Determination of Macromolecular Structures from Anomalous Diffraction of Synchrotron Radiation

WAYNE A. HENDRICKSON

Resonance between beams of x-ray waves and electronic transitions from bound atomic orbitals leads to a phenomenon known as anomalous scattering. This effect can be exploited in x-ray crystallographic studies on biological macromolecules by making diffraction measurements at selected wavelengths associated with a particular resonant transition. In this manner the problem of determining the three-dimensional structure of thousands of atoms is reduced to that of initially solving for a few anomalous scattering centers that can then be used as a reference for developing the entire structure. This method of multiwavelength anomalous diffraction has now been applied in a number of structure determinations. Optimal experiments require appropriate synchrotron instrumentation, careful experimental design, and sophisticated analytical procedures. There are rich opportunities for future applications.

DRAMATIC ADVANCES IN THE FIELD OF STRUCTURAL BIOLOGY have occurred in the last few years, especially for studies on biological macromolecules at the atomic level of detail. Reports of new structures appear with remarkable frequency along with an even greater production of highly informative follow-up studies on biochemical and biophysical properties. X-ray crystallography, nuclear magnetic resonance (NMR) spectroscopy, and theoretical simulation are the three dominant methods used in these analyses, and recent technical developments have greatly enhanced each of these methods. Major enhancements have come from advances in instrumentation, from new analytical methods, and through recombinant DNA technology.

In the case of crystallography, several developments in instrumentation contribute to the conspicuous success of this technique. These include area detectors, advanced computers, and graphical display units. However, synchrotron radiation is having a truly exceptional impact; it makes possible experiments that would otherwise be effectively unthinkable. Indeed, the applications go far beyond what could have been envisioned when synchrotron radiation was first brought to the attention of biological crystallographers about 20 years ago (*1*).

Until now, the major use of synchrotron radiation in macromolecular crystallography has taken advantage of the exceptional intensity of synchrotron x-rays to speed up data collection generally and to permit the investigation of samples (very small crystals or ones with very large unit cells) that could not be studied with conventional sources (*2*). A second category of application is time-resolved crystallography, which exploits the continuous spectral distribution and unique time structure of synchrotron radiation in Laue diffraction experiments (*3*). This area is in active development for following enzyme reactions and photochemical processes. The third major use of synchrotron radiation in macromolecular crystallography is the subject of this article. The continuously intense spectral output from a synchrotron allows determination of crystal structures directly from measurements of anomalous diffraction made at multiple wavelengths.

The method of multiwavelength anomalous diffraction (MAD) has now been used to solve several crystal structures, and a number of other MAD analyses are in progress. MAD experiments are more demanding of instrumentation than are conventional, fixed wavelength experiments, and the data-processing requirements are also somewhat different. Nevertheless, facilities for MAD experiments are now on the verge of being routinely available at some synchrotron beam lines and the necessary software is already available. This progression of MAD into a mature methodology has required a synergy between developments in instrumentation, experimental design, and analytical procedures. The method has been elaborated in a recent review (*4*). In this article I discuss the fundamentals of the MAD method, outline the essential characteristics of MAD experiments, and describe a few applications.

Anomalous Scattering and the Phase Problem

X-ray diffraction patterns from macromolecular crystals typically comprise many thousands of unique reflection intensities, and these can be measured quite accurately. Each intensity depends on the atomic parameters of all of the atoms in the crystal structure, and this is at once the great strength and the essential difficulty of the method. Although the observations permit an overdetermination of the atomic structure (provided that data extend to Bragg spacings at least as fine as 3 Å), the extraction of atomic parameters is complicated. In order to compute an image of the structure, phases are needed for the diffracted waves as well as their measured amplitudes. This phase problem is the central obstacle in crystal structure analysis.

In the case of small molecules, a direct solution of the phase problem is possible from statistical relations among the intensities. Unfortunately, these direct methods alone have until now proved too weak for macromolecules (*5*). Initial phases for new macromolecular structures have usually been developed by the method of isomorphous replacement with heavy atoms (*6*). Differences in diffracted intensities from the derivative and native crystals are used to locate the heavy atoms. The calculated contributions from these centers serve as reference waves for evaluation of phases for the

The author is in the Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032.

native crystal structure. In general, the phase information from a single isomorphous pair is ambiguous. Thus, multiple isomorphous replacements (MIR) or supplementary information such as from anomalous scattering, molecular averaging, or solvent flattening is used to resolve the ambiguity. Anomalous scattering measurements when made at appropriate multiple wavelengths can also provide a definitive experimental solution for the phase problem. This MAD method for evaluating phases is akin to the MIR method in requiring especially distinctive scattering centers, and it is similar to direct methods in that all of the necessary data can be recorded from a single crystalline species.

Details of the MAD approach to de novo phasing of macromolecular crystal structures are given in Box 1. Simply put, MAD experiments can be thought of as in situ isomorphous replacements in which physics rather than chemistry is used to effect the change in scattering strength at the site. Unlike MIR, a series of changes in common at a single constellation of sites suffices in general for definitive phase determination because of the phase shifts that accompany anomalous scattering. The MAD method has potential advantages for accuracy in phase evaluation in that isomorphism is perfect, relative scattering strength and phasing power increase with scattering angle, diffraction data can be measured from a single crystal (or set of equivalent crystals), and algebraically exact analysis is possible (7).

Box 1. Fundamentals of anomalous scattering and MAD phase evaluation. The diffraction pattern from a crystal has discrete reflections that vary in intensity with a strong dependence on the atomic structure of the crystal. The structural dependence of the intensity $I(\mathbf{h})$ for a reflection identified by indices $\mathbf{h}(h,k,l)$ is given by

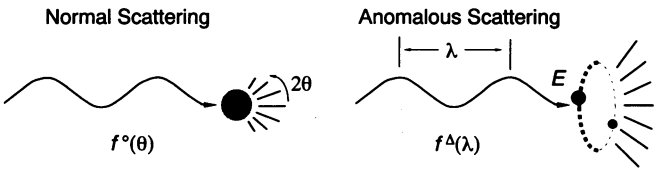
$$I(\mathbf{h}) \propto |F(\mathbf{h})|^2 \quad (1)$$

where

$$F(\mathbf{h}) = \sum_{j}^{\text{atoms}} f_j \exp(-B_j s^2)\exp(2\pi\, \mathbf{h} \cdot \mathbf{x}_j) \quad (2)$$

is the structure factor, f_j is the atomic scattering factor for the jth atom of the crystal, $s = \sin\theta/\lambda$ (where θ is the Bragg angle of reflection and λ is the x-ray wavelength), B_j is the jth atomic thermal or mobility parameter, and $\mathbf{x}_j$ (x,y,z) defines the jth atomic position vector.

