

# Multiple-scattering XAS calculations for the characterization of the binuclear (type 3) copper sites of hemocyanins and related model compounds

Elena Borghi<sup>a\*</sup> and Pier Lorenzo Solari<sup>b</sup>

<sup>a</sup>Dipartimento di Chimica, Università La Sapienza, p. le A. Moro 5, I-00185 Roma, Italy, and  
<sup>b</sup>European Synchrotron Radiation Facility, 6 Rue Jules Horowitz, BP 220, F-38043 Grenoble, France. E-mail: e.borghi@caspur.it

Received 2 October 2003  
Accepted 5 November 2004

Our previous studies on the met- and met-azido-Hc forms from *Octopus vulgaris* (mollusc) and *Carcinus aestuarii* (arthropod) at pH 7.5 and on some related binuclear models without X-ray diffraction have considered and resolved some fundamental aspects of their binuclear type-3 copper site (*i.e.* correct values of the Cu–Cu distances; apical distortion, when present, at the copper site; presence and type of bridging groups). In this contribution the multiple-scattering (MS) calculations performed in order to refine the EXAFS modulation of the absorption spectra are presented. It is only with a composite and advanced approach that some severe problems, mainly deriving from the presence of two absorbing atoms and from the fact that the metal–metal contribution in the absorption spectra overlaps with the Cu–His signals, have been overcome. Results are also presented which indicate the role of the MS calculations in the XANES edge region, and which show how it is possible to extract quantitative information from this zone of the spectrum in order to refine the structure of the site also in the case of a binuclear centre.

© 2005 International Union of Crystallography  
Printed in Great Britain – all rights reserved

**Keywords:** XAS; EXAFS; XANES; hemocyanin.

## 1. Introduction

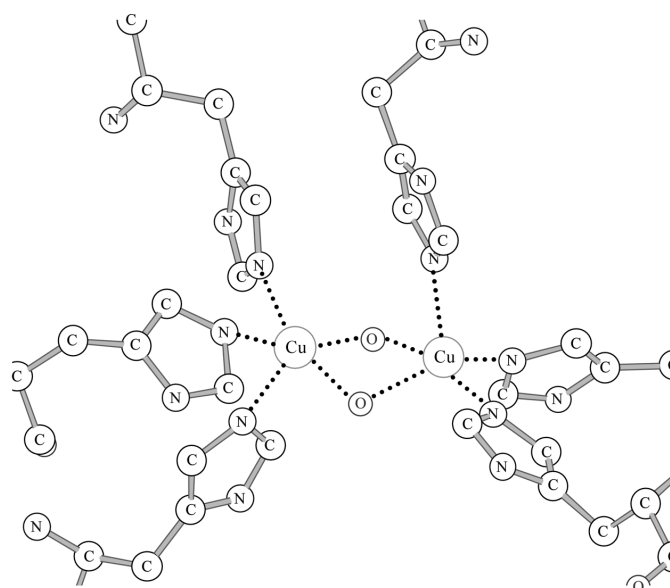
Type-3 copper proteins are an important class of metalloproteins characterized by a binuclear active site that is anti-ferromagnetically coupled and therefore EPR silent in the oxidized state (see Fig. 1). Three different members from this group are known: hemocyanins (Hcs), which are dioxygen transport proteins, tyrosinases (Tyr) and catechol oxidases (COs), which are oxygenase/oxidase enzymes.

The large dimensions of the molecules and/or the difficulties in their isolation have hampered their structural characterization with NMR spectroscopy and with X-ray diffraction (XRD). As a result, up to now only the crystal structures of four Hcs and one CO have been resolved (Volbeda & Hol, 1989; Magnus *et al.*, 1994; Cuff *et al.*, 1998; Klabunde *et al.*, 1998; Perbandt *et al.*, 2003).

X-ray absorption spectroscopy (XAS), owing to its ability to selectively probe the local structure of the metallic centres in materials without constraints on the physical state of the samples under investigation, was used very early to characterize the active site of type-3 copper proteins and related model compounds, particularly in the extended X-ray absorption fine structure (EXAFS) region (Brown *et al.*, 1980; Co *et al.*, 1981; Co & Hodgson, 1981; Woolery *et al.*, 1984*a,b*).

However, most of these early investigations were carried out in the frame of the single scattering approximation, which

has proven to be erroneous, especially in the case of active sites surrounded by histidines. Furthermore, the metal–metal distance in the binuclear copper site has been difficult to



**Figure 1**  
Structure of the active site of met-Hc from *Limulus polyphemus*, adapted from the PDB code 1LL1 (S. Liu, H. Ton-That & K. A. Magnus, unpublished).

characterize, since the Cu–Cu contribution is usually superposed with the contributions coming from the external shells of the histidines or other surrounding amino acids. Consequently, some inconsistencies between EXAFS and XRD results have appeared over the years.

The recent developments of new analysis approaches based on the spherical-wave multiple-scattering theory (MS) allow for an improved reliability in EXAFS studies, and indeed, nowadays, several advanced programs are able to satisfactorily analyse this region of the absorption spectra.

The low-energy region of the XAS spectra, the X-ray absorption near-edge spectroscopy (XANES) region, has also been used, although to a lesser extent, to characterize the active site of type-3 copper proteins and related model compounds. In this case, the quantitative or semi-quantitative analysis of this part of the spectra has become straightforward and possible; this only after that many structural investigations (Kau *et al.*, 1987; Sano *et al.*, 1992; Shadle *et al.*, 1993; Pickering & George, 1995) have correlated the position and the intensity of the characteristic features of the edge region for copper complexes with the general geometry and symmetry of the metal site, and that different software programs have been developed to calculate the theoretical XANES signal on the basis of full MS formalism (Durham, 1988). However, only recently with a new software procedure, *MXAN* (Benfatto & Della Longa, 2001), has it also become possible to carry out a full geometrical fitting for this region, in a reasonable time.

The ability to simulate or analyse this part of the spectra is of great importance, not only in the attempt to confirm or to replace the analysis of the EXAFS region but also, as we will show in the following, in the case of binuclear type-3 copper sites, in order to refine the structure of the sample and to discern the geometrical differences between two coordinating metal centres of the same chemical species.

Recently, we have reconsidered and resolved some fundamental aspects of the binuclear copper site for met- and met-azido derivatives of *Octopus vulgaris* (mollusc) and *Carcinus aestuarii* (arthropod) Hcs at pH 7.5 (Borghi *et al.*, 2002). The study described therein is part of a larger project in collaboration with B. Salvato's group in Padova and L. Casella's group in Pavia, carried out at the ESRF facility, which has also considered the XAS characterization of the hemocyanin derivatives from the same *phyla* at pH 5.5 and some related binuclear models without XRD.

In that paper we have reported results from the EXAFS single-scattering Fourier filter method for the first shell and the partial results from the MS EXAFS analysis of the whole spectra. A qualitative XANES approach has also been reported to confirm the analysis of the EXAFS region. The accuracy of the analysis of the data has been tested with mononuclear- and binuclear-related model compounds of the poly(benzimidazole) ligand system 2-BB (Casella *et al.*, 1996), L-55 and L-6,6 (Casella *et al.*, 1993).

