The Multiwavelength Anomalous Solvent Contrast (MASC) Method in Macromolecular Crystallography

Roger Fourme, William Shepard, Richard Kahn,* Guillaume l'Hermite† and Inès Li de La Sierra

LURE, Batiment 209D, Université Paris-Sud, 91405 Orsay, France

The wavelength dependence of anomalous scattering of X-rays, due to atoms randomly dispersed in the solvent phase of a macromolecular crystal, is a way of producing solvent-density contrast variation with perfect isomorphism. The largest contrast variations are obtained by tuning the X-ray wavelength near an absorption edge of the anomalous-scattering species. In this method, which we call MASC, the anomalous partial structure is an extended uniform electron density, in contrast to the few punctual ordered scatterers in the multiwavelength anomalous-dispersion (MAD) method. MASC is, in principle, applicable to the determination of the molecular envelope and of low-resolution structure-factor phases. Structure factors $^\lambda F(\pm\mathbf{h})$ of an anomalous pair of reflections are expressed as a sum of two terms independent and dependent on wavelength, respectively. Squaring $^\lambda F(\pm\mathbf{h})$ leads to a set of equations which can be solved to give $|G(\mathbf{h})|$ and $|^0F(\mathbf{h})|$, the modulus of the envelope and of the total 'normal' structure factors, respectively, and $\Delta\varphi = (\varphi_F - \varphi_G)$. The moduli $\{|G|\}$ behave like structure-factor amplitudes from small-molecule crystals, and the estimation of their phases can be carried out by statistical direct methods. Then, the phase of $^0F(\mathbf{h})$ and finally the conventional (*e.g.* in vacuum) protein structure factor $F_P(\mathbf{h})$ can be determined. As in the MAD method, the strength of MASC signals can be quantified by Bijvoet and dispersive ratios, for which practical expressions are derived in the case of zero contrast. The behaviour of these ratios at increasing resolution is discussed, using approximations for $\langle|G(\mathbf{h})|\rangle$ and $\langle|\Delta(\mathbf{h})|\rangle$, respectively, derived from Porod's law and assuming a random distribution of atoms in the solvent excluding volume. Expected values of anomalous ratios are calculated for a hypothetical MASC experiment based on the known three-dimensional structure of kallikrein A, using a solvent with 3.5 M ammonium selenate to ensure zero contrast, and wavelength tuning near the Se K-absorption edge. The main steps of a MASC experiment are discussed in the context of a MAD-like data collection optimized for accurate measurements of intensities of anomalous pairs at low resolution. Finally, the results of preliminary experiments on two protein crystals are reported. The first, a partial single-wavelength data collection, used anomalous scattering of selenium at the K edge and gave anomalous ratios with the expected behaviour. The second one, at three wavelengths, used anomalous scattering of ytterbium at the L_{III} edge. In this case, effects from solvent as well as from ordered lanthanide ions were demonstrated.

Keywords: contrast variation; low-resolution phase determination; multiple-wavelength anomalous solvent contrast; molecular envelope determination.

1. Introduction

The atomic scattering factor for X-rays includes the 'normal' component f^0, a real number which is the Fourier transform of the electron density of the atom, and a wavelength-dependent increment $^\lambda\delta$ which is the manifestation of the phenomenon called anomalous dispersion. This increment is a complex number with real and imaginary components denoted $^\lambda f'$ and $^\lambda f''$, respectively. Consequently, the total scattering factor is,

$$^\lambda f = f^0 + {}^\lambda f' + i{}^\lambda f'' = f^0 + {}^\lambda\delta. \tag{1}$$

Considering (1), the scattering from a real atom located at

r can be formally described as the sum of the scattering from two emitters at **r**: (i) an atom with scattering factor f^0 (normal scattering) and (ii) a punctual source with scattering factor $^\lambda\delta$ (anomalous scattering, which, as assumed usually, does not depend on resolution). Both the magnitude and the phase of diffracted beams can be modified to some extent merely by tuning the X-ray wavelength in the appropriate range, without any change of the crystal structure.

Anomalous scattering can be used to derive phase information in a variety of ways. In the most common category of applications, the basic structure (*B*) to be determined contains many ordered atoms, *e.g.* light atoms in a protein, for which only normal scattering is detectable. A few ordered atoms with detectable anomalous scattering (structure *A*) are – or can be – inserted into the basic structure. In

---

the case of single-wavelength intensity measurements (SIRAS method), data sets from two crystals of types (B) and $(B + A)$, respectively, are in principle sufficient for an unambiguous phase estimation, provided that the two crystals are fairly isomorphous. If measurements are performed at several wavelengths (MAD method), a single crystal of $(B + A)$ type is sufficient, thereby ensuring strict isomorphism. The MAD method has now been successfully applied in a number of determinations of macromolecular structure (for reviews see, for example, Fourme & Hendrickson, 1990; Hendrickson, 1991; Smith, 1991). When the crystal contains not only ordered molecules but also solvent, a phase with rapidly rearranging molecules, X-ray anomalous dispersion due to atoms dispersed in the solvent can be used as a way of producing contrast variation. This second category of applications, which is still in its infancy, is the subject of this article.

The basic arguments and the precise description of what later became known as contrast variation in small-angle scattering (Stuhrmann & Kirste, 1965) have been given by Bragg & Perutz (1952). They showed that intensities of low-resolution reflections from a haemoglobin crystal could be changed by altering the electron density of the crystal mother liquor. They noted that changes in centrosymmetric structure-factor amplitudes are proportional to the Fourier transform of solvent-accessible volumes in the crystal. Contrast-variation intensity changes provide information only about the macromolecular boundary. The resulting structure-factor differences were interpreted as amplitudes for the transform of an ellipsoidal envelope, and phases were estimated for centric reflections. Harrison and co-workers combined X-ray contrast-variation data with a three-dimensional reconstruction of the TBSV particle obtained by electron microscopy (Harrison, 1969; Jack, Harrison & Crowther, 1975). The contrast dependence of solution scattering on scattering functions arising separately from the solute shape and its internal fluctuations was explicitly related by Ibel & Stuhrmann (1975), who also showed that these functions could be recovered from a contrast-variation series. This analysis, and the use of isotopic substitution (H/D exchange) as a powerful technique for contrast variation, made it possible to use contrast-variation neutron diffraction measurements in solving macromolecular crystal structures at low resolution (Roth, Lewit-Bentley & Bentley, 1984; Bentley, Lewit-Bentley, Finch, Podjarny & Roth, 1984; Moras et al., 1983; Podjarny, Bhat & Zwick, 1987). Bricogne (unpublished work) has derived a formulation of contrast variation in which diffraction effects due to the macromolecular envelope are explicitly separated from those due to internal density fluctuations. For each low-resolution reflection, three parameters are derived from X-ray intensities measured from crystals equilibrated in mother liquors of different electron densities: the modulus of the envelope transform $G(h)$, and components X and Y (the latter with an ambiguity about the sign), relative to $G(h)$, of the structure-factor vector for the transform of intramolecular density fluctuations $\Delta(h)$. Sim-

ulations and X-ray contrast-variation measurements were performed on the basis of this formalism (Carter, Bricogne & Dumas, 1986; Dumas, 1988), which was also used in a direct phase determination for the molecular envelope of tryptophanyl-tRNA synthetase (Carter, Crumley, Coleman, Hage & Bricogne, 1990).

Anomalous dispersion is a way of producing contrast variation. It is different from other contrast-variation methods such as chemical or isotopic substitution, solvent exchange and nuclear spin-dependent scattering, since it contains an imaginary part of the scattering amplitude. For nuclear-resonant neutron scattering, which is described by the Breit–Wigner formula, there are only a few nuclei that show a resonance in the thermal neutron energy range (e.g. ^{133}Cd, ^{149}Sm, ^{157}Gd). In the case of X-ray anomalous scattering, the choice of elements is much wider, as discussed in the context of MAD applications (e.g. Hendrickson, 1991). Anomalous X-ray small-angle scattering has been measured from iron in ferritin (Stuhrmann, 1980), terbium in parvalbumin (Miake-Lye, Doniach & Hodgson, 1983), sulfur in crystalline purple membrane (Munk, 1988) and bacteriophages (Kühnholz, 1991), and phosphorus in ribosomes (Hütsch, 1993). There are few references concerning the measurement and the exploitation of Bijvoet differences due to anomalous-scattering effects of the solvent in macromolecular crystals, and they refer essentially to the single-wavelength case. Such differences have been observed by H. W. Wyckoff, who also suggested the use of anomalous scattering in the solvent (unpublished), Dumas (1988), Crumley (1989) and Carter et al. (1990). Bricogne and co-workers (Carter et al. 1986, 1990; Dumas, 1988) have pointed out that it is possible to take advantage of Bijvoet differences to supplement standard contrast-variation measurements. In a broader perspective, the use of anomalous scattering from solvent has been integrated in the 'Bayesian programme' which should lead to an integrated phasing procedure based on entropy maximization and likelihood ranking (Bricogne, 1993). Data have been collected from ribosomal crystals by Yonath and coworkers (of the whole ribosome, its complex with mRNA and tRNA, and its small and large subunits). The crystals were immersed in solutions containing up to 1 M gold thioglucose or ammonium selenate, corresponding to electron densities ranging from 0.37 to 0.45 e Å^{-3}, at one wavelength close to the L-absorption edges of gold; the significant Bijvoet differences have been exploited for extracting the structure factors of the envelope transform, and the feasibility of phasing of these data by direct methods is being assessed (A. Yonath, private communication). With respect to multiple-wavelength measurements, the possibility of using anomalous scattering by the solvent at several wavelengths was first explicitly mentioned by Bricogne (1993). A MAD-like analysis, using as a starting point the formulation used by Carter et al. (1990) to exploit X-ray contrast-variation data, was given by Fourme (1993), including expressions for the calculation of expected Bijvoet and dispersive ratios. We call this method MASC, an acronym for multiwavelength

