The *MXAN* procedure: a new method for analysing the XANES spectra of metalloproteins to obtain structural quantitative information

M. Benfatto,[a]* S. Della Longa[b] and C. R. Natoli[a]

[a]Laboratori Nazionali di Frascati dell'INFN, CP13, 00044 Frascati, Italy, and [b]Universita' dell'Aquila, Via Vetoio, loc. Coppito II, 67100 l'Aquila, Italy. E-mail: benfatto@lnf.infn.it

The first quantitative analyses are reported of the Fe *K*-edge polarized X-ray absorption near-edge structure (XANES) of a single crystal of the iron protein carbonmonoxy-myoglobin (MbCO) and of its cryogenic photoproduct Mb*CO. The CO—Fe–heme local structure has been determined using a novel fitting procedure, named *MXAN*, which is able to fit the XANES part (from the edge to about 200 eV) of experimental X-ray absorption data. This method is based on the comparison between the experimental spectrum and several theoretical spectra that are generated by changing the relevant geometrical parameters of the site around the absorbing atom. The theoretical spectra are derived in the framework of the full multiple-scattering approach. The *MXAN* procedure is able to recover information about the symmetry and atomic distances, and the solution is found to be independent of the starting conditions. The extracted local structure of Mb*CO includes an Fe—CO distance of 3.08 (7) Å, with a tilting angle between the heme normal and the Fe—C vector of 37 (7)° and a bending angle between the Fe—C vector and the C—O bond of 31 (5)°

Keywords: X-ray absorption near-edge spectroscopy (XANES); data analysis.

1. Introduction

The low-energy part of an XAS spectrum (the XANES region) is of great interest for biological studies since it is extremely sensitive to the structural details of the absorbing site (overall symmetry, distances and bond angles). Therefore, in principle, an almost complete recovery of the geometrical structure within 6–7 Å from the absorbing site can be achieved from the experimental data. Structural information on protein-metal sites can be obtained with atomic resolution, in any state of the protein sample (crystal, solution), and this information can then be either compared with known X-ray structures at high resolution or used to obtain information on proteins that have proven difficult to crystallize (Hasnain & Hodgson, 1999; Cruickshank, 1999). Moreover, electronic information about the metal site, like its charge and spin state, can be deduced from an accurate interpretation of the XAS spectrum. Furthermore, the effects of the atomic thermal disorder are limited, because the Debye–Waller-like term is almost independent of temperature at low energies. In fact, any multiple-scattering (MS) signal of order n can be written as a sine function whose arguments contain the term $2kR_{\rm tot}$, where $R_{\rm tot}$ is the total length of the n-order MS path. Consequently the associated total Debye–Waller factor contains the term $\exp(-2k^2\sigma^2)$, which gives the strongest contribution to the Debye–Waller factor and is almost equal to unity at low k values.

Therefore, in principle, different main conformational states of a protein-metal site can be investigated, and their structure can be extracted, if the site's conformational equilibrium is modulated (*e.g.* by pH or temperature changes) or destabilized (*e.g.* by light triggering) to obtain each of its different specific states (Della Longa *et al.*, 1998).

However, the quantitative analysis of the full XAS spectrum, including the edge, is a complex many-body problem that requires an adequate treatment and time-consuming algorithms in order to calculate the absorbing cross section in the framework of the full multiple-scattering approach. Because of these difficulties, analyses of the pre-edge and low-energy parts of XAS spectra (up to 50–100 eV) have been exploited so far only on qualitative grounds: either by comparison with model compounds or as an aid for EXAFS studies or more advanced investigations, such as those based on the analysis of contributions related to correlation functions of orders higher than two (Filipponi & Di Cicco, 1995).

Few attempts have been made to quantify the theoretical sensitivity of the low-energy part of the spectrum to the structural parameters. In the few examples that can be found in the literature, quantitative comparisons between experimental data and *ab initio* calculations are mostly related to known structural compounds (Binsted & Hasnain, 1996). This fact arose mainly because of the lack of a fitting procedure based on the full MS approach. Such a fitting procedure should allow the exact calculation of the photoabsorption cross section from the edge and should avoid any *a priori* selection of the relevant MS paths.

Recently some of us have proposed (Benfatto & Della Longa, 2001) and applied to several systems (Della Longa *et al.*, 2001) a new method for performing a quantitative analysis of the XANES energy range, *i.e.* from the edge up to 200 eV. The method, called *MXAN*, is based on a comparison between experimental data and many theoretical spectra that are calculated by varying selected structural parameters of an initial putative structure, *i.e.* a well defined initial geometrical configuration around the absorber. Hundreds of different geometrical configurations are needed to obtain the best fit to the experimental data. However, the XANES spectra that are related to these configurations are calculated in a reasonable time. The optimization in the parameter space is achieved by the minimization of the square residual function in the parameter space. The calculations are performed in the energy space without involving any Fourier-transform algorithm; polarized spectra can be easily analysed because the calculations are performed by the full MS approach.

2. The *MXAN* procedure

The *MXAN* procedure uses the set of programs developed by the Frascati theory group (Natoli & Benfatto, 1986; Tyson *et al.*, 1992 and references therein): in particular, *VGEN*, a generator of muffin-tin potentials, and the *CONTINUUM* code for the calculation of the full multiple-scattering cross section. The optimization in the space of the parameters is achieved using the *MINUIT* routines of the CERN library; a single best-fit procedure takes typically 8 h on a UNIX scalar α-VAX machine for a calculation involving six fitting parameters in a cluster of 35 atoms. The *MINUIT* routines minimize the square residual function

$$S^2 = n\sum_{i=1}^{m} w_i\left[\left(y_i^{\rm th} - y_i^{\rm exp}\right)\varepsilon_i^{-1}\right]^2 / \sum_{i=1}^{m} w_i, \qquad (1)$$

where n is the number of independent parameters, m is the number of data points, $y_i^{\rm th}$ and $y_i^{\rm exp}$ are the theoretical and experimental values of absorption, ε_i are the individual errors in the experimental data set, and w_i are statistical weights. For w_i = constant = 1, the square residual function S^2 becomes the statistical χ^2 function. The application of the software package to several test cases shows that the best-fit

solution is independent of the minimization strategy and of the starting conditions.

