X-ray structure analysis of a metalloprotein with enhanced active-site resolution using *in situ* x-ray absorption near edge structure spectroscopy

Alessandro Arcovito[†], Maurizio Benfatto[‡], Michele Cianci[§], S. Samar Hasnain[§], Karin Nienhaus[¶], G. Ulrich Nienhaus^{¶†‡§}, Carmelinda Savino[†], Richard W. Strange^{§‡}, Beatrice Vallone[†], and Stefano Della Longa^{‡,§§}

[†]Department of Biochemical Sciences and Consiglio Nazionale delle Ricerche, Institute of Molecular Biology and Pathology, University of Rome La Sapienza, P.le A. Moro 5, 00185 Rome, Italy; [‡]Laboratori Nazionali di Frascati, Istituto Nazionale di Fisica Nucleare, P.O. Box 13, 00044 Frascati, Italy; [§]Molecular Biophysics Group, Daresbury Laboratory, Central Laboratory of the Research Councils, Warrington WA4 4AD, United Kingdom; [¶]Institute of Biophysics, University of Ulm, Albert-Einstein-Allee 11, 89081 Ulm, Germany; ^{††}Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, IL 61801; and ^{§§}Dipartimento di Medicina Sperimentale, Università dell'Aquila, Via Vetoio, 67100 L'Aquila, Italy

Edited by Douglas C. Rees, California Institute of Technology, Pasadena, CA, and approved February 17, 2007 (received for review September 23, 2006)

X-ray absorption spectroscopy is exquisitely sensitive to the coordination geometry of an absorbing atom and therefore allows bond distances and angles of the surrounding atomic cluster to be measured with atomic resolution. By contrast, the accuracy and resolution of metalloprotein active sites obtainable from x-ray crystallography are often insufficient to analyze the electronic properties of the metals that are essential for their biological functions. Here, we demonstrate that the combination of both methods on the same metalloprotein single crystal yields a structural model of the protein with exceptional active-site resolution. To this end, we have collected an x-ray diffraction data set to 1.4-Å resolution and Fe K-edge polarized x-ray absorption near edge structure (XANES) spectra on the same cyanomet sperm whale myoglobin crystal. The XANES spectra were quantitatively analyzed by using a method based on the multiple scattering approach, which yielded Fe-heme structural parameters with $\pm(0.02\text{--}0.07)\text{-Å}$ accuracy on the atomic distances and $\pm 7^\circ$ on the Fe-CN angle. These XANES-derived parameters were subsequently used as restraints in the crystal structure refinement. By combining XANES and x-ray diffraction, we have obtained an cyanomet sperm whale myoglobin structural model with a higher precision of the bond lengths and angles at the active site than would have been possible with crystallographic analysis alone.

myoglobin | crystallographic restraints | protein crystallography | synchrotron radiation

Metalloproteins perform a wide range of biological functions essential for sustaining life; they are particularly abundant in bioenergetic pathways, e.g., photosynthesis and respiration (1, 2). Often, the metal ion is located at the active site of an enzyme, where an intricate coordination geometry fine-tunes the electronic properties of the metal so as to enable it to catalyze specific reactions. X-ray diffraction (XRD) on protein single crystals or high-field NMR on protein solutions are the standard techniques by which entire molecular structures of proteins including their active-site metal clusters can be analyzed. By using quantum-chemical refinement, precise structural models of metal sites have been obtained from XRD data (3). X-ray absorption spectroscopy (XAS) is an alternative technique for measuring the structure of metal sites (4, 5). Electrons are liberated from the metal ion by the photoelectric effect, and the interference effects of the electron waves are analyzed to map out the local environment of the metal. XAS can reveal structural parameters with a very high precision, comparable to that of small-molecule crystallography. Such detailed structural data are required for elucidating the function of metal centers in a metalloprotein (6), as structural changes to the metal coordination during redox or substrate-binding reactions are generally $<0.1\text{ Å}$ and hence remain unnoticed in standard protein crys-

tallography unless the resolution is exceptionally high. As yet, x-ray structures of proteins at true atomic resolution ($<1.2\text{ Å}$) are still scarce (7, 8).

The combined measurement and analysis of XRD and XAS data of one and the same single crystal is expected to yield a molecular structure with excellent metal-site definition. To demonstrate the feasibility of this approach, we have chosen myoglobin (Mb), a small muscle protein that binds diatomic ligands such as dioxygen, carbon monoxide, and nitric oxide (9–11). Mb serves as a model system in protein science and, therefore, has been referred to as the “hydrogen atom of biology” (11). It was the first protein to have its 3D structure solved by x-ray crystallography >50 years ago (12), and many biophysical techniques have been used since then to study structure–dynamics–function relationships in this protein (13–15). XAS is one of these techniques, focusing either on the extended x-ray absorption fine structure (EXAFS) (5) or the x-ray absorption near edge structure (XANES) (16). Indeed, the first attempt to obtain and interpret angle resolved XANES spectra from a carbonmonoxy Mb (MbCO) crystal was made >20 years ago, with the aim of accurately defining the position and orientation of the CO with respect to the heme plane (17). However, the quantitative analysis of XANES spectra was rather restricted in earlier work because of the difficult theoretical treatment of the XANES features. Consequently, the low-energy part of the spectrum was often discussed only qualitatively, either by comparison with model compounds or as an aid for EXAFS studies (18). Only recently was a method introduced that allows reliable modeling and fitting of XANES spectra (19).

