Journal of
Synchrotron
Radiation

ISSN 0909-0495

Received 9 October 2003
Accepted 22 October 2004

Integration of XAS and NMR techniques for the structure determination of metalloproteins. Examples from the study of copper transport proteins

Lucia Banci,[a] Ivano Bertini[a] and Stefano Mangani[a,b]*

[a]Magnetic Resonance Center (CERM) and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy, and [b]Department of Chemistry, University of Siena, Via Aldo Moro 53100 Siena, Italy. E-mail: mangani@unisi.it

Nuclear magnetic resonance (NMR) is a powerful technique for protein structure determination in solution. However, when dealing with metalloproteins, NMR methods are unable to directly determine the structure of the metal site and its coordination geometry. The capability of X-ray absorption spectroscopy (XAS) to provide the structure of a metal ion bound to a protein is then perfectly suited to complement the process of the structure determination. This aspect is particularly relevant in structural genomic projects where high throughput of structural results is the main goal. The synergism of the two techniques has been exploited in the structure determination of bacterial copper transport proteins.

Keywords: metalloprotein; XAS; NMR; structure determination; copper transport; structural genomics.

1. Introduction

X-ray absorption spectroscopy (XAS) studies of biological systems have gained new impetus and have achieved relevant success in recent years (George *et al.*, 1998). A new application of XAS is to couple it with the structure determination of a metalloprotein in solution obtained by NMR methods. NMR is one of the techniques of choice for protein structure determination (Riek *et al.*, 2002; Wüthrich, 2001, 2003) and it is intensively used in structural genomics projects. The NMR procedure for solving protein structures relies on obtaining constraints about different structural parameters, like interproton distances, dihedral angles and interatomic vectors, from a variety of NMR experiments (Wüthrich, 1995, 2001; Bertini, Luchinat & Rosato, 1996; Wider & Wüthrich, 1999). The constraints provide upper limits to the values of the above parameters, which can be used in distance–geometry calculations that eventually determine the protein structure. When dealing with metalloproteins, however, no classical constraints can be obtained about the metal center and NMR is not able to provide its structure. Often the metal ion is NMR silent. If not, HETCOR spectroscopy provides the group which is bound to the metal (Bertini, Luchinat & Rosato, 1996; Bertini & Luchinat, 1999; Bertini, Donaire *et al.*, 1996). Only when a paramagnetic ion is bound to a protein is it possible to place it within the protein frame with an accuracy of about 1–2 Å (Bertini, Donaire *et al.*, 1996; Bertini & Luchinat, 1999; Bertini, Luchinat & Rosato, 1996; Bertini *et al.*, 2001). The analysis of the EXAFS region of the X-ray absorption spectrum gives the complete local structure of a metal ion bound to a protein, namely the type and number of metal ligands and extremely accurate bond distances (within 0.02 Å or better). Further information about the electronic structure of the metal ion and its coordination geometry can be obtained from the analysis of the edge and pre-edge regions of the spectrum. In this way, the EXAFS data provide the missing constraints about the metal center, which can then be used to complete the structure determination by NMR. Alternatively, NMR can be used to determine the structure of the apoprotein and XAS provides the metal site structure with most, if not all, of the metal ligands. By comparison of results, a quite accurate model of the holo-protein might be determined.

The outlined approach allows the entire process of structure determination in solution to be speeded up and, in addition, it provides unique information on the electronic structure of the metal ion in the studied system. We have exploited this synergism between the NMR and XAS techniques for a rapid and complete elucidation of the structure of proteins involved in copper trafficking in bacteria. Examples taken from our work on the copper transport proteins CopC from *Pseudomonas syringae* and CopZ from *Bacillus subtilis* are presented below to illustrate the advantages of the combined techniques.

