

Crystallographic Structure Analysis of Lamprey Hemoglobin From Anomalous Dispersion of Synchrotron Radiation

Wayne A. Hendrickson,¹ Janet L. Smith,¹ R. Paul Phizackerley,² and Ethan A. Merritt²

¹Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, New York 10032; ²Stanford Synchrotron Radiation Laboratory, Stanford University, Stanford, California 94305

ABSTRACT The molecular structure of lamprey hemoglobin was previously determined and refined by conventional crystallographic analysis. In this study, the structural analysis has been repeated in the course of developing the method of multiwavelength anomalous diffraction (MAD) for phase determination. New experimental and analytical procedures that were devised to perform this determination should have general applicability. These include an experimental design to optimize signal strength and reduce systematic errors, experimental evaluation of anomalous scattering factors, and a least-squares procedure for analyzing the MAD data. MAD phases for the structure at 3 Å resolution are as accurate overall as the multiple isomorphous replacement (MIR) phases determined previously.

Key words: protein structure, diffraction, anomalous scattering, x-ray crystallography

INTRODUCTION

The remarkable success of crystallography in producing detailed pictures of biological macromolecules is founded almost entirely on the method of isomorphous replacement with heavy atoms. Other techniques such as molecular averaging, solvent flattening, or anomalous scattering are often used to supplement isomorphous replacements, and structures related to ones already known can be analyzed by isomorphous differencing or molecular replacement. However, the phase problem remains as the central conceptual difficulty in structural crystallography, and for macromolecules the method of multiple isomorphous replacement (MIR) has been unique in giving an ab initio evaluation of the required phase angles. Reliance on MIR phasing does have a price; namely, the preparation and interpretation of heavy atom derivatives is often rate limiting, and lack-of-isomorphism and intercrystal scaling errors often limit accuracy. One exception to the dominance of the MIR method came in the analysis of crambin from its anomalous scattering alone.¹ Synchrotron radiation affords the opportunity to exploit anomalous scattering to even greater advantage, and this report describes our first application of a new method for ab initio phase determination from multiwavelength anomalous diffraction (MAD) measurements.

Anomalous scattering arises from resonant interactions between x-ray waves and the bound electrons in atoms. This is a perturbation on the dominant,

normal scattering, which is proportional to electron density. X-ray scattering arises in general from vibrations excited in electrons by an incident x-ray wave. When the frequency of the incident wave approaches the natural frequency of a bound electron, the vibrations that are induced resonate with the intrinsic atomic oscillations, and scattering is affected in both amplitude and phase. Just as these atomic orbitals are characteristic of the atomic species, so too is the anomalous scattering distinctive for each kind of atom.

The possibilities for phase determination based on measurements of anomalous diffraction at multiple wavelengths have long been recognized.²⁻⁷ The actual feasibility of such proposals was first demonstrated in a pilot experiment conducted with conventional x-ray tubes,⁸ but it has also been clear from the outset that a continuously tunable x-ray source is needed to optimize the anomalous scattering effects in the context of phase determination. Computational tests⁹ and scattering factor measurements¹⁰ indicated that synchrotron radiation would indeed make this a very powerful method for macromolecular structure determination. For various technical reasons the realization of this bright promise has been rather slow to materialize, but several recent results suggest that the method is rapidly coming of age. Applications using a probability formalism like that of the MIR method have been reported for a parvalbumin,¹¹ a cytochrome c',¹² and an azurin,¹³ and these formed the basis for further analysis by molecular replacement. We have reported in abstract form on applications of the algebraic approach described here to lamprey hemoglobin,^{14,15} selenolanthionine,¹⁶ a bacterial ferredoxin,¹⁵ selenobiotinyl streptavidin,¹⁵ and Urechis hemoglobin.¹⁵ Structure determinations by this MAD phasing procedure have been completed for ferredoxin (H.M. Krishna Murthy, W.A. Hendrickson, W.H. Orme-Johnson, E.A. Merritt, and R.P. Phizackerley, unpublished results) and streptavidin (A.

Received May 6, 1988; revision accepted June 9, 1988.

Address reprint requests to W.A. Hendrickson, Howard Hughes Medical Institute, Department of Biochemistry and Molecular Biophysics, Columbia University, New York, NY 10032.

The present address for J.L. Smith is Department of Biological Sciences, Purdue University, West Lafayette, IN 47907; and the present address for E.A. Merritt is Department of Biological Structure, University of Washington, Seattle, WA 98195.

Pahler, W.A. Hendrickson, J.L. Smith, Y. Satow, E.A. Merritt, and R.P. Phizackerley, unpublished results); the structure of a cucumber basic blue protein has also been determined (J.M. Guss, E.A. Merritt, R.P. Phizackerley, B. Hedman, M. Murata, K.O. Hodgson, and H.C. Freeman, in press).

We chose lamprey hemoglobin as a proving ground for development of MAD phasing methods because it presents a challenging yet theoretically achievable problem. The expected diffraction signals, on the order of 2–3%, are within a range that has produced satisfactory Patterson maps in other cases,^{17–19} but the iron structure would not suffice for resolved anomalous phasing as used for crambin.¹ An additional advantage of lamprey hemoglobin is that MAD phasing results can be compared with MIR and least-squares refined model phases. The crystal structure of type D2 crystals was solved by MIR phasing from three heavy-atom derivatives,^{19,20} and this structure is now highly refined ($R = 0.14$) at 2 Å resolution.²¹ This test on lamprey hemoglobin in the monomeric state was also designed as a forerunner for the analysis of polymeric lamprey hemoglobin in the type Å8 modification,²² which is a structure of particular interest with regard to cooperativity.

MATERIALS AND METHODS

Protein Crystals

Crystals of lamprey hemoglobin were grown from 2.1 M phosphate at pH 6.8 by batch crystallization from component V in the cyanomet state of ligation as described previously.²² The particular crystals used in this study were selected from tubes set up some years before in Warner Love's laboratory. The crystal used in the initial 5.5 Å study measured $0.34 \times 0.62 \times 0.40$ mm along the principal axes of the prismatic form. Crystals of similar size were also used in the subsequent 3.0 Å work. All crystals analyzed were of type D2 with $a = 44.57$, $b = 96.62$, and $c = 31.34$ Å in space group $P2_12_12_1$.²²

Synchrotron Source

Experiments were conducted at the end station of beamline I-5 at the Stanford Synchrotron Radiation Laboratory (SSRL) using the area detector facility designed specifically by Phizackerley et al.²³ to explore MAD phasing techniques. The storage ring was operated in a dedicated mode at 3.0 GeV with electron currents up to 80 mA. A double-crystal monochromator was used to produce a tunable x-ray beam of approximately 1-eV bandwidth over the required range of wavelengths and was not focused. Si(111) monochromator crystals were used in all cases, but they differed in that a symmetric cut was used for the 5.5 Å data, whereas asymmetrically cut crystals were used for the 3 Å data in an effort to achieve a higher incident intensity. The asymmetric cut was designed to produce a beam height that would maximize flux through the collimator. Unfortunately, the antici-

pated enhancement was not realized and such monochromator crystals have not been used subsequently. The angle of asymmetry was 13° for 3 Å set 1 (optimal for a previous experiment at Cs L_{III}), and a cut angle of 9° (optimal for the Fe K edge) was used for sets 2, 3, and 4. The beam path from the source until just before the monochromator located at 20 m was evacuated, and the 5-m path between the monochromator and sample was mostly helium filled. The incident beam was defined by a pair of Nonius 0.8-mm collimators separated by 0.5 m on the incident arm of a Huber goniometer. Beam intensity was monitored by two ion chambers (95% helium, 5% nitrogen) placed between the two defining collimators. The energy of monochromatic x-rays was calibrated to be 7111.2 eV (1.7435 Å) at the first inflection point of x-ray absorption through an iron foil inserted between the two ion-chamber detectors. Wavelengths at other angular settings of the monochromator were calculated assuming that $d(111) = 3.1356$ Å for silicon. Rejection of the Si(333) harmonic was accomplished by a 5–10% detuning of the principal Si(111) reflection combined with the lower intensity of the source at the harmonic energy (21 keV) and a lower quantum detection efficiency of the area detector at this energy.