We have carried out combined XRD and XAS experiments using beamline 10 at the Daresbury Synchrotron Radiation Source (20), where we collected XRD data at 1.4-Å resolution and angle-resolved XANES spectra on the same cyanomet Mb (MbCN) single crystal. Here, we show that this combination of

Author contributions: A.A., S.S.H., G.U.N., R.W.S., and S.D.L. designed research; A.A., M.C., C.S., B.V., and S.D.L. performed research; M.B., S.S.H., K.N., G.U.N., R.W.S., and S.D.L. contributed new reagents/analytic tools; A.A., M.B., M.C., C.S., B.V., and S.D.L. analyzed data; and A.A., G.U.N., R.W.S., and S.D.L. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Abbreviations: XAS, x-ray absorption spectroscopy; XANES, x-ray absorption near edge structure; EXAFS, extended x-ray absorption fine structure; XRD, x-ray diffraction; Mb, myoglobin; MbCO, carbonmonoxy Mb; MbCN, cyanomet Mb.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2JHO).

^{††}To whom correspondence may be addressed. E-mail: uli@uiuc.edu, r.w.strange@dl.ac.uk, or dlunga@caspur.it.

This article contains supporting information online at www.pnas.org/cgi/content/full/0608411104/DC1.

© 2007 by The National Academy of Sciences of the USA

Fig. 1. Angle-resolved Fe K-edge XANES spectra of Mb single crystals. (A) An MbCN crystal was rotated around the crystallographic **b** axis. Black solid line indicates $\varepsilon \parallel \mathbf{c}^*$, $\alpha = 23^\circ$; black dotted line indicates $\varepsilon \parallel \mathbf{a}$, $\alpha = 90^\circ$; gray dotted line indicates $\alpha = 45^\circ$; and gray solid line indicates $\alpha = 56^\circ$. (B and C) Comparison between the polarized XANES spectra of MbCN (solid line) and MbCO (dotted line) for $\varepsilon \parallel \mathbf{c}^*$ (B) and $\varepsilon \parallel \mathbf{a}$ (C).

techniques can substantially enhance the precision with which the coordination geometries of metal sites can be determined, despite a conventional crystallographic resolution (21).

Experimental Results and Data Analysis

Fig. 1A shows angle-resolved XANES spectra of the MbCN crystal for various angles θ of crystal rotation around the vertically oriented **b** crystal axis. For crystal orientation and definition of angles, see *Materials and Methods*. The following discussion is based on the polarization angle α , i.e., the angle between the heme normal and the x-ray photon polarization vector ϵ , which is related to θ in a simple way (22). Details can be found in [supporting information \(SI\) Text](#) and [SI Fig. 5](#). The XANES spectra exhibit a pronounced dichroism that can be fully reproduced and modeled in the framework of multiple scattering theory. The spectrum polarized along the (reciprocal) **c*** crystal axis ($\alpha = 23^\circ$), which is nearly parallel to the Fe–CN bond vector, is reminiscent of the XANES spectrum of the ferricyanide complex $[\text{Fe}(\text{CN})_6]^{3-}$ (23). The overall shapes of the spectra are similar to those previously reported for MbCO (Fig. 1B and C) (17, 22, 24). The rising edges and the main peaks of the polarized

Fig. 2. Quantitative analysis of angle-resolved XANES spectra. Circles indicate experimental data, and solid lines indicate theoretical calculations. (A) Experimental $\varepsilon \parallel \mathbf{a}$ spectrum, assumed to be the $\varepsilon \parallel$ heme-plane spectrum ($\alpha = 90^\circ$), and best fit theoretical curve obtained by optimization of the planar parameter (heme core size). (B and C) Experimental spectra at intermediate angles ($\alpha = 45^\circ$ and 56°), superimposed with the calculated XANES spectra after the optimization procedure for the extreme angles. No fit is attempted, only the polarization term is changed here. (D) Experimental $\varepsilon \parallel$ heme-normal ($\alpha = 0^\circ$) spectrum, obtained as a weighted sum of the $\varepsilon \parallel \mathbf{c}^*$ and $\varepsilon \parallel \mathbf{a}$ spectra, and best fit of the XANES calculation obtained by MXAN via optimization of the axial parameters.

spectra of MbCN are blue-shifted with respect to MbCO by ≈ 1.5 eV, as expected for the chemical-shift difference between a ferrous and a ferric iron-heme site. Based on both this observation and the constant edge position throughout the experiment on MbCN, we conclude that photo-reduction was absent during data acquisition. Moreover, the overall qualitative similarities between the spectra of MbCN and MbCO suggest a similar ligand binding geometry.

