Cell. Mol. Life Sci. 65 (2008) 3040 – 3048
1420-682X/08/193040-9
DOI 10.1007/s00018-008-8189-9

Cellular and Molecular Life Sciences

Review

Metallomics and metalloproteomics

W. Shi and M. R. Chance*

Center for Synchrotron Biosciences, Case Center for Proteomics, Case Western Reserve University, 10900 Euclid Ave, BRB 113, Cleveland, Ohio 44106 (USA), Fax: +1 216 368 3812, e-mail: mark.chance@case.edu

Received 8 April 2008; received after revision 12 May 2008; accepted 15 May 2008
Online First 14 June 2008

Abstract. Metallomics and metalloproteomics are emerging fields addressing the role, uptake, transport and storage of trace metals essential for protein functions. The methodologies utilized in metallomics and metalloproteomics to provide information on the identity, quantity and function of metalloproteins are discussed. The most widely used approach is through inductively coupled plasma mass spectrometry to identify the metal bound to a protein, and electrospray ionization mass spectrometry to elucidate the structure, dynamics and function of a metal-protein complex. Other approaches include X-ray absorption and X-ray fluorescence spectroscopies, and bioinformatics sequence analysis. X-ray absorption spectroscopy utilizing a synchrotron radiation source is a powerful tool to provide a direct analysis of metal bound to proteins and proteomic metal distribution in biological matrices. With the advent of genome sequencing, a large database of protein primary structures has been established, and specific tools to identify metalloproteins in the genome sequences have been developed.

Keywords. Metallomics, metalloproteomics, metal, metalloproteins, mass spectrometry, X-ray absorption spectroscopy, computational modeling, structural genomics.

Introduction

About one-third of all proteins are associated with a metal. The function of these proteins, termed metalloproteins, depends on the interaction between the proteins and the bound metals [1, 2]. Examination of the Protein Data Bank (PDB, http://www.rcsb.org) shows that Zn is the most abundant, while Fe, Mg and Ca are also frequently observed [3]. Metalloproteins are one of the most diverse classes of the proteins, with the intrinsic metal atoms providing a catalytic, regulatory and structural role critical to protein function [2]. Transition metals such as Cu, Fe and Zn play important roles in life [4–6]. Zn, the most abundant transition metal in cells, plays a vital role in the functionalities of more than 300 enzymes, in the stabilization of DNA and in gene expression [7]. Other trace metals, such as Se, W and Mo, are essential in human health and also important in the environment. For example, Se is usually incorporated into antioxidant enzymes as selenocysteine (SeCys), the redox active site of SeCys-containing enzymes, and plays a key role in host oxidative defense. It is predicted that there are a total of 25 SeCys-containing proteins in humans [8–10]. The most widely known SeCys-containing enzymes are thioredoxin reductase and glutathione peroxidase. Both are ubiquitous and found in bacteria, plants and mammals, including humans [11].

In the past decade, there has been tremendous development in the fields of biology that end in -*omics*. The best-known discipline among them is

* Corresponding author.

genomics, which produces complete genome DNA sequences of a living organism. Other well-developed disciplines include structural genomics and proteomics. Structural genomics aims to generate high-resolution, three-dimensional structural models for macromolecules at genomic scale. Hundreds of protein structures are determined and deposited in the PDB yearly through several programs that are currently funded in North America, Europe and Japan [12–14]. Proteomics provides information about protein localization, structure and function, and most important, interactions with other proteins. Recent improvements in high-throughput sample separation and mass spectrometry impact positively on the proteomic characterization of proteins in systems biology [15, 16]. Metallomics and metalloproteomics are emerging fields addressing the role, uptake, transport and storage of trace metals essential for life. Metallomics is defined as analysis of the entirety of metal and metalloid species within a cell or a tissue type, whereas metalloproteomics is focused on exploration of the function of metals associated with proteins [17–19]. Compared to genomics and proteomics, metallomics and metalloproteomics are relatively new fields; however, genomics and proteomics have constructed a large amount data that can be utilized in metallomics and metalloproteomics studies to achieve rapid advancement.

There are three main approaches that are being developed in metallomics and metalloproteomics (Fig. 1). The first is the widely used analytic techniques of mass spectrometry, in particular, electrospray ionization mass spectrometry (ESI-MS) and inductively coupled plasma mass spectrometry (ICP-MS). These two techniques are ideal partners in the comprehensive structural and functional characterization of metalloproteins [20, 21]. The second approach is high-throughput X-ray absorption spectroscopy (HT-XAS) to provide a direct metal analysis of proteins and proteomic metal distribution in tissues and cells [3, 22]. The third approach is through computational bioinformatics analysis. With the advent of genome sequencing, a large database of protein primary structures has been established. At present, specific tools to identify metalloproteins in genome sequences have also been developed. Metal binding properties can be predicted based on amino acid sequence by taking into account known consensus sequences in a metal binding region [23, 24]. The methods have been applied in predicting Zn, Cu and other trace metals in proteins [4, 6, 10].

This review is focused on advances in the methodologies used in metallomics and metalloproteomics. In addition, application of different techniques synergically from the same field (i.e. ICP-MS and ESI-MS) and across fields (i.e. HT-XAS and bioinformatics) to structural and functional characterization of metalloproteins are also discussed here. We believe that synergic approaches will be the most important step in achieving a comprehensive study of metalloproteins in the future.

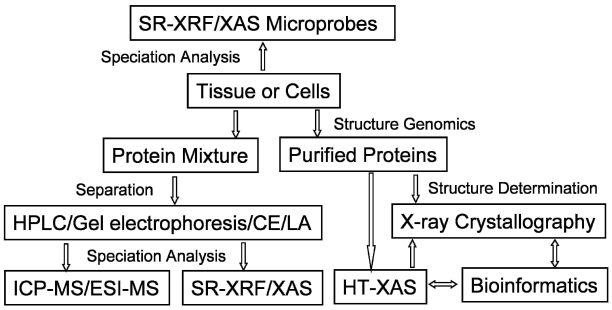

Figure 1. Metalloproteomics and metallomics flow chart. Organ, tissue and cells are subjected to speciation analysis using synchrotron radiation X-ray fluorescence microprobes. Protein mixtures can be separated using various techniques and subjected to speciation analysis using mass spectrometric techniques or synchrotron radiation X-ray fluorescence or X-ray absorption spectroscopies. Purified proteins through structure genomics pipeline are subjected to high-throughput X-ray absorption spectroscopy analysis for identification of metalloproteins. Bioinformatics analysis can identify metal binding sites based on primary structure of proteins. Synergic use of HT-XAS, sequence analysis and comparative modeling provides a comprehensive analysis of metalloproteins and can increase the throughput of structure determination in structure genomics.