anomalous solvent contrast, as suggested by one of the authors (WS) of the present article. The possibility of using multiple-wavelength anomalous-scattering effects from solvent in a way similar to MAD was independently pointed out by Yonath (1993). In this article, in contrast to the initial presentation (Fourme, 1993), a reciprocal rather than direct-space approach is preferred for the derivation of basic MASC equations. We found this approach somewhat more convenient to discuss the effects of solvent and of disorder affecting both the molecular envelope and protein atoms. The main steps of a MASC experiment are then discussed with reference to the experience gained from MAD measurements. Finally, the results of preliminary experiments on crystals of two proteins with an anomalous scattering species dispersed in the mother liquor are reported.*

2. Derivation of structural information from MASC measurements

2.1. Model description and application to the normal solvent-scattering case

The measurement of the integrated intensity of reflection h leads to a value proportional to the squared modulus of the time-averaged structure factor $F(\mathbf{h})$. This is a classical result, which is still valid at low resolution. (A complete derivation, showing in particular that the contribution of diffuse scattering in this resolution domain is negligible, has been deposited with the IUCr as supplementary material).† For a protein crystal, the time-dependent structure factor,

$$F(\mathbf{h},t) = \sum_{j=1}^{n} f_j \exp\left[i2\pi\mathbf{h}\cdot\mathbf{r}_j(t)\right], \quad (2)$$

can be written as the sum of two components $F_S(\mathbf{h},t)$ and $F_P(\mathbf{h},t)$, from the solvent and from the protein, respectively. Molecules in the solvent can move quasi-freely like in a liquid, while in the protein region atomic motions are limited about mean atomic positions in the unit cell. Contributions from F_S and F_P are clearly correlated because the solvent volume and the protein volume are mutually exclusive.

In the following, the time-averaged value of $F(\mathbf{h},t)$ is obtained in two steps. First, protein atoms are frozen at given positions and the solvent contribution is averaged for this particular protein configuration. Then, a final averaging on all possible instantaneous protein configurations leads to the average value, $F(\mathbf{h})$, of $F(\mathbf{h},t)$.

A particular protein configuration defines a protein domain, $D_P(t)$, which in turn appears as an exclusion domain for the solvent. The protein volume, U, is considered as a constant. Solvent molecules are supposed to be

* Definitions and notations: s denotes the scattering vector $(s = 2\sin\theta/\lambda)$; $d = 1/s$; h is a vector of the reciprocal lattice. $\langle x \rangle$ is the root-mean-square value of x, $\bar{x}$ the expected value and x^* the complex conjugate of x.

† A complete derivation has been deposited with the IUCr (Reference: HI007). Copies may be obtained through The Managing Editor, International Union of Crystallography, 5 Abbey Square, Chester CH1 2HU, England.

randomly distributed in the complementary domain, $D_S(t)$, of the cell C. The volume of $D_S(t)$ is $V_S = V - U$. The probability distribution for the position of one solvent molecule is thus $1/V_S$ anywhere inside $D_S(t)$ and 0 inside $D_P(t)$. In general, the solvent contains several molecular species, k, each of which has N_k molecules randomly distributed in the solvent volume V_S, which corresponds to respective concentrations $c_k = N_k/V_S$; the average contribution from the solvent region is thus,

$$\sum_{j=1}^{n_s} f_j \frac{1}{V_s} \int_{D_S(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r} = \sum_k c_k f_k \int_{D_S(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r}.$$

Taking into account that

$$\int_{D_S(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r} = \int_V \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r}$$
$$- \int_{D_P(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r},$$

and since the sum over the whole crystal of cell contributions, $\int_V \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r}$, vanishes except at the origin of reciprocal space ($\mathbf{h} = 0$), the average solvent contribution at h for the particular protein configuration is therefore,

$$-\sum_k c_k f_k \int_{D_P(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r} = -\sum_k c_k f_k\, G(\mathbf{h},t), \quad (3)$$

where $G(\mathbf{h},t)$ is defined as

$$G(\mathbf{h},t) = \int_{D_P(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})d^3\mathbf{r}. \quad (4)$$

$G(\mathbf{h},t)$ is the inverse Fourier transform at h of the indicator function of the solvent excluding volume, $\chi(t)$, which takes the value 1 within $D_P(t)$ and 0 elsewhere (Bricogne, 1974).

Defining $F_P(\mathbf{h})$ and $G(\mathbf{h})$ as the average value of $F_P(\mathbf{h},t)$ and $G(\mathbf{h},t)$, respectively, the average value, $F(\mathbf{h})$, of $F(\mathbf{h},t)$ is thus,

$$F(\mathbf{h}) = F_P(\mathbf{h}) - \sum_k c_k f_k\, G(\mathbf{h}). \quad (5)$$

At low resolution, and when the solvent contains no anomalous scatterers,

$$\sum_k c_k f_k \cong \rho_s, \quad (6)$$

where ρ_s is the mean electron density of the solvent; then, from (5),

$$F(\mathbf{h}) = F_P(\mathbf{h}) - \rho_s G(\mathbf{h}). \quad (7)$$

The average value of the indicator function departs from the extreme values 0 or 1 at the interface between protein and solvent regions. A simple model taking into account the breadth of the boundary is obtained by convoluting a fixed indicator function χ_0 by a normalized isotropic Gaussian distribution $1/[(2\pi)^{3/2}\sigma^3]\exp[-(\Delta r^2)/(2\sigma^2)]$. Defining $B = 8\pi^2\sigma^2$, the inverse Fourier transform of the average value of the indicator function is therefore,

$$G(\mathbf{h}) = G_0(\mathbf{h}) \exp(-Bs^2/4), \quad (8)$$

where $G_0(\mathbf{h})$ is the Fourier transform at h of χ_0.

At very low resolution, $F_P(\mathbf{h},t)$ tends towards the number of electrons within the protein domain,

$$F_P(\mathbf{h},t) \simeq \bar{\rho}_P V_P, \qquad (9)$$

where $\bar{\rho}_P$ denotes the mean value of the electron density, $\rho_P(\mathbf{r},t)$, within the protein domain. Since $\rho_P(\mathbf{r},t)$ can be considered as taking the value $\rho_P(\mathbf{r},t) = \bar{\rho}_P + [\rho_P(\mathbf{r},t) - \bar{\rho}_P]$ inside $D_P(t)$ and $\rho_P(\mathbf{r},t) = 0$ elsewhere, $F_P(\mathbf{h},t)$ may be written at any resolution as,

$$
\begin{aligned}
F_P(\mathbf{h},t) &= \bar{\rho}_P \int_{D_P(t)} \exp(i2\pi\mathbf{h}\cdot\mathbf{r})\mathrm{d}^3\mathbf{r} + \Delta(\mathbf{h},t) \\
&= \bar{\rho}_P G(\mathbf{h},t) + \Delta(\mathbf{h},t), \qquad (10)
\end{aligned}
$$

where $\Delta(\mathbf{h},t)$ is the Fourier transform of the internal electron-density fluctuations $[\rho_P(\mathbf{r},t) - \bar{\rho}_P]$. The average value $F_P(\mathbf{h})$ of $F_P(\mathbf{h},t)$ is,

$$F_P(\mathbf{h}) = \bar{\rho}_P G(\mathbf{h}) + \Delta(\mathbf{h}). \qquad (11)$$

Finally,

$$F(\mathbf{h}) = (\bar{\rho}_P - \rho_s)G(\mathbf{h}) + \Delta(\mathbf{h}). \qquad (12)$$

$\Delta(\mathbf{h})$ includes implicitly the disorder of protein atoms expressed as usual, i.e. generally by means of isotropic temperature factors. The term $(\bar{\rho}_P - \rho_s)$ is called the contrast (Stuhrmann & Kirste, 1965). At zero contrast (matching point), the structure factor reduces to the fluctuation term. Expression (12) has already been derived by Carter et al. (1990), using a Fourier transformation of electron density, with a slightly different treatment of disorder.

The model which leads to (12) is simple and makes a number of approximations (solvent of uniform electron density, description of excursions of the molecular boundary with a single parameter, etc.). Phillips (1980) used a model with essentially the same physical content for a careful refinement of the structure of oxymyoglobin. Only two parameters were adjusted by trial and error: the contour level ρ_e corresponding to the edge of the solvent continuum, and the 'B value' applied to smooth the edge of the model solvent electron-density boundary. The final B value used for solvent corrections was 40 Å^2. Taking into account solvent corrections restored an excellent agreement between low-resolution observed and calculated structure-factor amplitudes. Myoglobin is a small protein with highly ordered solvent, and its B value is likely to be on the lower border. But even in cases of structures with looser packing and a higher solvent content, $\exp(-Bs^2/4)$ will remain close to 1 in the low-resolution range which is considered here. For example, at a resolution of 10 Å and for $B = 100$ Å^2, $\exp(-Bs^2/4) = 0.78$.