Fig. 2 shows the results of the quantitative XANES analysis of

Table 1. Active-site geometrical parameters obtained by XANES and x-ray crystallography on a single P2₁ crystal of sperm whale MbCN

Parameter	XANES	XRD unrestrained	XRD restrained	XRD [†]	XANES MbCO [‡]
Resolution, Å		1.4	1.4	1.8	
R_{work}		18.0	18.0	18.1	
Fe–N _{pyrrol} , Å	2.01 (3)	2.04	2.04	2.03	2.00 (2)
Fe–N _{His93} , Å	2.07 (3)	2.08	2.08	2.02	2.06 (3)
Fe–CN, Å	1.87 (4)	2.22	1.92	2.02	1.83 (2)
Fe–CN, °	170 (7)	166	167	166	166 (4)
C–N, Å	1.10 (2)	1.10	1.11	1.06	1.07

Numbers in parentheses represent the errors of the last digit.

[†]Data taken from Protein Data Bank code 1EBC (26).

*Data on MbCO are included for comparison and refer to heme-bound CO instead of CN (23). As the C–O distance was kept fixed, no error is given.

Table 2. X-ray data collection and refinement statistics

Resolution, Å	50.00 - 1.40 (1.43–1.40)
R_{merge}	4.2 (17.6)
$\ \sigma(I) \ $	32.3 (5.5)
Completeness	93.9 (83.0)
Redundancy	6.1 (2.9)
Space group	P2 ₁
a, b, c , Å	34.4, 29.8, 63.6
α, β, γ , °	90.0, 104.9, 90.0
$R_{\text{work}}/R_{\text{free}}$	18.0 / 23.0
No. of atoms	
Protein	1,272
Ligands	1 CN [−] ion, 1 sulphate ion
Water	188
B factors, Å ²	
Protein main chain	24.2
CN ligand	20.8
Fe ion	20.1
Sulphate ion	18.3
Water	42.9
rmsd	
Bond lengths, Å	0.018
Bond angles, °	2.17
Cruickshank diffraction-component precision index (32)	
Coordinate error, Å	0.1046
Coordinate error, Å based on R_{free}	0.0871

Values in parentheses refer to the highest-resolution shell.

the polarized spectra using MXAN (19). We have calculated the absorption components linearly polarized along the heme normal, σ_{normal} ($\alpha = 0^\circ$), and along the heme plane, σ_{heme} ($\alpha = 90^\circ$). Fig. 2*A* displays the σ_{heme} experimental spectrum (open circles) together with the theoretical best fit (solid line) obtained by MXAN via optimization of a single parameter, the heme core size, leading to the determination of the Fe–N_{pyrrol} distance. Similarly, the σ_{normal} experimental spectrum (Fig. 2*D*) is fitted by MXAN via optimization of the axial parameters (Fe–N_{His-93}, Fe–CN and C–N distances, and Fe–C–N angle). All optimized parameters are compiled in Table 1. The similarity between the MbCN and MbCO ligand binding geometries is confirmed by a quantitative analysis of the MbCO XANES spectra, yielding essentially the same parameters (Table 1) (24). In Fig. 2*B* and *C*, we have plotted XANES experimental spectra of MbCN at intermediate polarization angles ($\alpha = 45^\circ$ and 56° , respectively) together with the theoretical XANES spectra, which were calculated from the spectra optimized at $\alpha = 0^\circ$ and 90° according

Fig. 3. Active-site structure of MbCN. (A) Electron density at the heme, shown at the 1.2- σ contour level. (B) Sketch of the active site, with XANES fit parameters included. The only parameter that is out of the XANES error range after XRD refinement is shown in red.

to Eq. 2. It is evident that all spectral features are reproduced. We note that a generalization of this procedure to other crystal systems is straightforward. Each of the two Mb molecules in the unit cell of the P2₁ crystal contributes equally to the XANES signal for any orientation of the x-ray polarization vector onto the ac^* plane (22). However, by including a single polarization factor for any measured orientation and fitting simultaneously all of the polarized XANES spectra acquired, the method becomes generally applicable to crystals showing a XANES dichroism weaker than the one in the P2₁ Mb crystal (e.g., crystals with higher symmetry).

From the XRD data, the MbCN crystal structure was determined and refined to an R_{work} of 18.0% ($R_{\text{free}} = 23.0\%$). All 153 protein residues and the Fe-heme group could be positioned in the electron density, with 92% of the residues in the most favored region of the Ramachandran plot and 8% in additionally allowed regions. The overall stereochemical G factor was 0.06. The parameters all fall within acceptable limits set by the PROCHECK program, which compares the current model against a database of structures (25). The final model also contains 188 water molecules and one sulfate ion (Table 2). The electron density at the heme site is shown in Fig. 3*A*.