The *MXAN* method is based on the muffin-tin approximation for the shape of the potential and uses the concept of complex optical potential, which is based on the local-density approximation of the self-energy of the excited photoelectron. The effects of non-MT corrections on the XANES calculation are not well understood; nevertheless, evidence exists that their influence, if present, is confined within the first 20–30 eV from the edge (Cabaret *et al.*, 1999; Joly, 2001; Benfatto & Della Longa, 2001). The application of the *MXAN* procedure to several test cases confirms this indication, since the non-MT corrections are found to weakly influence the structural determination (D'Angelo *et al.*, 2002). This result is not surprising as it is possible to demonstrate that the absorption cross section can be written as the sum of the MT-calculated cross section and other terms containing all of the non-MT corrections (Benfatto & Della Longa, 2001). These corrections depend on the system and go to zero as the energy increases. Their influence is strongly reduced if the fits are performed in the energy range from the edge up to 200 eV. In this way the geometrical arrangements restrain the numerical results of the fitting procedure.

The real part of the self-energy is calculated either by the X-α approximation or by using the Hedin–Lundqvist (HL) potential. To avoid over damping of the complex part of the HL potential at low energies in the case of covalent molecular systems, the *MXAN* method can alternatively account for all the inelastic processes by a convolution with a broadening Lorentzian function that has an energy-dependent width of the form $\Gamma(E) = \Gamma_c + \Gamma_{mfp}(E)$. The constant part Γ_c includes the core-hole lifetime and the experimental resolution, while the energy-dependent term represents all the intrinsic and extrinsic inelastic processes. The $\Gamma_{mfp}(E)$ function is zero below an onset energy E_s (which in extended systems corresponds to the plasmon excitation energy) and begins to increase from a value A_s following the universal functional form of the mean free path in solids. Both the onset energy E_s and the jump A_s are introduced into the $\Gamma_{mfp}(E)$ function *via* an arctangent functional form to avoid discontinuities and to simulate the electron–hole pair excitations. The numerical values of E_s and A_s are derived *via* a Monte Carlo fit at each step of the computation (*i.e.* for each configuration), a procedure similar to that used in optimization by simulated annealing (Kirkpatrick *et al.*, 1983). This type of approach can be justified on the basis of a multichannel multiple-scattering theory, as discussed by Natoli *et al.* (2003), where the theoretical background of the method is described in detail. Note that the above considerations are completely equivalent to the discussion presented by Natoli *et al.* (2003) because of the equivalence of the use of a suitable complex self-energy in the calculation and a convolution of the photoabsorption cross section, calculated with the real part of the self-energy, with a Lorentzian function of a well defined energy dependent width.

We have also tested how the choice of MT radii influences the structural determination. If we change the percentage of overlap from 0% to 12% our results are very weakly affected. The *MXAN* method introduces four non-structural parameters, the influence of which seems to be limited to an increase of a few percent in the error value in the determination of the structural parameters.

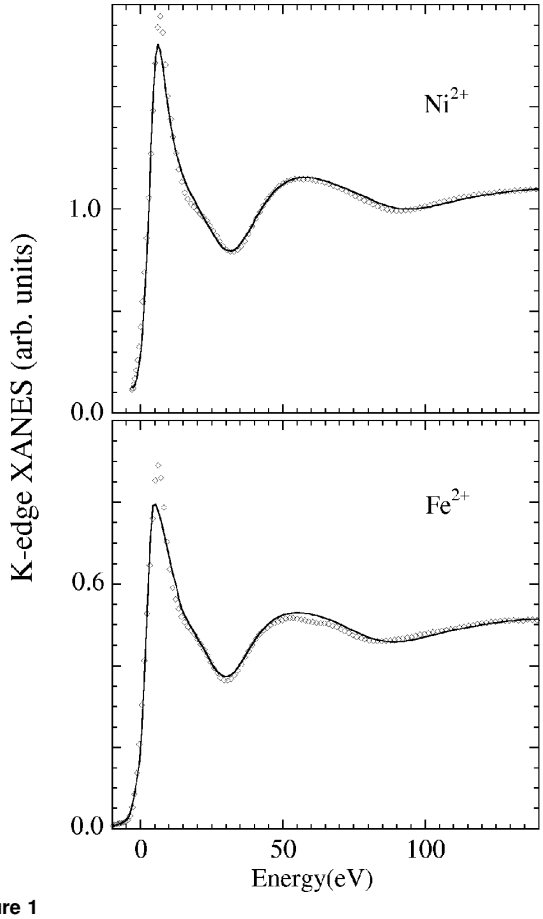

Figure 1
Upper frame: experimental (circles) and best-fit calculation (solid line) of the Ni K-edge XANES spectrum of the Ni^{2+} aqua ion. Lower frame: experimental (circles) and best-fit calculation (solid line) of the Fe K-edge XANES spectrum of the Fe^{2+} aqua ion.

3. Test cases: metal ions in solution

We show, as test cases, the *MXAN* analyses of the K-edges of Ni^{2+} and Fe^{2+} ions in aqueous solution. Both ions are often used as test cases because of the well defined formal valency of the ionic species and the very simple geometry around the absorber. The data at the K-edges were recorded in transmission mode using a Mylar cell at beam station 7.1 of the Daresbury Laboratory. A double-crystal Si(111) monochromator was used and the storage ring was operated at 2 GeV with an average current of about 150 mA. The samples were prepared to obtain a water solution of 5 mM and 50 mM of Ni^{2+} and Fe^{2+} ions, respectively. The pH was controlled in order to preserve the hydrated species. The background contribution from previous edges has been fitted with a linear function and subtracted from the raw data.

