf{h})||\Gamma_T(\mathbf{h})|\cos[\Delta\varphi(\mathbf{h})]$ and $p_4 = |{}^{\circ}F(\mathbf{h})||\Gamma_T(\mathbf{h})|\sin[\Delta\varphi(\mathbf{h})]$. One has the general relation

$$\sigma(p_i)^2 = (\mathbf{Q}^{-1})_{ii}\chi^2,$$

where

$$\chi^2 = \sum_j [I_j(\mathbf{h})_{\rm obs} - I_j(\mathbf{h})_{\rm calc}]^2/(N_p - 3) = R_{\mathbf{h}}/(N_p - 3)$$

is a sum over wavelengths and Bivoet mates for the reflection $\mathbf{h}$, with $I_j(\mathbf{h})$ the observed or calculated (from *MADLSQ*) intensity. N_p is the number of observations per reflection. $\mathbf{Q}$ is the hessian matrix (matrix of the second derivatives *versus* the p_i variables) of $R_{\mathbf{h}}$. From the terms $\sigma(p_i)$, one can compute $\sigma[|\Gamma_T(\mathbf{h})|]$ and $\sigma[\Delta\varphi(\mathbf{h})]$ with the following relations

$$\sigma[|\Gamma_T(\mathbf{h})|] = \sigma(p_2)/2|\Gamma_T(\mathbf{h})|$$

and

$$\sigma[\Delta\varphi(\mathbf{h})] = [p_3^2\sigma(p_4)^2 + p_4^2\sigma(p_3)^2]^{1/2}/(p_3^2 + p_4^2).$$

The program *GFROMF* (Carter & Bricogne, 1987; Carter *et al.*, 1990), which was originally designed for chemical contrast-variation data, has been modified to take into account anomalous dispersion. As *MADLSQ*, it determines the $|\Gamma_T(\mathbf{h})|$ values from MASC data. The results obtained with *GFROMF* are essentially the same as those obtained with *MADLSQ* (see below) but the knowledge of extra parameters such as $\mathring{\rho}_{sA}$ and $\bar{\rho}_P$, the mean electronic density in the ordered phase, is needed. With *MADLSQ*, the agreement between experimental and model $|\Gamma_T(\mathbf{h})|$ is satisfactory for $d \geq 20$ Å without considering ordered anomalous sites and up to at least 10 Å if one takes them into account. As a result, we will focus only on the experimental $|\Gamma_T(\mathbf{h})|_{\rm obs}$ obtained from *MADLSQ*.