Area Detector System

The area detector facility at SSRL features a Huber diffractometer and a multiwire proportional chamber (MWPC) detector. This system, described in detail earlier,²³ incorporates an MWPC detector that was built some years ago in a collaboration between SSRL, University of California at San Diego, and Lawrence Berkeley Laboratory.²⁴ The Huber goniometer (250-mm diameter chi circle) is mounted on a motorized table inside a radiation protection hutch. This goniometer permits an arbitrary positioning of the crystal and scanning about appropriate axes for data collection. An FTS temperature controller supplies a stream of dry nitrogen gas coaxially along the sample capillary. The detector consists of a multiwire proportional counter similar to the San Diego design,²⁵ filled with 90% xenon and 10% CO₂, and appropriate electronic circuitry. This chamber has an active area of 256 mm horizontally (128 pixels of 2 mm) by 281 mm vertically (256 pixels of 1.098 mm). The detector is mounted on a cross bar from a sturdy arm along which its position can be adjusted. For these experiments the detector was placed centrally in the horizontal direction, and it was raised vertically about the 2θ axis of the goniometer so that the lower edge was approximately at the incident beam position.

X-Ray Absorption Spectra

Anomalous scattering factors are related in a known way to atomic absorption coefficients. Thus, x-ray absorption spectra permit an experimental determination of scattering factors very near the absorption edge. Absorption spectra were recorded by fluores-

cence measurements²³ from the iron atoms in the same crystal as that used for diffraction measurements. A NaI (TI) scintillation counter was placed normal to the x-ray beam in the horizontal plane at approximately 1 cm from the crystal. The crystal was aligned with a principal axis parallel with the electric vector of the polarized (90%) synchrotron radiation. Spectra were recorded as the ratio of fluorescence intensity, $I_f(E)$, to incident beam intensity, $I_0(E)$, monitored by an ion chamber. Spectra were taken in four segments: pre-edge, 6.71 to 7.09 keV in monochromator increments of 0.025° (ca. 10 eV); edge, 7.09 to 7.19 keV in steps of 0.002° (ca. 1 eV); exafs, 7.19 to 7.51 keV in steps of 0.010°; and tail, 7.51 to 8.32 keV in steps of 0.040°.

Diffraction Measurements

Experiments were conducted on two occasions. In the first instance (April 1984), a single crystal was used to obtain complete sets of data to 5.5Å spacings at each of four wavelengths. The area detector was positioned at 560 mm and raised to 12° for this experiment. On the second occasion (June 1985), 3.0Å data were measured from three crystals with the detector at 480 mm and 18° elevation. Each crystal was carefully oriented so that Bijvoet pairs related by mirror symmetry fell on the detector essentially simultaneously. Data were then recorded in a manner similar to that used by Xuong et al.²⁶ The crystal was systematically moved about the φ or ω axis as in rotation photography with electronic pictures taken at each 0.01° step. Images from 15 successive steps were stored as frames of data on magnetic discs. Each frame was recorded for the four wavelengths in succession before proceeding to the next frame. Exposure times were adjusted by monitor dosages established to provide an average counting statistics error on the order of 1% in $|F|$.

Wavelengths for the diffraction measurements were chosen to optimize the phasing power. Since Bijvoet differences are proportional to $2f''$ and dispersive differences are proportional to $|f'(\lambda_i) - f'(\lambda_j)|$,²⁷ we sought to use the extreme values of f'' and of f' together with remote reference points having a low magnitude of f' . The edge extrema were located from the absorption spectra to be at 1.738Å for the absorption peak (maximal f'') and 1.740Å for the inflection point (minimal f'). Two remote points were chosen for these tests: one at a pre-edge point (1.800Å) and a second at a remote high-energy point (1.500Å for the 5.5Å test, 1.650Å for the 3Å test). The former has the advantage of closer proximity to the edge points whereas at 1.500Å the phasing signals are greater. The 1.650Å point was dictated by the limited range available with an asymmetrically cut monochromator crystal used in the 3Å experiment.

A complete sampling of reciprocal space for the 5.5Å data set was accomplished by a rotation of 104° about φ at $\chi = 0^\circ$ followed by rotations of 29° about

ω at $\chi = -90^\circ$ with φ set at two orthogonal settings giving Bijvoet symmetry. For the later 3Å experiments, with the detector set at 480 mm, spacings from the 97Å cell edge were still resolved. Then, in the worst case (1.800Å), data were measured to Bragg spacings of 6.9Å horizontally and up to 3.0Å vertically. Usable data were acquired from three crystals in four sets. The first crystal was mounted along *c*, and a full quadrant (set 1) of multiwavelength data was measured by rotation of 100° about φ at $\chi = 0$. The second crystal was mounted along *a*, and two sets of data were measured. Set 2 was a full quadrant by rotation (100°) about φ at $\chi = 0$, and set 3 was a fill-in at $\chi = -90^\circ$ with φ oriented such that the ω rotation (33°, from -15 to 18°) was about the *c*-axis. The third crystal was mounted along *b*, and set 4 was collected at $\chi = 0$ by a rotation of 64° from φ (*a* \perp x-rays) + 26° to φ (*c* \perp x-rays).

Crystals were maintained in a dry nitrogen stream at approximately 10°C. Temperature control in these early experiments was not ideal, and some crystals were lost because of distillation caused by temperature gradients.

Data Processing

Data analysis for this study divides into several distinct aspects. In many instances, the theoretical treatment and software used in this processing were developed during the course of this application. Categories of analysis include the processing of x-ray absorption spectra into anomalous scattering factors, reduction of detector images to integrated diffraction intensities, scaling and merging of intensity data into unique data sets, and phase determination and structure analysis.

Fluorescence spectra, $R = I_f(E)/I_0(E)$, were reduced to atomic absorption spectra by a program XASFIT to make background and scaling corrections such that experimental values for atomic absorption coefficients

$$\mu_a = sR - (a + b\Delta + c\Delta^2) \quad (1)$$

were fitted in a least-squares sense to theoretical values outside the edge region (6.68 to 7.40 keV). Here R is the fluorescence ratio, $\Delta = E - E_0$ where $E_0 = 7111.2$ eV; s is the scale factor; and a , b , and c parameterize the background. Theoretical values were computed using SPECTRUM, a variant of the FPRIME program of Cromer²⁸ adapted to provide a complete spectral profile of absorption and scattering factors for any atomic species. Another program, KRAMIG, was written to perform the transformation from f'' to f' values subject to the condition that scattering factors outside the chemically sensitive edge region should match the isolated atom values computed by SPECTRUM.

