University of Florence

International Doctorate in Structural Biology

Cycle XIX (2004-2006)


# Copper homeostasis: Genome analysis and characterization of proteins involved

Ph.D. thesis of

**Manuele Martinelli**

| Tutor | Coordinator |
|---|---|
| Prof. Lucia Banci | Prof. Claudio Luchinat |

S.S.D. CHIM/03

This thesis has been approved by the University of Florence, the University of Frankfurt and the Utrecht University

# Contents

# 1

# INTRODUCTION

## 1.1 The role of copper in biological systems

Metals play an essential role as trace elements in most of the biochemical processes at the basis of Life[1]. Current estimates indicate that about a third of all structurally characterized proteins contain metal ions either as a structural component or as a catalytic co-factor[2].

The use of metals in living organisms is, indeed, important mainly due to their redox properties and to the feasibility of metal ion transfer in different locations or compartments of the cell. The metals are implicated in various biological processes such as electron transfer reactions, oxygen transport, and in a large number of catalytic processes.

Many proteins (called metalloproteins) need to bind one or more metal ions to be able to perform their function, either because the metal ion is involved in the catalytic mechanism or because it stabilizes/determines the protein tertiary or quaternary structure. Metal ions are also very important for the structure and function (in the case of RNA) of nucleic acids. The intracellular concentration of several metals as well as their distribution among the various cell compartments and their incorporation into metalloproteins is tightly controlled[2,3,4]. A proper balance of the equilibria involved in these control processes is necessary for a healthy phenotype. Previously, it was thought that metal enzymes captured their essential cofactor by collision processes with free metal ions or metal complexes with low-molecular-weight ligands. Recently, several mechanisms how metals are transported and inserted into diverse protein location have been identified and this research is an emerging field in the scientific world[5,6]**.** Both eukaryotes and prokaryotes have evolved several mechanisms that ensure efficient metal homeostasis. From biochemical, inorganic, structural, and mechanistic studies a new family of soluble metal receptor, know as "metallochaperones", have been found to interact with metal ions and deliver them to several targets in the cell[6]. These proteins have the role to guiding and/or inserting a metal cofactor into the active site of a target enzyme

An important metal present as a co-factor in many proteins is copper. Biological systems did not utilize copper before the advent of atmospheric oxygen[7]. In the prevailing reducing conditions before this event, copper was in the water-insoluble Cu(I) state, in the form of highly insoluble sulfides, and was not available for biological processes. Cyanobacteria are thought to be responsible for the beginning of dioxygen ($O_2$) production about $10^9$ years ago. The appearance of a significant $O_2$ concentration in the atmosphere required another 200–300 million years because the oxygen produced was initially

consumed by the oxidation of ferrous iron in the oceans. The advent of oxygen was a catastrophic event for most living organisms and it irreversibly changed life on earth. In contrast to the oxidation of iron and its loss of bioavailability as insoluble Fe(III), the oxidation of insoluble Cu(I) led to soluble Cu(II). Whereas enzymes involved in anaerobic metabolism were designed to operate in the lower portion of the redox spectrum, the arrival of dioxygen created the need for a new redox active metal which could attain higher redox potentials. Copper, now bioavailable, was ideally suited to exploit the oxidizing power of dioxygen. Copper began to be used in energy-capture systems like cytochrome *c* oxidase. The arrival of copper also coincided with the development of multicellular organisms which had extracellular cross-linked matrices capable of resisting attacks by oxygen free radicals[7].

In this frame the case of copper represents the most studied example of metal homeostasis. Copper is an essential trace element that plays a vital role as a catalytic co-factor for a variety of metalloenzymes including superoxide dismutase (for protection against free radicals), cytochrome *c* oxidase (mitochondrial electron transport chain), tyrosinase (pigmentation), peptidylglycine alpha-amidating mono-oxygenase (PAM) (neuropeptide and peptide hormone processing) and lysyl oxidase (collagen maturation)[8,9,10]. At the same time, copper is toxic to both eukaryotic and prokaryotic cells, due to its ability to catalyze the generation of hydroxyl radicals (OH·) via the so called Fenton-type reaction:

$$Cu^{+} + H_2O_2 \Leftrightarrow Cu^{2+} + OH^{-} + OH\cdot \qquad (1)$$

Furthermore, oxidative stress carried out by partially reduced oxygen species (ROS), could be amplified by copper reactivity, leading to the impairment of essential molecules such as lipids, proteins and DNA[11]. Therefore, several genes and their respective gene products are dedicated specifically to the transport of copper across cellular membranes, the trafficking of copper to cupro-proteins, and the sequestration or elimination of any copper excess. Many of these genes and proteins involved in copper metabolism are regulated by intracellular copper concentrations, either at the level of transcription, protein localization, or stability. Finally, there must exist a responsive form of communication between the proteins responsible for copper homeostasis, the organelles in which they reside, the external environment and perhaps other cellular processes. At present, several of the participants in copper homeostasis have been identified and are in various stages of

structural and functional characterization[12]. How these proteins are regulated is less well understood, and the mechanisms for intracellular communication are not yet clear. It was indeed proven that essentially no free copper is available in the cytoplasm of eukaryotic cell being the total cytoplasmatic free copper concentration less than $10^{-18}$M[6]. This feature eliminates the possibility that weakly bound copper ions can freely diffuse in the cell until a physiologically important active site captures the metal as it was observed *in vitro* where many copper enzymes easily acquire their metal without an auxiliary protein.

## 1.2 Copper and diseases

Keeping in mind the dual aspect of copper (need and toxicity), it is expected that the alteration of copper homeostasis leads to errors of metabolism and damage to cells, resulting in pathological conditions[13]. This is well exemplified by two inherited diseases of copper metabolism in humans, Menkes and Wilson syndromes: in both cases, the impairment by mutation of two homologous copper transport ATPases (named ATP7A and ATP7B, respectively), which are selectively expressed in different tissues, leads either to the decrease or to the overload of copper in the cells.

Menkes disease (MD) is a fatal X-linked copper deficiency disorder due to mutations affecting Menkes gene (*MNK).* Defective export of copper by ATP7A from the intestinal cells is responsible for systemic copper deficiency; the copper content is particularly low in the brain because mutated ATP7A is also expressed at the blood–brain barrier. Individuals affected by Menkes disease usually die in early childhood. Wilson's disease (WD) is more frequent, with a later onset in life and is an autosomal recessive disorder. Mutations of the ATP7B exporting copper from the liver to the bile and circulation lead to liver failure as a result of the very high concentrations of copper accumulated in this organ. If this disease is not treated, death can result from liver failure. Both of the inherited diseases of copper homeostasis are accompanied by neurodegeneration.

Moreover, some very interesting recent findings have linked the imbalance of copper homeostasis to some neurodegenerative diseases as Alzheimer's disease (AD), Creutzfeldt-Jakob disease, Parkinson's disease (PD) and Amyotrophic Lateral Sclerosis disease (ALS) [15,16]. It is probable that aberrant reactivity of copper is a major source of production of ROS, which in turn is responsible for the more global oxidative stress parameters observed in these diseases. These disorders may all be classified as "conformational" in that they all show, as hallmarks, conformationally altered proteins, which precipitate, form aggregates,

and may be responsible of cell damage[17]. Direct interaction between copper and the proteins that are the pathological hallmark of these diseases has been reported, and thus, redox activity of the metal may also be involved in the process of protein misfolding. The molecular mechanisms leading to neurodegeneration in all these diseases are still unknown. However, metal-mediated oxidative stress could induce a cascade of events, including mitochondrial dysfunction that may be responsible for cell death.

A large body of evidence supports the role of metal accumulation in aged brain as a causal or at least ancillary factor in the etiology of neurodegenerative disorders typical of late life, namely AD and PD[16,18]. For example, amyloid precursor protein (APP) and the native Prion Protein which are altered in Alzheimer diseases and in Spongiform encephalopathies are both membrane proteins showing conserved copper-binding sites in their N-terminal region[19], this, together with other experimental clues, led to the suggestion that they may be implicated in copper transport and regulation at the neuron plasma membrane and synapses[19, 21]. APP binds Cu(II) by two conserved histidines at a cysteine-rich region, which shows high structural homology to copper chaperones[19]. APP seems to be a key modulator of neuronal copper homeostasis, because knockout of the protein results in increased copper levels[22], whereas its over-expression results in reduced copper levels in the mouse brain[23]. Copper bound to APP and β-amyloid can be reduced to Cu(I), with ROS generation[24,25]. It has to be remembered that copper concentration is higher in the neuropil of Alzheimer's patients and in the core and periphery of plaques[26], where β-amyloid also accumulates; thus a powerful source of oxidative stress is operative in these regions that show elevated markers of oxidative stress. Recent findings on the efficacy of copper-chelating agents in ameliorating behavioral and biochemical changes in AD confirm a role for copper in the pathogenesis of that disease and indicate novel therapeutic tools for the prevention and treatment of AD[27].

In PD, it has also been demonstrated that Cu(II) prompts α-synuclein to form self-oligomers and that the acidic C-terminus of the protein is essential for copper interaction[28]. Interestingly, fragments of α-synuclein represent the non-β-amyloid component of senile plaques in AD, thus reinforcing the role of copper in AD progression.

In ALS, mutations of the cuproenzyme SOD1 gene have been associated with the diseases progression[29]. Treatment with copper chelators has been shown to slow the progression of the degeneration in both animal and cell culture models of familiar ALS (fALS)[30]suggesting a noxious behavior of the metal when not properly bound to the protein.

Alternatively, in the oligomerization hypothesis, misfolded mutant SOD1 proteins oligomerize into increasingly higher molecular weight aggregates that ultimately lead to neurodegeneration[30,31].

Moreover, recently clinical trials have established that Cu privation by diet or by Cu chelators diminishes tumor's ability to mount an angiogenic response[9].

## 1.3 Copper routes in yeast

An accessible model for investigation of eukaryotic copper transport is *Saccharomyces cerevisiae.* It is an extremely attractive eukaryotic model system due to the ease with which it can be grown, the facility with which it can be genetically manipulate and moreover its complete genome has been completely sequenced.

Recently, biochemical and structural studies have shed light on copper pathways **(Fig.1)**, and the chemical mechanisms for copper shunting, while gene knockout experiments in mice have revealed fascinating aspects of the role of copper chaperones in mammalian physiology and development.

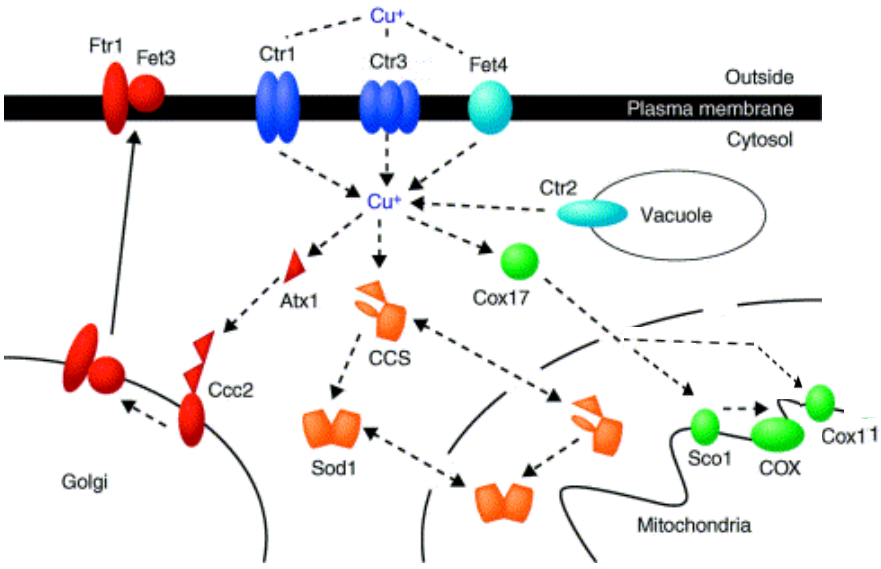



**Fig. 1:** *Copper transport and distribution in yeast.* (Reprinted from Puig S et.al Curr Opin Chem Biol. 2002 )

Depending on extracellular copper concentrations, copper is transported inside the cell by two high-affinity copper transporters Ctr1 and Ctr3 ($K_m$ for copper ~1–5 μM), or by a low-affinity Cu/Fe-transporter Fet4 ($K_m$ for copper ~35μM)[32]. Cu(I) is the substrate for the Ctr family members which are relatively small proteins containing three

transmembrane domains[33]. Prior to the uptake, copper, which is present in the extracellular space in the Cu(II) form, is thought to be reduced to Cu(I) by one or more cell-surface Fe(III)/Cu(II) reductases encoded by the FRE1 genes[34]. A conserved feature of some Ctr importers is an N-terminal segment that contains one or more "Mets" motifs (MxxM or MxM), which have been shown by deletion studies in yeast and human cells, to be important for survival under copper starvation. These "Mets" motifs are part of the extracellular domain and are involved in the acquisition of copper ions to facilitated import[35]. Unlike some other high affinity metal transporters (such as ATP7A and ATP7B), Ctr proteins do not require ATP for copper import[36]. Their transport ability is stimulated by extracellular $K^+$ and probably facilitated by the extremely low intracellular concentration of free copper.

The delivery of copper to target cuproenzymes depends on an elegant metallochaperone system **(Fig. 1)**. Copper is deployed to mitochondrial cytochrome *c* oxidase, via a pathway involving Cox17, Sco1 and Cox11 proteins[37-40], to cytosolic copper zinc superoxide dismutase (Cu,Zn SOD), via a pathway involving a copper chaperone for Cu,Zn SOD, called CCS[6,41], or to the multicopper oxidase Fet3 in the secretory pathway involving the P-type ATPase copper transporter CCC2[42] and the copper chaperone ATX1[43,44]. In addition, cytosolic concentrations of free copper are typically maintained at low levels by metal scavenging systems like metallothioneins (MT) and by the Ctr2 protein which is suggested to play a role in the mobilization of intracellular pools of copper in the vacuoles[45].