2.2. Application to the anomalous solvent scattering case

It is now assumed that anomalous scatterers of a single species are randomly dispersed in solvent and that there are no ordered anomalous scatterers attached to the macromolecules. In the low-resolution range, the scattering factor $f = f^0 + {}^\lambda f' + i^\lambda f''$ of these atoms can be considered as constant

to a first-order approximation. Formally, they contribute to the solvent electron density by a complex wavelength-dependent density noted ${}^\lambda \rho_{sA}$,

$$
{}^\lambda \rho_{sA} = {}^0\rho_{sA}\left(1 + \frac{{}^\lambda f'}{f^0} + i\frac{{}^\lambda f''}{f^0}\right), \qquad (13)
$$

where ${}^0\rho_{sA}$ is the normal part of the electron density of anomalous scatterers.

The solvent electron density can be resolved into two components,

$$
{}^\lambda \rho_s = {}^0\rho_s + {}^0\rho_{sA}\left(\frac{{}^\lambda f'}{f^0} + i\frac{{}^\lambda f''}{f^0}\right), \qquad (14)
$$

where ${}^0\rho_s$ is the total (i.e. taking into account the non-resonant contribution from anomalous scatterers) normal electron density of the solvent. Equation (12) becomes,

$$
\begin{aligned}
{}^\lambda F(\pm\,\mathbf{h}) = &\{\Delta(\mathbf{h}) + (\bar{\rho}_P - {}^0\rho_s)G(\mathbf{h})\} \\
&+ \left\{\left(\frac{{}^\lambda f'}{f^0} \pm i\frac{{}^\lambda f''}{f^0}\right)[-{}^0\rho_{sA}G(\mathbf{h})]\right\}. \qquad (15)
\end{aligned}
$$

The first term between brackets in (15) is the part of the structure factor, noted ${}^0F(\mathbf{h})$, which is λ-independent; it incorporates the fluctuation term and an envelope term. The second term between brackets is wavelength-dependent. Note that $-G(\mathbf{h})$ is the Fourier transform at $\mathbf{h}$ ($\mathbf{h} \neq 0$) of the indicator function for the solvent-accessible volume, which has a value of 1 within this region and 0 elsewhere. Defining,

$$\Gamma(\mathbf{h}) = -{}^0\rho_{sA}G(\mathbf{h}), \qquad (16)$$

then, from (15),

$$
{}^\lambda F(\pm\,\mathbf{h}) = {}^0F(\mathbf{h}) + \left(\frac{{}^\lambda f'}{f^0} \pm i\frac{{}^\lambda f''}{f^0}\right)\Gamma(\mathbf{h}). \qquad (17)
$$

This equation separates λ-dependent and λ-independent contributions to the total structure factor, following the seminal idea of Karle (1980), and the resolution has the same form as that used for the algebraic MAD method (Hendrickson, 1985) in the case of a single anomalous-scattering species,

$$
{}^\lambda F_T(\pm\,\mathbf{h}) = {}^0F_T(\mathbf{h}) + \left(\frac{{}^\lambda f'}{f^0} \pm i\frac{{}^\lambda f''}{f^0}\right){}^0F_A(\mathbf{h}), \qquad (18)
$$

where ${}^\lambda F_T$ is the total structure factor, ${}^0F_A$ is the normal scattering component from the partial structure A and ${}^0F_T$ is the total normal contribution (including ${}^0F_A$). That $\Gamma(\mathbf{h})$ plays the same role as ${}^\lambda F_A(\mathbf{h})$ in (18) has an obvious physical meaning: the anomalous partial structure (A) is now an extended uniform electron density, in contrast to the few punctual ordered scatterers used in the MAD method.

Multiplication of ${}^\lambda F(\pm\,\mathbf{h})$ (17) by its complex conjugate gives the following equations,

$$
|{}^\lambda F(\pm\,\mathbf{h})|^2 = |{}^0F|^2 + a(\lambda)|\Gamma|^2 + b(\lambda)|{}^0F||\Gamma|\cos(\Delta\phi)
$$

$$
\pm\, c(\lambda)|{}^0F||\Gamma|\sin(\Delta\phi), \qquad (19)
$$

where $a(\lambda) = ({}^\lambda f'^2 + {}^\lambda f''^2)/(f^0)^2$, $b(\lambda) = 2({}^\lambda f'/f^0)$ and $c(\lambda) = 2({}^\lambda f''/f^0)$ and $\Delta\phi$ is the phase difference $(\varphi^0_F - \varphi_\Gamma)$. Then, among various possible approaches, it is possible to use the

algebraic method of Hendrickson (1985), by which the set of equations pertaining to the anomalous pair ($\pm h$) can be solved for $|^0F|$, $|\Gamma|$ and $\Delta\phi$. Finally, taking into account the relationship between G and Γ, (16), one gets for each h: $|^0F|$, $|G|$ and $\Delta\varphi$.

2.3. Phase determination for the $\{|G|\}$ structure-factor amplitudes for the molecular envelope

As pointed out by Bricogne (unpublished) and applied successfully by Carter et al. (1990), the set of $\{|G|\}$ exhibit several properties similar to those that make small-molecule data sets amenable to solution by standard direct methods. (i) They represent a structure much simpler than the macromolecular structure. (ii) They have large normalized moduli. (iii) Due to the sharp decline in the mean amplitude of $|G|$ with increasing resolution, the data set $\{|G|\}$ is essentially complete at very moderate resolution limits. (iv) Since the χ_0 function is equal to its square, its sampled transform satisfies Sayre's equation (Sayre, 1952). Once phases of $\{|G|\}$ have been estimated by direct methods, an initial molecular envelope can be determined. Phase refinement can be carried out using additional information such as non-negativity of the indicator function, non-crystallographic symmetry, etc. This procedure is similar to the corresponding one in the MAD method, which includes the solution of structure (A) by Patterson or direct methods followed by a refinement of the parameters of this structure which provides a more accurate set of $\{^0F_A\}$. Then, phases of the $\{^0F\}$ are calculated from $\varphi_{0F} = \Delta\varphi + \varphi_G$. Finally, from the sets $\{^0F\}$ and $\{G\}$, the conventional protein structure factors $\{F_P\}$ (i.e. corresponding to atoms in vacuum) can be calculated as,

$$F_P(h) = {}^0F(h) + {}^0\rho_s G(h). \tag{20}$$

3. Strength of anomalous signals in the MASC method

3.1. Anomalous ratios

We define,

$$|\Delta F_{\pm h}| = \left| |{}^\lambda F(h)| - |{}^\lambda F(-h)| \right|$$

$$|\tilde{F}| = \tfrac{1}{2}[|{}^\lambda F(h)| + |{}^\lambda F(-h)|]$$

$$|\Delta F_{\Delta\lambda}| = \left| |{}^{\lambda_i}\tilde{F}| - |{}^{\lambda_j}\tilde{F}| \right|$$

$$\Delta f' = |{}^{\lambda_i}f' - {}^{\lambda_j}f'|.$$

As in the MAD method (Hendrickson, 1985), the strength of MASC signals can be quantitated by anomalous and dispersive ratios calculated from (19). But, in the MASC case, these ratios are expected to vary considerably with resolution (by perhaps an order of magnitude), being largest at very low resolution and decreasing rapidly at higher resolution. The resolution domain where the MASC method is likely to bring useful phase information is directly probed by the values of anomalous ratios. For this reason, the following discussion is focused on the asymptotic behaviour of anomalous ratios when increasing resolution. Accord-

ingly, expressions (21) to (26) are calculated with approximations based on the assumption that the anomalous contribution in (15) is relatively small: $|F(h)|$ can then be considered as wavelength-independent and is approximated by $|{}^0F(h)|$ in the denominator of the right-hand side of the various expressions. For the sake of simplicity, these equations will be used for the analysis of the whole data irrespective of the resolution, keeping in mind that the approximation deteriorates as $s \to 0$.

(i) Bijvoet ratio at a single wavelength λ,

$$\frac{\langle|\Delta F_{\pm h}|\rangle}{\langle|\tilde{F}|\rangle} \simeq \frac{(2^\lambda f'')}{f^0} \frac{\langle|\Gamma|\rangle}{\langle|{}^0F|\rangle} \langle|\sin(\Delta\phi)|\rangle. \tag{21}$$

(ii) Dispersive ratio involving two wavelengths λ_i and λ_j,

$$\frac{\langle|\Delta F_{\Delta\lambda}|\rangle}{\langle|\tilde{F}|\rangle} \simeq \frac{\Delta f'}{f^0} \frac{\langle|\Gamma|\rangle}{\langle|{}^0F|\rangle} \langle|\cos(\Delta\phi)|\rangle. \tag{22}$$

We point out the interest of adjusting the composition of the solvent so that the total normal electron density of the solvent matches the average density of the ordered part of the structure ($^0\rho_s = \bar{\rho}_p \simeq 0.42$ e Å^{-3}). In (15), the envelope term vanishes, leaving only $\Delta(h)$; the consequences are the following.