In Fig. 1 we report the comparison between the experimental data and the calculations related to the best-fit structures for both ions. These best fits correspond to an octahedral symmetry with an oxygen–metal distance of 2.03 (3) Å and 2.08 (2) Å for Ni^{2+} and Fe^{2+} solutions, respectively. H atoms are included in the calculations. The agreement between the experimental data and the best-fit theoretical curves is good for the whole energy range; the small discrepancies that remain in the intensity of the resonance at 15 eV are essentially due to the muffin-tin approximation for the shape of the potential. We note that the inclusion of the H atoms in the calculation does change the metal–oxygen distance determination: our value of 2.03 Å is compatible with the value of 2.00 Å that has been previously calculated without H atoms (Benfatto & Della Longa, 2001). In Fig. 2 we show a comparison between the experimental data related to the Ni^{2+} ion and the best-fit calculation that is obtained by using the complex HL potential. In this case the best fit structure corresponds

to an octahedron with a oxygen–metal distance of 2.04 (2) Å. Note that we obtain the right geometry despite the presence of relevant discrepancies between the experimental data and the best-fit calculation in the first 40 eV. These discrepancies are due to the peculiar behaviour of the complex part of the HL potential. The presence of such discrepancies demonstrates the importance of using the phenomenological approach to calculate the inelastic losses, in order to have full access to the structural information that the low-energy part of the XANES spectrum contains. The use of this approach is particularly important in biological systems, where most of the experiments involve data with a low S/N ratio and a limited k-range.

4. A biological application: the low-temperature photoproduct of MbCO

The photolysis of the protein carbonmonoxy-myoglobin (MbCO) at low temperature [*i.e.* the dissociation, induced by visible light, of the CO molecule from the Fe–heme complex in the hydrophobic core of myoglobin (Mb)] is a protein dynamic event (Petsko, 1994) that has been addressed by many theoretical and experimental studies (Schlichting *et al.*, 1994; Teng *et al.*, 1994; Vitkup *et al.*, 1997; Teng *et al.*, 1997; Meller & Elber, 1998; Chu *et al.*, 2000). This process is represented in a simplified scheme by the three-well model, MbCO → Mb*CO → Mb, in which MbCO is the initial main state of the protein, Mb is the final main state when the CO molecule is in the solvent and Mb*CO represents all the intermediate states when the CO molecule is in the protein matrix.

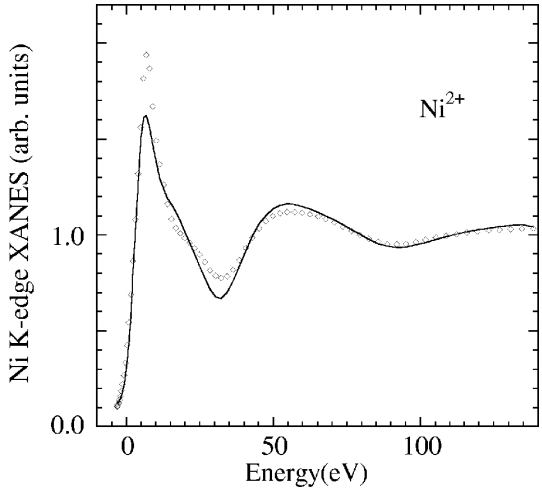

Figure 2
Comparison between the experimental data (circles) and best-fit calculation (solid line) of the Ni K-edge XANES spectrum of the Ni^{2+} aqua ion. In this case we have used the complex HL potential to account for the inelastic losses of the photoelectron.

The X-ray crystal structures of MbCO and Mb at a resolution of 1.15 Å (Kachalova *et al.*, 1999; Vojtechovsky *et al.*, 1999) show that concerted motions of the Fe atom, the heme and the helices E and F of the protein accompany the CO dissociation at room temperature. At room temperature the photodissociated CO molecule rebinds to the Fe atom on a microsecond time scale. At $T < 30$ K the CO rebinds very slowly (Iizuka *et al.*, 1974), so that an almost 100% conversion to the photoproduct can be achieved in the experimental time window of standard techniques, but the protein matrix remains frozen in the conformation of ligated MbCO. X-ray diffraction studies of the cryogenic Mb*CO photoproduct at $T = 20$ K (Schlichting *et al.*, 1994) and $T = 40$ K (Teng *et al.*, 1994) reported different positions of the CO molecule in the heme pocket. According to later X-ray studies (Teng *et al.*, 1997) it has been proposed that these different positions correspond to the migration of the CO molecule. Sketches of the MbCO (from Vojtechovsky *et al.*, 1999) and Mb*CO (from Schlichting *et al.*, 1994) structures are depicted in Fig. 3.

As shown by Bianconi *et al.* (1985), the XANES spectrum of sperm-whale MbCO in solution can be 'experimentally' decomposed into two almost perfectly polarized spectra by collecting measurements from the $P2_1$ single crystal formed by this protein. A pictorial view of the crystal unit cell, including the Fe–heme sites (with proximal histidine and the CO ligand) of the two symmetry-related myoglobin molecules, is shown in Fig. 4. The monoclinic unit cell is seen along a projection parallel to the b axis. It is possible to rotate the crystal to have the X-ray polarization vector $\boldsymbol{\varepsilon}$ oriented either parallel to the **a*** direction in the crystal, which forms an angle of 23° with the heme normal, or parallel to the c axis, which forms an angle of 86° with the heme normal, *i.e.* is almost parallel to the heme planes. As a consequence of this particularly favourable orientation and of the symmetry of the Fe–heme in the crystal, the $\boldsymbol{\varepsilon}$//**c** spectrum contains a 99.5% contribution from scattering pathways within the heme plane, and the $\boldsymbol{\varepsilon}$//**a*** spectrum contains a 15% contribution from scattering pathways within the heme plane and an 85% contribution from scattering pathways along the heme normal. The solution spectrum and the $\boldsymbol{\varepsilon}$//**c** and $\boldsymbol{\varepsilon}$//**a*** polarized spectra of the MbCO single crystal are reported in Fig. 5.