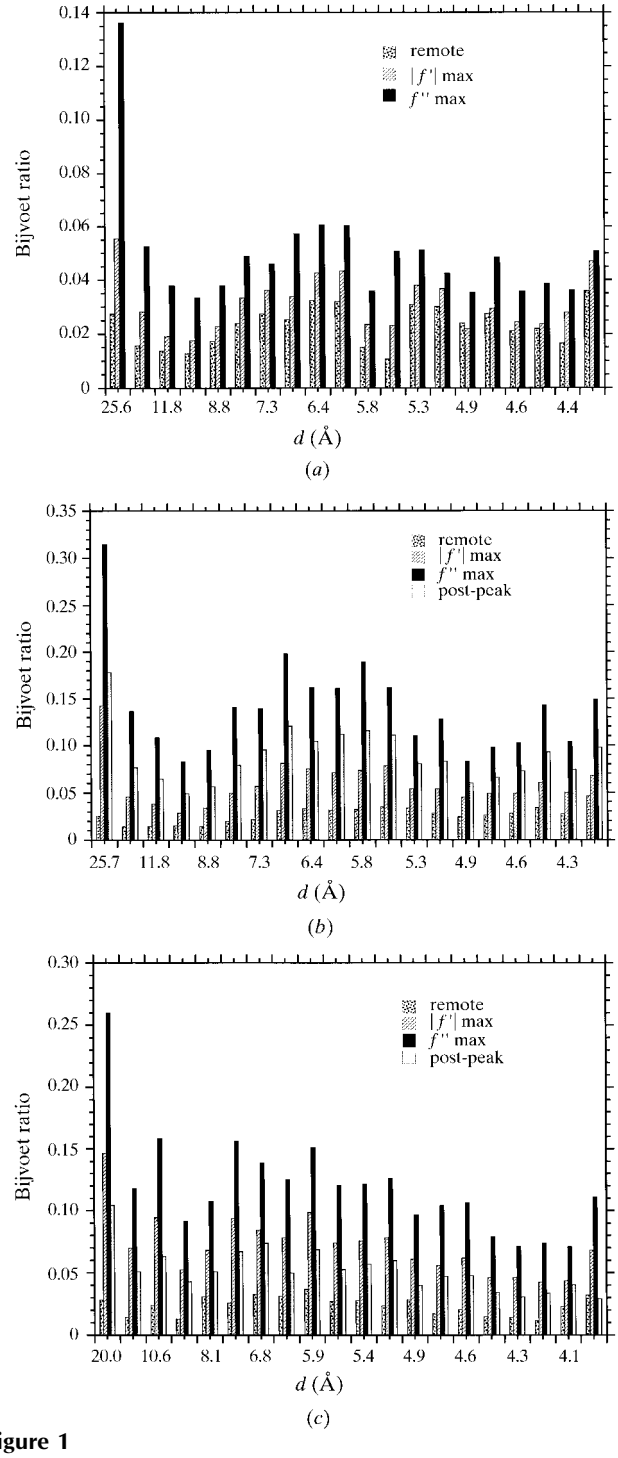

Figure 1
(*a*) Histogram of Bijvoet ratios for P64k in 3.5 *M* $(NH_4)_2SeO_4$. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$. (*b*) Histogram of Bijvoet ratios for XI in 2 *M* $(NH_4)_2SeO_4$. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta\lambda$. (*c*) Histogram of Bijvoet ratios for HEWL in 0.5 *M* $YbCl_3$. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$.

4. Determination of low-temperature models for the three proteins

4.1. From the room-temperature to low-temperature structures

Since all of the MASC data in this study were collected at low temperature, determination of low-temperature models was necessary in each case. The three-dimensional structures of the three test proteins have been solved and their coordinates at room temperature have been retrieved from the Protein Data Bank (PDB) (Bernstein *et al.*, 1977). Molecular replacement was performed with the package *AMoRe* (Navaza, 1994), taking the low-energy remote wavelength MASC data with *d* spacings <8 Å as input data and the three-dimensional structures at room temperature as starting models. These models were positioned in the low-temperature unit cell and the calculation terminated by a step of fast rigid-body refinement (Huber & Schneider, 1985). Molecular replacement was followed by a second round of rigid-body refinement in the same resolution range with the package *X-PLOR* (Brünger, 1992), considering protein domains as rigid bodies. The refinement was controlled by the use of the $R_{\rm free}$ factor (Brünger *et al.*, 1993).

4.2. Ordered anomalous scatterers

As significant anomalous signals are observed at medium resolution for the three proteins, there is evidence of ordered anomalous scatterers. An accurate description of the model should include these atoms, which contribute to normal and resonant scattering in the whole resolution range. The search

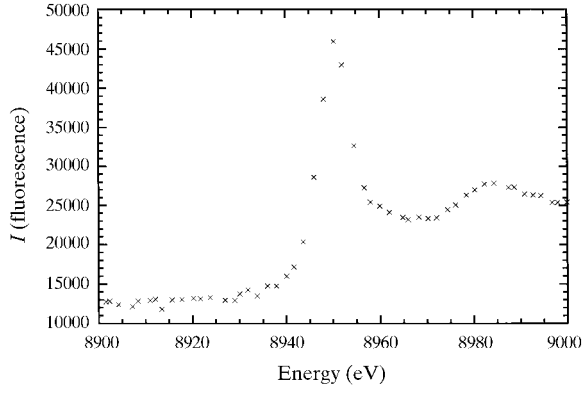

Figure 2
Fluorescence spectrum of a 0.1 *M* $YbCl_3$ solution recorded on the DW21b station near the Yb $L_{\rm III}$ edge (λ = 1.3862 Å) with an Si(111) monochromator and showing a strong white line.

for these atoms was based on Bijvoet or Friedel mates data at the peak wavelength. For HEWL, Yb was the anomalous scatterer with $f''_{max} \simeq 16\ e^-$. For XI and P64k, the anomalous scatterer was Se with $f''_{max} \simeq 9\ e^-$. Reflections possibly affected by MASC effects (d spacings ≥ 8 Å) were not used. The maximum resolution was 3.9 Å for HEWL and 4.1 Å for XI and P64k.

The phased anomalous Fourier difference map (Pepinski & Okaya, 1956) is a very sensitive method to detect residual anomalous scatterers in an otherwise known structure. On the basis of the low-temperature models obtained as described above, such maps were calculated and proved the existence of ordered sites in all cases. Examples of such sites are shown in Figs. 5, 6 and 7 for P64k, XI and HEWL respectively, where the envelope model is superimposed with the peaks of the phased anomalous Fourier difference map.

For HEWL, both an anomalous Patterson map using all anomalous pairs (Rossmann, 1961) and a Patterson map using all the $|\Gamma_T|^2$ for $d < 8$ Å (*i.e.* essentially the $|^\circ F_A|^2$ from *MADLSQ*) gave two Yb sites, the first one corresponding to self-vectors in several Harker sections of these maps and the second one obtained from cross-vectors with the first site. The program *SHARP* (De La Fortelle & Bricogne, 1997) was used to refine the occupancy and position of the strongest site. Residual maps based on log-likelihood gradient were produced by *SHARP* and confirmed the second site. In a second round of calculations with *SHARP*, the occupancies and positions of both sites were refined. No extra site was discovered in the residual maps. A few additional peaks, not detected by *SHARP*, were visible in the phased anomalous Fourier difference map, and were interpreted as sites with a marginal occupancy. In order to confirm these results, an 8 Å Fourier difference map was calculated with amplitudes based on the difference between $\{|\Gamma_T(\mathbf{h})|\}$ from *MADLSQ* and $\{|^\circ F_A|\}$ calculated from the two main sites and the $\{^\circ F_A\}$ phases. Four additional sites with heights above 4σ were thus discovered. To estimate the site occupancies, it was assumed that the phase of the total normal structure factor $\phi_{\circ F}$ is nearly the same as the phase of the protein in vacuum ϕ_P for $d < 8$ Å. This latter phase was calculated from the low-temperature model and enabled us to recover ϕ_{Γ_T} by using the relation