The biological problem connected to the structural characterization of these Hc derivatives and the biophysical implications of the XAS results obtained are illustrated in that

study.<sup>1</sup> In this paper we will only present and focus attention on the MS calculations that have been used to achieve the EXAFS results, and on the possibilities opened by the MS calculations of the XANES region.

## 2. EXAFS analysis: results and discussion

In the following we will describe in detail the EXAFS theoretical calculations that we have considered to achieve good fits of the experimental spectra of the met-Hc forms.

The analysis has been performed using the *GNXAS* set of programs (Filipponi *et al.*, 1995; Filipponi & Di Cicco, 1995, 2000). This code, which takes full account of the MS effects, is particularly suited for analysing disordered systems and has been successfully applied to investigate biological or biological-related samples [see Di Cicco (2003), and references therein].

As in other modern programs for EXAFS analysis, in *GNXAS* a theoretical signal is calculated *ab initio* from a model structure and is subsequently refined against the experimental signal in order to determine the structural parameters of the sample. In our study, the first geometrical model that has been considered in order to calculate the theoretical signals for the met-Hcs was based on the crystallographic structure of the  $[\text{Cu}_2(\text{L-5,5})(\text{OMe})_2](\text{ClO}_4)_2$  complex (Battaini *et al.*, 2003) and on the EXAFS results from the analysis of the experimental spectra of the  $[\text{Cu}_2(\text{L-5,5})(\text{OH})_2](\text{ClO}_4)_2$  and the  $[\text{Cu}_2(\text{L-6,6})(\text{OH})_2](\text{ClO}_4)_2$  complexes without crystallographic structures (Casella *et al.*, 1993; Borghi *et al.*, 2002; E. Borghi, L. Casella & P. L. Solari, unpublished results). Subsequently, a second geometrical model based on the crystallographic structure of the met- form of *Limulus polyphemus* Hc (S. Liu, H. Ton-That & K. A. Magnus, unpublished; PDB code 1LL1) has been considered (see Fig. 1). The quality of the fits, starting from the two different theoretical structures, was comparable. Nonetheless, smaller residual values were obtained for the calculations starting from the 1LL1 structure. Key distances and angles associated with the 1LL1 crystallographic structure are reported subsequently in Table 1.

In *GNXAS*, differently from other advanced analysis codes, the irreducible *n*-body expansion is used to calculate the theoretical absorption coefficient (Filipponi *et al.*, 1995). This means that the total theoretical EXAFS  $\chi(k)$  signal is not decomposed as a sum of signals related to the number of scattering events [*i.e.*  $\chi(2)$ ,  $\chi(3)$ ,  $\chi(4)$ ] but as a sum of signals related to the different two-body, three-body and four-body subgeometries inside the system of atoms [*i.e.*  $\gamma(2)$ ,  $\gamma(3)$  and  $\gamma(4)$ ]. Several structural parameters are associated with the different *n*-body signals [*i.e.* the multiplicity *N*, the distances *R*,

<sup>1</sup> The XAS experiments have been performed at the European Synchrotron Radiation Facility (ESRF) on the General-purpose Italian beam Line for Diffraction and Absorption (GILDA). The measurements were carried out in the fluorescence mode with a high-purity Ge 13-element detector. More details on the sample preparation and on the experimental set-up (including the requirements to prevent the radiation damage) are described by Borghi *et al.* (2002) and Solari (2002), respectively.

**Table 1**

 Results obtained for the met-Hc of *Octopus vulgaris* and *Carcinus aestuarii* starting from the 1LL1 structure.

Signal	Structural feature × multiplicity	met-Hc of <i>O. vulgaris</i>		met-Hc of <i>C. aestuarii</i>		XRD (1LL1)
		Distance or angle	$\sigma_{\text{dist}}^2$ or $\sigma_{\text{ang}}^2$	Distance or angle	$\sigma_{\text{dist}}^2$ or $\sigma_{\text{ang}}^2$	Distance or angle
$\gamma(2)$	Cu–O × 2	1.96 Å	0.004 Å <sup>2</sup>	1.91 Å	0.006 Å <sup>2</sup>	2.012 Å
	Cu–N <sub>1</sub> × 2	1.97 Å	0.005 Å <sup>2</sup>	1.97 Å	0.001 Å <sup>2</sup>	2.028 Å
	Cu–N' × 1	2.34 Å	0.010 Å <sup>2</sup>	2.41 Å	0.009 Å <sup>2</sup>	2.405 Å
$\gamma(3)$	Cu–Cu × 1	3.01 Å	0.005 Å <sup>2</sup>	3.00 Å	0.011 Å <sup>2</sup>	3.127 Å
	Cu–O–Cu × 2	100°	7° <sup>2</sup>	103°	9° <sup>2</sup>	102.1°
	Cu–C <sub>1</sub> /C <sub>2</sub> × 4	3.06 Å	0.014 Å <sup>2</sup>	3.04 Å	0.011 Å <sup>2</sup>	2.982 Å
	Cu–N <sub>1</sub> –C <sub>1</sub> /C <sub>2</sub> × 4	136°	6° <sup>2</sup>	132°	9° <sup>2</sup>	124.8°
	N <sub>1</sub> –C <sub>1</sub> /C <sub>2</sub> × 4	1.32 Å	0.011 Å <sup>2</sup>	1.35 Å	0.012 Å <sup>2</sup>	1.348 Å
	Cu–C <sub>1</sub> '/C <sub>2</sub> ' × 2	3.37 Å	0.015 Å <sup>2</sup>	3.40 Å	0.011 Å <sup>2</sup>	3.347 Å
	Cu–N'–C <sub>1</sub> '/C <sub>2</sub> ' × 2	127°	2° <sup>2</sup>	125°	1° <sup>2</sup>	124.3°
	N'–C <sub>1</sub> '/C <sub>2</sub> ' × 2	1.40 Å	0.009 Å <sup>2</sup>	1.39 Å	0.003 Å <sup>2</sup>	1.346 Å
$\gamma(4)$	Cu–N <sub>2</sub> /C <sub>3</sub> × 6	4.14 Å	0.014 Å <sup>2</sup>	4.11 Å	0.012 Å <sup>2</sup>	4.219 Å
	Cu–N <sub>1</sub> –C <sub>1</sub> /C <sub>2</sub> –N <sub>2</sub> /C <sub>3</sub> × 6	180°	27° <sup>2</sup>	178°	8° <sup>2</sup>	166.8°
	N <sub>1</sub> –C <sub>1</sub> /C <sub>2</sub> –N <sub>2</sub> /C <sub>3</sub> × 6	118°	16° <sup>2</sup>	111°	12° <sup>2</sup>	107.7°
	C <sub>1</sub> /C <sub>2</sub> –N <sub>2</sub> /C <sub>3</sub> × 6	1.25 Å	0.001 Å <sup>2</sup>	1.29 Å	0.003 Å <sup>2</sup>	1.337 Å
Other $\gamma(2)$	Cu–N/C <sup>+</sup> × 0.6	4.00 Å	0.010 Å <sup>2</sup>	4.00 Å	0.010 Å <sup>2</sup>	–
	Cu–N <sub>x</sub> × 3	5.00 Å	0.011 Å <sup>2</sup>	4.99 Å	0.011 Å <sup>2</sup>	4.518 Å
Residual, experiment (theory)		1.44 × 10 <sup>−6</sup> (1.20 × 10 <sup>−6</sup> )		1.89 × 10 <sup>−6</sup> (1.74 × 10 <sup>−6</sup> )		

