Journal of
Synchrotron Radiation

ISSN 0909-0495

Received 6 November 2008
Accepted 20 March 2009

Biological X-ray absorption spectroscopy and metalloproteomics

Isabella Ascone[a]* and Richard Strange[b]

[a]Synchrotron SOLEIL, BP 48, 91192 Gif-sur-Yvette Cedex, France, and [b]Molecular Biophysics Group, School of Biological Sciences, University of Liverpool, Liverpool L69 7ZB, UK.
E-mail: isabella.ascone@synchrotron-soleil.fr

In the past seven years the size of the known protein sequence universe has been rapidly expanding. At present, more then five million entries are included in the UniProtKB/TrEMBL protein database. In this context, a retrospective evaluation of recent X-ray absorption studies is undertaken to assess its potential role in metalloproteomics. Metalloproteomics is the structural and functional characterization of metal-binding proteins. This is a new area of active research which has particular relevance to biology and for which X-ray absorption spectroscopy is ideally suited. In the last three years, biological X-ray absorption spectroscopy (BioXAS) has been included among the techniques used in post-genomics initiatives for metalloprotein characterization. The emphasis of this review is on the progress in BioXAS that has emerged from recent meetings in 2007–2008. Developments required to enable BioXAS studies to better contribute to metalloproteomics throughput are also discussed. Overall, this paper suggests that X-ray absorption spectroscopy could have a higher impact on metalloproteomics, contributing significantly to the understanding of metal site structures and of reaction mechanisms for metalloproteins.

Keywords: metalloproteomics; metal site; metalloproteins; X-ray absorption spectroscopy; BioXAS.

1. Abbreviations

EXAFS: extended X-ray absorption fine structure.
XANES: X-ray absorption near-edge structure.
NEXAFS: near-edge X-ray absorption fine structure.
XAFS: X-ray absorption fine structure.
XAS: X-ray absorption spectroscopy.
BioXAS: biological X-ray absorption spectroscopy.

2. Introduction

Metal ions are essential for several fundamental biological processes in both eukaryotic and prokaryotic organisms. Metalloproteins are responsible for several metabolic processes, such as biological energy conversion in photosynthesis and respiration, and signalling processes which govern gene expression and regulation (Holm *et al.*, 1996).

Genomes of many organisms have already been sequenced and the number of metalloproteins is at least one-third of the total number of proteins (Lu, 2006). Some authors even indicate a higher percentage, up to 50% (Thomson & Gray, 1998; Strange *et al.*, 2005). Bioinformatics will contribute to refine the accuracy of these estimates (see §2).

Emerging post-genomics areas, involving metal studies, include metallomics and metalloproteomics. Metallomics is a comprehensive analysis of the entirety of metal and metalloid species within a cell or tissue type (Gao *et al.*, 2007). Metallome in phytoplankton and bacteria has been determined: trace metals consisting of transition metals plus zinc are present in a stoichiometric molar formula as follows, $Fe_1Mn_{0.3}Zn_{0.26}Cu_{0.03}Co_{0.03}Mo_{0.03}$ (Barton *et al.*, 2007). Metalloproteomics is the study of a part of the metallome; it is focused on investigating the distributions, structure and function of all metalloproteins in a proteome (Gao *et al.*, 2007). Metalloproteomics has also been defined as the structural and functional characterization of metal-binding proteins and their structural metal-binding moieties (Jakubowski *et al.*, 2004).

BioXAS can play an important role in metalloproteomics. With this technique, scientists can determine both molecular and electronic structural details of metal sites in metalloproteins, biomimetic complexes or other biological systems, giving an insight into metalloenzyme reaction mechanisms.

In the last 25 years, X-ray absorption spectroscopy has been used as a powerful tool for local structural and electronic determinations in a large panel of scientific domains. To give a quantitative indication of the scientific activities involving this

spectroscopy, we searched the number of publications as a function of time in the ISI Web of Knowledge (WOS database), updated on 26 October 2008. This evaluation clearly depends on the completeness of the database over a long period of time and the panel of journals included in the database. Nevertheless, a trend emerges, especially in the last ten years, where the majority of journals have an electronic version. Fig. 1(*a*) shows the number of publications using X-ray absorption spectroscopy as a function of time (five-year periods), obtained searching 'X-ray absorption' and related keywords (see *Abbreviations*) within article titles, article keywords and abstracts contained in databases. Figs. 1(*a*) and 1(*b*) show an increase in the number of publications as a function of time, in particular for those concerning BioXAS in the January 2003 to December 2007 period.

In spite of the high intrinsic interest of this technique for metalloprotein studies, the number of BioXAS publications has progressed only recently, owing to several limiting factors, many of which have now been overcome.

Three international workshops (BioXAS Study Weekends) have considered the development of this technique in the post-genomic frame (Ascone, Fourme & Hasnain, 2003; Ascone, Fourme *et al.*, 2005). Two special issues of the *Journal of Synchrotron Radiation* (Volume 10, Part 1, 2003 and Volume 12, Part 1, 2005) related the presentations of the first two meeting, held in Orsay, France, on 30 June–1 July 2001 and on 29–30 June 2003.

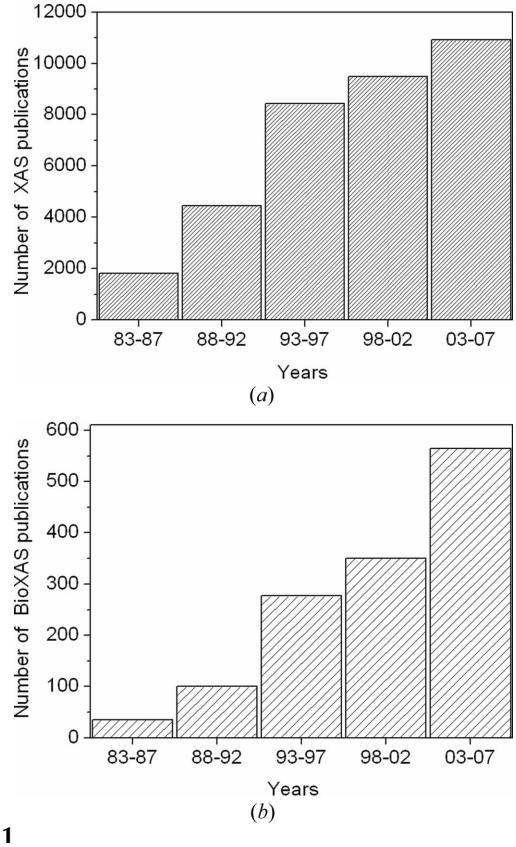

Figure 1
The number of XAS (*a*) and BioXAS (*b*) publications as a function of time in the ISI Web of Knowledge (WOS database). XAS publications have been selected using XAS-related keywords (EXAFS, XAS, XANES, X-ray absorption, NEXAFS or XAFS). BioXAS publications have been obtained restricting the research to the following keywords: protein*, metalloprotein*, metalloenzyme* or biological.

One aim of the present review is to give an updated overview of 'BioXAS and metalloproteomics' that emerged from the last international workshop held at the Synchrotron-Soleil Facility (France) on 10–11 August 2007 and from the XXI Congress of the International Union of Crystallography at Osaka, Japan, on 23–31 August 2008.