The atomic scattering factor describes the coherent scattering from an isolated atom relative to the Thomson scattering expected from a free electron. This factor includes a component that depends strictly on the electron density distribution in the atom and a dispersive (frequency-dependent) component that is related to bound electronic states of the atom. Anomalous dispersion of x-ray optical properties accompanies the resonant absorption of energy in the transition from a bound atomic orbital to an electronic state in the continuum. This process leads to an anomalous scattering increment, f^{Δ}, to the normal scattering factor, f°. The f° component is purely real, is independent of wavelength, and falls off with scattering angle because of diffuseness in outer electron shells. In contrast, f^{Δ} is complex, has a marked wavelength dependence due to the resonance, and is virtually independent of scattering angle because of its origin in core electrons:



The total scattering factor is then

$$f = f^{\circ} + |f^{\Delta}|\, e^{i\delta} = f^{\circ} + f' + if'' \quad (3)$$

where f' and f'' are the real and imaginary components of the anomalous scattering, respectively.

Three means are available for evaluating anomalous scattering factors for use in MAD analyses: (i) Quantum mechanical calculations on isolated elemental atoms (*37*) are accurate for wavelengths (energies) away from absorption edges. However, these do not apply near an edge where values depend in general on transitions to unoccupied molecular orbitals. (ii) Fittings can be made in the analysis of diffraction measurements recorded at specific wavelengths (*9, 22, 26*). (iii) Values can be obtained from x-ray absorption spectra because the imaginary component relates to the atomic absorption coefficient μ_a according to $f''(E) \propto E \cdot \mu_a(E)$, where E is the energy of the x-ray wave and the spectrum of real values follows by Kramers-Kronig transformation from f'' values (*38*).

The impact of anomalous scattering on diffraction measurements can be evaluated by substituting the scattering factor expression, Eq. 3, into the structure factor equation, Eq. 2, and then obtaining the squared modulus, which is of direct relevance to the actual intensity measurements, Eq. 1. The number of terms in the resulting expression depends on the number of distinctive kinds of anomalous scatterers (*7, 14*). However, in the commonly encountered case of a single kind of anomalous scatterer, this takes a simple form that separates known wavelength-dependent factors from unknown wavelength-independent variables. The total structure factor obtained at a particular wavelength λ for a particular reflection $\mathbf{h}$ is denoted ${}^{\lambda}F(\mathbf{h})$. The wavelength-invariant contribution from just the normal f° components of the scattering factors is given by ${}^{\circ}F_T = |{}^{\circ}F_T| \exp i^{\circ}\varphi_T$, and the part contributed solely by the normal scattering component of the anomalous centers is given by ${}^{\circ}F_A = |{}^{\circ}F_A| \exp i^{\circ}\varphi_A$. Then

$$|{}^{\lambda}F(\pm\mathbf{h})|^2 = |{}^{\circ}F_T|^2 + a(\lambda)|{}^{\circ}F_A|^2 + b(\lambda)|{}^{\circ}F_T||{}^{\circ}F_A|\cos({}^{\circ}\varphi_T - {}^{\circ}\varphi_A) \pm c(\lambda)|{}^{\circ}F_T||{}^{\circ}F_A|\sin({}^{\circ}\varphi_T - {}^{\circ}\varphi_A) \quad (4)$$

where, in the notation of Eq. 3,

$$a(\lambda) = (f'^2 + f''^2)/f^{\circ 2} \quad (5a)$$

$$b(\lambda) = 2(f'/f^{\circ}) \quad (5b)$$

$$c(\lambda) = 2(f''/f^{\circ}) \quad (5c)$$

Reflections $\mathbf{h}$ and $-\mathbf{h}$ are known as Friedel mates, and in the absence of anomalous scattering they have identical intensities. The actual difference

$$\Delta F_{\pm\mathbf{h}} = |{}^{\lambda}F(\mathbf{h})| - |{}^{\lambda}F(-\mathbf{h})| \quad (6)$$

between the structure amplitudes of Friedel mates, or respective rotational symmetry equivalents of them, is designated as the Bijvoet difference. On the basis of Eq. 4, it depends on $\sin(\Delta\varphi = {}^{\circ}\varphi_T - {}^{\circ}\varphi_A)$ and on $f''(\lambda)$. The difference

$$\Delta F_{\Delta\lambda} = |\overline{{}^{\lambda_i}F}| - |\overline{{}^{\lambda_j}F}| \quad (7)$$

where $|\overline{{}^{\lambda}F}| = [|{}^{\lambda}F(\mathbf{h})| + |{}^{\lambda}F(-\mathbf{h})|]/2$, is designated as the dispersive difference. It depends on $\cos(\Delta\varphi)$ and on $|f'(\lambda_i) - f'(\lambda_j)|$ (Eq. 4). Clearly, Bijvoet and dispersive differences provide orthogonal phase information and are thus complementary; moreover, the strength of these differences depends in a predictable way on the magnitudes of anomalous scattering factors.

The diffraction measurements at different wavelengths and for Bijvoet mates are each described in the form of Eq. 4. Given that f' and f'' have been evaluated so that $a(\lambda)$, $b(\lambda)$, and $c(\lambda)$ are known, and provided that sufficient appropriate measurements have been made, this system of equations can be solved for the desired unknown parameters, $|{}^{\circ}F_T|$, $|{}^{\circ}F_A|$, and $\Delta\varphi = {}^{\circ}\varphi_T - {}^{\circ}\varphi_A$. In practice, this is done by first solving for the four wavelength-independent factors in Eq. 4 and then imposing the identity that $\cos^2\theta + \sin^2\theta = 1$ as a LaGrange constraint in a nonlinear least-squares procedure (*14, 16*). The set of derived $|{}^{\circ}F_A|$ coefficients can then be used to deduce the locations of anomalous scattering centers. From these coordinates, ${}^{\circ}\varphi_A$ can be calculated and thus the desired phase ${}^{\circ}\varphi_T$ is determined. A number of enhancements to this basic analytical procedure have been made.

