The defect structure as described by the microdomain model does not consist of small precipitates of the Φ_1 phase. The intensities of the Φ_1 phase as observed by Marxreiter, Boysen, Frey & Vogt (1990) do not correspond to the intensities of the corresponding diffuse scattering. This confirms the results of Morinaga *et al.* (1980) who compared the intensities of the diffuse scattering with the intensity calculated from the partial structure model of the Φ_1 phase. The structure of the Φ_1 phase does not contain single oxygen vacancies. Furthermore, in the Φ_1 phase the calcium ions are second-nearest neighbors of the vacancies while they are preferentially nearest neighbors in the cubic solid solution.

This work was supported by funds of the BMFT.

References

ALLPRESS, J. G. & ROSSELL, H. J. (1975). *J. Solid State Chem.* **15**, 68-78.

ALLPRESS, J. G., ROSSELL, H. J. & SCOTT, H. G. (1975). *J. Solid State Chem.* **14**, 264-273.

ANDERSEN, N. H., CLAUSEN, K., HACKETT, M. A., HAYES, W., HUTCHINGS, M. T., MACDONALD, J. E. & OSBORN, R. (1985). *Proc. 6th Risø Int. Symp. on Metallurgy and Materials Science*, edited by F. W. POULSEN, N. H. ANDERSEN, K. CLAUSEN, S. SKAARUP & O. T. SORENSEN, pp. 279-284. London: Institution of Mining and Metallurgy.

ANDERSEN, N. H., CLAUSEN, K., HACKETT, M. A., HAYES, W., HUTCHINGS, M. T., MACDONALD, J. E. & OSBORN, R. (1986). *Physica (Utrecht)*, **136B**, 315-317.

BUTLER, V., CATLOW, C. R. A. & FENDER, B. E. F. (1983). *Radiat. Eff.* **73**, 273-277.

CARTER, R. E. & ROTH, W. L. (1963). In *Proc. Electromotive Force Measurements in High Temperature Systems*, edited by C. B. ALCOCK, pp. 125-144. London: Institution of Mining and Metallurgy.

FABER, J. JR, MUELLER, M. H. & COOPER, B. R. (1978). *Phys. Rev. B*, **17**, 4884-4888.

HELLMANN, J. R. & STUBICAN, V. S. (1983). *J. Am. Ceram. Soc.* **66**, 260-264.

HORIUCHI, H., SCHULTZ, A., LEUNG, P. C. W. & WILLIAMS, J. M. (1984). *Acta Cryst.* B**40**, 367-372.

HUDSON, B. & MOSELEY, P. T. (1976). *J. Solid State Chem.* **19**, 383-389.

HULL, S., FARLEY, T. W. D., HACKETT, M. A., HAYES, W., OSBORN, R., ANDERSEN, N. H., CLAUSEN, K., HUTCHINGS, M. T. & STIRLING, W. G. (1988). *Solid State Ionics*, **28–30**, 488-492.

KOESTER, L. & YELON, B. (1983). Neutron Diffraction Newsletter.

LORENZ, G. (1988). PhD thesis, Univ. München, Federal Republic of Germany.

LORENZ, G., FRAY, F., SCHULZ, H. & BOYSEN, H. (1988). *Solid State Ionics*, **28–30**, 497-502.

MARXREITER, J. (1988). Personal communication.

MARXREITER, J., BOYSEN, H., FREY, F. & VOGT, T. (1990). *Mater. Res. Bull.* **25**, 435-442.

MORINAGA, M., COHEN, J. B. & FABER, J. JR (1979). *Acta Cryst.* A**35**, 789-795.

MORINAGA, M., COHEN, J. B. & FABER, J. JR (1980). *Acta Cryst.* A**36**, 520-530.

NEDER, R. B., FREY, F. & SCHULZ, H. (1990). *Acta Cryst.* A**46**, 792-798.

OSBORN, R., ANDERSEN, N. H., CLAUSEN, K., HACKETT, M. A., HAYES, W., HUTCHINGS, M. T. & MACDONALD, J. E. (1985). *Mater. Sci. Forum*, **7**, 55-62.

SHANNON, R. D. (1976). *Acta Cryst.* A**32**, 751-767.

STEELE, D. & FENDER, B. E. F. (1974). *J. Phys. C*, **7**, 1-11.

SUBBARAO, E. C. (1981). In *Advances in Ceramics*, edited by A. H. HEUER & L. W. HOBBS, Vol. 3, pp. 1-24.

THORNBER, M. R., BEVAN, D. J. M. & GRAHAM, J. (1968). *Acta Cryst.* B**24**, 1183-1190.

Acta Cryst. (1990). A**46**, 809-820

Effect of the Anisotropy of Anomalous Scattering on the MAD Phasing Method

BY ERIC FANCHON* AND WAYNE A. HENDRICKSON

Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, 630 West 168th Street, New York, New York 10032, *USA*

(*Received* 26 *January* 1990; *accepted* 22 *May* 1990)

Abstract

The analysis of X-ray diffraction intensities is complicated by the anisotropy of anomalous scattering (AAS) that can occur due to resonance associated with transitions between core electrons and valence molecular orbitals. Substantial AAS has been observed directly in diffraction data near the *K* edge of selenium in selenolanthionine [Templeton & Templeton (1988). *Acta Cryst.* A**44**, 1045-1051] and in pleiochroism of X-ray absorption in selenobiotinyl streptavidin [Hendrickson, Pähler, Smith, Satow, Merritt & Phizackerley (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190-2194]. The impact of AAS on the multiple-wavelength anomalous diffraction (MAD) method for phase determination is of particular interest in the context of this chemical state of selenium in the light of a general method that has been developed to incorporate selenomethionine into

* Present address: Laboratoire de Cristallographie, CNRS, 66X, 38042 Grenoble CEDEX, France.

0108-7673/90/100809-12$03.00

proteins for use in MAD phasing [Hendrickson, Horton & LeMaster (1990). *EMBO J.* **9**, 1665–1672]. The first step of the MAD phasing method necessarily assumes that the anomalous-scattering factors are isotropic and our first aim here is to evaluate the effect of this approximation on initially determined phases. To obtain ultimate phases free from the effects of anisotropy, a least-squares procedure has been written in which *global* parameters (*i.e.* pertaining to the whole data set) are refined simultaneously with *local* parameters (*i.e.* pertaining to a given node **h**). The AAS is taken explicitly into account by considering f' and f'' as *tensors* instead of scalars [Templeton & Templeton (1982). *Acta Cryst.* A**38**, 62–67], and the components of the $\mathbf{f}'$ and $\mathbf{f}''$ tensors are among the refinable global parameters. The effectiveness of this procedure is tested with data simulated from the refined atomic model of selenobiotinyl streptavidin. The application of this procedure to actual Photon Factory measurements is also described. The results show that AAS does not cripple the MAD method, and that phases uncorrupted by these effects can be recovered.

1. Introduction

The effectiveness of the multiple-wavelength anomalous diffraction (MAD) method for solving protein structures has been evidenced recently by a number of achievements [for a review see Moffat (1988)]. The most notable of these are determinations of the completely novel structure of streptavidin (Hendrickson, Pähler, Smith, Satow, Merritt & Phizackerley, 1989) and the structure of cucumber basic protein (Guss, Merritt, Phizackerley, Hedman, Murata, Hodgson & Freeman, 1988), a member of the plastocyanin family that had resisted alternative methods of analysis. The MAD method exploits the frequency dependence of the anomalous scattering factor $f^{\Delta}(\omega) = f' + if''$ of an atom when the angular frequency ω of the incident beam is in the vicinity of an allowed electronic transition ω_{if} from a bound orbital to a state of the continuum (resonant scattering). The selective variation of the scattering factors for a small subset of the atoms is the basis for the solution of the phase problem.

The prerequisite of the MAD method is the presence in the structure of a suitable anomalous scatterer; that is, an atom having an absorption edge in the 0·5–2·5 Å wavelength region. Appropriate anomalous centers such as Fe or Zn can be present in native proteins. Another possibility is to replace a metal atom present in the native state by a more suitable atom. This has been done by Kahn, Fourme, Bosshardt, Chiadmi, Risler, Dideberg & Wery (1985) for a calcium-binding protein in which Ca has been replaced by Tb. Suitable analogs of ligands, cofactors or substrates can also be used, as can most heavy-atom derivatives ordinarily prepared as isomorphous replacements. Nevertheless, in most cases there is no heavy atom available for replacement or direct use. A general solution has been proposed to solve this problem for proteins (Hendrickson, 1985, 1987; Hendrickson, Horton & LeMaster, 1990). The idea is to replace sulfur by selenium in all methionines occurring in the protein by recombinant technology. It is thus possible to use the K edge of selenium for *ab initio* phase determination of any protein containing methionine in its sequence.