We first present X-ray absorption spectroscopy indicating methodological/theoretical advances in data analysis and data reduction (§3). Progress in identification and expression of metalloproteins (§4) and new experimental set-ups for BioXAS data acquisition at synchrotrons (§5) are described. The last section is dedicated to the recent applications of XAS to a range of important problems in bioinorganic chemistry and in biology (§6).

3. Methodological/theoretical advances in data analysis and data reduction

The increasing rate of BioXAS publications is due to many factors illustrated in this review, one of these being connected to the progress both in the theory and in *ab initio* codes for calculations of X-ray absorption spectra (Natoli *et al.*, 2003; Rehr, 2006).

An X-ray absorption spectrum is measured when X-rays have sufficient energy to eject one or more core electrons from the absorbing atom. Around this energy value there is an abrupt increase in the absorption coefficient, the so-called 'absorption edge'.

X-ray absorption spectra can be roughly divided into three regions with reference to the absorption edge: the pre-edge, near-edge (XANES) and post-edge (EXAFS) regions. Fig. 2 shows an example of an experimental X-ray absorption spectrum. For each region, specific procedures and codes to simulate experimental data have been developed.

The pre-edge region in metal absorption spectra may be simulated to obtain the electronic structure of the metal atom (Joly, 2003). Pre-edge features in ligand absorption spectra (*e.g.* Cl or S) can be related to the occupation of a molecular orbital and to the covalency of the ligand–metal bond. Readers are referred to a review on this item concerning metalloproteins (Solomon *et al.*, 2005).

The EXAFS region has been largely exploited to obtain quantitative information (metal site ligation, bond lengths, coordination number), and many reviews in the literature, in the past (Lee *et al.*, 1981) and more recently (Levina *et al.*, 2005), described details of EXAFS data analysis. Modern EXAFS data analysis programs include full multiple-scattering calculations (MS) (Binsted *et al.*, 1992; Di Cicco, 2003; Rehr & Ankudinov, 2003) which complete the more simple single-scattering analysis, widely used previously. MS is important in analysis of protein metal sites, for example the rigid imidazole ring of a histidine ligand gives rise to diagnostic multiple-scattering signals, and MS can also enable

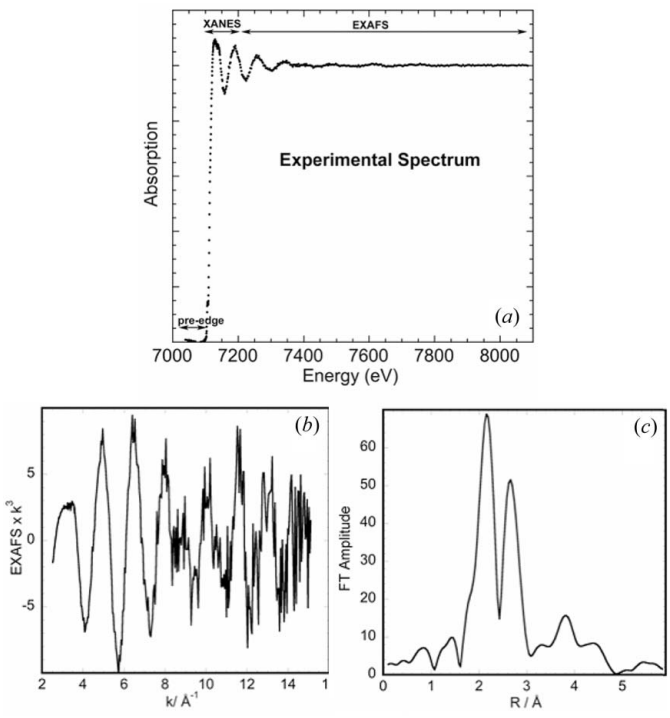

Figure 2
X-ray absorption spectrum measured at the Fe *K*-edge of the iron protein from *Klebsiella pneumoniae* nitrogenase (unpublished data). The different regions of the XAS spectrum are indicated in (*a*). The EXAFS obtained after background subtraction and normalization is shown in (*b*). The EXAFS spectrum is normally plotted in *k*-space in Å^{-1}, so that the Fourier transform (*c*) gives the radial distribution in Å corresponding to the distances of the backscattering atoms from the metal.

three-dimensional modelling of metalloprotein active sites when EXAFS is combined with protein crystal structure data.

In contrast with EXAFS, most analyses of the XANES region have focused on qualitative interpretations (Penner-Hahn, 2005). Only recently have technical procedures advanced sufficiently to extract quantitative structural information from XANES spectra (Benfatto *et al.*, 2001; Smolentsev & Soldatov, 2007). Methods for XANES quantitative analysis imply the use of an initial model which is rather close to the real structure. Structural parameters (angles and distances) are selected and refined. Initial models can be obtained from a biomimetic complex or from the protein metal site obtained by X-ray diffraction data, even at low resolution. This approach has been used for several protein metal sites, in the solution state or using protein crystals (Arcovito *et al.*, 2007; Marino *et al.*, 2007; Ascone, Nobili *et al.*, 2006).

BioXAS studies have specific advantages in exploiting XANES spectra for quantitative determinations, either complementing EXAFS analysis or as an alternative. XANES possesses the following characteristics.

(i) Spectroscopic features depend on the geometry of the metal site. Experimental data (Penner-Hahn, 2005) and simulations (Benfatto *et al.*, 2001; Smolentsev & Soldatov, 2007) indicate that intensity and position of peaks depend on the number and on the type of ligands. In addition, distance length variations induce modifications in the shape of theoretical XANES spectra, in spite of the same number and type of ligands (Ascone, Nobili *et al.*, 2006; Yalovega *et al.*, 2007). This last fact makes it unlikely that XANES could be routinely used as a 'fingerprint' method for rapidly interpreting or identifying coordination sites in 'unknown' systems.

(ii) Information on metal electronic structure may be obtained. The energy of the X-ray absorption edge varies with the oxidation state of the absorbing atom; as a first approximation, the edge position shifts toward higher energies upon oxidation. An accurate analysis, taking into account the shift also due to the ligand/coordination changes, indicates the reduction/oxidation of atoms, which can exist in multiple oxidation states.

(iii) Sensitivity to second and third atomic shells around the absorber is higher than for EXAFS, owing to the mean-free-path value for the photoelectron.

(iv) The XANES signal is almost two orders of magnitude more intense than the EXAFS signal, making it possible to use lower protein concentrations.

The most recent theoretical advances are expected to improve XANES interpretation. Up to now, simulation of spectra have been based on multiple-scattering theory to solve the Schrödinger equation, with potentials calculated in the so-called 'muffin tin' approximation, which cannot properly describe protein metal sites very often having high anisotropy. New full-potential multiple-scattering schemes (Hatada *et al.*, 2007) will potentially increase the structural and electronic information that may be obtained from examining the low-energy region of the absorption spectrum. Further advances are expected from the inclusion of electronic correlations in the description of the absorption process (Kruger & Natoli, 2004).

However, additional efforts have to be made to significantly increase the number of groups using EXAFS and, particularly, XANES analysis for metalloproteins studies.