Diffraction data were reduced to integrated intensities from the stored detector images using a pro-

gram developed at SSRL by E.A.M. based on initial software from San Diego.²⁶ This program refines alignment parameters, corrects for background, performs integrations, and applies Lorentz and polarization factor corrections. Integrations were carried out from rotation profiles of the counts summed up within prediction boxes on successive frames, and background values were estimated from a weighted average of counts within these boxes of frames preceding and following the reflection. Reflections typically extended over two or three frames. Absorption corrections for the 5.5Å data set were based on a ψ -scan made with CuK $_{\alpha}$ radiation on a conventional diffractometer after return from the synchrotron. An adjustment for the differences in linear absorption coefficient at different wavelengths was made according to a formula for transmission factors

$$T(\psi, \lambda_i) = \exp [\mu(\lambda_i) / \mu(\lambda_0) \ln T(\psi, \lambda_0)] \quad (2)$$

that derives from an assumption of a flat plate morphology. A small correction for radiation damage was based on a series of frames repeated at the beginning and end of the run for the 5.5Å data set. No absorption or radiation damage corrections were made for the 3.0Å set.

Another important aspect of the data processing concerned local scaling to minimize residual systematic errors. The procedures used here were based on those developed in the course of the crambin work,¹ and these have been described elsewhere.^{29,30} First, the data from mirror-related Bijvoet mates were brought to a common scale by means of parameterized local scaling factors determined and applied using ANOSCL and SCALE2. Then a new program, WVLSCL, based on the same procedures as used in ANOSCL, was written to bring the data sets at different wavelengths to a common scale. An estimate of expected average intensity from Wilson's statistics was used to place the data onto an absolute scale. In this case, the differences in scale due to variation in iron-scattering factors with wavelength were negligible (10^{-3}). All data, Bijvoet pairs and different wavelengths, that pertained to the phase evaluation for a particular reflection were merged together by WVLSCL for input to the phase-determining program MADLSQ. Redundant information from different evaluations of symmetry equivalents was subsequently combined using another program, OMERGE.

The final step in the processing related directly to the structure determination. MADLSQ was written to implement the least-squares procedure that is described below for the evaluation of structure factors from multiwavelength anomalous diffraction data. Patterson functions were computed using a Fourier synthesis program, PATTERN, with a provision for eliminating outlier data (see below). Trial refinements of the iron coordinates against the fitted $|\mathbf{F}_A|$

coefficients were carried out using ANOLSQ, but phase calculations here were done with ASCALC using the iron coordinates and thermal parameter determined in the 2Å stereochemically restrained refinement.²¹ A trivial program combined the outputs from MADLSQ, after merging by OMERGE, with the results from ASCALC to produce phases. These data were then used in phase comparisons and to produce coefficients for synthesis into an electron-density distribution via a space-group-specific program, FORIER.

The computations at SSRL were done on the PDP-11/34 system then controlling the area detector facility and on the VAX 11/780 there. Initial processing of the 5.5Å data set was performed at the Naval Research Laboratory on the TI-ASC computer system. Subsequent analysis of the 5.5Å set and all work on the 3.0Å set were done on the VAX 11/780 in the Cancer Center Computing Facility at Columbia.

RESULTS

Anomalous Scattering Factors

A knowledge of the anomalous scattering factors is essential for the derivation of phase information from diffraction measurements at multiple wavelengths. The anomalous scattering is a complex increment, f^{Δ} , to the normal scattering factor f° which is pure real and wavelength invariant. Thus the total atomic scattering factor is given as

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

Calculated scattering factors based on atomic wave functions³¹ are found to be quite accurate at x-ray energies somewhat away from the absorption edge. However, the characteristics very near the edge, including the precise energy of the edge, are dependent on the molecular environment, and this, in principle, is not known in advance of the structure analysis. Thus, experimental measures of the scattering factors are needed to choose the optimal wavelengths and to analyze the diffraction data.

Fortunately, the needed scattering factors are quite readily accessible from x-ray absorption data. After appropriate scaling and background corrections, the x-ray fluorescence measured from iron atoms in the crystal yields the spectrum, as a function of energy E , of atomic absorption coefficients, $\mu_a(E)$. These data are related to scattering factors as described by James.³² The imaginary part of the anomalous scattering, f'' , is directly related to the atomic absorption coefficient

$$f''(E) = \frac{mc}{4\pi e^2 h^2} E \mu_a(E) \quad (4)$$

and the real part, f' , is in turn related to the imaginary part by the Kramers-Kronig transformation

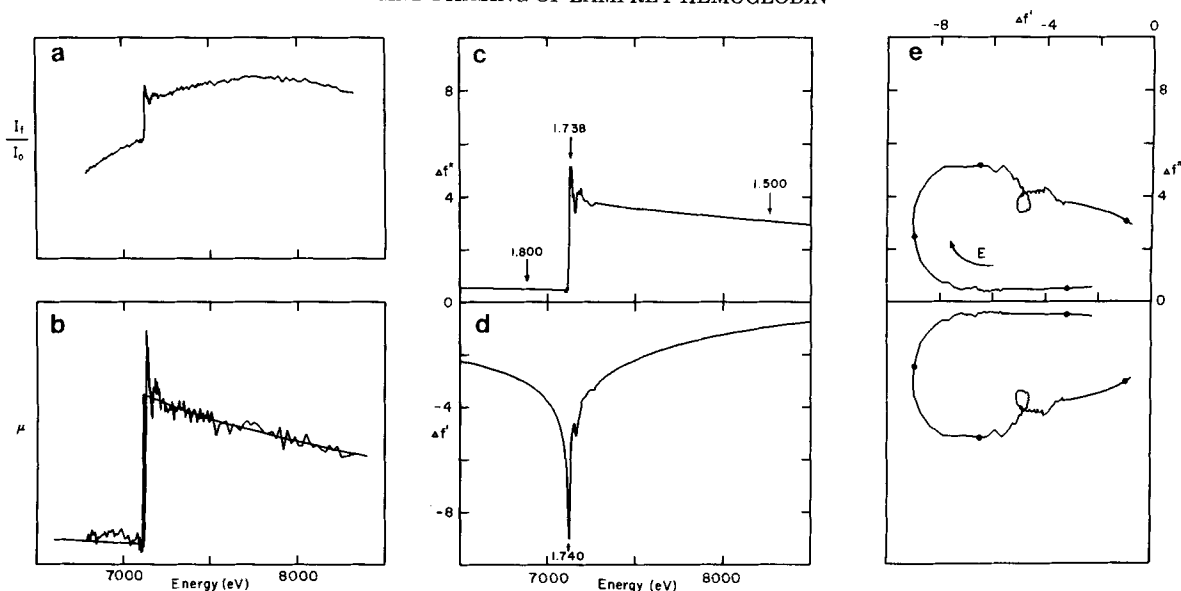


Fig. 1. Experimentally determined anomalous scattering factors. **a**: Normalized x-ray fluorescence spectrum from a crystal oriented with the polarization vector approximately parallel with the **b** axis. **b**: Spectrum of atomic absorption coefficients deduced from **a** by fitting to theoretical values (smooth curve) after background and scaling corrections. **c,d**: Imaginary and real components, respectively, of Fe scattering factors in type D2 lamprey

hemoglobin with **E** parallel **b**. Experimental data similar to those in **a** and **b** are the basis for near edge values, whereas calculated values are used elsewhere. Wavelengths of diffraction measurements in the initial 5.5Å experiment are indicated by arrows. **e**: Herzenberg-Lau representation of scattering factors data shown in **c** and **d**. Wavelength points of diffraction measurements are indicated by solid circles.

$$f'(E) = \frac{2}{\pi} \int_0^{\infty} \frac{E' f''(E')}{(E^2 - E'^2)} dE'. \quad (5)$$

The symbols m , c , e , and \hbar refer to fundamental physical constants with the usual meaning. With appropriate numerical procedures, the Cauchy principal value of this latter integral can be evaluated from limited spectral data and scaled to theoretical scattering factors at points remote from the edge.