Mass spectrometry approaches

ICP-MS

Quantitative determination of trace metals in biological systems can be achieved by using ICP-MS. This is a robust and highly sensitive technique that is capable of the quantitative detection of a wide range of metals and several non-metals, such as sulfur, phosphorus and iodine, at low concentrations [18]. It is based on coupling together an inductively coupled plasma as an ionization method with a mass spectrometer to separate and detect the ions. Nebulized sample enters the ICP collision cell and breaks down into atoms. A significant proportion of the atoms are ionized, and the positive ions are extracted into a mass spectrometer. Thus, the elements present in the sample can be identified on the basis of the atomic mass spectra generated from the mass spectrometer. ICP-MS detects the element regardless of its molecular environment and can quantitatively determine the amounts of elements in biological systems by using inorganic element standards. Unlike atomic absorption spectroscopy, which can only measure a

single element at a time, ICP-MS has the capability to detect multiple elements simultaneously.

ICP-MS has been used in speciation analysis by coupling with a front-end chromatograph separation technique. High-performance liquid chromatography (HPLC) is a well-established, reliable and reproducible separation technique for a wide range of samples. ICP-MS coupled with different types of HPLC, particularly multi-dimentional capillary and nano-flow HPLC, have been discussed in recent reviews [25, 26]. Other hyphenated ICP-MS techniques for identifying and quantifying metal binding proteins include preparative native continuous polyacrylamide gel electrophoresis (QPNC-PAGE), capillary electrophoresis (CE) and gel electrophoresis laser ablation (LA) Isolation of metalloproteins with non-covalent bound metal requires a mild, non-denaturing condition. Treatment of metalloprotein with denaturing SDS gel electrophoresis will result in a loss of metal. The traditional detection of heteroatom-containing protein in gels is through the use of radioactive isotopes [27]. A cheaper alternative way to scan gels for the presence of a metal is by laser ablation, where a focused laser beam is used for evaporation of solid sample [28]. The laser ablation can be carried out in gel and the ablated sample is directly blown into the ICP collision cell by a stream of Ar gas. Laser ablation is a simple, clean way of transferring sample from gel to ICP-MS and offers a reliable method in quantitative detection of trace elements. LA-ICP-MS has been applied successfully to detect selenium-containing protein and metal-protein complexes in protein bands [29–32]. The method is particularly useful in detection of metal-protein complexes that are formed in response to metal stress because of their abundance [33]. LA-ICP-MS has also been proposed to map the distribution of Cu, Zn and other elements in thin sections of human brain tissue [34, 35]. The limitations in using LA-ICP-MS in such applications is difficulties in the laser ablation of biological matrices, the lack of suitable quantification procedures and the high cost of obtaining a powerful laser ablation system with sufficient spatial resolution [35].

Capillary electrophoresis (CE) is a high-resolution separation technique useful in metalloproteomics research. Application of CE coupling with ICP-MS was reviewed recently, and different modes of capillary electrophoresis for different classes of molecules (peptides, proteins and DNA) were discussed [36]. The principle of separation in CE is based on the charges and molecular masses of the molecules. Separation of protein samples with CE requires a minimum amount of sample as little as 0.2 nL. Choice of a proper buffer system/condition and control of other factors such as temperature are essential to have a successful separation and reproducible outcome. Commercial availability and efficient design of the nebuliser and interface of CE-ICP-MS makes CE-ICP-MS highly sensitive and applicable in metalloproteomics research. Most CE-ICP-MS applications are focused on metallothionein (MT) isoform characterization [37, 38]. MT isoform distribution in brain tissue from patients with Alzheimer's disease was analyzed with the method [37]. Quantitative characterization of MT isoforms in term of Cu:Zn:Cd stoichiometry were carried out. Other CE-ICP-MS applications include determinations of Zn:S stoichiometry in carbonic anhydrase and Cu:Zn:S ratio in superoxide dismutase [39].

ESI-MS

ESI-MS is a molecular MS technique to determine metal binding selectivity and stoichiometry. In principle, eletrospray is a simple and gentle method for the ionization of polar molecules. Sample dissolved in appropriate solvent is pumped through a thin capillary with a high potential. Small charged droplets are sprayed from the ES capillary into a bath gas at atmospheric pressure and travel down a potential gradient towards the mass spectrometer vacuum system. As the droplets fly through the passage, solvent is extracted from the droplets, which break up into smaller droplets as well as single or multiple charged molecules. The mild method of ionization makes ESI-MS an important technique in biological studies, which often require that non-covalent metal-protein or protein-protein interactions remain intact during an experiment. The major drawback of the technique is that it requires volatile salt buffers such as ammonium acetate and is not tolerant of other non-volatile salt buffers or other reagents that might be necessary to keep the metal-protein complex soluble and properly folded. Size exclusion chromatography or a capillary electrophoresis may be necessary to exchange the protein storage buffer to a buffer system suitable for ESI-MS experiments [18].

Non-covalently bound species such as metal cofactors can be observed in ESI-MS spectra when the interactions between metal and protein are sufficiently strong to survive disruption. ESI-MS has been used to study interactions between identical subunits in protein oligomer and in metal-protein complexes. It is easy to detect metal binding to protein by ESI-MS. The comparison of the molecular mass measured under denaturing conditions and the mass measured under non-denaturing conditions gives directly the mass of possibly bound metal. Large protein complexes can readily be observed by ESI-MS. Yeast alcohol dehydrogenase (ADH) is an example of a large oligomeric metal-protein complex. The non-

covalently bound tetrameric complex of yeast ADH (MW = 147 000) has been observed by measuring ESI-MS spectra under non-denaturing conditions [40]. The tetrameric species consists not only of four protein subunits, but also eight non-covalently bound Zn^{2+} ions. The experiment provides a mass accuracy sufficient to determine the metal content.