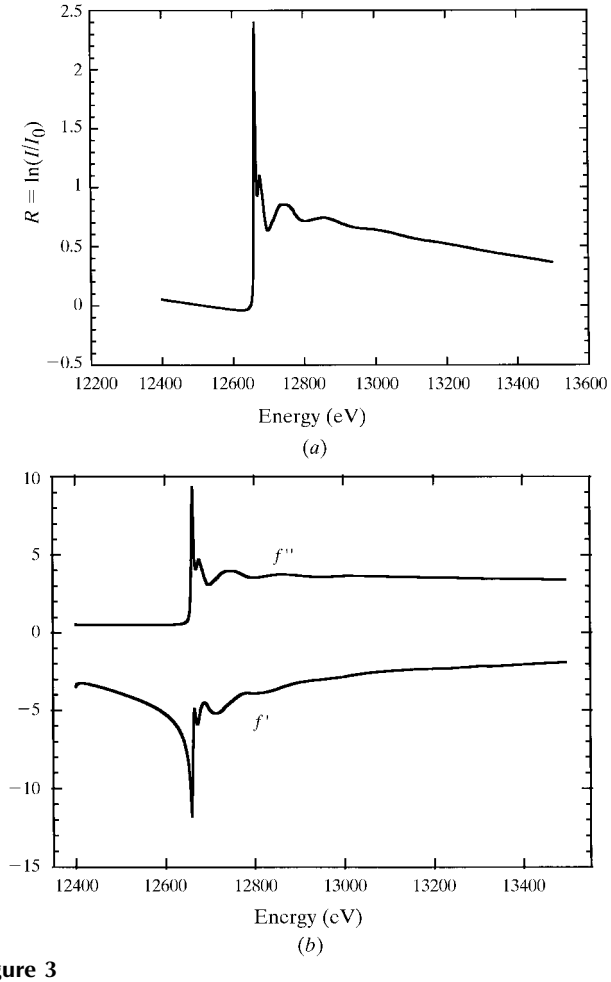

Figure 3
(*a*) EXAFS spectrum of a 0.1 *M* $(NH_4)_2SeO_4$ solution recorded on the D13 station of the DCI-LURE storage ring at Orsay, France at the selenium *K* edge (λ = 0.9793 Å) with a double Si(311) monochromator. Communicated by F. Villain. (*b*) Experimental curves of the real and imaginary anomalous scattering factors $^\lambda f'$ and $^\lambda f''$ of Se in SeO_4^{2-} around the Se *K* edge, extracted from the EXAFS spectrum of Se (Fig. 3*a*).

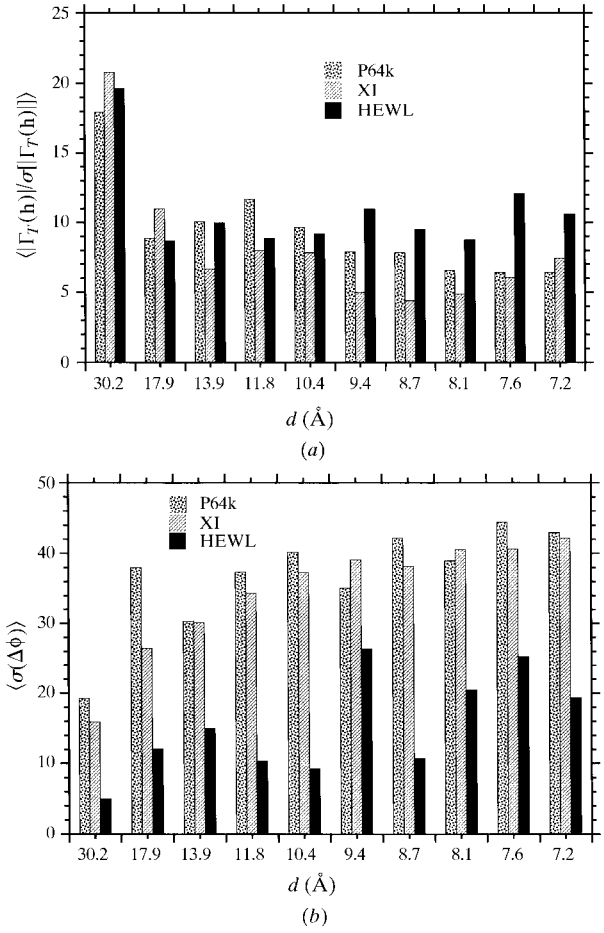

Figure 4
(*a*) Histograms of $\langle|\Gamma_T(\mathbf{h})|/\sigma|\Gamma_T(\mathbf{h})|\rangle$ for P64k, XI and HEWL. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$. A cutoff of 100.0 for P64k and XI and of 40.0 for HEWL was applied. (*b*) Histograms of $\sigma[\Delta\varphi(\mathbf{h})]$ for P64k, XI and HEWL. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$.