Homologs of yeast copper proteins have also been discovered in other eukaryotes including *Arabidopsis thaliana*, *Caenorhabditis elegans*, mice, rats, sheep, humans and in prokaryotes. The latter lack the intracellular compartmentalization that is typical of eukaryotes (cyanobacteria being a notable exception); thus organelle-specific carriers of metals such as Cox17 may not be essential. In bacteria, for example, Cu,Zn SOD is a periplasmic enzyme in contrast with eukaryotes, where it is located mainly in the cytosol. Furthermore, bacteria do not express prokaryotic homologues to CCS and Cox17 but still a homologue to ATX1 protein, called CopZ, is present.

## 1.4 Copper delivery to Cytochrome c oxidase

A relevant example of a copper-dependent enzyme is the cytochrome *c* oxidase (CcO) which is the terminal enzyme in the energy transducing respiratory chain.

CcO is a membrane-bound redox-driven proton pump[46]. It is located in the inner mitochondrial membrane, where it catalyzes reduction of molecular oxygen to water and pumps protons across the membrane[47,48]. The energy stored by the proton gradient is subsequently utilized for ATP synthesis. The reduction of $O_2$ to $H_2O$ in the catalytic center of CcO generates energy that is necessary for proton translocation from the mitochondrial matrix (a region of low proton concentration and negative electrical potential) to the intermembrane space (which is in contact with the cytosol, a region of high proton concentration and positive electrical potential). Mammalian CcO consists of 13 polypeptide subunits, 3 of which (Cox1-Cox3) are encoded by the mitochondrial genome with the remaining 10 subunits encoded by the nuclear genome[49,50]. In addition to the subunits, over 30 distinct proteins are important for the assembly of CcO[51]. A number of these accessory proteins are important in the processing and translation of Cox1-Cox3 mRNA transcripts, in chaperoning the assembly process, and in the synthesis or delivery of cofactors. The cofactors in CcO include two copper sites ($Cu_A$ and $Cu_B$), two heme A moieties, a magnesium and a zinc ion (**Fig.2 and Fig.3**)[52].

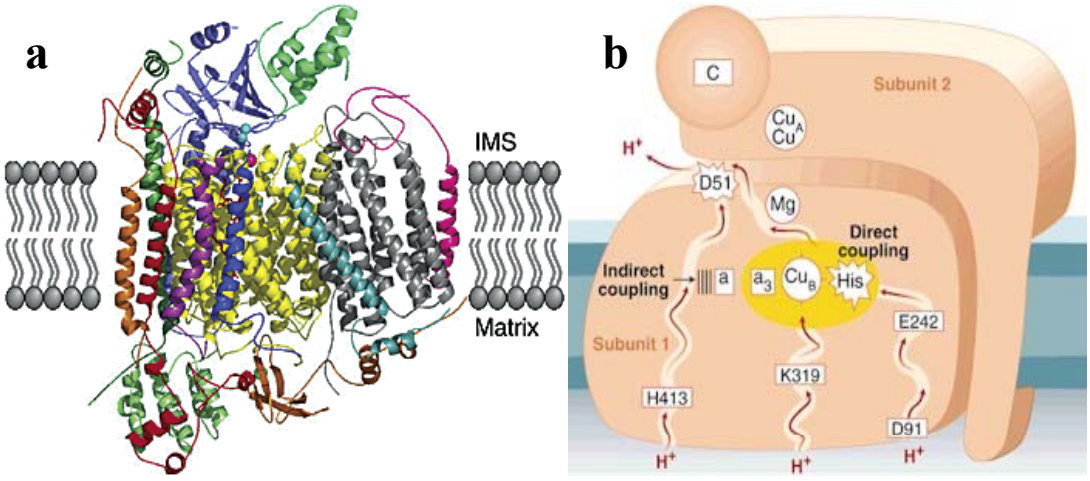



**Fig. 2:** *The structure of bovine CcO is shown with the 13 subunits colored differently (***a***). The cofactors are largely obscured by the polypeptide chains. Panel* **b** *shows the cofactors.*

(Panel **a** has been reprinted from Carr HS et al. (2003) Acc Chem Res. May;36(5):309-16)

Little is known about the delivery of Zn and Mg ions to the mitochondrion or whether specific accessory factors are necessary for the insertion of these ions into CcO. More information is available concerning copper delivery and insertion. A total of three copper ions have to be inserted into two subunits Cox1 and Cox2 which contain the $Cu_B$ and the $Cu_A$ centers, respectively. The former has one copper ion buried 13 Å below the membrane surface (**Fig. 3**), while the second contains a binuclear copper site located into the inner membrane space of the mitochondria in eukaryotes, in the periplasm of Gram-negative bacteria, and in the outer leaflet of the cell membrane of Gram-positive bacteria.

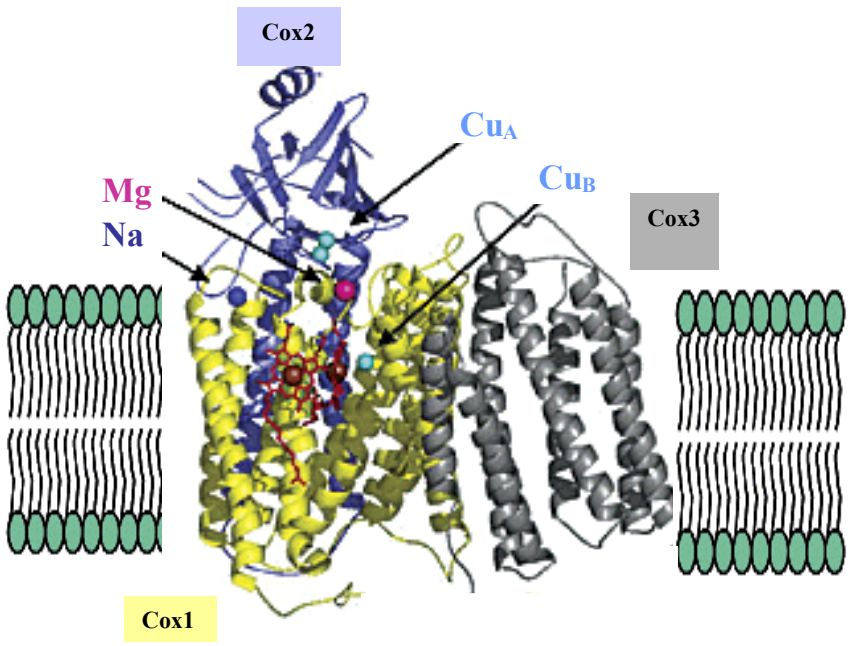



**Fig.3**: *Structural representation of the mitochondrial-encoded subunits that make up the catalytic core of the enzyme, Cox1 (yellow), Cox2 (blue) and Cox3 (grey)within their cofactors.*
*(* Reprinted from Carr HS et.al. (2003) Acc Chem Res. May;36(5):309-16 )

As both Cox1 and Cox2 are synthesized inside the mitochondria, the three Cu atoms must be imported from the cytoplasm or from the mitochondrial matrix. CcO is metallated by a pathway that consists of Cox17 and the co-chaperones, Sco1 and Cox11. Cox17 delivers copper to Cox11 for subsequent insertion into the $Cu_B$ site in Cox1, and to Sco1 for insertion into the $Cu_A$ site of Cox2 (**Fig. 4**). Initially, the most obvious candidate for delivery of copper to the mitochondrion was Cox17 protein which exists indeed in both the cytoplasm and mitochondrial inner membrane space (IMS)[53]. Moreover it is known that

yeast lacking Cox17 (*cox17*Δ yeast) are respiratory-deficient due to a complete lack of CcO activity[54]. It can be restored by the addition of 0.4% copper salts to the growth medium. This result is consistent with Cox17 functioning in copper delivery to CcO. The simple prediction was that Cox17 would shuttle Cu(I) into the IMS to be used in the assembly of CcO. This theory was strengthened by the observation that Cox17 is a copper-binding protein[55]. However, Cox17 tethered to the mitochondrial inner membrane by a fusion to the transmembrane domain of the inner membrane protein Sco2 was able to reverse the respiratory defect of *cox17*Δ cells and restored normal CcO activity[56]. These results suggest that the function of Cox17 is confined to the mitochondrial intermembrane space[57].

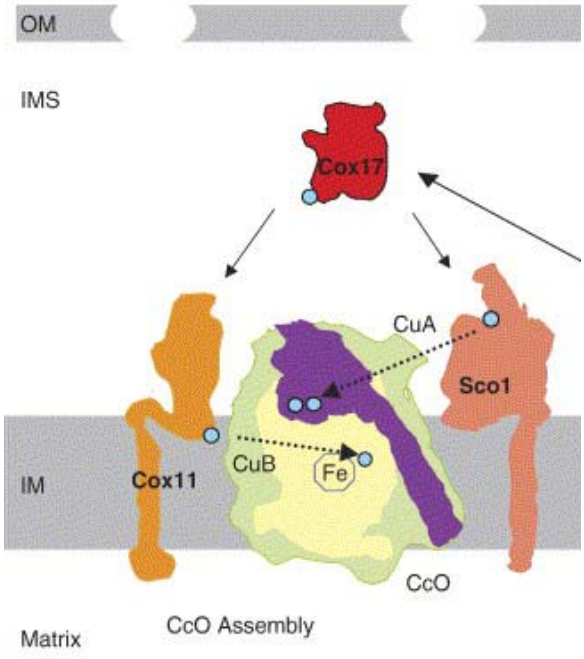



**Fig.4:** *Scheme of Cu metalation of COX1 and COX2 mediated by Cox11 and Sco1, respectively. Cox17* is postulated to transfer Cu(I) to Sco1 prior to the transfer to the $Cu_A$ site in COX2 and to transfer Cu(I) to Cox11 prior to transfer to the buried $Cu_B$ site in COX1. (Reprinted from Cobine PA et.al (2006) Biochim Biophys Acta. Jul;1763(7):759-72.)

In yeast, Cox11 and Sco1 appear to be co-metallochaperones assisting Cox17 in the metalation of CcO (**Fig. 4**). Sco1 and Cox11 proteins are known to be implicated in Cu insertion reaction into the $Cu_A$ site of COX2 and into the $Cu_B$ site of COX1 subunits, respectively.

Yeast Sco1 is a 30kDa protein anchored by a single trans-membrane segment of about 17 aminoacids to the inner mitochondrial membrane. Protein sequence analysis has

identified a potential metal binding region, CxxxCP, that is conserved in all mitochondrial Sco proteins and their bacterial homologues. Both cysteines in yeast Sco1 are essential for its function[58]. X-ray absorption spectroscopy, combined with functional studies on yeast Sco1 (Sco1p) and human Sco1(HSco1), suggested that the two cysteines present in the conserved motif and a conserved histidine in the C-terminal end are involved in copper(I) binding[59]. However, the X-ray structure of Sco1p exhibits a thioredoxin-like (Trx-like) fold, similar to the one observed by NMR spectroscopy for a homologue from *Bacillus subtilis*[60]. Upon superposition of the core elements defining the Trx-fold, the conserved cysteines of the CXXXC motif, found in all Sco homologues are localized in a position analogous to that of the active site CXXC cysteines in Trx and other related oxidoreductases. The striking structural similarities between HSco1 and Trx, and especially the position of the reactive cysteines, suggested that Hsco1 might have disulfide reductase activity. The fact that Sco1 specifically delivers copper to the $Cu_A$ site in the COX2 subunit has been proposed on the basis of site specific mutations on the yeast protein[61]. All these data confirm that HSco1 assists the copper insertion in the $Cu_A$ center but leave open the possibility that it might be a protein with multiple functions. Some eukaryotic genomes contain another protein, highly similar to Sco1, called Sco2, which conserves the CXXXCP metal binding motif and whose role remains unresolved. In contrast to Sco1 mutations, deletion of yeast-Sco2 (Sco2p), an a homologue of Sco1, does not affect respiratory growth[62]. Like Sco1p, also Sco2p is anchored to the membrane by a single transmembrane (TM) domain in the N-terminal part, and presumably possesses the same topology as Sco1p[63]. Over-expression of Sco2p cannot substitute for the function of Sco1p, but can rescue the respiratory deficiency of a cox17Δ mutant, albeit less efficiently than over-expression of Sco1p and only in the presence of higher concentrations of exogenous copper. The C-terminal part of Sco1p can be replaced by the respective Sco2p portion, the resulting chimeric proteins are functional and able to overcome the respiratory defect of *sco1*Δ[58].

Additionally, Sco1 has been shown to interact with COX2 [64]. At this moment, it is not clear how Hsco1 and Hsco2 cooperate in the copper insertion into $Cu_A$. Are they directly involved in the copper transfer from Cox17 to Cox2 or do they assist this process via reduction of the conserved CXXXC motif of $Cu_A$ prior to the copper transfer?

If Sco proteins are primarily important in Cu(I) insertion into COX2, then how is Cu provided to COX1? This pathway seems to require another mitochondrial inner membrane protein, Cox11 (**Fig. 4**). Indeed, recently Hiser *et al.* demonstrated that the $Cu_B$ site is

absent in CcO purified from *Rhodobacter sphaeroides* lacking Cox11 gene[65]. Moreover, Tzagoloff *et.al.*[66] have shown that some Cox11 mutants lack CcO activity; however, RNA and protein synthesis of the core subunits I and II are normal, suggesting that Cox11 functions post-translationally to generate active CcO[66].