In the first stages of the refinement, the positions of the Fe ligands, i.e., the porphyrin ring, the axial CN[−] anion, and the His-93 imidazole group, were not restrained. In the final step, the geometrical information provided by XANES was used to restrain the geometry of the Fe ligands. The model was then subjected to restrained refinement by using REFMAC (26). In the final structure, the R_{work} and R_{free} values did not increase by imposing the geometry derived by the MXAN best fit (Table 2). The Fe–CN distance was within <0.1 Å of the XANES model, and the Fe–N_{His-93} distance was refined to a value of 2.08 Å as compared with the imposed value of 2.07 Å. Fig. 3*B* shows a sketch of the Fe–heme site, with the relevant structural parameters derived by XANES analysis via the MXAN package.

In Fig. 4, we demonstrate the sensitivity of XANES spectroscopy to the geometry of the Fe–heme site by comparing how spectra calculated for alternative geometries fit the experimental XANES spectrum. For clarity, only the $\epsilon \parallel c^*$ spectra are presented. In Fig. 4*A*, the experimental data (open circles) are shown together with the model calculation (solid line) obtained by using the coordinates from the MbCN x-ray structure at 1.8 Å (Protein Data Bank ID code 1EBC), the best resolution for this derivative before this work (27). In Fig. 4*B*, we plot a model calculation based on our present x-ray structure analysis at 1.4 Å without restraints by the XANES data. The large discrepancies between the experimental data and the calculated spectrum are evident. The spectrum in Fig. 4*C* is a calculation using the structural model obtained by refinement of the XRD data, restrained by the XANES parameters. The MXAN best fit is shown in Fig. 4*D*. This comparison demonstrates clearly that the XANES analysis enables discrimination between different structural models of the metal site and can be used as a restraining factor.

Discussion

The first attempt to interpret polarized XANES spectra of a single MbCO crystal (17) generated a vigorous debate about the role of XAS in structural biology, especially in comparison to protein crystallography (28–30). Since those early days, XANES spectroscopy has matured into a powerful tool, aided by significant advances in theory, experimental techniques, and instrumentation, culminating in the recent development of simultaneous XRD and absorption experiments on the same crystal (20).

XANES spectroscopy is a well established technique for determining the redox state of a metal compound (31). It can be a valuable tool to monitor x-ray photoreduction during XRD data collection. However, XANES and XRD can be combined

Fig. 4. Comparison of the experimental XANES spectrum ($\varepsilon \parallel$ heme-normal, $\alpha = 0^\circ$) with calculated spectra. Experimental and calculated spectra are shown as open circles and solid lines, respectively. (A) Calculated spectrum using the coordinates of the MbCN structure at 1.8-Å resolution [Protein Data Bank ID code 1EBC (26)]. (B and C) Calculated spectra using the coordinates of the present MbCN structure at 1.4-Å resolution, unrestrained (B) and restrained (C) with the XANES data. (D) MXAN best fit. The minimization of the χ^2 function was performed by variation of four axial parameters. Normalized χ^2 values are reported for each calculation.

in a much more powerful way, as we have demonstrated in this work, in which we exploit the high 3D spatial resolution of XANES for determining the structural parameters of the metal site in MbCN. As discussed by Cruickshank (32), reliable standard uncertainties on coordinates and bond lengths are needed for any discussion of nondictionary distances between atoms of different residues, between protein and solvent atoms, or between metal atoms and their ligands. Considering the error on the coordinates for the XRD data collection reported in Table 2 (diffraction-component precision index = 0.1046 Å), the Fe–ligand distances derived from the XANES approach on the same crystal are by an order of magnitude more accurate than the ones derived from crystallographic refinement at 1.4-Å resolution. We use the XANES-derived parameters as restraints in the crystallographic refinement, and the combination of x-ray absorption and scattering methods thus provides atomic precision for bond distances and bond angles at the metal site (21). Such a precision is essential for understanding the electronic structure of the metal site, including orbital occupancies and oxidation states, as was recently emphasized by Gray and coworkers (6). They used EXAFS spectroscopy to determine the iron coordination distances and, in particular, the length of the Fe(IV)=O bond in chloroperoxidase compound II in solution with an error <0.1 Å. As the geometry of the heme iron coordination greatly affects the chemical nature of the oxoferryl complex, its precise analysis is fundamental to understanding the reactivity of thiolate-ligated heme enzymes, including chloroperoxidase, cytochrome P450, and nitric oxide synthase.

Based on this example, it is instructive to compare the closely related EXAFS and XANES techniques. Metalloproteins are dilute systems, as the x-ray absorption arises merely from the metal centers. To achieve the reported precision on the heme axial

parameters, Gray and coworkers (6) collected EXAFS data with good signal-to-noise ratio up to photoelectron wave vector magnitudes of $k = 16 \text{ \AA}^{-1}$. In our XANES experiment, the k range was limited to $k < 7 \text{ \AA}^{-1}$ (i.e., the first 200 eV of the x-ray absorption spectrum). Although a multiple scattering EXAFS analysis is in principle capable of providing full 3D information, it requires experimental data over a much wider k range (typically, $k > 13 \text{ \AA}^{-1}$). However, data collection on a single crystal over this extended k range would take much longer. In our experimental setup, we would need several hours instead of ≈ 30 min per orientation. We are able to collect a complete XRD/XANES data set on a single crystal within a few hours, which is necessary to avoid crystal damage even at cryogenic temperature (33).