$\langle|{}^0F|\rangle$ in each resolution shell is minimum. This is beneficial for data collection, because the dynamical range of intensities on the detector is reduced. Further, for a given concentration of anomalous scatterers in the solvent (i.e. for a given value of $^0\rho_{sA}$), $\langle|\Gamma|\rangle/\langle|{}^0F|\rangle$, and accordingly the anomalous ratios, are maximized.

The phases of $^0F(h)$ and $\Gamma(h)$ are least correlated, because $G(h)$ no longer contributes to $^0F(h)$ [equation (15)] [phases of $\Delta(h)$ and $G(h)$ are certainly not completely uncorrelated since the domain of electron-density fluctuations is bounded by the molecular envelope; this point will be investigated using simulations on known protein structures]. Then, we assume, to a first approximation, that $\langle|\cos(\Delta\phi)|\rangle = \langle|\sin(\Delta\phi)|\rangle = 1/(2)^{1/2}$.

Assuming that the matching point is achieved, one obtains expressions (23) to (26),

$$\frac{\langle|\Delta F_{\pm h}|\rangle}{\langle|\tilde{F}|\rangle} \simeq \frac{1}{(2)^{1/2}} \rho_{sA} \frac{(2^\lambda f'')}{f^0} \frac{\langle|G|\rangle}{\langle|\Delta|\rangle}, \tag{23}$$

$$\frac{\langle|\Delta F_{\Delta\lambda}|\rangle}{\langle|\tilde{F}|\rangle} \simeq \frac{1}{(2)^{1/2}} \rho_{sA} \frac{\Delta f'}{f^0} \frac{\langle|G|\rangle}{\langle|\Delta|\rangle}. \tag{24}$$

Since,

$$^0\rho_{sA} = NMf^0 \times 10^{-27} = 6.02 \times 10^{-4} Mf^0,$$

with $N = $ Avogadro's number and $M = $ molarity of the anomalous-scattering species (mol l^{-1}), we obtain practical expressions,

$$\frac{\langle|\Delta F_{\pm h}|\rangle}{\langle|\tilde{F}|\rangle} \simeq 4.26 \times 10^{-4} M(2^\lambda f'') \frac{\langle|G|\rangle}{\langle|\Delta|\rangle}, \tag{25}$$

$$\frac{\langle|\Delta F_{\Delta\lambda}|\rangle}{\langle|\tilde{F}|\rangle} \simeq 4.26 \times 10^{-4} M\Delta f' \frac{\langle|G|\rangle}{\langle|\Delta|\rangle}. \tag{26}$$

As in the MAD method, the expected accuracy in phase determination is related to the magnitudes of Bijvoet and dispersive ratios, which should both be large (for a quantitative evaluation, see Narayan & Ramaseshan, 1981). Since these ratios are proportional to $2f''$ and $\Delta f'$, respectively, wavelengths close to an absorption edge of the anomalous scatterers should be selected as is commonly carried out in MAD experiments: λ_1 at f'' maximum (peak); λ_2 at $|f'|$ maximum (edge); λ_3 remote from, and on the high-energy side of, the edge; and possibly another remote wavelength λ_4 on the low-energy side of the edge. Useful anomalous scatterers are not restricted to heavy atoms: for K-edge studies, all elements whatever their atomic number Z are essentially alike; the same is true for L edges. Obviously, specific features of each edge should be exploited to increase anomalous ratios, e.g. white lines at the K edge of selenium or L_{III} edge of lanthanides.

These ratios are also proportional to M, the concentration of the anomalous species. In practice, the maximum value of M will be set by the specific response of each crystal to the addition of the anomalous scatterer in the mother liquor or by the solubility of the compound containing anomalous scatterers. If, for this concentration, the total normal electron density of the solvent is still below the average protein electron density, the solvent electron density could be increased by addition of, for example, sucrose in order to reach the matching point, thus maximizing anomalous ratios. Finally, ratios are proportional to $\langle |G| \rangle = \langle |G_0| \rangle \exp(-Bs^2/4)$ and inversely proportional to $\langle |\Delta| \rangle$. It is of interest to discuss the behaviour of $\langle |\Delta| \rangle$ and $\langle |G_0| \rangle$ at increasing resolution, i.e. where the MASC signals are weakest, because it will set the practical limits of the MASC method as a tool for the phase problem.

3.2. Behaviour of anomalous ratios at increasing resolution

We have obtained an approximation for the expected value, $\overline{|\Delta(\mathbf{h})|^2}$, of $|\Delta(\mathbf{h})|^2$ assuming a random distribution of atoms in the ordered domain U (the complete derivation has been deposited with the IUCr as a supplementary publication),*

$$\overline{|\Delta(\mathbf{h})|^2} = n_t f_{\text{eff}}^2 \left\{ \left[\sum_g \exp(-i2\pi\,\mathbf{t_g}\cdot\mathbf{h}) \frac{G(\mathbf{h} - {}^t R_g \mathbf{h})}{U} \right] - n_g \frac{|G(\mathbf{h})|^2}{U^2} \right\},$$

(27)

where n_t is the number of atoms in the whole protein domain D_p in the unit cell, $f_{\text{eff}}^2 = (1/n_t)\sum_{j=1}^{n_t} |f_j|^2$ (for a protein, $f_{\text{eff}} \simeq 6.7$ e at $s = 0$). Taking into account the crystal symmetry, the n_t atoms within the whole protein domain are generated from a basic set of n atoms in the asymmetric unit. Those atoms are labelled using the index k. The position, $\mathbf{r}_j$, of the atom j is deduced, by the symmetry operation S_g, from the position, $\mathbf{r}_k$, of the atom k, by,

$$\mathbf{r}_j = \mathbf{t}_g + R_g \mathbf{r}_k,$$

(28)

where R_g is the orthogonal transformation (corresponding

---

* See supplementary material footnote.

transformation in reciprocal space, ${}^t R_g$) and $\mathbf{t}_g$ is a non-primitive translation related to the symmetry operation S_g.

When $\mathbf{h} \to 0$, $\overline{|\Delta(\mathbf{h})|^2} \to 0$, since $G(\mathbf{h}) \to U$. This is what is expected, because at very low resolution the diffraction pattern does not depend on electron-density fluctuations. At higher resolution, $G(\mathbf{h})$ tends rapidly towards zero, and, because $G(0) = U$, the only non-vanishing contribution to $\overline{|\Delta(\mathbf{h})|^2}$ comes from all terms in expression (27) with ${}^t R_g \mathbf{h} = \mathbf{h}$. The set of corresponding symmetry operations S_g define the isotropy subgroup of reflection $\mathbf{h}$ (Bricogne, 1991). Except if $\mathbf{h}$ is a systematically absent reflection, $\exp(-i2\pi\,\mathbf{t_g}\cdot\mathbf{h}) = 1$. Defining $m_{\mathbf{h}}$ as the number of elements of the isotropy subgroup of reflection $\mathbf{h}$ ($m_{\mathbf{h}} = 1$ for a general reflection in a primitive space group),

$$\overline{|\Delta(\mathbf{h})|^2} \simeq m_{\mathbf{h}} n_t f_{\text{eff}}^2,$$

(29)

$$\langle |\Delta(\mathbf{h})| \rangle \simeq (m_{\mathbf{h}} n_t)^{1/2} f_{\text{eff}}.$$

(30)

An approximation for $\langle |G_0(\mathbf{h})| \rangle$ at increasing values of s can be obtained on the basis of Porod's law (Porod, 1951), commonly used in small-angle scattering studies. When identical randomly oriented particles of uniform electron density are immersed in a phase of uniform electron density with a sharp interface between the two phases, the scattered intensity obeys a law which is asymptotically valid for increasing values of s,

$$I(s) \to (2\pi)^{-3} \Delta\rho A s^{-4},$$

(31)

where A is the area of the interface and $\Delta\rho$ is the difference between the electron densities of the two phases. The transposition of Porod's derivation to the model of the molecular envelope used in this article is straightforward, because this model features identical domains of uniform value ($\chi_0 = 1$), bounded by a closed surface (the molecular envelope), immersed in a phase of uniform value ($\chi_0 = 0$) and with a sharp interface between the two phases. $G_0(\mathbf{s})$ is the inverse Fourier transform at $\mathbf{s}$ of the indicator function of the solvent excluding volume. Because of the periodic arrangement in the crystal, G_0 is sampled at reciprocal lattice nodes ($\mathbf{s} = \mathbf{h}$). $\langle |G_0(\mathbf{h})| \rangle$ tends towards the following limit,

$$\langle |G_0(\mathbf{h})| \rangle \to (2\pi)^{-3/2} A_t^{1/2} s^{-2},$$

(32)

where A_t (in Å^2) is the area of the interface in the crystal unit cell and averaging is performed in resolution shells. Owing to the asymptotic nature of Porod's law, estimating the average power of $\langle |G_0(\mathbf{h})| \rangle$ with (32) is acceptable at moderately high resolution.