In recent years, the use of both X-ray dynamic focusing and the high-energy-resolving 24-element-array fluorescence detector has allowed the collection of polarized XANES spectra of a small single crystal ($0.2 \times 0.2 \times 0.3$ mm) of MbCO and its photoproduct Mb*CO at 20 K. These spectra were obtained under continuous illumination by a white light source at the BM32 beamline of the European Synchrotron Radiation Facility. Details of the experiment have been reported previously (Della Longa *et al.*, 1999). The experimental conditions were similar to those of Schlichting *et al.* (1994), who reported a well defined single photolized species with the CO molecule located in a 'primary' docking site above the pyrrol C atom of the heme, at a distance of about 3.6 Å from the Fe atom. The experimental data obtained by Della Longa *et al.* (1999) are shown in Fig. 6:

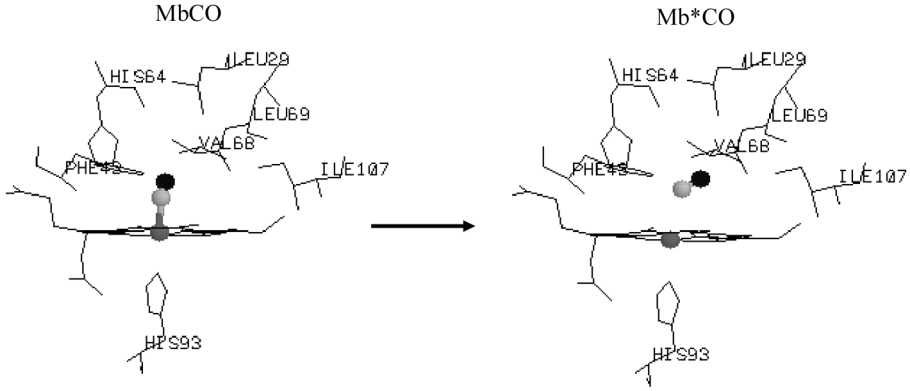

Figure 3
A sketch of the reported MbCO (from Vojtechovsky *et al.*, 1999) and Mb*CO (from Schlichting *et al.*, 1994) structures. The CO molecule after the Fe—C bond breaking, which is induced by light at $T = 20$ K, reaches a 'primary' docking site above the pyrrol C of the heme at about 3.6 Å from the Fe atom.

Table 1
Comparison between different structural determinations of the Fe-heme site geometry in MbCO.

Distance units: Å. Angle units: °. The resolution of the X-ray diffraction experiments is indicated. The values in parentheses are the statistical errors as evaluated by the fitting procedure.

Experiment	Fe—Np	Fe—Nhis	Fe—C	α	β	C—O
XRD† 1.5 Å	1.97	2.19	1.92	3	38	1.17
XRD‡ 1.1 Å	1.98	2.06	1.73	4	7	1.12
XRD§ 1.1 Å	1.98	2.06	1.82	9	9	1.09
EXAFS¶	2.01	2.20	1.93	–	–	–
XANES	2.00 (2)	2.06 (3)	1.83 (2)	–	14 (4)	1.07

† PDB code 1MBC. ‡ PDB code 1BZR. § PDB code 1A6G. ¶ Powers *et al.* (1987) solution sample.

Table 2
Comparison between different structural determinations of the Fe–heme site geometry in Mb*CO.

Distance units: Å. Angle units: °. The resolution of the X-ray diffraction experiments is indicated. The values in parentheses are the statistical errors as evaluated by the fitting procedure.

Experiment	Fe—Np	Fe— Nhis	Fe—C	α	β	C—O
XRD† 1.7 Å	2.00	2.11	2.84	23	45	1.21
XRD‡ 1.5 Å	1.97	2.25	3.60*	27*	54	1.12
EXAFS§	2.03	2.22	–	–	–	–
XANES	2.03 (2)	2.05 (2)	3.08 (7)	37 (7)	31 (5)	1.24

† PDB code 1AJH. ‡ PDB code 1ABS. § Powers *et al.* (1987) solution sample.

dashed lines depict the polarized spectra of MbCO. In the upper frame the polarization vector $\boldsymbol{\varepsilon}$ is oriented parallel to the heme plane. The main effect is a red-shift of the maxima at about 7200 and 7270 eV in Mb*CO relative to MbCO. This result suggests a change in the frequency of the main oscillating component of the signal, in agreement with the partial elongation, reported by X-ray diffraction and previous EXAFS data in solution, of the distance between the Fe and the N atoms that belong to the pyrrolic ring (see Tables 1 and 2). In the lower frame, the polarization vector $\boldsymbol{\varepsilon}$ is oriented parallel to **a***. The photolysis has dramatic effects on the low-energy range of the spectrum. After the Fe—CO bond is broken, a completely new spectrum appears with a different edge position and different prominent features.

In order to perform a quantitative analysis of these spectra we have used the *MXAN* procedure to fit the experimental data in terms of a selected set of parameters. In this case they are the first-shell distances d(Fe—N_p), d(Fe—C) and d(Fe—N_{his}), the tilting angle α between the Fe—C vector and the heme normal, the bending angle β between the Fe—C vector and the C—O bond, and the C—O bond length. During the fit, the imidazole ring of the histidine and the pyrrolic rings of the porphyrin rigidly follow the motion of the N_{his} and N_p atoms, respectively. The Fe–heme site and an indication of the parameters used are depicted in Fig. 7. Fig. 7 includes the heme plane, *i.e.* the iron and the porphyrin planar group, the CO ligand, and the imidazole group of the histidine protein residue. In this figure N_p indicates the pyrrolic N atoms of the porphyrin, N_{his} indicates the imidazole N atom of the histidine, α indicates the angle between the Fe—C vector and the heme normal (tilting angle), and β is the angle between the Fe—C vector and the C—O bond (bending angle).

