DOI: 10.1002/ejic.200700186

From Genes to Metalloproteins: A Bioinformatic Approach

Ivano Bertini*[a] and Antonio Rosato[a]

Keywords: Iron / Zinc / Copper / Heme proteins / Metalloproteins / Bioinformatics

Genome sequencing projects have provided researchers with detailed information on every gene present in a large variety of organisms. Each of these genes is responsible for producing one or more proteins (depending on processes such as alternative splicing and post-translational modification) that constitute the machineries needed to sustain cellular life. The ensemble of proteins that an organism is capable of producing is called the proteome. A significant share of each organism's proteome requires metal ions or metal-containing cofactors to carry out its physiological function (metalloproteome). Experimental methods for the characterization of the metalloproteome of even simple prokaryotic organisms are still not routinely available, therefore significant insights can currently be only obtained using bioinformatic approaches. Herein we review the results obtained from the bioinformatic analyses of a variety of organisms, including humans, in which the distribution of iron-, copper-, and zinc-binding proteins across the kingdoms of life were investigated. It was found that among the metals, zinc is required by the largest number of proteins in most organisms (on average, 10 % of the proteome in eukaryota is predicted to be zinc-binding), whereas copper is the least common. In addition, the majority of non-heme iron-binding proteins have homologs in each taxonomic domain (archaea, bacteria, eukaryota), whereas zinc- and copper-binding proteins are substantially differentiated in eukaryota with respect to prokaryotic organisms.

Introduction

The term bioinorganic chemistry (or biological inorganic chemistry or inorganic biochemistry) refers to the discipline in which the interactions between inorganic ions or complexes and molecules of biological interest are studied.[1] This is a vast field of science whose importance stems from the fact that life originated and developed on the Earth's crust, an eminently inorganic environment. A core part of bioinorganic chemistry is focused on the study of metalloproteins, which are defined as those proteins that bind tightly (with an association constant greater than 10^8 M^{-1}) to one or more metal ions or metal-containing cofactors,

[a] Magnetic Resonance Center (CERM) and Department of Chemistry – University of Florence,
via L. Sacconi 6, 50019 Sesto Fiorentino, Italy
Fax: +39-055-457-4271
E-mail: ivanobertini@cerm.unifi.it

Ivano Bertini (http://www.cerm.unifi.it/bertini.html) received his doctorate in chemistry from the University of Florence in 1964 and became full professor of Chemistry in 1975 at the same University. He is the Director of the Magnetic Resonance Center (CERM http://www.cerm.unifi.it/) of the University of Florence. His main research interests are NMR spectroscopy, the expression and preparation of metalloproteins, their structural characterization, and the investigation of their interactions with emphasis on understanding cellular processes at the molecular level. He has authored over 600 papers and many books. He received the Chugaev Diploma of the Kurnakov Institute of the Academy of Science, USSR, in 1981, the Golden Medal of the Magnetic Resonance Group of the Italian Chemical Society in 1991, the Prize Accademia dei Lincei, Italy in 1993, the Bijvoet Medal, Utrecht, NL, in 1998, the Sapio NMR Prize, Italy, in 1999, the Cannizzaro Medal of the Italian Chemical Society and the Basolo medal in 2006. Amongst the special lectures: A. D. Little Lecturer at MIT, Cambridge, MS, USA, in 1997, E. L. Mütterties Lecturer at Berkeley, CA, USA, in 1997, and FECS lecturer, Athens, Greece, in 2002. He has received three honorary doctorates and is a member of the Accademia Nazionale dei Lincei and of the Academia Europaea.

Antonio Rosato (http://www.cerm.unifi.it/rosato/) obtained his doctorate from the University of Florence (Italy) in 1998. He is now Associate Professor of Chemistry at the same University. His research interests focus on the theory and practice of NMR spectroscopy applied to paramagnetic systems, metal-mediated protein–protein interactions, and on the application and development of bioinformatic tools to browse gene banks and genome sequences, with a particular interest in metalloproteins. He has co-authored more than 60 articles in peer-reviewed journals.

and of metal–protein adducts, which are defined as adducts formed by metal ions or metal-containing cofactors with proteins (with an association constant lower than 10^8 M^{-1}). The metal may play a physiologically relevant role in both metalloproteins and metal–protein adducts. For simplicity, from here on we will collectively refer to "true" metalloproteins and metal–protein adducts as metalloproteins. From a more biologically oriented point of view, metalloproteins can be defined as those proteins requiring one or more metal ions or metal-containing cofactors to carry out their biological function. The two definitions overlap to a large extent. Note that the metal may be required not only because it is the center of catalysis, as in many metalloenzymes, but also for structural reasons (i.e. to stabilize the metalloprotein's fold) or regulatory ones (e.g. as in calmodulin).

In the early years of bioinorganic chemistry, the quantification of the interactions of metal ions with living organisms focused mainly on determining the quantity of metal present in a healthy organism per kilogram of weight (in the case of higher organisms such as mammals or plants) or per cell (in the case of unicellular organisms). These investigations provided valuable information on the importance of metal ions for the correct functioning of cells as well as on several diseases caused by defects in the metabolism of metal ions and metal-containing cofactors. The success of genome sequencing initiatives, however, has added a new dimension to bioinorganic chemistry, namely the quantification of *how many proteins are metalloproteins*. This aspect deals with identifying how many and which metabolic pathways require or involve metal ions in addition to establishing how much of a metal is needed to support the life of an organism. These two facets are highly complementary and are both essential in order for bioinorganic chemistry to fully contribute to the Life Sciences.

Genome sequencing projects have provided the amino acid sequences of proteins encoded in the genome (besides the possible effects of alternative splicing or of post-translational modifications, which contribute to effectively expanding the proteome) of an ever increasing number of organisms. The large majority of these proteins lack any experimental functional information, as functional characterization of proteins at the whole proteome level is still not routinely possible. To cope with this limitation, bioinformatic tools have been extensively exploited to predict functional features from the analysis of amino acid sequences.[2–5] In addition, projects in structural biology and structural genomics (the structural characterization of all proteins in a proteome) have prompted the development of several bioinformatic approaches to predict function from structure.[6–9] The functional information obtained either experimentally or through bioinformatics is generally summarized in an annotation, which is associated with the various individual proteins in any sequenced genome. The annotation of gene products is still, however, a far from trivial process. Predicting function from sequence, even when structural information is available,[10–12] often remains an elusive endeavor that is further complicated by difficulties in the exact identification of the coding sequence, especially in higher organisms.[13,14]

The concepts outlined in the preceding paragraph also apply to metalloproteins,[15] although these systems are even more challenging because metal-binding ability is not easily detected from standard sequence analysis. On the other hand, the presence of a motif suitable for protein binding to metal ion(s) can be exploited for more reliable identification and classification (including functional inferences) of these proteins.[15–18] Genome-wide identification and classification of metalloproteins is also difficult experimentally, although some systematic efforts in this direction have been described.[19–21] Since 2004, specific bioinformatic approaches have been developed, refined, and applied at the whole genome level to identify the fraction of an organism's proteome that (putatively) requires metal ions for its function, in other words to identify all metalloproteins encoded by an organism's genome.[18,22–24] The ensemble of the metalloproteins of an organism is hereafter called the *metalloproteome*. Herein we review the results of these efforts and provide an overview of the use of different essential metal ions by organisms in the three domains of life, including humans.