Using approximations (30) and (32), and as $A_t/n_t = A/n$ (where A and n are the area of the envelope and the number of atoms, respectively, for a single particle), then practical expressions for anomalous ratios are,

$$\frac{\langle |\Delta F_{\pm \mathbf{h}}| \rangle}{\langle |\tilde{F}| \rangle} \simeq 0.27 \times 10^{-4} \exp(-Bs^2/4) M \frac{2^i f''}{f_{\text{eff}}} \left(\frac{A}{n} \right)^{1/2} s^{-2}$$

(33)

$$\frac{\langle |\Delta F_{\Delta \lambda}| \rangle}{\langle |\tilde{F}| \rangle} \simeq 0.27 \times 10^{-4} \exp(-Bs^2/4) M \frac{\Delta f'}{f_{\text{eff}}} \left(\frac{A}{n} \right)^{1/2} s^{-2}.$$

(34)

3.2.1. *The case for globular proteins.* In the case of a protein with molecular weight M_w, n approximates to kM_w, with $k \simeq 0.07\,\mathrm{Da}^{-1}$. Furthermore, if the shape of the macromolecule is globular (where the shape can be approximated by a sphere), $A \simeq k'M_w^{2/3}$ with $k' \simeq 11.3\,\text{Å}^2$ $\mathrm{Da}^{-2/3}$ (values for k and k' were calculated using data for 12 globular proteins from Table II in Pavlov & Fedorov, 1983), then, from (33) and (34),

$$\frac{\langle|\Delta F_{\pm h}|\rangle}{\langle|\widetilde{F}|\rangle} \simeq 3.44 \times 10^{-4} \exp(-Bs^2/4)M\frac{2^{\lambda}f''}{f_{\mathrm{eff}}}\,(M_w^{1/12}s)^{-2} \quad (35)$$

$$\frac{\langle|\Delta F_{\Delta\lambda}|\rangle}{\langle|\widetilde{F}|\rangle} \simeq 3.44 \times 10^{-4} \exp(-Bs^2/4)M\frac{\Delta f'}{f_{\mathrm{eff}}}\,(M_w^{1/12}s)^{-2}. \quad (36)$$

In conclusion, in the resolution range where Porod's law holds, anomalous ratios are, for given values of other parameters, governed by the product $(M_w^{1/12}s)$. This provides an approximate 'scaling law' for ratios expected from globular proteins of different molecular weights.

3.3. *Expected values of anomalous ratios*

Assuming that the solvent electron density has been adjusted for contrast matching, (25) and (26) can be used to calculate expected anomalous ratios in a MASC experiment on a crystal with a known three-dimensional structure. In addition to f'', $\Delta f'$ and M values, $\langle|G|\rangle$ and $\langle|\Delta|\rangle$ values as a function of resolution are required. Such data have been calculated by Carter, Bricogne & Dumas (1986) for kallikrein A, a medium-sized globular protein ($M_w = 23.5\,\mathrm{kDa}$). The crystal structure was solved by Bode *et al.* (1983) at 2 Å resolution and structural parameters are available from the Protein Data Bank (PDB, Bernstein *et al.*, 1977). The space group is $P4_12_12$, the parameters of the unit cell are $a = 90.2$, $c = 159.4$ Å, and there are two molecules in the asymmetric unit. The kallikrein data sets were generated using atomic coordinates from the PDB, a B value of 10 Å^2 for each atom and included 124 water molecules; Fourier transforms were computed using a $1 \times 1 \times 1$ Å grid pitch. The molecular boundary, determined on the basis of these atomic coordinates, was submitted to smoothing according to a procedure which ensures, to first approximation, that the sum of the indicator maps of the protein and solvent regions is equal to 1 at each grid point. Since this procedure is roughly equivalent to convoluting by a three-dimensional Gaussian function as used in our model, values for $\langle|G|\rangle$ tabulated by Carter *et al.* (1986) were used without further correction. Expected anomalous ratios were calculated for a hypothetical MASC experiment on a kallikrein A crystal with a mother liquor containing ammonium selenate, and exploiting optimally the anomalous scattering of selenium close to the Se K-absorption edge. In the case of ammonium selenate, absorption measurements gave us $2f''_{\max} = 14.0\,\mathrm{e}$ and $\Delta f'_{\max} = 8.6\,\mathrm{e}$, and the selenate concentration ensuring contrast matching is about $3.5\,M$, depending of the exact mother liquor composition (§5.1.). Three wavelengths are selected close to the Se K edge, λ_1 at $f''_{\max}$, λ_2 at $|f'|_{\max}$ and λ_3 at a point remote from the edge on

Table 1

Expected anomalous ratios in a hypothetical MASC experiment on kallikrein A, using anomalous scattering from selenium in $3.5\,M$ ammonium selenate (ensuring contrast matching).

Resolution shell	d	$\langle|D|\rangle$	$\langle|G|\rangle$	$\langle|G|\rangle/\langle|D|\rangle$	$Br_{\max}$	$dr_{\max}$
0.0215–0.0289	39.68	456.5	17954.9	39.3	82.0	50.4
0.0289–0.0362	30.72	553.3	10481.6	18.9	39.5	24.2
0.0362–0.0436	25.06	475.8	7738.5	16.3	34.0	20.9
0.0436–0.0510	21.14	342.5	6678.8	19.5	40.7	25.0
0.0510–0.0584	18.28	365.0	4244.3	11.6	24.2	14.9
0.0584–0.0658	16.10	410.5	3018.5	7.4	15.4	9.5
0.0658–0.0732	14.39	504.1	3033.5	6.0	12.5	7.7
0.0732–0.0806	13.00	661.1	2564.9	3.9	8.1	5.0
0.0806–0.0880	11.86	704.0	2130.8	3.0	6.3	3.8
0.0800–0.0954	10.91	867.2	1667.2	1.9	4.0	2.4
0.0954–0.1028	10.09	767.9	1598.4	2.1	4.4	2.7
0.1028–0.1102	9.39	759.7	1489.8	2.0	4.2	2.6
0.1102–0.1176	8.78	670.1	1292.0	1.9	4.0	2.4
0.1176–0.1250	8.24	671.7	1197.7	1.8	3.8	2.3

Notes: Limits of the resolution shell (Å^{-1}); d (Å), resolution corresponding to the midpoint of the shell; $\langle|\Delta|\rangle$ (electrons), r.m.s. value $\langle|\Delta(h)|\rangle$ in the resolution shell; $\langle|G|\rangle$ (Å^3), r.m.s. value of $|G(h)|$ in the resolution shell; $Br_{\max}$ (%), values of Bijvoet ratios at $f''_{\max}$; $dr_{\max}$ (%), values of dispersive ratios at $\Delta f'_{\max}$. $\langle|\Delta|\rangle$ values are taken from Carter *et al.* (1986), who used coordinates of the atomic model of kallikrein A (Bode *et al.*, 1983). Statistics are on 829 independent reflections at 8 Å resolution.

the high-energy side of the edge. Values in resolution shells of the maximum Bijvoet and dispersive ratios expected for this particular experiment are listed in Table 1. The variation of Bijvoet ratios is shown in Fig. 1.

A few macromolecular structures have been solved by the MAD method with anomalous ratios in the range 2.3–3.5% (Hendrickson, 1991). If we consider these ratios as practical limits in a MASC experiment, then it can be concluded from results in Table 1 and Fig. 1 that useful phase information could be derived up to a resolution of about 8–10 Å for this 23.5 kDa protein, although a more realistic limit is probably about 12 Å. Solvent effects become huge at very low resolution.
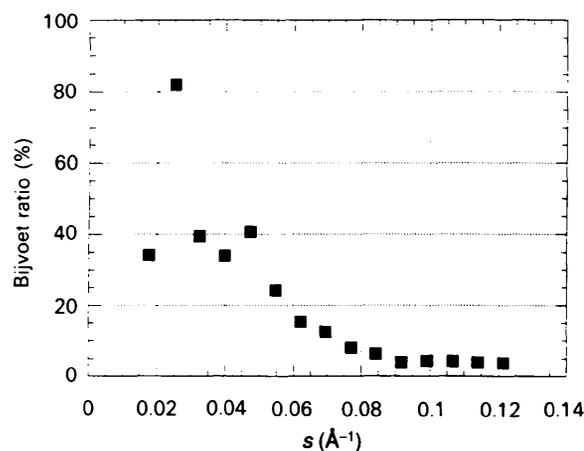
Figure 1

Bijvoet ratios expected in a hypothetical experiment using a crystal of kallikrein A with $3.5\,M$ ammonium selenate (ensuring contrast matching) at λ_{peak} ($2f'' = 14\,\mathrm{e}$). Values of ratios are obtained on the basis of equation (25), using values of $\langle|G|\rangle$ and $\langle|\Delta|\rangle$ calculated by Carter *et al.* (1986) from the known three-dimensional structure of kallikrein A.

Larger anomalous ratios – or alternatively, similar ratios for a lower concentration of the anomalous scattering species – may be obtained using the L-absorption edges of other species instead of the K edge. Resonances at the L_{III}-absorption edges of lanthanides are of particular interest, with f'' and $|f'|$ values reaching about 30 e when using a narrow bandpass monochromator (Templeton, Templeton, Phizackerley & Hodgson, 1982).