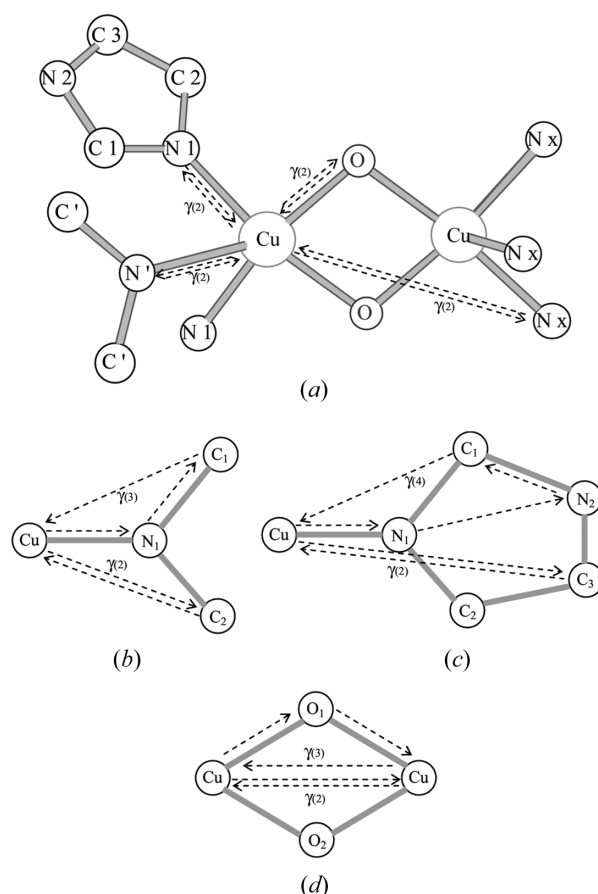
the angles  $\theta$ , the dihedral angles  $\psi$  and their associated Debye–Waller (DW) factors,  $\sigma^2$ ] and can be obtained with best fits [for more details see, for example, Di Cicco (2003)].

The data analysis of these type-3 copper proteins is made particularly delicate by the concomitant presence of two inequivalent copper sites contributing to the absorption signal. The GNXAS analysis package is particularly suitable in this respect since, given an atomic cluster, the program finds all inequivalent absorbers, calculates the associate signals and compares their sum to the experimental signal. The ability of the method, even in the multi-absorber case, in reproducing known structures in test cases has been amply documented in the literature.

In our case, considering that the structural differences for the neighbourhood of the two absorbing copper centres can be expected small, average values of the structural parameters related to the two copper centres have been taken into account in the calculations of the theoretical signals and subsequent fits. Therefore, in reality, all the paths described in the following with respect to the Cu1 centre are also considered with respect to the Cu2 centre.

In our study, the signals considered for the first shell of atoms surrounding the absorber were, respectively, a two-body signal  $\gamma(2)_{\text{Cu-O}}$  with a multiplicity 2, to take into account the single and multiple scattering coming from the two O atoms belonging to the hypothesized hydroxo groups, and a two-body signal  $\gamma(2)_{\text{Cu-N}_1}$  with a multiplicity 2, to take into account the single and the multiple scattering coming from the two N<sub>1</sub> atoms belonging to the equatorial imidazole rings. In addition, in order to evidence the presence of an apical distortion, a distinct two-body signal  $\gamma(2)_{\text{Cu-N}'}$  with a multiplicity 1 was considered, to take into account the single and the multiple scattering coming from the N' atom belonging to the axial imidazole ring (see Fig. 2a).

For the second shell of atoms surrounding the copper centre, a separation was again made between the signals


**Figure 2**

Schematic representation of the scattering paths considered in the EXAFS fit.

coming from the equatorial imidazoles and those coming from the axial imidazole. Thus, for both groups, the effective three-body scattering signal  $\eta(3)$ , related to the Cu–N<sub>1</sub>–C<sub>1</sub>/C<sub>2</sub> or to the Cu–N'–C<sub>1</sub>'/C<sub>2</sub>' scattering paths, consisted of a two-body

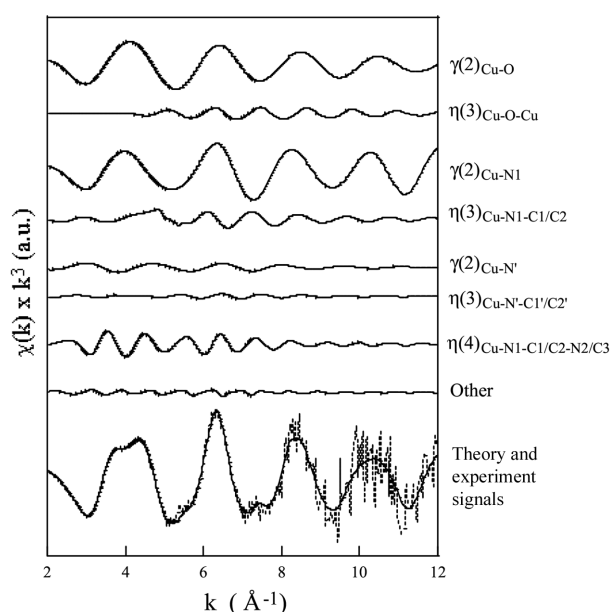
long-bond signal  $\gamma(2)$  and a three-body  $\gamma(3)$  signal (see Fig. 2b). To reduce the number of variables, no distinctions were made between the left side scattering paths and the right side paths of the imidazole ring, so the multiplicity of each path was just doubled. In the best fits obtained for the model complexes, the distinction of the  $\gamma(2)$  and  $\eta(3)$  signals for the axial imidazole proved to be feasible (E. Borghi & P. L. Solari, unpublished results), so it seemed to us possible to distinguish these contributions also for the proteins.

In addition to these signals coming from the imidazoles, in the second shell we have also considered an effective three-body scattering signal  $\eta(3)$ , consisting of a two-body (long-bond) and a three-body signal (with double multiplicity, one for each side) taking into account the copper–copper contribution through the two O atoms (*i.e.* related to the three-body Cu–O–Cu geometry) (see Fig. 2d).