2. The solution structure of CopC

The import and delivery of Cu(I) within the boundaries of the bacterial cytosol is strongly dependent on the use of Cu(I) transporters and metallo-chaperones. This general mechanism is useful and important for handling potentially toxic metal ions in environments rich in reductants as the cytosol (O'Halloran & Culotta, 2000; Rosenzweig & O'Halloran, 2000; Puig & Thiele, 2002). However, copper is used and also transported in oxidizing cellular compartments like the periplasmic space of Gram-negative bacteria where it is expected to predominantly exist in the Cu(II) oxidation state. Such an oxidizing environment could pose a serious threat to the organism if redox-active metals such as copper are freely engaged in redox chemistry that generates reactive oxygen species.

We have started the study of the copper transport in Gram-negative bacteria by using the model constituted by copper-resistant strains present in copper-contaminated soils like *Pseudomonas syringae* pathovar (pv.) *tomato*.

The copper-resistant strains of the bacterium *Pseudomonas syringae* are characterized by bearing a 35 kb plasmid which encodes for four proteins (CopABCD) involved in the control of the copper homeostasis (Puig & Thiele, 2002). Among these proteins, CopC is located in the periplasmic space of the bacterium and is able to bind copper with the probable function of shuttling the toxic metal ion in and out of the cytoplasm depending on the metal concentration in the cell.

The solution structure of the metal-free form of CopC has been obtained by NMR methods (Arnesano *et al.*, 2002) and the protein shows a greek-key β-barrel fold similar to that adopted by blue copper proteins like plastocyanin, which are involved in electron-transfer processes and host a type-I copper site (Guss & Freeman, 1983). In contrast, CopC was shown to be able to bind copper(II) in a type-II binding site probably identified in a $His_2GluAsp$ site. A second possible copper(I) binding site is consituted by a Met-rich region present in CopC (Arnesano *et al.*, 2002). In order to quickly identify the nature of the copper binding sites in CopC, an X-ray absorption study of this protein was started.

Two samples were prepared under nitrogen atmosphere from two 3.0 m*M* apo-CopC solutions in 100 m*M* phosphate buffer at pH 7.0; one equivalent of Cu(II) (as $CuSO_4$) was added to each sample. In one of the samples, Cu(II) was reduced by addition of one equivalent of sodium ascorbate. XAS data were collected at DESY (Hamburg, Germany) at the EMBL bending-magnet beamline D2. The experimental and data-analysis procedures have already been reported (Arnesano, Banci, Bertini, Mangani & Thompsett, 2003). Fig. 1 shows the edge region of the X-ray absorption spectra taken from the two samples while Fig. 2 reports the EXAFS with the respective Fourier transforms.

These data illustrate how powerful the XAS might be in helping to understand the chemical behavior of a metal ion. A simple visual inspection of the two EXAFS spectra clearly indicates that the Cu(I) and Cu(II) ions are bound to different binding sites present in the protein. The data analysis shows that Cu(II) is bound to two histidine residues (1.99 Å) and two

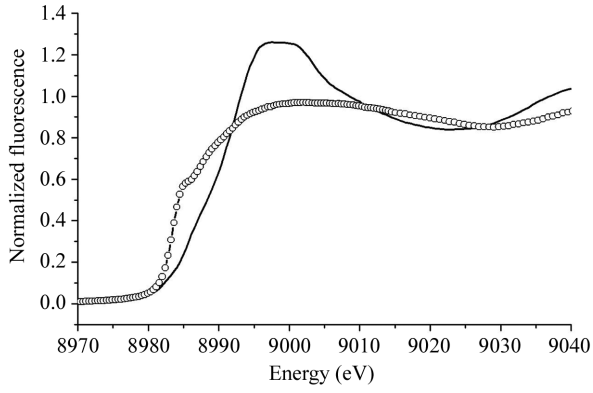

Figure 1
Comparison of the Cu *K*-edge regions of Cu(II)–CopC (solid line) and Cu(I)–CopC (open circles).