For comparison with XANES, we have also performed an EXAFS analysis with the XAS data in the narrow k range of 3 to $7 \text{ \AA}^{-1}$. This procedure is described in *SI Text*. Only the Fe–CN and Fe–N_{pyrrol} distances were refined because they had significantly different final values in the XANES and XRD analyses. The EXAFS structural parameters thus obtained were in complete agreement with the XANES results (see *SI Fig. 6* and *SI Table 3*). However, the accuracy of the parameters was rather low because of the limited k range and the well known statistical correlation between the edge position, which is a free parameter in EXAFS analysis, and the bond lengths. By contrast, the XANES analysis within the same k interval yielded the complete 3D structure of the metal site at atomic resolution. This property renders the XANES technique particularly suitable for solving the structures of metalloprotein active sites in crystals, which is especially useful if the crystals do not diffract to high resolution.

Because of the difficulties related to the treatment of the potential and inelastic losses, earlier work attempted only qualitative analyses of the low-energy parts of the XAS spectra, either by comparing with model compounds or complementing EXAFS studies. Only a few attempts were made to exploit the sensitivity of XANES to structural parameters, mostly related to known structures (35). With the MXAN fitting procedure based on the full multiple scattering approach, the photo-absorption cross-section can be calculated beginning at the absorption edge. Here, we have demonstrated that the combination of XANES and XRD on a single protein crystal and the use of the XANES structural parameters as restraints in the refinement of the x-ray structure allow determining the structure of a metalloprotein with very high active-site resolution. This approach offers substantial advantages for the crystal structure determination of metalloproteins, providing more accurate geometries of the metal sites than each method alone. The combination of XRD and XAS will also be advantageous for the study of redox-induced conformational dynamics by using time-resolved or cryo-trapping experiments, for example on heme enzymes such as cytochrome P450 (34).

Conclusions

We conclude with a few remarks on the current state of the art of the XANES method. Analysis of the low energy range of the XAS spectrum is revealing, as both structural and electronic information can be extracted from this region. Consequently, it is important to further advance XANES theory, thereby avoiding phenomenological treatments of some aspects, as are still implemented in the MXAN procedure. Recent theoretical efforts have addressed vibrational effects on the high energy part of the XANES spectrum (36). These contributions should be included in theoretical modeling if the fitting range is extended beyond the near edge, because both vibrational damping and electronic damping caused by the imaginary part of the optical potential are to be considered. More difficult theoretical challenges include the substitution of the muffin-tin approximation for the shape of the molecular potential, the implementation of a multichannel excitation multiple scattering theory, and the determination of

the final-state optical potential (36). As long as these theoretical limitations persist, any nonstructural parameters of the theory have to be extracted from model systems that provide suitable approximations of metal site geometries in biological molecules. Based on this strategy, the combination of XRD and XAS experiments on single crystals is an attractive structure determination method, and we suggest that it should be implemented as a standard technique at modern synchrotron facilities.

Materials and Methods

Protein Crystal Preparation. Mb was extracted from skeleton muscle tissue of a sperm whale stranded at Sankt Peter Ording (Germany) in 1998, purified, and set up for batch crystallization in a solution containing 3.2 M ammonium sulfate and 0.05 M phosphate buffer (pH 6.5) according to Kendrew and Parrish (37). Monoclinic ($P2_1$) single crystals of (ferric) aquomet Mb appeared within 2 weeks. The cyanomet derivatives were obtained by soaking the crystals for 30 min in mother liquor containing 30 mM NaCN. The crystals were suspended in a cryoprotectant solution containing the mother liquor and 25% glycerol and mounted in a cryoloop (38) before starting the experiments. After crystal screening, a thickness between 300 and 500 μm yielded optimal diffraction quality and signal-to-noise ratio for XANES measurements.

XRD Data Collection and Refinement. Data in the $P2_1$ space group were collected at 100 K up to a resolution of 1.4 $\text{\AA}$ on beam line 10.1 at the Daresbury Synchrotron Radiation Source using a wavelength of 0.94 $\text{\AA}$. After processing with the program HKL2000 (39), a total of 24,194 unique reflections were obtained with an agreement factor R_{merge} of 4.2% for equivalent reflections. The structure was solved by molecular replacement using MOLREP (40), taking the structure of sperm whale met MbCN at 1.8- $\text{\AA}$ resolution (27) as a search model (Protein Data Bank ID code 1EBC). The best solution yielded one monomer in the asymmetric unit with a correlation coefficient of 55.5 and an R factor of 34.5% after rigid-body refinement. Restrainted refinement of the model was performed with REFMAC5 (26); manual rebuilding was carried out with COOT (41). Water molecules were added by using ARP/wARP (42) and positioned only if well defined positive peaks were present in both the $2F_o - F_c$ and $F_o - F_c$ electron density maps, and if they were in hydrogen-bonding distance with either protein atoms or other presumed water molecules. Data collection and refinement statistics are summarized in Table 2.