4. Experimental

4.1. Sample preparation

Since our practical experience is limited at the present stage, only a few tentative guidelines can be given. A sufficient supply of material and a protocol for the growth of good quality crystals are prerequisites. The anomalous scatterers are included either as part of the crystallization liquor, or added to the mother liquor once crystals have been grown. Ideally, the goals are as follows.

(a) Molar or multimolar concentrations of anomalous scatterers should be achieved without substantial degradation of the crystal.

(b) Resonant atoms should be randomly dispersed in the solvent. Hence, the group of atoms to which they are attached should be of fairly small volume.

(c) Resonant atoms should not have ordered attachment to the protein. This can be checked by inspection of Bijvoet differences at medium and high resolution, which should not differ significantly from zero. If attachment is observed, chemical tricks such as caging of resonant atoms might be useful.

(d) The selected absorption edge should preferably be at a wavelength as short as possible, in order to minimize X-ray absorption effects.

(e) If the matching point is not achieved for the maximal concentration of anomalous scatterers, the solvent electron density, as discussed previously, should be increased in order to reach it.

4.2. Instrumentation and methods for collection and reduction of MASC diffraction data

A MASC data collection is essentially a MAD experiment performed in the low-resolution range, so that requirements of a synchrotron X-ray source and a high-resolution tunable monochromator are the same for both experiments (Fourme & Hendrickson, 1990). In the case of MASC measurements, the setup must be equipped as a low-angle scattering instrument, with a beam-stop system subtending a very small solid angle, and careful reduction of various sources of parasitic scattering. At very low resolution, experimental Bragg intensities are on average quite strong due to the Lorentz factor, cover a wide dynamical range and peak over a fairly high background. Among detectors which are currently available, proportional chambers are attractive as photon counters with a very wide dynamical range and negligible instrumental noise; their high duty

cycle (the ratio of exposure time to total elapse time) allows a small rotation per frame, thus improving the signal-to-noise ratio (Kahn, Fourme, Bosshard & Saintagne, 1986). In the case of crystals with very large unit-cell parameters – which may be important in view of the potential of MASC in the determination of the envelope of complex macromolecular systems – gas detectors are generally not appropriate since their spatial resolution is limited; then, the current alternative is an imaging-plate system. CCD-based detectors (Westbrook, Deacon, Strauss, Naday & Sweet, 1989; Moy, 1991) also have characteristics (especially a good duty cycle), which make them promising for multiple-wavelength measurements. Using only one crystal for the whole data collection reduces systematic errors; further, as discussed later, the addition of a high concentration of anomalous scatterers in the solvent may reduce the lifetime of irradiated crystals. For these reasons, shock cooling of the sample and data collection at cryogenic temperatures should be most useful. The quasi-amorphous structure of the solvent in the liquid state should be essentially preserved in the quenched sample, so that the flat solvent hypothesis should remain valid; this point will be investigated in forthcoming studies.

The main steps of any multiwavelength experiment, including monochromator calibration, determination of the wavelength dependence of $^{\lambda}f''$ and $^{\lambda}f'$, wavelength selection, data collection, data analysis and data reduction are always essentially the same and have been thoroughly described (see e.g. Hendrickson, Smith, Phizackerley & Merritt, 1988; Fourme & Hendrickson, 1990). With respect to a MAD experiment, MASC data collection may be somewhat simpler and faster as low-resolution reflections are stronger and less sensitive to crystal degradation than high-resolution reflections; further, reflections at a resolution higher than 5–6 Å which are recorded together with low-resolution data are insensitive to anomalous-dispersion effects from the solvent and are quite useful to scale data sets at the various wavelengths. Absorption of X-rays by the solvent and its wavelength dependence is a matter of serious concern in MASC, due to molar or multimolar concentration of anomalous scatterers in the sample. As mentioned previously, the anomalous-scattering species will, as far as possible, be chosen in order to perform MASC data collection at short wavelengths; nevertheless, taking proper care of absorption will probably be a major difficulty in obtaining accurate MASC measurements, and sensible dispersive differences in particular.

5. Preliminary results on two test cases

5.1. The 64 kDa outer-membrane protein, P64K, of Neisseria meningitidis

P64K is a constituent of the outer membrane of Neisseria meningitidis, a bacterium which is responsible for one in every three cases of bacterial meningitis in the world (De Voe, 1982). It is present in the majority of pathogenic Neisseria strains and is capable of inducing immunologically

active antibodies in its natural host. It has been cloned and expressed in *Escherichia coli* (Silva *et al.*, 1992) so that it can be produced with a high level of purity and in useful amounts for vaccine preparations.

The crystallization and a preliminary investigation of P64K by X-ray diffraction methods have been reported by Li de La Sierra *et al.* (1993). Crystals are grown by the hanging-drop diffusion method, using reservoir solutions composed of 0.1 M potassium phosphate and 2 M ammonium sulfate (pH 7.0). The space group is $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = 136.84$ and $c = 78.44$ Å; there is one molecule in the asymmetric unit, with a solvent content of about 55%. Crystals diffract to at least 2.9 Å resolution using synchrotron radiation. The crystal structure is currently being investigated at LURE, and a preliminary model has been obtained for about 80% of the macromolecule.

Crystals grown as described previously were equilibrated first with a reservoir containing a solution of 3.4 M ammonium sulfate, then another reservoir containing 3.4 M ammonium selenate. Finally, crystals were immersed and kept for several days in a solution of 3.4 M ammonium selenate and 0.1 M potassium phosphate. From density measurements, the calculated electron density of this solution is 0.411 e Å^{-3}; assuming that the mean electron density of the protein is 0.417 e Å^{-3} (taken from Cantor & Schimmel, 1982), contrast matching was thus achieved approximately (a better value is 3.5 M, the concentration used for subsequent experiments as well as for calculations of expected anomalous ratios). A single-wavelength data set was collected on one crystal, using the W32 setup on the wiggler beamline of the positron storage ring DCI at LURE. This instrument features double-focusing optics and a single-axis goniometer with an imaging-plate system (a prototype instrument with a circular plate, diameter 180 mm,

built by J. Hendrix and A. Lentfer at the EMBL Outstation, Hamburg). This experiment was a step in the preparation of a full multiwavelength MASC experiment on this protein. It was performed in 1993 shortly before the winter shutdown of the storage ring DCI, in conditions which were not optimal in terms of crystal orientation, resolution range, completeness of data and strength of anomalous signals.

Because of the sample morphology, the crystal was oriented with a along the spindle axis instead of the optimal setting with c along this axis. Data-collection conditions and parameters were as follows: temperature 294 K; wavelength 0.903 Å; angular range 1.5° per image. The crystal-to-plate distance was set at 370 mm, giving a useful resolution range between 3.7 and 27 Å. When compared with crystals grown in ammonium sulfate, the lifetime of crystals under the wiggler beam is decreased by at least a factor of 2. We found later that this situation may be improved by exchanging ammonium sulfate with ammonium selenate only a few hours before data collection and keeping the crystal at 273 K during data collection. A better solution would be cryocooling, and this is currently under investigation. Because of crystal degradation, data obtained for a total rotation of only 21° were kept for analysis with the *MOSFLM* package (Leslie, Brick & Wonacott, 1986) and *CCP*4 (SERC Daresbury Laboratory, 1979) suite of programs. From these data, $\langle |\Delta F| \rangle / \langle |F| \rangle$ ratios in resolution shells were calculated for acentric pairs of reflections; the noise level was estimated by calculating similar ratios for centric reflections (Fig. 2). From $d \simeq 8$ Å to the maximum resolution ($d = 3.7$ Å), the anomalous signal is too weak with respect to the noise level to be detected. It can thus be concluded that the anomalous signals are due entirely to

Figure 2
Results of an effective data collection using a P64K crystal with 3.4 M ammonium selenate at $\lambda = 0.903$ Å ($2f'' = 6.6$ e). Breakdown, as a function of resolution, of experimental ratios $\langle |\Delta F_{\pm h}| \rangle / \langle |\tilde{F}| \rangle$ for pairs of acentric reflections (open squares) and for pairs of centric reflections (filled triangles). The latter give an estimate of the intrinsic accuracy of the measurements.
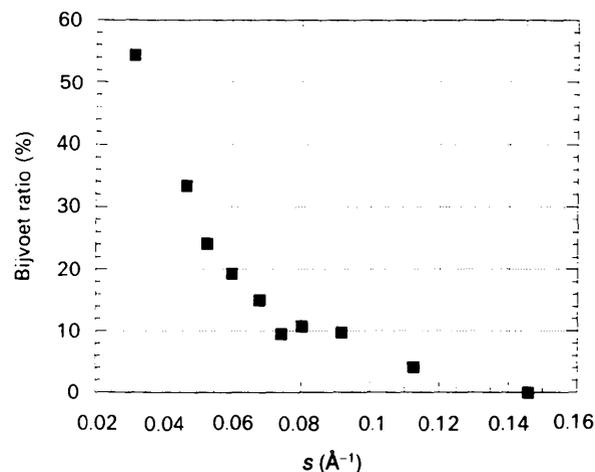
Figure 3
Bijvoet ratios expected in a hypothetical experiment using a crystal of P64K with 3.5 M ammonium selenate (ensuring contrast matching) at λ_{peak} ($2f'' = 14$ e). Values of ratios are extrapolated, on the basis of equation (25), from experimental values of $\langle |\Delta F_{\pm h}| \rangle / \langle |\tilde{F}| \rangle$ ratios for acentric reflections (Fig. 2) corrected for noise. Dispersive ratios expected using λ_{edge} and λ_{remote} data sets ($\Delta f' = 8.6$ e) would be obtained by multiplying these Bijvoet ratios by $8.6/14 = 0.641$.