In selenomethionyl proteins, the selenium atom is covalently bonded to two sp^3 carbon atoms. Recent diffraction results from Templeton & Templeton (1988) show a very pronounced anisotropy of anomalous scattering (AAS) for the Se K edge of selenolanthionine, an organic compound that also involves the C–Se–C system. It is actually the greatest AAS yet observed at a K edge: at $E = 12\,654{\cdot}9$ eV the effective f'' lies between 0·8 and 7·7 e, depending on the orientation of the crystal in the beam. Similarly, f' lies between $-14{\cdot}7$ and $-8{\cdot}8$ e at this energy. Such anisotropy is also evident in the pleiochroism of X-ray absorption spectra measured from selenobiotinyl streptavidin (Hendrickson *et al.*, 1989) and from selenomethionyl *E. coli* thioredoxin (Hendrickson *et al.*, 1990). The anisotropy originates from the non-spherical symmetry of upper-energy states involved in the electronic transition giving rise to the anomalous-dispersion effect. The pertinent final states are the anti-bonding valence molecular orbitals. The anomalous-dispersion effect thus depends on the relative orientation of the incident and diffracted electric fields with respect to these molecular orbitals.

The basics of the MAD method are briefly reviewed in § 2. A general formalism to deal with AAS in the context of the MAD method is presented in § 3. It is noted that the initial stage of MAD phasing for unknown constellations of anomalous centers necessarily rests on the assumption that the anomalous factors are isotropic, hence the AAS cannot be included in the MAD process in the initial stage. In particular, the diffracted intensity in the presence of AAS is no longer proportional to a single $|F|^2$ value, but is instead the sum of four terms that are strictly speaking not *structure* factors since they are functions of orientation with respect to the polarization directions as well as atomic structure. As a consequence, the polarization correction cannot be applied separately. We present a procedure for improving initial phases obtained through the isotropic approximation, and describe a least-squares refinement program written for that purpose. It permits the combined refinement of local parameters (phases and moduli) and global parameters describing the anomalous scatterer (AS) substructure, including anomalous-scattering parameters. Results are presented in § 4. Our first aim is to quantify the

magnitude of the systematic error on phases introduced by the isotropic approximation. Simulated diffraction data calculated from the refined structure of streptavidin into which a known amount of anisotropy is incorporated are used for this purpose. Our second aim is to test the refinement procedure by using the same simulated data set. Starting from biased phases obtained through the isotropic approximation, we attempt to recover the true phases with the least-squares refinement procedure. Finally, the real streptavidin data, replete with AAS, are analyzed by the new procedures.

2. Outline of the MAD phasing method

An algebraic solution of the phase problem exploiting multiple-wavelength data was first proposed by Karle (1980) and implemented by Hendrickson (1985) in a somewhat different form. This will now be briefly described for later reference; see Hendrickson (1985) and Hendrickson, Smith, Phizackerley & Merritt (1988) for more details. The total structure factor at a given wavelength and for a particular reciprocal node **h** is written as

$$G_T(\mathbf{h}, \lambda) = F_T(\mathbf{h}) + \sum_q D_{Aq}(\mathbf{h}, \lambda), \qquad (1)$$

where the 'normal' scattering (f^0) contributions to F_T from *all* atoms are separated from the anomalous scattering (f^Δ) contributions to D_{A_q} from each kind q of anomalous scatterer ($1 \le q \le N_q$). The subscript indicates the set of atoms considered: T for total and A_q for anomalous scatterers of type q. The symbol F refers to structure factors in which only the *normal* scattering contributions are considered: D refers to *anomalous* contributions only; and G refers to the total structure factor. (These distinctions are adopted here in anticipation of the next section where, in the presence of AAS, D_A and hence G_T are shown to be intimately dependent on orientation of polarization directions and thus, strictly, are not factors of structure alone.)

In each partial structure factor D_{A_q} one can factor $f'_a + if''_q$ and write

$$D_{A_q}(\mathbf{h}, \lambda) = \frac{f'_q + if''_q}{f^0_q} F_{A_q}(\mathbf{h}). \qquad (2)$$

Denoting by φ_T and φ_{A_q} the phases of F_T and F_{A_q}, respectively, the square modulus of $G_T(\mathbf{h}, \lambda)$ can be expressed as a function of $|F_T|$, $|F_{A_q}|$ and $\Delta\varphi_q = \varphi_T - \varphi_{A_q}$. In other words, since the index q runs over the N_q types of AS, there are $2N_q + 1$ variables to be determined for each node **h**. The f'_q and f''_q factors, which are involved as coefficients, are presumed known. The measurement of the diffracted intensities at several (N_λ) wavelengths and for the two Friedel mates **h** and $-\mathbf{h}$, or their equivalents (hereafter referred to as $\sigma\mathbf{h}$, $\sigma = \pm 1$), provides $2N_\lambda$ observations. Supposing $2N_\lambda > 2N_q + 1$, the problem reduces to solving a (non-linear) system of $2N_\lambda$ equations with $2N_q + 1$ unknowns. This is accomplished by first solving a *linear* system in $(N_q + 1)^2$ unknowns and then imposing trigonometric constraints to reduce to the solution in $2N_q + 1$ non-linear variables (Hendrickson, 1985).

The calculations proceed as follows:

(1) Data reduction is performed; in particular, the data from Friedel (or mirror) mates and at different wavelengths must be reduced to a common scale.

(2) For each node **h**, *MADLSQ* is used to calculate the initial values of $|F_T|$, $|F_{A_q}|$ and $\Delta\varphi_q$ and to refine these values against the $2N_\lambda$ observations associated with that node.

(3) If necessary, the phasing results for symmetry-equivalent nodes are reduced to a unique set by *MERGIT*.

Then, for each kind q of AS:

(4) The positions $\mathbf{r}_j$ of the anomalous centers of type q must be determined from the $\{|F_{A_q}(\mathbf{h})|\}$ set by Patterson or direct methods.

(5) *ASLSQ* is used to refine these positions against $\{|F_{A_q}(\mathbf{h})|\}$.

(6) Finally, the phases $\varphi_{A_q}(\mathbf{h})$ are calculated from the positions obtained in step 5, and the desired phases are simply given by $\varphi_T = \Delta\varphi_q + \varphi_{A_q}$. These calculations and figure-of-merit estimates are done by *MADFAZ*.

Another step replacing steps 5 and 6 has been projected by Hendrickson (1985) whereby the *local* parameters $[|F_T(\mathbf{h})|, \varphi_T(\mathbf{h})]$ are refined together with the *global* parameters defining the AS partial structure (position, thermal factors, occupancy) against the reduced intensity observations $\{I(\sigma\mathbf{h}, \lambda) = |G_T(\sigma\mathbf{h}, \lambda)|^2\}$. This avoids the stepwise propagation of errors that occurs in the two-step procedure in which $(|F_T|, |F_A|, \Delta\varphi)$ are refined against $\{I(\sigma\mathbf{h}, \lambda)\}$, and then the positions (x_j, y_j, z_j) are refined against $\{|F_A(\mathbf{h})|\}$. It was also proposed to include the anisotropy of anomalous scattering at this point. We will now discuss this in more detail.

3. Introduction of the anisotropy of the anomalous scattering in the phasing procedure

We use the optical model of Templeton & Templeton (1982) which has been tested experimentally (Templeton & Templeton, 1986). In this model the scalar quantities f' and f'' are replaced by second-rank tensors (tensors as well as vectors will be denoted by bold letters).

Let us first define the polarization reference frames. The wave vectors of the incident and diffracted beam are denoted **k** and **k'**. The incident polarization directions are denoted by two unit vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ which are such that $(\hat{\mathbf{u}}, \hat{\mathbf{v}}, \hat{\mathbf{k}})$ forms an orthonormal system of axes. Similarly, $(\hat{\mathbf{u}}', \hat{\mathbf{v}}', \hat{\mathbf{k}}')$ forms an orthonormal

system of axes associated with the diffracted field. It is not necessary to make reference to the diffracting plane in the definition of vectors $\hat{\mathbf{u}}'$, $\hat{\mathbf{v}}'$ and $\hat{\mathbf{u}}$, $\hat{\mathbf{v}}$. In cases where the incident beam is linearly polarized and the diffracting plane is fixed, it is computationally simpler to choose $\hat{\mathbf{u}}'$ perpendicular to the diffracting plane and $\hat{\mathbf{v}}'$ parallel to it. In other words the s, p (or σ, π) notation (Templeton & Templeton, 1982) corresponds to a particular choice. In cases where the incident beam is not linearly polarized and/or an area detector is used other choices might be more convenient. Typically, $\hat{\mathbf{u}}$ is then chosen along the predominant direction of polarization.