Crystal Orientation and XANES Data Collection. Because of the asymmetry of the heme site, the XANES signal with the polarization axis of the x-ray beam, ϵ , parallel to the c^* axis of the reciprocal lattice (nearly parallel to the heme normal) differs from the one obtained with ϵ parallel to the a axis (in the heme plane). To measure this dichroism, we aligned the MbCN crystal with the c^* axis parallel to the polarization axis of the x-ray beam, ϵ , by means of the χ circle, which can be rotated by 90° and is normally used for automatic crystal mounting. XANES data were measured immediately before and after XRD data collection on the same crystal by rotating the crystal around the b axis, i.e., by collecting angle-resolved XANES spectra with the x-ray polarization vector ϵ advancing from ($\epsilon \parallel a$) to ($\epsilon \parallel c^*$) by varying the θ angle of the diffractometer (see SI Fig. 5).

Fe K-edge x-ray absorption spectra were collected in fluorescence mode by using an energy-resolving four-element monolithic Ge detector of very high purity. The energy resolution at the Fe K_{α} fluorescence ($\approx 6,400$ eV) was ≈ 170 eV. Four XANES spectra were collected for each orientation between 7,100 and 7,300 eV in 0.5-eV steps with a dwell time of 1 s. The fluorescence count rate jumped from $\approx 2,200$ counts per s per element below the edge (7,100 eV) to $\approx 85,000$ counts per s per element

above the edge (7,200 eV), giving a total noise-to-jump ratio of 1.4×10^{-4} and 8.8×10^{-4} , respectively. Measurement of a single XANES spectrum took only 7 min, so that collection of the entire XRD and XANES angle-resolved data set was carried out in <4 h of x-ray irradiation. In all experimental spectra presented here, the energy was calibrated by referencing to the absorption edge of a metallic Fe foil.

Polarized XAS Spectra of MbCN. In the general case, assuming that the incident beam is linearly polarized along a direction denoted by the vector ϵ , each spectrum j (i.e., each crystal orientation) samples a particular direction ϵ_j . The fluorescence intensity σ_j emitted as a function of the energy E from a small crystal is proportional to the quadratic form:

$$\sigma_j(E) \propto \epsilon_j^T \sigma(E) \epsilon_j, \quad [1]$$

where $\sigma(E)$ denotes the absorption tensor, the superscript T indicates the transpose and $\sigma_j(E)$ denotes the effective scalar linear absorption coefficient along the direction ϵ_j . In the case of an Fe-heme site with approximate C_4 symmetry, it is advantageous to calculate the XAS components linearly polarized along the symmetry axis (the heme normal), $\sigma_{\text{normal}}(E)$, and along an arbitrary direction in the C_4 plane (treating the heme plane as a circular absorber), $\sigma_{\text{heme}}(E)$, in the fit procedure. Introducing an angle α between the polarization vector ϵ and the heme normal of one Mb molecule, the linearly polarized spectrum $\sigma_\alpha(E)$ is approximated by a simple function of σ_{normal} and σ_{heme} ,

$$\sigma_\alpha(E) \propto \sigma_{\text{normal}}(E) \cos^2 \alpha + \sigma_{\text{heme}}(E) \sin^2 \alpha. \quad [2]$$

For a sperm whale Mb single crystal in the space group $P2_1$, the two heme sites of the unit cell give equivalent contributions to the XAS spectrum for any x-ray polarization vector perpendicular to crystal axis b . Eq. 2 results in the following relationships for the spectra with polarization vector ϵ oriented parallel to crystal axes c^* and a (22),

$$\begin{aligned} \sigma_{c^*}(E) &\propto 0.85 \cdot \sigma_{\text{normal}}(E) + 0.15 \cdot \sigma_{\text{heme}}(E), \\ \sigma_a(E) &\propto 0.005 \cdot \sigma_{\text{normal}}(E) + 0.995 \cdot \sigma_{\text{heme}}(E). \end{aligned} \quad [3]$$

The relationships between the crystal axes and the x-ray polarization can be found in SI Text and SI Fig. 5. In the XANES analysis, σ_{normal} and σ_{heme} are fitted to the experimental data. According to Eq. 3, the experimental spectra for σ_{heme} and σ_{normal} are assumed to be proportional to σ_a and $(\sigma_{c^*} - 0.15 \sigma_a)/0.85$, respectively.

XANES Theory and Fit Procedure. Recently, a method was introduced for the calculation of XANES spectra (19) that is based on the full multiple scattering approach in the framework of the muffin-tin potential approximation. It also takes into account inelastic processes by way of a Lorentzian broadening function. This approach is implemented in the MXAN program package for modeling and fitting experimental spectra. MXAN has been successfully applied to K-edge spectra of a number of transition metal compounds (23, 43–45) and also has been used to interpret XANES spectra taken on MbCO crystals (24) and solutions (46) in low-temperature photolysis experiments.