For the external shell of the imidazole rings, a unique effective four-body  $\eta(4)$  signal (with multiplicity 6), related to the Cu–N<sub>1</sub>–C<sub>1</sub>/C<sub>2</sub>–N<sub>2</sub>/C<sub>3</sub> and Cu–N'<sub>1</sub>–C<sub>1</sub>'/C<sub>2</sub>'–N<sub>2</sub>'/C<sub>3</sub>' (for the axial position) scattering paths, was considered (see Fig. 2c). This consisted of a two-body long-bond signal, in two three-body signals, and a four-body signal.

Finally, two additional two-body signals,  $\gamma(2)_{\text{Cu-N/C}}$  and  $\gamma(2)_{\text{Cu-N}}$ , were added between 3.5 Å and 4.5 Å to take into account possible C or N atoms coming from other amino acidic residues and in particular from the N atoms which are coordinated to the opposite Cu atom (see Fig. 2a).

The use of the MS approach, particularly in association with the *n*-body expansion, and the concomitant fit over the whole EXAFS range, improves the refinement of the data because of the improved correctness of the theoretical calculation and in particular because it enables a natural constrain of the structural parameters that are varied during the fit.



**Figure 3**  
Detail of the theoretical signals considered for the final fit of the spectrum of met-Hc from *Carcinus aestuarii*.

Indeed, the inclusion of the MS signals in the refinements constrains the parameters of each *n*-body path to be linked through a well defined geometry. Thus, if some of these parameters are already linked through chemical groups constrains, such as those found in amino acidic residues, this may lead to a decrease in the number of degrees of freedom of the fitting function.

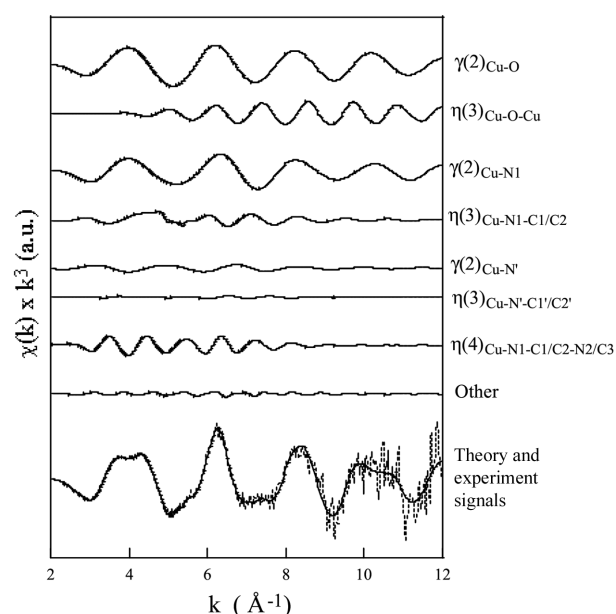
In our study the use of the MS signals enables us, for example, to constrain the copper–copper distance and the copper–oxygen shell distance between them, and it is ultimately what enables us to distinguish between the oxygen and the nitrogen donors in the first coordination shell.

The results of the fit for the EXAFS signals, including details of the different theoretical signals used, are shown in Figs. 3 and 4.

The refined results, including the previously unpublished values of the angular parameters, for the met-form of *O. vulgaris* and *C. aestuarii*, obtained by taking as a starting structure the XRD data of the met-form of *Limulus polyphemus* (1LL1), are presented in Table 1. The associated errors, as estimated from standard statistical analysis, are of the order of 0.02 Å for inner-shell distances and of the order of 0.1 Å for outer-shell distances.

The results obtained in our fits indicate that, for both species, the first shell of the met-Hcs is five coordinated, with two O atoms and three N atoms, as expected from the hypothesized structure. The Cu–O distance appears to be longer in the case of the met-Hc of *O. vulgaris* with respect to that of *C. aestuarii*, while for the equatorial Cu–N<sub>1</sub> contribution the distance is comparable.

The inclusion of a specific signal related to the apical N' atom is positively accounted for in the fit. For both species a distance of about 2.4 Å is found. While this value is slightly



**Figure 4**  
Detail of the theoretical signals considered for the final fit of the spectrum of met-Hc from *Octopus vulgaris*.

smaller than the average value (*i.e.* 2.5 Å) observed in the L-5,5-(OMe)<sub>2</sub> structure, it is in better agreement with the trend that has been evinced qualitatively from the XANES spectra (Borghi *et al.*, 2002) and with the average value from the 1LL1 structure. The high value of the associated DW factor denotes the high degree of disorder that is combined with this coordination atom.

In the second shell, the Cu–Cu contribution seems to be at an average distance of about 3.0 Å for both met-Hc species. The value of the average angle associated with the Cu–O–Cu paths is about 100°, in rough agreement with the average value obtained in the crystallographic 1LL1 structure. These values are also close to those obtained in previous crystallographic and EXAFS studies of hydroxo binuclear complexes (Blackburn *et al.*, 1988; Zippel *et al.*, 1996).

The contributions coming from the two external shells of the imidazole rings are refined to the expected distances. For the second shell, the splitting between the equatorial imidazoles' contribution and the apical imidazole contribution leads to a pronounced separation in their distances in line with the values coming from the crystallographic 1LL1 structure. For the contribution coming from the third shell of the imidazole ring, the refinement leads to a correct distance value of about 4.1 Å.

The Cu–N/C<sup>+</sup> contribution (which was associated with a variable coordination number) gives in this case a small contribution.

### 3. XANES simulations: results and discussion

In our paper (Borghi *et al.*, 2002) we have already discussed in detail how a qualitative analysis of the XANES region of our spectra can provide some general information on the geometric properties of the binuclear copper(II) sites of the considered samples. We have seen that all the spectra show the P feature (related to the 'dipole-forbidden'  $1s \rightarrow 3d$  transition) occurring in the pre-edge region. The position and the intensity of this feature mark the copper valence state as Cu(II) and the presence of a non-centrosymmetric cluster with N/O donors (Sano *et al.*, 1992). The A, B, C, D features (and their related  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  peaks) typical of dipole-allowed  $1s \rightarrow 4p$  transitions occur in the rising-edge region (A/ $\alpha$  and B/ $\beta$ ) and in the main-peak region (C/ $\gamma$  and D/ $\delta$ ), and their shapes appear complicated. In the higher-energy region the presence of the E/ $\epsilon$  and F/ $\phi$  features with observable intensity is an indicator of an anisotropic character in the symmetry of some sites. These features were clearly evident in the spectra of the L-5,5 complexes' family, but in the other spectra (proteins and models) could not be accurately estimated by a qualitative approach (for more details readers are referred to the cited article).

Generally it is difficult to rationalize the XANES region with only a qualitative analysis, given that small differences in the geometry of coordination (*i.e.* variation in bonding angle or movement of axial ligand around the metal centre) may produce important changes in the shape and in the intensity of the edge features and that, in our case, the XAS spectra of

these proteins contain structural information coming from the two copper ions. Hence, we have started to calculate the XAS signals in the edge region in order to perform quantitative simulations of the experimental absorption spectra.