The fluorescence spectrum and resulting atomic absorption data are shown in Figure 1a and b, respectively. Anomalous scattering factors processed from such data, in this case for the crystal oriented with the electric vector, **E**, parallel with the **b** axis, are shown in Figure 1c and d. In fact, these type D2 crystals are distinctly dichroic with respect to x-ray absorption. As shown in Figure 2, this dichroism reflects the unique, albeit appreciably oblique, orientation of the heme plane with respect to the crystal axes. Although the absorption edge positions are essentially the same, the scattering factor values are somewhat affected. This is a potentially complicating factor in our analysis, but the polarization effects are relatively small here and orientationally averaged scattering factors have been used. The position of wavelengths chosen for diffraction measurements are shown in the scattering factor spectra of Figure 1c and d, and these values are also shown in Figure 1e in the representation of Herzenberg and Lau⁷ to illustrate diffraction strength.

Diffraction Data

As described previously,²⁷ the strength of anomalous scattering signals can be estimated from scattering factors and the atomic composition of the crystal. The pertinent observations are the Bijvoet differences

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

and the dispersive differences

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

where $^{\lambda}F(h)$ denotes the crystallographic structure factor **F** at wavelength λ for a reflection with indices of **h**, and \bar{F} stands for the average of $|F(h)|$ and $|F(-h)|$. Averages of these anomalous diffraction differences are then normalized and expressed as Bijvoet difference ratios at a given wavelength, $\langle |\Delta F_{\pm h}| \rangle / \langle |F| \rangle$, and as the dispersive difference ratio between pairs of wavelengths, $\langle |\Delta F_{\Delta\lambda}| \rangle / \langle |F| \rangle$.

Anomalous diffraction estimates for lamprey hemoglobin are shown in Table I for the wavelengths used in these experiments and with the scattering factors measured in the x-ray absorption experiments. As can be seen, the expected anomalous signals are on the order of 3% in the most favorable cases. The anomalous diffraction ratios actually found in the data measured for the 5.5Å set are shown in Table II. There is clear evidence for variations in signal strength with wavelength as expected from Table I,

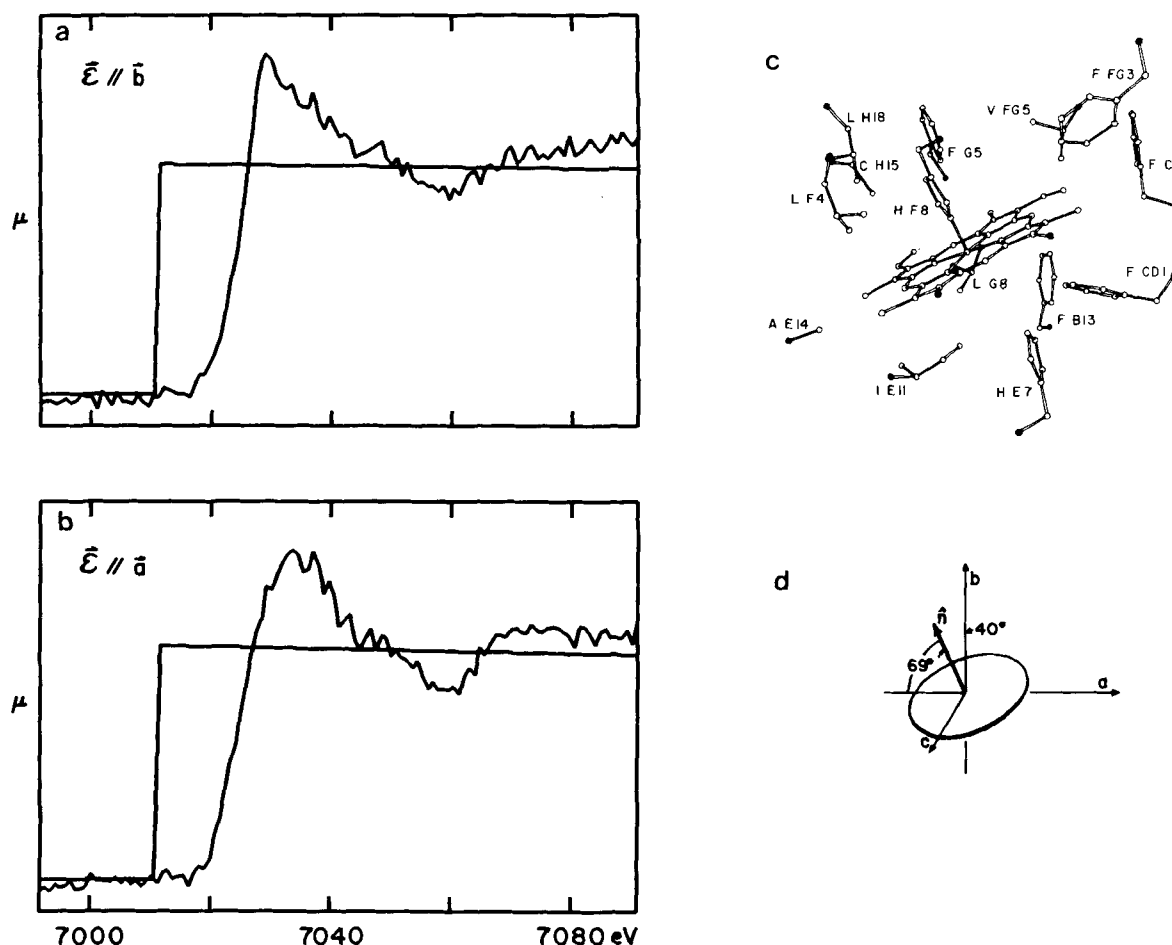


Fig. 2. Anisotropy in iron x-ray absorption spectra. **a,b**: Near edge atomic absorption coefficients from spectra with **E** parallel with **b** and **a** respectively. Theoretical absorption coefficients are shown in the smooth lines. **c**: Heme environment in lamprey

hemoglobin as viewed down the **c** axis. **d**: Orientation of the heme normal relative to the type D2 crystal axes. Thus, the **E // b** spectrum is dominated by heme normal absorption, whereas the **E // a** spectrum is dominated by in-plane absorption.

TABLE I. Estimates of Anomalous Diffraction Contributions

λ_i/λ_j	Expected diffraction ratios (%)					Scattering factors (e)	
	1.8000	1.7402	1.7380	1.6500	1.5000	f'	f''
1.8000	0.3	1.8	1.0	0.3	0.7	-3.17	0.50
1.7402		1.5	0.8	2.1	2.4	-8.97	2.48
1.7380			3.2	1.3	1.7	-6.49	5.19
1.6500				2.2	0.4	-2.18	3.56
1.5000					1.9	-0.95	3.07

Diagonal elements are computed as $q(2f'')$ and off-diagonal elements are given by $q[f'(\lambda_i) - f'(\lambda_j)]$ where $q = (2N_p Z_{\text{eff}})^{-1/2}$ according to Hendrickson.²⁷ The effective normal atomic scattering, Z_{eff} , was set to 6.7e, and the number of nonhydrogen atoms, N_p in lamprey hemoglobin, is 1,194. Scattering factors at the edge position of 1.7402Å and 1.7380Å were determined experimentally from data presented in Figure 1, and other values were computed by the Cromer program.²⁸ These estimates pertain to zero scattering angle and would be increased by factors of approximately 1.14 and 1.41 at 5.0 and 3.0Å Bragg spacings, respectively.

but these signals are embedded in a noise level also on the order of 3% as indicated by the Bijvoet difference ratios for the centric reflections. Centric Bijvoet differences should theoretically be zero, and therefore these latter statistics are equivalent to R_{sym} values based on $|F|$.