Application of ESI-MS in the structural and fuctional characterization of metalloproteins was recently reviewed [41]. Examples of the unique features and ability of ESI-MS were addressed in terms of direct characterization of metal-protein interactions, determination of metal binding site components and the oligomerization state of protein complexes. Determination of metal binding site constituents can be achieved by chemical modification of the possible metal binding amino acids followed by enzymatic digestion and ESI-MS analysis [42, 43]. ESI-MS in conjunction of hydrogen deuterium exchange and enzymatic digestion techniques can be used to address the dynamics of metalloproteins in terms of the conformational changes upon binding of a metal cofactor. The approach has been successfully applied to several metalloproteins, including Fe^{3+} binding of ferric uptake regulatory protein and Ca^{2+} binding of calmodulin [44, 45].

Similar to ICP-MS, ESI-MS can also be coupled with a separation technique such as HPLC, CE and two-dimensional gel electrophoresis (discussed in detail later) to be used in identification of metalloproteins in metalloproteomics and metallomics studies [36, 38, 46]. For example, isoforms of metallothionein were identified and quantified from liver extracts based on their Cu, Zn and Cd stoichiometry using online and offline HPLC or CE-ESI-MS methods [46, 47].

Complementary studies of metalloprotein with atomic and molecular mass spectrometry

Using ICP-MS and ESI-MS synergically to approach metalloproteomics studies was recently proposed [20, 21]. ICP-MS is a sensitive and quantitative atomic mass spectrometry method that provides information about the metals bound to proteins, whereas ESI-MS is a molecular mass spectrometry to elucidate the metal selectivity, stoichiometry and metal-protein complex structure. In theory, the two techniques are well matched and can offer complementary approaches in a comprehensive metalloprotein research to obtain information about the metal bound to the proteins as well as the structure, dynamics and function of a metal-protein complex. However, in practice, the complementary use of the two techniques is still limited by the low concentration of metalloproteins in a given biological sample and the lack of accuracy in molecular mass determination of a large metal-protein complex to derive the metal stoichiometry by ESI-MS.

The synergic approach using ICP-MS and ESI-MS has been applied successfully to the characterization of metallothionein. In 2000, Mounicou et al. used the combination of CE-ICP-MS and CE-ESI-MS to identify the MT isoforms and determine metal stochiometries of the different isoforms in commertially available MT preps [48]. The similar complementary approach was used later to characterize the MT isoforms from liver extracts based on the Cu : Zn : Cd stoichiometric ratio [46, 47].

Proteomic technique in metallomics and mealloproteomics

Proteomic technology can be used to identify and classify metalloproteins based on their structure and function at the cellular level [49]. Techniques in quantitative proteomics such as two-dimensional gel electrophoresis (2-D gel) and liquid chromatography (LC) linked to mass spectrometry are powerful tools in metalloprotein separation and identification. The total soluble protein mixture from a particular organ, tissue or cell line can be separated using 2-D gel technique based on the differences in isoelectric point and molecular mass. The gel spots with individual proteins were extracted for enzymatic digestion by protein proteases into peptide fragments, and the mixture was analyzed by mass spectrometry. The metalloprotein-specific technique immobilized metal affinity chromatography (IMAC) is designed to isolate metalloproteins [50]. IMAC, combined with 2-D gel technique, has been utilized to study metal-related diseases such as Wilson disease, which causes Cu accumulation in the liver. The human hepatoma cell line, Hep G2, which preserves many basic characteristics of normal liver cells, was selected for these studies. Sarkar and colleagues used the 2-D gel/IMAC method to separate and identify 67 highly abundant Cu and Zn binding proteins in Hep G2, including albumin and enolases that are known to bind metal [51, 52]. They also discovered several proteins with previously unknown metal binding functions, including protein disulfide isomerase and peroxiredoxin. These proteins may have implications in Wilson disease.

A combinational approach of 2-D gel and LC with ESI-MS was used to study ferric uptake regulation in *Shewanella oneidensis*, a useful bacterium known for its potential to serve as a nuclear waste clean-up agent [53]. Studies using a similar approach were also carried out to investigate the metal reduction properties of *Geobacter sulfurreducens*, another bacterium which possesses promising nuclear waste clean-up capabilities [54].

There are many metal-related diseases caused by genetic disorders of metal metabolism involving transition metals such as Fe, Zn and Cu. These diseases have been reviewed, and the approaches of quantitative proteomics and metallomics/metalloproteomics will be highly useful to identify proteins implicated in these diseases and to reveal the structure and function of these proteins under normal and diseased states [49].

Synchrotron radiation XAS and XRF approaches

High-throughput X-ray absorption spectroscopy

The application of synchrotron radiation-based high-throughput X-ray absorption spectroscopy (HT-XAS) in the structural characterization of metalloproteins has been proposed [3, 12, 55]. The HT-XAS approach utilizes the same physical principle of excitation of metal atoms as the X-ray absorption spectroscopies, i.e. Extended X-ray absorption fine structure (EXAFS) and X-ray absorption near-edge spectroscopy (XANES). X-ray photons of appropriate energy, produced by a synchrotron source, are capable of ejecting electrons from the first electron shell (1 s) surrounding the nucleus of a metal atom into the continuum, thus leaving the atom in an excited state. To stabilize the atom, a passage of another electron from a higher-energy shell fills the 'hole', and the energy difference between the binding electron energies of the two orbitals is released through X-ray fluorescence emission. The characteristics of fluorescence for each element are known. HT-XAS relies on the detection of a fluorescence signal to identify the metal present in the sample. The method can be quantitative by using inorganic metal compounds as standards.