$\phi_{\Gamma_T} = \phi_{\circ F} - \Delta\varphi$. The map with coefficients $|\Gamma_T|_{\rm obs} \exp(i\phi_{\Gamma_T})$ was then calculated for $d < 8$ Å, which gave as its first six positive peaks the six sites already obtained. The four residual site occupancies estimated from the heights of the peaks lie between 0.2 and 0.3. These results were validated by the comparison between experimental and model $|\Gamma_T(\mathbf{h})|$ (see below).

For P64k, one Se site was recovered from an anomalous Patterson map (discarding pairs with the largest anomalous differences) and no extra sites were found. Attempts to use the same method as with HEWL to find other sites were unsuccessful, the agreement between experimental and model $|\Gamma_T(\mathbf{h})|$ remaining poor for $d < 8$ Å.

For XI, both anomalous Patterson maps and Patterson maps based on the $|\Gamma_T|^2$ for $d < 8$ Å were not able to provide recognizable sites, however the phased anomalous map suggests possible sites at the surface of the macromolecule (see Fig. 6). None of these sites could be confirmed after refinement with *SHARP*.

5. Comparison of experimental and model $\{|\Gamma_T(\mathbf{h})|\}$

5.1. Derivation of model $\{|\Gamma(\mathbf{h})|\}$

The *CCP*4 program *HGEN* (S. Phillips) was used to incorporate H atoms in the first low-temperature models obtained. Maps corresponding to these latter models were then calculated with another *CCP*4 program, *SFALL* (E. Dodson & T. Baker). The voxel size used for the grid was $(0.5\ \text{Å})^3$ and remained the same for the calculation of the envelopes. Two different packages were used to get the indicator function. The first is a modified version of the program *INDICAT* (C. W. Carter), where a threshold defining the boundary between the solvent and protein domains is considered. Each voxel in the three-dimensional structure map of the macromolecular, built with *SFALL* from the PDB file, is viewed as a solvent voxel (density set to 0), a protein voxel (density set to 1) or a cavity voxel corresponding to a hole in the protein (density set to 1) according to the number of surrounding voxels above the threshold and the density in the voxel itself. The second is the *CCP*4 program *MAPMASK*. From the expected fraction occupied by the protein domain, *MAPMASK* builds a mask which can be converted into a map, with the same grid as before.

Once the indicator function is available, model $\{G(\mathbf{h})\}$ is calculated by inverse Fourier transform. Model $\{|\Gamma(\mathbf{h})|\}$, the normal structure factors of the disordered anomalous scatterers, are computed with the relation: $\Gamma(\mathbf{h}) = -{}^0\rho_{\rm sA} G(\mathbf{h}) \exp(-Bs^2/4)$. The value of ${}^0\rho_{\rm sA}$ is calculated from the concentration of the anomalous species in the mother liquor. The value of B is not known, but the exponential term remains close to 1 at very low resolution even for B values as large as 100 Å^2. Accordingly, B was taken to be initially equal to 0 for $d \geq 10$ Å.

5.2. Comparison between experimental and model $\{|\Gamma_T(\mathbf{h})|\}$

The first step was a comparison of $\{|\Gamma_T(\mathbf{h})|\}$ produced by *MADLSQ* and $\{|\Gamma(\mathbf{h})|\}_{\rm mod}$, *i.e.* those calculated with the contribution from disordered anomalous scatterers only. Experimental and model amplitudes were compared by computing the breakdown of the R factor [defined as $R = \sum_h \big||\Gamma_T(\mathbf{h})|_{\rm obs} - |\Gamma(\mathbf{h})|_{\rm mod}\big| / \sum_h |\Gamma_T(\mathbf{h})|_{\rm obs}$] as a function of resolution with the *CCP*4 program *RSTATS*. Results are shown in Table 4. Histograms of model and experimental amplitudes for $d \geq 10$ Å corresponding to the best R factors are shown in Figs. 8(*a*), 8(*b*) and 8(*c*) for P64k, XI and HEWL, respectively. The average amplitudes have their largest values for $d \geq 20$ Å. $\{|\Gamma(\mathbf{h})|_{\rm mod}\}$ are lower for XI than for P64k in this resolution range for two reasons: firstly, ${}^0\rho_{\rm sA}$ computed from the concentration in anomalous scatterers is smaller for the former protein than for the latter and, secondly, two reflections with very large $|\Gamma(\mathbf{h})|_{\rm mod}$ are lacking in the case of XI in