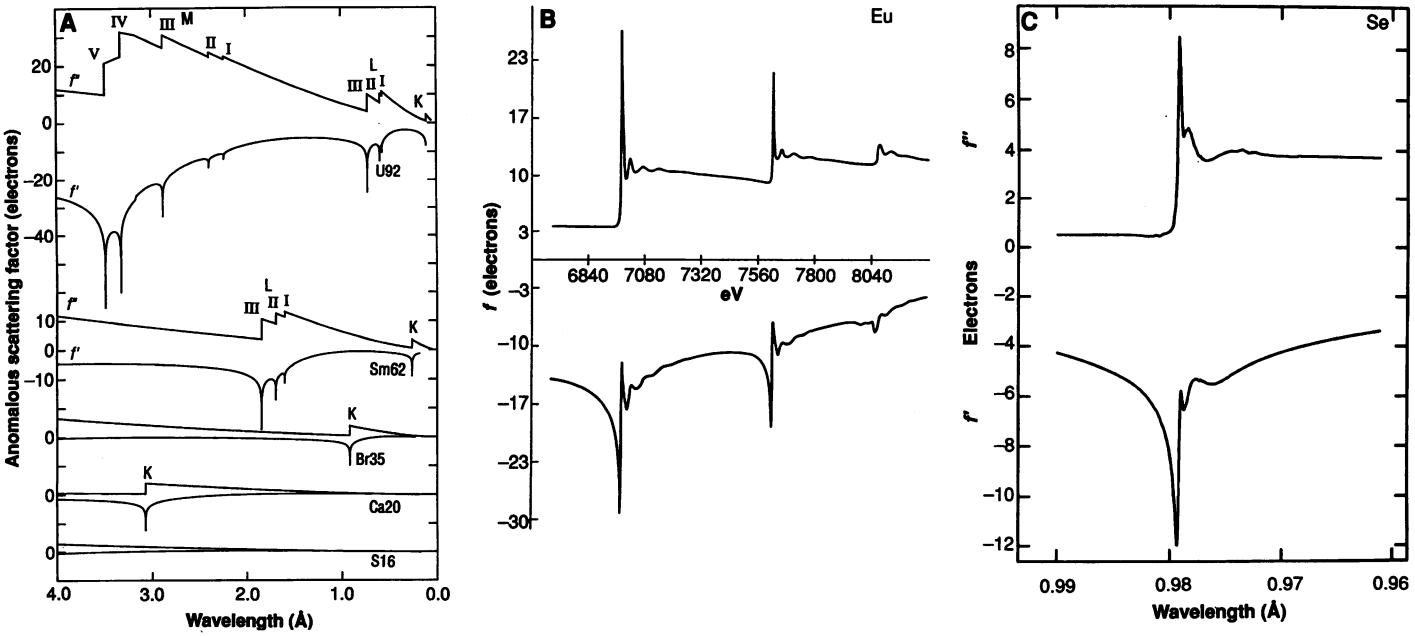


Fig. 1. Anomalous scattering factor spectra for selected elements. (**A**) Calculated spectra (*37*) for certain isolated atoms. For each element, the imaginary component (f'') is drawn in the upper curve and the real component (f') is in the lower curve. Origins for the five elements are displaced vertically as indicated. The figure is adapted from (*14*). (**B**) Experimental values derived from an x-ray absorption spectrum of Eu-(PhAcAc)$_3$ (Ph, phenyl, and Ac, acetyl). The resonances from left to right are associated with the L_{III}, L_{II}, and L_I transitions. The L_{III} maximum in f'' occurs at 6982.2 eV, which corresponds to 1.7757 Å in wavelength. The figure is adapted from (*8*). (**C**) Experimental values derived from an x-ray absorption spectrum of selenomethionyl thioredoxin from *Escherichia coli*. Reproduced from (*34*).

Anomalous scattering is intimately associated with the resonant absorption of x-rays that occurs when the frequency of the incident radiation approaches the frequency of oscillations in a bound electronic orbital (see Box 1). This anomalous dispersion (frequency dependence) of scattering is, as the resonance phenomenon suggests, most pronounced in the immediate vicinity of the absorption edge. The resonances associated with K and L_{III} absorptions are of greatest interest for MAD experiments, and these must of course be within the accessible x-ray spectrum to be useful. The anomalous scattering profiles for isolated atoms of a few elements are shown in Fig. 1A. These examples illustrate a number of points: (i) Edge positions for a given orbital occur at systematically increasing energy (shorter wavelength) as the atomic number increases. (ii) Apart from the energy of transition, all K edges are essentially alike and all L edges are alike. (iii) L_{III} edges, which are associated with the six $2p$ electrons, have anomalous scattering factor magnitudes on the order of three times greater than those for K edges, which are associated with the two $1s$ electrons. As shown in Fig. 1, B and C, the anomalous scattering profiles that are actually observed from molecules are typically more strongly featured than the calculated ones for isolated atoms. These resonant "white-line" features, which correspond to transitions to unoccupied molecular orbitals, can be threefold greater than those expected from isolated atoms (*8, 9*).

An accessible spectral range for MAD experiments can be considered to be the window from ~0.3 to ~3.0 Å in wavelength (4 to 40 keV in energy). This range includes the K edges from atomic number Z = 20 (Ca) to Z = 58 (Ce) and L_{III} edges from Z = 51 (Sb) to Z = 92 (U). Thus, all elements at least as heavy as calcium are possible candidates for MAD experiments. Diffraction experiments have also been conducted at the S K edge (λ = 5.02 Å), but these require vacuum chambers and very thin samples (*10*). Even experiments at the Ca K edge would be compromised by absorption. Experience at the high-energy extreme is very limited, but such experiments should be readily feasible. All elements normally used as heavy atoms in MIR plus many that are too light for use as MIR derivatives are all well suited for MAD experiments. Such elements either occur naturally in macromolecules or can be introduced in a variety of ways (Table 1).

X-ray Sources

MAD experiments require sources capable of producing x-rays at various suitable wavelengths. With conventional x-ray tubes the options are limited primarily to the characteristic lines from usable target materials. Nevertheless, a pioneering demonstration of feasibility was performed on *Chironomus* hemoglobin by using two x-ray tubes (*11*), and instruments have been developed to use multiple lines from mixed targets or L emissions (*12*). Indeed, MAD data measured from a Pt derivative with Au L-line x-rays were used together with MIR data to determine the structure of a CD4 fragment (*13*).

The bremsstrahlung continuum from x-ray tubes is another possible source for MAD experiments, and wavelengths selected from the continuum emitted by a Mo anode were used to solve the structure of selenolanthionine (*14*), a small molecule. The bremsstrahlung intensity is much weaker than that in characteristic lines, but it can be optimized by using a high-Z anode material such as Au.

Although MAD experiments are possible at conventional "home" sources, the sporadic availability of characteristic emissions and the relative weakness of the bremsstrahlung (Fig. 2) are limitations. On the other hand, the spectral brightness of synchrotron radiation (Fig. 2) is well suited for MAD work. The bending-magnet radiation from several existing sources (CHESS at Cornell, NSLS at Brookhaven, SSRL at Stanford, the Photon Factory in Tsukuba, LURE in Orsay, DESY in Hamburg, and SRS at Daresbury) can provide adequate flux for many experiments. Wigglers give enhanced flux and extend the spectrum to higher energy and can make third-generation, low-energy sources suitable (such as ALS at Berkeley, MAX-II in Lund, SRRC in Hsinchu). Undulators on the high-energy, low-emittance sources that are under construction

Table 1. Categories of MAD application. An accessible x-ray range from ~0.3 to ~3.0 Å wavelengths covers the resonant transition from atomic numbers 20 (Ca) to 58 (Ce) for K absorptions and from 51 (Sb) to 92 (U) for the stronger L_{III} absorptions. Apart from Ca, all of the likely candidates listed below fall within a range of wavelengths from 0.6 to 1.8 Å for which experiments are especially convenient.

Category	Typical element
Metalloproteins	
Transition metals	Fe, Cu, Zn, Mn
Other	Ca, Mo
Metal replacements	
Lanthanides for Ca^{2+}, Mg^{2+}	Tb, Ho, Yb
Mercury for zinc	Hg
Ligand analogs	
Arsenates, selenoethers, organohalides	As, Se, Br
Heavy-atom complexes	
Common protein derivatives	Pt, Au, Hg, U
Cluster compounds	Ta, W
Building-unit analogs	
Selenomethionyl residues	Se
Brominated nucleotides	Br

(APS at Argonne, ESRF in Grenoble, and SPring-8 near Kobe) would yield exceptionally bright x-ray radiation over a narrower but tunable range.