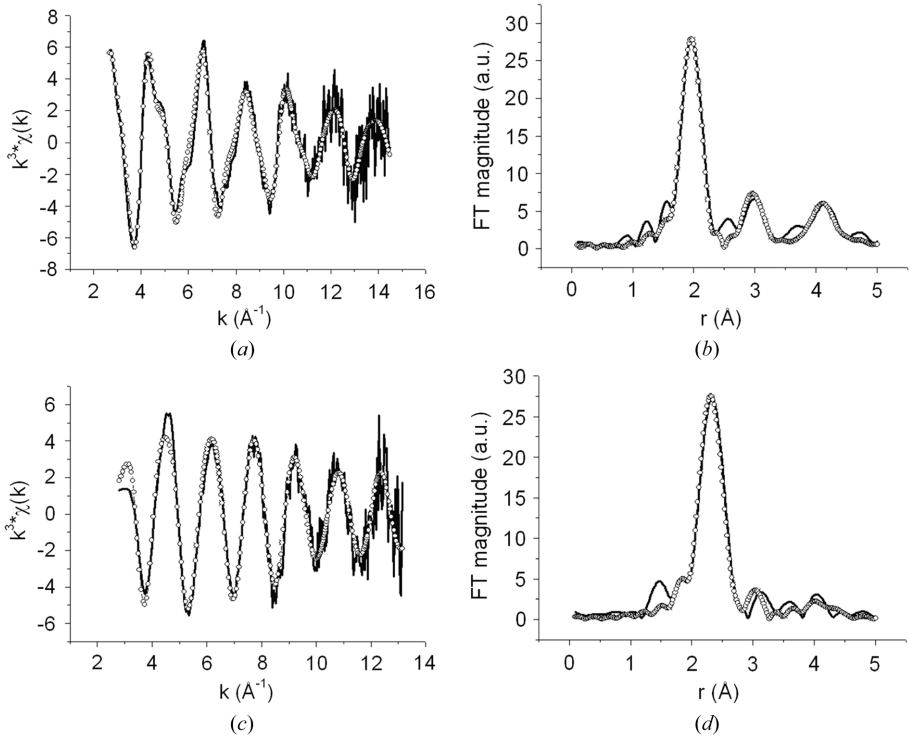

Figure 2
Experimental (solid line) and simulated (open circles) EXAFS spectra (*a*) and their Fourier Transforms (*b*) of Cu(II)–CopC, and the corresponding data of Cu(I)–CopC, (*c*) and (*d*), respectively.

oxygen ligands at short distance (1.97 Å), and that the coordination is completed by one/two further light (O/N) ligands, most probably from water molecules at 2.83 Å. In the other sample the Cu(I) ion is bound to three S atoms at 2.30 Å and one histidine at 1.95 Å. By applying this information to the structure of the apo-CopC, the location of the binding sites in the protein becomes evident since only two sites are present in CopC which fulfill the requirements obtained from the EXAFS analysis. Cu(II) is bound to the suggested site located at the N-terminus of CopC (Fig. 3, in red), while Cu(I) is bound to the opposite end of the β-barrel where the Met-rich region is present (Fig. 3, in blue). It is worth noting that the two copper binding sites found in CopC are two novel Cu(I) and Cu(II) binding motifs.

Furthermore, the XAS experiment has shown in a very simple and elegant way that the same copper ion migrates from one site to the other depending on the copper oxidation state. This finding discloses a new functional aspect of CopC that may act as a redox-driven switch in copper transport (Arnesano, Banci, Bertini, Mangani & Thompsett, 2003). The application of the constraints obtained from EXAFS has later provided valuable help in the structure determination of the paramagnetic Cu(II)-bound form of CopC (Arnesano, Banci, Bertini, Felli *et al.*, 2003).