Yeast Cox11 is a 34kDa protein that resembles Sco1, with a single transmembrane helix of about 24 aminoacids just downstream of the N-terminal mitochondrial anchored targeting sequence. The C-terminal domain protrudes into the inner membrane space and, as with the soluble C-terminal domain of Sco1, binds a copper(I) ion. Recently, it has been shown that the soluble C-terminal domain of yeast Cox11 forms a dimer[67]. The protein binds one copper atom per monomer and in the dimer the two copper(I) ions are in close proximity.

## 1.5 CcO deficiency linked to HSco1/2 pathogenic mutations

Mutations in HSco1 and HSco2 genes cause severe, tissue-specific COX deficiencies owing to a failure in CcO assembly: HSco2 mutations are predominately associated with early-onset hypertrophic cardiomyopathy and encephalopathy[68,69,70], whereas one missense mutation (P174L) in HSco1 is founds to be associated with fatal infantile hepatoencephalomyopathy[71], a progressive neonatal disorder predominantly affecting the liver. Proline 174 , adjacent to the conserved CXXXC domain of HSco1, is completely conserved in eukaryotes. Overexpression of the P174L-Hsco1 mutant in Sco1pΔ yeast mutants impaired CcO assembly and induced loss of CcO activity [72]. It has been shown that the pathogenicity of the mutant does not result from its instability, but rather from an impaired function[73].

In HSco2, the E140K pathogenic mutation (**Fig. 5**) is adjacent to the postulated CXXXC conserved hypothetical copper binding motif. Also S225F pathogenic mutation is adjacent to the conserved His224, while the R171W mutation is close to a DXXXD conserved motif, which might also be involved in copper binding too. These observations suggest that the mutations probably impair the copper binding on HSco2 and/or the copper transfer process. On the other hand, using a policonal antibody against Hsco2, severely decreased level of HSco2 in fibroblast and myoblast from patients with different Sco2 mutations were observed suggesting an unstable gene product[70].

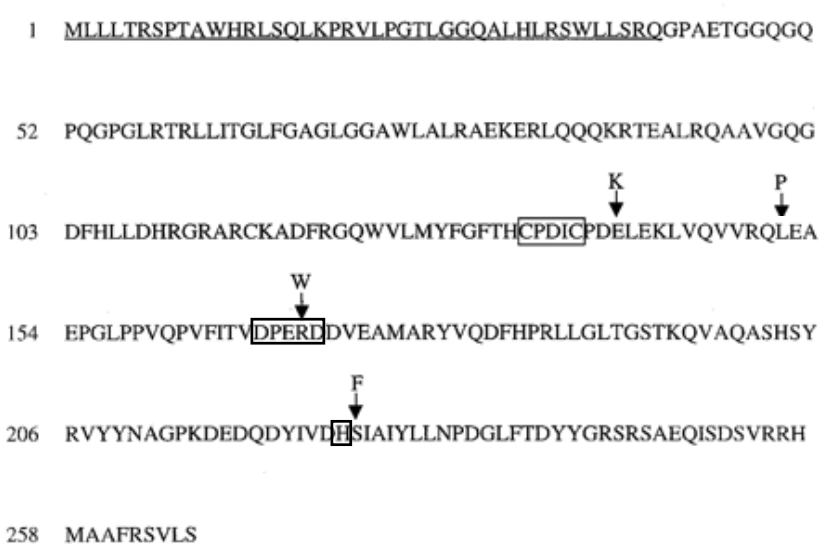



**Fig. 5***: Amino acid sequence of HSco2 showing the mitochondrial leader sequence (underlined), the copper-binding site (boxed) and the known missense mutations ( E140K; L151P; R171W; S225F)*

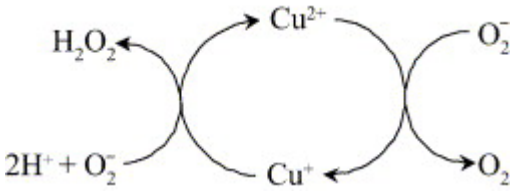
Both Hsco1 and Hsco2 are ubiquitary expressed and the reason for "tissue specificity" in the two disorders is unknown. Remarkably, COX deficiency in cells from patients with HSco2 mutations[59,64] and in bacteria harboring a null mutation in YpmQ (the SCO-like homolog in *Bacillus subtilis*[57], can be rescued by the addition of copper to the growth medium, but the mechanism by which this rescue occurs is unknown.

## 1.6 Cu,Zn Superoxide dismutase

Superoxide radical anions are physiologically produced in controlled amounts in animals and plants from a one-electron reduction of dioxygen occurring in several metabolic pathways[75-77]. Eukaryotes have evolved with two intracellular superoxide dismutase (SOD) enzymes capable to catalyze the disappearance of superoxide[78]. One, a manganese containing enzyme (SOD2), is located in the mitochondrial matrix[79] in close proximity to a primary endogenous source of superoxide, the mitochondrial respiratory chain. The other, a copper- and zinc-containing protein (SOD1), is an abundant cytosolic enzyme[80] that has also been found in the nucleus, peroxisomes, and lysosomes. Cu,Zn superoxide dismutase (Cu,Zn SOD) (known earlier as erythrocuprein) has long been known as the major copper containing protein in erythrocytes[81], but its enzymatic function was not discovered until 1969, when McCord and Fridovich found that 'erythrocuprein' dismutes superoxide[82]. SOD1 catalyzes a disproportionation reaction turning two molecules of superoxide into one molecule of $O_2$ and one of $H_2O_2$:

$$Cu^{2+} + O_2^- \leftrightarrow Cu^+ + O_2$$
$$Cu^+ + 2H^+ + O_2^- \leftrightarrow Cu^{2+} + H_2O_2$$

generating a cycling 'ping-pong' mechanism:




Cu,Zn SOD is found in almost all eukaryotic cells and in very few prokaryotes. Structural studies reveal that all eukaryotic SOD1s are homodimeric proteins that contain one copper, one zinc, and one intrasubunit disulfide bond per monomer. A considerable sequence homology is present among Cu,Zn SOD proteins from various species so far examined, within the mammalian species, human, rat, pig, and horse the homology is around 80%.

The X-ray structure of the oxidized form of SOD1 has been available since 1982 for the bovine enzyme[83,84]. Nowadays several other structures are available[85,86]. The solution structure of monomeric ad dimeric Cu,Zn,SOD were solved in our laboratory[87,88] (**Fig.6**).

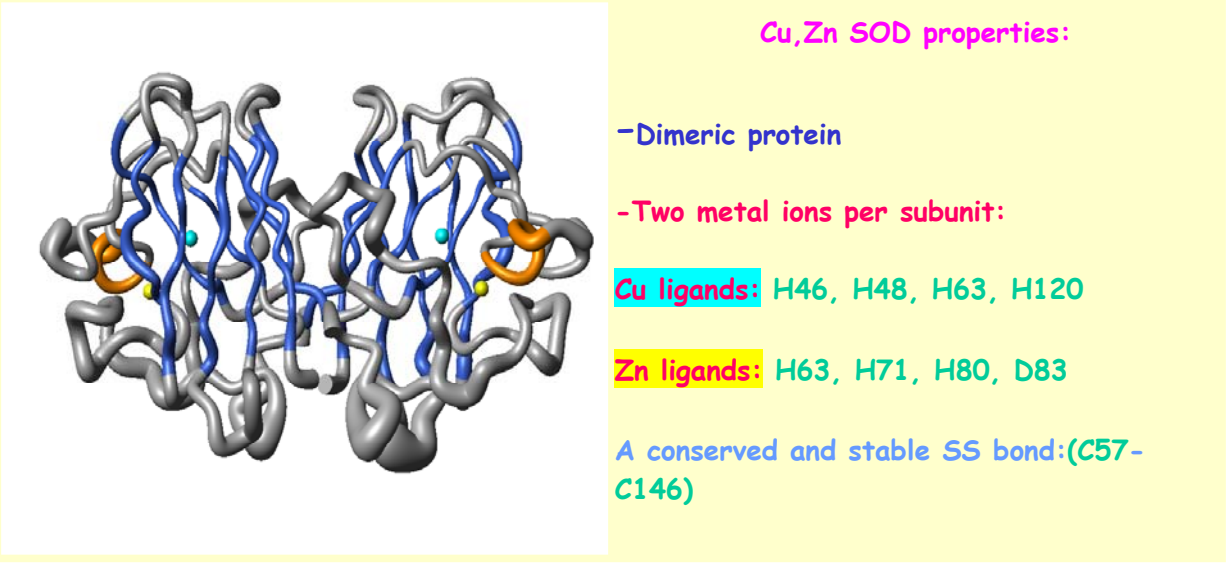



**Fig. 6:** *Tube representation of the family of 30 conformers of human reduced native SOD.* Elements of secondary structure are highlighted (blue: β structure, orange: α structure). The Cu ion is shown in cyan, the Zn ion in yellow. *(*Reprinted from Lucia Banci, Ivano Bertini, Fiorenza Cramaro, Rebecca Del Conte and Maria Silvia Viezzoli, Eur J Biochem. 2002 *)*

The human Cu,Zn SOD is constituted by of two identical subunits bearing the same amino acidic sequence of about 150 residues. Within each subunit the polypeptide chain is folded in a characteristic eight stranded β-barrel. The latter is formed by two four-stranded antiparallel β-sheets which are connected by turns and loops in a Greek key β-barrel motif **(Fig. 7)**. The eight β-strands forming the walls of the β-barrel display an overall right handed twist. The first sheet is formed by strands 1, 2, 3 and 6 and the second by strands 4, 5, 7 and 8. The loops connecting the secondary structure elements can be divided in two groups: the odd loops are located on the opposite side of the barrel with respect to regions involved in the subunit-subunit interface, while the even loops are in part located at the subunit-subunit interface. Inspection of the SOD1 structure reveals that loop IV, containing, Cys57, which forms the disulfide bridge with residue Cys146 in strand β8, can influence the conformation of the catalytically important residue, Arg143, through a hydrogen-bonding network[(89)]. Portions of this loop also contribute to the dimer interface[(90)], leading to the possibility that the disulfide bond influences the protein dimerization and thereby the SOD1 quaternary structure. This disulfide bond is conserved among all the structurally determined Cu,Zn SOD1. Crystallographic studies of SOD1 suggest a structural role for the disulfide in guiding the substrate into the active site[(91)]. Furthermore, the effect of the disulfide on dismutase activity and quaternary structure is

not clear, even though altered or aggregated forms of the protein are thought to be at the center of a familial form of amyotrophic lateral sclerosis (fALS)[92].

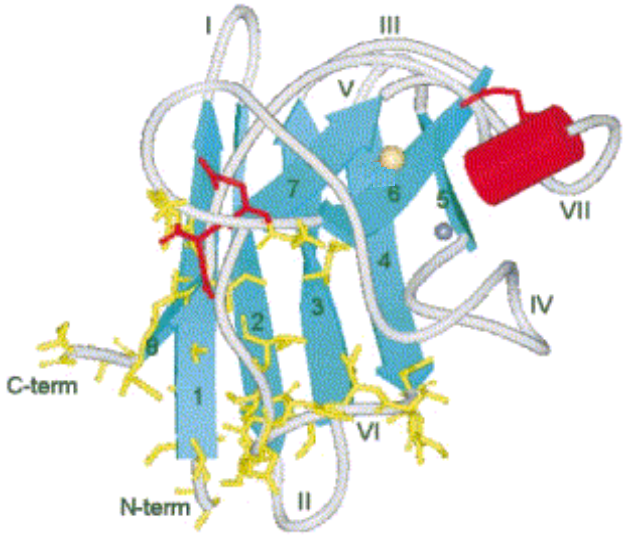



**Fig. 7**: *Schematic view of the Q133M2SOD structure displaying the secondary structure elements numbered by Getzoff et al. Proteins. 1989;5(4):322-36. The side chains of the residues homologous to those involved in the subunit- subunit interfaee in the WT dimer are shown as yellow sticks. The mutated residues E50, E51, Q133 are shown as red sticks.*
*Orange, Yellow and gray spheres of arbitrary radius represent copper and zinc ions.*
(Reprinted from Ferraroni et.al Journal of Molecular Biology1999)

The two monomers of SOD1 are held together by hydrophobic interactions. The contact surface between the monomers comprises the N terminus, strand β1, the C terminus, strand β8, the two loop regions:loop IV and loop VI. The extension of the contact region explains the high stability of the dimer to thermal and chemical denaturation. SOD1 is one of the most thermally stable enzymes known in mesophilic organisms. Dismutase activity declines at 80 °C with a corresponding melting temperature, Tm, above 90 °C[93]. The protein is stable in the presence of strong denaturants, and the activity is observed in 4% SDS or 10 M urea[94].

Mutations of residues in the interface region can lead to disruption of the quaternary structure, resulting in the formation of stable SOD1 monomers. Monomeric forms of the enzyme have been obtained by substituting hydrophobic residues at the subunit–subunit interface with hydrophilic residues. In particular, Phe50 and Gly51 have been substituted by two Glu residues, yielding a soluble single subunit. This mutant displays an activity of about 10% respect to the wild–type human Cu,ZnSOD[95]. Therefore, to partially restore the activity, Glu133 has been neutralized by substituting it with Gln[96]. This particular

monomeric CuZnSOD is referred as F50E/G51E/E133Q monomeric triple mutant or CuZnQ133M2SOD mutant.