The requirements in beam-line optics for MAD experiments differ somewhat from those needed for fixed-wavelength data collection. In addition to the obvious requirement for tunability, it is highly desirable to have relatively high-energy resolution (2×10^{-4} to 5×10^{-4} $\Delta E/E$) to match the intrinsic widths of a few electron volts for the resonant transitions. Many different optical configurations have been used (Fig. 3). Most commonly, the beam path includes a double-crystal monochromator to provide a fixed exit angle for the beam and the use of a mirror or a bent second crystal or both to focus the beam into the small sample crystal.

Experimental Procedures

The strategy used in a MAD experiment must be adapted to the data accuracy requirements of the situation, to the particular crystal parameters and properties, and to the idiosyncrasies of the source and the experimental apparatus. Obviously, the data must be measured at a level of accuracy commensurate with the expected signals. As can be seen from the tabulated examples (Table 2), these signals range from being quite small (~2%) for K-edge scatterers at low concentration to being quite large at the white lines of abundant L-edge scatterers (~25%). From the successful precedent of crambin (*15*) we know that even small signals can be measured with sufficient accuracy in well-designed experiments. Many of the procedures now used routinely were first worked out in the MAD application to lamprey hemoglobin (*16*).

The first step in an experimental design is to select wavelengths that optimize the signals. Since the scattering factors that determine these signals depend on the initially unknown chemical environment of the site and on the energy resolution of the beam, an experimental evaluation is essential for each case. X-ray absorption spectra measured by fluorescence detection are used directly to identify (at the inflection point and the peak of absorption) the wavelengths of scattering-factor extrema needed for optimal signals. These data can be used later to evaluate the scattering factors themselves (*16*). Computer simulations and analytical assessments have been made to guide the choice of wavelengths (*17*); however, practical considerations typically modify this choice. The number of wavelengths required is dictated by information content. In principle, two could suffice, but in practice three to five have been used to increase overdeterminacy.

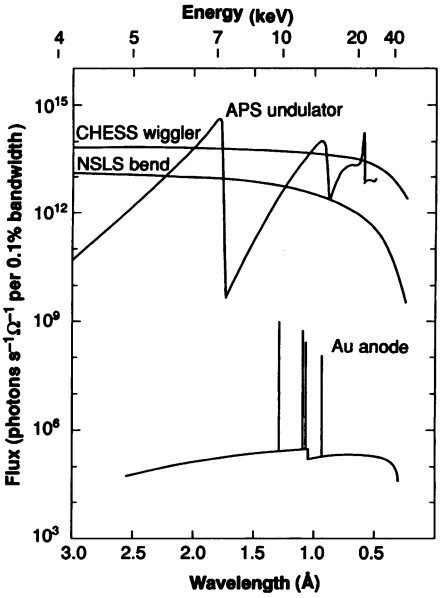


Fig. 2. Spectral flux from representative x-ray sources. The unit of angular acceptance Ω differs among the sources as defined below. The undulator spectrum is estimated for APS undulator A operated at 7.0 GeV and 100 mA. This spectrum was calculated by using previously prescribed procedures (*39*) as the flux through a pinhole (Ω = 50 μrad by 50 μrad) at an undulator gap corresponding to deflection parameter K = 1.5. The positions of the sharp harmonic features can be varied continuously by adjusting this gap. Note that the advantage of undulator radiation is much greater in brightness than in flux. The spectrum for the CHESS F line wiggler is estimated for operation at 5.5 GeV and 100 mA. This spectrum is adapted from figure 5.1 of (*39*) with adjustment for the change in characteristics of the new CHESS wiggler (*40*) from 6 poles to 25 poles and for the increase in standard operating current from 80 to 100 mA. The spectrum for NSLS bending magnet radiation is estimated for operation at 2.5 GeV and 250 mA and is also adapted from figure 5.1 of (*39*) with adjustment to the current standard operating current at NSLS. The acceptance aperature for the wiggler and bending magnet sources is the conventional horizontal opening of Ω = 1 mrad with full vertical integration. The spectrum from a gold rotating anode has been estimated from measurements and calculations of the emission from a tungsten source (*41*). The reported data have been adjusted to a diffraction tube takeoff angle of 3°. An operating current of 200 mA from a rotating anode is assumed. Intensities from the bremsstrahlung continuum are adjusted to a 0.1% bandwidth. The intensities reported here are in units of angular acceptance Ω = 1 mrad by 1 mrad rather than per steradian. The characteristic line intensities have been adjusted according to the reported ratio of AuL_α to WL_α intensities (*41*), and the bremsstrahlung has been adjusted in the ratio of atomic numbers. The energy and wavelength of electromagnetic waves are related by $E = hc/\lambda$ or numerically by E (keV) = 12.3985/λ (Å).

Having chosen the wavelengths, a data collection regimen is established to optimize accuracy. For cases where the expected signals are relatively small, such a regimen has components that are not always present in conventional macromolecular data collections. One is to assure counting times sufficient that random errors do not limit accuracy, and the second is to optimize precision in differences by controlling systematic errors. Absorption, radiation damage, crystal variability, and detector variability can all contribute to systematic error. Thus, one attempts to measure under similar conditions all of the data $\{|^{\lambda}F(\pm\mathbf{h})|\}$ from different wavelengths λ and Bijvoet mates ±**h**, that pertain to a single phase. In order to minimize absorption differences, Bijvoet measurements are usually made to take advantage of mirror symmetry by diad-axis rotation or to exploit centrosymmetry by using the inverse beam geometry. In order to minimize effects of radiation damage, all related measurements are made close together in time, which requires frequent and reproducible change of wavelength. Multiple crystals or crystal orientations can be used, but each individual phasing set should come from one orientation to assure precision in differences.

Users of diffraction facilities at synchrotrons usually have little control over the x-ray optics or detection apparatus at a beam line. It is important nevertheless to take care in monitoring the performance of this equipment in the course of a particular experiment.

For example, periodic x-ray absorption scans of a metal foil need to be made to calibrate wavelength stability. Obviously, checks of alignment of components are also essential. In addition, counting rates or exposure times must be within the dynamic range of the detector. Successful MAD experiments have been carried out with a variety of detectors, including multiwire proportional counters, imaging phosphors, and scintillation detectors. Photographic film, however, is probably not a sufficiently accurate medium for most K-edge MAD experiments.

Analytical Methods

The possibility for phase determination based on anomalous diffraction measurements at multiple wavelengths was appreciated long before such measurements could readily be made (*18*). Probability methods adopted from MIR (*19*) and used in the analysis of crambin from single-wavelength anomalous scattering (*15*) were a significant impetus. However, although MAD data can be analyzed with procedures largely analogous with those used for MIR data, the lack of an actual "native" state makes this awkward. Spurred by the imminent feasibility of synchrotron MAD experiments and an algebraic formalism by Karle (*7*), a powerful analytic alternative to the MIR-based approach was developed (*14*), and this has been used in most of the recent applications.