The error level in diffraction measurements was not limited by counting statistics. Counting time was adjusted such that an average error contributed by counting statistics would be on the order of 1% in $|F|$ (approximately 2,500 counts). Some of the additional error seen in the 5.5Å data arose from nonuniformities in the area detector as used at that time. Parameterized local scaling, which is smoothly varying, removed some of the systematic error; however, it soon became evident that there were substantial additional systematic errors that varied discontinuously across the detector surface. Special scaling adjustments were made to compensate for these errors, but it is clear from the results that much error remains.

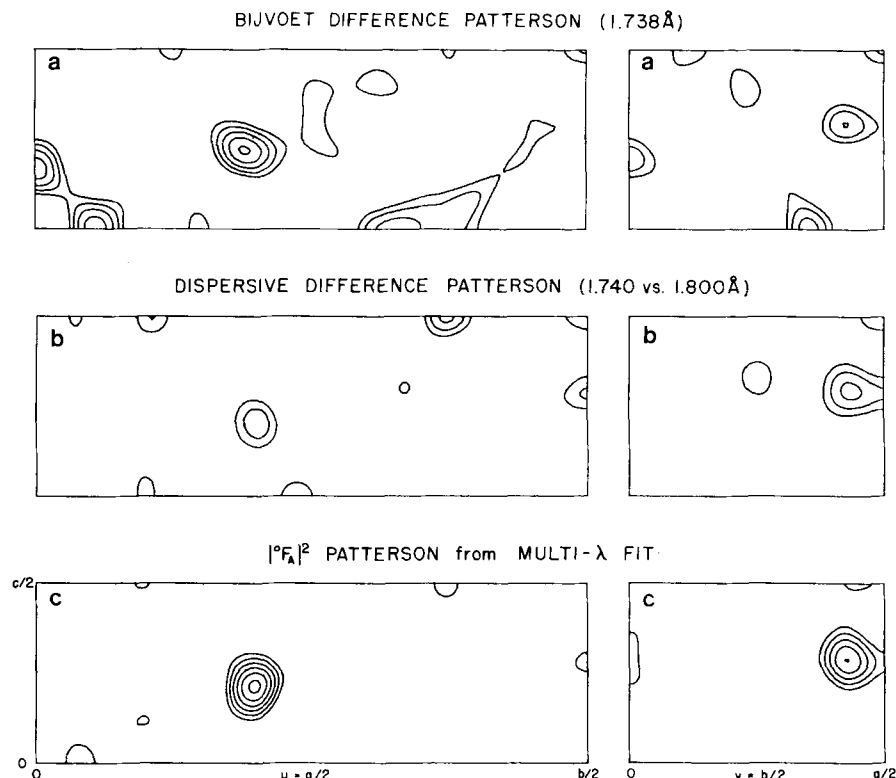


Fig. 3. Harker sections of Patterson syntheses from the 5.5Å data. The panels at left show $u = 1/2$ sections and those at right show $v = 1/2$ sections. Contours are drawn at equal intervals starting at a minimal level of twice the interval. **a:** Bijvoet-difference Patterson map from the 1.738Å data after local scaling. This synthesis included 131 coefficients out of 329 general (noncentric) reflections meeting the criteria of $|F| > 10$ (F), $\Delta F > 3 \sigma(\Delta F)$,

and $\Delta F < 3$ r.m.s. (ΔF). **b:** Dispersive difference Patterson map between data at 1.740 and 1.800Å after all local scaling operations. This synthesis included the 164 reflections of a possible 528 that survived the quality cuts as used above. **c:** The MAD Patterson map based on $|\phi_F|$ coefficients derived from all measurements. This synthesis included the 502 of 528 reflections having $|\phi_F| < 175$.

TABLE II. Observed Diffraction Ratios for the 5.5Å Data Set*

λ_i/λ_j	1.8000	1.7402	1.7380	1.5000
1.8000	3.2% (2.9)	4.0%	3.3%	4.9%
1.7402		3.6% (3.0)	1.8%	6.0%
1.7380			5.1% (2.8)	5.3%
1.5000				4.3 (3.5)

*Diagonal: r.m.s. ($\Delta F_{\pm h}$)/r.m.s. ($|F|$) (centric in parentheses); Off-diagonal: r.m.s. ($\Delta F_{\Delta h}$)/r.m.s. ($|F|$).

Data Analysis

With reduced and scaled structure factor magnitudes in hand, the first task in the data analysis is to locate the positions of the anomalous scattering centers. It can be shown that Patterson syntheses with coefficients as the square of Bijvoet or dispersive differences (eqs. 6 and 7) are, for relatively weak anomalous signals, good approximations of the Patterson

map of the anomalous scattering centers.³³ Harker sections of such syntheses are shown in Figure 3a and b. The Bijvoet-difference Patterson from the $\lambda = 1.738$ Å data was correctly interpreted by one of us who was unaware of the known iron position in these crystals. However, because of the previously mentioned systematic errors, this is a very noisy map. Indeed, there are noise features that are stronger than the true peaks. The best dispersive difference Patterson maps are somewhat better than the Bijvoet difference maps, but still they are of rather poor quality.

As is evident from the tables, for four wavelengths there are ten possible sets of difference coefficients from which Patterson maps could be synthesized. Obviously, a synthesis that somehow incorporates all of the available information in a properly weighted manner would be expected to show less noise than that in the individual components. A naive approach would be to average the various coefficients. Instead we used a least-squares approach that simultaneously solves for phase differences. The resulting coefficients are optimal and involve no theoretical approximations. As can be seen in Figure 3c, an extremely clean map is obtained.

Our analysis is based on an algebraic approach introduced by Karle.³⁴ This analysis refers everything to the normal scattering terms of the structure factors. The complete structure factor at a particular wavelength λ , including anomalous scattering contributions, is designated $|F(h)|$, whereas the wavelength invariant contribution due solely to the normal scattering of all atoms is denoted by $|F_T|$. Then in the special case of a single kind of anomalous scatterer that contributes a scattering component of $|F_A|$,

$$|F(\pm h)|^2 = |F_T|^2 + a(\lambda) |F_A|^2 + b(\lambda) |F_T| |F_A| \cos(\varphi_T - \varphi_A) \pm c(\lambda) |F_T| |F_A| \sin(\varphi_T - \varphi_A), \quad (8)$$

where, using the notation of eq. 3,

$$a(\lambda) = (f'' / f')^2, \quad (9a)$$

$$b(\lambda) = 2(f'' / f'), \quad (9b)$$

and

$$c(\lambda) = 2(f''' / f'). \quad (9c)$$

In the present study there are eight observational equations for each h ,

$$y_0 = |F(\pm h)| \quad (10)$$

and these are represented by three independent parameters in eq. 8. We evaluate these desired parameters by a fitting procedure that first solves the linear least-square problem

$$\phi = \sum_{\lambda} \sum_h (y_0 - y_c)^2, \quad (11)$$

where

$$y_c = p_1 + a(\lambda)p_2 + b(\lambda)p_3 \pm c(\lambda)p_4, \quad (12)$$

and then uses this set of p values as the starting point for the nonlinear problem

$$\Phi = \phi + \gamma g \quad (13)$$

which invokes, via the undetermined multiplier γ , a Lagrange constraint

$$g = p_1 p_2 - p_3^2 - p_4^2 = 0 \quad (14)$$

that imposes the identity $\sin^2 \theta + \cos^2 \theta = 1$ and, implicitly, the positivity of structure factor magnitudes. The three fundamental parameters are then simply

$$\begin{aligned} |F_T| &= p_1^{1/2}, \\ |F_A| &= p_2^{1/2}, \text{ and} \\ \Delta\varphi &= (\varphi_T - \varphi_A) = \tan^{-1} (p_4/p_3). \end{aligned}$$

Standard deviations in these parameters are produced from the propagation of errors given by the least squares solution.