Each SOD1 monomer contains one Cu and one Zn ion. The Zn ion is tetrahedrally coordinated to three histidyl imidazoles (His63, His71, His80) and to an aspartyl carboxylate group (Asp83). The Cu ion, in the oxidized enzyme, is coordinated by four histidines (His46, His48, His63, His120) and, weakly, by a water molecule in a distorted square planar pyramid. The water molecule occupies the apical position toward the opening of the cavity. The two metal ions are bridged, in the oxidized enzyme, by His63 that keeps them apart of about 6 Å. His63 binds Zn ion and Cu ion through the $N\delta_1$ and $N\varepsilon_2$ atoms, respectively. In the reduced Cu(I) form, a trigonal planar arrangement is obtained through the migration of the copper ion of about 1 Å which is accompanied by protonation of the histidine ligand that is no more linked to the Cu ion (**Fig. 8**).

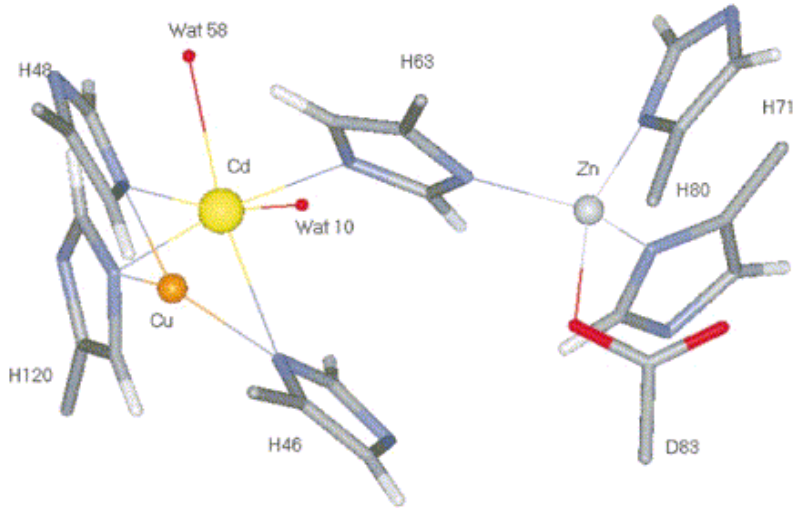



**Fig. 8:** *Overall view of the metal coordination in the active site in Cu,Zn SOD1*. The zinc ion is in magenta and the copper(I) ion is in orange and cadmiun (mimicking the Cu(II) site) is in yellow. *(*Reprinted from Ferraroni et.al Journal of Molecular Biology1999)

In SOD1, the zinc is completely buried in the protein and its role appears to be primarily structural[83]. Removal of the Zn ion indeed, results in a diminished thermal stability, i.e., the zinc-depleted protein denaturates at a lower temperature[97]. However, in conditions where the copper ion remains bound to the copper site, the removal of Zn ion does not significantly reduce SOD1 activity. The Cu ion instead has a solvent-exposed surface of 5.2 $Å^2$ and lies at the bottom of a narrow channel that is large enough to admit only water, small anions and similarly small ligands (**Fig. 9**). Its role in the enzyme appears

to be mainly catalytic. The redox properties of copper ion permit the disproportionation of superoxide anion to hydrogen peroxide and dioxygen.

The insertion of copper into Cu,Zn SOD requires a copper metallochaperone called CCS (copper chaperone for superoxide dismutase)[41,45,98]. Yeast mutants lacking CCS express a form of SOD protein that is essentially copper depleted[98] but contains a single atom of zinc per subunit[41,99]. Mice with disruptions in CCS exhibit marked reductions in SOD activity, stressing the conserved requirement for CCS in activation of eukaryotic Cu,Zn SOD[100].

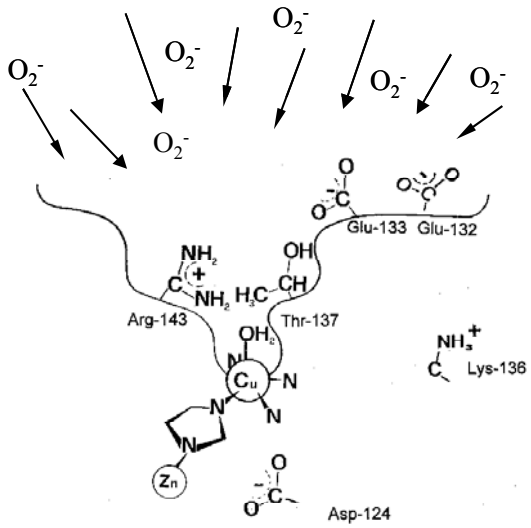



**Fig. 9:** *Schematic diagram of a cross section of the active site channel in Cu,Zn SOD1*

CCS is the largest copper metallochaperone identified to date. Whereas other copper chaperones as ATX1 and COX17 represent single domain proteins, CCS folds into three functionally distinct protein domains[100-102]. The N-terminal domain I of CCS bears striking homology to ATX1, including a MXCXXC copper-binding motif. Surprisingly, a CCS molecule lacking this domain can still insert copper into superoxide dismutase *in vivo*[101]. It has been therefore proposed that this ATX1-like domain is only needed to maximize CCS function under extreme copper-limiting conditions[5]. The central domain of CCS (domain II) is homologous to Cu,ZnSOD[100,102]. This domain physically interacts with SOD[103] and it was suggested to secure the enzyme during copper insertion. The C-terminal domain III of CCS is quite small (30 amino acids), yet is extremely crucial for activating SOD *in vivo*. This domain is highly conserved among CCS molecules from

diverse species and includes an invariant CXC motif that can bind copper[101]. Domain III is disordered in the crystal structure; however, it is predicted to lie in the vicinity of the N-terminal domain I. Models have been proposed in which domain III, perhaps in concert with the N-terminal domain I, directly inserts copper into the active site of SOD[104]. The mechanism of copper movement from an all-sulfur coordination environment (in CCS) to the all-nitrogen site in SOD remains an open issue. On the contrary, domain II of CCS is proposed to participate in target recognition rather than in direct transfer of copper.

.

## 1.7 SOD1 and amyotrophic lateral sclerosis disease

Amyotrophic lateral sclerosis (ALS) is a devastating, fatal neurodegenerative disease that specifically targets the motor neurons in the spinal cord, brain stem, and cortex[105]. It typically has an adult onset, starting first with weakness in arms or legs and proceeding relentlessly to total paralysis and death. With a lifetime risk of approximately 1 in 2000, it is the most common motor neuron disease[106]. Unfortunately, there is no known cure or effective treatment for ALS at present. Since the disease is selective for motor neurons, intellect is usually not affected. Patients generally die of respiratory failure within two to five years for first appearance of the diseases symptoms. Most instances (90%–95%) of ALS have no apparent genetic link and are termed sporadic ALS (sALS). In the remaining 5%–10% of cases there is a family history of the disease, and the disease is termed familial ALS (fALS).

To date, at least 105 different mutations in the sod1 gene have been linked to fALS[92]. The majority of these mutations cause single amino acid substitutions at one of at least 64 different locations but, some cause frameshifts, truncations, deletions, or insertions[92]. (http://www.alsod.org) Mutations are scattered throughout the sequence of the protein (**Fig. 10)**. The vast majority of the mutations are genetically dominant.

The autosomal dominant nature of SOD1-associated fALS suggests a toxic gain of function for mutant SOD1, and pioneering fALS transgenic mouse studies provided strong support for this hypothesis[107].

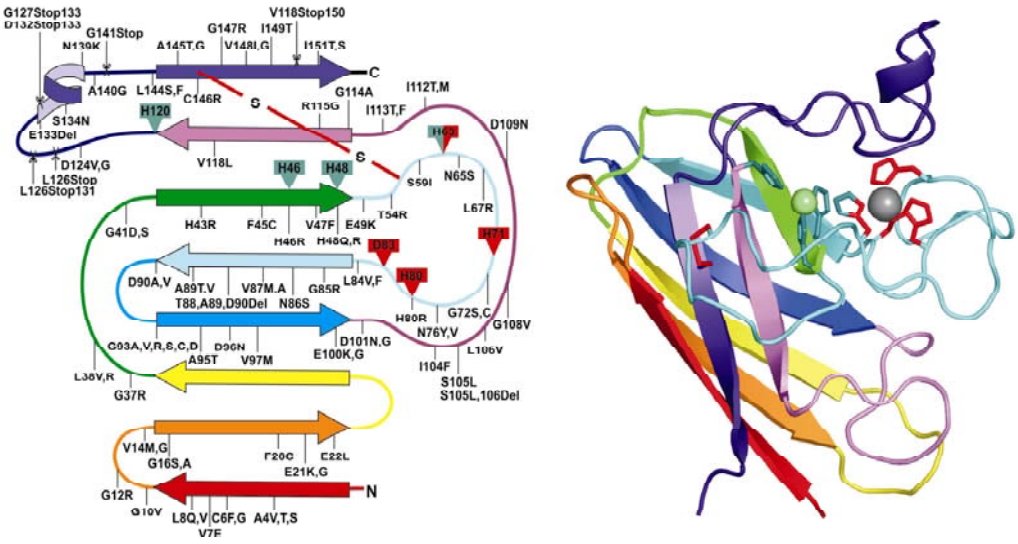



**Fig. 10***: Secondary structural representation of SOD1 showing the locations of fALS-associated mutations (left) and a monomer of SOD1 (right) colored to match the drawing on the left.* Copper ligands are shown in green and zinc ligands shown in red. Copper and zinc ions are shown as green and grey spheres, respectively, and the intrasubunit disulfide bond is shown in red. Point mutation, deletions, and insertions are indicated with a line, whereas mutations that cause C-terminal truncations are shown as scissor cuts at the point of the stop codon. *( Reprinted from Valentine JS et al (2005) Annual Review of Biochemistry)*

Two hypothesis can explain the toxicity of ALS mutant Cu,Zn SOD proteins: the oxidative damage hypothesis and the oligomerization hypothesis[92,108]. The first one maintains that fALS mutants acquire one or more toxic properties, catalysing reactions with oxidants such as peroxynitrite[98] and possibly hydrogen peroxide[109]. These reactive nitrogen and oxygen species cause toxicity spoiling proteins, nucleic acids and lipids. The oxidative damage hypothesis seems to require copper (or some other redox active metal ion) bound to the fALS mutant Cu,Zn SOD protein to promote the oxidative reaction. This former hypothesis has been recently abandoned as the second is supported by the finding of proteinaceous inclusions rich in mutant SOD1 in tissues from ALS patients, ALS-SOD1 transgenic mice, and in cell culture model systems[106]. These finding led many investigators to think that SOD1-associated fALS is a protein conformational disorder, similar to Alzheimer's disease, Parkinson's disease, Huntington's disease, transmittable spongiform encephalopathies, and other neurodegenerative diseases in which protein aggregates are found[30,110]. The visible inclusions in SOD1-linked fALS contain neurofilament proteins, ubiquitin, and a variety of other components in addition to SOD1, but it is not known if copper, zinc, or any other metal ions are present in the inclusions or are involved in their formation. It is also not known if the SOD1 polypeptide has been fragmented or otherwise covalently modified in the processes leading to aggregate

formation. The most likely explanation is that inclusions, and other visible protein aggregates, common to all of these the neudogenerative diseases, represent an end stage of a molecular cascade of several steps, and that earlier steps in the cascade may be more directly tied to pathogenesis than the inclusions themselves **(Fig. 11).**

The relatively large fibrils or insoluble inclusions observed in ALS may instead be the result of a protective mechanism that forms inclusions when the burden of misfolded or damaged proteins exceeds the capacity of the protein degradation machinery to eliminate them[111]. High-molecular-weight oligomerized species of SOD1, which may be more closely related to the toxic form, are found in the spinal cords of mice expressing mutant SOD1 well before disease onset[112]. This clearly indicates that the pathogenic SOD1 proteins must have some feature distinct from the wild-type protein that facilitates their self-association.

Currently many structural and biochemical data about metallated and apo WT-SOD1 and mutants are known. However all the efforts performed in order to correlate the biophysical properties with the survival data or with any other of the biological properties failed. It is not clear at this stage which factors represent the driving force for oligomerization of SOD1 and why ALS affect selectively only the motor neurons. Moreover, at the moment, there is no valid model able to explain a common mechanism for both familiar and sporadic ALS.

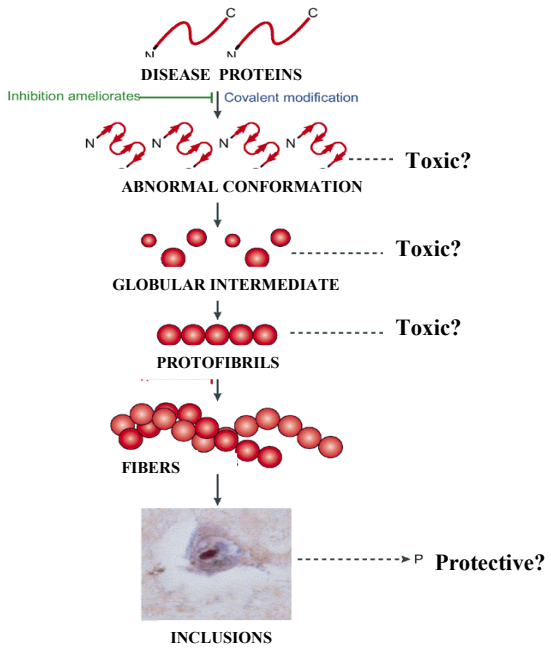



**Fig. 11:** *Hypothetical common pathway of protein aggregation in neudegenerative diseases.* An initiating event in aggregation may be covalent modification of the disease protein, for example by cleavage or phosphorylation, facilitating conversion of the protein to an abnormal conformation. Oligomeric (globular) intermediates may form, and then protofibrillar structures are assembled. Amyloid fibers can then form, possibly through association of protofibrillar intermediates, resulting in aggregates or inclusions visible in the light microscope. The intermediate species are hypothesized to be more toxic than either the precursor protein or the aggresomes and inclusions. (Reprinted from Ross CA et.al Nat Med. 2004 Jul;10 Suppl:S10-7.)