Data analysis automation is essential for routine BioXAS contributions to metalloproteomics research. One outcome of recent BioXAS meetings is the proposal to develop a 'black box' approach for use by an expanding BioXAS community. This approach has been successfully adopted by the community of biocrystallographers who have developed programs with a high level of automatic control for input parameters (physical and chemical) and for program output (structure). This approach facilitates correct usage and removes unnecessary technical obstacles from the non-expert's view. This effort is performed in the frame of Collaborative Computational Project Number 4 (CCP4) and programs are free for academic use. The availability of such highly specialized programs, brought together in a well thought out graphical interface, has also contributed to the exponential growth of the number of structures deposited in the Protein Data Bank. The CCP4 model has inspired similar developments in NMR, and it is suggested that the biological XAS community may pursue similar collaborations (Winn, 2003). The idea is to eliminate or minimize the explicit role of theoretical parameters in the fitting procedures of XAS programs, instead using only structural parameters and making the process more

amenable to researchers from outside the physical sciences community. Some progress has already been achieved in the frame of the *FEFF* project (Rehr & Albers, 2000) with the *ab initio* mean-free-paths calculations (Sorini *et al.*, 2006) and theoretical X-ray absorption Debye–Waller factors (Vila *et al.*, 2007).

4. Progress in identification and expression of metalloproteins

Sequencing of genomes has increased rapidly in recent years (Grabowski *et al.*, 2007) and the sizes of protein databases increases consequently: in December 2007 the UniProtKB/TrEMBL protein database contained more than five million sequence entries, fivefold more than in 2004 (Apweiler *et al.*, 2004; Bairoch & Apweiler, 2000). At present, the genome of communities of organisms is analyzed to understand their roles and interactions in an ecosystem. In the frame of the Sorcerer II Global Ocean Sampling expedition, microbial communities in water samples have been analyzed and more than six million proteins have been predicted (Yooseph *et al.*, 2007). Many methods to predict metal-binding proteins are based on sequence similarity to known metalloproteins, in particular looking at the amino acids involved in metal binding. Bioinformatics is now able to identify metalloproteins in archaea, bacteria and eukarya (Dupont *et al.*, 2006; Andreini *et al.*, 2006*b*). Zinc-binding proteins are between 5–6% (bacteria and archaea) and 8.8% (eukaryota) (Andreini *et al.*, 2006*b*), in particular 10% in the human proteome (Andreini *et al.*, 2006*a*). The non-heme iron-binding proteins in archaea are on average 7.1% $\pm$ 2.1% (Andreini *et al.*, 2007).

If we consider that bioinformatics takes into account only the known binding sequences and that in bacteria the zinc content is about one-quarter of that of iron [stoichiometric molar formula is $Fe_1Mn_{0.3}Zn_{0.26}Cu_{0.03}Co_{0\cdot03}Mo_{0.03}$ (Barton *et al.*, 2007)], the estimated number of metalloproteins in an organism [one-third of the total number of proteins (Lu, 2006)] is a lower limit. Then, it is clear that in the next ten years there will be a growing need for metalloprotein characterization.

The high-throughput production of metalloproteins involves many specific issues (Jenney *et al.*, 2005). These issues include the following.

(i) The accuracy of bioinformatics research in the selection and prediction of suitable targets and hypothetical metalloproteins.

(ii) Obtaining a properly expressed and folded functional protein that contains the correct metal.

(iii) The co-expression of chaperone proteins that are often required for proper folding of the target and/or insertion of the metal atom into the target.

(iv) The choice of a suitable expression system and whether anaerobic purification is necessary.

The provision of high-quality protein in adequate quantities is a prerequisite for BioXAS studies. BioXAS already takes advantage of expression and purification protocols optimized in structural genomics programmes, as for example the case of metalloproteins from the *Mycobacterium tuberculosis* genome (Hall *et al.*, 2005). The Production Module of the Southeast Collaboratory for Structural Genomics (SECSG) in collaboration with the University of Georgia (USA) has developed a protocol for analyzing recombinant metalloproteins with X-ray absorption spectroscopy for identifying metal centres, especially novel centres with no homologs. Bottlenecks in the pipeline from genome to metalloproteome, allowing determination of metal-site structures, have been discussed in terms of high-throughput X-ray absorption spectroscopy (HTXAS) (Scott *et al.*, 2005). Many of the bottlenecks identified may be overcome by developing automation for several tasks, particularly in the XAS pipeline, including manipulating microarray samples using robots, and rapid data collection and processing using intelligent decision-making algorithms. The level of automation demanded by HTXAS makes it radically different from the traditional BioXAS approach, which generally uses larger volumes and more concentrated samples with slower data collection and analysis procedures. The Italian structural genomics effort at the Magnetic Resonance Center (CERM) of the University of Florence is focused on metalloproteins. Their methodological approach is based on NMR with X-ray crystallography and X-ray absorption spectroscopy (Arnesano *et al.*, 2006).

A high-throughput method for measuring transition metal content, based on quantitation of X-ray fluorescence signals, has been used in the New York Structural GenomiX Research Consortium (Shi *et al.*, 2005).

5. Data collection: advances in beamlines for BioXAS

BioXAS data acquisitions cannot be performed with laboratory X-ray sources, such as rotating anodes or X-ray tubes, so that synchrotron radiation is the unique source available for such experiments. Owing to the low signal of protein samples, BioXAS applications have high experimental requirements. They need optimized beamlines having intense synchrotron photon flux, stable optics, sensitive fluorescence detectors and specific sample environments (Ascone, Meyer-Klaucke & Murphy, 2003). Recent technological and methodological advances have strongly increased the potential of XAS to characterize the structure and function of metalloproteins. BioXAS beamlines suitable for metalloproteomics have intense insertion-device or bending-magnet sources on second- or third-generation synchrotrons, respectively, which reduce the data-acquisition time as well as the quantity of the sample. For metalloproteins, which have a low solubility or which form oligomers at concentrations of <50 μM, insertion devices on last-generations synchrotron are required.

New set-ups also have an impact on the quality of EXAFS data. EXAFS spectra are typically measured over a limited k range {X-ray energy is usually converted to photoelectron wavevector k, which is the inverse photoelectron wavelength, $k = [2m_e(E - E_0)/\hbar^2]^{1/2}$}.

Sensitive detectors and intense sources have considerably increased the energy range of EXAFS spectra. In the past decade, typical values of k_{max} for metalloproteins studies were

between 11 and 14 Å^{-1}. Recent experiments (Rich *et al.*, 1998; Corbett *et al.*, 2005, 2006) on appropriate samples push this limit up to 16–18 Å^{-1}. With the optimization of experimental set-ups (sensitive detectors and more stable sources and optics) allowing lower noise in the data, the k_{max} value may be increased up to 25 Å^{-1} as for human sulfite oxidase (Harris *et al.*, 2006). This means that the number of independent data points (given by $\sim 2\Delta k \Delta R/\pi$) is increased with a positive impact on bond-length resolution and on the number of parameters allowed for fitting.