The following development is valid within the framework of the kinematical theory (ideally imperfect crystal) in the two-beam situation. We will first consider the case of an incident beam linearly polarized along a direction $\hat{\mathbf{e}}$ and show that, due to AAS, the diffracted beam is no longer linearly polarized. It is thus necessary to consider the two mutually perpendicular components $\hat{\mathbf{u}}'$ and $\hat{\mathbf{v}}'$ which we identify collectively as $\hat{\mathbf{e}}'$. Denoting $\mathbf{f}=\mathbf{f}'+i\mathbf{f}''$ as the anomalous-scattering tensor, the *effective* AS factor for a particular atom at a given orientation of the crystal is $^t\hat{\mathbf{e}}\mathbf{f}\hat{\mathbf{e}}'$ (Templeton & Templeton, 1982, 1985) where superscript t stands for 'transpose of'. What matters here is the relative orientation of the principal axes of $\mathbf{f}$ (molecular axes) with respect to the polarization vectors $\hat{\mathbf{e}}$ and $\hat{\mathbf{e}}'$. For example, if the crystal is rotated around the scattering vector $\mathbf{h}$ (ψ scan), $\hat{\mathbf{e}}$ and $\hat{\mathbf{e}}'$ are fixed in the laboratory system but the molecular axes rotate so that the tensorial product $^t\hat{\mathbf{e}}\mathbf{f}\hat{\mathbf{e}}'$ changes during the scan. Similarly, during a data collection, the crystal is rotated to bring various nodes $\mathbf{h}$ into diffraction positions, and an effective AS factor must be calculated for each of these.

The anomalous contribution to the diffraction for a pair of directions $(\hat{\mathbf{e}}, \hat{\mathbf{e}}')$ is

$$D_A(\hat{\mathbf{e}}\hat{\mathbf{e}}'|\sigma\mathbf{h}, \Omega, \lambda) = \sum_j m_j \sum_s {}^t\hat{\mathbf{e}}\mathbf{f}_{js}\hat{\mathbf{e}}' T_{js} \exp(2\pi i\sigma\mathbf{h}\cdot\mathbf{r}_{js}), \quad (3)$$

where index s refers to symmetry elements, and j to independent atoms; m_j, $\mathbf{r}_{js}$ and T_{js} represent the occupancy, position and thermal factor of atom (j, s), respectively; $\mathbf{f}_{js}$ is the complex tensor $\mathbf{f}'_{js}+i\mathbf{f}''_{js}$. The parameter Ω represents any additional parameter needed to specify completely the orientation of the crystal with respect to polarization vectors. In the conventional diffractometer setting, where the scattering plane is fixed (*e.g.* horizontal) and reflections are measured in the bisecting geometry ($\theta=\omega$), no such parameter is needed. Matters are more complicated in the case of an area detector since each reflection defines its own scattering plane. Additional information specifying the crystal orientation must then be provided.

The complete structure factor, generalized to account for anisotropy, is, for a given $(\hat{\mathbf{e}}, \hat{\mathbf{e}}')$ pair,

$$G_T(\hat{\mathbf{e}}\hat{\mathbf{e}}'|\eta) = \hat{\mathbf{e}}\cdot\hat{\mathbf{e}}' F_T(\sigma\mathbf{h}) + D_A(\hat{\mathbf{e}}\hat{\mathbf{e}}'|\eta), \quad (4)$$

where η designates $(\sigma\mathbf{h}, \Omega, \lambda)$ to shorten the notations. Here the different types of anomalous scatterers are not separated and the subscript A refers to all of the anomalous scatterers in the structure. The $\hat{\mathbf{e}}'$ vector being either $\hat{\mathbf{u}}'$ or $\hat{\mathbf{v}}'$, two structure factors must be calculated. Given a linearly polarized incident field having $\mathbf{E}\sim E_e\hat{\mathbf{e}}$, the diffracted electric field is then

$$\mathbf{E}' = E_e K(\Psi)[G_T(\hat{\mathbf{e}}\hat{\mathbf{u}}'|\eta)\mathbf{u}' + G_T(\hat{\mathbf{e}}\hat{\mathbf{v}}'|\eta)\hat{\mathbf{v}}'] \times \exp(i\omega t - i\mathbf{k}'\cdot\mathbf{R}),$$

where $K(\Psi)$ depends on a set of parameters Ψ and accounts for absorption, Lorentz and other physical and experimental factors, except for polarization effects. In the presence of AAS the phases of the two structure factors are different and the scattered beam consequently has *elliptical* polarization. The normalized intensity of this diffracted beam is $I(\eta)=|\mathbf{E}'|^2/(K^2|\mathbf{E}|^2)$ and, as the $\hat{\mathbf{u}}'$ and $\hat{\mathbf{v}}'$ components are mutually perpendicular ($\hat{\mathbf{u}}'\cdot\hat{\mathbf{v}}'=0$),

$$I(\eta) = |G_T(\hat{\mathbf{e}}\hat{\mathbf{u}}'|\eta)|^2 + |G_T(\hat{\mathbf{e}}\hat{\mathbf{v}}'|\eta)|^2. \quad (5)$$

We now consider the general case of an *elliptically* polarized *incident* beam. This is pertinent since the photons emitted from synchrotron bending magnets have such characteristics outside the plane of the ring. Moreover, other relevant situations derive directly from the general result. The incident electric field is then proportional to the sum of two components

$$\mathbf{E}\sim E_u\hat{\mathbf{u}} + E_v\hat{\mathbf{v}}, \quad (6)$$

where E_u and E_v are complex numbers with different phases δ_u and δ_v, respectively. The polarization state of an electromagnetic plane wave is specified in general by the four Stokes parameters (*e.g.* Jackson, 1975), given here in terms of the linear polarization basis $(\hat{\mathbf{u}}, \hat{\mathbf{v}})$ as

$$\begin{aligned} s_0 &= |E_u|^2 + |E_v|^2 \\ s_1 &= |E_u|^2 - |E_v|^2 \\ s_2 &= 2|E_u||E_v|\cos(\delta_v - \delta_u) \\ s_3 &= 2|E_u||E_v|\sin(\delta_v - \delta_u). \end{aligned} \quad (7)$$

An observable electromagnetic field is an ensemble average from the superposition of photons not all of which will be coherent. Coherence depends on the physical event at the origin of the photon field, and on subsequent optical events. It is thus more useful to define the Stokes parameters as ensemble averages of the quantities defined by (7) (we do not introduce a new notation for these 'macroscopic' Stokes parameters since they are the only kind we will refer to in the following). The observable Stokes parameters

are subject to the constraint

$$s_0^2 \geq s_1^2 + s_2^2 + s_3^2,$$

where the equality obtains in the case of strict phase coherence. They characterize the state of polarization of a partially coherent beam. In the following we will make use of the ratios s_i/s_0 $(i=1,2,3)$, the total intensity of the incident beam s_0 being included in the scale factor. The ratios s_1/s_0 and $-s_3/s_0$ define the usual degree of *linear* polarization (P_L) and degree of *circular* polarization (P_C), respectively. It is important to realize that all these parameters make sense only with reference to a specified reference basis.

The diffracted electric field in the general case of an elliptically polarized incident wave (6) is then

$$\begin{aligned}\mathbf{E}' = K(\Psi)[&\{E_u G_T(\hat{\mathbf{u}}\hat{\mathbf{u}}'|\eta) + E_v G_T(\hat{\mathbf{v}}\hat{\mathbf{u}}'|\eta)\}\hat{\mathbf{u}}' \\ &+ \{E_u G_T(\hat{\mathbf{u}}\hat{\mathbf{v}}'|\eta) \\ &+ E_v G_T(\hat{\mathbf{v}}\hat{\mathbf{v}}'|\eta)\}\hat{\mathbf{v}}'] \exp(i\omega t - i\mathbf{k}'\cdot\mathbf{R}),\end{aligned}$$

and the normalized intensity of such a diffracted beam is of the form

$$I(\eta) = I_{u'}(\eta) + I_{v'}(\eta), \tag{8a}$$

where, for $\hat{\mathbf{e}}' = \hat{\mathbf{u}}'$ or $\hat{\mathbf{v}}'$,

$$\begin{aligned}I_{e'}(\eta) = &\tfrac{1}{2}(1 + s_1/s_0)|G_T(\hat{\mathbf{u}}\hat{\mathbf{e}}'|\eta)|^2 \\ &+ \tfrac{1}{2}(1 - s_1/s_0)|G_T(\hat{\mathbf{v}}\hat{\mathbf{e}}'|\eta)|^2 \\ &+ s_2/s_0\{\mathscr{R}e\, G_T(\hat{\mathbf{u}}\hat{\mathbf{e}}'|\eta)\, \mathscr{R}e\, G_T(\hat{\mathbf{v}}\hat{\mathbf{e}}'|\eta) \\ &+ \mathscr{I}m\, G_T(\hat{\mathbf{u}}\hat{\mathbf{e}}'|\eta)\, \mathscr{I}m\, G_T(\hat{\mathbf{v}}\hat{\mathbf{e}}'|\eta)\} \\ &- s_3/s_0\{\mathscr{I}m\, G_T(\hat{\mathbf{u}}\hat{\mathbf{e}}'|\eta)\, \mathscr{R}e\, G_T(\hat{\mathbf{v}}\hat{\mathbf{e}}'|\eta) \\ &- \mathscr{R}e\, G_T(\hat{\mathbf{u}}\hat{\mathbf{e}}'|\eta)\, \mathscr{I}m\, G_T(\hat{\mathbf{v}}\hat{\mathbf{e}}'|\eta)\}.\end{aligned} \tag{8b}$$

The analog of optical activity in the X-ray domain can be studied by considering the case of circular polarization ($P_C = \pm 1$, $P_L = s_2/s_0 = 0$) in (8).