On the Reasons Why Identifying Metal-Binding Sites is Difficult

Metal ions can bind to metalloproteins at a variety of different sites. However, the specificity of a metal-binding site for a metal may be poor as the same site can often bind different metal ions with similar affinities (the physiologically relevant metal ion may sometimes actually not be the tightest binder), a phenomenon that has allowed researchers to obtain structural details of metal-binding sites through a combination of spectroscopic tools and metal replacement.[25,26] This can also be true for polymetallic cofactors such as iron-sulfur clusters.[27] When preference for a given metal ion indeed occurs at a binding site, it can be tuned through site-directed mutagenesis, typically by substituting one or two protein residues.[28]

Even when 3D data are available from X-ray crystallography, the identification of the presence of a metal-binding site or the identification of the metal bound at a site is often far from trivial.[29,30] There have been several examinations of the whole protein data bank (PDB) in order to identify, classify, and attempt to devise general rules for confirming the presence of metal-binding sites.[31–34] However, these efforts have not provided guidelines that can be applied safely and systematically, even with structural data in hand. When the protein structure is solved in its apo form, a further obstacle to the prediction of the presence of a metal-binding site may arise because of structural rearrangements between the apo and holo forms of the protein, especially, but not exclusively, those involving the side chain atoms of the amino acidic metal ligands.[35]

X-ray fluorescence analysis is a suitable strategy to experimentally identify the nature of an unknown metal ion bound to a protein.[36] Another suitable approach is the study of X-ray anomalous diffraction.[37] NMR analysis can also be informative.[38] The exact details of metal coordination can often be obtained upon substitution of the metal with a paramagnetic metal ion through the application of purposefully designed NMR experiments[39,40] and/or of advanced computational approaches.[41] Alternatively, NMR-active diamagnetic metals, such as Ag, Cd, or Hg, can be exploited to directly detect proton–metal correlations in heteronuclear two-dimensional experiments.[42–45]

Another possible approach to the identification of metalloproteins is high-throughput screening.[19–21] These methods, which are based on XAS or colorimetry, can identify unprecedented metalloproteins and provide hints about the identity of the bound metal. They do not, however, provide detailed information on the nature, identity, and location of amino acidic ligands within the protein sequence.

Sequence-Based Bioinformatic Identification of Metalloproteins

Because of the difficulties mentioned in the preceding section, it is important to develop methods that can predict metal-binding capabilities and metal-binding sites based only on the experimental information directly obtainable from genome sequencing, namely the amino acid sequence of the proteins. In other words, the scientific community working in bioinorganic chemistry would benefit from bioinformatic tools that can identify metalloproteins within proteomes.

Bioinformatic resources designed for bioinorganic chemists are relatively limited with respect to the plethora of resources available to the general biologists' community.[15] In particular, no web servers are currently available for the prediction of metal-binding capabilities based on protein sequence. On the other hand, a few algorithms suitable for the identification of metalloproteins in proteomes have been published; these are addressed below.

Before discussing the methods available for the prediction of metal-binding properties based on sequence, it is useful to briefly describe the application to metalloproteins of protein structure modeling based on homology (so-called homology modeling or comparative modeling).[46,47] Homology modeling has been applied to metalloproteins with success in a number of instances, yielding interesting insights into the determinants of structure and function of the systems investigated.[16,48–58] If one is uncertain about the metal-binding properties of a given protein, homology modeling is a potential method for building a 3D structural model that may subsequently be inspected for the presence of putative metal-binding sites. This approach, however, presents various potential pitfalls, some of which are linked to the intrinsic shortcomings of the homology modeling approach, such as the difficulty in correctly modeling the conformation of amino acid side-chains and loop regions,[46] where metal-binding sites are often located. However, it has been suggested that, especially in systems binding metal-containing cofactors of a relatively large size, such as iron-sulfur clusters or iron porphyrins, modeling of a metalloprotein structure in its apo form can lead to significant distortions of even the global protein fold with respect to the structure of the holo protein.[16,48] Such distortions would prevent reliable use of the modeled structure for the identification of binding sites.

The simplest approach to assigning metal-binding capabilities to a protein is to look for the presence of a known metal-binding pattern (MBP). MBPs can be automatically extracted from the PDB and can be described in detail using the COMe formalism.[17] Simpler formalisms essentially report the identity of ligands, the spacing between them, and, sometimes, the presence of other conserved residues within or near the MBP. The simplest representation of an MBP is in the format AX_nBX_mC..., where A, B, and C are the metal-binding amino acids and n, m,... are the number of amino acid residues in between two subsequent ligands. A non-specialized database of functional patterns is PROSITE.[59] However, simply looking for the occurrence of an MBP within a sequence can be misleading[18] and may lead to the assignment of an excessively high number of metalloproteins. A simple example is given by the case of proteins that bind metals with pairs of cysteines, which in many systems are involved in the formation of disulfide bridges rather than in metal binding. Based on the presence of cysteine pairs (with a given spacing), these proteins would be erroneously assigned as metalloproteins (note that in fact they may be able to bind metals in vitro in their reduced state). A more refined, less error-prone approach takes into account the presence of local sequence similarity around the MBP[18] with respect to a known metalloprotein. The presence of local sequence similarity is indicative of local structural similarity, which implies similar conformation for the amino acidic ligands of the MBP and thus formation of the metal-binding site.[18] The main weakness of this approach lies in the detection of metalloproteins containing an MBP site where the spacing between ligands is very large. As an example, consider a case where two ligands are close to one another, say within five amino acids, a third one is 30 amino acids away, and a fourth one is another 30 amino acids away; these types of situations are relatively easy to overlook. To enhance sensitivity and reliability, the analysis of MBPs can be successfully complemented by analyzing the occurrence of protein domains known to be metal-binding taken from domain libraries such as Pfam[60] or SCOP.[61] In the case of human zinc-binding proteins, it has been shown that these two approaches independently provide metal-binding predictions that are in good agreement with one another, and at the same time their combination enhances coverage in terms of zinc-binding proteins detected.[22] The combined use of MBPs and protein domains for metalloprotein identification is also useful in the reduction of the number of false positives (i.e. of non-metalloproteins that are erroneously

assigned as metalloproteins).[23,24] The use of only metal-binding protein domains does not provide the same benefits.

Alternatively, machine learning methods have shown better performance than an approach based on the simple detection of the occurrence of an MBP within a sequence.[62] These methods can provide indications at the level of an individual cysteine or histidine residue and predict whether it is bound to a metal or not. Metalloproteins are identified as proteins containing one or more metal-bound cysteines or histidines.[62] Although this approach may be promising, especially because it may identify unprecedented MBPs, significant improvement is still needed, especially as other amino acids (aspartic acid, glutamic acid, tyrosine, etc.) must be taken into account as possible ligands.