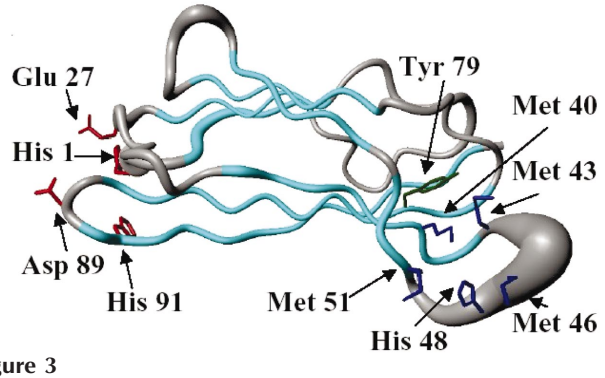

Figure 3
(*a*) Solution structure of Cu(I)–CopC represented as a tube whose radius is proportional to the backbone RMSD of each residue. β-strands are colored in cyan. Conserved residues in the Cu(I) binding site are shown in blue. Conserved residues in the Cu(II) binding site are shown in red. Tyr 79 is shown in green. The figure was generated with *MOLMOL* (Koradi *et al.*, 1996).

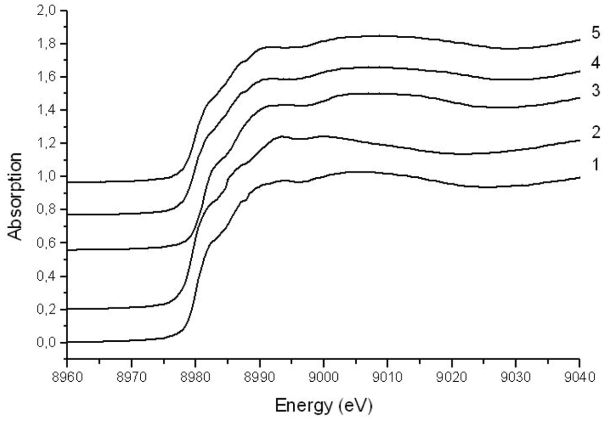

Figure 4
Comparison of the Cu *K*-edge regions of (1) DTT-reduced 3 m*M* Cu(I)–CopZ; (2) ascorbate-reduced Cu(I)–CopZ; (3) DTT-reduced 3 m*M* Cu(I)–CopZ + reduced glutathione; (4) DTT-reduced 3 m*M* Cu(I)–CopZ + 6 m*M*Na-citrate; (5) DTT-reduced 3 m*M* Cu(I)–CopZ + Na-acetate.

3. The solution structure of CopZ

B. subtilis CopZ is a cytosolic small protein (about 8 kDa) involved in the copper homeostasis of this Gram-positive bacterium. CopZ brings the MXCXXC sequence motif which is characteristic of other Cu(I) binding proteins like the human Hah1, the yeast Atx1 and plant Cch (Harrison *et al.*, 2000). The NMR-XAS analysis of the Cu(I)–CopZ adduct conducted under different conditions, where parameters like ionic strength, different reducing agents and exogenous potential ligands were varied, has allowed us to obtain new information about the chemistry of CopZ (Banci, Bertini, Del Conte *et al.*, 2003). As an example, Fig. 4 reports the edges of some of the CopZ samples where the different Cu(I) coordination environments are evident. Figs. 5 and 6 compare the behaviour of Cu(I)–CopZ solutions obtained from different copper reducing agents like DTT (dithiotreithol) and sodium ascobate. The two spectra provide direct evidence that, besides the two cysteine sulfurs provided by the protein, an exogenous ligand has entered the Cu(I) coordination sphere. In the case of the DTT-reduced sample, the ligand is a thiolate sulfur, most probably from the DTT molecule itself, whereas