When the $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ components of the incident beam have no phase correlation, s_2 and s_3 are zero. This is a special case of particular interest since it obtains for a beam taken symmetrically from the plane of a synchrotron source. The total diffracted intensity is then given by

$$\begin{aligned}I(\eta) = &(1 + P_L)/2\,\{|G_T(\hat{\mathbf{u}}\hat{\mathbf{u}}'|\eta)|^2 + |G_T(\hat{\mathbf{u}}\hat{\mathbf{v}}'|\eta)|^2\} \\ &+ (1 - P_L)/2\,\{|G_T(\hat{\mathbf{v}}\hat{\mathbf{u}}'|\eta)|^2 + |G_T(\hat{\mathbf{v}}\hat{\mathbf{v}}'|\eta)|^2\}.\end{aligned} \tag{9}$$

In the event that $P_L = 1$, (9) reduces to (5) - the case of linear polarization with $\hat{\mathbf{e}} = \hat{\mathbf{u}}$. Note that AAS affects the diffracted intensities even in the case of a non-polarized ($P_L = s_2/s_0 = P_C = 0$) incident beam.

At wavelengths for which *all* anomalous scatterers can be considered isotropic, (8) reduces to

$$\begin{aligned}I(\eta) = &\{\tfrac{1}{2}(1 + P_L)[(\hat{\mathbf{u}}\cdot\hat{\mathbf{u}}')^2 + (\hat{\mathbf{u}}\cdot\hat{\mathbf{v}}')^2] \\ &+ \tfrac{1}{2}(1 - P_L)[(\hat{\mathbf{v}}\cdot\hat{\mathbf{u}}')^2 + (\hat{\mathbf{v}}\cdot\hat{\mathbf{v}}')^2] \\ &+ s_2/s_0[(\hat{\mathbf{u}}\cdot\hat{\mathbf{u}}')(\hat{\mathbf{v}}\cdot\hat{\mathbf{u}}') \\ &+ (\hat{\mathbf{u}}\cdot\hat{\mathbf{v}}')(\hat{\mathbf{v}}\cdot\hat{\mathbf{v}}')]\}|G_T(\mathbf{h}, \lambda)|^2.\end{aligned} \tag{10}$$

The factor in front of $|G_T(\mathbf{h}, \lambda)|^2$ corresponds to the general formula of the polarization factor given by equation 10 of Vaillant (1977). As noted by this author the s_3 parameter does not figure in this expression and therefore cannot be determined from the measurement of integrated intensities in such situations. AAS offers an opportunity to make this characterization *via* (8) in a restricted energy range. It should be noted that in the general case (8), or even in the case of (9), a single polarization *factor* cannot be defined and the polarization correction should not be applied when processing data to which this formalism is to be applied.

If we consider now the MAD phasing process, the complications are threefold. First, as for anisotropic thermal factors, the AS factors '$\mathbf{ef}_{js}\mathbf{e}$' cannot be factored in (3) because their values depend on the orientation of each atom (j, s) in the cell. The only exception to this is the special case in which all anomalous scatterers of a given kind (including symmetry-related ones) have the same orientation of their tensors. Second, unlike temperature-factor anisotropy, the effects of AAS depend on crystal orientation about the diffraction vector. This reinforces the factorization problem since polarization-dependent structure factors must be calculated for each of the four pairs of polarization directions, and the diffracted intensity is no longer proportional to a single $|F|^2$ value. Finally, the AS tensors are unknown at the outset. Even if the principal values of $\mathbf{f}_j'$ and $\mathbf{f}_j''$ tensors are known from previous experiments, the orientations of the molecular orbitals in the particular structure under study are unknown.

The first point is the fundamental difficulty since the method developed previously (§ 2) is based on the factorization of f^A. It is therefore not possible to extract directly the values of the partial normal structure factors $|F_{A_q}(\mathbf{h})|$ from the multiwavelength data. The approach considered here is to decompose the problem into two parts; first the isotropic approximation is made by using averages of the principal values of $\mathbf{f}_j'$ and $\mathbf{f}_j''$ in order to initiate the phasing procedure. The steps 2–6 are performed according to the description of § 2 and result in a set of values biased to some extent by the AAS. At that point the positions of the AS are determined and it is thus possible to use (3), (4) and (9) in a least-squares refinement program aimed at refining the $\mathbf{f}_j'$ and $\mathbf{f}_j''$ tensor components

along with local parameters. Since the orientations of molecular orbitals are still unknown, the refinement must start from spherical (isotropic) tensors derived from the scalar values used in the first step. This is analogous to what is done for anisotropic temperature factors in small-molecule structures. As the $\mathbf{f}'_j$ and $\mathbf{f}''_j$ tensors components are changed, the local parameters $|F_T(\mathbf{h})|$ and $\varphi_T(\mathbf{h})$ must also be readjusted. Global and local parameters have thus to be refined simultaneously.

The scheme just sketched shows that a complete description including the AAS can be integrated in the additional step alluded to at the end of § 2 (which would be step 7), the set of refinable global parameters being extended to anomalous-scattering tensor components. Upon convergence, this should give a phase set $\{\varphi_T\}$ free from the AAS effect. Note that all of this is based on the hypothesis that the isotropic approximation of the first stage suffices to yield the positions of anomalous centers and then gives phases sufficiently reliable that they can initiate the refinement accounting for AAS.

A program called *MADDST* has been written to perform this task. It is space-group general and can deal with any number of different types of anomalous scatterers. The refinement is performed on normalized intensities. The $|F_T(\mathbf{h})|$ and $\varphi_T(\mathbf{h})$ parameters are refined by minimizing the local residual

$$\mathscr{L}(\mathbf{h}) = \sum_{\sigma} \sum_{\lambda} w(\eta)[I(\eta) - I_{\text{obs}}(\eta)]^2,$$

where $I(\eta)$ is usually given by (9), or by (8) in the more general case. The weights are either unitary weights or $w = 1/\sigma^2(I_{\text{obs}})$. The global parameters are refined by minimizing $G = \sum_{\mathbf{h}} \mathscr{L}(\mathbf{h})$. When doing this, the correlations between local and global parameters are neglected. This choice is dictated by the fact that the size of the complete normal matrix is prohibitive. A similar situation occurs in the combined refinement of phases and heavy-atom parameters in the MIR method. It is in particular the case of Sygusch's (1977) method in which phases are not estimates as in the classical Dickerson scheme (Dickerson, Weinzierl & Palmer, 1968), but extra refinable parameters. Bricogne (1982) pointed out that neglect of correlations between heavy-atom parameters and phases leads to slow convergence and possibly biased results, because these correlations are liable to induce indirect correlations between heavy-atom positions in two unrelated derivatives. In the MAD method the problem does not exist for AS positions since they are common to all 'derivatives' (*i.e.* wavelengths). But it might be a concern for the anomalous-scattering parameters at different wavelengths, although our test refinements do not indicate that this is a problem. In any case the general formalism developed by Bricogne (1982) can also be adapted to the MAD phase refinement.

The anomalous-scattering (symmetric) tensors are described by six components f_{mn} in much the same way as anisotropic temperature factors are. This leads to 12 parameters per independent atom and per wavelength. An alternative description in terms of principal values and orientational parameters (Eulerian angles) has also been implemented. This permits a reduction of the number of parameters since the principal values are common to all the AS of the same kind, and the orientational parameters are common to both $\mathbf{f}'$ and $\mathbf{f}''$ and for all wavelengths. Other advantages of the second description are that the principal values at wavelength λ are characteristic of the element and its chemical state, and thus these can be reused in the study of another structure (provided that equivalent experimental conditions are used, in particular the same energy resolution); moreover, the orientation at a wavelength at which the atom is weakly anisotropic would be poorly defined. The use of Eulerian angles common to several wavelengths removes this problem. On the other hand it is difficult to handle symmetry restriction on the Eulerian angles if the atom is on a special position.