It must be remembered that the performance of any predictive method depends not only on the quality and ingenuity of the underlying algorithm(s) but also on the quality and amount of experimental data available to perfect the computational approach. Figure 1 shows the number of predictions for zinc-binding proteins based on the approach described by Andreini et al.,[18] which is based solely on structural information available from the PDB, as a function of the number of structures of zinc proteins available. It can clearly be seen that a dramatic increase in the number of predicted human zinc proteins occurs after a "critical mass" of available structures is reached. However, the subsequent growth of the number of structures of zinc proteins in the PDB that occurred after the mid-90s yielded only a relatively small additional increase in the number of predicted zinc-binding proteins (a 10-fold increase in structures corresponds to slightly less than a twofold increase in predicted sequences). This is presumably a consequence of the relatively low structural diversity additionally contained in the newest structures.

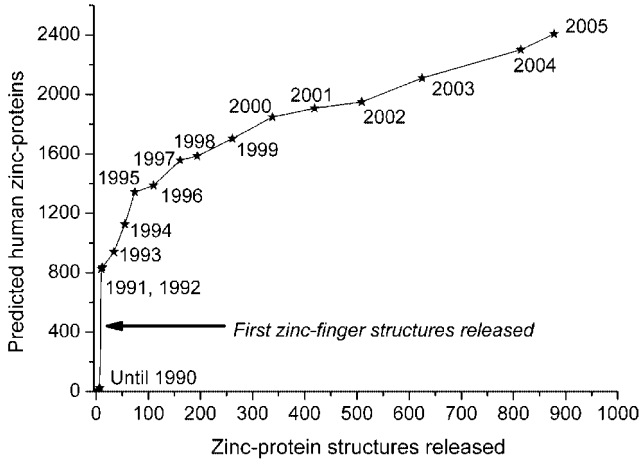

Figure 1. Number of human zinc proteins predicted as a function of the number of experimental zinc protein structures available from the PDB. The abscissa of each point corresponds to the number of structures availble at the end of the indicated year.

The Metalloproteomes

The Zinc Metalloproteome

Zinc-binding proteins are found in essentially all organisms[24] and are quite abundant in eukaryotic organisms such as humans.[22] Zinc-binding patterns most commonly include four (70% of instances in the PDB) or three (23% of instances in the PDB) protein ligands (not distinguishing between mono- and bidentate carboxylate ligands). In zinc-binding patterns with three protein ligands, the metal ion often completes its first coordination sphere by binding a water or buffer molecule. The most common amino acidic ligands for zinc are cysteine and histidine: 68% of the patterns contain at least one cysteine while 65% contain at least one histidine. Figure 2 shows a distribution of the various types of zinc-binding patterns grouped on the basis of the number, identity, and order in sequence of protein ligands, in predicted human zinc-binding proteins.[22]

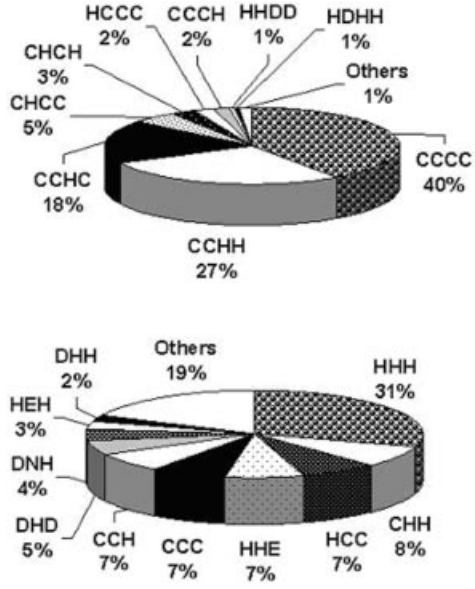

Figure 2. Distribution of four- and three-ligand MBP types detected in putative human zinc proteins. MBPs are grouped based on their amino acid composition and on the order of the protein ligands in the sequence. For example, the MBPs HX_3HX_5C and HX_8HX_4C are grouped together as HHC; all MBPs containing four cysteines are grouped together as CCCC.

Human zinc enzymes are present in all enzyme classes, with hydrolases being the most common zinc-dependent enzymes.[22] Indeed, it is well known that a zinc-activated water molecule is responsible for nucleophilic attack on the substrate in these systems.[63,64] However, the largest percentage of all human zinc-binding proteins for which a function can be proposed (>40%) are transcription factors. These consist mainly of zinc fingers; note, however, that zinc fingers do not interact only with DNA, but can also mediate interactions with RNA or proteins.[65] As the human zinc proteome makes up approximately 10% of the entire proteome, it follows from the above figures that 4%

of the human proteome is involved in zinc-dependent regulation of gene expression. Of course, to have a full picture of the importance of gene regulation in humans, one should add to this figure all the proteins involved in regulation that do not depend on zinc.

Some types of zinc-binding patterns are specifically associated with a given function, such as the CCHH type (C = cysteine, H = histidine), which is characteristic for transcription factors and contributes crucially to the stabilization of their 3D structure, whereas others can be associated with very different functional groups spanning different enzyme classes.[22] Even when contained in enzymes, the zinc ion is not necessarily the center of catalysis, but, as in the case of transcription factors, it may act as a contributor to the stability of the structure of the protein. Zinc-binding sites are thus useful modules that can be exploited not only for catalysis but also for structural purposes, for example to drive/stabilize the fold of a protein.

As mentioned, zinc-binding proteins are present in organisms from all domains of life (archaea, bacteria, eukaryotes), although they constitute a significantly diverse proportion of their proteomes.[24] General trends can be readily identified that point to archaeal and bacterial proteomes having a similar content of zinc proteins (5% to 6%), while the zinc proteomes of eukaryotic organisms constitute a much more important fraction of the whole proteome (around 9%). The incidence of eukaryotic zinc proteomes is higher due mainly to increased needs for regulation of gene expression, cell compartmentalization, and cell differentiation in multicellular organisms. Indeed, eukaryotic zinc-binding proteins involved in regulation of expression contribute nearly all the additional zinc proteins with respect to simpler organisms. Eukaryotes have recruited new zinc-binding proteins that are not needed, and indeed are not encoded, by archaea and bacteria specifically for the above purpose. As a result, the majority of the eukaryotic zinc proteome is unique to eukaryotes, in other words the majority of eukaryotic zinc proteins do not have homologs in archaeal or bacterial proteomes (Figure 3).

The role of zinc in the latter organisms is mainly as the catalytic center of hydrolytic enzymes, while very little use is made of zinc in regulatory proteins.[24] An interesting example is that of a metal-binding helix-turn-helix domain present in some bacterial and archaeal metal-sensing repressors where transcription is regulated as a function of intracellular metal concentration, such as that in the zinc-sensing SmtB protein from *Synechococcus* PCC 7942.[66,67] Intriguingly, an analysis of the zinc-proteome size of archaeal and bacterial organisms based on their range of growth temperature (i.e. by dividing them into hyperthermophilic, thermophilic, mesophilic, or psychrophilic organisms) suggests that the proteome of organisms living at higher temperatures is enriched in putative zinc-binding proteins. The zinc proteome of the 11 hyperthermophilic organisms studied in a previous paper[24] constitutes 7.0% ± 1.1% of the entire proteome, compared to 6.0% ± 1.0% of five thermophilic organisms, 5.3% ± 1.0% of 34 mesophilic organisms, and 4.5% ± 0.2% of two psychrophilic organisms. This effect may be due to an increased use of zinc for structural stability by organisms living at higher temperatures.