The combination of X-ray crystallography and X-ray absorption spectroscopy has proved extremely useful (Hasnain & Hodgson, 1999). The crystallographic phasing methods, namely multiple-wavelength anomalous dispersion (MAD), require a tunable monochromator as in the case of XAS experiments. Conceptually new experimental set-ups, where both MAD and high-quality XAS data can be collected on the same crystalline sample, were presented during BioXAS 2003. Currently this arrangement is implemented on the Stanford Synchrotron Radiation Laboratory (SSRL) beamline 9-3, a wiggler side-station dedicated to general user biological XAS (Latimer *et al.*, 2005). This beamline includes the developments in automation that have been made for high-throughput crystallography at SSRL (Cohen *et al.*, 2002); this development allows high-throughput XAS measurements. Moreover, XAS measurements can benefit from the plane-polarized nature of the synchrotron radiation beam. XAS spectra of disordered samples (*e.g.* solutions) have isotropic absorption while single-crystal samples, in which the metal sites are oriented, exhibit distinct dichroism. At the *K*-edge, $N^* = 3N\cos^2\theta$, where θ is defined as the angle between the polarization vector and an absorber-scatter vector, N is the coordination number and N^* is the effective polarization-dependent coordination number. Sample orientations are chosen to enhance the contribution of specific backscatterers to the XAS spectrum to have a better insight into the structure.

Polarized XAS has been used to better characterize cyanomet sperm whale myoglobin (Arcovito *et al.*, 2007) and the Mn_4Ca cluster within photosystem II (Yano *et al.*, 2006).

Intense sources may photoreduce metallic centres; a number of different methods to control or avoid this have been proposed for protein samples, in solution and lyophilized (Ascone, Meyer-Klaucke & Murphy, 2003; Ascone, Zamponi *et al.*, 2005; Ascone *et al.*, 2000). An assessment of the metalloprotein redox state is impossible to obtain crystallographically, but this can easily be done by XAS also on oriented single crystals (Yano *et al.*, 2005) using recent protein crystallography/BioXAS beamlines. The new beamlines dedicated to high-throughput crystallography at third-generation synchrotron radiation facilities submit samples to very high X-ray flux. Detection and monitoring of this phenomenon by single-crystal XAS, during crystallographic experiments, has been suggested (Ascone, Girard *et al.*, 2006; Holton, 2007).

Photoreduction was observed by XAS in crystals of photosystem II (Yano *et al.*, 2005) and of $[Fe_2S_2]$-containing metalloprotein putidaredoxin (Corbett *et al.*, 2007). In both cases the photoreduction measured at about 100 K decreases for lower temperatures (10 K or 40 K, respectively). Use of liquid-helium-based rather than liquid-nitrogen cooling for metalloprotein crystallography should limit the photoreduction of sensitive sites (Corbett *et al.*, 2007). As well as using XAS, the effects of radiation damage at metal centres during crystallographic data collection can also be followed by *in situ* optical measurements on metalloprotein crystals (Hough *et al.*, 2008; Ellis *et al.*, 2008). Fig. 3 shows an image of this experimental set-up. These multi-technique experiments have shown that X-ray-induced photoreduction occurs on a more rapid time scale than that normally required for complete crystallographic data collection and happens long before X-ray damage to the protein crystals leads to loss of diffractive power. For a given metalloprotein, identification of the metal oxidation state and coordination environment is needed at the same time, and on the same crystal, that crystallographic data are recorded. An intriguing suggestion is to collect complete crystallographic data sets at several energies around the absorption edge of a metal and use these data to obtain (or reconstruct) the XANES for the metal by refining the energy-dependent anomalous scattering factors for each atom (Einsle *et al.*, 2007). Individual sites in multi-metal-containing proteins

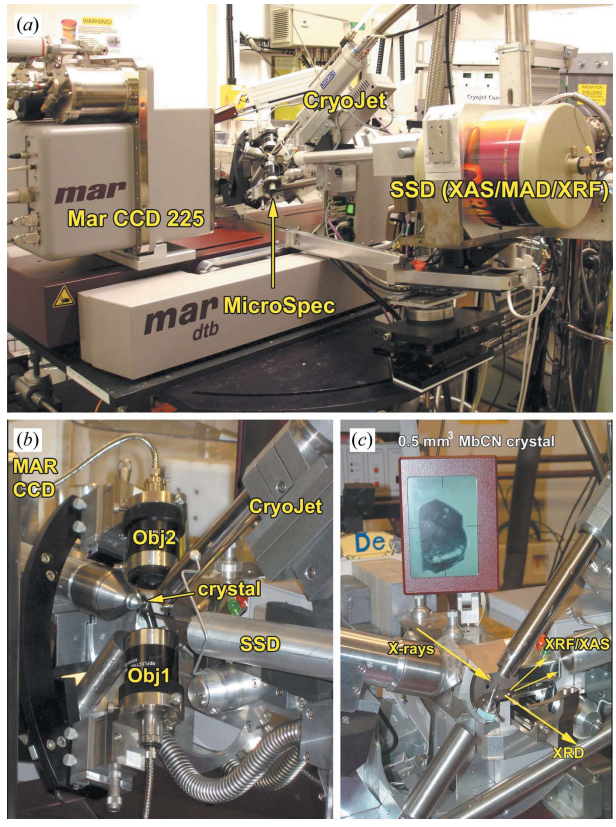

Figure 3
A BioXAS experimental set-up (*a*) allowing both crystallographic data collection and *in situ* optical measurements on metalloprotein crystals (Hough *et al.*, 2008; Ellis *et al.*, 2008), with a solid-state detector (SSD) for measuring X-ray fluorescence (XRF) and X-ray absorption (XAS or XANES) data. (*b*) Close up of the microspectrophotometer system. (*c*) Arrangement for polarized single-crystal XANES experiments with the microspectrophotometer removed (Arcovito *et al.*, 2007).

(*e.g.* ceruloplasmin) could be picked out and identified in this way. The challenge is to apply the method without radiation damage to the metal site or indeed the protein during the time required for several complete crystal data sets to be collected. Use of rapid data collection methods is essential here.

In the last study weekend meeting, new beamlines fully or partially dedicated to BioXAS experiments were presented. These are built or are under construction on third-generation synchrotron radiation sources characterized by high flux and microfocused beam: Canadian Light Source, Synchrotron-Soleil, Diamond Light Source. Their development has benefited from earlier generations of XAS beamlines that have already contributed significantly to BioXAS research. Some of these beamlines, like those at the SRS Daresbury Laboratory, UK, EMBL Hamburg, Germany, and D21-LURE, France, have ceased operation. Others, like those at Stanford and the X3b beamline at the Brookhaven NSLS, are still active. In order to contribute to metallogenomics, all these beamlines should be equipped with highly specific protein sample environments.

The Synchrotron-Soleil Facility (France) has several X-ray absorption beamlines that, in principle, are suitable for BioXAS experiments. SAMBA and LUCIA (Flank *et al.*, 2006) are multipurpose XAS beamlines covering the 5–40 keV and 0.8–8 keV energy range, respectively. The LUCIA microsource (2×2 µm) is suitable for XAS imaging. Most BioXAS experiments are performed at energies higher than 5 keV; LUCIA provides softer X-rays, which probe the local structural and electronic environment of relevant elements for biology (S, Na, Mg, Ca). Such spectroscopic studies on biological systems are quite recent (Akabayov *et al.*, 2005). PROXIMA 1 (Girard *et al.*, 2006) is a Soleil beamline for macromolecular crystallography that will be optimized to combine MAD/SAD and BioXAS measurements on single crystals with X-ray energies in the 5–15 keV range (Ascone, Girard *et al.*, 2006).