Previously, some MS simulations of the XANES region have already been performed using the *G4XANES* package (Li *et al.*, 1991) for the *Carcinus maenas*-Hc (Della Longa *et al.*, 1993) and for the *Agaricus bisporus*-Tyr proteins (Della Longa *et al.*, 1996). We point out that these simulations on Hc-forms have assigned the polarization dependence of the XANES features to its metal centre (*i.e.* peaks A and B to a final state  $4p_z$ , C and D peaks to the  $4p_{xy}$  state). Moreover, they have shown that a longer bond distance of the apical ligand is correlated to a red shift of the B/ $\beta$  peak and to a variation of the intensity ratio of the A/ $\alpha$  and B/ $\beta$  peaks. So, the  $d$  (Å) versus  $E - E_0$  (eV) correlation for the B/ $\beta$  peak, where  $E_0$  is fixed as usual at 8979.0 eV, has been used by us as a qualitative criterion to estimate differences in the axial symmetry for the Hcs derivatives and for the related model compounds (refer to Borghi *et al.*, 2002).

For the simulations of the XANES region we have used the *CONTINUUM* set of programs developed at the INFN Laboratori Nazionali di Frascati on the basis of the multiple-scattering formalism (Tyson *et al.*, 1992). Using a known model structure the Cu *K*-edge total absorption coefficient is computed. As the previous calculations for the Hc protein have shown, the size of the scattering cluster is a key factor in the XANES simulations: the A, B and C peaks arise from the first shell and the D peak reflects the whole geometry of the copper site. Usually for an appreciable intensity of all the features we needed at least a three-shell cluster (including usually the three imidazole rings around the copper site). The D feature (usually) shows an intensity that strongly depends on the number of shells considered.

The Coulomb part of the potential is built by using the atomic charge densities from the Clementi–Roetti tables (Clementi & Roetti, 1974). The exchange correlation part of the potential is calculated using the complex Hedin–Lundqvist potential (Hedin & Lundqvist, 1971) with the  $Z + 1$  approximation for the core-hole relaxation. Muffin-tin radii are chosen according to the Norman criterion (Norman, 1976) with a 10% overlap allowed between the contiguous spheres. The calculated spectra are further convoluted with a Lorentzian curve for the core-hole lifetime (1.55 eV for the Cu *K* edge) and the energy resolution of the monochromator [0.3 eV for Si(311) crystal]. We would like to point out that the muffin-tin atomic potentials do not allow accurate quantitative calculations of the near-edge structures (Joly *et al.*, 1999, and references therein), while for higher photoelectron energies (from 20 eV above the edge) the muffin-tin approximation can be safely used. The imaginary part of the Hedin–Lundqvist potential, which controls the damping of the excited photoelectron, performs badly in the near-edge region of the spectrum. For this reason the sensitivity to structural parameters is reduced, so all features in the spectra are correctly reproduced in energy position, although their intensities are either overestimated or underestimated.

Here we will show how, following MS theoretical calculations with different model clusters taking into account structural differences, it is possible to reproduce and to differentiate the features of each Cu(II) site, and to simulate the experimental spectrum of a binuclear site.

As a first step in our study, we have started with the XANES simulations of the mononuclear complexes of the 2-BB ligand [*i.e.* the cationic complexes with XRD data  $[\text{Cu}(2\text{-BB})(\text{N}_3)]^+$ ,  $[\text{Cu}(2\text{-BB})(\text{MeOH})(\text{ClO}_4)]^+$ ,  $[\text{Cu}(2\text{-BB})(\text{NO}_2)]^+$  (Casella *et al.*, 1996). In the three complexes the ligand coordinates the Cu(II) ion, through its imidazole (N1 and N3) and amino (N5) nitrogen atoms, with different coordination numbers and geometries. However, the quite high flexibility of the 2-BB ligand makes the distances and the bond angles in the first shell very close, while the N1–Cu–N3 angle changes, reflecting the different symmetries [*i.e.*  $150.5^\circ$  for the four-coordinated pseudo tetrahedral 2-BB-N<sub>3</sub>;  $171.6^\circ$  for the five-coordinated 2BB-(MeOH)(ClO<sub>4</sub>) with a compressed trigonal bipyramidal geometry;  $156.8^\circ$  for the five-coordinated 2BB-NO<sub>2</sub> with a long (2.438 Å) Cu–O bond and a quite square planar stereochemistry].

These theoretical calculations have been performed in order to understand if, using *CONTINUUM*, it is possible to exemplify the fine structural differences of this family of compounds, in which the bond distances and angles of the 2-BB ligand do not feel the variation of stereochemistry at the metal centre. This may also give important indications for the study of the Hc proteins, since the geometrical changes of the 2-BB ligand may model the flexibility of the protein matrix in accomplishing its biological activity.

The results of these calculations, which are presented in Fig. 5, indeed show that the fine structural differences between the compounds produce changes in the shape of the XANES region and that using *CONTINUUM* it is possible to describe them.

In order to simulate with MS calculations the experimental spectra of the Hcs, we have started by focusing our attention on the structure of the oxy-form of *Limulus polyphemus* Hc by Magnus *et al.* (1994, PDB code 1OXY) which shows large differences for the structural parameters of the first shells of

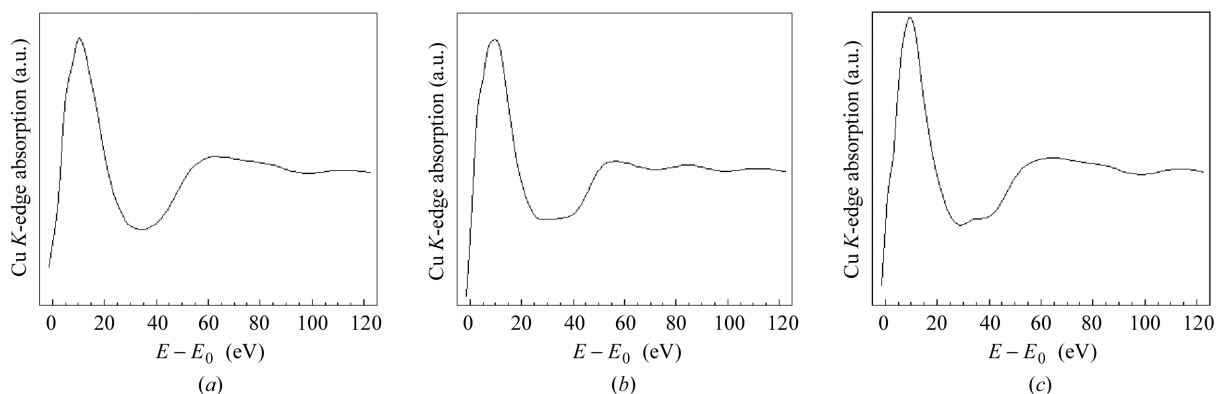
two Cu(II) ions that are present in the binuclear type-3 site. The local cluster for Cu1 is made up of a first shell with two O atoms at 2.071 and 2.175 Å, as well as three N atoms at 1.917, 2.097 and 2.445 Å. The local cluster for Cu2 shows a first shell with two O atoms at 1.716 and 1.956 Å, as well as three N atoms at 2.078, 2.227 and 2.296 Å. The Cu1–Cu2 distance is 3.594 Å.