Published work utilizing HT-XAS in metalloproteomics so far has all been closely associated with structural genomics. The aim of structural genomics is to experimentally determine 3-D structure for at least one protein from each of the protein families, thus providing comparative modeling templates for the other members of the families. Targets are selected based on strategy, and proteins are cloned, expressed, purified and their structure determined by X-ray crystallography or nuclear magnetic resonance (NMR) spectroscopy [13, 14]. The success rate from purified protein to the final experimental 3-D structures in the large structural genomics centers in the Protein Structure Initiative are ~10–15% [56]. However, the large set of purified proteins represents a useful resource for other biochemical and biophysical characterization of proteins.

An initial test of a high-throughput XAS methodology was carried out at the Stanford Synchrotron Radiation Laboratory using purified gene product samples from *Pyrococcus furiosus*, and distributions of Co, Ni, Cu and Zn were mapped [55]. The work was done in collaboration with the Southeast Collaboratory for Structural Genomics (SECSG), where a large fraction of 2200 open reading frames in *Pyrococcus furiosus* have been cloned and purified by high-throughput techniques for structure determination for X-ray crystallography or NMR [57]. Small amounts of samples (3 μl) from a test sample set were loaded on a 25-well sample holder and frozen in liquid nitrogen. Metal distribution was determined by recording X-ray fluorescence emission from multiple elements using a multi-channel solid-state fluorescence detector. Distributions of Co, Ni, Cu and Zn were determined, and the information can be used for further analysis by XANES and EXAFS.

An HT-XAS analysis using a large set of proteins and quantitative analysis of six metals (Mn, Fe, Co, Cu, Ni and Zn) was reported by our laboratory [3, 12]. Protein samples were purified through the high-throughput protein production pipeline from the New York Structural Genomics Research Consortium (NYSGXRC). Approximately 200 μg of protein were loaded in the 16-well sample holders and dried overnight, and each sample was screened for the 6 metals. 654 samples from the NYSGXRC were analyzed, and over 10% of total samples showed the presence of transition metal atoms in stoichiometric amounts. The method was shown to be ~94% accurate in predicting the presence or absence of a transition metal based on 50 crystal structures from the sample set. Bioinformatics functional annotation was carried out for the metalloproteins identified by the HT-XAS method. In many cases, the metal binding information provides a distinct, new annotation for protein with unknown function and an improved annotation for proteins with vaguely understood functions.

In X-ray absorption spectroscopy, XANES provides information about the oxidation state and symmetry around the absorbing atom, while EXAFS provides precise distance information, as well as coordination numbers and the chemical identity of neighbors. The accuracy of metal and ligand positions defined by EXAFS is typically ± 0.02 Å for near neighbors, comparable to atomic-resolution crystal structures. There is an emerging trend of integrated XAS and X-ray crystallographic data acquisition for metalloproteins during the same experiment [58]. The combination of X-ray crystallography and ultra-high-resolution EXAFS is a powerful and complementary approach to study structure-function relationship in metalloproteins. In cases where subtle structural

changes occur in a reaction catalyzed by a metalloenzyme, this approach is particularly useful to derive the mechanism of the reaction [59].

Synchrotron radiation X-ray imaging and microprobes

Advances in recent years in methods and instrumentation for third generation synchrotron microprobe beamlines have allowed high-resolution spatial speciation analysis by synchrotron radiation X-ray fluorescence and X-ray absorption spectroscopies (SR-XRF and SR-XAS) in individual cells and cellular compartments [22]. The X-ray microbeams can penetrate sample in depth (1000 μm) and have high spatial resolution in speciation analysis (0.1–1 μm).

Intracellular distribution of trace elements can be determined by SR-XRF microprobe analysis. For example, a synchrotron X-ray microprobe with Kirkpatrick-Baez mirror optics was utilized for trace element imaging of yeast and human cells [60]. SR-XAF microanalysis was also instrumental in determining the location of Cu in mitochondria and Golgi apparatus in animal cells [61].

SR-XAS is particularly useful in obtaining information on metal distribution and other metal details such as metal oxidation states in cells. Time-dependent uptake and distribution of Cr (IV) in human lung cells was determined by SR-XAS [62]. Oxidation states of Cr in human ovarian cells were also observed by SR-XAS [63]. It was noted that Cr (IV) is reduced to Cr (III) in human ovarian cells when the cells are exposed to soluble chromate compounds in vitro, but remains oxidized when the cells are exposed to chromate compounds with low solubility.

SR-XRF or SR-XAS can also be coupled with a separation technique such as chromatography or gel electrophoresis to analyze metal content in a biological sample of interest. Protein bands from plant samples of *Citrus sinensis* were scanned for metal protein content (Ca, Cu, K and Zn) using SR-XRF following gel electrophoresis separation [64]. Quantitative determination of metal binding proteins was carried out by SR-XRF in addition to other methods such as flame atomic absorption spectrometry and flame atomic emission spectrometry. In another study, metalloprotein distribution and cancer development in human liver tissues was investigated [65, 66]. Proteins from tumor and surrounding non-tumor tissues were separated by gel electrophoresis. Zn, Fe, Cu and Mn contents in protein bands were measured with SR-XRF, and the metal contents in protein bands from tumor tissues were generally less than those from corresponding non-tumor tissues.

Bioinformatics approach

Since the report of the first complete sequence of an entire genome from *Haemophilus influenzae* in 1995, a huge amount of genome sequence data has become available [67]. As of March 2008, there 446 genome sequencing projects had been reported for eukaryotic genomes alone, with 23 completed and 224 in progress (http://www.ncbi.nlm.nih.gov/Genomes/). With the complete genome sequence of a living organism known, the next step is to annotate primary structures of all proteins from the organism followed up by the determination of high-resolution 3-D structures of these proteins through experimental or computational methods. Identification of metal binding sites from these proteins can be helpful in increasing the throughput of the structural determination process during protein expression, purification and structure solution steps [3]. In addition, metal binding information provides clues to protein function and this is particularly useful in annotation for hypothetical proteins with previously unknown function.