This least-squares analysis of MAD data leads to a rather direct solution of the phase problem. First, the $|F_A|$ values are used to solve for the structure of anomalous scattering centers. This can readily be done from Patterson syntheses for simple structures (see Fig. 3c) or by direct methods for more complicated cases. These $|F_A|$ values also serve as the basis for refinement of the atomic parameters. Then by Fourier inversion this anomalous scatterer structure gives the calculated values of φ_A and hence $\varphi_T = \Delta\varphi + \varphi_A$.

Finally, these phases for the normal components of the structure factor together with the magnitudes $|F_T|$, also determined in the least-squares fitting, provide the coefficients for the desired macromolecule. Only experimental errors of measurement prevent this from being an exact solution.

Phasing Results

Application of the MAD least-squares fitting procedure described above to the lamprey hemoglobin problem proceeded very smoothly. Convergence was rapid in all cases and the process was found to be exceptionally robust. An example for a particular reflection is given in Table III to illustrate the magnitudes of the various parameters that are involved. This is a particularly favorable example that leads to a standard deviation in phase of 9° . In addition to the fundamental parameters, $|F_T|$, $|F_A|$, and $(\varphi_T - \varphi_A)$ and their standard deviations, the r.m.s. residual, $Q = (\phi/n)^{1/2}$, produced by MADLSQ for each reflection is also somewhat useful in assessing accuracy.

The phasing results from these experiments can be checked for accuracy in two ways. First, the MAD phases can be compared with those determined by MIR²⁰ or with those calculated from the least-squares model (LSQ) obtained by restrained refinement.²¹ Second, MAD phases of reflections replicated with the different data sets can be compared. In the case of the 5.5Å data set, MAD phases were determined for 512 of the 528 possible reflections, and the average phase discrepancy between the MIR and MAD sets was 47.0° . The MIR vs. MAD phase discrepancy for the 5.5Å subset (505 reflections) of the merged 3Å data was 42.9° , and LSQ vs. MAD gave 43.3° as compared with an LSQ vs. MIR agreement of 40.8° . Full comparisons of MAD phases in the merged data from the 3Å experiment with the MIR and LSQ results are

summarized in Figure 4. A more detailed elaboration of these comparisons is given in Table IV, where replicated data from the four constituent data sets are also compared. It is clear from the results in Table IV, as well as from the observed diffraction ratios for this set, that data set 1 is markedly inferior to data sets 2, 3, and 4. This difference is associated with a change from a monochromator crystal having an 13° asymmetric cut to one of 9° cut. In addition, temperature control was a particular problem with this crystal. When data from set 1 are excluded in the merging, the MAD phases and MIR phases are in equally close agreement with the LSQ results. We presume that inaccuracies in set 1 arose from inadequate bathing of the first crystal in the x-ray beam or from incipient

thermal deterioration. Another important result shown in Table IV is the validation of various measures of accuracy in the MAD least-squares analysis. Both the internal consistency of $|\varphi_A|$ and $^\circ\varphi_T$ values and the agreement with LSQ results improve when data that produced either 1) excessively large $|\varphi_A|$ values (outliers), 2) low ratios of $|\varphi_A|$ to $\sigma(|\varphi_A|)$, 3) high values of $\sigma(\Delta\varphi)$, or 4) poor fitting residuals, Q , are eliminated.

Data from the first experiment have been used to compute a Fourier synthesis at 5.5\AA resolution and to compare this result with that obtained with MIR phasing. This electron density synthesis was computed with coefficients of $m|\varphi_T| \exp(i\varphi_T)$ where $^\circ\varphi_T = \Delta\varphi_{\text{MAD}} + ^\circ\varphi_A(\text{calc})$ and $m = \cos(\sigma(\Delta\varphi))$. Identical sections from the Fourier map at 5.5\AA resolution are shown in Figure 5. This is a superposition of special sections through a stack of planes that contains the A and B helices and cuts the G and H helices in a slightly oblique cross section. Obviously the MIR map is superior, but it is very encouraging that the MAD map is as clean as it is when the substantial systematic errors in this particular data set are considered.

DISCUSSION

The possibility recognized long ago² for phase determination based on measurements of anomalous dif-

TABLE III. MAD Phasing Results for Reflection (10, -9, 4)

λ	Observed		Calculated		$a(\lambda)$	$b(\lambda)$	$c(\lambda)$
	$^\circ F(h)$	$^\circ F(\bar{h})$	$^\circ F(h)$	$^\circ F(\bar{h})$			
1.6500	122	99	124	96	0.04	-0.21	0.35
1.7382	133	91	128	86	0.17	-0.64	0.51
1.7400	115	93	117	97	0.21	-0.88	0.24
1.8000	111	103	111	107	0.03	-0.31	0.05

$$|\varphi_T| = 112 \pm 4e; |\varphi_A| = 84 \pm 10e; (^\circ\varphi_T - ^\circ\varphi_A) = 72 \pm 9^\circ.$$

TABLE IV. Measures of Accuracy in MAD-Extracted Parameters

Parameter	Data Set	All	$ \varphi_A < 200$	$ \varphi_T < 200$ $^\circ F_T > 10\sigma F$	$ \varphi_A < 200$ $^\circ F_A > 2\sigma F$	$ \varphi_A < 200$ $\sigma_{\Delta\varphi} < 50^\circ$	$ \varphi_A < 200$ $Q < 20$	All conditions
Number of reflections (fraction retained)	1	1,312	1,026 (78%)	895 (68%)	504 (38%)	820 (63%)	827 (63%)	442 (34%)
	2-4	2,779	2,765 (99%)	2,682 (97%)	1,618 (58%)	2,347 (85%)	2,695 (97%)	1,584 (57%)
$R(\varphi_T \text{ vs. } F_P(\text{obs}))$	1		0.127	0.126	0.136	0.132	0.139	0.131
	2-4		0.080	0.077	0.086	0.082	0.082	0.084
$R(\varphi_A \text{ vs. } F_A(\text{calc}))$	1		0.512	0.504	0.433	0.463	0.482	0.408
	2-4		0.426	0.421	0.362	0.395	0.418	0.355
$\Delta\varphi(^\circ\varphi_T \text{ vs. } \varphi_{\text{LSQ}})$	1		68.2°	66.5°	58.5°	64.4°	64.3°	56.7°
	2-4		50.6°	49.7°	44.1°	47.5°	49.8°	43.5°
Number of common reflections	1 vs. 2-4		956	826	306	640	753	264
	2 vs. 3 vs. 4		732	699	295	523	695	282
$R(\varphi_T(i) \text{ vs. } \varphi_T(j))$	1 vs. 2-4		0.093	0.091	0.109	0.099	0.101	0.109
	2 vs. 3 vs. 4		0.084	0.079	0.091	0.090	0.083	0.084
$R(\varphi_A(i) \text{ vs. } \varphi_A(j))$	1 vs. 2-4		0.534	0.527	0.399	0.438	0.532	0.373
	2 vs. 3 vs. 4		0.446	0.436	0.329	0.370	0.431	0.306
$\Delta\varphi(^\circ\varphi_T(i) \text{ vs. } ^\circ\varphi_T(j))$	1 vs. 2-4		72.4°	70.3°	52.0°	65.7°	68.0°	48.8°
	2 vs. 3 vs. 4		49.0°	48.1°	31.8°	42.3°	46.7°	31.7°