For the *MXAN* analysis we used a cluster of 32 atoms, *i.e.* the porphyrin macrocycle, the histidine imidazole and the CO molecule. The molecular potential for the nearest atoms, *i.e.* the FeN_5CO cluster, was obtained by the self-consistent field (SCF) method, which imposes the formal valence of each atom (Pedio *et al.*, 1994; Díaz-Moreno *et al.*, 2000). Muffin-tin radii are chosen based on the Norman

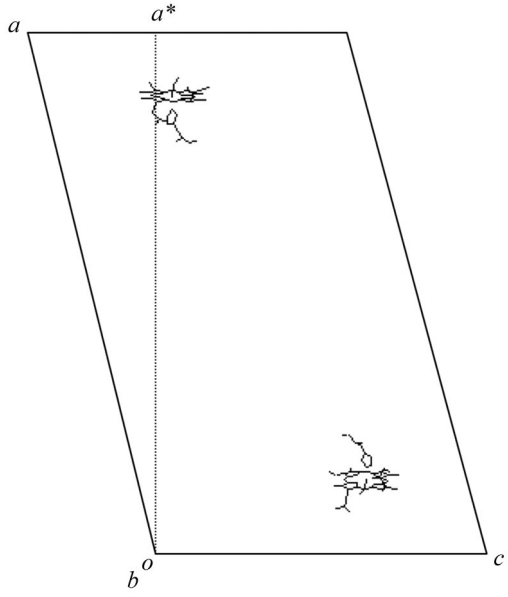

Figure 4
Pictorial view of the arrangement of the two Fe–heme sites (including the proximal histidine imidazole and the CO ligand) in the unit cell of a MbCO single crystal. The view is projected along the *b* axis of the monoclinic crystal.

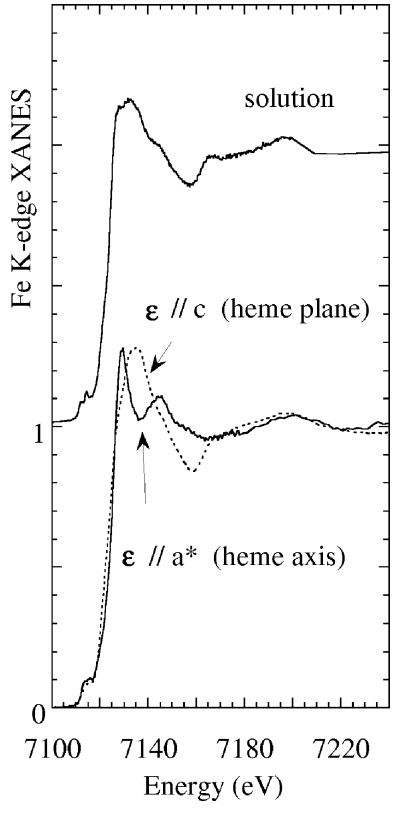

Figure 5
Fe *K*-edge XANES spectra of sperm-whale MbCO in solution (upper curve); polarized XANES spectra from an MbCO single crystal (lower frame).

criterion, and the exchange-correlation part of the potential was determined by the *X*-α approximation. The inelastic losses have been included using the method described previously. No statistical weight has been used here, so the square residual function S^2 coincides with the statistical χ^2 function. A constant experimental error, which corresponds to a noise-to-signal ratio of 0.01 in the normalized spectrum, was chosen.

Starting from a chosen configuration of the absorbing site, the *MXAN* procedure reached the best-fit conditions in reasonable time by the minimization of χ^2 in the space of selected structural parameters. The $\boldsymbol{\varepsilon}$//**c** spectra are used to fit the Fe—N_p distance, while the $\boldsymbol{\varepsilon}$//**a*** spectra are used to fit all the other parameters. The values of χ^2/n that are found at the best-fit conditions are 3.46, 3.08, 2.09 and 1.85, respectively, for MbCO $\boldsymbol{\varepsilon}$//**c**, Mb*CO $\boldsymbol{\varepsilon}$//**c**, MbCO $\boldsymbol{\varepsilon}$//**a*** and Mb*CO $\boldsymbol{\varepsilon}$//**a***. The corresponding structural determinations at the best-fit conditions for the MbCO and Mb*CO species are reported in Tables 1 and 2, respectively. Alignment of the theoretical and experimental spectra has been obtained with an energy shift of 7115.5 (0.3) eV. A value of $\Gamma_c = 2.7$ (0.1) eV results from the fit procedure, in good agreement with the sum of the known value of core-hole lifetime (1.2 eV) and experimental resolution, which is about 1.4 eV. This high value is due to the experimental conditions, *i.e.* the crystal being placed in a glass capillary and embedded in a helium atmosphere during the data-acquisition runs. The plasmon-like excitation that is included in the fit has an onset at 15 (4) eV and an amplitude of 10 (3) eV.

For brevity we report here only the calculations with the polarization vector of the light $\boldsymbol{\varepsilon}$ parallel to the heme normal. The $\boldsymbol{\varepsilon}$//**c** experimental data and calculations are reported by Della Longa *et al.* (2001). In Fig. 8 we show the comparisons between the experimental spectrum of MbCO, the calculation related to the *MXAN* best-fit structure and other calculations related to different structural determinations of the heme-site geometry obtained by X-ray diffraction. The values of the selected parameters are reported in Table 1. They include the low-resolution data by Kuriyan *et al.* (1986; PDB code 1MBC), the old EXAFS determination in solution by Powers *et al.* (1987), and two high-resolution X-ray structures measured by Kachalova *et al.* (1999) at room temperature (PDB code 1BZR) and by Vojtechovsky *et al.* (1999) at $T = 100$ K (PDB code 1A6G). These latter two structures are essentially identical, apart from the CO binding geometry. Our geometrical determination corresponding to the best-fit condition is in complete agreement with the results of Vojtechovsy *et al.* (1999), if we assume that a resolution of 1.15 Å in an X-ray diffraction determination should correspond to an uncertainty in the atomic position of about 0.05 Å. The small residual discrepancies between the XANES fit and experiments could be due to either an inhomogeneity of the protein ensemble or the approximations in the description of the inelastic losses in the framework of the phenomenological approach. The *MXAN* fit is highly sensitive to the first-neighbour distances, to the bending angle β and to the C—O bond length. The errors in the table are calculated by *MINUIT* on a statistical basis, taking into account all of the parameter correlations (but not the non-linearities). As an example, a contour plot pertaining to a three-parameter (Fe—C; Fe—N_{his}; β) *MXAN* analysis of the $\boldsymbol{\varepsilon}$//**a*** XANES spectrum of MbCO is depicted in Fig. 9. The fit includes the weakly correlated variables β and d(Fe—C). Other structural parameters, like the Fe-atom displacement from the heme plane and the heme doming, were considered in