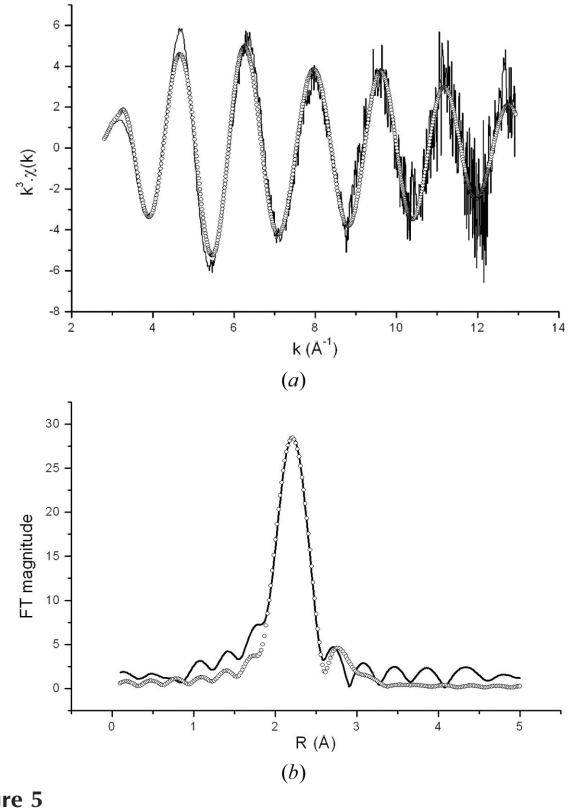

Figure 5
(*a*) EXAFS spectrum and (*b*) Fourier transform of DTT-reduced 3 m*M* CopZ solution. The best fit (open circles) and the experimental (continuous line) data are reported.

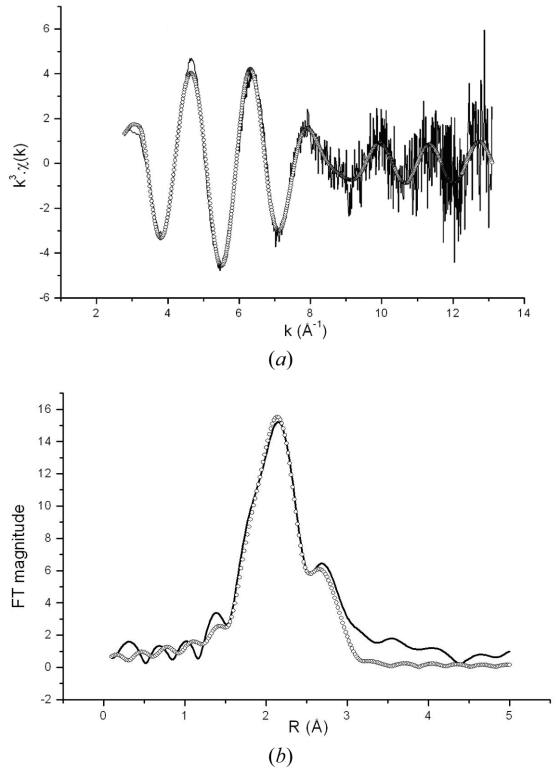

Figure 6
(*a*) EXAFS spectrum and (*b*) Fourier transform of ascorbate-reduced 3 m*M* CopZ solution. The best fit (open circles) and the experimental (continuous line) data are reported.

in the case of the ascorbate-reduced sample an oxygen donor is found bound to copper. The presence of an exogenous metal-bound thiol in the DTT-reduced Cu(I)–CopZ has been supported by ^{15}N HSQC NMR measurements of the sample after exhaustive washing of the reductant from the protein solution with buffer. Under these conditions a different CopZ complex is formed, and only after the reintroduction of DTT into the sample is the original DTT–Cu(I)–CopZ NMR spectrum obtained (Banci, Bertini, Del Conte *et al.*, 2003).

The EXAFS analysis of the ascorbate-reduced sample also shows that the prevalent Cu(I)–CopZ species present in solution is a dimeric form which involves a close outer-shell Cu—Cu interaction (2.81 Å) as evidenced by the peak present in the spectrum Fourier transform at this distance (see Fig. 6*b*). The tendency of CopZ to form dimers has also been observed in solutions where an excess of Na-citrate or Na-acetate was present. NMR measurements of ^{15}N longitudinal and transversal relaxation rates have shown that the dimerization of CopZ also depends on the protein concentration, but in this case different aspecific dimers are formed which do not involve Cu—Cu interactions. The CopZ dimerization through the copper-binding loop may reflect a functional aspect of the protein relevant for copper transfer. Indeed, CopZ forms adducts with the partner CopA protein as demonstrated by the solution structure of the CopZ–CopA complex (Banci, Bertini, Ciofi-Baffoni *et al.*, 2003).