The theoretical foundation for analyzing MAD experiments is laid out in Box 1. This analysis supposes that all of the data have been placed on a common scale. Once the integrated intensities have been extracted from the measurements, the first step in the analysis is then to perform this scaling. Ordinarily, data from area detectors are used. We have developed procedures for data handling in the course of our analyses of lamprey hemoglobin (*16*), a chromomycin:DNA complex (*20*), ribonuclease H (*21*), and the binding domain of an animal lectin (*22*). The data are first placed on an approximate common scale with conventional merging routines. Then the data from each orientation and each wavelength are separated out, without merging of symmetry equivalents, and a parameterized local scaling (*15*) is made to reduce systematic errors between Bijvoet mates. After this, a second local scaling is made to reduce errors between wavelengths (*16*) and to place the data on an approximately absolute scale that takes into account the wavelength dependence of expected average intensities. At each step, an appropriate rejection of outliers is imposed.

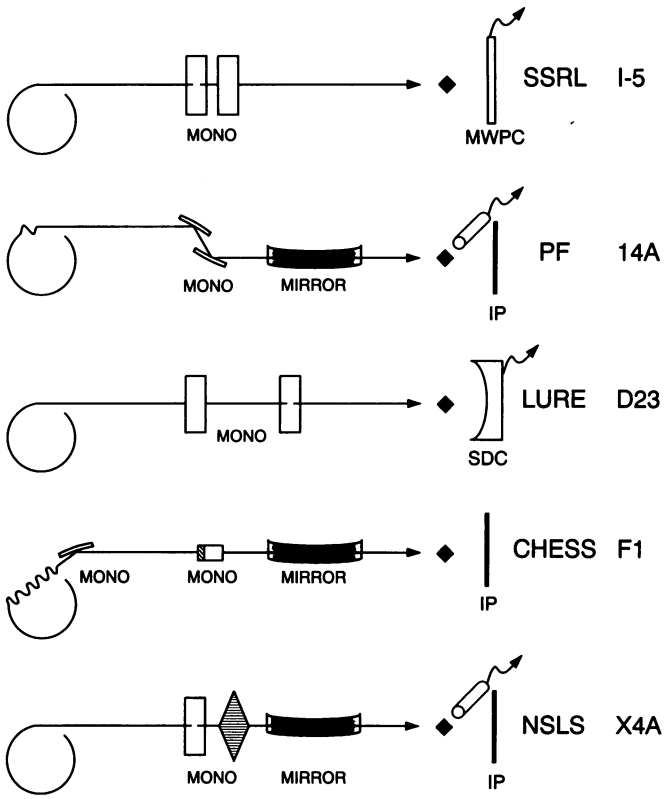


Fig. 3. Schematics of optical configurations at selected beam lines suitable for MAD experiments. The specific optical components are described in reports on the respective beam lines: SSRL I-5 (*29*), Photon Factory 14A (*32*), LURE D23 (*31*), CHESS F1 (*35*), and NSLS X4A (*36*). The components are illustrated here in a highly schematic form. The source of 14A of the Photon Factory is a single-pole vertical wiggler, and the source at CHESS F1 is a multipole wiggler; otherwise, all of the sources are bending magnets. The monochromators (MONO) at the SSRL, Photon Factory and LURE lines illustrated are flat, double-crystal monochromators. The first monochromator at CHESS is a bent triangular crystal intended to produce high flux; a monolithic channel-cut monochromator is inserted to produce high energy resolution. The monochromator at NSLS X4A is a saggitally focusing double-crystal monochromator. The mirrors are metal-coated bent-glass plates. Varying detectors are used. These include a multiwire proportional counter (MWPC) at SSRL, a multiwire spherical drift chamber (SDC) at LURE, imaging phosphor (IP) detector systems at the Photon Factory, CHESS, and NSLS, and single-reflection scintillation detectors at the Photon Factory and NSLS.

Table 2. MAD phasing strength. Expected anomalous diffraction signals are estimated as ratios of averaged diffraction differences relative to averaged total diffraction. Thus, the maximal Bijvoet diffraction ratio is $\text{rms}(\Delta F_{\pm h})/\text{rms}(|F|) = q \cdot f''_{\max}$ (rms, root-mean-square) and the maximal dispersive diffraction ratio is $\text{rms}(\Delta F_{\Delta\lambda})/\text{rms}(|F|) = q \cdot |f'(\lambda_i) - f'(\lambda_j)|_{\max}$, where $q \simeq (N_A/2N_P)^{1/2}/Z_{\text{eff}}$, where N_P is the number of non-hydrogen atoms in the molecule, N_A is the number of anomalous scatterers per molecule, and Z_{eff} is the effective normal scattering at zero scattering angle, 6.7 electrons. N_P can be estimated to be 7.7 atoms per amino acid residue. Except for values reported for a hypothetical selenomethionyl protein, evaluations are computed for particular cases at energy resolutions for use in the actual experiments cited in references. Diffraction ratios for the hypothetical selenomethionyl protein were computed at the extrema of averaged pleiotropic spectra measured from *Escherichia coli* thioredoxin (*34*) to account for the polarization effects.

Molecule	Residues	N_A	N_P	Maximal bijvoet ratio (percent)	Maximal dispersive ratio (percent)	Energy resolution (eV)	Reference
Crambin*	46	6 S	350	1.5			(*15*)
Streptavidin*	252	2 Se	1850	2.5	2.3	~12	(*25*)
Ribonuclease H*	155	3 Se	1190	3.9	3.5	~12	(*21*)
Escherichia coli thioredoxin	108	1 Se	830	5.8	3.5	~3	(*34*)
Hypothetical SeMet	300	5 Se	2310	7.7	4.7	~3	(*34*)
DNA:chromomycin*	32/4	4 Br	980	4.6	3.7	~12	(*20*)
Hemocyanin Og unit	798	4 Cu	6360	2.0	1.8	~8	(*42*)
Animal lectin domain*	230	4 Ho	1980	23.6	8.9	~4	(*22*)

*Solved structures.

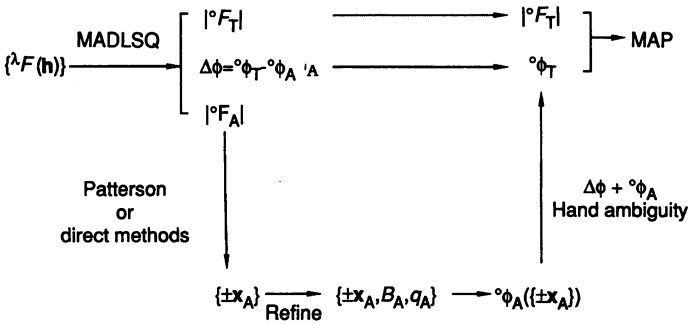


Fig. 4. Flow diagram for the analysis of MAD data in a structure determination. Symbols are defined in the text and in Box 1.