## 1.8 Aims and topics of the research

During the three years of PhD my research was focused on the investigation of the mechanisms that ensure the copper homeostasis in both eukaryotes and prokaryotes. My attention was devoted to proteins belonging to two different copper trafficking pathways in cells, i.e. copper delivery to cytosolic superoxide dismutase and to cytochrome *c* oxidase. Molecular mechanisms of copper insertion into these two enzymes, essential for cells life, still remain a puzzling issue.

Copper incorporation in Cox1 and Cox2 subunits of CcO is an essential step during the enzyme maturation for proper assembly of the complex. The recent identification of accessory proteins involved in this process, Cox17, Cox11 and HSco1/2 provides an exciting chance to elucidate the mechanism of assembly of this complex. Therefore, the aim of our research was the investigation and integrate comparison of all genetic, structural and biochemical data (including red-ox, structural and metal binding properties) in order to clarify the respective functional and molecular roles of these proteins within the assembly CcO mechanism. In the absence of direct experimental data, some clues on the properties of Cox11, Sco and related proteins, and on their involvement in copper delivery to CcO, have been initially gained from their genomic context, including operon structures and conserved domain fusions in prokaryotes.[114]. Therefore, we have performed a genome-wide search in prokaryotic organisms for sequences sharing similarity with human Cox17, Sco1/2, and Cox11, and we extended the analysis to genes close to the found proteins. Afterwards, wild-type human Sco1 and Sco2, and P174L-HSco1 pathogenic mutant were cloned and expressed and their solution and/or crystal structures determined for the first time. The copper acquisition mechanism of HSco1 and its pathogenic P174L mutant from the mithocondrial chaperone Cox17, was finally investigated in order to understand how copper transfer occurs and how the mutation can affect it.

The eukaryotic copper,zinc superoxide dismutases are remarkably stable dimeric proteins that maintain an intrasubunit disulfide bond in the reducing environment of the cytosol and are active under a variety of stringent denaturing conditions. The structural interplay of conserved disulfide bond and metal-site occupancy in human copper,zinc superoxide dismutase (Cu,Zn SOD1) is of increasing interest as these post-translational modifications are known to dramatically alter the catalytic chemistry, the subcellular localization, and the susceptibility of the protein to aggregation. The structural destabilization and/or an increase in mobility of superoxide dismutase mutants is suggested

to play a role in the pathology of amyotrophic lateral sclerosis disease (ALS) which is may be classified as an "conformational" disorder. These class of disordes generally, through an abnormal folding pathway, result in protein aggregation and high-molecular weight protein complexes formation. Therefore, at the beginning, the stability of the quaternary structure of WT-SOD1 in relation to its metallation and oxidation states was analyzed. The results of this first set of studies showed that reduction of the C57–C146 disulfide bond and the metallation are critical for the integrity of the dimer interface and shifts the equilibrium toward marginally stable monomers. It has been proposed that protein monomerization plays a role in formation of SOD1 misfolded intermediates, leading to protein aggregation in ALS. Therefore, the propensity for oligomerization of mature, i.e., disulfide-intact, WT human SOD1 under the relatively mild conditions likely to be encountered in the protein *"in vivo"* and the roles of metallation in hindering or promoting such oligomerization were then investigated.

# 1.9 Reference List


(1) Bertini, I.; Sigel, A.; Sigel, H. (2001) *Handbook on Metalloproteins*; New York,; pp 1-1800.

(2) Finney, L.A. & O'Halloran, T.V. (2003) *Science* **300**, 931-936.

(3) Changela, A.; Chen, K.; Xue, Y.; Holshen, J.; Outten, C. E.; O'Halloran, T. V.; Mondragon, A. (2003) *Science* **301**, 1383-1387.

(4) Bertinato, J.; L'Abbe, M. R.(2003) *J. Biol. Chem*. **278**, 35071-35078.

(5) O'Halloran, T.V. & Culotta, V.C. (2000) *J.Biol.Chem.* **275**, 25057-25060.

(6) Rae, T.; Schmidt, P.J.; Pufahl, R.A.; Culotta, V.C. & O'Halloran, T.V. (1999) *Science* **284**, 805-808.

(7) Crichton, R.R. & Pierre, J.L. (2001) *Biometals* **14**, 99-112.

(8) Pena, M.M.; Lee, J.; Thiele, D.J. (1999) *J. Nutr.* **129,** 1251–1260.

(9) Harris, E.D.; (2000) *Annu. Rev. Nutr.* **20**, 291–310.

(10) Uauy, R.; Olivares. M.; Gonzalez, M. (1998) *Am. J. Clin. Nutr.* **67**, 952S–959S.

(11) Rotillio, G.; Rossi, L.; De Martino, A.; Da Costa Ferreira,A.M.; Ciriolo, M.R. *J.Braz.Chem.Sco.* **6**,221-227.

(12) Puig, S & Thiele,D.J. (2002) *Curr. Opin. Chem. Biol*. **6**, 171–180.

(13) Brewer, G. J. (2003). *Curr. Opin. Chem. Biol.* **7**,207–212.

(14) Menkes, J.H.; Alter, M.; Stegleder, G.; Weakley, D.R.; Sung, J.H. (1962) *Pediatrics* **29**,764–779.

(15) Strausak, D.; Mercer, J.F.; Hermann, H.D.; Stremmel, W; Multhaup, G. (2001) *Brain Res.Bull.* **55**, 175-185.

(16) Waggoner, D.J.; Bartnikas, T.B.; Gitlin, J.D. (1999) *Neurobiol.Dis.* **6**, 221-230.

(17) Dobson CM. (2004) *Semin Cell Dev Biol.* **15** (1), 3-16.

(18) Rotilio, G.; Ciriolo, M. R.; Carri, M. T.; Rossi, L. (2002) *Handbook of Copper Pharmacology and Toxicology*, Humana Press, Totowa, NJ. 277–296.

(19) Stockel, J.; Safar, J.; Wallace, A.C.; Cohen, F.E.; Prusiner, S.B. (1998). *Biochemistry* **3,** 7185–7193.

(20) Barnham, K.J.; McKinstry, W.J.; Multhaup, G.; Galatis, D.; Morton, C.J.; Curtain, C.C.; Williamson, N.A.; White, A.R.; Hinds, M.G.; Norton, R.S.; Beyreuther, K.; Masters, C.L.; Parker, M.W.; Cappai, R. (2003) *J.Biol.Chem.* **278**, 17401-7.

(21) Brown, D.R. (2001) *Trends Neurosci*. **24**, 85–90.

(22) White, A.R.; Reyes, R.; Mercer, J.F.; Camakaris, J.; Zheng, H.; Bush, A.I.; Multhaup, G.; Beyreuther, K.; Masters, C.L.; Cappai, R. (1999). *Brain Res*. **842**, 439–444.

(23) Maynard, C.J.; Cappai, R.; Volitakis, I.; Cherny, R.A.; White, A.R.; Beyreuther, K.; Masters, C.L.; Bush, A.I.; Li, Q.X. (2002) *J. Biol. Chem.* **277**, 44670–44676.

(24) Huang, X.; Cuajungco, M.P.; Atwood, C.S.; Hartshorn, M.A.; Tyndall, J.D.A.; Hanson, G.R.; Stokes, K.C.; Leopold, M.; Multhaup, G.; Goldstein, L.E.; Scarpa, R.C.; Saunders, A.J.; Lim, J.; Moir, R.D.; Glabe, C.; Bowden, E.F.; Masters, C.L.; Fairlie, D.P.; Tanzi, R.E.; Bush, A.I. (1999). *J. Biol. Chem.* **274**, 37111–37116.

(25) Hesse, L.;Beher, D.; Masters, C. L.; Multhaup, G. (1994) *FEBS Lett*. **349**,109–116.

(26) Lovell, M.A.; Robertson, J.D.; Teesdale, W.J.; Campbell, J.L.; Markesbery, W.R. (1998) *J. Neurol. Sci.* **158**, 47–52.

(27) Cherny, R.A.; Atwood, C.S.; Xilinas, M.E.; Gray, D.N.; Jones,W.D.; McLean, C.A.; Barnham, K.J.; Volitakis, I.; Fraser, F.W.; Kim, Y-S.; Huang, X.; Goldstein, L.E.; Moir, R.D.; Lim, J.T.; Beyreuther, K.; Zheng, H.; Tanzi, R.E.; Masters, C.L., Bush,A.I. (2001) *Neuron* **30**,665–676.

(28) Uversky, V.N.; Li, J.; Fink, A.L. (2001) *J. Biol. Chem*. **276**,44284–44296.

(29) Rosen, D. R. (1993) *Nature* **362***,* 59–62.

(30) Valentine, J.S.; Hart, P.J. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 3617–3622.

(31) Rakhit, R.; Cunningham, P.; Furtos-Matei, A.; Dahan, S.; Qi, X.; Crow, J.P.; Cashman, N.R.; Kondejewski, L.H.; Chakrabartty,A (2002) *J. Biol. Chem.* **277**,47551–47556.
(32) Hassett, R.; Dix, D.R. ; Eide, D.J.;Kosman, D.J. (2000) *Biochem J.* **15**,351 Pt 2:477-84.
(33) Zhou, B.; Gitschier, J. (1997) *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7481–7486.
(34) Hasset, R.; Kosman, D.J. (1995) *J Biol Chem.* **270**:128-34.
(35) Guo, Y.; Smith, K.; Lee, J.;. Thiele, D.J.; Petris, M.J. (2004) *J. Biol. Chem.* **279**, 17428–17433.
(36) Lee, J.; Pena, M.M.; Nose, Y.;. Thiele, D.J (2002*) J. Biol. Chem.* **277**, 4380–4387.
(37) Hamza, I.; Gitlin, J.D. (2002) *J.Bioenerg.Biomemb.* **34**, 381-388.
(38) Amaravadi, R.; Glerum, D.M.; Tzagoloff, A. (1997) *Hum.Genet.* **99**, 329-333.
(39) Beers, J.; Glerum, D.M.; Tzagoloff, A. (2002) *J.Biol.Chem.* **277**, 22185-22190
(40) Punter, F.A.; Glerum, D.M. (2003) *J.Biol.Chem.* **278**, 30875-30880.
(41) Culotta, V.C.; Klomp, L.W.; Strain, J.; Casareno, R.L.; Krems, B.; Gitlin, J.D. (1997) *J.Biol.Chem.* **272**, 23469-23472.
(42) Huffman, D.L.; O'Halloran, T.V. (2001) *Annu Rev Biochem* **70**, 677-701.
(43) Lin, S.J.; Pufahl, R.; Dancis, A.; O'Halloran, T.V.; Culotta, V.C. (1997) *J.Biol.Chem.* **272**, 9215-9220.
(44) Pufahl, R.; Singer, C.P.; Peariso, K.L.; Lin, S.-J.; Schmidt, P.J.; Fahrni, C.J.; Culotta, V.; Penner-Hahn, J.E.; O'Halloran, T.V. (1997) *Science* **278,** 853-856.
(45) Portnoy, M.E.; Schmidt, P.J.; Rogers, R.S.; Culotta, V.C.; (2001) *Mol.Genet.Genomics* **265**, 873-882.
(46) Wikstrom, M. (1984) *Nature* **308**, 558-560.
(47) Wikstrom, M. (1998) *Curr.Opin.Struct.Biol.* **8**, 480-488.
(48) Michel, H.; Behr, J.; Harrenga, A.; Kannt, A. (1998) *Annu.Rev.biophys.biomol.Struct.* **27**, 329-356.
(49) Capaldi, R. A. (1990) *Annu. Rev. Biochem.* **59,** 569–596.
(50) Poyton, R.O.; McEwen, J.E. (1996) *Annu. Rev. Biochem.* **65,** 563–607.
(51) Tzagoloff, A.; Dieckmann, C. L; (1990) *Microbiol. Rev*. **54**, 211–225.
(52) Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Hakashima, R.; Yaono, R.; Yoshikawa, S. (1995) *Science* **269**, 1069–1074.
(53) Beers, J.; Glerum D.M.; Tzagoloff, A .(1997) *J. Biol. Chem.* **272**,33191–33196.
(54) Glerum, D.M.; Shtanko, A.; Tzagoloff, A. (1996) J. *Biol. Chem.* **271**, 14504–14509.
(55) Heaton, D.N.; George, G.N.; Garrison, G.; Winge, D.R. (2001) *Biochem* **40**, 743–751.
(56) Maxfield, A.B.. Heaton, D.N.; Winge, D.R. (2004) *J. Biol. Chem.* **279**, 5072–5080.
(57) Horng, O.C.; Cobine, P.A.; Maxfield, A.B.; Carr H.S.; Winge, D.R. (2004) *J. Biol. Chem.* **279,** 35334–35340.
(58) Rentzsch, A.; Krummeck-Weiss, G.; Hofer, A.; Bartuschka, A.; Ostermann, K.; Rodel, G. (1999) *Curr Genet* **35**, 103-108.
(59) Nittis, T.; George, G.N.; Winge, D.R. (2001) *J.Biol.Chem.* **276**, 42520-42526.
(60) Balatri, E.; Banci, L.; Bestini, I.; Cantini, F.; Ciofi-Baffoni. S. (2003) *Structure.* **11**,1431-43.
(61) Dickinson, E.K.; Adams, D.L.; Schon, E.A.; Glerum, D.M. (2000) *J.Biol.Chem.* **275**, 26780-26785.
(62) Lode, A.; Paret C.; Rödel, G. (2002) *Yeast* **19**,909–922.
(63) Glerum, D.M.; Shtanko, A.; Tzagoloff, A. (1996) J. *Biol. Chem.* **271**, 20531–20535.
(64) Lode, A.; Kuschel, M.; Paret, C.; Rodel, G. (2000) *FEBS Lett* **485**, 19-24.
(65) Hiser, L.; Di Valentin, M.; Hamer, A.G.; Hosler, J.P. (2000) *J.Biol.Chem.* **275**, 619-623.
(66) Tzagoloff, A.; Capitanio, N.; Nobrega, M.P.; Gatti, D. (1990) *EMBO J.* **9**, 2759-2764.
(67) Carr, H.S.; George, G.N.; Winge, D.R.; (2002) *J Biol Chem* **277**, 31273-31242.