So, starting from this crystallographic structure we have performed two different calculations with the two different model clusters of the two Cu(II) centres. In each simulation we have considered subclusters of atoms of increasing size, from one to three shells.

The results of these simulations on the Cu1 and Cu2 sites are reported in Fig. 6. From Figs. 6(a) and 6(b) it appears that at least three shells are necessary to obtain the expected features with an appreciable intensity. Fig. 6(c) shows the three-shell simulation for the two sites (Cu1 and Cu2) and their sum (Cu1 + Cu2). Fig. 6(d) shows a comparison between the experimental spectrum of the oxy-Hc from *Carcinus aestuarii* and the (Cu1 + Cu2) three-shell simulation of 1OXY (PDB code).

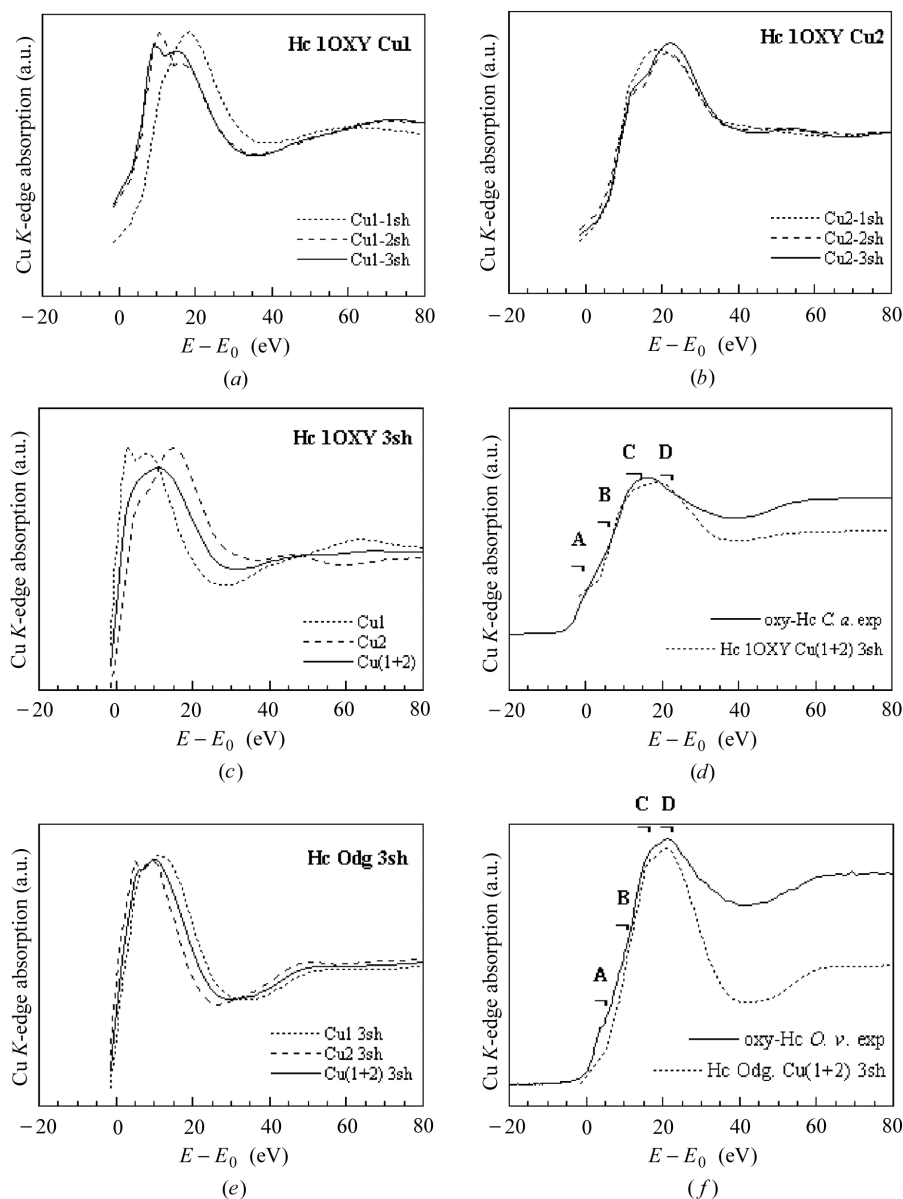
The results of these calculations demonstrate the strength of the theoretical approach using *CONTINUUM* for this copper 3 site. In each simulation the main features are correctly reproduced at the correct energy position and it is possible to differentiate them very well. It is a clear indication that, by means of MS calculations starting from suitable model clusters, the corresponding features of two different sites may be reproduced and differentiated.

From Fig. 6(d) we see a good agreement between the experimental spectrum of the oxy-Hc from *C. aestuarii* and the theoretical (Cu1 + Cu2) signal. The main features, A to D, of oxy-Hc from *C. aestuarii* are correctly reproduced in position, but the intensities are overestimated or underestimated. These results, confirming the validity of our approach using the *CONTINUUM* code, indicate that the simulations have to be improved by using alternative geometries of the two copper sites in order to account for the residual differences between the oxy-Hc of *C. aestuarii* and the oxy-Hc of *L. polyphemus*.



**Figure 5**

Simulations of the Cu *K*-edge absorption of the mononuclear complexes of the 2-BB ligand with different coordination geometries. Clusters of atoms corresponding to the entire cationic complexes have been considered. (a)  $[\text{Cu}(2\text{-BB})(\text{N}_3)]^+$  cation complex, (b)  $[\text{Cu}(2\text{-BB})(\text{MeOH})(\text{ClO}_4)]^+$  cation complex, (c)  $[\text{Cu}(2\text{-BB})(\text{NO}_2)]^+$  cation complex.


**Figure 6**

Simulation of the Cu  $K$ -edge absorption of the PDB code 1OXY. (a) Cu1 site with one- to three-shell model clusters. (b) Cu2 site with one- to three-shell model clusters. (c) Comparison of the three-shell simulations of Cu1, Cu2 and (Cu1 + Cu2) sites. (d) Comparison between the experimental spectrum of oxy-Hc of *C. aestuarii* and the (Cu1 + Cu2) three-shell simulation. Simulation of the Cu  $K$ -edge absorption of the PDB code 1JS8, fragment Odg. (e) Comparison of the three-shell simulations of Cu1, Cu2 and (Cu1 + Cu2) sites. (f) Comparison between the experimental spectrum of oxy-Hc of *O. vulgaris* and the (Cu1 + Cu2) three-shell simulation of Hc Odg.

A similar analysis was performed on the Cu  $K$ -edge of the oxy-Hc from *Octopus vulgaris* starting from the XRD data of the PDB code 1JS8, *i.e.* the oxy-form of a functional unit of Hc from *Octopus dofleini*, the fragment Odg (Cuff *et al.*, 1998). We show in Fig. 6(e) the third-shell simulation for the two sites (Cu1 and Cu2) and their sum (Cu1 + Cu2), and in Fig. 6(f) a comparison of the experimental data and the simulated XANES spectra based on a three-shell cluster. In this case the main features are also correctly reproduced.