What follows is the MAD analysis itself (Fig. 4). A least-squares fitting is made to extract from each phasing set the independent variables, namely, the respective structure-factor moduli $|^{\circ}F_T|$ and $|^{\circ}F_A|$ and the phase difference $\Delta\varphi = {}^{\circ}\varphi_T - {}^{\circ}\varphi_A$ for the normal scattering contributions from all atoms (T) and from the anomalous centers alone (A). At this point, symmetry equivalents are merged to produce a unique set of phased reflections (*21*). The structure of the anomalous scattering centers $\{\mathbf{x}_A\}$ is then deduced from the merged $|^{\circ}F_A|$ data by Patterson or direct methods and refined to yield the calculated set of reference wave phases, $\{^{\circ}\varphi_A{}^{calc}\}$, and hence the desired set $\{^{\circ}\varphi_T = \Delta\varphi^{obs} + {}^{\circ}\varphi_A{}^{calc}\}$. Since in general the enantiomeric state of a structure is indeterminate from its normal diffraction data, the $\{\mathbf{x}_A\}$ structure contains an inherent ambiguity. This ambiguity is usually readily resolved by chemical reasonableness or by symmetry considerations in the Fourier syntheses from alternative $^{\circ}F_T$ coefficients.

In many instances it is useful to combine phase information from the MAD experiment with that from other sources, such as molecular averaging, solvent flattening, MIR, or a partially refined model. A phase-probability formulation (*23*) facilitates this. Another enhancement for the analysis relates to the complication of anisotropy of anomalous scattering, which is a general consequence of anisotropy in molecular orbitals and is manifest in pleiochroic x-ray absorption (*24, 25*). The effects of anisotropy in anomalous scattering can be accounted for in a phase-refinement procedure (*26*).

Applications

At present several macromolecular structures have been solved by the MAD method for phase determination. The first reports of complete analyses—those on a terbium-substituted parvalbumin (*27*) and on lamprey hemoglobin (*16*)—can be considered as thorough methodology tests, because in both cases protein models determined by other means were used to corroborate the results of MAD phasing. These studies were very important for the development of the experimental and analytical procedures of the method. However, the real task of any method is in its actual practical application. In 1987 (*28*) reports of such applications began to appear, and there has been rapid development since then. These results are summarized in Table 3 along with MAD experiments presently being analyzed in my laboratory. Many other such analyses are in progress elsewhere.

The studies reported in Table 3 include experiments conducted with a variety of x-ray sources and detectors to collect MAD data from several different anomalous scattering elements that have been introduced in various ways. In fact, new structures have now been determined in each of the categories introduced in Table 1. Below I describe briefly each category of MAD application and give a representative example.

The application of MAD phasing to naturally occurring metalloproteins presents an important new opportunity for structure determination. Most metalloproteins contain transition metals, and although these are relatively light atoms, their K edges are well located for MAD experiments. The low mobility typically found at these sites makes them strong phasing labels. Cucumber basic protein is a particularly significant example from among the several metalloproteins in Table 3. This small (10 kD) blue copper protein from cucumber seedlings contains one copper atom at its redox center. The structure was determined from data measured at four wavelengths with the multiwire area detector facility at SSRL (*29*). A correction for coincidence losses due to detector overloading was crucial in the analysis. The structure was interpreted from a 3.0 Å resolution map after refinement of the MAD phases by solvent flattening (*30*). This structure had resisted solution by molecular replacement or MIR methods, despite what proves to be a strong

Table 3. Applications of the MAD method for phase determination. Abbreviations: LURE, Laboratoire pour l'Utilisation du Rayonnement Electromagnétique, Orsay; SSRL, Stanford Synchrotron Research Laboratory; PF, Photon Factory, Tsukuba; CHESS, Cornell High Energy Synchrotron Source; brems., bremsstrahlung continuum; MWPC, multiwire proportional counter; and IP, imaging phosphor.

Molecule	Anomalous element	X-ray source	Detector	Map resolution (Å)	Reference
Parvalbumin*	Tb	LURE	MWPC	2.3	(*27*)
Lamprey hemoglobin*	Fe	SSRL	MWPC	3.0	(*16*)
Selenolanthionine†	Se	Mo brems.	Scintillator	1.0	(*14*)
Ferredoxin*	Fe	SSRL	MWPC	5.0	(*43*)
Urechis hemoglobin†	Fe	SSRL	MWPC	3.0	(*44*)
Cytochrome c'*	Fe	PF	Scintillator	6.0	(*45*)
Streptavidin†	Se	PF, SSRL	Scintillator	3.1	(*25*)
Cucumber basic protein†	Cu	SSRL	MWPC	3.0	(*30*)
Thioredoxin‡	Se, Cu	SSRL	MWPC	3.0	(*34*)
Interleukin-1α*	Se	PF	Scintillator	3.1	(*46*)
DNA:chromomycin†	Br	PF	IP	2.85	(*20*)
Ribonuclease H†	Se	PF	IP	2.2	(*21*)
Hemocyanin Og unit‡	Cu	PF	IP	3.0	(*42*)
CD4 fragment*	Pt	Au lines	MWPC	3.5	(*13*)
Cytochrome c-553†	Fe	PF	IP		(*47*)
Animal lectin domain†	Ho	LURE	MWPC	2.5	(*22*)
Tenascin FnIII domain‡	Se	CHESS	IP	2.5	(*48*)

*Complete phase determinations performed as tests, done only at low resolution, or used in conjunction with other information to develop the structure. †Structures solved by direct phase evaluation from MAD data. ‡Other complete data sets under evaluation from my laboratory.

similarity in folding and Cu coordination to other members of the plastocyanin family. Other metalloprotein candidates for MAD phasing include heme proteins, iron-sulfur proteins, and an ever increasing group of zinc proteins.

Light metal centers can often be replaced by heavier ones. Notably, the replacement of the group II ions Ca^{2+} and Mg^{2+} by lanthanides yields very strong anomalous scattering at the L_{III} edges. Our recent structural analysis of the calcium-dependent carbohydrate recognition domain from an animal lectin, mannose-binding protein, illustrates this well. A recombinant fragment of this protein was crystallized as a dimer with Ho^{3+} ions replacing the two Ca^{2+} sites in each protomer. Data were collected at three wavelengths on the spherical drift chamber detector at LURE (*31*). As expected from the lanthanide white line features shown in Fig. 1B, the measured signals were extraordinarily strong (Table 2). As a result, phases were determined with sufficient accuracy that the model fitted to the MAD-phased map gave an initial *R* value of 0.359 for all observations to 2.5 Å spacings (*22*). Because of an imperfection in wavelength control in this experiment, the ability to refine scattering factors was crucial. MAD experiments with lanthanide replacements or replacements of transition metals by heavier L_{III}-edge scatterers, such as replacement of zinc by mercury, can prove very powerful.