Metal binding sites can be predicted based on a known 3-D structure with reasonable accuracy because of the conserved nature of a metal binding site and its usual compact size. A method based on the free energy calculated from empirical force field was developed to predict the metal binding sites of Mg^{2+}, Ca^{2+}, Zn^{2+}, Mn^{2+} and Cu^{2+} and their affinities [68]. The method has a 90–97% success rate depending on the type of metal for a test set of 50 crystal structures. The free energies of Ca^{2+} and Zn^{2+} derived from the algorithm correlate well with the experimental data, allowing one to discriminate high- and low-affinity binding sites. Another study was reported to take advantage of low-resolution structural data in combination with sequence profile information to detect metal binding residue clusters with an accuracy rate of ~95% [69]. The method is useful to identify metal binding sites when no metal binding sequence motifs are obvious. Metal binding site prediction based on sequence information alone is more challenging compared with the prediction with incorporation of 3-D structure information. A bioinformatics approach to predict the metal binding ability of proteins to Zn, non-heme Fe and Cu in several organisms has been reported [6, 23]. In this approach, metalloproteins are identified through the combined search of known metal binding domains and of local sequence similarity to known metal binding motifs. The search can be applied to the whole genome sequence of any organism and provides an estimate of the complete ensemble of metalloproteins in the organism. A similar approach using structural bioinformatics and whole genome sequences have been carried out to

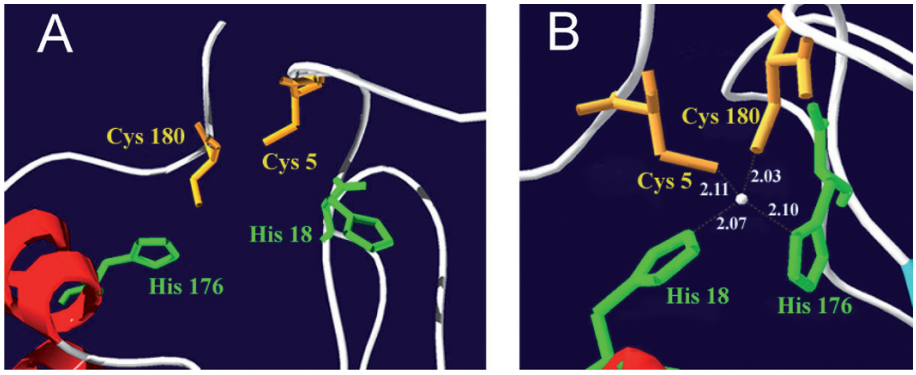

Figure 2. Accuracy evaluation of homology models [3]. (*A*) Existing structure model from Modbase for a NYSGXRC target. (*B*) New structure model for the same target using a new template with Zn^{2+} bound. In this case, the metal binding information helped to choose a more appropriate template to build a structure model with higher accuracy in predicting the metal binding site/active site. The models were based on the related PDB entries 1MLZ and 1NKU, respectively.

identify Fe, Zn, Mn and Co binding proteins [70]. Structural information was extracted from the Structural Classification of Proteins (SCOP) and the PDB [71]. Metalloproteins were identified in 23 archaea, 233 bacteria and 57 eukaryotic genome sequences. Based on the abundance of metalloproteins found in these organisms, it is suggested that eukaryotic organisms are likely to be evolved in a highly oxygenated environment with increased bioavailability of trace metals, while the opposite is true for prokaryotic organisms.

A two-stage machine-learning algorithm to predict histine and cysteine amino acids that participate in binding of transition metals and iron complexes was reported [72]. The computational program utilizes sequence information by taking into count position-specific evolutionary profiles as well as other factors such as protein length and amino acid composition. This method provides information on transition metal binding site components and is highly complementary to the HT-XAS technique, which identifies the metal bound to a protein [3, 12].

Comparative structure modeling is a very powerful tool in deriving protein function from 3-D structures. MODBASE is a database of structural models [73] based on template-target alignments and can be accessed at the MODBASE website (http://www.alto.compbio.ucsf.edu/modbase-cgi/index.cgi). In general, the accuracy of constructed models depends on the sequence similarity between the model sequence and the templates. The models based on >30% sequence identity are suitable for many applications, including functional analysis and screening of small-molecule databases for potential drug discovery. The metal binding information can be used as an independent criterion in accessing the accuracy of an existing homology model (Fig. 2) [3].

Databases devoted to metal binding include the Metalloprotein Database (MDB, http://metallo.scripps.edu) [74], which contains quantitative information on all the metal containing sites available from structures in the PDB, and PROMISE, which provides structural and functional information on metalloproteins and is now discontinued [75]. Other related databases include BIND (http://bind.ca), which is designed to store full descriptions of interactions with ligands, including metals. It is built up from peer-reviewed literature and only published results are accepted. In addition, databases from the large structural genomics centers usually include biophysical characterization of the targets and may contain metal binding information. These servers and databases are extremely useful in providing the broad scientific communities easy access to the existing data on structural and functional characterization of metalloproteins.

1 Tainer, J. A., Roberts, V. A. and Getzoff, E. D. (1991) Metal-binding sites in proteins. Curr. Opin. Biotechnol. 2, 582–591.
2 Degtyarenko, K. (2000) Bioinorganic motifs: towards functional classification of metalloproteins. Bioinformatics 16, 851–864.
3 Shi, W., Zhan, C., Ignatov, A., Manjasetty, B. A., Marinkovic, N., Sullivan, M., Huang, R. and Chance, M. R. (2005) Metalloproteomics: high-throughput structural and functional annotation of proteins in structural genomics. Structure 13, 1473–1486.
4 Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2006) Counting the zinc-proteins encoded in the human genome. J. Proteome Res. 5, 196–201.
5 Andreini, C., Banci, L., Bertini, I., Elmi, S. and Rosato, A. (2007) Non-heme iron through the three domains of life. Proteins 67, 317–324.
6 Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2008) Occurrence of copper proteins through the three domains of life: a bioinformatic approach. J. Proteome Res. 7, 209–216.
7 Andreini, C., Banci, L., Bertini, I. and Rosato, A. (2006) Zinc through the three domains of life. J. Proteome Res. 5, 3173–3178.
8 Kryukov, G. V. and Gladyshev, V. N. (2004) The prokaryotic selenoproteome. EMBO Rep. 5, 538–543.
9 Kryukov, G. V., Castellano, S., Novoselov, S. V., Lobanov, A. V., Zehtab, O., Guigo, R. and Gladyshev, V. N. (2003) Characterization of mammalian selenoproteomes. Science 300, 1439–1443.
10 Gladyshev, V. N., Kryukov, G. V., Fomenko, D. E. and Hatfield, D. L. (2004) Identification of trace element-containing proteins in genomic databases. Annu. Rev. Nutr. 24, 579–596.