At Diamond Light Source (UK) a new beamline (ID20) is under construction. It is optimized for operation between 4 and 35 keV based upon a multipole wiggler source. This versatile X-ray spectrometer will be dedicated to ultra-dilute systems, including biological samples (http://www.diamond.ac.uk/default.htm).

A phase III beamline project has been approved (in 2006) at the Canadian Light Source (http://www.lightsource.ca/). It is focused on a suite of two beamlines (BioXAS-1 and BioXAS-2) specifically optimized for biological and health-related studies. BioXAS-1 is a spectroscopic set-up operating in the 5–28 keV energy range while BioXAS-2 will be dedicated to imaging (5–15 keV). The imaging scale varies from 50 µm to less then 1 µm for investigation of tissues and sub-cellular regions.

6. Applications of XAS in bioinorganic chemistry and in biology

This section illustrates recent studies in which BioXAS has contributed to the understanding of biological issues, providing structural and electronic details during protein function. The majority of these studies have been on protein solutions, usually at liquid-nitrogen or liquid-helium temperatures. The interest in combining BioXAS with crystallography and applications using polarized single-crystal experiments are highlighted. This is not an exhaustive review of all recent absorption studies; it intends to show, focusing on some biological process (*e.g.* metal trafficking in the cell, substrate interactions with protein metal site and catalytic), that this spectroscopy is an appropriate tool for metalloproteomics.

6.1. Accurate structural determinations of metal sites combining BioXAS and X-ray diffraction

There is an increasing number of higher/atomic resolution structures reported in the Protein Data Bank. However, of some 250 structures of Cu-containing proteins, about 2% reported at better than 1.2 Å resolution and 90% of structures had a resolution lower than 1.5 Å (Strange *et al.*, 2005). An understanding of the complex biological processes at the molecular level requires an atomic resolution that increases the quality of electron density maps revealing details of the metal site. EXAFS measurements are used to determine the metal–ligand distances with greater accuracy than is usually possible crystallographically owing to the limited resolution. Crystallographic coordinates have been used to fit EXAFS data, refining models of complex clusters, as for the FeMo co-factor site in Kp1 nitrogenase (Strange *et al.*, 2003) or the FeMo co-factor precursor bound to NifEN protein (Corbett *et al.*, 2006).

Even when an atomic crystallographic structure is obtained, an unwanted metal may be bound to the metal site requiring an XAS study. For example, the periplasmic copper/silver-binding protein (CusF) was characterized by crystallography at 1.0 Å resolution but crystals of Cu(I)-loaded CusF had not been obtained. By combining XAS and diffraction data, an unusual metal coordination for both Cu(I) and Ag(I) complexes was determined (Loftin *et al.*, 2007).

XAS spectroscopy is particularly appropriate when the metal is bound to flexible protein regions, or to chain termini metal-sites that are disordered or are determined with low accuracy. This can be illustrated by the example of the viral non-structural proteins (NS3) for which crystallographic structures show a Zn binding site with high temperature factors (Di Marco *et al.*, 2000; Yao *et al.*, 1999). In this case, the information obtained by protein X-ray diffraction has been used to refine distances, fitting XANES spectra with *MXAS* software. This kind of analysis, performed on the protein in solution (Ascone, Nobili *et al.*, 2006), takes advantage of the fact that XANES spectra are poorly sensitive to the Debye–Waller factors (Rehr & Albers, 2000). The type and numbers of amino acids which coordinate the metal are known by protein crystallography, and a rigid-body refinement of metal/amino acids distances is performed, reducing the number of fitting parameters.

6.2. BioXAS contribution to the understanding of catalytic processes

Catalytic processes controlled by metalloenzymes which have fundamental and technological relevance can be investigated. One of the objectives of biochemists and chemists is to better understand factors that modulate metalloproteins activity in order to develop metalloenzymes with novel structures and functions (Lu, 2006). This approach has an impact on biocatalysis and biosensor technology, such as in the degradation of pollutants and biomass, and in drug and food processing (Maglio *et al.*, 2007).

A survey of the reactivity associated with nitroxyl (HNO or NO^-), the reduced form of nitric oxide (NO), has been reported (Farmer & Sulc, 2005). X-ray absorption spectroscopy has contributed with other techniques, such as H-1 NMR and resonance Raman, to the characterization of the coordination chemistry of the nitroxyl ligand with hemes and synthetic coordination complexes.

Reactive intermediates have been trapped using rapid-freeze-quench of samples at various stages along a reaction pathway. This procedure allowed the investigation of the zinc active site in bacterial alcohol dehydrogenase during single substrate turnover, with reaction times varying from 2 ms to 110 ms (Kleifeld *et al.*, 2003).

Transmembrane proteins are important cellular components performing several crucial functions for the cell as receptors or transporters of molecules and ions. Owing to their physiological relevance they are of considerable interest to the pharmaceutical industry as a target for drugs. Three-dimensional structures of transmembrane proteins at the atomic level remain a particular challenge for crystallography (often they are insoluble and then difficult to crystallize) and for NMR (protein size limits the structural refinement step) (Lacapere *et al.*, 2007).

For BioXAS studies, membrane metalloproteins are excellent candidates, as membranes may be ordered in one dimension (z), where the z axis is collinear with the membrane normal, and the x and y axes remain disordered. This approach, enhancing the contribution of specific back-scatterers, has been used to characterize the Mn_4Ca cluster in plant photosystem II membranes (Pushkar *et al.*, 2007). XAS measurements with quantum mechanics/molecular mechanics data provided structural models of catalytic water oxidation in photosystem II (Dau *et al.*, 2001; Sproviero *et al.*, 2007). A new model for the Mn_4Ca cluster has been proposed through coupling of X-ray diffraction and single-crystal EXAFS (Yano *et al.*, 2006). Solution XANES combined with crystallographic data may allow structural models of catalysis to be revealed better than the crystal structure alone, a recent example being the nucleophilic attack mechanism proposed for degradation of histidine at the Zn site of imidazoloneproprionase (Yang *et al.*, 2008). Furthermore, structural parameters obtained by simulations of XANES solution spectra may be used to restrain metal sites during refinement of crystal structures (Strange *et al.*, 2003). In both applications, improvements in the local (metal site) accuracy and precision of the crystal structure are obtained that can aid interpretation of ligand binding and catalysis.

6.3. Metal trafficking

Environments that contain large amounts of heavy metals (*e.g.* mercury, copper, cadmium or lead) may be toxic to living organisms at certain concentrations or inappropriate oxidation states. Nevertheless, they are essential, in small quantities, for protein structure and function. Opella and co-workers (Opella *et al.*, 2002) have reviewed the structural biology of proteins which sequester and transport metals, showing that X-ray absorption is appropriate for the investigation of protein-mediated mechanisms in heavy-metal homeostasis. Further research is required to enhance the knowledge on transfer mechanisms of metal ions between proteins.