For the calculation of the MbCN XANES spectra, altogether 32 atoms, i.e., the porphyrin macrocycle, the His-93 imidazole, and the CN^- ion were considered. The cluster size and the value of λ_{max} , the maximum orbital angular momentum in the spherical harmonic expansion of the scattering path operators, were chosen based on a convergence criterion. The molecular potentials of the central iron and its next neighbors, i.e., the FeN_5CN cluster, were obtained by the spin-unpolarized, self-consistent

field method imposing the formal valence of each atom. Spin polarization effects and partial metal–ligand charge donations were not considered in this calculation. Spin polarization corrections are only relevant for photoelectron energies on the order of 1 eV, and changes in the oxidation state of the central metal ion result in a constant shift of the edge that has no influence on the structural fit.

The least-squares minimization of the χ^2 function was performed in the space of the following five structural parameters: heme–core size, leading to the Fe–N_{pyrrol} distance, the Fe–C and Fe–N_{HIS} distances, the Fe–C–N bending angle, and the C–N bond length. The His-93 imidazole and the heme pyrrol rings were treated as perfectly rigid bodies. The fit is only weakly sensitive to the Fe displacement from the heme plane, to the tilting angle between the heme normal and the Fe–C vector, and to the tilting angle of the histidine residue, as reported for MbCO (24).

The MXAN procedure uses only four nonstructural parameters, one for the choice of the atomic spheres in the muffin-tin molecular potential and the other three parameters relating to the phenomenological damping (19). To assess the influence of these parameters on the structural fit, we have systematically varied them in XANES analyses on related model systems and proteins (23, 24). Their influence on the structural parameters was found to be weak.

After measuring XAS of the single crystal, we fitted only one parameter (the Fe–N_{pyrrol} distance) in the polarized XANES spectrum $\sigma_{\text{heme}}(E)$ and four parameters in the polarized spectrum $\sigma_{\text{normal}}(E)$. This procedure strongly reduces the possibility that the fit becomes trapped in local minima of the structural parameter space. Quality and stability of the fit were further improved by choosing a wide energy interval (130–180 eV) that

is much larger than all of the corrections to the potential determination (nonmuffin-tin effects, exchange, and correlation potential), which are on the order of a few electron volts (36). There are well known limitations to the muffin-tin approximation, especially concerning metal sites with an open coordination sphere, e.g., 2-, 3-, or 4-fold square-planar systems (47, 48). For photoelectron energies <10 eV, these systems exhibit pronounced features at the preedge and in the lower rising edge, probing unoccupied or partially occupied bound states that are difficult to model. The MXAN method has not been tested and validated for these coordination geometries. A possible strategy could be to omit preedge features from the analysis by fitting the data starting from photoelectron energies of 3–8 eV above the rising edge, without losing any structural information. In fact, the main objective of the MXAN method is not to accurately fit preedge features, but to extract structural parameters of the compound under study by using an appropriate spectral range starting at the rising edge, spanning the low-energy region of the full multiple scattering regime, i.e., the XANES continuum states, and extending up to the low-energy part of the EXAFS region.

We thank the Central Laboratory of the Research Councils for provision of facilities at Daresbury Laboratory. R.W.S., S.S.H. and M.C. were supported by Biotechnology and Biological Sciences Research Council Grants 719/B15474 and 719/REI20571, the Medical Research Council, the Engineering and Physical Sciences Research Council, Northwest Development Agency Project Award N0002170, and the Central Laboratory of the Research Council. G.U.N. was supported by Deutsche Forschungsgemeinschaft Grant Ni 291/3 and the Fonds der Chemischen Industrie. B.V., C.S. and A.A. were supported by Ministero dell'Università della Ricerca Grants 2005050270-004 and RBLA03B3KC-004.