Another category of replacement that can be very effective is in the replacement of a natural ligand or cofactor by an analogous derivatized ligand. This strategy has often been effective in MIR experiments; for MAD the range of labeling elements can be extended to the lighter K-edge scatterers. The structure analysis of streptavidin as its selenobiotinyl complex is an example (Fig. 5A). In this case, the data were measured on a single counter diffractometer on a tunable beam line (*32*) at the Photon Factory in Japan. Streptavidin is a tetramer and the asymmetric unit of the crystal contained one-half of the molecule. The polypeptide chain was traced independently for the two protomers in a MAD-phased map at 3.3 Å resolution (*25*), but the fitting was done after molecular averaging with probabilistic phase combination (*23*). Brominated organic ligands could be quite generally useful for this class of MAD experiments.

Perhaps the greatest opportunity for future MAD experiments lies in the use of conventional heavy-atom derivatives. All of the commonly used substituting elements have L_{III} edges in a readily accessible region near 1 Å wavelengths. The one MAD experiment that has been done with conventional heavy atoms was not performed on a synchrotron. MAD data from a Pt derivative of the D1D2 fragment of CD4 were measured on a multiwire area detector with characteristic lines from a Au anode that bracket the L_{III} edge of Pt. These data were combined with MIR phases to yield an interpretable map. Had it been possible to conduct this experiment on a suitable synchrotron beam line, more definitive initial phasing could have been expected. Since isomorphism is not required, MAD phasing of heavy-atom derivatives both increases the accuracy of phases and extends the range of useful derivatives. Potential disorder at heavy-atom sites remains a problem.

The final category of MAD applications is an exciting one by virtue of its potential for generality. Brominated nucleic acid bases can usefully label nucleic acids without appreciably perturbing their structures. Our analysis of the complex of the antitumor drug chromomycin with a duplex of octanucleotides in which one thymine was replaced by 5-bromouracil illustrates this. The structure of this complex was solved from imaging phosphor data measured at the Photon Factory in Japan (*32, 33*). The DNA structure was directly interpretable (*20*), and phase combination between the MAD probability distributions (*23*) and those from a partially refined model were used to complete the drug structure.

In the case of proteins, the incorporation of selenomethionine in place of methionine residues provides a general vehicle for incorporating MAD labels into proteins (*34*). The analysis of ribonuclease H (Fig. 5B) is a successful example of this approach. The recombinant protein was grown in bacteria with complete incorporation of selenomethionine, and the structure was solved from imaging-phosphor data measured at the Photon Factory in Japan (*32, 33*). In this case the phases were accurate enough to permit an initial interpretation into maps at 2.2 Å resolution (*21*). Higher energy resolution would have given even stronger signals (Fig. 1C and Table 2). The extension to larger proteins with many Se sites poses an exciting challenge for future applications.

Prospects

The MAD method can fairly be said to have emerged with vitality from its long gestation. A lack of readily available, satisfactory instrumentation has certainly impeded practical realization of the promise of MAD phasing in macromolecular structure determination. Even now, these experiments remain rather complex relative to the routine measurements of conventional crystallography. Nevertheless, impressive results have already been obtained in a number of

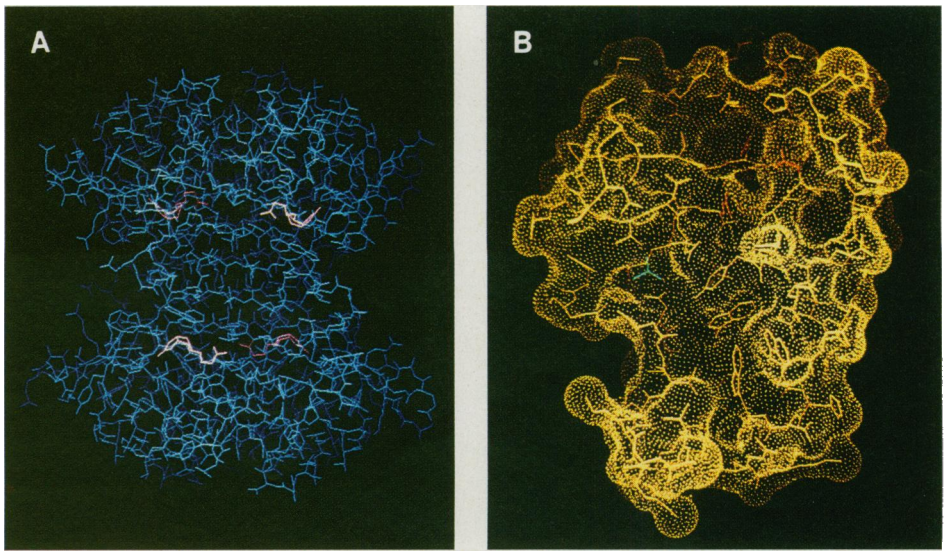


Fig. 5. Illustrations of novel structures determined directly and exclusively by the MAD method. (**A**) Tetramer of core streptavidin, with protein bonds drawn in blue. The selenobiotin groups used for MAD phasing are shown in red. (**B**) Ribonuclease H from *Escherichia coli*, which was solved by use of selenomethionine labels. The molecular surface and most atomic bonds are drawn in yellow, those for catalytic site residues are drawn in red, and a sulfate ion is in green. [Reprinted from (*21*) with permission © AAAS]

cases, effective experimental protocols have been developed, and the analytical methods and computational procedures for processing MAD data are now quite sophisticated. Recent successes of the method are also spurring instrumentation development; for example, an innovative MAD compatibility has been introduced at CHESS (*35*). Moreover, a dedicated beam line designed expressly for MAD experiments is nearly operational at NSLS (*36*). Thus, straightforward access to this new technology may be expected in the near future.

The realm of potential applications of MAD phasing is clearly broad and rich. There is little doubt that given the availability of suitable instrumentation, MAD becomes the method of choice for many problems in biological crystallography. Indeed, with general methods for introducing suitable labels (for example, production of recombinant selenomethionyl proteins or site-directed mutations to cysteines for reaction with mercurials) one might hope to enjoy the assurance, as in small-molecule crystallography, that once a crystal is grown the structure is all but solved. The challenge is to produce at existing and developing synchrotron sources both the necessary instrumentation and also convenient modes of access and use. A long-awaited trip to the synchrotron used to yield packs of films to be processed; now one can return from a few synchrotron centers with processed diffraction amplitudes. In the future, the hope is to return from an on-call trip to one of several MAD beam lines with both the amplitudes and phases needed to solve another exciting biological problem.