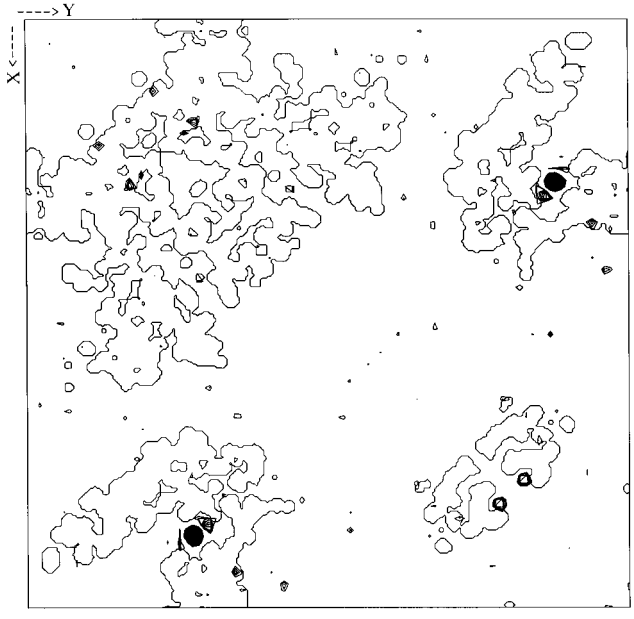

Figure 5
Section $z = 0$ of the phased anomalous Fourier-difference map of P64k in 3.5 *M* $(NH_4)_2SeO_4$ contoured at 3σ above the mean density superimposed with an envelope model section of P64k at low temperature.

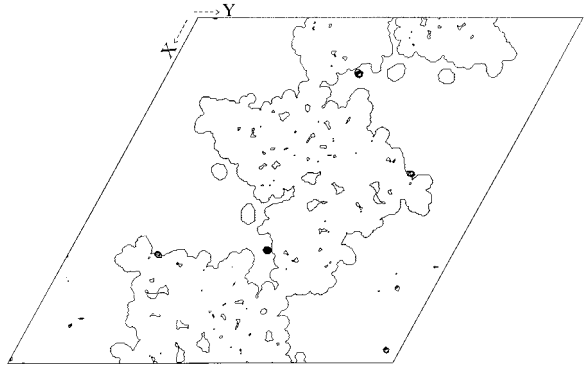

Figure 6
Section $z = 0.167$ of the phased anomalous Fourier-difference map of XI in 2 *M* $(NH_4)_2SeO_4$ contoured at 3σ above the mean density superimposed with an envelope model section of XI at low temperature.

Table 4
Results obtained from *MADLSQ*.

The model used and parameter values are indicated. Refined values of $^{\lambda}f'$ and $^{\lambda}f''$ are from *MADLSQ*. *R* factor = $\sum_{\mathbf{h}} \big||F_{\rm obs}| - |F_{\rm model}|\big| / \sum_{\mathbf{h}} |F_{\rm obs}|$.

Protein	P64k Se 3.5 *M*		XI Se 2 *M*		HEWL Yb 0.5 *M*	
	MAPMASK, low-temperature model		*MAPMASK*, low-temperature model		*INDICAT*, low-temperature model	
$^{\lambda}f'$ and $^{\lambda}f''$						
Remote long λ	−3.89	0.51	−3.89	0.51	−14.20	3.77
Inflection point	−9.69	3.90	−9.95	2.62	−22.54	9.22
Peak maximum	−10.69	9.20	−10.76	8.33	−16.35	5.41
Remote short λ	—	—	−2.09	4.25	−12.93	7.11
R factor $\|\Gamma(\mathbf{h})\|$ ($d \geq 20$ Å) (%)						
Excluding sites.	32.3		22.8		39.3	
Including sites.	—		—		10.5	

this resolution range. Between 10 and 20 Å, model amplitudes are systematically lower than the experimental values. This was expected, since the contribution of ordered sites is neglected in the model, whereas contributions from ordered and disordered anomalous scatterers are similar in this resolution range. At very low resolution ($d \geq 20$ Å), *R* factors vary from 0.393 for HEWL to 0.228 for XI and decrease with an increasing molecular weight. For HEWL, it will be shown below that this result can be considerably improved by taking into account ordered sites. Considering the simplicity of the binary description of the crystal structure and errors in both experimental and model amplitudes, these results are considered as satisfactory.

In cases where ordered anomalous scatterers have been detected, including their contribution should enhance the agreement between experimental and model amplitudes. This was performed for HEWL. The six Yb sites, detected as described previously, were included in the model to give

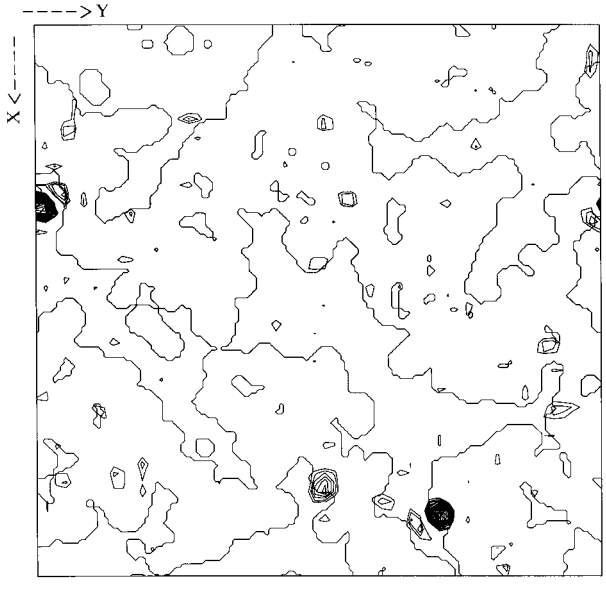

Figure 7
Section $z = 0.094$ of the phased anomalous Fourier-difference map of HEWL in 0.5 *M* $YbCl_3$ contoured at 3σ above the mean density superimposed with an envelope model section of HEWL at low temperature. The Yb atoms are located near the surface and/or in crevices.