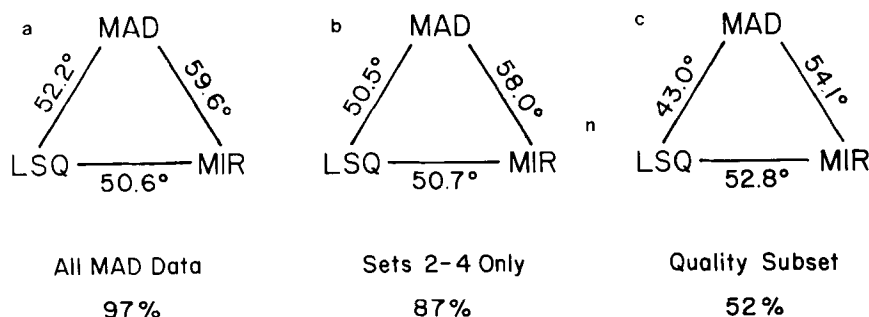


Fig. 4. Phasing comparisons from the 3Å data sets. MIR, LSQ, and MAD refer respectively to phase sets from multiple isomorphous replacement,²⁰ calculations from the refined model,²¹ and from the multiwavelength anomalous diffraction results reported here. Phase discrepancy values are averages of the lesser of $|\Delta\varphi|$ or $360^\circ - |\Delta\varphi|$. MIR phases were available for all 2,987 possible reflections with d spacing greater than 3Å, but reflections with $d > 10\text{\AA}$ were omitted from the LSQ set. Various subsets of the MAD phases were used. Each leg of a

triangle includes only those reflections in common to all three data sets. **a:** All MAD data. MAD phases were determined for 2,898 reflections, of which 2,779 are common with the LSQ set and thus included in the averages. **b:** MAD data exclusively from sets 2-4, excluding the defective results from set 1. This subset includes 2,586 reflections, of which 2,479 are included in the averages. **c:** MAD data from the 2-4 subset that meets all quality criteria applied in Table IV. This subset includes 1,552 reflections, of which 1,507 were used in these comparisons.

fraction at multiple wavelengths can clearly be achieved in practice. Even in this case where the expected peak signals are only at the 2-3% level, we have obtained MAD phases as accurate overall as the MIR phases from which the lamprey hemoglobin structure was previously interpreted.¹⁹ Metalloproteins for which expected anomalous diffraction ratios are as strong as those found here should be amenable to analysis by the MAD phasing procedure. Moreover, even stronger signals can be expected from the L absorption edges of heavy atoms, such as platinum, gold, mercury, and lanthanides, in derivatized proteins. Since the analytical procedures developed here are definitive in character, only measurement errors should limit the accuracy of results. Thus, with stronger signals, one can expect to obtain MAD phases that are more accurate than MIR phases.

In the course of this study on lamprey hemoglobin we have introduced a number of experimental and analytical procedures to facilitate MAD phase determination. First, the choice of wavelengths was designed to optimize both the Bijvoet and the dispersive diffraction ratios.²⁷ This required that measurements be made right at the resonant absorption edge and also at remote points in accord with the analysis of Narayan and Ramaseshan.³⁵ The approach taken here contrasts with previously reported experiments in which the wavelengths used were either exclusively very near the edge¹¹ or else avoided the extreme edge positions.^{12,13} Since the position and strength of scattering at absorption edges depends on the chemical environment, we also developed the fluorescence measurement techniques and computational procedures needed for extracting scattering factors from x-ray absorption spectra. Another important development concerns the reduction of systematic errors. We chose an experimental design whereby Bijvoet mates would be measured such that equivalent absorption paths would be traversed and all data pertinent to a given phase would be measured close

together in time. The adaptation of parameterized local scaling¹ to this problem also played a critical role. As a result, significant signals were measured at the 2-3% level in spite of absolute accuracy in data at only the 8-9% level (Table IV). Finally, and perhaps most importantly, we have reduced Karle's algebraic approach³⁴ for analyzing MAD data to effective practice. This innovation provides, in a completely general way, a definitive evaluation of the fundamental diffraction parameters of the problem.

The experiments described here were conducted as the synchrotron instrumentation was still undergoing development, and in some instances (c.f. Fig. 3 and Table IV) the results were less than might optimally be expected. Very much cleaner single-wavelength anomalous Patterson maps have been produced from diffractometer data,^{17,18} and extremely clean maps have also been obtained with area-detector data from orthorhombic myoglobin crystals (unpublished results) and from lamprey hemoglobin crystals (M. Hatada and B. Graves, personal communication). It is reasonable to expect that if data of comparable accuracy were available at multiple wavelengths, MAD phases evaluated by the procedures described here would produce truly exceptional Fourier syntheses. Some of the later experiments (sets 2-4 of the 3Å data) approach this level, but even here the R_{sym} values are high (8%). Subsequent improvements to the SSRL instrumentation have produced data sets of much higher quality ($R_{\text{sym}} = 4.5\%$ based on $|F|^2$ for cucumber basic protein). Thus, despite the complications of a synchrotron setting, we see no fundamental barrier to performance levels at least comparable to those obtained with area-detector systems in conventional laboratories.

ACKNOWLEDGMENTS

This work was supported in part by grant No. GM34102 from the National Institutes of Health. The lamprey hemoglobin crystals used in this study were

LAMPREY HEMOGLOBIN at 5.5Å RESOLUTION

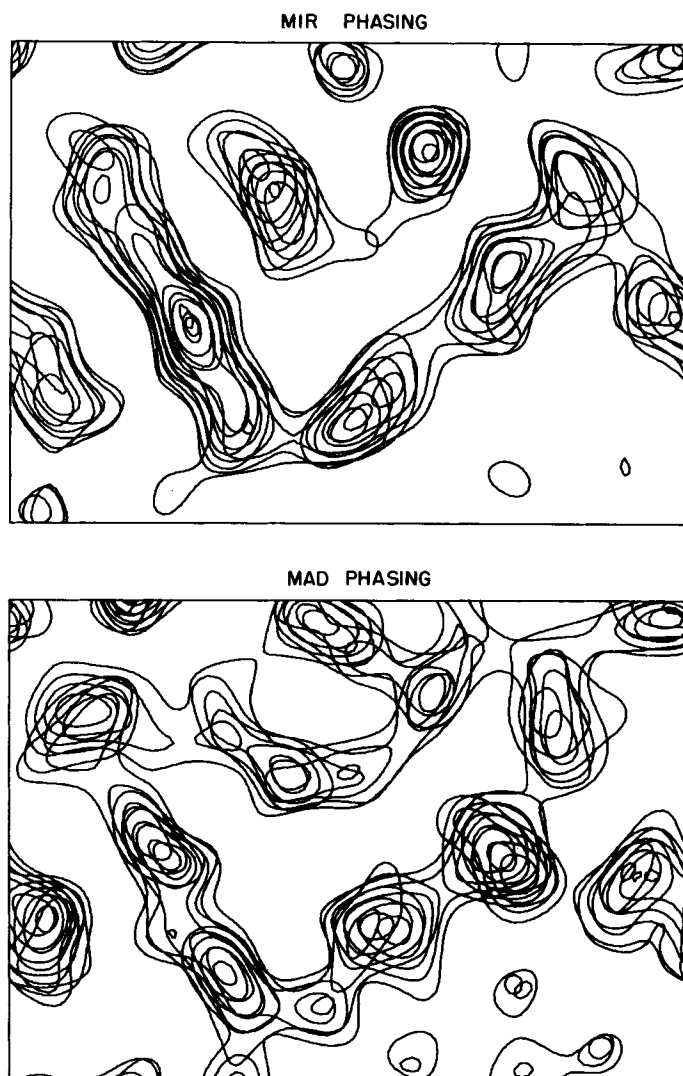


Fig. 5. Portion of electron-density maps for lamprey hemoglobin at 5.5Å resolution. The plane of section was chosen especially to include the A and B helices. Portions of the G and H helices

as well as the CD corner (at right) are also included. Both panels include the same parts of the map. Upper panel: MIR results; lower panel: MAD results.

grown at Johns Hopkins University, and we thank Warner Love for supplying them. The early calculations were carried out at the Naval Research Laboratory, and we thank Jerome Karle for his hospitality. The synchrotron facilities used at SSRL are supported by the Department of Energy, Office of Basic Energy Sciences, and by the National Institutes of Health, Division of Research Resources.