11 Wu, Q., Huang, K. and Xu, H. (2003) Effects of long-term selenium deficiency on glutathione peroxidase and thioredoxin reductase activities and expressions in rat aorta. J. Inorg. Biochem. 94, 301–306.
12 Chance, M. R., Fiser, A., Sali, A., Pieper, U., Eswar, N., Xu, G., Fajardo, J. E., Radhakannan, T. and Marinkovic, N. (2004) High-throughput computational and experimental techniques in structural genomics. Genome Res. 14, 2145–2154.
13 Manjasetty, B. A., Turnbull, A. P., Panjikar, S., Bussow, K. and Chance, M. R. (2008) Automated technologies and novel techniques to accelerate protein crystallography for structural genomics. Proteomics 8, 612–625.
14 Manjasetty, B. A., Shi, W., Zhan, C., Fiser, A. and Chance, M. R. (2007) A high-throughput approach to protein structure analysis. Genet. Eng. (N Y) 28, 105–128.
15 Bantscheff, M., Schirle, M., Sweetman, G., Rick, J. and Kuster, B. (2007) Quantitative mass spectrometry in proteomics: a critical review. Anal. Bioanal. Chem. 389, 1017–1031.
16 Aebersold, R. and Mann, M. (2003) Mass spectrometry-based proteomics. Nature 422, 198–207.
17 Szpunar, J. (2000) Bio-inorganic speciation analysis by hyphenated techniques. Analyst 125, 963–988.
18 Szpunar, J. (2005) Advances in analytical methodology for bioinorganic speciation analysis: metallomics, metalloproteomics and heteroatom-tagged proteomics and metabolomics. Analyst 130, 442–465.
19 Bertini, I. and Cavallaro, G. (2008) Metals in the "omics" world: copper homeostasis and cytochrome c oxidase assembly in a new light. J. Biol. Inorg. Chem. 13, 3–14.
20 Sanz-Medel, A. (2005) From metalloproteomics to heteroatom-tagged proteomics. Anal. Bioanal. Chem. 381, 1–2.
21 Bettmer, J. (2005) Metalloproteomics: a challenge for analytical chemists. Anal. Bioanal. Chem. 383, 370–371.
22 Lobinski, R., Moulin, C. and Ortega, R. (2006) Imaging and speciation of trace elements in biological environment. Biochimie. 88, 1591–1604.
23 Andreini, C., Bertini, I. and Rosato, A. (2004) A hint to search for metalloproteins in gene banks. Bioinformatics 20, 1373–1380.
24 Bertini, I. and Rosato, A. (2003) Bioinorganic chemistry in the postgenomic era. Proc. Natl. Acad. Sci. U S A 100, 3601–3604.
25 Schaumloffel, D. (2004) Capillary liquid separation techniques with ICP MS detection. Anal. Bioanal. Chem. 379, 351–354.
26 Montes-Bayon, M., DeNicola, K. and Caruso, J. A. (2003) Liquid chromatography-inductively coupled plasma mass spectrometry. J. Chromatogr. A 1000, 457–476.
27 Chery, C. C., Dumont, E., Cornelis, R. and Moens, L. (2001) Two-dimensional gel electrophoresis of selenized yeast and autoradiography of 75Se-containing proteins. Fresenius J. Anal. Chem. 371, 775–781.
28 Neilsen, J., Abildtrup, A., Christensen, J., Watson, P., Cox, A. and McLeod, C. W. (1998) Laser ablation inductively coupled plasma-mass spectrometry in combinaton with gel electrophoresis: a new strategy for speciation of metal-binding serum proteins. Spectrochim. Acta, Part B: At. Spectrosc. 53, 339–345.
29 Binet, M. R., Ma, R., McLeod, C. W. and Poole, R. K. (2003) Detection and characterization of zinc- and cadmium-binding proteins in Escherichia coli by gel electrophoresis and laser ablation-inductively coupled plasma-mass spectrometry. Anal. Biochem. 318, 30–38.
30 Ballihaut, G., Tastet, L., Pecheyran, C., Bouyssiere, B., Donard, O. F. X., Grimaud, R. and Lobinski, R. (2005) Biosysthesis, purification and analysis of selenomethionyl calmodulin by gel electrophoresis-laser ablation ICP MS and capillary HPLC-ICP MS peptide mapping following in-gel trptic digestion. J. Anal. At. Spectrom. 20, 493–499.
31 Fan, T. W. M., Pruszkowski, E. and Shuttleworth, S. (2002) Speciation of selenoproteins in se-contaminated wilelife by gel electrophoresis and laser ablation-ICP-MS. J. Anal. At. Spectrom. 17, 1621–1623.