[68] Papadopoulou, L.C.; Sue, C.M.; Davidson, M.M.; Tanji, K.; Nishino, I.; Sadlock, J.E.; Krishna, S.; Walker, W.; Selby, J.; Glerum, D. M.; Van Coster, R.; Lyon, G.; Scalais, E.; Lebel, R.; Kaplan, P.; Shanske, S.; De Vivo, D.C.; Bonilla, E.; Hirano, M.; DiMauro, S.; Schon, E.A (1999) *Nat. Genet.* **23**, 333-337.
[69] Jaksch, M.; Ogilvie, I.; Yao, J.B.; Kortenhaus, G.; Bresser, H.G.; Gerbitz, K.D.; Shoubridge, E.A. (2000) *Hum. Mol. Genet.* **9,** 795-801.
[70] Jaksch, M.; Paret, C.; Stucka, R.; Horn, N.; Muller-Hocker, J.; Horvath, R.; Trepesch, N.; Stecker, G.; Freisinger, P.; Thirion, C.; Muller, J.; Lunkwitz, R.; Rödel, G.; Shoubridge, E. A.; Lochmuller, H. (2001) *Hum. Mol. Genet.* **10,** 3025-3035.
[71] Valnot, I.; Osmond, S.; Gigarel, N.; Mehaye, B.; Amiel, J.; Cormier-Daire, V.; Munnich, A.; Bonnefont, J.P.; Rustin, P.; and Rötig, A. (2000) *Am. J. Hum. Genet.* **67,** 1104-1109.
[72] Arner, E. S.; Holmgren, A.; (2000) *Eur J Biochem* **267**, 6102-6109.
[73] Williams, J. C., Sue, C., Banting, G. S., Yang, H., Glerum, D. M., Hendrickson, W. A. & Schon E. A. (2005) *J. Biol. Chem.* **280**, 15202-15211.
[74] Horng, Y. C., Cobine, P. A., Maxfield, A. B., Carr, H. S. & Winge, D. R. (2004) *J. Biol. Chem.* **279**, 35334-35340.
[75] Davies, K.J. (1995) *Biochem.Soc.Symp.* **61**, 1-31.
[76] Halliwell, B. (1994) *Nutr.Rev.* **52**, 253-265.
[77] Richter, C.; Gogvadze, V.; Laffranchi, R.; Schlapbach, R.; Schweizer, M.; Suter, M.; Walter, P.; Yaffee, M. (1995) *Biochim.Biophys.Acta* **1271**, 67-74.
[78] Fee, J.A.; McCord, J.M.; Fridovich, I.. (1977) eds. Michelson, 173-192.
[79] Weisiger, R. A., and Fridovich, I (1973) *J. Biol. Chem.* **248**, 4793–4796.
[80] Crapo, JD.; Oury, T.; Rabouille, C.; Slot, JW.; Chang, LY (1992) *Proc. Natl. Acad. Sci. USA* **89**, 10405-10409.
[81] Markowitz, H.; Cartwright G.E.; Wintrobe, M.M.(1959*) J. Biol. Chem.* **234**, 40–45.
[82] McCord, J. M.; Fridovich, I. (1969) *J. Biol. Chem.* **244**, 6049-6055.
[83] Tainer, J.A.; Getzoff, E.D.; Richardson, J.S.; Richardson, D.C. (1983) *Nature* **306**, 284-287.
[84] Tainer, J.A.; Getzoff, E.D.; Beem, K.M.; Richardson, J.S.; Richardson, D.C. (1982) *J.Mol.Biol.* **160**, 181-217.
[85] Parge, H.E.; Hallewell, R.A.; Tainer, J.A. (1992) *Proc.Natl.Acad.Sci.USA* **89**, 6109-6114.
[86] Ferraroni, M.; Rypniewski, W.; Wilson, K.S.; Viezzoli, M.S.; Banci, L.; Bertini, I.; Mangani, S. (1999) *J.Mol.Biol.* **288**, 413-426.
[87] Banci, L.; Benedetto, M.; Bertini, I.; Del Conte, R.; Piccioli, M.; Viezzoli, M.S. (1998) *Biochemistry* **37**, 11780-11791.
[88] Banci, L.; Bertini, I.; Cramaro, F.; Del Conte, R.; Viezzoli, M.S. (2002) *Eur J.Biochem.* **269**, 1905-1915.
[89] Fisher, C.L.; Cabelli, D.E.; Tainer, J.A.; Hallewell, R.A.; and Getzoff, E.D. (1994) *Proteins* **19**, 24–34).
[90] Bertini, I.; Mangani, S.; and Viezzoli, M. S. (1998) *Advanced Inorganic Chemistry* (Sykes, A. G., ed) 127–250, Academic Press, San Diego
[91] Fisher, C.L.; Cabelli, D.E.; Tainer, J.A.; Hallewell, R.A.; Getzoff, E.D. (1994) *Proteins* **19**, 24–34.
[92] Cleveland, D.W.; Rothstein, J.D. (2001) *Nat.Rev.Neurosci.* **2**, 806-819.
[93] Roe, J.A.; Butler, A.; Scholler, D.M.; Valentine, J.S.; Marky, L.; Breslauer, K.J. (1988) *Biochemistry* 27, 950–958.
[94] Forman, H.J.; and Fridovich, I. (1973) *J. Biol. Chem.* **248**, 2645–2649.
[95] Bertini, I.; Piccioli, M.; Viezzoli, M.S.; Chiu, C.Y.; Mullenbach, G.T. (1994) *Eur.J.Biophys.* **23**, 167-176.

[96] Banci, L.; Bertini, I.; Chiu, C.Y.; Mullenbach, G.T. Viezzoli, M.S. (1995) *Eur.J.Biochem.* **234**, 855-860.
[97] Estévez, A.G.; Crow, J.P.; Sampson, J.B.; Reiter, C.; Zhuang, Y.; Richardson, G.J.; Tarpey, M.M.; Barbeito, L.; Beckman, J.S. (1999) *Science* **286**, 5449.
[98] Gamonet, F.; Lauquin, G.J.M. (1998) *Eur.J.Biochem.* **251**, 716-723
[99] Wong, P.C.; Waggoner, D.; Subramaniam, J.R., Tessarollo, L.; Bartnikas, T.B.; Culotta, V.C.; Price, D.L.; Rothstein, J.; Gitlin, J.D. (2000) *Proc.Natl.Acad.Sci.USA* **97**, 2886-2891.
[100] Lamb, A.L.; Wernimont, A.K., Pufahl, R.A.; Culotta, V.C.; O'Halloran, T.V.; Rosenzweig, A.C. (1999) *Nature Struct.Biol.* **6**, 724-729.
[101] Schmidt, P.J.; Rae, T.D.; Pufahl, R.A.; Hamma, T.; Strain, J.; O'Halloran, T.V.; Culotta, V.C. (1999) *J.Biol.Chem.* **274**, 23719-23725.
[102] Hall, L.T.; Sanchez, R.J.; Holloway, S.P.; Zhu, H.; Stine, J.E.; Lyons, T.J.; Demeler, B.; Schirf, V.; Hansen, J.C.; Nersissian, A.M.; Valentine, J.S.; Hart, P.J. (2000) *Biochemistry* **39**, 3611-3623.
[103] Casareno, R.L.; Waggoner, D.J.; Gitlin, J.D. (1998) *J.Biol.Chem.* **273**, 23625-23628.
[104] Zhu, H.; Shipp, E.; Sanchez, R.J.; Liba, A.; Stine, J.E.; Hart, P.J.; Gralla, E.B.; Nersissian, A.M.; Valentine, J.S. (2000) *Biochemistry* **39**, 5413-5421.
[105] Rowland, L.P.; Shneider, N.A. (2001) *N. Engl. J. Med.* **344**, 1688–700.
[106] Bruijn, L.I.; Miller, T.M.; Cleveland D.W (2004) *Annu. Rev. Neurosci.* **27**, 723–49.
[107] Gurney, M.E.; Pu, H.; Chiu, A.Y.; Dal Canto, M.C.; Polchow, C.Y.; (1994) *Science* **264**, 1772–75.
[108] Valentine, J.S. (2002) *Free Radic.Biol.Med.* **33**, 1314-1320.
[109] Yim, M.B.; Kang, J.H.; Yim, H.S.; Kwak, H.S.; Chock, P.B., Stadtman, E.R. (1996) *Proc.Natl.Acad.Sci.USA* **93**, 5709-5714.
[110] Ross, C.A.; Poirier, M.A. (2004) *Nat Med.* **10**, Suppl:S10-7.
[111] Sherman, M.Y.; Goldberg, A.L. (2001) *Neuron* **29**, 15–32.
[112] Johnston, J.A.; Dalton, M.J.; Gurney, M.E.; Kopito, R.R.; (2000) *Proc. Natl. Acad. Sci. USA* **97**, 12571–76.
[113] Huynen, M.; Snel, B.; Lathe, W.; Bork, P. (2000) *Curr. Opin. Struct. Biol.* **10**, 366-370.
[114] Galperin, M.Y.; Koonin, E.V. (2000) *Nat. Biotechnol,* **18**, 609-613.

# 2

# METHODOLOGICAL ASPECT

## 2.1 Structural and Functional Genomics approach

The genome sequencing projects developed in the last few years have made available the primary sequence of an enormous number of new genes. The draft sequence of the human genome[1,2] is now available together with the complete genomes of many other organisms (Web site: www.ncbi.nlm.nih.gov). This wealth of genomic data can be fully exploited only through elucidation of the three-dimensional structures of the corresponding proteins. The final objective of the "classic" Stuctural Genomics is the structure-based drug design for the care of diseases.

The European perception of Structural Genomics is extended to the study of the protein function. This can be achieved by the identification, production and characterization of partner proteins and by protein-protein interaction studies. The various Structural Genomics initiatives commencing worldwide differ in the choice of protein targets in an effort to provide a complete coverage of protein structures and relative associated function.

So, there are actually two different possible philosophies that distinguish the various research consortia.

-**The "classic" Structural Genomics approach:** It is High-Throughput approach whose goal is to provide a complete repertoire of protein folding by determining a representative structure. In this case the effort for the target selection is limited.

Entire genomes or very large protein families (like for example: soluble human proteins under 100 residues) are selected, cloned, expressed in parallel using automatic systems. The great advantage of this approach is that a huge number of genes can be screened for protein expression and solubility. However, a common cloning and expression protocol has to be used for all the genes limiting in this way, the rate of success (usually 20-30%) for each target.

- **The Functional Genomics approach:** this approach focuses the effort on a limited number of proteins that belong specific families linked into diseases or involved in some specific metabolic pathways. Large bibliographic and bioinformatics research is at the basis of the target selection. A limited number of targets are cloned but, in order to maximize the rate of success, many simultaneous cloning and expression strategies have to be performed.

For both strategies a multidisciplinary approach is necessary; a large number of steps, regarding different scientific fields like bioinformatics, biochemistry and molecular biology have to cooperate in order to proceed to the target selection, protein expression, structural and functional characterization. The general strategy of the research consists of the following phases: target selection performing genome browsing in bacteria and eukaryotes; domain definition; gene cloning and protein expression; structural determination by NMR and/or X-ray; modeling; protein-protein interaction studies.

## 2.2 Genome browsing

One of the challenges faced by the molecular biology community today is to make sense of the wealth of data that have been produced by the genome sequencing projects. In the past decade, bioinformatics has become an integral part of research and development in biomedical sciences. Bioinformatics has nowadays an essential role both in deciphering genomic, transcriptomic and proteomic data, generated by high-throughput experimental technologies, and in organizing information gathered from traditional biology. It is an interdisciplinary research area at the interface between the biological and computational science. The ultimate goal of bioinformatics is to uncover the biological information hidden in the wide collection of data and to obtain a better insight into the fundamental biology of organisms. A number of data banks with databases and programs have been created with the aim of providing the scientific community with tools for searching gene banks, for the analysis of protein sequences and for the prediction of a variety of protein properties.

In this research we have focused our attention on finding new metalloproteins. In the initial stage of the search a number of metalloprotein sequences were selected on the basis of previously genetic and biochemical studies. Several different bioinformatic approaches are possible in relation to which kind of target we are interested in searching. One possible strategy is performing a BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) search in order to explore all of the available sequence databases (NCBI at http://www.ncbi.nlm.nih.gov/) searching for an unknown protein that contain a conserved consensus motifs for the metal binding, For example, HSco1 and HSco2 show the same CXXXC hypothetical copper binding motif, conserved in the $Cu_A$ centre of CcO, while Atx1 presents the same MXCXXC motif found also in CCC2. The consensus sequence gives a unique perspective

to structural genomics of metalloproteins based on genome browsing, as an intriguing and broad variety of functions and metal binding properties are found in proteins sharing the same fold and the same consensus sequence. Performing a ClustalW (http://www.ebi.ac.uk/clustalw/) alignment and a Seqspace (http://industry.ebi.ac.uk/SeqSpace/) run of the obtained sequences from the genome browsing, we selected an identified macro-group of genes characterized by common features. In some cases, it is possible to observe the occurrence of distinctive domains and the repetition of metal consensus sequences within the same gene. The collected sequences can be further divided in subfamily on the basis of their domain organization. Phylogenetic comparative analysis, and a gene-context analysis (STRING at http://string.embl.de/), for the genes having prokaryotic homologues, could be useful to divide the genes of the macro-families into classes.