REFERENCES

1. Hendrickson, W.A., Teeter, M.M. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulphur. *Nature* 290:107-113, 1981.
2. Okaya, Y., Pepinsky, R. New formulation and solution of the phase problem in x-ray analysis of noncentric crystals containing anomalous scatterers. *Phys. Rev.* 103:1645-1647, 1956.
3. Mitchell, C.M. Phase determination by the two-wavelength method of Okaya and Pepinsky. *Acta Cryst.* 10:475-476, 1957.
4. Ramaseshan, S., Venkatesan, K., Mani, N.V. The use of anomalous scattering for the determination of crystal structures-KMnO₄. *Proc. Indian Acad. Sci.* 46A:95-111, 1957.
5. Raman, S. Theory of the anomalous dispersion method of determining the structure and absolute configuration of non-centrosymmetric crystals. *Proc. Indian Acad. Sci.* 50A:95-107, 1959.
6. Karle, J. Anomalous scatterers in x-ray diffraction and the use of several wavelengths. *Appl. Optics* 6:2132-2135, 1967.
7. Herzenberg, A., Lau, H.S.M. Anomalous scattering and the phase problem. *Acta Cryst.* 22:24-28, 1967.
8. Hoppe, W., Jakubowski, V. The determination of phases of erythrocrucorin using the two-wavelength method with iron as anomalous scatterer. In: "Anomalous Scattering." (S. Ramaseshan and S.C. Abrahams, eds.) Copenhagen: Munksgaard, 1975:437-461.
9. Phillips, J.C., Hodgson, K.O. The use of anomalous scattering effects to phase diffraction patterns from macromole-

- cules. *Acta Cryst.* A36:856–864, 1980.
10. Templeton, L.K., Templeton, D.H., Phizackerley, R.P., Hodgson, K.O. L_3 -edge anomalous scattering by gadolinium and samarium measured at high resolution with synchrotron radiation. *Acta Cryst.* A38:74–78, 1980.
 11. Kahn, R., Fourme, R., Bosshard, R., Chaimdi, M., Risler, J.L., Dideberg, O., Wery, J.P. Crystal structure study of *Opsanus tau* parvalbumin by multiwavelength anomalous diffraction. *FEBS Lett.* 179:133–137, 1985.
 12. Harada, S., Yasui, M., Murakawa, K., Kasai, N., Satow, Y. Crystal structure analysis of cytochrome c' by the multiwavelength anomalous diffraction method using synchrotron radiation. *J. Appl. Cryst.* 19:448–452, 1986.
 13. Korzun, Z.R. The tertiary structure of azurin from *Pseudomonas denitrificans* as determined by Cu resonant diffraction using synchrotron radiation. *J. Mol. Biol.* 196:413–419, 1987.
 14. Hendrickson, W.A. Measurement and use of anomalous x-ray scattering. *Acta Cryst.* A40:C-3, 1984.
 15. Smith, J.L., Pahler, A., Murthy, H.M.K., Hendrickson, W.A. Multiple-wavelength phase determination in protein crystallography. *Acta Cryst.* A43:C-10, 1987.
 16. Hendrickson, W.A., Troup, J.M., Swepston, P.N., Zdansky, G. Structure of d-selenolanthionine determined directly from multiwavelength anomalous diffraction of bremsstrahlung. Abstracts of the American Crystallographic Association, Series 2, 14:48, 1986.
 17. Argos, P., Mathews, F.S. Uses of a native anomalous scatterer in a protein structure determination. *Acta Cryst.* B29:1604–1611, 1973.
 18. Sheriff, S., Hendrickson, W.A. Location of iron and sulfur atoms in myohemerythrin from anomalous-scattering measurements. *Acta Cryst.* B43:209–212, 1987.
 19. Hendrickson, W.A., Love, W.E. Structure of lamprey haemoglobin. *Nature; New Biol.* 232:197–203, 1971.
 20. Hendrickson, W.A., Love, W.E., Karle, J. Crystal structure analysis of sea lamprey hemoglobin at 2 Å resolution. *J. Mol. Biol.* 74:331–361, 1973.
 21. Honzatko, R.B., Hendrickson, W.A., Love, W.E. Refinement of molecular model for lamprey hemoglobin from *Petromyzon marinus*. *J. Mol. Biol.* 184:147–164, 1985.
 22. Hendrickson, W.A., Love, W.E., Murray, G.C. Crystal forms of lamprey hemoglobin and crystalline transitions between ligand states. *J. Mol. Biol.* 33:829–842, 1968.
 23. Phizackerley, R.P., Cork, C.W., Merritt, E.A. An area detector data acquisition system for protein crystallography using multiple-energy anomalous dispersion techniques. *Nucl. Instr. Meth.* A246:579–595, 1986.
 24. Phizackerley, R.P., Cork, C.W., Hamlin, R.C., Nielsen, C.P., Vernon, W., Xuong, Ng.H., Perez-Mendez, V. Progress report on the development of an area detector data acquisition system for x-ray crystallography and other diffraction experiments. *Nucl. Instr. Meth.* 172:393–395, 1980.
 25. Cork, C., Hamlin, R., Vernon, W., Xuong, Ng.H., Perez-Mendez, V. A xenon-filled multiwire area detector for X-ray diffraction. *Acta Cryst.* A31:702, 1975.
 26. Xuong, Ng.H., Freer, S.T., Hamlin, R., Nielsen, C., Vernon, W. The electronic stationary picture method for high-speed measurement of reflection intensities from crystals with large unit cells. *Acta Cryst.* A34:289–296, 1978.
 27. Hendrickson, W.A. Analysis of protein structure from diffraction measurement at multiple wavelengths. *Trans. Am. Cryst. Assoc.* 21:11–21, 1985.
 28. Cromer, D.T. Calculation of anomalous scattering factors at arbitrary wavelengths. *J. Appl. Cryst.* 16:437, 1983.
 29. Smith, J.L., Hendrickson, W.A. Resolved-anomalous phase determination in macromolecular crystallography. In: "Computational Crystallography." (D. Sayre, ed.) Oxford: Oxford University Press, 1982: 209–222.
 30. Hendrickson, W.A. Anomalous dispersion in phase determination for macromolecules. In: "Crystallographic Computing 3." (G.M. Sheldrick, C. Kruger, and R. Goddard, eds.) Oxford: Oxford University Press, 1985: 277–285.
 31. Cromer, D.T., Liberman, D. Relativistic calculation of anomalous scattering factors for x-rays. *J. Chem. Phys.* 53:1891–1898, 1970.
 32. James, R.W. *The Optical Principles of the Diffraction of X-Rays*. London: Bell. 1948: 135–149.
 33. Rossmann, M.G. The position of anomalous scatterers in protein crystals. *Acta Cryst.* 14:383–388, 1961.
 34. Karle, J. Some developments in anomalous dispersion for the structural investigation of macromolecular systems in biology. *Int. J. Quant. Chem.* 7:357–367, 1980.
 35. Narayan, R., Ramaseshan, S. Optimum choice of wavelengths in the anomalous scattering techniques with synchrotron radiation. *Acta Cryst.* A37:636–641, 1981.