Organisms have developed specific systems to control metal trafficking in order to maintain the homeostatic balance of intracellular copper. X-ray absorption has contributed with other techniques (Banci *et al.*, 2005) to the elucidation of the copper–protein interaction (Arnesano *et al.*, 2003; Pufahl *et al.*, 1997; DiDonato *et al.*, 2000).

The metal sites of ferric uptake regulator (Fur) from *P. aeruginosa* have been investigated in solution and in the crystalline state by XAS and crystallography, respectively (Pohl *et al.*, 2003).

Toxicity associated with childhood lead poisoning may be attributable to interactions of Pb(II) with proteins containing thiol-rich structural zinc-binding sites. Lead toxicity stems from its ability to mimic Zn, interfering for example with the essential enzyme, aminolevulinic acid dehydratase (ALAD). Lead containing ALAD structure has been determined by crystallography (Erskine *et al.*, 2000). XAS structural studies of Pb(II)/Zn peptides indicated that the Pb(II) coordination sphere is significantly different from that of Zn(II) bound to the same peptides providing models for protein sites and critical insights into the mechanism by which lead alters the protein activity (Magyar *et al.*, 2005).

6.4. Micro-XAS and imaging

There is a panel of analytical techniques, including XAS, for biological trace-element imaging, and identification/quantification of chemical species in the biological environment (Lobinski *et al.*, 2006).

The speciation and distribution of heavy metal within leaves of plants has been considered in order to understand their metal biochemistry and to select species for specific functions. X-ray absorption spectroscopy and XAS imaging were used to examine *in vivo* the hyper accumulator *Iberis intermedia*, capable of accumulating expensive metal-like thallium (Scheckel *et al.*, 2004), and fern *Pteris vittata,* which accumulates unusually high levels of arsenic (Pickering *et al.*, 2006). Accumulation of selenium in *A. bisulcatus* plant (Pickering *et al.*, 2003) and caesium distribution in *Arabidopsis thaliana* (Isaure *et al.*, 2006) plant have also been considered. Plants accumulating toxic metals are useful for phytoremediating contaminated sites.

The distribution of chromium and endogenous elements within A549 human lung adenocarcinoma epithelial cells, following treatment with Cr(VI), was studied by micro-XAS. Chromium oxidation states have been determined in cells (Harris *et al.*, 2005). Micro-XANES under cryogenic conditions at subcellular levels has already been experimented at ESRF (Bacquart *et al.*, 2007). The distribution of elements was determined in primary invasive ductal carcinoma of breast (Farquharson *et al.*, 2008) and in colorectal liver metastases (Gurusamy *et al.*, 2008).

The chemical form of mercury in fish intended for human consumption has also been determined (Harris *et al.*, 2004).

The impact of a model water dispersion of nanoparticles (7 nm CeO_2 oxide) on Gram-negative bacteria has been studied, defining the conditions for which contact is lethal (Thill *et al.*, 2006).

It is now possible, using intense third-generation synchrotrons like that of the APS at Argonne, to produce extremely intense X-ray beams with a submicrometre beam size (0.25 μm). X-ray fluorescence maps of thin sections of human ovarian cancer cells treated with anti-cancer active complexes containing platinum have been collected localizing platinum in the cell nucleus (Hall *et al.*, 2006). An original experimental set-up was developed at ESRF (Grenoble) to perform chemical element imaging with a 90 nm spatial resolution (Ortega *et al.*, 2007). This enables chemical element imaging in subcellular compartments (Bohic *et al.*, 2008).

7. Conclusions

The studies described in this review show that BioXAS is a flexible technique where biological samples can be studied in a variety of forms, including oriented single protein crystals, disoriented microcrystals, frozen or room-temperature solutions and proteins inserted in membranes. Experimentally, BioXAS is now relatively straightforward to perform at dedicated synchrotron radiation beamlines and, using current analytical tools, is capable of providing key insights into the electronic states and atomic structures of metalloproteins, localized to the metal site at ultra-high resolution. We have pointed out a need for further development of the analytical tools towards a more unified, user-friendly and expert-free implementation, along the lines of collaborative programmes like CCP4 that have proved so successful for protein crystallography.

In the context of metalloproteomics where a large amount of proteins will be expressed, BioXAS is expected to give a valuable contribution to our understanding and knowledge of metalloproteins.

References

Akabayov, B., Doonan, C. J., Pickering, I. J., George, G. N. & Sagi, I. (2005). *J. Synchrotron Rad.* **12**, 392–401.

Andreini, C., Banci, L., Bertini, I., Elmi, S. & Rosato, A. (2007). *Proteins*, **67**, 317–324.

Andreini, C., Banci, L., Bertini, I. & Rosato, A. (2006*a*). *J. Proteome Res.* **5**, 196–201.

Andreini, C., Banci, L., Bertini, I. & Rosato, A. (2006*b*). *J. Proteome Res.* **5**, 3173–3178.

Apweiler, R. *et al.* (2004). *Nucl. Acids Res.* **32**, D115–119.

Arcovito, A., Benfatto, M., Cianci, M., Hasnain, S. S., Nienhaus, K., Nienhaus, G. U., Savino, C., Strange, R. W., Vallone, B. & Della Longa, S. (2007). *Proc. Natl. Acad. Sci. USA*, **104**, 6211–6216.

Arnesano, F., Banci, L., Bertini, I., Capozzi, F., Ciofi-Baffoni, S., Ciurli, S., Luchinat, C., Mangani, S., Rosato, A., Turano, P. & Viezzoli, M. S. (2006). *Coord. Chem. Rev.* **250**, 1419–1450.

Arnesano, F., Banci, L., Bertini, I., Mangani, S. & Thompsett, A. R. (2003). *Proc. Natl. Acad. Sci. USA*, **100**, 3814–3819.

Ascone, I., Fourme, R., Hasnain, S. & Hodgson, K. (2005). *J. Synchrotron Rad.* **12**, 1–3.

Ascone, I., Fourme, R. & Hasnain, S. S. (2003). *J. Synchrotron Rad.* **10**, 1–3.

Ascone, I., Girard, E., Gourhant, P., Legrand, P., Roudenko, O., Roussier, L. & Thompson, A. W. (2006). *13th International Conference on X-ray Absorption Fine Structure (XAFS13)*, edited by B. Hedman & P. Pianetta, pp. 872–874. New York: American Institute of Physics.

Ascone, I., Meyer-Klaucke, W. & Murphy, L. (2003). *J. Synchrotron Rad.* **10**, 16–22.

Ascone, I., Nobili, G., Benfatto, M. & Congiu-Castellano, A. (2006). *13th International Conference on X-ray Absorption Fine Structure (XAFS13)*, edited by B. Hedman & P. Pianetta, pp. 319–321. New York: American Institute of Physics.

Ascone, I., Sabatucci, A., Bubacco, L., Di Muro, P. & Salvato, B. (2000). *Eur. Biophys. J. Biophys. Lett.* **29**, 391–397.

Ascone, I., Zamponi, S., Cognigni, A., Marmocchi, F. & Marassi, R. (2005). *Electrochim. Acta*, **50**, 2437–2443.

Bacquart, T., Deves, G., Carmona, A., Tucoulou, R., Bohic, S. & Ortega, R. (2007). *Anal. Chem.* **79**, 7353–7359.