MADDST calculates the derivatives of the intensity with respect to the following parameters:

(1) local parameters: $|F_T(\mathbf{h})|$ and $\varphi_T(\mathbf{h})$ (only $|F_T|$ if $\mathbf{h}$ is centric);

(2) global parameters:

global scale factor, three normalized Stokes parameters s_i/s_0 (only the refinement of $s_1/s_0 = P_L$ has been tested in the current version);

position, occupancy and temperature factor (isotropic or anisotropic) for anomalous scatterers;

isotropic $f'_q(\lambda), f''_q(\lambda)$ factors for wavelengths remote from the absorption edge of species q;

the components $[\mathbf{f}'_{js}(\lambda)]_{mn}$ and $[\mathbf{f}''_{js}(\lambda)]_{mn}$ (description 1) or the principal values and Eulerian angles if description 2 is chosen.

In principle, the polarization state of the incident beam depends on the wavelength. This is because the synchrotron emission is itself dependent on it, but it is also due to the change in monochromator angle when selecting different wavelengths. These are expected to be small effects (at least for bending magnets and wigglers), and the Stokes parameters are treated as wavelength *in*dependent in the current version of *MADDST*. In typical actual practice, the diffracting crystal is centered in the orbital plane of the synchrotron, and the parameters s_2 and s_3 are zero. This assumption, of course, depends on the beam stability achieved at a particular ring. Although the degree of linear polarization is known from the theory of synchrotron emission, it has been recognized (Materlik & Suortti, 1984) that it should be measured for each particular experimental configuration. A precise method to achieve this has been published recently (Staudenmann & Chapman, 1989).

The calculation involves the evaluation of the exact orientation of the crystal in the beam. For that purpose the orientation matrix of the crystal must be given as input. The program has been developed for a four-circle diffractometer equipped with a single counter, and also for area detector systems. To facilitate modifications, all features depending on the particular experimental configuration (namely the calculation of polarization vectors) are grouped in a subroutine.

The program is written in Fortran 77 and runs on Vax 11/750 and Convex C220 machines.

4. Results

To test the hypothesis proposed previously (namely that the isotropic approximation is good enough to initiate the refinement of tensor components and phases to exact values), the complex of selenobiotinyl streptavidin was chosen for application of the AAS formalism developed above. Streptavidin is a protein that binds tightly to biotin. It contains no methionine in its sequence, but biotin has a sulfur atom which is replaced by selenium in selenobiotin. The structure was determined based on MAD phasing using the Se K edge (Hendrickson *et al.*, 1989), and this has now been refined at 2 Å resolution to $R = 0{\cdot}17$ (Pähler & Hendrickson, 1990). The reasons for this choice are first the presence of a selenium atom covalently bonded to two carbons as would be the case in selenomethionyl proteins, and second the availability of refined atomic coordinates which permits the calculation of reference phases. Moreover, the fact that this structure has actually been solved using multiwavelength data offers the opportunity to examine AAS in a real situation. Although the large energy bandwidth in this experiment (approximately 10 eV) attenuates the effect of anisotropy, AAS can be observed directly in the two absorption spectra that were recorded in two different orientations: the first one with **a** parallel to the electric field **E**, the second one with **E**‖**c** (Figs. 1*a*, *b*). The corresponding spectra recorded at SSRL with a 3-4 eV energy bandwidth (Figs. 1*e*, *f*) have a quite different aspect. In particular the 'white line' feature has mostly disappeared in the Photon Factory experiment. This point will be discussed later.

Only features of the streptavidin analysis relevant to this study are presented here. Details are to be found in Hendrickson *et al.* (1989). The space group is $I222$ and the cell parameters are $a = 95{\cdot}27$, $b = 105{\cdot}41$, $c = 47{\cdot}56$ Å. There is only one kind of anomalous center ($N_q = 1$) and thus diffraction data were measured at only three wavelengths ($\lambda_1 = 0{\cdot}9000$; $\lambda_2 = 0{\cdot}9795$; $\lambda_3 = 0{\cdot}9809$ Å). Data were collected at the Photon Factory using a conventional diffractometer operated in the bisecting mode with radiation produced by a vertical wiggler. Thus, the scattering plane is horizontal and, to first approximation, the incident beam is linearly polarized along the vertical axis. In this configuration the polarization correction is not necessary. The $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ vectors of the polarization basis are chosen vertical and horizontal, respectively, so that the degree of linear polarization is $P_L = 1$. The number of reciprocal nodes is 4618 in the 70 to 3·1 Å range of Bragg spacings. This corresponds to 27 677 observations. In general six observations (two Friedel mates at three wavelengths) are associated to a given **h**. No equivalent reflections have been collected. An empirical absorption correction based on a ψ scan, and a parameterized local scaling procedure (Hendrickson & Teeter, 1981) were also applied. This local scaling procedure preserves any anisotropy present in a data set. Since the polarization correction was not applied, it was not necessary here to reprocess the data to undo it.

Table 1. *Orientation of the two selenobiotin molecules: angles (°) between molecular axes and cell axes*

	$\mathbf{x}_{mol}(1)$	$\mathbf{y}_{mol}(1)$	$\mathbf{z}_{mol}(1)$	$\mathbf{x}_{mol}(2)$	$\mathbf{y}_{mol}(2)$	$\mathbf{z}_{mol}(2)$
a	28·7	91·6	61·4	154·4	91·8	117·5
b	68·1	126·1	135·8	63·1	81·4	151·5
c	72·4	36·2	120·5	84·4	171·2	96·8

The molecular model has been refined at 2·0 Å resolution against a fourth data set collected at Cu $K\alpha$ wavelength. The molecular axes defining the Se environment (principal axes of the AS tensors) were calculated from atomic coordinates as in the case of selenolanthionine (Templeton & Templeton, 1988). Thus, $\mathbf{z}_{mol}$ is defined as the bisector of the two Se–C interatomic vectors, $\mathbf{x}_{mol}$ is normal to the C–Se–C plane and $\mathbf{y}_{mol}$ completes the orthonormal basis set. There are two independent selenobiotin molecules per asymmetric unit, and consequently two independent Se atoms. These two selenobiotin molecules are such that $\mathbf{x}_{mol}$ is nearly parallel (within 30°) to the **a** axis, as can be seen in Table 1. The space group being $I222$, all the biotin molecules have $\mathbf{x}_{mol}$ nearly parallel to the **a** axis.

(*a*) *Calculations with simulated (error-free) data*

The $\{(|F_T(\mathbf{h})|, \varphi_T(\mathbf{h}))\}$ set of local parameters was first calculated from the refined atomic model for reference. This set will be referred to as the MODEL set.

To evaluate the effect of the AAS on the MAD phases a set of diffracted intensities (containing the same 4618 nodes as in the real data set) was generated incorporating a known amount of anisotropy. The local parameter values are those of the MODEL set; the positions, temperature factors and the orientational parameters of the anomalous scatterers are from the refined model, and the principal values of the AS tensors are as extracted from Table 4 of Templeton & Templeton (1988). The $\lambda_1 = 0{\cdot}9000$ Å

wavelength is far from the Se K absorption edge so that scalar f' and f'' parameters are used: $f'(\lambda_1) = -1{\cdot}622$ and $f''(\lambda_1) = 3{\cdot}285$ e. The wavelengths chosen for this simulation do not correspond exactly to those of the actual Photon Factory experiment since our goal here was just to evaluate the magnitude of the error introduced by the neglect of AAS in a typical situation. Moreover the edge wavelength as measured in the selenolanthionine experiment (Templeton & Templeton, 1988) is shifted by several eV with respect to the one defined in the Photon Factory experiment, so that nominal wavelength values cannot be compared directly. The values at λ_2 are such that the observed anisotropy is large for both f' and f''.