$\{|\Gamma_T(\mathbf{h})|\}$ values. The agreement between model and experimental $\{|\Gamma_T(\mathbf{h})|\}$ was greatly improved in the entire resolution range (up to 3.9 Å), as shown in Fig. 9, with a *R* factor equal to 0.269. For *d* spacings ≥ 10 Å (64 reflections), the *R* factor is 0.323. Comparison of Figs. 8(*c*) and 9 shows than the MAD effects dominate the MASC effects for HEWL, even for $d \geq 20$ Å, since the model $\{|\Gamma_T(\mathbf{h})|\}$, taking the ordered sites into account, are about three times larger than the model $\{|\Gamma(\mathbf{h})|\}$ in this resolution range. This is not surprising for several reasons. Firstly, the concentration of ytterbium is relatively low (0.5 *M*) compared with the concentration of selenium for the two other proteins. Secondly, six ytterbium sites have been found, which is substantial, especially for such a small protein as HEWL. Finally, the very low resolution ($d \geq 20$ Å) shell is far from being complete (only seven reflections out of 13), which makes the comparison between model $\{|\Gamma_T(\mathbf{h})|\}$ and model $\{|\Gamma(\mathbf{h})|\}$ more difficult in this resolution range.

In the case of HEWL with a model including the ordered ytterbium sites, trials were made to refine the *B*-factor values in the expression

$$\Gamma_T(\mathbf{h}) = -^{0}\rho_{\rm sA} G(\mathbf{h}) \exp(-Bs^2/4) +^{\circ} F_A(\mathbf{h})$$

for different resolution ranges by minimizing a residual between experimental and model $\{|\Gamma_T(\mathbf{h})|\}$. No conclusive results were obtained.

6. Conclusions

Recording high-quality MASC data has been shown to be possible but requires some special precautions: a movable beamstop of a small size in order to record all low-resolution reflections, rinsing of crystals before data collection, the use of small oscillation angles at the peak wavelength and, finally, the use of cryocooled samples. The anomalous signals obtained in our MASC experiments are very large at low resolution, as predicted by theory, but remain significant at medium resolution, which proves the existence of additional MAD effects. The first step of the MASC methodology, *i.e.* the derivation of the $\{G(\mathbf{h})\}$ moduli from experimental MASC data can be considered as straightforward for $d \geq 20$ Å if one considers that, in this resolution range, for sufficiently high concentrations of anomalous scatterers, experimental $\{|\Gamma_T(\mathbf{h})|\}$ are

proportional to $\{|G(\mathbf{h})|\}$. For d spacings between 10 and 20 Å, this approximation is no longer valid and experimental $\{|\Gamma_T(\mathbf{h})|\}$ have to be phased directly. As the $\{\Gamma_T(\mathbf{h})\}$ are a combination of the $\{G(\mathbf{h})\}$ and $\{^{\circ}F_A(\mathbf{h})\}$ terms, if the phases of the $\{^{\circ}F_A(\mathbf{h})\}$ are known from medium-resolution data (*cf.* above), this is equivalent to finding the phases of the $\{G(\mathbf{h})\}$.

The agreement between experimental and model $\{|\Gamma_T(\mathbf{h})|\}$ may still be improved with a better description of the envelope model and of the solvent state. A 'flat' model for the solvent is obviously oversimplified. First, the existence of hydration shells at the surface of a biological macromolecule has already been demonstrated (Cheng & Schoenborn, 1990). These shells contribute to the ordered part of the solvent and should be taken into account by considering the existence of an interface width between the macromolecule and the solvent (Phillips, 1980; Schoenborn, 1988). Secondly, the solvent is in fact a liquid or an amorphous solid (at low temperature) that may have some regularities and the study of diffraction effects should take this into account. These regularities may also come from the attraction of anomalous scatterers by hydrophobic or hydrophilic residues or nucleotides in the biological macromolecule. For instance, domains exhibiting positive charges at their surface should attract more anomalous scatterers if the latter belong to a negatively charged ion (*e.g.* SeO_4^{2-}). The converse is true if the anomalous scatterer is a cation such as Yb^{3+}. All anomalous scatterers studied in this article are ions which may bind to proteins, which are locally charged entities. The use of neutral anomalous scatterers which should not bind to the macromolecule in a MASC experiment is currently under investigation. The following step, phasing the $\{G(\mathbf{h})\}$ terms, is also under study and will be described in a forthcoming paper (R. Kahn *et al.*, in preparation). Recording of both low and high-resolution reflections in X-ray biocrystallography experiments is now proved to be useful. To confirm potentialities of the MASC method, it must be applied to unknown macromolecular structures or assemblies such as protein–protein or protein–nucleic acid complexes, or viruses, for which it was originally intended.