Bairoch, A. & Apweiler, R. (2000). *Nucl. Acids Res.* **28**, 45–48.

Banci, L., Bertini, I. & Mangani, S. (2005). *J. Synchrotron Rad.* **12**, 94–97.

Barton, L. L., Goulhen, F., Bruschi, M., Woodards, N. A., Plunkett, R. M. & Rietmeijer, F. J. (2007). *Biometals*, **20**, 291–302.

Benfatto, M., Congiu-Castellano, A., Daniele, A. & Della Longa, S. (2001). *J. Synchrotron Rad.* **8**, 267–269.

Binsted, N., Strange, R. W. & Hasnain, S. S. (1992). *Biochemistry*, **31**, 12117–12125.

Bohic, S., Murphy, K., Paulus, W., Cloetens, P., Salome, M., Susini, J. & Double, K. (2008). *Anal. Chem.* **80**, 9557–9566.

Cohen, A. E., Ellis, P. J., Miller, M. D., Deacon, A. M. & Phizackerley, R. P. (2002). *J. Appl. Cryst.* **35**, 720–726.

Corbett, M. C., Hu, Y., Fay, A. W., Ribbe, M. W., Hedman, B. & Hodgson, K. O. (2006). *Proc. Natl. Acad. Sci. USA*, **103**, 1238–1243.

Corbett, M. C., Latimer, M. J., Poulos, T. L., Sevrioukova, I. F., Hodgson, K. O. & Hedman, B. (2007). *Acta Cryst.* D**63**, 951–960.

Corbett, M. C., Tezcan, F. A., Einsle, O., Walton, M. Y., Rees, D. C., Latimer, M. J., Hedman, B. & Hodgson, K. O. (2005). *J. Synchrotron Rad.* **12**, 28–34.

Dau, H., Iuzzolino, L. & Dittmer, J. (2001). *Biochim. Biophys. Acta*, **1503**, 24–39.

Di Cicco, A. (2003). *J. Synchrotron Rad.* **10**, 46–50.

DiDonato, M., Hsu, H. F., Narindrasorasak, S., Que, L. Jr & Sarkar, B. (2000). *Biochemistry*, **39**, 1890–1896.

Di Marco, S., Rizzi, M., Volpari, C., Walsh, M. A., Narjes, F., Colarusso, S., De Francesco, R., Matassa, V. G. & Sollazzo, M. (2000). *J. Biol. Chem.* **275**, 7152–7157.

Dupont, C. L., Yang, S., Palenik, B. & Bourne, P. E. (2006). *Proc. Natl. Acad. Sci. USA*, **103**, 17822–17827.

Einsle, O., Andrade, S. L. A., Dobbek, H., Meyer, J. & Rees, D. C. (2007). *J. Am. Chem. Soc.* **129**, 2210–2211.

Ellis, M. J., Buffey, S. G., Hough, M. A. & Hasnain, S. S. (2008). *J. Synchrotron Rad.* **15**, 433–439.