solvent effects, and that there are no ordered selenate anions in the crystal structure. Below 8 Å, the anomalous signal is detectable and increases rapidly as resolution decreases to 27 Å. From the crystallographic data on P64K crystals, the volume occupied by one macromolecule is 19 724 Å^3 and the radius R of a sphere with the same volume is 27 Å: the low-resolution limit of this experiment is such that $R/d \geq 1$. The wavelength of the present experiment, 0.903 Å, is remote from the Se K edge and the corresponding value of f'' for Se is 3.3 e (Sasaki, 1984). The absorption spectrum of a thin pellet of ammonium selenate was recorded close to the K edge, using the D21 spectrometer at LURE (energy resolution ~ 3 eV). These data were exploited as usual (Hendrickson *et al.*, 1988) to get the variation of f'' and f' as a function of wavelength; as in other selenium compounds (*e.g.* Hendrickson, Horton & LeMaster, 1990), they exhibit a 'white line' of edge absorption. The maximum value of $2f''$ is 14.0 e and, assuming a remote wavelength at 0.90 Å ($f' = 1.62$ e), the maximum value of $\Delta f'$ is 8.6 e. Bijvoet ratios at $\lambda = 0.903$ Å were corrected for noise, as estimated from measurements on centric reflections. On the basis of (25) and (26), these corrected ratios were used to obtain expected maximum values of Bijvoet and dispersive ratios in an optimized MASC experiment with P64K crystals in 3.5 M ammonium selenate. The results are shown in Figs. 2 and 3.

5.2. HEW lysozyme

As another test case, we chose hen egg-white lysozyme (HEWL), since the crystals grow easily from NaCl solutions which could be replaced with YbCl$_3$ in order to exploit the anomalous scattering of Yb^{3+} at wavelengths close to the L_{III}-absorption edge. Crystals were grown in the tetragonal form (space group $P4_32_12$) using batch-crystallization methods from solutions containing 0.35–0.55 M ytterbium chloride and 20 mg ml^{-1} protein (pH $\simeq 5.0$). In order to increase the ytterbium ion concentration in the mother liquor, the crystals were transferred to sitting-drop solutions where the reservoir solution could be increased progressively up to 1.2 M YbCl$_3$, at which point crystals crack. A crystal soaked in 0.8 M YbCl$_3$ was finally used for a multiwavelength data collection.

Data collection was performed with the D23 instrument, installed on a bending magnet line of the storage ring DCI. This instrument features a sagittally focusing double-crystal monochromator and a multi-circle diffractometer. The area detector is a spherical drift proportional chamber (Kahn *et al.*, 1986), which has a proven capability for accurate measurements (Weis, Kahn, Fourme, Drickamer & Hendrickson, 1991). In order to measure low-resolution data, the conventional beam stop was replaced by smaller one (diameter 2 mm) stuck to a thin strip of Kapton which was placed across the entrance window of the detector. A cone filled with helium was placed between the crystal and the beam stop. The detector was set with the sensing planes of the multiwire proportional chamber normal to the X-ray beam. At the wavelength of the experiment ($\lambda \simeq 1.39$ Å)

Table 2

Multiwavelength X-ray data-collection statistics of lysozyme in the 0.8 M YbCl$_3$ test case.

Wavelength (Å)	Total No. of reflections	Total No. of unique reflections	R_{sym}*	Average intensity
1.3875 (peak)	5727	2575	0.086	73.3
1.3881 (edge)	5438	2520	0.052	132.2
1.3808 (remote)†	2641	2178	0.036	178.5

* $R_{sym} = \sum_{hkl} |I_j - \langle I \rangle| / \sum_{hkl} |I_j|$, for all measurements of I_{hkl} above 3σ, omitting anomalous differences. † Statistics for only the first half data set (24°) are shown for the remote wavelength.

and using the modified beam stop, the resolution range extends from ~180 to ~3.4 Å.

With the monochromator of the D23 instrument, tuning the wavelength of the monochromatic beam reflected by the second crystal is slow and tedious, and so is the recording of the fluorescence emission from the sample. In contrast, it is easy to record the absorption spectrum of a thin sample placed on the path of the beam reflected by the first crystal. The fluorescence from a sample of YbCl$_3$ was recorded only once and the angular difference between the point of strongest fluorescence (f''_{max}) and the maximum absorption of a copper foil (Cu K-absorption edge: $\lambda = 1.3808$ Å) was determined; the latter was used as a reference for subsequent monochromator settings. X-ray diffraction data were collected at three wavelengths from a single crystal. To optimize the anomalous (f''_{max}) and dispersive ($|f'_{max}|$) signals, these two wavelengths were determined from the X-ray fluorescence spectrum and correspond to the maximum (f''_{max} at $\lambda_{peak} = 1.3875$ Å) and the inflection point ($|f'_{max}|$ at $\lambda_{edge} = 1.3881$ Å) of the spectrum. The copper absorption edge ($\lambda = 1.3808$ Å) was used as λ_{remote}. The data were recorded as 0.1° rotation images in two blocks of 24° for each wavelength. The programs *MADNES* (Pflugrath & Messerschmidt, 1989) and *XDS* (Kabsch, 1988) were used to process the images (see Table 1). Further data treatment was completed using programs from the *CCP4* and *ANOSYS/MADSYS* (Weis, Smith & Hendrickson, 1991) suites. Local scale factors (Matthews & Czerwinski, 1975) were applied to the three wavelength-dependent data sets. Data-collection statistics are presented in Table 2.

The experimental Bijvoet ratios at each wavelength in resolution shells are shown in Fig. 4. In addition, ratios for pairs of centric reflections (the curve at λ_{remote} is shown in Fig. 4) gave an estimate of the intrinsic accuracy of measurements, and experimental Bijvoet ratios were accordingly corrected. Assuming that the value of f'' at λ_{remote} is 10.46 e (Sasaki, 1984) and that corrected Bijovet ratios are proportional to f'', values of f'' at λ_{peak} and λ_{edge}, respectively, derived from the corrected Bijvoet ratios, are 24 and 10 e, respectively. Taking into account the X-ray beam bandpass (about 4 eV) on the amplitude of the 'white line', these values are in agreement with expected values. Measurements of Bijvoet differences are reasonably accurate because absorption effects are very similar for both mates

of each anomalous pair. Large dispersive ratios have also been observed. In this case, however, absorption effects have a detrimental effect on errors, which are relatively large especially for differences involving the λ_{peak} data set. Effective absorption corrections, and not only local scaling
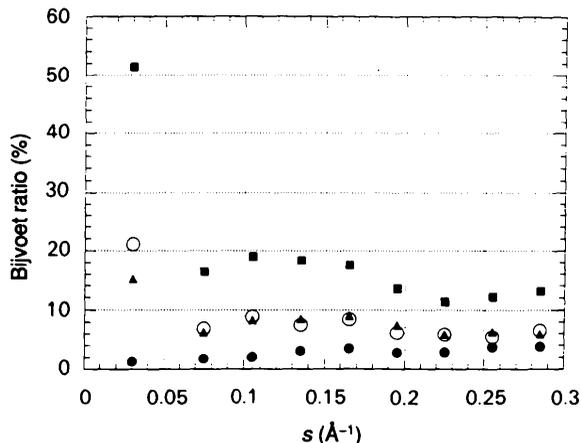
Figure 4

Lysozyme with 0.8 M YbCl$_3$. Breakdown of experimental ratios $\langle|\Delta F_{\pm h}|\rangle/\langle|\widetilde{F}|\rangle$ according to resolution, for acentric reflections (filled squares, open circles and filled triangles for λ_{peak}, λ_{edge} and λ_{remote}, respectively). These ratios were also calculated for centric reflections, to obtain an estimate of the intrinsic accuracy of measurements; only the breakdown at λ_{remote} is shown for clarity (filled circles).

procedures, will be required to exploit the potential of such measurements fully.

The largest anomalous-scattering effects are at low resolution and with the largest values for the λ_{peak} data, as already shown in the case of P64K, but here the anomalous-scattering effects are present in the whole resolution range at the three wavelengths (Fig. 4). The interpretation is that there is a superposition of effects due to disordered Yb^{3+} ions in the solvent and from some ordered ions bound to specific sites. Indeed, three metal sites were located from a phased anomalous Fourier map (Pepinsky & Okaya, 1956), using phases (calculated from coordinates of the tetragonal form of lysozyme from the PDB) incremented by 90°. Three peaks above 6σ in height appear in the map (a section is shown in Fig. 5). The largest of these (12σ in height) lies along the twofold diagonal axis between two protein molecules making contacts with the two side chains of Lys13 and the two C-terminal residues. The other two sites (7σ in height), are separated by 3.6 Å and lie in an intermolecular pocket surrounded by the residues 21, 35, 46, 52, 57 and 59. Interestingly enough, an anomalous Patterson map did not reveal any of these lanthanide sites.