These intensities were used as if they were experimental data. They were fed into *MADLSQ* which gave $|F_T|$, $|F_A|$ and $\Delta\varphi$ for all nodes except for eight that could not be phased due to lack of signal or insufficient number of observations. The AS factor values used in these calculations are isotropic equivalents of those in Table 2: $f'_{iso}(\lambda_2) = -11{\cdot}90$, $f''_{iso}(\lambda_2) = 4{\cdot}53$, $f'_{iso}(\lambda_3) = -6{\cdot}83$, $f''_{iso}(\lambda_3) = 0{\cdot}47$. The AS atomic positions are presumed to be known and step 3 is skipped. The refinement of positional par-

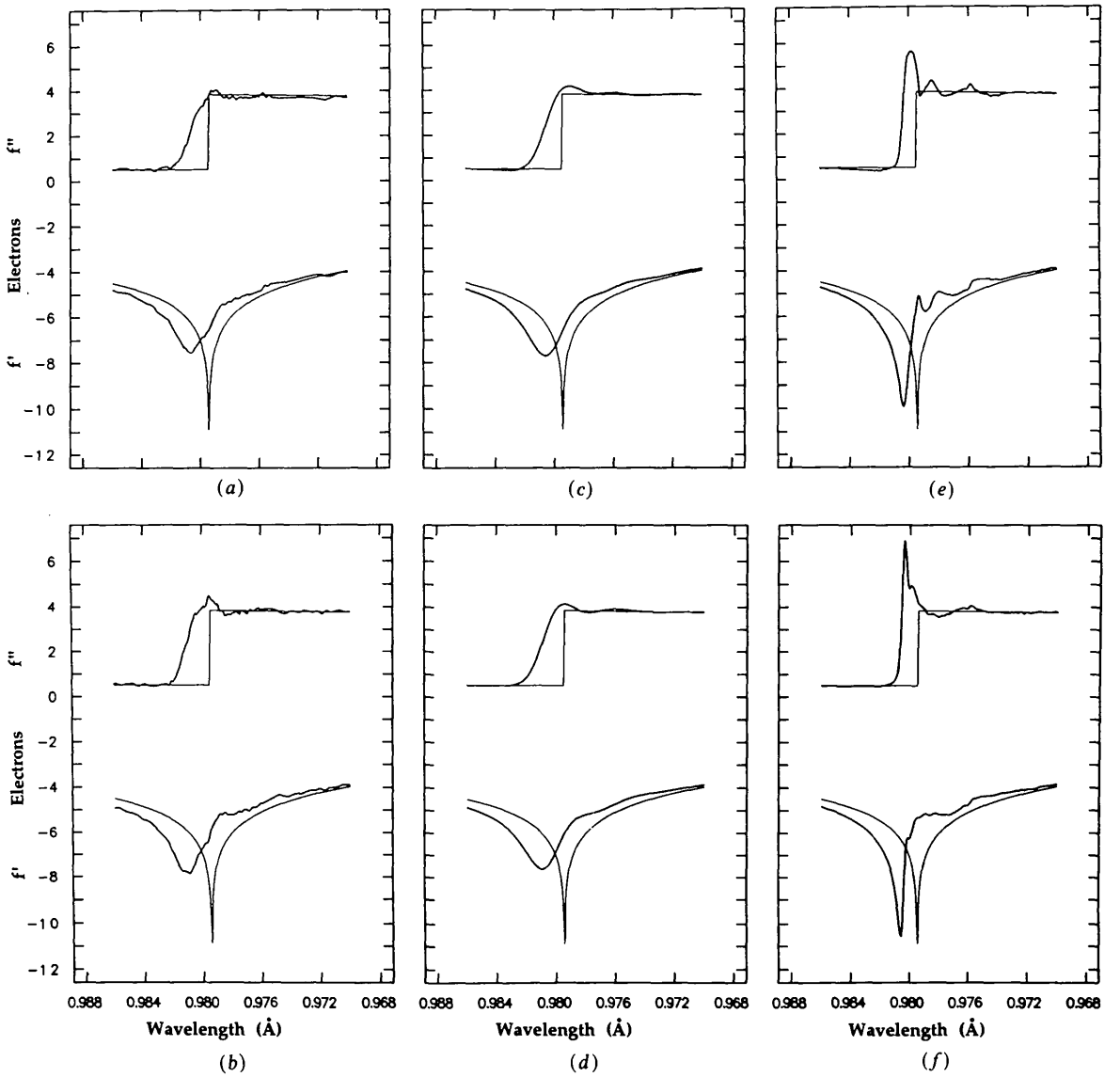

Fig. 1. Anisotropy in anomalous-scattering factors for selenium in selenobiotinyl streptavidin. Top row (*a*), (*c*), (*e*): f' and f'' calculated from absorption spectra recorded with **E**‖**a**. Bottom row (*b*), (*d*), (*f*): f' and f'' calculated from absorption spectra recorded with **E**‖**c**. Curves (*a*) and (*b*) are calculated from the Photon Factory (KEK, Japan) absorption spectra. Curves (*e*) and (*f*) are calculated from the original Stanford (SSRL) spectra, corresponding to an evaluated 3–4 eV energy spread. Note that a 17·5 eV translation has been applied with respect to the originally published spectra (Hendrickson *et al.*, 1989), in order to bring these curves in correspondance with the (*a*) and (*b*) curves. Curves (*c*) and (*d*) are from Stanford (SSRL) absorption spectra after convolution with a Gaussian ($\sigma = 11$) and translation (see text). Theoretical curves as computed by the Cromer (1983) program are shown as smooth lines.

Table 2. *Principal values (in electrons) of the AS tensors used in the simulation*

These are extracted from Table 4 of Templeton & Templeton (1988).

	f'_x	f'_y	f'_z	f''_x	f''_y	f''_z
λ_2	−8·8	−14·7	−12·2	0·8	7·7	5·1
λ_3	−4·8	−8·1	−7·6	−0·8	1·0	1·2

ameters by *ASLSQ* (temperature factors were not refined) led to the following displacements: 0·047 Å for Se(1) and 0·019 Å for Se(2). The standard and weighted agreement factors (based on $|F|$) are 13·9 and 16·5%, respectively. The bias introduced by the isotropic approximation on $|F_A|$ values is not negligible but the localization of AS from $\{|F_A(\mathbf{h})|\}$ would not be expected to be impaired. Next, φ_A phases were computed, together with $\varphi_T = \Delta\varphi + \varphi_A$. This gave what will be referred to as the SIMUL set.

Control calculations based on a data set corresponding to isotropic AS confirmed that no unexpected bias had been introduced by *MADLSQ*; that is, theoretical values were obtained for $|F_T|$, $|F_A|$ and $\Delta\varphi$. On the other hand, the isotropic approximation of actual anisotropy introduces systematic errors in principle in each step: in the *MADLSQ* step; in *ASLSQ* through the shift in positions due to approximate $|F_A(\mathbf{h})|$ (actually quite small as we just saw); and in the calculation of φ_A from isotropic AS factors.

The average discrepancy between the MODEL phases and the SIMUL phases is $\langle|\Delta\Phi|\rangle = 8{\cdot}7°$ for 4610 nodes, and $\langle|\Delta\Phi|\rangle = 9{\cdot}8°$ for the 3814 acentric nodes among them. These values can be compared with the discrepancy between the real MAD phases and MODEL phases which are 58·0° for the whole data set and 57·1° for the acentric subset. Table 3 shows the dependence of $\langle|\Delta\Phi|\rangle$ on $(\sin\theta)/\lambda$ when comparing the MODEL and SIMUL phase sets. $\langle|\Delta\Phi|\rangle$ is slowly increasing above $(\sin\theta)/\lambda = 0{\cdot}02$, which is what is expected from the fact that the normal contribution is decreasing with $(\sin\theta)/\lambda$ while the anomalous part is essentially constant, thus its relative weight increases.

We now turn to our second point which is to test the ability of *MADDST* to lead to unbiased phases. *MADDST* was run with the SIMUL set as starting point for the local parameters. The global and local parameters were refined simultaneously without particular strategy. Fixed values were assigned to the following global parameters: the degree of linear polarization P_L, the overall scale factor, the occupancies, the isotropic thermal factors and the isotropic f' and f'' for λ_1. The global parameters that were allowed to vary are the tensorial components of AS factors for λ_2 and λ_3 representing 48 variables (12×2 independent atoms $\times2$ wavelengths) and the six positional coordinates for the Se atoms. There are 8425 local parameters. After four cycles the average phase discrepancy had been reduced to 2·17° (4610 nodes). Further refinement (32 cycles) lead to a 0·89° discrepancy. This residual value is mainly due to centric phases since these are not refined, and thus phases that have wrong values initially remained incorrect. All of the acentric phases had converged towards the right value as had the orientation parameters and principal values. The final approach was slow; it took about 36 cycles to reach the minimum. After the 32nd cycle the description of the anomalous-scattering tensors was switched to description 2, which uses principal values and Eulerian angles. Corresponding principal values determined independently from Se(1) and Se(2) agree within 0·07 e. Similarly, only two of the eight sets of Eulerian angles that can be calculated - **f'** and **f''** at two wavelengths for Se(1) and Se(2) - are independent. Comparisons show that corresponding values agree within 3°. The identities of principal axes were assigned to correspond with the molecular axial system on the basis of the known principal values given above.