The met-Hc forms of *Octopus vulgaris* and *Carcinus aestuarii* at pH 7.5 show fine differences in the XANES region

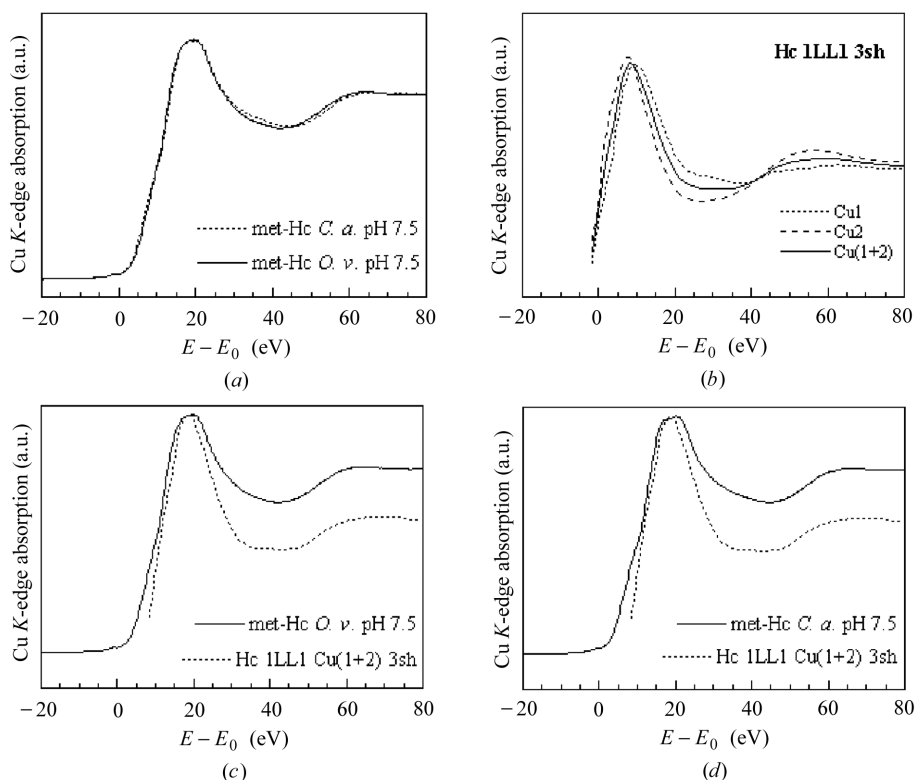
[refer to Borghi *et al.* (2002) and see Fig. 7(a)]. In order to simulate the experimental XANES spectra of these met-Hc forms we have considered, as for the MS EXAFS analysis, the crystallographic structure of the met-Hc form from *Limulus polyphemus* (S. Liu, H. Ton-That & K. A. Magnus, unpublished; PDB code 1LL1). This met-form of arthropod is characterized by a structure with a Cu–Cu distance of 3.13 Å and by a bis( $\mu$ -O-atom) bridge with different Cu–O–Cu angles (97.5° and 106.6°) inside the Cu<sub>2</sub>O<sub>2</sub> core. The additional anisotropic character of the symmetry of the Cu1 site (*i.e.* a short Cu–O distance and a long Cu–N axial distance) and the differences for the structural parameters of the first shells of two Cu(II) ions are shown in Table 2.

We have performed two different calculations with the two different structural models of the Cu1 and Cu2 sites as a function of cluster size (one to three shells), and then we have considered the sum of the two three-shell simulations, which is shown in Fig. 7(b) in comparison with the two different three-shell calculations. Figs. 7(c) and 7(d) show comparisons between the simulated signal and the experimental spectrum of the met-Hc from *O. vulgaris* and from *C. aestuarii*, respectively. These results indicate that the simulation of the met-Hc forms of *O. vulgaris* and *C. aestuarii*, starting from the analogous form of *L. polyphemus*, needs to be improved by considering alternative geometries of the two copper sites. This is in order to take into account whether the residual differences observed between the experimental spectra and the theoretical calculation can be due to fine structural differences present in the

met-form of proteins of the same *phylum*.

In conclusion, the resolution of the fine structural aspects of these binuclear sites is difficult, and the simulation using the present version of the *CONTINUUM* code is particularly delicate, owing to the limits of use of the Hedin–Lundqvist potential in the near-edge region. Nevertheless, our simulations of the structural data in this energy region seem to successfully reproduce all the expected features, according to the indications of the PDB codes.

Moreover, our MS XANES calculations are showing, and anticipating too, that, with a simulated two-centre spectrum, it



**Figure 7**

(a) Comparison of the experimental spectra in the XANES region of the met forms of *O. vulgaris* and *C. aestuarii* at pH 7.5. (b) Simulation of the Cu *K*-edge absorption of the PDB code 1LL1: comparison of the three-shell simulations of Cu1, Cu2 and (Cu1 + Cu2) sites. (c) Comparison between the experimental spectrum of met-Hc of *O. vulgaris* at pH 7.5 and the (Cu1 + Cu2) three-shell simulation of 1LL1. (d) Comparison between the experimental spectrum of met-Hc of *C. aestuarii* at pH 7.5 and the (Cu1 + Cu2) three-shell simulation of 1LL1.

**Table 2**

Structural data of the met-Hc form of *Limulus polyphemus*.

Starting coords file: PDB code 1LL1			
First-shell cluster for the Cu <sub>1</sub> site		First-shell cluster for the Cu <sub>2</sub> site	
Distances (Å)		Distances (Å)	
0	0.00000	Cu_Cu__1	
1	1.88052	O__HOH__51	
2	1.96839	NE2_HIS__324	
3	2.07634	O__HOH__124	
4	2.12445	NE2_HIS__364	
5	2.46666	NE2_HIS__328	
0	3.12816	Cu_Cu__2	
Angles (°)		Angles (°)	
ang 1-0-2	162.1	ang 1-0-2	95.1
ang 1-0-3	74.2	ang 1-0-3	92.1
ang 2-0-3	105.8	ang 2-0-3	153.2
ang 1-0-4	89.3	ang 1-0-4	142.0
ang 2-0-4	85.2	ang 2-0-4	71.2
ang 3-0-4	157.4	ang 3-0-4	87.3
ang 1-0-5	96.7	ang 1-0-5	115.3
ang 2-0-5	100.5	ang 2-0-5	96.4
ang 3-0-5	104.9	ang 3-0-5	103.7
ang 4-0-5	92.0	ang 4-0-5	101.6

would be possible to reproduce the experimental data of the binuclear (type 3) copper sites of hemocyanins and related model compounds.