When a protein of particular interest is identified, it becomes the subject of further bioinformatics investigations in order to predict features like stability, solubility, hydrophobicity, secondary and tertiary structures.

## 2.3 Domain definition

From each class of proteins, several domains are selected for expression. Multiple sequence alignment provides a number of homologous genes from different organisms; the gene cloning could be carried out on various sequences. Several preliminary analysis are thus necessary to select the protein construct with the highest probability of giving rise to a soluble and folded protein.

In our laboratory the domain definition is performed taking in consideration the following aspects:

- The presence of transmembrane regions. It is very important to know if we are dealing with a completely soluble protein, with an integral transmembrane protein or with a protein with a short peptide inserted into the phospholipidic bilayer, that is for the most part soluble. While in the first case the full length chain represent probably the best construct, in the latter one it is possible to attempt the expression of a recombinant protein without the transmembrane tail. In this case, in order to maximize the successful rate several possible constructs of the soluble part(s) are designed.

- The presence of N-terminal signal peptide (Predicted by SignalP at www.cbs.dtu.dk/services/SignalP/) sequence that enable the expression of the protein in the periplasmic or mithocondrial area.
- The N-terminal part with designed in respect to the N-end rules[3]. This rule suggest a relation between the metabolic stability of a protein and the identity of its N-terminal residue.
- The presence in the sequence of intrinsically unstructured/disordered regions (Predicted by IUPred at http://iupred.enzim.hu/)
- The presence in the DNA coding sequences of codons that are "rare" to the host selected for the expression of the recombinant protein. (for *E.coli* expression system are predicted at http://nihserver.mbi.ucla.edu/RACC/)
- The domain is usually cut in correspondence to regions that show in ClustalW alignment (www.ebi.ac.uk/clustalw/), a low degree of similarity to the protein family they belong to, and where no stable secondary structural elements are localized . (Predicted at http://npsa-pbil.ibcp.fr/).

In the current work this approach was used to design the HSco1 domains. Hsco1 contains a N-terminal signal peptide, in the regions 1-31 and a transmembrane region from 94 to 110. ClustalW alignment with all other Sco1 homologues shows the presence, in HSco1, of a highly charged 14aa sequence localized between the TM region and the CXXXC conserved copper binding motif, that is not present in prokaryotic Sco1 homologues. Two different domains were designed, one with and the other without this charged region. As it can be read in chapter 3.2 both the two domain result in soluble expression but the short one is monomer while the long is dimer.

## 2.4 Gene cloning

In the Functional genomic approach, which currently represent our philosophy, the strategy for cloning the target protein has to be carefully designed since it will influence the behaviour of the protein, e.g. yield, solubility, folding, etc. Although protein expression is no longer considered a major limiting step and protein purification techniques have improved dramatically in the past decade, the problem of producing soluble proteins for purification has continued to be a major bottleneck in the field. Unfortunately many proteins belonging to interesting families, especially the human proteins, are extremely

difficult to be produce as soluble proteins in *E.coli* as they are often accumulated as inclusion bodies. Even if in the recent years, the increasing knowledge of additives can assist the "in vitro" refolding process, and numerous proteins have been refolded into their active forms, yet the successful rate of this strategy is very low. Considerable efforts are currently underway to make alternative hosts more accessible and affordable, and eukaryotic systems including mammalian, yeast and insect cell expression are becoming easier and less expensive to be used[4,5]. Cell-free protein synthesis has also great potential for overcoming some of the problems of soluble and membrane proteins expression, but it still remains a work in progress for the time being[6]. However especially for NMR purpose, which requires the production of high yield of labelled $^{15}N$ and $^{13}C$ samples, the *E.coli* expression system is nowadays, the most widely used.

Many efforts are currently focused on the optimization of the *E.coli* expression system. Factors such as reduced temperature[7], changes in the *E.coli* expression strain[8], different promoters or induction conditions[9] and co-expression of molecular chaperones and folding modulators[10] have all been examined and, in some specific cases, they lead to enhancements of soluble protein production. Many years ago it was discovered that some affinity tags are able to enhance the solubility of some of the partner proteins to which they are attached[11]; even if the number of fusion partners is increasing progressively during the years none of these tags work universally with each partner protein.

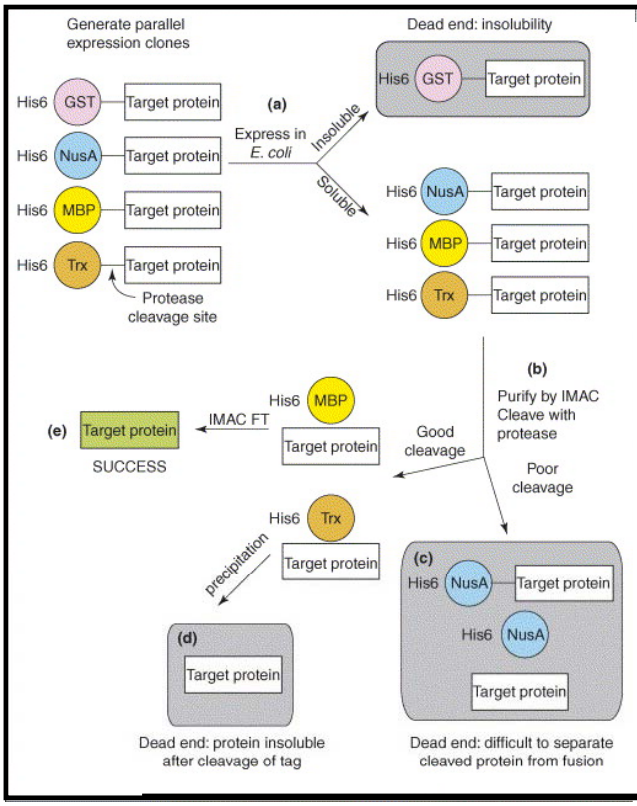



**Fig. 12:** *Schematic representation of the pathway from protein expression to purification using solubility tags.* Four arbitrary tagged versions of the protein of interest are generated in E. coli.

**(a)** After expression, some fusions will remain in the insoluble fraction and be lost from the pathway.
**(b)** Soluble fusions are purified by IMAC using the attached His6 tag. A protease is then used to cleave the fusion tag from the partner protein.
**(c)** Some fusions will not cleave efficiently, and will leave behind a mixture of cleaved and uncleaved proteins that cannot be easily separated.
**(d)** Other fusions will cleave efficiently, but when separated from the solubility tag the partner protein will become insoluble and precipitate.
**(e)** However, a well-behaved fusion will remain in solution and can be purified by a second IMAC step to remove the His6-tagged solubility tag and protease, leaving only the target protein in the flow through (IMAC FT).

IMAC=immobilized metal affinity chromatography. (Reprinted from Esposito D et.al (2006) Current Opinion in Biotechnology 17, I 353-358)

The best way to maximize the probability of obtaining a soluble and correctly folded target protein is to proceed with a parallel cloning and expression of it with a high number of fusion partner (**Fig.12**)

Anyway, the experience shows that in some cases the classic approach, that is to express the native protein without any tag results the only successful one. While classical approaches do not require further sub cloning, fusion partners' impact on protein solubility leads to sub cloning the gene of interest in a library of expression vectors that becomes laborious when handling a large number of genes. Recently, Landy and co-workers[12] have described a cloning method (Gateway technology) that enables rapid cloning of one or more genes into virtually any expression vector using site-specific and conservative recombination (LR reaction) eliminating the requirement to work with restriction enzymes and ligase **(Fig.13).**

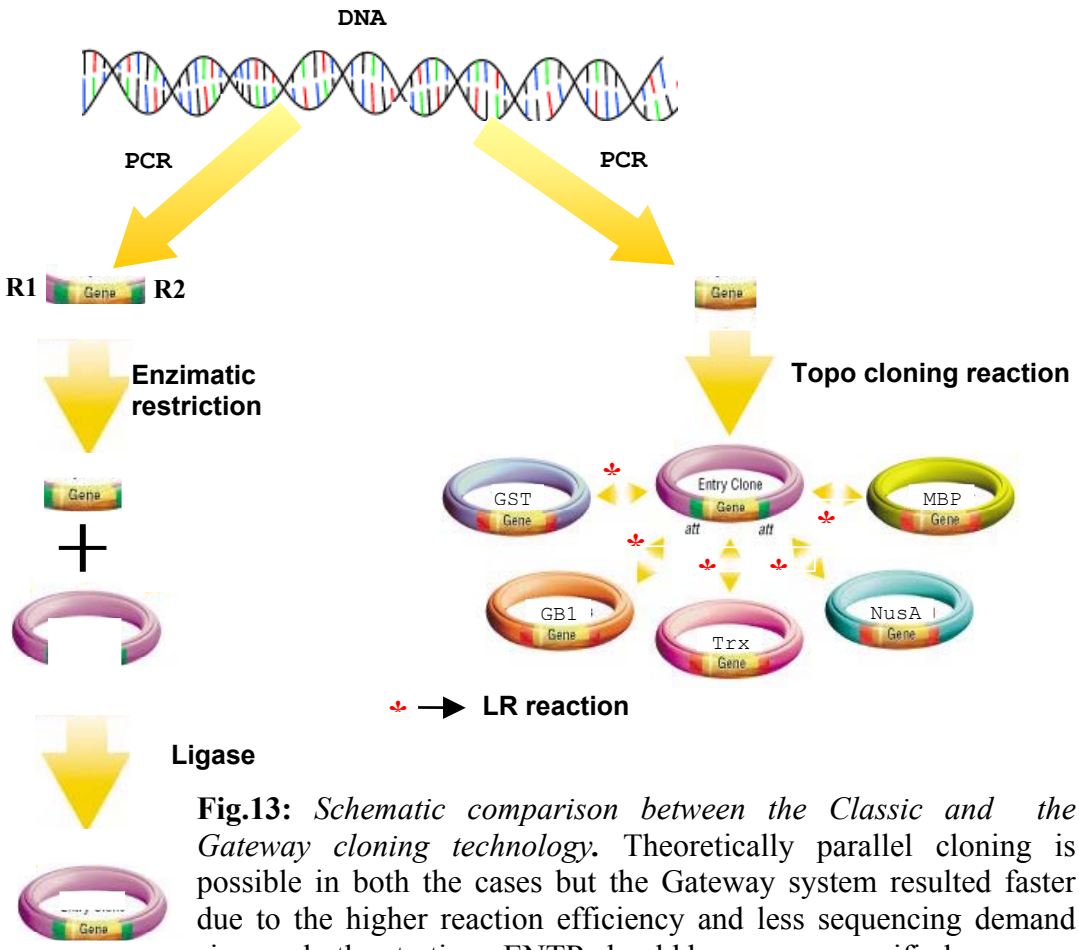



**Fig.13:** *Schematic comparison between the Classic and the Gateway cloning technology.* Theoretically parallel cloning is possible in both the cases but the Gateway system resulted faster due to the higher reaction efficiency and less sequencing demand since only the starting pENTR should be sequence verified.

On the basis of the former considerations, in our laboratory, we select at list two different domains for each new target. These are cloned with the classic methods to

express the protein in the native form, and with Gateway system (Invitrogen) in order to express it with five different N-Terminal fusion tag (GST, MBP,NusA,Trx,GB1).

Human Sco1 and human Cox17 were cloned according to the general procedure described above.

## 2.5 Protein expression

At first, expression and solubility screening, in small scale culture (1-10ml LB) is performed generally using at list 3 different *E.coli* strains (BL21Gold (DE3), Rosetta(DE3) and Origami(DE3 ) and inducing the protein expression at two different temperature (37°C and 25°C). The result of this first test is analysed and on the basis of these data it is decided whether or not proceeding to the scale-up and testing the expression in minimum media. A second screening is sometimes performed in order to refine expression conditions or, in the case in which all the test are negative, to redefine the strategy.

1) If no expression is observed we can take into consideration the following choice:
- Redefine the domain.
- Use an expression vector containing an inducible promoter different from T7(for example ARA or Cold inducible promoters)[9]
- Test other *E.coli* strains. In some cases, using promoters different from T7, as for instance LacIq , better results were obtained with *E.coli* strain lacking the DE3 episome.
- Consider the possibility to move to an eukaryotic expression system.

2) If all the proteins are expressed in the insoluble fraction as inclusion bodies the choice are:
- Change the expression condition (temperature, induction time, medium ecc)
- Redefine the domain.
- Proceed with an " in vitro" refolding screening
- Clone again the domain in a "classic vector" codifying for the same fusion partners but lacking the His-tag. In some case it was observed that the presence of an N-or C-terminal His-tag[13] and/or the sequences corresponding at the Gateway recombination cassette, has a noticeable negative affect on protein solubility.

- Consider the possibility to move to an eukaryotic expression system.