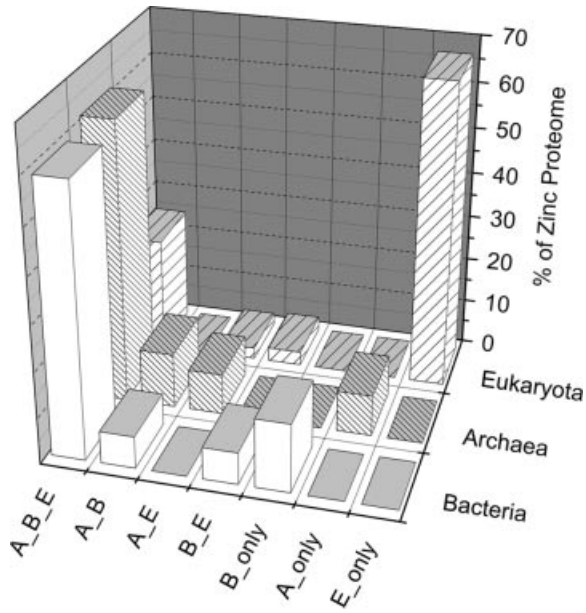

Figure 3. Fraction (with respect to the entire zinc-proteome) of predicted zinc-binding proteins in each domain of life having homologues in all three, two, or only one domain of life (A: archaea; B: bacteria; E: eukaryota). As an example, 60% of archaeal zinc-binding proteins have bacterial and eukaryotic homologues.

The Non-Heme-Iron Metalloproteome

In biological systems, iron can be incorporated into metalloproteins in a variety of cofactors, usually in the oxidation states +3 or +2 (other oxidation states may be generated during the catalytic cycle of iron enzymes). The simplest iron-binding protein is probably rubredoxin, a small electron-transfer protein that binds a single iron ion through the side-chain sulfur atoms of four cysteines. Due to its simplicity and ease of preparation and purification, the three-dimensional structure of rubredoxin has been studied in great detail.[68–71] However, iron-binding proteins are normally much more complicated than rubredoxin, as discussed below. The chemical complexity of most iron-containing cofactors (as well as, at least in the case of heme, their poor solubility), together with the need to control the extremely versatile reactivity of this metal to prevent undesired reactions, have resulted in the evolution of complex molecular machineries dedicated to the uptake/assembly and intracellular distribution of the cofactors.[72–76]

Many proteins bind iron in polymetallic clusters composed of iron and sulfide ions that can have various stoichiometries (e.g., Fe_2S_2, Fe_3S_4, Fe_4S_4).[77] These proteins are therefore called iron-sulfur proteins (or ferredoxins) and typically employ the side chains of cysteines or cysteines together with histidines to bind the iron-sulfur complex, although some exceptions exist. Clearly, the properties of these systems depend crucially on the bound cluster, but

also on the protein environment surrounding it.[78–80] A variety of heterometallic clusters containing iron together with other metals can also be formed and bound by proteins.[81] Finally, in very many metalloproteins iron is contained at the center of a variously substituted porphyrin (heme). The porphyrin provides four nitrogen ligands to the metal ion while the fifth coordination position is filled by a protein ligand and the sixth is filled by either a protein ligand or water, or remains vacant. The most common of these heme-containing systems are called cytochromes and are distinguished by the type of heme bound, which, in turn, is determined by the chemical nature of the substituents of the porphyrin. In some cytochrome variants, called *c*-type cytochromes, the heme substituents react with the protein to form covalent bonds between the polypeptide chain and the cofactor. Proteome-level analyses of the occurrence of heme-containing proteins cannot be carried out by exploiting MBPs as these involve only one or two protein ligands. In addition, the extensive hydrophobic interactions between the porphyrin moiety and the protein chain provide a significant contribution in driving binding, which must be taken into account. For these reasons, the heme proteome must be analyzed using approaches based on fold recognition together with the analysis of the occurrence of known protein sequence motifs, as done recently for mitochondrial-type mono-heme cytochromes *c*.[82] For proteins containing *b*-type heme, initial analyses based on the known folds (e.g., the globin fold or the peroxidase fold) indicate that their percentage in eukaryotic proteomes is in the range 0.3–1.0% (unpublished data from our laboratory). The corresponding figures for bacteria and archaea are also quite variable but have a clear tendency toward being smaller than for eukaryotes (in the range 0.0–0.5%).

The methods described in the section on "Sequence-based bioinformatic identification of metalloproteins" exploiting known MBPs can be successfully applied to the analysis of proteins binding non-heme iron.[23] Of the various types of MBPs relevant to this section, the most common ones contain only cysteine residues. This is presumably due to the fact that cysteine-only MBPs are involved in binding iron-sulfur clusters by most ferredoxins. Approximately half of all MBPs that bind non-heme iron are formed by various combinations of histidine, glutamic acid, and aspartic acid residues, without cysteines.

At variance with what is observed for zinc-binding proteins, in the case of non-heme iron-proteins there is no expansion of the fraction of the proteome that is metal-binding in eukaryotes with respect to simpler organisms. For some organisms, such as yeast, a contraction is actually observed.[23] Non-heme iron proteins constitute, on average, 7.1% ± 2.1% of archaeal proteomes, 3.9% ± 1.6% of bacterial proteomes, and only 1.1% ± 0.4% of eukaryotic proteomes (Figure 4).

Because eukaryotic proteomes are somewhat larger than those of simpler organisms, possibly with the exception of yeast, the above numbers mainly indicate that the non-heme