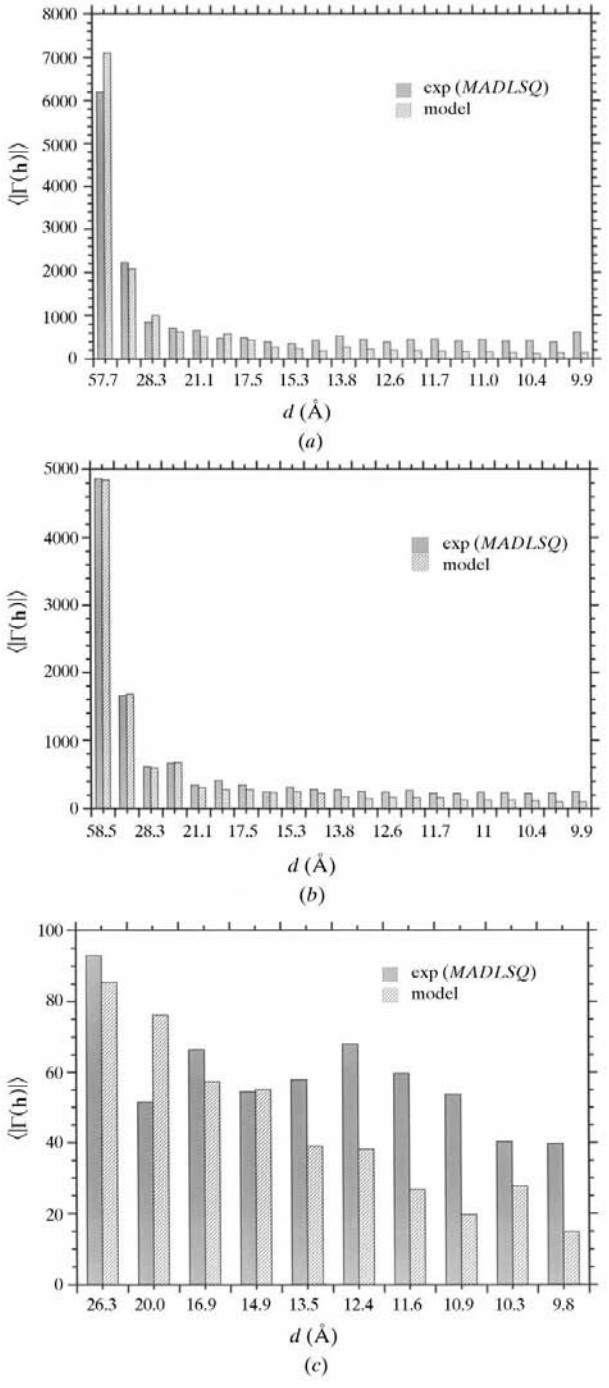

Figure 8
(*a*) P64k in 3.5 *M* $(NH_4)_2SeO_4$. Comparison of model $|\Gamma(\mathbf{h})|$ and experimental $|\Gamma_T(\mathbf{h})|$ obtained from *MADLSQ* for $d \geq 10$ Å. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$. (*b*) XI in 2 *M* $(NH_4)_2SeO_4$. Comparison of model $|\Gamma(\mathbf{h})|$ and experimental $|\Gamma_T(\mathbf{h})|$ obtained from *MADLSQ* for $d \geq 10$ Å. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$. (*c*) HEWL in 0.5 *M* $YbCl_3$. Comparison of model $|\Gamma(\mathbf{h})|$ and experimental $|\Gamma_T(\mathbf{h})|$ obtained from *MADLSQ* for $d \geq 10$ Å. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$.

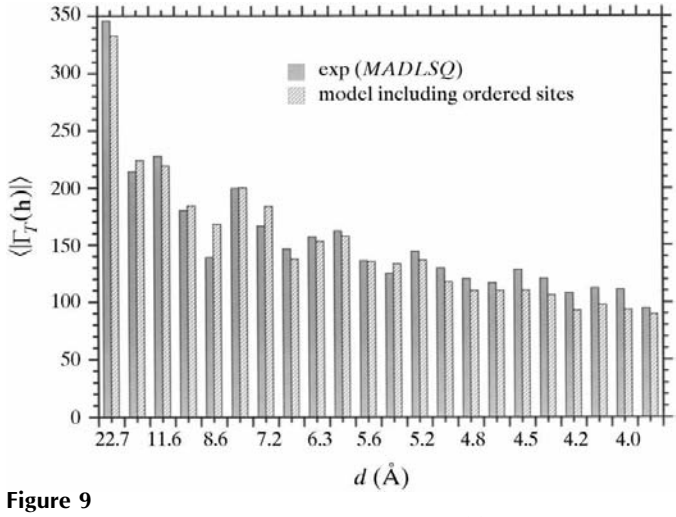

Figure 9
HEWL in 0.5 *M* $YbCl_3$. Comparison of model $|\Gamma_T(\mathbf{h})|$ taking into account ordered anomalous scatterers sites and experimental $|\Gamma_T(\mathbf{h})|$ obtained from *MADLSQ* for $3.9 \leq d \leq 33.5$ Å. Resolution bins correspond to Δs^2 = constant where $s = 2\sin\theta/\lambda$.