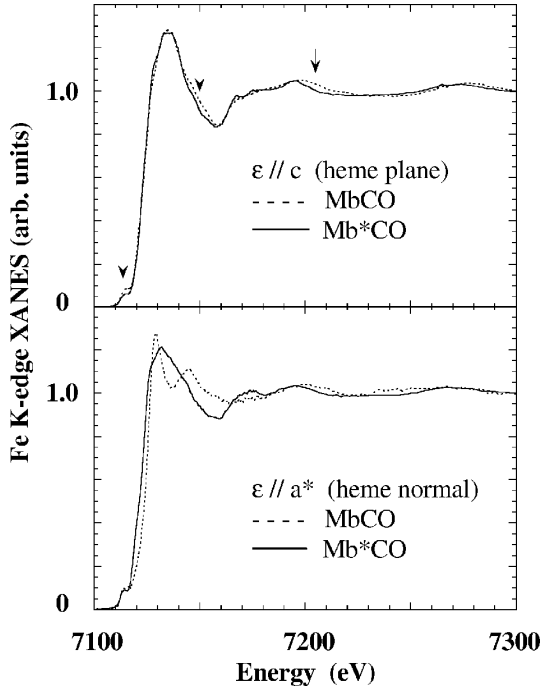

Figure 6
Comparison between the polarized XANES spectra of MbCO (dashed line) and Mb*CO (solid line). Arrows in the upper frame emphasize the differences between the $\boldsymbol{\varepsilon}$//**c** spectra.

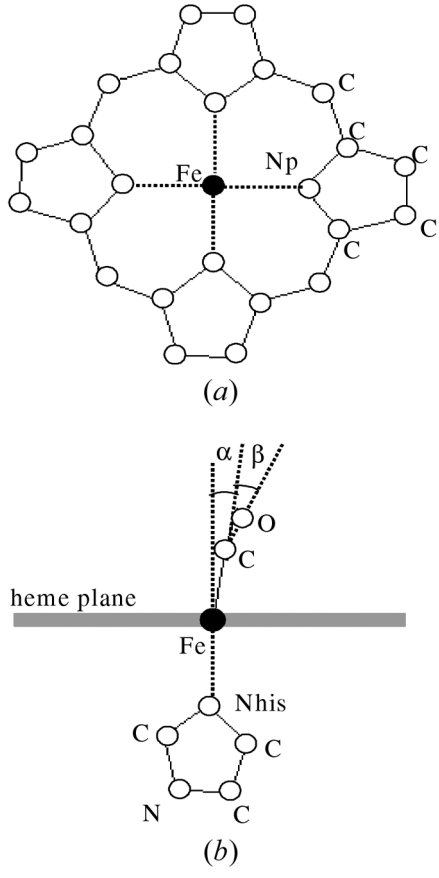

Figure 7
Pictorial view of the Fe site in MbCO. (*a*) Projection along the heme normal. (*b*) Projection along the heme plane. N_p indicates the pyrrolic N atoms of the porphyrin while N_{his} indicates the imidazole N atom of the histidine protein residue.

a preliminary analysis but were not used in the last run because the fit turned out not to be sensitive to them. The low sensitivity to the tilting angle α is explained by its low value in this case.

The Mb*CO species has been analysed using the same procedure and the same set of structural parameters as were used for MbCO. The values of the structural parameters at the best-fit conditions are reported in Table 2, together with the values that have been determined by X-ray diffraction data. These latter values are at lower resolution as a result of an uncertainty in the atomic position of about 0.2 Å. Our results are in rather good agreement with the X-ray diffraction data of Teng *et al.* (1997) (PDB code 1AJH) but differ more significantly from those of Schlichting *et al.* (1994) (PDB code 1ABS). The fit is sensitive to the distances between the Fe and pyrrolic N atoms and the ligated imidazole N atom of histidine. The fit is also highly sensitive to the distance and orientation of the unligated CO molecule, and therefore we can determine its distance with a statistical error of 0.07 Å and the angles α and β with an error of 7° and 5°, respectively. The fit is sensitive to the Fe—C tilting angle α in this case because of the higher value of the angle. Moreover, the C—O distance is still evaluated as accurately as it was in the MbCO case.

In Fig. 10 we report the comparison between the experimental data of Mb*CO, the calculations related to the geometry at the *MXAN* best-fit condition and the structural X-ray determinations reported in Table 2. A comparison of these curves with those depicted in Fig. 8 shows that the dramatic difference between MbCO and Mb*CO in the $\boldsymbol{\varepsilon}$//**a*** spectra, which is a result of the change in the coordination symmetry, is fully reproduced by the calculation at the best-fit conditions. Moreover, it is noticeable that small changes in the structural geometry, essentially the CO position and orientation, induce such relevant differences in the XANES calculation, even in the case of the unligated Mb*CO. This fact confirms that it is possible to use the fitting procedure of XANES data to access structural