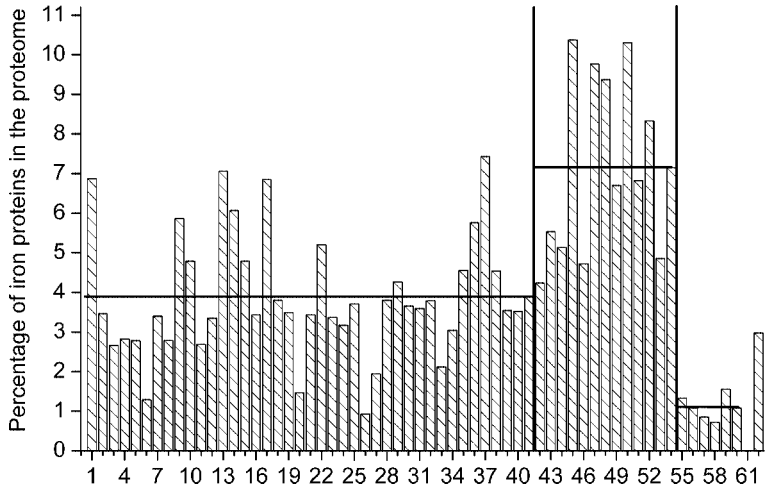

Figure 4. Percentage of iron-binding proteins identified in each proteome analyzed. The average values for all bacterial, archaeal, and eukaryotic organisms are also shown. Vertical lines separate the three domains and horizontal lines show the average values within each domain. The horizontal axis reports the data for, from left to right: 1. *Aquifex aeolicus*, 2. *Bacillus subtilis*, 3. *Bacteroides thetaiotaomicron*, 4. *Bdellovibrio bacteriovorus*, 5. *Bifidobacterium longum*, 6. *Borrelia burgdorferi*, 7. *Caulobacter crescentus*, 8. *Chlamydia trachomatis*, 9. *Chlorobium tepidum*, 10. *Clostridium acetobutylicum*, 11. *Coxiella burnetii*, 12. *Deinococcus radiodurans*, 13. *Desulfotalea psychrophila*, 14. *Desulfovibrio vulgaris*, 15. *Escherichia coli*, 16. *Fusobacterium nucleatum*, 17. *Geobacter sulfurreducens*, 18. *Haemophilus influenzae*, 19. *Helicobacter pylori*, 20. *Mesoplasma florum*, 21. *Mesorhizobium loti*, 22. *Methylococcus capsulatus*, 23. *Neisseria meningitides*, 24. *Nitrosomonas europaea*, 25. *Nostoc sp*, 26. Onion yellows, 27. *Rhodopirellula baltica*, 28. *Prochlorococcus marinus*, 29. *Pseudomonas aeruginosa*, 30. *Ralstonia solanacearum*, 31. *Rickettsia prowazekii*, 32. *Shewanella oneidensis*, 33. *Streptococcus pyogenes*, 34. *Streptomyces coelicolor*, 35. *Synechocystis sp*, 36. *Thermoanaerobacter tengcongensis*, 37. *Thermotoga maritime*, 38. *Thermus thermophilus*, 39. *Vibrio parahaemolyticus*, 40. *Xanthomonas campestris*, 41. All bacteria, 42. *Aeropyrum pernix*, 43. *Pyrobaculum aerophilum*, 44. *Sulfolobus tokodaii*, 45. *Archaeoglobus fulgidus*, 46. *Halobacterium sp*, 47. *Methanocaldococcus jannaschii*, 48. *Methanopyrus kandleri*, 49. *Methanosarcina acetivorans*, 50. *Methanothermobacter thermautotrophicus*, 51. *Thermoplasma acidophilum*, 52. *Pyrococcus abyssi*, 53. *Nanoarchaeum equitans*, 54. All archaea, 55. *Saccharomyces cerevisiae*, 56. *Drosophila melanogaster*, 57. *Caenorhabditis elegans*, 58. *Homo sapiens*, 59. *Arabidopsis thaliana*, 60. All eukaryota, 61. Empty, 62. All organisms.

iron proteome does not expand with the increase of proteome size that accompanies the increase in complexity of higher eukaryotic organisms. Indeed, in absolute terms, the number of non-heme iron proteins predicted in eukaryotes and in the bacteria with the largest proteomes span the same range. Figure 5 shows that the majority of non-heme iron proteins have homologues in all domains of life and that only about 10% of the non-heme iron proteome is specific to an individual domain. Taken together, these data point to non-heme iron being a cofactor of a set of metalloproteins that carry out an ensemble of basic functions which have changed very little with evolution. It can thus be concluded that the protein repertoire that sustains the increased complexity of eukaryotic organisms, and which is unique to these organisms, comprises very few predicted non-heme iron proteins.

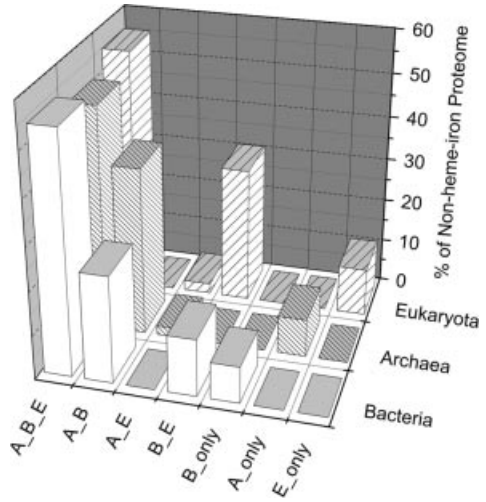

Figure 5. Fraction (with respect to the entire non-heme iron-proteome) of predicted iron-binding proteins in each domain of life having homologues in all three, two or only one domain of Life (A: Archaea; B: Bacteria; E: Eukaryota).

Overall, the non-heme iron proteome of organisms appears to be evolutionarily more ancient than the zinc proteome, with variations in its composition driven much more by loss of genes due to changes in environmental conditions that made some functions unnecessary than by recruitment of new iron proteins to cope with new functional needs (e.g. due to the evolution of multicellular organisms).[23]

The Copper Proteome

Copper is the least pervasive of all the metallic cofactors described in this review. Nevertheless, copper plays a crucial role in a number of fundamental biological processes such as aerobic respiration (in the enzyme cytochrome *c* oxidase). Copper is a redox-active metal that can shuttle between the +2 and +1 states. It is potentially toxic to the cell, as it can lead to the formation of reactive oxygen species (ROS), in particular hydroxyl radicals, through a Fenton reaction. Consequently, nature has evolved complex pathways to control the intracellular concentration and distribution of copper,[83–85] parts of which have been structurally characterized.[86] Recent structural studies of proteins involved in the above pathways in eukaryotic systems are available, for example of proteins involved in copper delivery to cytochrome *c* oxidase[87] or in copper delivery to the trans-Golgi network.[88,89]

Copper ions are most commonly bound by cysteine and histidine side-chains, followed by methionine, and, to a lesser extent, glutamic and aspartic acid. Copper-binding proteins are typically predicted (by applying essentially the same approach as for non-heme iron proteins) to account for less than 1% of an organism's proteome (unpublished data from our laboratory). In particular, eukaryotic proteomes contain between 0.2% and 0.5% of copper proteins, with an average of 0.3%. These figures correspond to eukaryotic organisms encoding a few tens up to around one hundred copper-binding proteins. Among eukaryotes, the plant *Arabidopsis thaliana* seems to be richer in copper enzymes (as well as in non-heme iron proteins, see preceding section) than the average. The copper proteome of archaea and bacteria is composed mainly of proteins that have homologs either in all domains of life or at least in the two prokaryotic domains (unpublished data from our laboratory, Figure 6). Eukaryotes instead contain a significant share of eukaryote-specific copper proteins. These proteins are mainly copper-dependent oxidoreductases, whose number could be higher with respect to what is found in less complex organisms because of the compartmentalization of the eukaryotic cell. In agreement with this, a parallel increase, albeit less important in numerical terms, of copper-ion transporters is also observed.

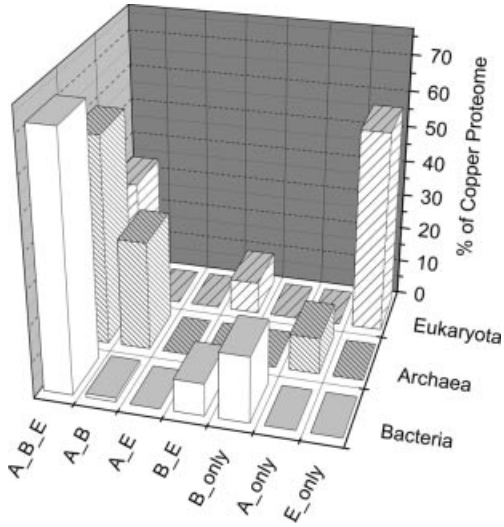

Figure 6. Fraction (with respect to the entire copper proteome) of predicted copper-binding proteins in each domain of life having homologues in all three, two, or only one domain of life (A: archaea; B: Bacteria; E: Eukaryota).