We thank A. Thompson, A. Gonzales and D. Abernathy for their help during the data collection at the ESRF and F. Rey for his gift of XI crystals. The high-resolution absorption spectrum of ammonium selenate was recorded with the help of F. Villain at LURE.

References

Bentley, G. A., Lewit-Bentley, A., Finch, J. T., Podjarny, A. D. & Roth, M. (1984). *J. Mol. Biol.* **176**, 55–75.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Blake, C. C. F., Koenig, D. F., Mair, G. A., North, A. C. T., Phillips, D. C. & Sarma, V. R. (1965). *Nature (London)*, **206**, 757–767.

Bricogne, G. (1974). *Acta Cryst.* A**30**, 395–405.

Brünger, A. T. (1992). *X-PLOR. Version 3.1. A System for X-ray Crystallography and NMR*, Yale University Press, New Haven, CT, USA.

Brünger, A. T. (1993). *Acta Cryst.* D**49**, 24–36.

Carter, C. W. Jr & Bricogne, G. (1987). *GFROMF: A computer program for scaling and estimating envelope structure factors from contrast variation data*. Department of Biochemistry, CB#7260, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599–7260, USA.

Carter, C. W., Crumley, K. V., Coleman, D. E., Hage, F. & Bricogne, G. (1990). *Acta Cryst.* A**46**, 57–68.

Cheng, X. & Schoenborn, B. P. (1990). *Acta Cryst.* B**46**, 195–208.

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

De La Fortelle, E. & Bricogne, G. (1997). In *Methods in Enzymology. Macromolecular Crystallography*, edited by R. M. Sweet & C. W. Carter Jr. New York: Academic Press.

Dumas, C. (1988). Thèse de doctorat en sciences naturelles. Université Paris-Sud, Centre Scientifique d'Orsay, France.

Evans, P. R. (1993). *Proceedings of the CCP*4 *Study Weekend: Data Collection and Processing*, pp. 114–122. Warrington: Daresbury Laboratory.

Fourme, R., Shepard, W. & Kahn, R. (1996). *Prog. Biophys. Mol. Biol.* **64**, 167–199.

Fourme, R., Shepard, W., Kahn, R., L'Hermite, G. & Li De La Sierra, I. (1995). *J. Synchrotron Rad.* **2**, 36–48.

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

Hendrickson, W. A. (1991). *Science*, **254**, 51–58.

Hendrickson, W. A., Smith, J. L., Phizackerley, R. P. & Merritt, E. A. (1988). *Proteins*, **4**, 77–88.

Huber, R. & Schneider, M. (1985). *J. Appl. Cryst.* **18**, 165–169.

Leslie, A. G. W. (1992). *MOSFLM Users Guide*, Version 4.30. MRC Laboratories, Cambridge, England.

Li De La Sierra, I., Pernot, L., Prange, T., Saludjian, P., Schiltz, M. Fourme, R. & Padron, G. (1997). *J. Mol. Biol.* **269**, 129–141.

Moras, D., Lorber, B., Romby, P., Ebel, J.-P., Giegé, R., Lewit-Bentley, A. & Roth, M. (1983). *J. Biol. Struct. Dyn.* **1**, 209–223.

Navaza, J. (1994). *Acta Cryst.* A**50**, 157–163.

Pebay-Peroula, E., Garavito, R. M., Rosenbusch, J. P., Zulauf, M. & Timmins, P. A. (1995). *Structure*, **3**, 1051–1059.

Pepinski, R. & Okaya, Y. (1956). *Proc. Natl Acad. Sci. USA*, **42**, 286–292.

Phillips, S. E. V. (1980). *J. Mol. Biol.* **142**, 531–554.

Podjarny, A., Bhat, T. N. & Zwick, M. (1987). *Annu. Rev. Biophys. Biophys. Chem.* **16**, 351–374.

Rey, F., Jenkins, J., Janin, J., Lasters, I., Alard, P., Matthyssens, G. & Wodak, S. (1988). *Proteins*, **4**, 165–172.

Rossmann, M. G. (1961). *Acta Cryst.* **14**, 383–388.

Rossmann, M. G., Leslie, A. G. W., Abdel-Meguid, S. S. & Tsukihara, T. (1979). *J. Appl. Cryst.* **12**, 570–581.

Roth, M., Lewit-Bentley, A. & Bentley, G. A. (1984). *J. Appl. Cryst.* **17**, 77–84.

Sasaki, S. (1984). In *Anomalous Scattering Factors for Synchrotron Radiation Users Calculated using the Cromer and Liberman Method*. National Laboratory for High Energy Physics, KEK, Tsukuba, Japan.

Schoenborn, B. P. (1988). *J. Mol. Biol.* **201**, 741–749.

Teng, T. Y. (1990). *J. Appl. Cryst.* **23**, 387–391.

Timmins, P. A., Poliks, B. & Banaszak, L. (1992). *Science*, **257**, 652–655.

Winkler, F. K., Schutt, C. E. & Harrisson, S. C. (1979). *Acta Cryst.* A**35**, 901–911.