Erskine, P. T., Duke, E. M. H., Tickle, I. J., Senior, N. M., Warren, M. J. & Cooper, J. B. (2000). *Acta Cryst.* D**56**, 421–430.
Farmer, P. J. & Sulc, F. (2005). *J. Inorg. Biochem.* **99**, 166–184.
Farquharson, M. J., Al-Ebraheem, A., Falkenberg, G., Leek, R., Harris, A. L. & Bradley, D. A. (2008). *Phys. Med. Biol.* **53**, 3023–3037.
Flank, A. M., Cauchon, G., Lagarde, P., Bac, S., Janousch, M., Wetter, R., Dubuisson, J. M., Idir, M., Langlois, F., Moreno, T. & Vantelon, D. (2006). *Nucl. Instrum. Methods Phys. Res. B*, **246**, 269–274.
Gao, Y. X., Chen, C. Y. & Chai, Z. F. (2007). *J. Anal. Atom. Spectrom.* **22**, 856–866.
Girard, E., Legrand, P., Roudenko, O., Roussier, L., Gourhant, P., Gibelin, J., Dalle, D., Ounsy, M., Thompson, A. W., Svensson, O., Cordier, M.-O., Robin, S., Quiniou, R. & Steyer, J.-P. (2006). *Acta Cryst.* D**62**, 12–18.
Grabowski, M., Joachimiak, A., Otwinowski, Z. & Minor, W. (2007). *Curr. Opin. Struct. Biol.* **17**, 347–353.
Gurusamy, K. S., Farquharson, M. J., Craig, C. & Davidson, B. R. (2008). *Biometals*, **21**, 373–378.
Hall, J. F., Ellis, M. J., Kigawa, T., Yabuki, T., Matsuda, T., Seki, E., Hasnain, S. S. & Yokoyama, S. (2005). *J. Synchrotron Rad.* **12**, 4–7.
Hall, M. D., Alderden, R. A., Zhang, M., Beale, P. J., Cai, Z., Lai, B., Stampfl, A. P. J. & Hambley, T. W. (2006). *J. Struct. Biol.* **155**, 38–44.
Harris, H. H., George, G. N. & Rajagopalan, K. V. (2006). *Inorg. Chem.* **45**, 493–495.
Harris, H. H., Levina, A., Dillon, C. T., Mulyani, I., Lai, B., Cai, Z. H. & Lay, P. A. (2005). *J. Biol. Inorg. Chem.* **10**, 105–118.
Harris, H. H., Pickering, I. J. & George, G. N. (2004). *Science*, **303**, 764–766.
Hasnain, S. S. & Hodgson, K. O. (1999). *J. Synchrotron Rad.* **6**, 852–864.
Hatada, K., Hayakawa, K., Benfatto, M. & Natoli, C. R. (2007). *Phys. Rev. B*, **76**, 060102.
Holm, R. H., Kennepohl, P. & Solomon, I. (1996). *Chem. Rev.* **96**, 2239–2314.
Holton, J. M. (2007). *J. Synchrotron Rad.* **14**, 51–72.
Hough, M. A., Antonyuk, S. V., Strange, R. W., Eady, R. R. & Hasnain, S. S. (2008). *J. Mol. Biol.* **378**, 353–361.
Isaure, M. P., Fraysse, A., Deves, G., Le Lay, P., Fayard, B., Susini, J., Bourguignon, J. & Ortega, R. (2006). *Biochimie*, **88**, 1583–1590.
Jakubowski, N., Ryszard, L. & Moens, A. L. (2004). *J. Anal. At. Spectrom.* **19**, 1–4.
Jenney, F. E. Jr, Brereton, P. S., Izumi, M., Poole II, F. L., Shah, C., Sugar, F. J., Lee, H.-S. & Adams, M. W. W. (2005). *J. Synchrotron Rad.* **12**, 8–12.
Joly, Y. (2003). *J. Synchrotron Rad.* **10**, 58–63.
Kleifeld, O., Frenkel, A., Martin, J. M. L. & Sagi, I. (2003). *Nat. Struct. Biol.* **10**, 98–103.
Kruger, P. & Natoli, C. R. (2004). *Phys. Rev. B*, **70**, 245120.
Lacapere, J.-J., Pebay-Peyroula, E., Neumann, J.-M. & Etchebest, C. (2007). *Trends Biochem. Sci.* **32**, 259–270.
Latimer, M. J., Ito, K., McPhillips, S. E. & Hedman, B. (2005). *J. Synchrotron Rad.* **12**, 23–27.
Lee, P. A., Citrin, P. H., Eisenberger, P. & Kincaid, B. M. (1981). *Rev. Mod. Phys.* **53**, 769.
Levina, A., Armstrong, R. S. & Lay, P. A. (2005). *Coord. Chem. Rev.* **249**, 141–160.
Lobinski, R., Moulin, C. & Ortega, R. (2006). *Biochimie*, **88**, 1591–1604.
Loftin, I. R., Franke, S., Blackburn, N. J. & McEvoy, M. M. (2007). *Protein Sci.* **16**, 2287–2293.
Lu, Y. (2006). *Inorg. Chem.* **45**, 9930–9940.
Maglio, O., Nastri, F., Martin de Rosales, R. T., Faiella, M., Pavone, V., DeGrado, W. F. & Lombardi, A. (2007). *C. R. Chim.* **10**, 703–720.
Magyar, J. S., Weng, T. C., Stern, C. M., Dye, D. F., Rous, B. W., Payne, J. C., Bridgewater, B. M., Mijovilovich, A., Parkin, G., Zaleski, J. M., Penner-Hahn, J. E. & Godwin, H. A. (2005). *J. Am. Chem. Soc.* **127**, 9495–9505.
Marino, S., Hayakawa, K., Hatada, K., Benfatto, M., Rizzello, A., Maffia, M. & Bubacco, L. (2007). *Biophys. J.* **93**, 2781–2790.
Natoli, C. R., Benfatto, M., Della Longa, S. & Hatada, K. (2003). *J. Synchrotron Rad.* **10**, 26–42.
Opella, S. J., DeSilva, T. M. & Veglia, G. (2002). *Curr. Opin. Chem. Biol.* **6**, 217–223.
Ortega, R., Cloetens, P., Devos, G., Carmona, A. & Bohic, S. (2007). *PLoS ONE*, **2**, e925.
Penner-Hahn, J. E. (2005). *Coord. Chem. Rev.* **249**, 161–177.
Pickering, I. J., Gumaelius, L., Harris, H. H., Prince, R. C., Hirsch, G., Banks, J. A., Salt, D. E. & George, G. N. (2006). *Environ. Sci. Technol.* **40**, 5010–5014.
Pickering, I. J., Wright, C., Bubner, B., Ellis, D., Persans, M. W., Yu, E. Y., George, G. N., Prince, R. C. & Salt, D. E. (2003). *Plant Physiol.* **131**, 1460–1467.
Pohl, E., Haller, J. C., Mijovilovich, A., Meyer-Klaucke, W., Garman, E. & Vasil, M. L. (2003). *Mol. Microbiol.* **47**, 903–915.
Pufahl, R. A., Singer, C. P., Peariso, K. L., Lin, S. J., Schmidt, P. J., Fahrni, C. J., Culotta, V. C., Penner-Hahn, J. E. & O'Halloran, T. V. (1997). *Science*, **278**, 853–856.
Pushkar, Y., Yano, J., Glatzel, P., Messinger, J., Lewis, A., Sauer, K., Bergmann, U. & Yachandra, V. (2007). *J. Biol. Chem.* **282**, 7198–7208.
Rehr, J. J. (2006). *Radiat. Phys. Chem.* **75**, 1547–1558.
Rehr, J. J. & Albers, R. C. (2000). *Rev. Mod. Phys.* **72**, 621.
Rehr, J. J. & Ankudinov, A. L. (2003). *J. Synchrotron Rad.* **10**, 43–45.
Rich, A. M., Armstrong, R. S., Ellis, P. J. & Lay, P. A. (1998). *J. Am. Chem. Soc.* **120**, 10827–10836.
Scheckel, K. G., Lombi, E., Rock, S. A. & McLaughlin, N. J. (2004). *Environ. Sci. Technol.* **38**, 5095–5100.
Scott, R. A., Shokes, J. E., Cosper, N. J., Jenney, F. E. & Adams, M. W. W. (2005). *J. Synchrotron Rad.* **12**, 19–22.
Shi, W., Zhan, C., Ignatov, A., Manjasetty, B. A., Marinkovic, N., Sullivan, M., Huang, R. & Chance, M. R. (2005). *Structure*, **13**, 1473–1486.
Smolentsev, G. & Soldatov, A. V. (2007). *Comput. Mater. Sci.* **39**, 569–574.
Solomon, E. I., Hedman, B., Hodgson, K. O., Dey, A. & Szilagyi, R. K. (2005). *Coord. Chem. Rev.* **249**, 97–129.
Sorini, A. P., Kas, J. J., Rehr, J. J., Prange, M. P. & Levine, Z. H. (2006). *Phys. Rev. B*, **74**, 165111–165118.
Sproviero, E. M., Gascon, J. A., McEvoy, J. P., Brudvig, G. W. & Batista, V. S. (2007). *Curr. Opin. Struct. Biol.* **17**, 173–180.
Strange, R. W., Ellis, M. & Hasnain, S. S. (2005). *Coord. Chem. Rev.* **249**, 197–208.
Strange, R. W., Smith, B. E., Eady, R. R., Lawson, D. & Hasnain, S. S. (2003). *J. Synchrotron Rad.* **10**, 197.
Thill, A., Zeyons, O., Spalla, O., Chauvat, F., Rose, J., Auffan, M. & Flank, A. M. (2006). *Environ. Sci. Technol.* **40**, 6151–6156.
Thomson, A. J. & Gray, H. B. (1998). *Curr. Opin. Chem. Biol.* **2**, 155–158.
Vila, F. D., Rehr, J. J., Rossner, H. H. & Krappe, H. J. (2007). *Phys. Rev. B*, **76**, 014301.
Winn, M. D. (2003). *J. Synchrotron Rad.* **10**, 23–25.
Yalovega, G., Smolentsev, G., Soldatov, A., Chan, J. & Stillman, M. (2007). *Nucl. Instrum. Methods Phys. Res. A*, **575**, 162–164.
Yang, F., Chu, W., Yu, M., Wang, Y., Ma, S., Dong, Y. & Wu, Z. (2008). *J. Synchrotron Rad.* **15**, 129–133.
Yano, J., Kern, J., Irrgang, K. D., Latimer, M. J., Bergmann, U., Glatzel, P., Pushkar, Y., Biesiadka, J., Loll, B., Sauer, K., Messinger, J., Zouni, A. & Yachandra, V. K. (2005). *Proc. Natl. Acad. Sci. USA*, **102**, 12047–12052.
Yano, J., Kern, J., Sauer, K., Latimer, M. J., Pushkar, Y., Biesiadka, J., Loll, B., Saenger, W., Messinger, J., Zouni, A. & Yachandra, V. K. (2006). *Science*, **314**, 821–825.
Yao, N., Reichert, P., Taremi, S. S., Prosize, W. W. & Weber, P. C. (1999). *Structure*, **7**, 1353–1363.
Yooseph, S. *et al.* (2007). *PLoS Biol.* **5**, 432–466.