In reality, a bigger source of discrepancy concerning intensities is due to the imaginary part of the Hedin-Lundqvist potential, even if this discrepancy could be attributed to a lack of refinement of the starting structural parameters, which is not attempted in this case, as was done instead in the EXAFS region. To overcome the overestimation or underestimation of the intensities of all features in the spectra it would be necessary to consider an energy-dependent damping. This would result in a much better simulation of the data with increased sensitivity to the structural parameters (Joly *et al.*, 1999, and references therein). This is what is performed by the new *MXAN* software (Benfatto & Della Longa, 2001), which combines the *CONTINUUM* code with a *MINUIT* minimization procedure (CERN library) to optimize the structural parameters by fitting the simulated spectrum onto the experimental data, fitting at the same time an energy-dependent term due to intrinsic and extrinsic inelastic processes. Furthermore, while with *CONTINUUM* it is not possible to optimize at the same time the signals from two different

absorbers, the use of the *MXAN* program makes it possible to consider the presence of a second absorbing site.

Since our present work on the XANES region anticipates in a way the good performance of *MXAN* once a multi-centre fitting procedure is implemented, we are looking forward to this.

We acknowledge A. Filipponi and A. Di Cicco for the use of the *GNXAS* program, and M. Benfatto for the use of the *CONTINUUM* program. We acknowledge the European Synchrotron Radiation Facility and the General Italian Beamline for Diffraction and Absorption for provision of synchrotron radiation facilities.

## References

- Battaini, G., Casella, L., Gullotti, M., Monzani, E., Nardin, G., Perotti, A., Randaccio, L., Santagostini, L., Heinemann, F. W. & Schindler, S. (2003). *Eur. J. Inorg. Chem.* **6**, 1197–1205.
- Benfatto, M. & Della Longa, S. (2001). *J. Synchrotron Rad.* **8**, 1087–1094.
- Blackburn, N. J., Strange, R. W., Farooq, A., Haka, M. S. & Karlin, K. D. (1988). *J. Am. Chem. Soc.* **110**, 4263–4272.
- Borghini, E., Solari, P. L., Beltramini, M., Bubacco, L., Di Muro, P. & Salvato, B. (2002). *Biophys. J.* **82**, 3254–3268.



- Brown, J. M., Powers, L., Kincaid, B., Larrabee, J. A. & Spiro, T. G. (1980). *J. Am. Chem. Soc.* **102**, 4210–4216.
- Casella, L., Carugo, O., Gullotti, M., Doldi, S. & Frassoni, M. (1996). *Inorg. Chem.* **35**, 1101–1113.
- Casella, L., Carugo, O., Gullotti, M., Garofani, S. & Zanella, P. (1993). *Inorg. Chem.* **32**, 2056–2067.
- Clementi, E. & Roetti, C. (1974). *Atom. Data Nucl. Data Tables*, **14**, 177–478.
- Co, M. S. & Hodgson, K. O. (1981). *J. Am. Chem. Soc.* **103**, 3200–3201.
- Co, M. S., Hodgson, K. O., Eccles, T. K. & Lontie, R. (1981). *J. Am. Chem. Soc.* **103**, 984–986.
- Cuff, M. E., Miller, K. I., van Holde, K. E. & Hendrickson, W. A. (1998). *J. Mol. Biol.* **278**, 855–870.
- Della Longa, S., Ascone, I., Bianconi, A., Bonfigli, A., Congiu Castellano, A., Zarivi, O. & Miranda, M. (1996). *J. Biol. Chem.* **271**, 21025–21030.
- Della Longa, S., Bianconi, A., Palladino, L., Simonelli, B., Congiu Castellano, A., Borghi, E., Barteri, M., Beltramini, M., Rocco, G. P., Salvato, B., Bubacco, L., Magliozzo, R. S. & Peisach, J. (1993). *Biophys. J.* **65**, 2680–2691.
- Di Cicco, A. (2003). *J. Synchrotron Rad.* **10**, 46–50.
- Durham, P. J. (1988). *Theory of XANES*, in *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEFAXS, XANES*, edited by R. Prinz and D. Koningsberger, pp. 53–84. New York: J. Wiley and Sons.
- Filipponi, A. & Di Cicco, A. (1995). *Phys. Rev. B*, **52**, 15135–15149.
- Filipponi, A. & Di Cicco, A. (2000). *Task Q.* **4**, 575–669.
- Filipponi, A., Di Cicco, A. & Natoli, C. R. (1995). *Phys. Rev. B*, **52**, 15122–15134.
- Hedin, L. & Lundqvist, B. I. (1971). *J. Phys. C*, **4**, 2064–2083.
- Kau, L. S., Spira-Solomon, D. J., Penner-Hahn, J. E., Hodgson, K. O., & Solomon, E. I. (1987). *J. Am. Chem. Soc.* **109**, 6433–6442.
- Klabunde, T., Eicken, C., Sacchetti, J. C. & Krebs, B. (1998). *Nature Struct. Biol.* **5**, 1084–1090.
- Joly, Y., Cabaret, D., Renevier, H. & Natoli, C. R. (1999). *Phys. Rev. Lett.* **82**, 2398–2401.
- Li, C., Pompa, M., Della Longa, S. & Bianconi, A. (1991). *Physica C*, **178**, 421–431.
- Magnus, K. A., Hazes, B., Ton-That, H., Bonaventura, C., Bonaventura, J. & Hol, W. G. (1994). *Proteins*, **19**, 302–309.
- Norman, J. G. Jr (1976). *Mol. Phys.* **31**, 1191–1198.
- Perbandt, M., Guthohrlein, E. W., Rypniewski, W., Idakieva, K., Stoeva, S., Voelter, W., Genov, N. & Betzel, C. (2003). *Biochemistry*, **42**, 6341–6346.
- Pickering, I. J. & George, G. N. (1995). *Inorg. Chem.* **34**, 3142–3152.
- Sano, M., Komorita, S. & Yamatera, H. (1992). *Inorg. Chem.* **31**, 459–463.
- Shadle, S. E., Penner-Hahn, J. E., Schugar, H. J., Hedman, B., Hodgson, K. O. & Solomon, E. I. (1993). *J. Am. Chem. Soc.* **115**, 767–776.
- Solari, P. L. (2002). PhD thesis, Université Joseph Fourier, Grenoble, France.
- Tyson, T. A., Hodgson, K. O., Natoli, C. R. & Benfatto, M. (1992). *Phys. Rev. B*, **46**, 5997–6019.
- Volbeda, A. & Hol, W. G. (1989). *J. Mol. Biol.* **209**, 249–279.
- Woolery, G. L., Powers, L., Winkler, M., Solomon, E. I., Lerch, K. & Spiro, T. G. (1984b). *Biochim. Biophys. Acta*, **788**, 155–161.
- Woolery, G. L., Powers, L., Winkler, M., Solomon, E. I. & Spiro, T. G. (1984a). *J. Am. Chem. Soc.* **106**, 86–92.
- Zippel, F., Ahlers, F., Werner, R., Haase, W., Nolting, H. F. & Krebs, B. (1996). *Inorg. Chem.* **35**, 3409–3419.