With the exception of hemocyanins, which have a role in dioxygen transport, the vast majority of copper enzymes can be assigned as catalyzing redox reactions. The second most common role for copper proteins is in metal-ion transport (e.g. these proteins are devoted to shuttling copper ions

within the cell, or to the uptake or excretion of copper). Copper transport is generally coupled to ATP hydrolysis.

Because of the extensively discussed (potential) toxicity of copper ions, systems devoted to controlling copper trafficking with the cells are present in a very large majority of all organisms that encode copper-dependent proteins. Some of these mechanisms have been so successful that they have significant similarities across a wide variety of prokaryotic as well as eukaryotic organisms. These include, for example, the chaperone/metalloATPase system for copper removal from the cytosol[53] or the Sco proteins, which are devoted to the assembly of cytochrome *c* oxidase and possibly of other copper-dependent oxidases.[90,91]

Concluding Remarks

As mentioned for iron and copper, the uptake, transport, release, and incorporation of metal cofactors in proteins may require a specialized machinery whose malfunction may even lead to illness.[92] This could be expected for particularly large or chemically complex metal-containing cofactors such as heme or iron-sulfur clusters, but also holds for individual metal ions. As an example, the covalent attachment of heme to *c*-type cytochromes requires up to eight different proteins in bacteria (so called cytochrome *c* maturation genes, *c*cm*ABCDEFGH*).[93] Of the proteins encoded by these genes, one (ccmE) is devoted to transporting the insoluble heme cofactor, which is bound to ccmE through a very unusual heme–protein covalent bond, in the bacterial periplasm.[94] The structure of ccmE, however, has only been solved only in the apo form.[74,95] As regards individual metal ions, the most intensely studied soluble transporters (called metallochaperones) are probably those that handle copper,[96–99] although metallochaperones are also known for other metal ions such as nickel[100] or arsenic.[101]

The presence of delivery systems for metal ions serves two purposes, namely to prevent metals from binding to adventitious sites on proteins different from the intended targets and to incorporate the appropriate metal ion at each site regardless of the affinity of other metals for the same site. This second aspect is particularly relevant because of the need to overcome the intrinsic trend in the variation of stability of the metal complexes along the *d* series (the so-called Irving–Williams series) in order to incorporate the metal needed for the biochemical function of the system and not the most thermodynamically favored metal at each site. In other words, metallochaperones allow the incorporation of metal ions into metalloproteins to take place under kinetic rather than thermodynamic control, thereby ensuring proper maturation and thus functioning of metalloenzymes and metalloproteins in general. From the chemical point of view, this unfortunately poses some problems, because predictions cannot rely on thermodynamic properties to define which metal is bound at which site, even if the presence of a metal-binding site is known. In addition, experimental procedures for protein purification may lead to exchange of the biologically relevant metal ion with other ions contained in the buffers that bind to the protein more tightly (e.g. metallophosphatases).[102,103] On the other hand, nature can exploit this property to incorporate different metals in a protein, depending on their availability.[104]

In summary, there are now methods available that allow the prediction of an organism's metalloprotein content with some confidence. A systematic application of these methods shows significant differences in the use of metals by archaeal, bacterial, and eukaryotic organisms, the latter of which have significantly expanded their zinc proteome over evolutionary time to cope with the increased need for regulatory proteins arising from increased cell complexity and cellular differentiation. On the other hand, non-heme iron-binding proteins have not evolved significantly from simpler to more complex organisms; in fact, their number has actually diminished, with the exception of *Arabidopsis thaliana*, in eukaryotes, possibly reflecting reduced iron bioavailability. Finally, copper proteins are relatively uncommon in all organisms across the three domains of life and their number and distribution does not show significant trends. It should be kept in mind, however, that at the level of the individual protein metal binding is also dependent on the action of suitable metal-delivery pathways, which may always leave some uncertainty regarding the correctness of predictions but also of experimental data when the metalloprotein has been expressed heterologously or manipulated in some ways.

[1] I. Bertini, A. Sigel, H. Sigel, *Handbook on Metalloproteins*, 1st ed., Marcel Dekker, New York, **2001**.
[2] E. Marcotte, M. Pellegrini, M. Thompson, T. Yeates, D. Eisenberg, *Nature* **1999**, *402*, 83–86.
[3] M. Y. Galperin, E. V. Koonin, *Nat. Biotechnol.* **2000**, *18*, 609–613.
[4] R. L. Tatusov, E. V. Koonin, D. J. Lipman, *Science* **1997**, *278*, 631–637.
[5] A. Valencia, *Curr. Opin. Struct. Biol.* **2005**, *15*, 267–274.
[6] J. M. Thornton, A. E. Todd, D. Milburn, N. Borkakoti, C. A. Orengo, *Nat. Struct. Biol.* **2000**, *7* (suppl.), 991–994.
[7] J. Moult, E. Melamud, *Curr. Opin. Struct. Biol.* **2000**, *10*, 384–389.
[8] J. C. Whisstock, A. M. Lesk, *Q. Rev. Biophys.* **2003**, *36*, 307–340.
[9] R. L. Marsden, J. A. Ranea, A. Sillero, O. Redfern, C. Yeats, M. Maibaum, D. Lee, S. Addou, G. A. Reeves, T. J. Dallman, C. A. Orengo, *Philos. Trans. R. Soc., B* **2006**, 425–440.
[10] J. D. Watson, R. A. Laskowski, J. M. Thornton, *Curr. Opin. Struct. Biol.* **2005**, *15*, 275–284.
[11] R. A. Laskowski, J. D. Watson, J. M. Thornton, *J. Struct. Funct. Genomics* **2003**, *4*, 167–177.
[12] D. Pal, D. Eisenberg, *Structure* **2005**, *13*, 121–130.
[13] M. R. Brent, *Genome Res.* **2005**, *15*, 1777–1786.
[14] A. J. Windsor, T. Mitchell-Olds, *Curr. Opin. Biotechnol.* **2006**, *17*, 161–167.
[15] I. Bertini, A. Rosato, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 3601–3604.
[16] I. Bertini, C. Luchinat, A. Provenzani, A. Rosato, P. R. Vasos, *Proteins Struct. Funct. Genet.* **2002**, *46*, 110–127.
[17] K. Degtyarenko, S. Contrino, *BMC Struct. Biol.* **2004**, *4*, 3.
[18] C. Andreini, I. Bertini, A. Rosato, *Bioinformatics* **2004**, *20*, 1373–1380.