Table 3. *Comparison of the MODEL and SIMUL phase sets: variation of* $\langle|\Delta\Phi|\rangle$ (°) *with* $(\sin\theta)/\lambda$ (Å^{-1})

$(\sin\theta)/\lambda$	0	0·032	0·064	0·096	0·128	0·160	0·192
$\langle|\Delta\Phi|\rangle$		3·6	8·1	7·7	9·1	8·8	11·1

An alternative program derived from *MADDST* was also tested. Instead of refining moduli and phases, trial phases are chosen at regular points between 0 and 360°, and for each trial value the $|F_T|$ modulus is refined. For centric phases the search is limited to the two possible values 0 and 180°. The phase giving the lowest local residual is kept. In this procedure the phases previously estimated by *MADLSQ* are discarded. The global parameters were not refined. This procedure is analogous to what was done by Guss *et al.* (1988) after phasing with *MADLSQ* and to the *ABCD* procedure used with streptavidin (Pähler, Smith, & Hendrickson, 1990). All of the centric phases now achieved correct values and the true solution was recovered exactly. The only difference with the refinement-only pathway lies in the fact that centric phases have all been flipped to their proper values.

(*b*) *Photon Factory data*

Our starting point consists of the AS parameters $\{\mathbf{r}_j, B_j\}$ and the moduli and phases for streptavidin as determined by Hendrickson *et al.* (1989) with *MADLSQ/ASLSQ/MADFAZ*. Isotropic AS factors were set initially at previously assigned values: $f'(\lambda_1) = -1{\cdot}622$, $f''(\lambda_1) = 3{\cdot}285$, $f'(\lambda_2) = -6{\cdot}203$, $f''(\lambda_2) = 3{\cdot}663$, $f'(\lambda_3) = -8{\cdot}198$, $f''(\lambda_3) = 2{\cdot}058$. The average phase discrepancy between these MAD phases and those from the stereochemically restrained model is $\langle|\Delta\Phi|\rangle = 58{\cdot}0°$ for the 4608 phased nodes.

Table 4. *Principal values (in electrons) of the AS tensors obtained from refinement against diffraction data (e.s.d.'s are given in parentheses)*

	f'_x	f'_y	f'_z	f''_x	f''_y	f''_z
$\lambda = 0{\cdot}9795$ Å	−6·2 (1)	−6·1 (2)	−6·7 (2)	3·5 (1)	4·5 (3)	3·5 (1)
$\lambda = 0{\cdot}9809$ Å	−8·2 (1)	−7·6 (2)	−7·5 (2)	2·1 (2)	2·2 (3)	1·4 (2)

There are 257 centric phases which are 180° apart, giving $\langle|\Delta\Phi|\rangle = 62{\cdot}4°$ for the 796 centric nodes. The positions refined by *ASLSQ* are very close to the positions in the final protein model.

The AS factors were first kept isotropic, and the positions refined against a subset of observations. Then the local parameters were refined. After refinement, including the phase search procedure, the average discrepancy was reduced to $\langle|\Delta\Phi|\rangle = 57{\cdot}5°$. This slight improvement of 0·5° overall was accompanied by an average phase shift of 16·8° on 3814 acentric nodes. As a comparison, molecular averaging improved the phases by 2·5° (4958 nodes) on this problem (Hendrickson *et al.*, 1989). This is not much better considering that this procedure introduces external information, and it shows that the MAD phases are of fairly good quality.

In order to evaluate the anisotropy of anomalous scattering, we first refined independently the six components of each tensor. This amounts to 48 parameters since anisotropy is expected at two wavelengths. Then principal values and Eulerian angles were extracted and the refinement was pursued for a few cycles. The assignment of molecular axes to principal values is based on their relative values as listed in Table 2. The global parameters are held fixed at their initial value. No phase improvement was observed. Table 4 contains the refined principal values obtained.

If we repeat the simulation procedure as described above (that is generation of an anisotropic error-free data set followed by MADLSQ phasing), but now using these anisotropic AS values from analysis of the Photon Factory data at low energy resolution in place of the high-resolution parameter of Templeton & Templeton (1988), the average phase discrepancy $\langle|\Delta\Phi|\rangle$ is 1·8°. The systematic error introduced by the AAS is thus negligible compared with other sources of errors.

Table 4 shows that the anisotropy is weaker here than that observed by Templeton & Templeton (1988), but it is present. To assess the significance of our results, the directions of principal axes of the two molecules are compared with those obtained from the refined atomic model for selenobiotinyl streptavidin. This is summarized in Table 5. A 30° rotation relates the two systems for Se(1), whereas it takes a 40° rotation in the case of Se(2). Considering that a weak asphericity of tensors is associated with poorly defined angles this is reasonable, although far from a perfect match.

Table 5. *Angles (°) between molecular axis calculated from atomic model and corresponding axis obtained from refinement*

	Se(1)	Se(2)
(x_{model}, x_{ref})	29·3	33·3
(y_{model}, y_{ref})	28·5	24·6
(z_{model}, z_{ref})	13·7	32·2

Table 6. *Comparison of values calculated from absorption spectra (Absn) with values from MADDST refinement (Diffn)*

$\lambda = 0{\cdot}9795$ Å:	$\langle f'$ (Diffn)$\rangle$	f' (Absn)	$\langle f''$ (Diffn)$\rangle$	f'' (Absn)
a	−6·43	−6·56	3·60	3·85
c	−6·13	−5·84	4·35	4·42
$\lambda = 0{\cdot}9809$ Å:	$\langle f'$ (Diffn)$\rangle$	f' (Absn)	$\langle f''$ (Diffn)$\rangle$	f'' (Absn)
a	−7·76	−7·48	1·69	1·69
c	−7·60	−7·88	2·14	2·64

Comparison can also be made with results from absorption spectra. As noted above, two absorption spectra were recorded at the Photon Factory: the first one with **a** parallel to the electric field **E**, the second one with **E**$\|$**c**. The local programs *XASFIT* and *KRAMIG* (Hendrickson *et al.*, 1988) permit one to calculate the $f'(\omega)$ and $f''(\omega)$ curves from the absorption spectra (Fig. 1). The values obtained in that way are averages over the unit cell since the absorption coefficient μ is a macroscopic quantity. To obtain comparable quantities from the *MADDST* tensors, averages $\langle\mathbf{f}'\rangle$ and $\langle\mathbf{f}''\rangle$ over all Se atoms in the cell must first be calculated, and then the tensorial products ${}^t\mathbf{a}\langle\mathbf{f}'\rangle\mathbf{a}$ *etc.* are formed. Results are summarized in Table 6. The refined diffraction values are within 0·3 e of the values from absorption spectra, which is on the order of the estimated standard deviation values from the refinement except for $\langle\mathbf{f}''\rangle_c$ at 0·9809 Å.

The orientation of the principal-axes system and the principal values from diffraction data seem consistent with atomic model and absorption spectra, respectively. This indicates that the anisotropy, although very small, is significant. Nevertheless, refinement of local parameters taking into account this small anisotropy does not improve the phases, which suggests that in this case of low energy resolution the error introduced by neglect of anisotropy is negligible compared to other errors.

The difference between the high anisotropy observed by Templeton & Templeton (1988), and the low anisotropy observed here can be related to the different energy spreads of the incident beam: ~2 eV with a Si(220) monochromator at SSRL for the selenolanthionine experiment as against on the order of 10 eV in the Photon Factory experiment. This hypothesis can be tested because absorption spectra of selenobiotinyl streptavidin have been recorded both at Stanford with a 3–4 eV energy spread from a

Ge(111) monochromator [Hendrickson *et al.*, (1989), reproduced on Figs. 1(*e*), (*f*)] and at the Photon Factory (Figs. 1*a*, *b*). The energy distribution from the monochromator is taken to be Gaussian:

$$p(E) = \exp[-(E - E_0)^2/2\sigma^2].$$

As noted earlier, the SSRL and Photon Factory experiments also appear to differ in energy calibration. Thus, a quantitative comparison of the spectra required an energy translation ΔE as well as convolution with the point spread function. The SSRL absorption spectra were fitted to the Photon Factory spectra using various values of σ and ΔE. For the $\mathbf{a}\|\mathbf{E}$ spectrum the best match was obtained for $\sigma = 12$ and a 17 eV translation. These values were 10 and 18 eV, respectively, for the $\mathbf{c}\|\mathbf{E}$ spectrum. The $f'(\omega)$ curves (Figs. 1*c*, *d*) calculated from such convolved spectra (with intermediate values of $\sigma = 11$ and 17·5 eV) are very similar to Photon Factory curves (Figs. 1*a*, *b*). This indicates that a large energy spread averages out the AS anisotropy and also reduces the extremes of AS features. Incidentally, we observed that the $\sigma = 11$ convolution on the $\mathbf{c}\|\mathbf{E}$ spectrum produces a 12 eV shift toward higher energy in the f'' peak, and a 5 eV shift toward lower energy in the f' extremum (which corresponds to the inflection point of the f'' curve). This would indicate that the inflection point is a better reference point for energy calibration.