32 Becker, J. S., Zoriy, M., Pickhardt, C., Damoc, E., Juhacz, G., Palkovits, M. and Przybylski, M. (2005) Determination of phosphorus-, copper-, and zinc-containing human brain proteins by LA-ICPMS and MALDI-FTICR-MS. Anal. Chem. 77, 5851–5860.
33 Becker, J., Boulyga, S., Becker, J. S., Pickhardt, C., Damoc, E. and Przybylski, M. (2003) Structural identification and quantification of protein phosphorylations after gel electrophoresis separation using Fourier transform ion cyclotron resonance mass spectrometry and LA ICP MS. Int. J. Mass Spectrom. 228, 985–997.
34 Becker, J. S., Matusch, A., Depboylu, C., Dobrowolska, J. and Zoriy, M. V. (2007) Quantitative imaging of selenium, copper, and zinc in thin sections of biological tissues (slugs-genus arion) measured by laser ablation inductively coupled plasma mass spectrometry. Anal. Chem. 79, 6074–6080.
35 Becker, J. S., Zoriy, M. V., Pickhardt, C., Palomero-Gallagher, N. and Zilles, K. (2005) Imaging of copper, zinc, and other elements in thin section of human brain samples (hippocampus) by laser ablation inductively coupled plasma mass spectrometry. Anal. Chem. 77, 3208–3216.
36 Prange, A. and Profrock, D. (2005) Application of CE-ICP-MS and CE-ESI-MS in metalloproteomics: challenges, developments, and limitations. Anal. Bioanal. Chem. 383, 372–389.
37 Prange, A., Schaumloffel, D., Bratter, P., Richarz, A. N. and Wolf, C. (2001) Species analysis of metallothionein isoforms in human brain cytosols by use of capillary electrophoresis hyphenated to inductively coupled plasma-sector field mass spectrometry. Fresenius J. Anal. Chem. 371, 764–774.
38 Prange, A. and Schaumloffel, D. (2002) Hyphenated techniques for the characterization and quantification of metallothionein isoforms. Anal. Bioanal. Chem. 373, 441–453.
39 Hagege, A., Baldinger, T., Martin-Jouet, M., Zal, F., Leroy, M., Leize, E. and Dorsselaer, A. V. (2004) Assessment of the contribution of inductively coupled plasma mass spectrometry to metalloprotein analysis: a novel approach for studies of multiproteic complexes. Rapid Commun. Mass Spectrom. 18, 735–738.
40 Potier, N., Barth, P., Tritsch, D., Biellmann, J. F. and Van Dorsselaer, A. (1997) Study of non-covalent enzyme-inhibitor complexes of aldose reductase by electrospray mass spectrometry. Eur. J. Biochem. 243, 274–282.
41 Kaltashov, I. A., Zhang, M., Eyles, S. J. and Abzalimov, R. R. (2006) Investigation of structure, dynamics and function of metalloproteins with electrospray ionization mass spectrometry. Anal. Bioanal. Chem. 386, 472–481.
42 Apuy, J. L., Busenlehner, L. S., Russell, D. H. and Giedroc, D. P. (2004) Ratiometric pulsed alkylation mass spectrometry as a probe of thiolate reactivity in different metalloderivatives of Staphylococcus aureus pI258 CadC. Biochemistry 43, 3824–3834.
43 Gonzalez de Peredo, A., Saint-Pierre, C., Adrait, A., Jacquamet, L., Latour, J. M., Michaud-Soret, I. and Forest, E. (1999) Identification of the two zinc-bound cysteines in the ferric uptake regulation protein from Escherichia coli: chemical modification and mass spectrometry analysis. Biochemistry 38, 8582–8589.
44 Gonzalez de Peredo, A., Saint-Pierre, C., Latour, J. M., Michaud-Soret, I. and Forest, E. (2001) Conformational changes of the ferric uptake regulation protein upon metal activation and DNA binding; first evidence of structural homologies with the diphtheria toxin repressor. J. Mol. Biol. 310, 83–91.
45 Zhu, M. M., Rempel, D. L., Zhao, J., Giblin, D. E. and Gross, M. L. (2003) Probing Ca2+-induced conformational changes in porcine calmodulin by H/D exchange and ESI-MS: effect of cations and ionic strength. Biochemistry 42, 15388–15397.
46 Nischwitz, V., Michalke, B. and Kettrup, A. (2003) Identification and quantification of metallothionein isoforms and superoxide dismutase in spiked liver extracts using HPLC-ESI-MS offline coupling and HPLC-ICP-MS online coupling. Anal. Bioanal. Chem. 375, 145–156.