6. Conclusions

MASC is a way of producing contrast variation by a physical change induced in a single sample, instead of conventional contrast variation by multiple solvent exchange.
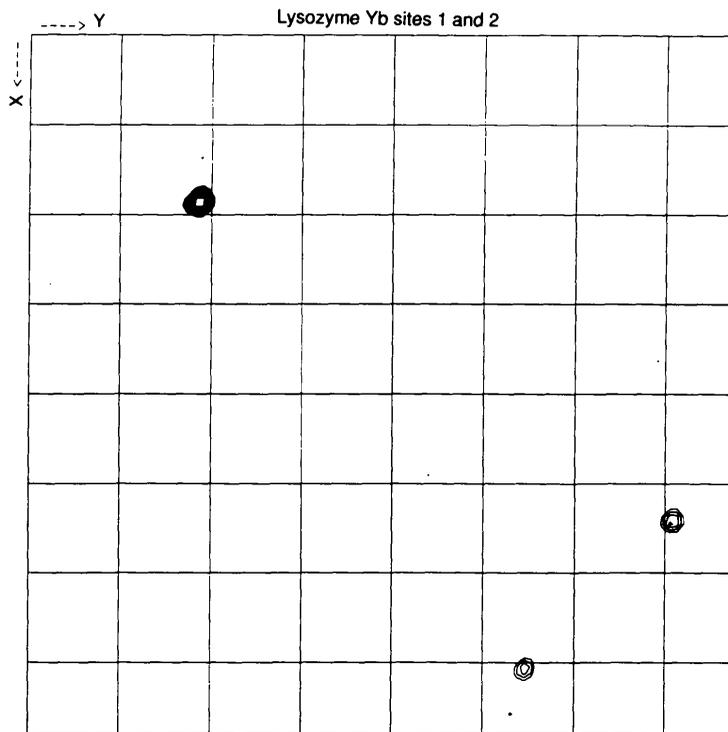
Figure 5

Section at $z = 0$ of the phased anomalous Fourier map (Pepinsky & Okaya, 1956) of lysozyme with 0.8 M YbCl$_3$ showing two of the three ytterbium sites. Contours start at 3σ and are plotted for every 1σ thereafter.

Using molar or multimolar concentrations of anomalous scatterers dispersed in the solvent of a macromolecular crystal and MAD-like techniques, it has been shown that anomalous signals have the expected magnitudes, which are appropriate to undertake the determination of $|{}^0F|$, $|G|$ and $\Delta\varphi = (\varphi_{0F} - \varphi_G)$ for a number of low-resolution reflections. This determination has yet to be carried out for a practical case. Then, the steps for the derivation of structural information from these data will be essentially the same as those for the determination of the envelope of tryptophanyl tRNA synthetase (Carter *et al.*, 1990), from the determination of the phases for $\{|G(\mathbf{h})|\}$ by direct methods to the determination of the molecular envelope and low-resolution phasing. The potential of the MASC method for complex structures depends critically on the initial phase determination by direct methods for $\{|G(\mathbf{h})|\}$. The molecular envelope has a much simpler structure than the initial macromolecular structure. The conversion of unitary structure factors obtained from the $\{|G(\mathbf{h})|\}$ to normalized structure factors makes use of the square root of an effective number of 'atoms' N. In the case of tryptophanyl tRNA synthetase, the value of $N^{1/2} = 6.3$ was used in the normalization process, corresponding to ~80 effective scatterers in the P321 asymmetric unit (74 000 Da); phases were obtained using conventional direct methods. More complex cases could probably be tackled, especially through the application of a new multisolution phasing method based on entropy maximization and likelihood ranking (for a review, see Bricogne, 1993).

Ideas and comments from G. Bricogne, C. W. Carter and C. Dumas were seminal during the sabbatical year of C. W. Carter at LURE in 1986. Discussions with P. Vachette at LURE were also most useful to clarify relations between solution scattering and the MASC method.

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. J., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). *J. Mol. Biol.* **112**, 535–542.

Bode, W., Chen, Z., Bartels, K., Kutzbach, C., Scmidt, G., Kastner, G. S. & Bartunik, H. D. (1983). *J. Mol. Biol.* **164**, 237–282.

Bragg, W. L. & Perutz, M. F. (1952). *Acta Cryst.* **5**, 277–289.

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

Bricogne, G. (1991). *Acta Cryst.* A**47**, 803–829.

Bricogne, G. (1993). *Acta Cryst.* D**49**, 37–60.

Cantor, C. R. & Schimmel, P. R. (1980). *Biophysical Chemistry, Part II, Techniques for the Study of Biological Structure and Function*. San Francisco: W. H. Freeman.

Carter, C. W., Bricogne, G. & Dumas, C. (1986). LURE Internal Report. LURE, Univ. Paris Sud, Orsay, France.

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

Crumley, K. V. (1989). MSc thesis, Univ. of North Carolina, Chapel Hill, USA.

DeVoe, I. W. (1982). *Microbiol. Rev.* **46**, 162–170.

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

Fourme, R. (1993). *Conference at the Third European Workshop on Crystallography of Biological Macromolecules*, Como, Italy, 24–28 May.

Fourme, R. & Hendrickson, W. A. (1990). *Biophysics and Synchrotron Radiation*, edited by S. S. Hasnain, pp. 156–175. Chichester: Ellis Horwood.

Harrison, S. C. (1969). *J. Mol. Biol.* **42**, 457–483.

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

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

Hendrickson, W. A., Horton, J. R. & LeMaster, D. M. (1990). *EMBO J.* **9**, 1665–1672.

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

Hütsch, M. (1993). PhD thesis, Univ. of Hamburg, Germany.

Ibel, K. & Stuhrmann, H. B. (1975). *J. Mol. Biol.* **93**, 255–265.

Jack, A., Harrison, S. C. & Crowther, R. A. (1975). *J. Mol. Biol.* **97**, 163–172.

Kabsch, W. (1988). *J. Appl. Cryst.* **21**, 916–924.

Kahn, R., Fourme, R., Bosshard, R. & Saintagne, V. (1986) *Nucl. Instrum. Methods*, **246**, 596–603.

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

Kühnholz, O. (1991). *J. Appl. Cryst.* **24**, 811–814.

Leslie, A. G. W., Brick, P. & Wonacott, A. T. (1986). *Daresbury Lab. Q. Protein Crystallogr.* **18**, 33–39.

Li de La Sierra, I., Prangè, T., Fourme, R., Padron, G., Fuentes, P., Musacchio, A. & Madrazo, J. (1994). *J. Mol. Biol.* **235**, 1154–1155.

Matthews, B. W. & Czerwinski, E. W. (1975). *Acta Cryst.* A**31**, 480–487.

Miake-Lye, R. C., Doniach, S. & Hodgson, K. O. (1983). *Biophys. J.* **41**, 287–301.

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

Moy, J. P. (1991). *Proceedings of the European Workshop on X-ray Detectors for Synchrotron Radiation*, Aussois, France, 30 September–4 October, edited by A. H. Walenta, pp. 221–227. Univ. of Siegen, Germany.

Munk, B. (1988). PhD thesis, Univ. of Hamburg, Germany.

Narayan, R. & Ramaseshan, S. (1981). *Acta Cryst.* A**37**, 636–641.

Pavlov, M. Yu & Fedorov, B. A. (1983). *Biopolymers*, **22**, 1507–1522.

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

Pflugrath, J. & Messerschmidt, A. (1989). *MADNES*. EEC Version, 17 January 1989. *EEC Cooperative Workshop on Position Sensitive Detectors*. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA.

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.

Porod, G. (1951). *Kolloid Z.* **124**, 83–114.

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.

Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

SERC Daresbury Laboratory (1979). *CCP4. A Suite of Programs for Protein Crystallography*. SERC Daresbury Laboratory, Warrington WA4 4AD, England.

Silva, R., Selman, M., Guillen, G., Herrera, L. S., Fernández, J. R., Novoa, L. I., Morales, J., Morera, V., González, S., Tamargo, B., del Valle, J. A., Caballero, E., Alvarez, A., Coizeau, E., Cruz, S. & Musacchio, A. (1992). European Patent Application EP 0 474 313 A2.

Smith, J. L. (1991). *Curr. Op. Struct. Biol.* **1**, 1002–1011.

Stuhrmann, H. B. (1980). *J. Appl. Cryst.* A**36**, 996–1001.

Stuhrmann, H. B. & Kirste, R. (1965). *Z. Phys. Chem.* **46**, 247–250.

Templeton, L. K., Templeton, D. H., Phizackerley, R. P. & Hodgson, K. O. (1982). *Acta Cryst.* A**38**, 74–78.

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

Weis, W. I., Smith, J. & Hendrickson, W. A. (1991). *ANOSYS/MADSYS* programs. Columbia Univ., New York, NY 10032, USA.

Westbrook, E. M., Deacon, M. L. Strauss, M. G., Naday, I. & Sweet, R. M. (1989). Crystallogr. Assoc. Meet., Series 2, Vol. 17, Abstract PA108.

Yonath, A. (1993). *XVI Congress of the International Union of Crystallography*, Beijing, China, 21–29 August.