[19] R. A. Scott, J. E. Shokes, N. J. Cosper, F. E. Jenney, M. W. W. Adams, *J. Synchrotron Rad.* **2005**, *12*, 19–22.
[20] M. Hogbom, U. B. Ericsson, R. Lam, H. M. Bakali, E. Kuznetsova, P. Nordlund, D. B. Zamble, *Mol. Cell Proteomics.* **2005**, *4*, 827–834.
[21] W. Shi, C. Zhan, A. Ignatov, B. A. Manjasetty, N. Marinkovic, M. Sullivan, R. Huang, M. R. Chance, *Structure (Cambridge, MA, U. S.)* **2005**, *13*, 1473–1486.
[22] C. Andreini, L. Banci, I. Bertini, A. Rosato, *J. Proteome Res.* **2006**, *5*, 196–201.
[23] C. Andreini, L. Banci, I. Bertini, S. Elmi, A. Rosato, *Proteins* **2007**, *67*, 317–324.
[24] C. Andreini, L. Banci, I. Bertini, A. Rosato, *J. Proteome Res.* **2006**, *5*, 3173–3178.
[25] L. Banci, M. Piccioli, *Cobalt(II) and Nickel(II) Substituted Proteins* (Eds.: D. M. Grant, R. K. Harris), *Encyclopedia of Magnetic Resonance*, *Encycl. of Nuclear Magn. Reson.*, **1996**, p. 1365.
[26] L. Banci, I. Bertini, E. Gallori, C. Luchinat, F. Paoletti, M. Polsinelli, M. S. Viezzoli, *J. Inorg. Biochem.* **1987**, *30*, 77–85.
[27] S. Benini, S. Ciurli, C. Luchinat, *Inorg. Chem.* **1995**, *34*, 417–420.
[28] I. Bertini, I. Gelis, N. Katsaros, C. Luchinat, A. Provenzani, *Biochemistry* **2003**, *42*, 8011–8021.
[29] P. Muller, S. Kopke, G. M. Sheldrick, *Acta Crystallogr., Sect. D* **2003**, *59*, 32–37.
[30] J. W. Schymkowitz, F. Rousseau, I. C. Martins, J. Ferkinghoff-Borg, F. Stricher, L. Serrano, *Proc. Natl. Acad. Sci. U. S. A* **2005**, *102*, 10147–10152.
[31] E. Pidcock, G. R. Moore, *J. Biol. Inorg. Chem.* **2001**, *6*, 479–489.
[32] M. M. Yamashita, L. Wesson, G. Eisenman, D. Eisenberg, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 5648–5652.
[33] M. M. Harding, *Acta Crystallogr., Sect. D.* **2006**, *62*, 678–682.
[34] M. M. Harding, *Acta Crystallogr., Sect. D.* **2004**, *60*, 849–859.
[35] M. Babor, H. M. Greenblatt, M. Edelman, V. Sobolev, *Proteins* **2005**, *59*, 221–230.
[36] C. Zhan, E. V. Fedorov, W. Shi, U. A. Ramagopal, R. Thirumuruhan, B. A. Manjasetty, S. C. Almo, A. Fiser, M. R. Chance, A. A. Fedorov, *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2005**, *61*, 959–963.
[37] J. T. Bolin, A. E. Ronco, T. V. Morgan, L. E. Mortenson, N. H. Xuong, *Proc. Natl. Acad. Sci. USA* **1993**, *90*, 1078–1082.
[38] G. J. Kornhaber, D. Snyder, H. N. Moseley, G. T. Montelione, *J. Biomol. NMR* **2006**, *34*, 259–269.
[39] I. Bertini, C. Luchinat, M. Piccioli, *Methods Enzymol.* **2001**, *339*, 314–340.
[40] I. Bertini, Y.-M. Lee, C. Luchinat, M. Piccioli, L. Poggi, *ChemBioChem* **2001**, *2*, 550–558.
[41] D. F. Hansen, J. J. Led, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 1738–1743.
[42] L. M. Utschig, J. W. Bryson, T. V. O'Halloran, *Science* **1995**, *268*, 380–385.
[43] L. M. Utschig, J. G. Wright, G. Dieckmann, V. L. Pecoraro, T. V. O'Halloran, *Inorg. Chem.* **1995**, *34*, 2497–2498.
[44] K. Zangger, G. Oz, J. D. Otvos, I. M. Armitage, *Protein Sci.* **1999**, *8*, 2630–2638.
[45] S. S. Narula, D. R. Winge, I. M. Armitage, *Biochemistry* **1993**, *32*, 6773–6787.
[46] M. A. Marti-Renom, A. C. Stuart, A. Fiser, R. Sanchez, F. Melo, A. Sali, *Annu. Rev. Biophys. Biomol. Struct.* **2000**, *29*, 291–325.
[47] A. Sali, T. L. Blundell, *J. Mol. Biol.* **1993**, *234*, 779–815.
[48] L. Banci, I. Bertini, A. Rosato, G. Varani, *J. Biol. Inorg. Chem.* **1999**, *4*, 824–837.
[49] L. Banci, I. Bertini, S. Ciurli, A. Dikiy, J. Dittmer, A. Rosato, G. Sciara, A. R. Thompsett, *ChemBioChem* **2002**, *3*, 299–310.
[50] I. Bartalesi, I. Bertini, P. Hajieva, A. Rosato, P. R. Vasos, *Biochemistry* **2002**, *41*, 5112–5119.
[51] C. Andreini, L. Banci, I. Bertini, C. Luchinat, A. Rosato, *J. Proteome Res.* **2004**, *3*, 21–31.
[52] C. Andreini, L. Banci, I. Bertini, S. Elmi, A. Rosato, *J. Proteome Res.* **2005**, *4*, 881–888.
[53] F. Arnesano, L. Banci, I. Bertini, S. Ciofi-Baffoni, E. Molteni, D. L. Huffman, T. V. O'Halloran, *Genome Res.* **2002**, *12*, 255–271.
[54] F. Musiani, B. Zambelli, M. Stola, S. Ciurli, *J. Inorg. Biochem.* **2004**, *98*, 803–813.
[55] S. Ciurli, F. Musiani, *Photosynth. Res.* **2005**, *85*, 115–131.
[56] B. Zambelli, M. Stola, F. Musiani, K. De Vriendt, B. Samyn, B. Devreese, J. Van Beeumen, P. Turano, A. Dikiy, D. A. Bryant, S. Ciurli, *J. Biol. Chem.* **2005**, *280*, 4684–4695.
[57] S. C. Tosatto, G. M. Giacometti, G. Valle, P. Costantini, *Biochem. Biophys. Res. Commun.* **2006**, *339*, 277–283.
[58] J. S. Oakhill, B. J. Sutton, A. R. Gorringe, R. W. Evans, *Protein Eng. Des. Sel.* **2005**, *18*, 221–228.
[59] N. Hulo, A. Bairoch, V. Bulliard, L. Cerutti, E. De Castro, P. S. Langendijk-Genevaux, M. Pagni, C. J. Sigrist, *Nucleic Acids Res.* **2006**, *34*, D227–D230.
[60] R. D. Finn, J. Mistry, B. Schuster-Bockler, S. Griffiths-Jones, V. Hollich, T. Lassmann, S. Moxon, M. Marshall, A. Khanna, R. Durbin, S. R. Eddy, E. L. Sonnhammer, A. Bateman, *Nucleic Acids Res.* **2006**, *34*, D247–D251.
[61] A. Andreeva, D. Howorth, S. E. Brenner, T. J. Hubbard, C. Chothia, A. G. Murzin, *Nucleic Acids Res.* **2004**, *32*, D226–D229.
[62] A. Passerini, M. Punta, A. Ceroni, B. Rost, P. Frasconi, *Proteins* **2006**, *65*, 305–316.
[63] J. E. Coleman, *Curr. Opin. Chem. Biol.* **1998**, *2*, 222–234.
[64] I. Bertini, C. Luchinat, M. S. Viezzoli, "Metal substitution as a tool for the investigation of zinc proteins" in *Zinc Enzymes* (Eds.: I. Bertini, C. Luchinat, W. Maret, M. Zeppezauer), Birkhäuser, Basel, **1986**, pp. 27–47.
[65] J. M. Matthews, M. Sunde, *IUBMB Life* **2002**, *54*, 351–355.
[66] J. L. Erbe, K. B. Taylor, L. M. Hall, *Nucleic Acids Res.* **1995**, *23*, 2472–2478.
[67] W. J. Cook, S. R. Kar, K. B. Taylor, L. M. Hall, *J. Mol. Biol.* **1998**, *275*, 337–346.
[68] B. F. Volkman, S. J. Wilkens, A. L. Lee, B. Xia, W. M. Westler, R. Berger, J. L. Markley, *J. Am. Chem. Soc.* **1999**, *121*, 4677–4683.
[69] R. Bau, D. C. Rees, D. M. Kurtz, Jr., R. A. Scott, H. S. Huang, M. W. W. Adams, M. K. Eidsness, *J. Biol. Inorg. Chem.* **1998**, *3*, 484–493.
[70] I. Bertini, D. M. Kurtz, Jr., M. K. Eidsness, G. Liu, C. Luchinat, A. Rosato, R. A. Scott, *J. Biol. Inorg. Chem.* **1998**, *3*, 401–410.
[71] Z. Dauter, K. S. Wilson, L. C. Sieker, J.-M. Moulis, J. Meyer, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 8836–8840.
[72] L. Thony-Meyer, *Biochemistry* **2003**, *42*, 13099–13105.
[73] L. Thony-Meyer, *Biochem. Soc. Trans.* **2002**, *30*, 633–638.
[74] F. Arnesano, L. Banci, P. D. Barker, I. Bertini, A. Rosato, X. C. Su, M. S. Viezzoli, *Biochemistry* **2002**, *41*, 13587–13594.
[75] R. Lill, R. Dutkiewicz, H. P. Elsasser, A. Hausmann, D. J. Netz, A. J. Pierik, O. Stehling, E. Urzica, U. Muhlenhoff, *Biochim. Biophys. Acta* **2006**, *1763*, 652–667.
[76] S. S. Mansy, J. A. Cowan, *Acc. Chem. Res.* **2004**, *37*, 719–725.
[77] H. Beinert, R. H. Holm, E. Münck, *Science* **1997**, *277*, 653–659.
[78] C. Luchinat, F. Capozzi, D. Bentrop, *Iron-Sulfur Proteins* (Eds.: I. Bertini, A. Sigel, H. Sigel), *Handbook on Metalloproteins*, Marcel Dekker, New York, **2001**.
[79] I. Bertini, C. Luchinat, A. Rosato, *NMR Spectra of Iron-Sulfur Proteins* (Eds.: A. G. Sykes, R. Cammack), *Advances in Inorganic Chemistry*, Academic Press, San Diego, **1999**, p. 251.
[80] I. Bertini, S. Ciurli, C. Luchinat, *Struct. Bonding (Berlin)* **1995**, *83*, 1–54.
[81] L. Banci, I. Bertini, C. Luchinat, P. Turano, *Special Cofactors and Metal Clusters* (Eds.: I. Bertini, B. H. Gray, J. S. Valentine,