The above results suggest that physiological thiols may be needed *in vivo* to complete the Cu(I) coordination sphere in order to prevent homodimer formation. A saturated copper coordination sphere, like the Cu(I)–S_3 or Cu(I)–(S_2O) forms, may be the stable adduct needed for the safe copper transport to the partner proteins which avoids the formation of possible dead-end products like a homodimer. Our results also suggest that such an arrangement is not thermodynamically as stable as to prevent the interaction with the partner protein when appropriate conditions are met.

This work was supported by the PRIN2003 from the Italian MIUR and by the European Community (SPINE no. QLG2-CT-2002-00988, 'Structural Proteomics in Europe'). The authors also acknowledge funding from the EU for support of the work at EMBL Hamburg through the HPRI-CT-1999-00017 grant.

References

Arnesano, F., Banci, L., Bertini, I., Felli, I. C., Luchinat, C. & Thompsett, A. R. (2003). *J. Am. Chem. Soc.* **125**, 7200–7208.
Arnesano, F., Banci, L., Bertini, I., Mangani, S. & Thompsett, A. R. (2003). *Proc. Natl. Acad. Sci. USA*, **100**, 3814–3819.
Arnesano, F., Banci, L., Bertini, I. & Thompsett, A. R. (2002). *Structure*, **10**, 1337–1347.
Banci, L., Bertini, I., Ciofi-Baffoni, S., Del Conte, R. & Gonnelli, L. (2003). *Biochemistry*, **42**, 1939–1949.
Banci, L., Bertini, I., Del Conte, R., Mangani, S. & Meyer-Klaucke, W. (2003). *Biochemistry*, **42**, 2467–2674.
Bertini, I., Donaire, A., Felli, I. C., Luchinat, C. & Rosato, A. (1996). *Magn. Reson. Chem.* **34**, 948–950.
Bertini, I. & Luchinat, C. (1999). *Curr. Opin. Chem. Biol.* **3**, 145–151.
Bertini, I., Luchinat, C. & Parigi, G. (2001). *Solution NMR of Paramagnetic Molecules.* Amsterdam: Elsevier.
Bertini, I., Luchinat, C. & Rosato, A. (1996). *Prog. Biophys. Mol. Biol.* **66**, 43–80.
George, G. N., Hedman, B. & Hodgson, K. O. (1998). *Nature Struct. Biol. Suppl.* **5**, 645–647.
Guss, J. M. & Freeman, H. C. (1983). *J. Mol. Biol.* **169**, 521–563.
Harrison, M. D., Jones, C. E., Solioz, M. & Dameron, C. T. (2000). *Trends. Biochem. Sci.* **25**, 29–32.
Koradi, R., Billeter, M. & Wüthrich, K. (1996). *J. Mol. Graph.* **14**, 51–55.
O'Halloran, T. V. & Culotta, V. C. (2000). *J. Biol. Chem.* **275**, 25057–25060.
Puig, S. & Thiele, D. J. (2002). *Curr. Opin. Chem. Biol.* **6**, 171–180.
Riek, R., Fiaux, J., Bertelsen, E. B., Horwich, A. L. & Wüthrich, K. (2002). *J. Am. Chem. Soc.* **124**, 12144–12153.
Rosenzweig, A. C. & O'Halloran, T. V. (2000). *Curr. Opin. Chem. Biol.* **4**, 140–147.
Wider, G. & Wüthrich, K. (1999). *Curr. Opin. Struct. Biol.* **9**, 594–601.
Wüthrich, K. (1995). *Acta Cryst.* D**51**, 249–270.
Wüthrich, K. (2001). *Nature Struct. Biol.* **8**, 923–925.
Wüthrich, K. (2003). *Angew. Chem. Int. Ed. Engl.* **42**, 3340–3363.