5. Concluding remarks

The effects of AAS add considerable complexity to the description of X-ray diffraction in general and to the analysis of MAD data in particular. Observed intensities can no longer be decomposed into structure-independent and structure-dependent factors. Instead, the 'structure factors' now also depend on the orientation of the anisotropic scattering tensors with respect to directions of polarization of the incident and diffracted X-ray beams. In the general situation, the intensity formula comprises 12 terms and even in the more usual circumstance of phase incoherence there are separate terms for the four pairs of incident and diffracted polarization directions. Nevertheless, this rather general analysis of AAS in the context of MAD phase determination is quite straightforward as implemented in *MADDST*.

That AAS is significant is very evident from the data on selenolanthionine (Templeton & Templeton, 1988), and this chemical situation is very relevant to studies under way on selenomethionyl proteins (Hendrickson *et al.*, 1990). Thus it is encouraging from the simulation on selenobiotinyl streptavidin that, in the isotropic *MADLSQ* analysis, average phase errors were limited to 9° despite anisotropy such that at one wavelength f'' components range from 0·8 to 7·7 e. Moreover, in this situation of perfect data, the starting point given by the isotropic approximation proved adequate for the determination of AS tensors and for a full recovery of phases freed of AAS effects. In experiments at lower energy resolution in the monochromatic beam (10 *vs* 2 eV), although AAS is still evident in pleiochroic absorption spectra, the effect of AAS on the MAD analysis is mostly averaged out - here the isotropic approximation only imparts a 2° error on the phases.

This study supports the viability of a strategy for phase determination based on exploiting the extreme anomalous-scattering factors associated with 'white lines' at high energy resolution (1–3 eV) despite the complications of AAS that attends these features in certain cases. The error introduced in the initial isotropic assumption of *MADLSQ* will, of course, increase with the concentration of AS centers. Alternative strategies of using a reduced energy resolution and/or avoiding wavelengths of greatest anisotropy could be advantageous in such situations.

In this paper, the absorption correction problem has not been addressed. The linear absorption coefficient can also be anisotropic (pleiochroism). In that case (9) should be rewritten as

$$\begin{aligned} I(\eta) \sim\ & (1 + P_L)/2\{{}_{uu'}A \,.\, |G_T(\hat{\mathbf{u}}\hat{\mathbf{u}}'|\eta)|^2 \\ & + {}_{uv'}A \,.\, |G_T(\hat{\mathbf{u}}\hat{\mathbf{v}}'|\eta)|^2\} \\ & + (1 - P_L)/2\{{}_{vu'}A \,.\, |G_T(\hat{\mathbf{v}}\hat{\mathbf{u}}'|\eta)|^2 \\ & + {}_{vv'}A \,.\, |G_T(\hat{\mathbf{v}}\hat{\mathbf{v}}'|\eta)|^2\}, \end{aligned}$$

where

$${}_{ee'}A = \int \exp[-{}^t\hat{\mathbf{e}}\boldsymbol{\mu}\hat{\mathbf{e}}t_1(\mathbf{r}) - {}^t\hat{\mathbf{e}}'\boldsymbol{\mu}\hat{\mathbf{e}}'t_2(\mathbf{r})]\, d^3\mathbf{r}$$

is the transmission factor, and $t_1(\mathbf{r})$, $t_2(\mathbf{r})$ are the path lengths associated with diffracting volume centered on $\mathbf{r}$ and $\boldsymbol{\mu}$ is the anisotropic absorption tensor. Obviously this is another contribution to the systematic error induced by the isotropic approximation.

Another type of scan can be performed to demonstrate directly the AAS, without interference with the 'geometric' effect of the absorption correction (different path lengths in the crystal). If there is a linearly polarized incident beam, a given node $\mathbf{h}$ could be recorded in different conditions by simply rotating the polarization direction around the incident beam. Presented in this way that means the 'goniometer+detector' ensemble must be rotated, which is not easily achieved. But one can equivalently rotate the crystal alone around the incident-beam axis. In this case the scattering plane rotates and the diffraction spot describes a circle in the detector plane. The only requirement for this type of scan is therefore the availability of an appropriate area detector and a conventional four-circle goniometer.

It has been proposed that the AAS could be exploited to locate the anomalous scatterers in macromolecular structures (Templeton & Templeton,

1986). This seems difficult because in the general case this would necessitate knowing the orientation of each independent AS atom. Only in special cases is it possible to evaluate the position from the modulation of an azimuthal scan. However, another application could be to use the AAS to resolve the twofold ambiguity inherent in a single-wavelength experiment. Actually, the variation of the effective AS factors with orientation could be exploited in a way similar to the MAD method. The data would be measured at a single wavelength near the absorption edge, but at different orientations as described above. No wavelength variations being needed, the experimental setup is simplified and the problem of scaling data collected at different wavelengths is eliminated. However, the implementation is more complicated since the effective AS factors vary in an unpredictable way as long as the orientation of molecular axes is unknown, and the signal-to-noise ratio may be lower than in the MAD experiment.

We thank Lieselotte and David Templeton for sharing their results on selenolanthionine ahead of publication, and Jean-Louis Staudenmann and Boris W. Batterman for helpful discussions. This work was supported in part by grant no. GM 34102 from the National Institutes of Health.

References

BRICOGNE, G. (1982). In *Computational Crystallography*, edited by D. SAYRE, pp. 223–230. Oxford: Clarendon Press.

CROMER, D. T. (1983). *J. Appl. Cryst.* **16**, 437.

DICKERSON, R. E., WEINZIERL, J. E. & PALMER, R. A. (1968). *Acta Cryst.* B**24**, 997–1003.

GUSS, J. M., MERRITT, E. A., PHIZACKERLEY, R. P., HEDMAN, B., MURATA, M., HODGSON, K. O. & FREEMAN, H. C. (1988). *Science*, **241**, 806–811.

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

HENDRICKSON, W. A. (1987). In *Crystallography in Molecular Biology*, edited by D. MORAS *et al.*, pp. 81–87. New York: Plenum.

HENDRICKSON, W. A., HORTON, J. R. & LEMASTER, D. M. (1990). *EMBO J.* **9**, 1665–1672.

HENDRICKSON, W. A., PÄHLER, A., SMITH, J. L., SATOW, Y., MERRITT, E. A. & PHIZACKERLEY, R. P. (1989). *Proc. Natl Acad. Sci. USA*, **86**, 2190–2194.

HENDRICKSON, W. A., SMITH, J. L., PHIZACKERLEY, R. P. & MERRITT, E. A. (1988). *Proteins: Struct. Function Genet.* **4**, 77–78.

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

JACKSON, J. D. (1975). *Classical Electrodynamics*, 2nd ed. New York: Wiley.

KAHN, R., FOURME, R., BOSSHARDT, R., CHIADMI, M., RISLER, J. L., DIDEBERG, O. & WERY, J. P. (1985). *FEBS Lett.* **179**, 133–137.

KARLE, J. (1980). *Int. J. Quant. Chem. Symp.* **7**, 356–367.

MATERLIK, G. & SUORTTI, P. (1984). *J. Appl. Cryst.* **17**, 7–12.

MOFFAT, K. (1988). *Nature (London)*, **336**, 422–423.

PÄHLER, A. & HENDRICKSON, W. A. (1990). In preparation.

PÄHLER, A., SMITH, J. L. & HENDRICKSON, W. A. (1990). *Acta Cryst.* A**46**, 537–540.

STAUDENMANN, J.-L. & CHAPMAN, L. D. (1989). *J. Appl. Cryst.* **22**, 162–172.

SYGUSCH, J. (1977). *Acta Cryst.* A**33**, 512–518.

TEMPLETON, D. H. & TEMPLETON, L. K. (1982). *Acta Cryst.* A**38**, 62–67.

TEMPLETON, D. H. & TEMPLETON, L. K. (1985). *Acta Cryst.* A**41**, 365–371.

TEMPLETON, D. H. & TEMPLETON, L. K. (1986). *Acta Cryst.* A**42**, 478–481.

TEMPLETON, L. K. & TEMPLETON, D. H. (1988). *Acta Cryst.* A**44**, 1045–1051.

VAILLANT, F. (1977). *Acta Cryst.* A**33**, 967–970.