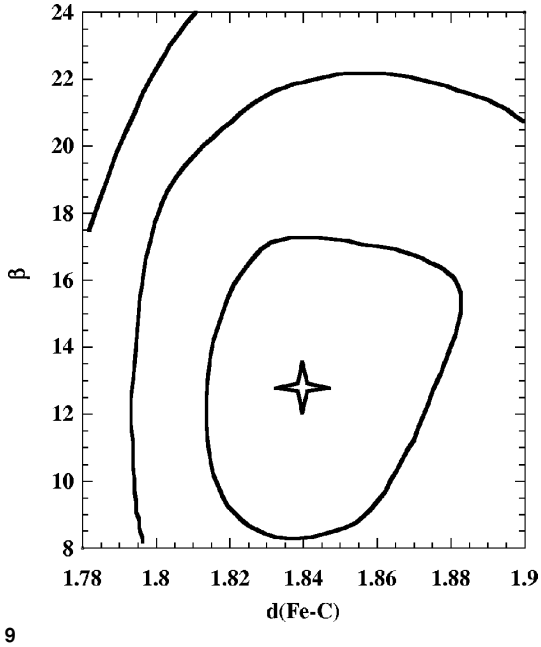

Figure 9
β versus d(Fe—C) contour plot related to a three-parameter (Fe—C; Fe—N_{his}; *β*) *MXAN* analysis of the $\boldsymbol{\varepsilon}$//**a*** XANES spectrum of MbCO. The distance is expressed in Å and the angle in degrees. The inner ellipsoid corresponds to the 1σ (95%) confidence level.

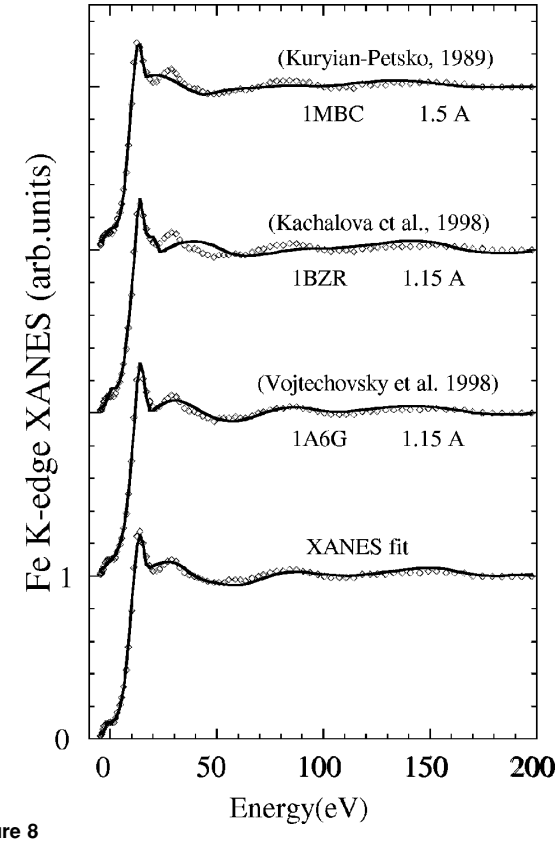

Figure 8
Comparison between the $\boldsymbol{\varepsilon}$//**a*** polarized experimental data of MbCO (circles) and the calculations (solid curves) derived from three different reported structures, *i.e.* by Kuriyan *et al.* (1986; PDB code 1MBC), by Kachalova *et al.* (1999; PDB code 1BZR) and by Vojtechovsky *et al.* (1999; PDB code: 1A6G). The bottom theoretical curve is the *MXAN* best fit.

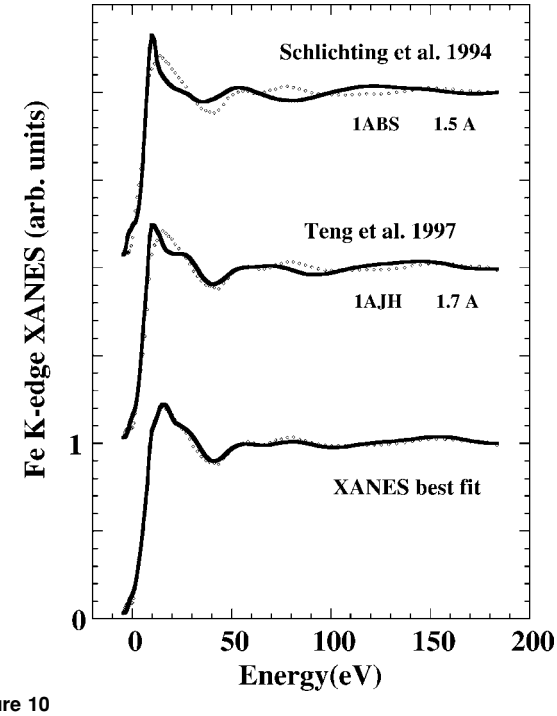

Figure 10
Comparison between the $\boldsymbol{\varepsilon}$//**a*** polarized experimental data of Mb*CO (circles) and the calculations (solid curves) derived from two different reported structures, *i.e.* that by Schlichting *et al.* (1994; PDB code 1ABS) and Teng *et al.* (1994; PDB code 1AJH). The bottom theoretical curve is the *MXAN* best fit.

details of the low-temperature photoproduct of MbCO with high precision. The sensitivity to the position of the CO molecule is particularly high in the XANES region because the scattering power of light elements is large and comparable with that of heavy atoms in the EXAFS energy range. At higher energy, *i.e.* in the EXAFS region, the signal from light atoms decreases so much that it is hard to obtain more information than the first-shell distances. The ability to perform a quantitative analysis of the XANES energy region allows us to determine some relevant structural parameters that are normally impossible to recover by the standard analysis of the EXAFS data.

5. Conclusions

The method that we present here seems to open the way to systematic XAS investigations of the time (over long time scales), temperature (4–60 K) and illumination dependence of important structural parameters in the dynamics of myoglobin. XAS quantitative results could complement other independent studies, such as those that follow events related to the migration of the ligand through the protein matrix. An accurate study of the thermal and light-induced changes in the polarized XAS and a comparison between polarized XAS of ligated and unligated Fe–heme states could be attempted at low temperatures and/or by fast XANES collection, in order to probe both the CO migration from the primary docking site and the relaxation at the proximal side. If heme relaxation and ligand rebinding are two distinct processes (Franzen & Boxer, 1997), polarized XAS promises to allow the observation, in the same experiment, of some of these complementary aspects of a protein dynamic event. XAS could be further applied to check the controversial hypothesis of a proximal control of the rebinding barrier, the measurement of which relies on changes in the position of the iron–histidine complex of the order of 0.1 Å or less. The effect of light on the Fe–heme structure could be investigated to assess the possible effects on the Fe—His bond (Ahmed *et al.*, 1991; Chu *et al.*, 1995).