47 Polec, K., Perez-Calvo, M., Garcia-Arribas, O., Szpunar, J., Ribas-Ozonas, B. and Lobinski, R. (2002) Investigation of metal complexes with metallothionein in rat tissues by hyphenated techniques. J. Inorg. Biochem. 88, 197–206.
48 Mounicou, A., Polec, K, Chassaigne, H., Potin-Gautier, M. and Lobinski, R. (2000) Characterization of metal complexes with metallothioneins by capillary zone eletrophoresis (CZE) with ICP-MS and electrospray (ES)-MS detection. J. Anal. At. Spectrom. 15, 635–642.
49 Kulkarni, P. P., She, Y. M., Smith, S. D., Roberts, E. A. and Sarkar, B. (2006) Proteomics of metal transport and metal-associated diseases. Chemistry 12, 2410–2422.
50 Roelofsen, H., Balgobind, R. and Vonk, R. J. (2004) Proteomic analyzes of copper metabolism in an in vitro model of Wilson disease using surface enhanced laser desorption/ionization-time of flight-mass spectrometry. J. Cell. Biochem. 93, 732–740.
51 She, Y. M., Narindrasorasak, S., Yang, S., Spitale, N., Roberts, E. A. and Sarkar, B. (2003) Identification of metal-binding proteins in human hepatoma lines by immobilized metal affinity chromatography and mass spectrometry. Mol. Cell. Proteomics 2, 1306–1318.
52 Smith, S. D., She, Y. M., Roberts, E. A. and Sarkar, B. (2004) Using immobilized metal affinity chromatography, two-dimensional electrophoresis and mass spectrometry to identify hepatocellular proteins with copper-binding ability. J. Proteome Res. 3, 834–840.
53 Thompson, D. K., Beliaev, A. S., Giometti, C. S., Tollaksen, S. L., Khare, T., Lies, D. P., Nealson, K. H., Lim, H., Yates, J., 3rd, Brandt, C. C., Tiedje, J. M. and Zhou, J. (2002) Transcriptional and proteomic analysis of a ferric uptake regulator (fur) mutant of Shewanella oneidensis: possible involvement of fur in energy metabolism, transcriptional regulation, and oxidative stress. Appl. Environ. Microbiol. 68, 881–892.
54 Khare, T., Esteve-Nunez, A., Nevin, K. P., Zhu, W., Yates, J. R., 3rd, Lovley, D. and Giometti, C. S. (2006) Differential protein expression in the metal-reducing bacterium Geobacter sulfurreducens strain PCA grown with fumarate or ferric citrate. Proteomics 6, 632–640.
55 Scott, R. A., Shokes, J. E., Cosper, N. J., Jenney, F. E. and Adams, M. W. (2005) Bottlenecks and roadblocks in high-throughput XAS for structural genomics. J. Synchrotron Radiat. 12, 19–22.
56 Bonanno, J. B. Almo, S. C., Bresnick, A., Chance, M. R., Fiser, A., Swaminathan, S., Jiang, J., Studier, F. W., Shapiro, L., Lima, C. D., Gaasterland, T. M., Sali, A., Bain, K., Feil, I., Gao, X., Lorimer, D., Ramos, A., Sauder, J. M., Wasserman, S. R., Emtage, S., D'Amico, K. L. and Burley, S. K. (2005) New York-Structural GenomiX Research Consortium (NYSGXRC): a large scale center for the protein structure initiative. J. Struct. Funct. Genomics 6, 225–232.
57 Wang, B. C., Adams, M. W., Dailey, H., DeLucas, L., Luo, M., Rose, J., Bunzel, R., Dailey, T., Habel, J., Horanyi, P., Jenney, F. E., Jr., Kataeva, I., Lee, H. S., Li, S., Li, T., Lin, D., Liu, Z. J., Luan, C. H., Mayer, M., Nagy, L., Newton, M. G., Ng, J., Poole, F. L., 2nd, Shah, A., Shah, C., Sugar, F. J. and Xu, H. (2005) Protein production and crystallization at SECSG – an overview. J. Struct. Funct. Genomics 6, 233–243.
58 Latimer, M. J., Ito, K., McPhillips, S. E. and Hedman, B. (2005) Integrated instrumentation for combined polarized single-crystal XAS and diffraction data acquisition for biological applications. J. Synchrotron Radiat. 12, 23–27.
59 Hasnain, S. S. and Strange, R. W. (2003) Marriage of XAFS and crystallography for structure-function studies of metalloproteins. J. Synchrotron Radiat. 10, 9–15.
60 Ortega, R., Bohic, S., Tucoulou, R., Somogyi, A. and Deves, G. (2004) Microchemical element imaging of yeast and human cells using synchrotron X-ray microprobe with Kirkpatrick-Baez optics. Anal. Chem. 76, 309–314.
61 Yang, L., McRae, R., Henary, M. M., Patel, R., Lai, B., Vogt, S. and Fahrni, C. J. (2005) Imaging of the intracellular topography of copper with a fluorescent sensor and by synchrotron x-ray fluorescence microscopy. Proc. Natl. Acad. Sci. U S A 102, 11179–11184.
62 Harris, H. H., Levina, A., Dillon, C. T., Mulyani, I., Lai, B., Cai, Z. and Lay, P. A. (2005) Time-dependent uptake, distribution and biotransformation of chromium(VI) in individual and bulk human lung cells: application of synchrotron radiation techniques. J. Biol. Inorg. Chem. 10, 105–118.
63 Ortega, R., Fayard, B., Salome, M., Deves, G. and Susini, J. (2005) Chromium oxidation state imaging in mammalian cells exposed in vitro to soluble or particulate chromate compounds. Chem. Res. Toxicol. 18, 1512–1519.
64 Verbi, F. M., Arruda, S. C., Rodriguez, A. P., Perez, C. A. and Arruda, M. A. (2005) Metal-binding proteins scanning and determination by combining gel electrophoresis, synchrotron radiation X-ray fluorescence and atomic spectrometry. J. Biochem. Biophys. Methods 62, 97–109.
65 Liu, Y., Li, L., Gao, Y., Chen, C., Li, B., He, W., Huang, Y. and Chai, Z. (2007) Distribution of metalloproteins in hepatocellular carcinoma and surrounding tissues. Hepatogastroenterology 54, 2291–2296.
66 Gao, Y., Chen, C., Chai, Z., Zhao, J., Liu, J., Zhang, P., Heb, W. and Huang, Y. (2002) Detection of metalloproteins in human liver cytosol by synchrotron radiation X-ray fluorescence combined with gel filtration chromatography and isoelectric focusing separation. Analyst 127, 1700–1704.
67 Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J. F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C., Gocayne, J. D., Scott, J., Shirley, R., Liu, L., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D., Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. and Venter, J. C. (1995) Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. Science 269, 496–512.
68 Schymkowitz, J. W., Rousseau, F., Martins, I. C., Ferkinghoff-Borg, J., Stricher, F. and Serrano, L. (2005) Prediction of water and metal binding sites and their affinities by using the Fold-X force field. Proc. Natl. Acad. Sci. U S A 102, 10147–10152.
69 Sodhi, J. S., Bryson, K., McGuffin, L. J., Ward, J. J., Wernisch, L. and Jones, D. T. (2004) Predicting metal-binding site residues in low-resolution structural models. J. Mol. Biol. 342, 307–320.
70 Dupont, C. L., Yang, S., Palenik, B. and Bourne, P. E. (2006) Modern proteomes contain putative imprints of ancient shifts in trace metal geochemistry. Proc. Natl. Acad. Sci. U S A 103, 17822–17827.
71 Murzin, A. G., Brenner, S. E., Hubbard, T. and Chothia, C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. J. Mol. Biol. 247, 536–540.
72 Passerini, A., Punta, M., Ceroni, A., Rost, B. and Frasconi, P. (2006) Identifying cysteines and histidines in transition-metal-binding sites using support vector machines and neural networks. Proteins 65, 305–316.
73 Pieper, U., Eswar, N., Braberg, H., Madhusudhan, M. S., Davis, F. P., Stuart, A. C., Mirkovic, N., Rossi, A., Marti-Renom, M. A., Fiser, A., Webb, B., Greenblatt, D., Huang, C. C., Ferrin, T. E. and Sali, A. (2004) MODBASE, a database of annotated comparative protein structure models, and associated resources. Nucleic Acids Res. 32, D217-D222.
74 Castagnetto, J. M., Hennessy, S. W., Roberts, V. A., Getzoff, E. D., Tainer, J. A. and Pique, M. E. (2002) MDB: the Metalloprotein Database and Browser at The Scripps Research Institute. Nucleic Acids Res. 30, 379–382.
75 Degtyarenko, K. N., North, A. C., Perkins, D. N. and Findlay, J. B. (1998) PROMISE: a database of information on prosthetic centres and metal ions in protein active sites. Nucleic Acids Res. 26, 376–381.