REFERENCES AND NOTES

1. G. Rosenbaum, K. C. Holmes, J. Witz, *Nature* **230**, 434 (1971).
2. T. J. Greenhough and J. R. Helliwell, *Prog. Biophys. Mol. Biol.* **41**, 67 (1983).
3. K. Moffat, D. Szebenyi, D. Bilderback, *Science* **223**, 1423 (1984); J. Hajdu and L. N. Johnson, *Biochemistry* **29**, 1669 (1990).
4. R. Fourme and W. A. Hendrickson, in *Synchrotron Radiation and Biophysics*, S. S. Hasnain, Ed. (Horwood, Chichester, 1990), pp. 156–175.
5. G. Bricogne, *Acta Crystallogr.* **A40**, 410 (1984); J. Karle, *ibid.* **A45**, 765 (1989).
6. T. L. Blundell and L. N. Johnson, *Protein Crystallography* (Academic Press, London, 1976).
7. J. Karle, *Int. J. Quantum Chem. Symp.* **7**, 357 (1980).
8. R. C. Lye, J. C. Phillips, D. Kaplan, S. Doniach, K. O. Hodgson, *Proc. Natl. Acad. Sci. U.S.A.* **77**, 5884 (1980).
9. L. K. Templeton, D. H. Templeton, R. P. Phizackerley, K. O. Hodgson, *Acta Crystallogr.* **A38**, 74 (1982).
10. H. B. Stuhrmann, in *Synchrotron Radiation and Biophysics*, S. S. Hasnain, Ed. (Horwood, Chichester, 1990), pp. 223–242.
11. W. Hoppe and V. Jakubowski, in *Anomalous Scattering*, S. Ramaseshan and S. C. Abrahams, Eds. (Munksgard, Copenhagen, 1975), pp. 437–461.
12. G. T. DeTitta, D. C. Swenson, F. Han, W. A. Pangbow, *Abstr. Am. Crystallogr. Assoc. Ser. 2* **18**, 20 (1990); Ng.-h. Xuong, D. Sullivan, C. Nielsen, X. Dai, V. Ashford, *ibid.*, p. 20.
13. S.-E. Ryu *et al.*, *Nature* **348** (1990).
14. W. A. Hendrickson, *Trans. Am. Crystallogr. Assoc.* **21**, 11 (1985).
15. _____ and M. M. Teeter, *Nature* **290**, 107 (1981).
16. W. A. Hendrickson, J. L. Smith, R. P. Phizackerley, E. A. Merritt, *Proteins* **4**, 77 (1988).
17. J. C. Phillips and K. O. Hodgson, *Acta Crystallogr.* **A36**, 856 (1980); R. Narayan and S. Ramaseshan, *ibid.* **A37**, 636 (1981).
18. Y. Okaya and R. Pepinsky, *Phys. Rev.* **103**, 1645 (1956); C. M. Mitchell, *Acta Crystallogr.* **10**, 475 (1957); S. Ramaseshan, K. Venkatesan, N. V. Mani, *Proc. Indian Acad. Sci.* **46A**, 95 (1959); S. Raman, *ibid.* **50A**, 95 (1959); J. Karle, *Appl. Opt.* **6**, 2132 (1967); A. Herzenberg and H. S. M. Lau, *Acta Crystallogr.* **22**, 24 (1967).
19. D. M. Blow and F. H. C. Crick, *Acta Crystallogr.* **12**, 794 (1959); W. A. Hendrickson and E. E. Lattman, *ibid.* **B26**, 136 (1970).
20. C. M. Ogata, W. A. Hendrickson, L. Gao, Y. Satow, Y. Amemiya, in preparation.
21. W. Yang, W. A. Hendrickson, R. J. Crouch, Y. Satow, *Science* **249**, 1398 (1990).
22. W. I. Weis, R. Kahn, R. Fourme, K. Drickamer, W. A. Hendrickson, in preparation.
23. A. Pähler, J. L. Smith, W. A. Hendrickson, *Acta Crystallogr.* **A46**, 537 (1990).
24. L. K. Templeton and D. H. Templeton, *ibid.* **A44**, 1045 (1988).
25. W. A. Hendrickson *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 2190 (1989).
26. E. Fanchon and W. A. Hendrickson, *Acta Crystallogr.* **A46**, 809 (1990).
27. R. Kahn *et al.*, *FEBS Lett.* **179**, 133 (1985).
28. A. Pähler, W. A. Hendrickson, M. A. Gawinowicz Kolks, C. Argaraña, C. Cantor, *J. Biol. Chem.* **262**, 13933 (1987).
29. R. P. Phizackerley, C. W. Cork, E. A. Merritt, *Nucl. Instrum. Methods* **A246**, 579 (1986).
30. J. M. Guss *et al.*, *Science* **241**, 806 (1988).
31. R. Kahn, R. Fourme, R. Bosshard, V. Saintagne, *Nucl. Instrum. Methods* **A246**, 596 (1986).
32. Y. Satow and Y. Iitaka, *Rev. Sci. Instrum.* **60**, 2390 (1989).
33. Y. Amemiya *et al.*, *Nucl. Instrum. Methods* **A266**, 645 (1988).
34. W. A. Hendrickson, J. R. Horton, D. M. LeMaster, *EMBO J.* **9**, 1965 (1990).
35. K. Finkelstein *et al.*, in preparation.
36. J.-L. Staudenmann, W. A. Hendrickson, R. Abramowitz, *Rev. Sci. Instrum.* **60**, 1939 (1989).
37. D. T. Cromer and D. Liberman, *J. Chem. Phys.* **53**, 1891 (1970).
38. R. W. James, *The Optical Principles of the Diffraction of X-rays* (Bell, London, 1948), pp. 135–149.
39. G. K. Shenoy, P. J. Viccaro, D. M. Mills, *Argonne Natl. Lab. Rep. ANL-88-9* (1988); G. K. Shenoy, personal communication.
40. B. W. Batterman, personal communication.
41. D. B. Brown and J. V. Gilfrich, *J. Appl. Phys.* **42**, 4044 (1971); _____, M. C. Pecharar, *ibid.* **46**, 4537 (1975).
42. M. E. Cuff and W. A. Hendrickson, unpublished results.
43. H. M. Krishna Murthy, W. A. Hendrickson, W. H. Orme-Johnson, E. A. Merritt, R. P. Phizackerley, *J. Biol. Chem.* **263**, 18430 (1988).
44. P. R. Kolatkar *et al.*, *Acta Crystallogr.*, in press.
45. S. Harada, M. Yasui, K. Murakawa, N. Kasai, Y. Satow, *J. Appl. Crystallogr.* **19**, 448 (1986).
46. B. J. Graves *et al.*, *Biochemistry* **29**, 2679 (1990).
47. A. Nakagawa, Y. Higuchi, N. Yasuaka, Y. Katsube, T. Yagi, *J. Biochem.* **108**, 701 (1990).
48. D. J. Leahy, H. P. Erickson, W. A. Hendrickson, unpublished results.

49. I especially thank J. L. Smith for her pivotal role in developing MAD methodology. I also thank M. E. Cuff, E. Fanchon, M. H. Hatada, J. R. Horton, P. R. Kolatkar, H. M. Krishna Murthy, C. M. Ogata, A. Pähler, W. E. Royer, Jr., S.-E. Ryu, J. M. Troup, W. I. Weis, and W. Yang for their many contributions to the method and its application; Y. Amemiya, K. Finkelstein, R. Fourme, R. Kahn, E. A. Merritt, R. P. Phizackerley, Y. Satow, J.-L. Staudenmann, and Ng.-h. Xuong for the instrumentation used in our MAD experiments; and G. K. Shenoy and B. W. Batterman for calculations and advice in preparing the figure on spectral fluxes. Supported in part by NIH grant no. GM-34102.