In conclusion, we have presented a new method for a quantitative analysis of XANES spectra of inorganic metal compounds and metalloproteins at the metal *K*-edge based on a fitting procedure in the framework of full multiple-scattering calculations. In this way we achieve full access to the whole XANES energy range and hence a complete structural determination of the local geometry around the absorber. When applied to the MbCO low-temperature photolysis process, our analysis has allowed a complete determination of the iron–ligand geometry in MbCO and Mb*CO at a resolution comparable with that of recent X-ray diffraction studies.

References

Ahmed, A. M., Campbell, B. F., Caruso, D., Chance, M. R., Chavez, M. D., Courtney, S. H., Friedman, J. M., Iben, I. E. T., Ondrias, M. R. & Yang, M. (1991). *Chem. Phys.* **158**, 329–351.

Benfatto, M. & Della Longa, S. (2001). *J. Synchrotron Rad.* **8**, 1087–1094.

Bianconi, A., Congiu Castellano, A., Durham, P. J., Hasnain, S. S. & Phillips, S. (1985). *Nature* (*London*), **318**, 685–687.

Binsted, N. & Hasnain, S. S. (1996). *J. Synchrotron Rad.* **3**, 185–196.

Cabaret, D., Joly, Y., Renevier, H. & Natoli, C. R. (1999). *J. Synchrotron Rad.* **6**, 258–260.

Chu, K., Ernst, R. M., Frauenfelder, H., Mourant, J. R., Neinhaus, G. U. & Philipp, R. (1995). *Phys. Rev. Lett.* **74**, 2607–2610.

Chu, K., Vojtchovsky, J., McMahon, B. H., Sweet, R. M., Berendzen, J. & Schlichting, I. (2000). *Nature* (*London*), **403**, 921–923.

Cruickshank, D. W. J. (1999). *Acta Cryst.* D**55**, 583–601.

D'Angelo, P., Benfatto, M., Della Longa, S. & Pavel, N. V. (2002). *Phys. Rev B*, **66**, 064209-1–062409-7.

Della Longa, S., Arcovito, A., Girasole, M., Hazemann, J. L. & Benfatto, M. (2001). *Phys. Rev. Lett.* **87**, 155501.

Della Longa, S., Arcovito, A., Vallone, B., Congiu Castellano, A., Kahn, R., Vicat, J., Soldo, Y. & Hazemann, J. L. (1999). *J. Synchrotron Rad.* **6**, 1138–1147.

Della Longa, S., Pin, S., Cortes, R., Soldatov, A. & Alpert, B. (1998). *Biophys. J.* **75**, 3154–3162.

Díaz-Moreno, S., Muñoz-Páez, A. & Chaboy, J. (2000). *J. Phys. Chem. A*, **104**, 1278–1286.

Filipponi, A. & Di Cicco, A. (1995). *Phys. Rev. B*, **52**, 15135–15149.

Franzen, S. & Boxer, S. G. (1997). *J. Biol. Chem.* **272**, 9655–9660.

Hasnain, S. S. & Hodgson, K. O. (1999). *J. Synchrotron Rad.* **6**, 852–864.

Iizuka, T., Yamamoto, H., Kotani, M. & Yonetani, T. (1974). *Biochim. Biophys. Acta*, **371**, 126–139.

Joly, Y. (2001). *Phys. Rev. B*, **63**, 125120-1–125120-10.

Kachalova, G. S., Popov, A. N. & Bartunik, H. D. (1999). *Science*, **284**, 473–476.

Kirkpatrick, S., Gelatt, C. D. Jr & Vecchi, M. P. (1983). *Science*, **220**, 671–680.

Kuriyan, J., Wilz, S., Karplus, M. & Petsko, G. A. (1986). *J. Mol. Biol.* **192**, 133–154.

Meller, J. & Elber, R. (1998). *Biophys. J.* **74**, 789–802.

Natoli, C. R. & Benfatto, M. (1986). *J. Phys. Paris Colloq.* **47**, 11–23.

Natoli, C. R., Benfatto, M., Della Longa, S. & Hatada, K. (2003). *J. Synchrotron Rad.* **10**, 26–42.

Pedio, M., Benfatto, M., Aminpirooz, S. & Haase, J. (1994). *Phys. Rev. B*, **50**, 6596–6602.

Petsko, G. A. (1994). *Nature* (*London*), **371**, 740–741.

Powers, L., Chance, B., Chance, M., Campbell, B., Khalid, J., Kumar, C., Naqui, A., Reddy, K. S. & Zhou, Y. (1987). *Biochemistry*, **26**, 4785–4796.

Schlichting, I., Berendzen, J., Phillips, G. N. Jr & Sweet, R. M. (1994). *Nature* (*London*), **371**, 808–812.

Teng, T.-Y., Srajer, V. & Moffat, K. (1994). *Nature Struct. Biol.* **1**, 701–705.

Teng, T.-Y., Srajer, V. & Moffat, K. (1997). *Biochemistry*, **36**, 12087–12100.

Tyson, T. A., Hodgson, K. O., Natoli, C. R. & Benfatto, M. (1992). *Phys. Rev. B*, **46**, 5997–6019.

Vitkup, D., Petsko, G. A. & Karplus, M. (1997). *Nature Struct. Biol.* **4**, 202–208.

Vojtechovsky, J., Chu, K., Berendzen, J., Sweet, R. M. & Schlichting, I. (1999). *Biophys. J.* **77**, 2153–2174.

Received 19 February 2002 • Accepted 29 August 2002