E. I. Stiefel), *Biological Inorganic Chemistry: Structure & Reactivity*, University Science Books, Sausalito (CA), **2006**, p. 43.
[82] I. Bertini, G. Cavallaro, A. Rosato, *Chem. Rev.* **2006**, *106*, 90–115.
[83] L. A. Finney, T. V. O'Halloran, *Science* **2003**, *300*, 931–936.
[84] M. D. Harrison, C. E. Jones, C. T. Dameron, *J. Biol. Inorg. Chem.* **1999**, *4*, 145–153.
[85] T. V. O'Halloran, V. C. Culotta, *J. Biol. Chem.* **2000**, *275*, 25057–25060.
[86] L. Banci, A. Rosato, *Acc. Chem. Res.* **2003**, *36*, 215–221.
[87] L. Banci, I. Bertini, S. Ciofi-Baffoni, I. Leontari, M. Martinelli, P. Palumaa, R. Sillard, S. Wang, *Proc. Natl. Acad. Sci. U. S. A* **2007**, *104*, 15–20.
[88] L. Banci, I. Bertini, F. Cantini, N. DellaMalva, T. Herrmann, A. Rosato, K. Wuthrich, *J. Biol. Chem.* **2006**, *281*, 29141–29147.
[89] L. Banci, I. Bertini, F. Cantini, I. C. Felli, L. Gonnelli, N. Hadjiliadis, R. Pierattelli, A. Rosato, P. Voulgaris, *Nat. Chem. Biol.* **2006**, *2*, 367–368.
[90] F. Arnesano, L. Banci, I. Bertini, M. Martinelli, *J. Proteome Res.* **2005**, *4*, 63–70.
[91] L. Banci, I. Bertini, G. Cavallaro, A. Rosato, *J. Proteome Res.* **2007**, *6*, 1568–1579.
[92] D. Strausak, J. F. Mercer, H. D. Hermann, W. Stremmel, G. Multhaup, *Brain Res. Bull.* **2001**, *55*, 175–185.
[93] L. Thony-Meyer, *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 337–376.
[94] H. Schulz, H. Hennecke, L. Thony-Meyer, *Science* **1998**, *281*, 1197–1200.
[95] E. Enggist, L. Thony-Meyer, P. Guntert, K. Pervushin, *Structure* **2002**, *10*, 1551–1557.
[96] A. L. Lamb, A. K. Wernimont, R. A. Pufahl, V. C. Culotta, T. V. O'Halloran, A. C. Rosenzweig, *Nat. Struct. Biol.* **1999**, *6*, 724–729.
[97] A. L. Lamb, A. S. Torres, T. V. O'Halloran, A. C. Rosenzweig, *Nat. Struct. Biol.* **2001**, *8*, 751–755.
[98] I. Anastassopoulou, L. Banci, I. Bertini, F. Cantini, E. Katsari, A. Rosato, *Biochemistry* **2004**, *43*, 13046–13053.
[99] E. Balatri, L. Banci, I. Bertini, F. Cantini, S. Ciofi-Baffoni, *Structure* **2003**, *11*, 1431–1443.
[100] H. Remaut, N. Safarov, S. Ciurli, J. J. Van Beeumen, *J. Biol. Chem.* **2001**, *276*, 49365–49370.
[101] Y. F. Lin, A. R. Walmsley, B. P. Rosen, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 15617–15622.
[102] D. Namgaladze, H. W. Hofer, V. Ullrich, *J. Biol. Chem.* **2002**, *277*, 5962–5969.
[103] M. R. Swingle, R. E. Honkanen, E. M. Ciszak, *J. Biol. Chem.* **2004**, *279*, 33992–33999.
[104] M. Schmidt, B. Meier, F. Parak, *J. Biol. Inorg. Chem.* **1996**, *1*, 532–541.

Received: February 8, 2007
Published Online: May 16, 2007