- Hill HAO, Sadler PJ, Thomson AJ (1999) *Metal Sites in Proteins and Models* (Springer, Berlin).
- Cowan JA (1996) *Inorganic Biochemistry: An Introduction* (Wiley, New York).
- Nilsson K, Hersleth HP, Rod TH, Andersson KK, Ryde U (2004) *Biophys J* 87:3437–3447.
- Koningsberger DC, Prins R (1988) *X-Ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS, and XANES* (Wiley, New York).
- Rehr JJ, Albers RC (2000) *Rev Mod Phys* 72:621–654.
- Green MT, Dawson JH, Gray HB (2004) *Science* 304:1653–1656.
- Sheldrick GM, Schneider TR (1997) *Methods Enzymol* 227:319–343.
- Abola A, Bernstein FC, Bryant SH, Koetzle TF, Weng J (1987) in *Protein Data Bank in Crystallographic Databases-Information Content, Software Systems, Scientific Applications*, eds Allen FH, Bergerhoff G, Sievers R (Data Commission of the International Union of Crystallography, Bonn), pp 107–132.
- Antonini E, Brunori M (1971) *Hemoglobin and Myoglobin in Their Reactions with Ligands* (North-Holland, Amsterdam).
- Frauenfelder H, Wolynes PG (1985) *Science* 229:337–345.
- Frauenfelder H, McMahon BH, Fenimore PW (2003) *Proc Natl Acad Sci USA* 100:8615–8617.
- Kendrew JC, Dickerson RE, Strandberg BE, Hart RG, Davies DR, Phillips DR, Phillips DC, Shore VC (1960) *Nature* 185:422–427.
- Nienhaus GU, Young RD (1996) in *Encyclopedia of Applied Physics*, ed Trigg GL (Wiley, New York), Vol 15, pp 163–184.
- Nienhaus GU (2005) *Protein-Ligand Interactions* (Humana, Totowa, NJ).
- Parak FG, Nienhaus GU (2002) *Chem Phys Chem* 3:249–254.
- Natoli CR, Benfatto M, Della Longa S, Hatada K (2003) *J Synchr Rad* 10:26–42.
- Bianconi A, Congiu-Castellano A, Durham PJ, Hasnain SS, Phillips S (1985) *Nature* 318:685–687.
- Shulman RG, Eisenberger P, Blumberg WE, Stombaugh NA (1975) *Proc Natl Acad Sci USA* 72:4003–4007.
- Benfatto M, Della Longa S (2001) *J Synchr Rad* 8:1087–1094.
- Cianci M, Antonyuk S, Bliss N, Bailey MW, Buffey SG, Cheung KC, Clarke JA, Derbyshire GE, Ellis MJ, Enderby MJ, et al. (2005) *J Synchr Rad* 12:455–466.
- Hasnain SS, Hodgson KO (1999) *J Synchr Rad* 6:864–875.
- Della Longa S, Arcovito A, Vallone B, Castellano AC, Kahn R, Vicat J, Soldo Y, Hazemann JL (1999) *J Synchr Rad* 6:1138–1147.
- Hayakawa K, Hatada K, D'Angelo P, Della Longa S, Natoli CR, Benfatto M (2004) *J Am Chem Soc* 126:15618–15623.
- Della Longa S, Arcovito A, Girasole M, Hazemann JL, Benfatto M (2001) *Phys Rev Lett* 87:155501.
- Morris AL, MacArthur MW, Hutchinson EG, Thornton JM (1992) *Proteins* 12:345–364.
- Murshudov GN, Vagin AA, Dodson EJ (1997) *Acta Crystallogr D* 53:240–255.
- Bolognesi M, Rosano C, Lusso R, Borassi A, Rizzi M, Wittenberg JB, Boffi A, Ascenzi P (1999) *Biophys J* 77:1093–1099.
- Galloway J (1985) *Nature* 318:602.
- Fermi G, Perutz MF, Shulman RG (1987) *Proc Natl Acad Sci USA* 84:6167–6168.
- Shulman RG (1987) *Proc Natl Acad Sci USA* 84:973–974.
- Yachandra VK (1995) *Methods Enzymol* 246:638–675.
- Cruickshank DW (1999) *Acta Crystallogr D* 55:583–601.
- Teng TY, Moffat K (2002) *J Synchr Rad* 9:198–201.
- Schlichting I, Berendzen J, Chu K, Stock AM, Maves SA, Benson DE, Sweet RM, Ringe D, Petsko GA, Sligar SG (2000) *Science* 287:1615–1622.
- Binsted N, Hasnain SS (1996) *J Synchr Rad* 3:185–196.
- Natoli CR, Benfatto M, Della Longa S, Hatada K (2003) *J Synchr Rad* 10:26–42.
- Kendrew JC, Parrish RG (1956) *Proc R Soc London Ser A* 238:305–324.
- Teng TY (1990) *J Appl Crystallogr* 23:387–391.
- Otwinowski Z, Minor W (1997) *Methods Enzymol* 276:307–326.
- Vagin A, Teplyakov A (1997) *J Appl Crystallogr* 30:1022–1025.
- Emsley P, Cowtan K (2004) *Acta Crystallogr D* 60:2126–2132.
- Lamzin VS, Wilson KS (1993) *Acta Crystallogr D* 49:129–147.
- Benfatto M, D'Angelo P, Della Longa S, Pavel NV (2002) *Phys Rev B* 65:174205.
- D'Angelo P, Benfatto M, Della Longa S, Pavel NV (2002) *Phys Rev B* 66:064209.
- Frank P, Benfatto M, Szilagyi RK, D'Angelo P, Della Longa S, Hodgson KO (2005) *Inorg Chem* 44:1922–1933.
- Arcovito A, Lamb DC, Nienhaus GU, Hazemann JL, Benfatto M, Della Longa S (2005) *Biophys J* 88:2954–2964.
- Munoz-Paez A, Diaz-Moreno S, Sanchez Marcos E, Rehr JJ (2000) *Inorg Chem* 39:3784–3790.
- Chen K, Yuldasheva S, Penner-Hahn JE, O'Halloran TV (2003) *J Am Chem Soc* 125:12088–12089.