3) If from the expression screening we obtain positive results we select the more promising conditions and move to the scale-up, which is performed according to the following scheme:

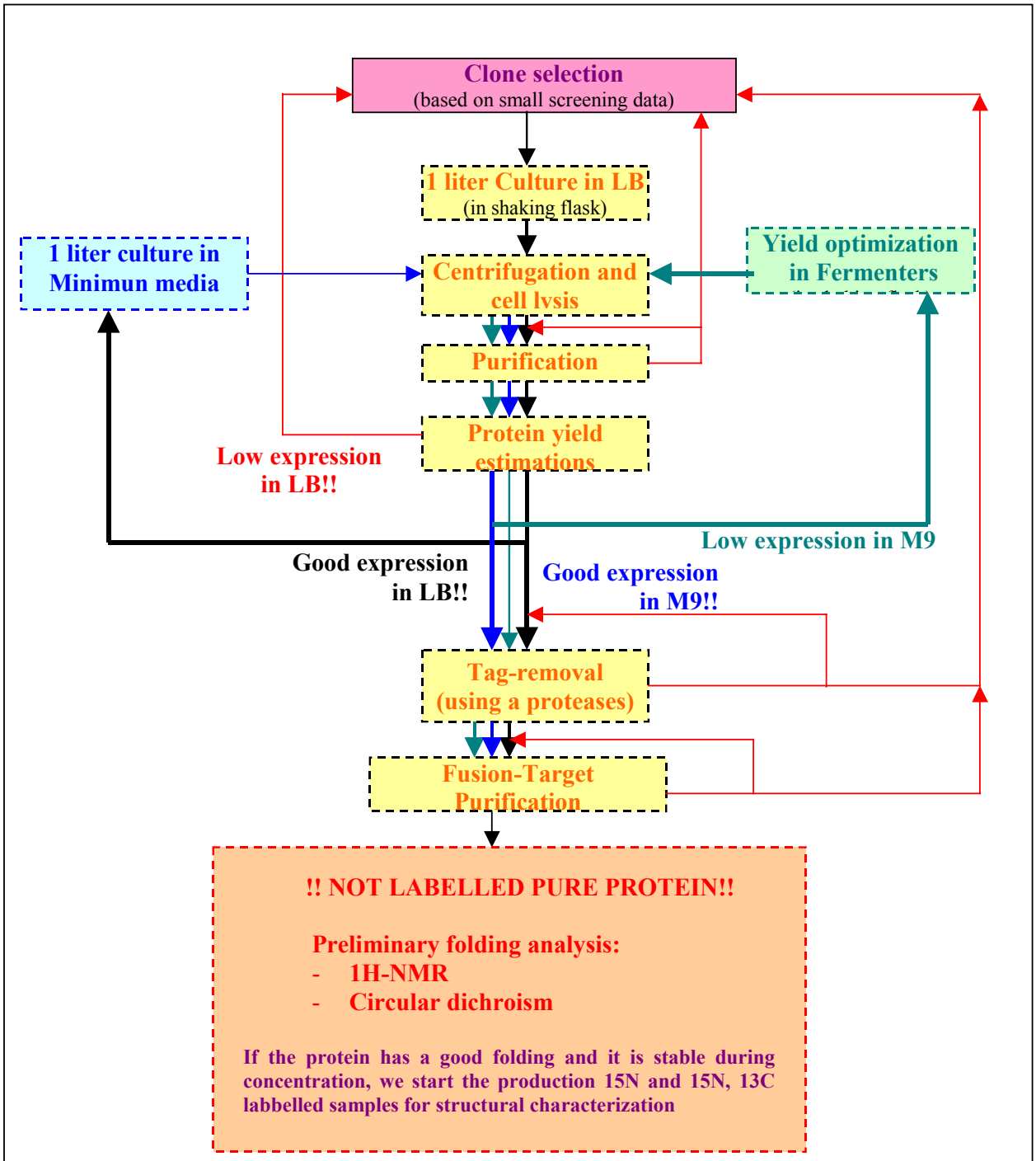



**Fig.14:** *Typical scale-up optimization process.* Expression, purification, tag-removal by proteases are optimized on the unlabelled samples from the LB culture. In parallel, simulation of culture in minimal media(M9) are performed to optimize the labelling procedure. Fermentors are used only for M9 when low yields are obtained from shaking-flask.

The set-up of a new expression protocol generally requires the optimization of different steps. In **Fig.14,** it is also shown, the frequent situation in which it is necessary to return to the cloning selection and restart from the beginning if the optimisation of some intermediate step fail.

For HSco1 proteins expression, competent *E.coli* BL21-Gold(DE3) cells were used. The $^{15}$N-labeled $^{15}$N $^{13}$C-labeled proteins were obtained by growing cells in minimal medium (M9) containing 3g/l $^{13}$C-Glucose and 1g/l $^{15}$N-$(NH_4)_2SO_4$ whereas LB medium was used for the non-labelled protein. For the latter form, the cells were grown in shaking flaks at 37°C until OD=0.7 and then induced with 0,7mM IPTG for 1h a 37°C and 16h at 30°C. A longer induction at 37°C results in expression of Hsco1 in the insoluble fraction. The same procedure repeated in minimal medium results in a dramatic loss of protein expression level. A good yield (>40mg/l of the GST-Hsco1 fused proteins) was recovered, optimizing the growth and expression condition in 1.8 liter Minifors fermenter (Infors HT).

For Cox17 protein expression, competent *E.coli* Origami(DE3)plyss cells were used. The $^{15}$N-labeled $^{15}$N $^{13}$C-labeled proteins were obtained by growing cells in M9 containing 3g/l $^{13}$C-Glucose and 1g/l $^{15}$N-$(NH_4)_2SO_4$ whereas LB medium was used for the non-labelled protein. For the latter form, the cells were grown in shaking flask at 37°C until OD=0.7; protein expression was induced with 0,7mM IPTG at 25°C over night. For Cox17 the shaking flask culture in minimal medium provide a really good yield (>70mg/l of the GST-Cox17 fused proteins), and the optimization with fermenters was not necessary.

## 2.6 Protein purification

The strategy of purification depends mainly upon the location of the expressed protein within the host, indeed, the protein can be transported in the periplasmic space or expressed like a soluble or insoluble (inclusion bodies) protein within the cytoplasm. In each cases, the isolation was performed in a different way. All the purifications involve several chromatographic steps performed adjusting the parameters according to the different physical chemical and biological characteristics of the proteins. Ion exchange and size exclusion chromatography are commonly used to purify proteins expressed in their native states. Some tags, like MBP or GST, are specific affinity tag, which not only facilitate soluble expression but also increase the efficiency of protein purification. In some cases, other solubility tags have been combined with a simple His-tag, allowing the fusion

partner to maintain its solubilizing functionality and also simultaneously as an affinity tag. Immobilized metal ion affinity chromatography (IMAC) is currently the most used affinity technique, it exploits the interaction between chelated transition metal ions (generally $Zn^{2+}$ or $Ni^{2+}$) and side-chains of specific amino acids (mainly histidine) on the protein (**Fig. 15**)

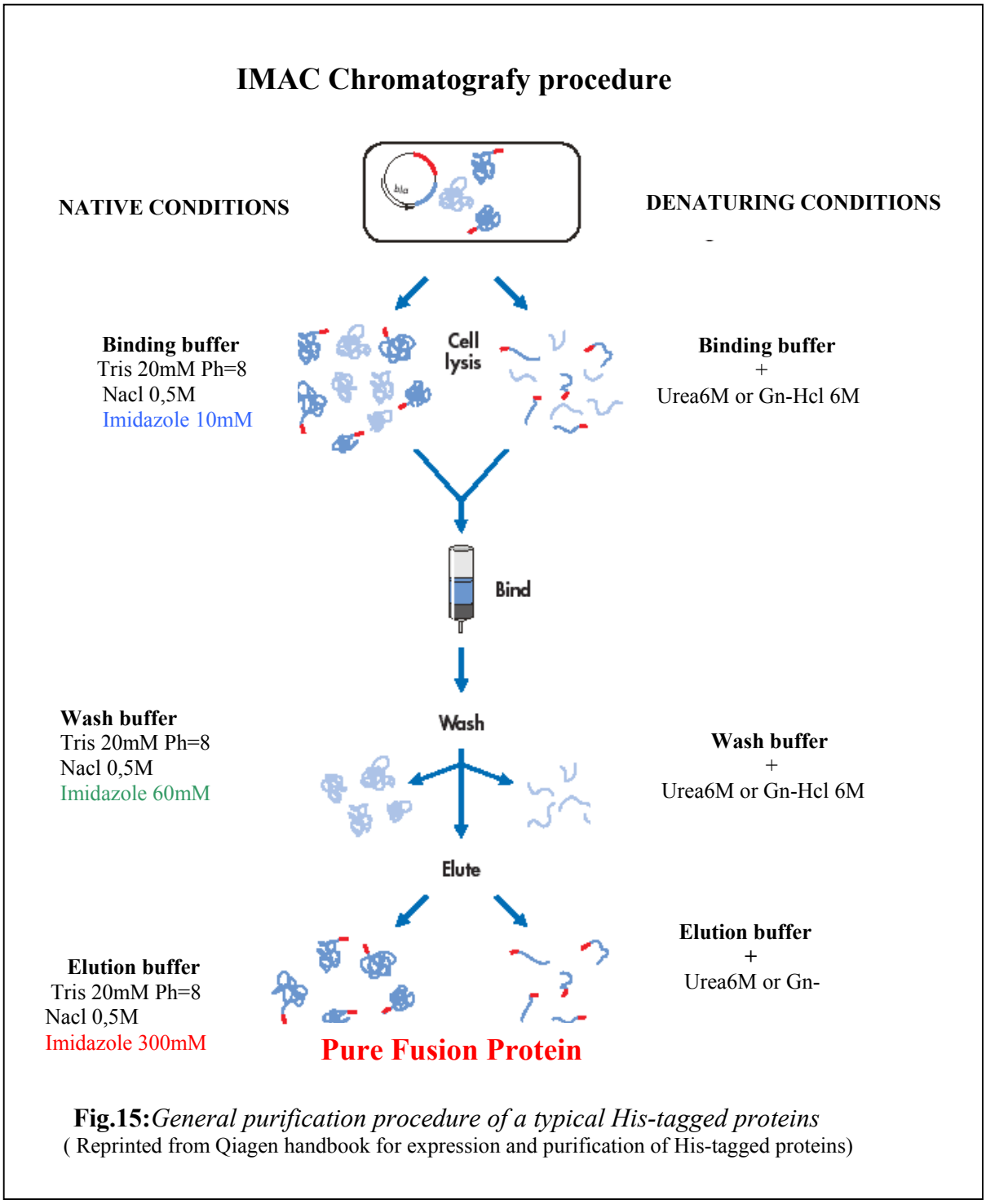



**Fig.15:***General purification procedure of a typical His-tagged proteins*
( Reprinted from Qiagen handbook for expression and purification of His-tagged proteins)

In IMAC chromatography, the target protein is usually washed from the impurities and then eluted using increasing concentration of imidazole, which acts like a competitive

agent. As an alternatives to imidazole elution, histidine-tagged proteins can be eluted from the coloumn by several other methods or combinations of methods:

- for example a lowering of pH within the range of 2.5–7.5 can be used,. At pH values below 4, metal ions will be stripped off the medium.
- EDTA in the elution buffer strips the metal ions from the column.

Enzymatic digestion using a specific protease is then necessary to remove the fusion partner from the target protein and a second IMAC purification is generally performed in order to separate the two proteins. TEV, Fatt.Xa, Thrombin, Prescission Protease, recombinant Enterokinase are some example of proteases that are normally used for the cleavage of fusion proteins. The protease specific recognition site is selected and cloned in the vector codifying for the protein sequence at the cloning step. For each protease/fusion protein pilot experiments should be done to find out suitable conditions.

For both HSco1 and HCox17, the first purification step was performed using IMAC by His-trap column charged with $Zn^{2+}$. Both proteins were eluted in Tris 20mM, NaCl 0.5M and imidazol 300mM pH=8. After concentration the samples were loaded on a PD-10 desalting column in order to exchange the buffer in Tris 50mM, EDTA 0.5mM DTT 1mM. His-GST tag was cleaved with 2ul of Ac-TEV protease/1mg of fusion protein (Invitrogen, Carlsbad, CA) under over night incubation at room temperature. For Hsco1 a second IMAC purification was necessary. At this point, the HSco1, lacking the fusion part containing the his-tag do not bind the column, while, the His-GST tag is eluted at 300mM of imidazole. On the contrary the cut HCox17, in the second IMAC purification, in not eluted in the unbound fraction. It remains bound and it is eluted in the same fraction with of His-GST tag. Exploiting the high difference in molecular weight between HCox17 and GST, this problem is overcome performing a size exclusion chromatography using a 20mM Tris, 1mM DTT, 5mM EDTA, 300mM NaCl buffer, in order to minimize possible protein-protein interaction.

## 2.7 Samples preparation

Copper and zinc metalloproteins, which in many cases bind metal atoms through Cys residues are not easy to handle. Establishing a good "in vitro" metallation protocol, could is a huge effort. Disulfide status is important and can affect sample metallation and stability. Disulfide reduction prior to metallation have to be performed. Unfortunately, the most commonly used reducing agents, like Dithiothreitol DTT, are characterized by a short life time and by a high affinity for metals. Therefore, after disulfide reduction and prior to metallation , the reducing agents have to be removed almost completely from the samples. The re-oxidation of the cysteine residues can induce no specific intra-molecular disulfide bonds formation followed by to protein aggregation and precipitation. Many different trials were done before founding the metallation protocol for Hsco1. Protein reduction and metallation were carried out under nitrogen atmosphere in an anaerobic chamber to prevent oxidation of the cysteine residues. (DTT) was added to the protein in a 10mM concentration to reduce the cysteine residues in the CXXXC motif prior to metal reconstitution. The buffer was than exchanged into 50 mM phosphate (Pi) buffer at pH 7.2 to remove the excess of DTT using PD-10 desalting column. In order to obtain Cu(I), Cu(II), Ni(II) metallated protein, progressive titration, monitored by UV-vis or $^{15}$N-HSQC experiment were performed, using the metals respectively in the $[Cu(I)(CH_3CN)_4]PF_6$, $Cu_2SO_4$ and $NiCl_2$ forms. The titration was performed using a maximum protein concentration of 0.2mM. Intensive precipitation was observed in all the tests performed with higher protein concentrations. The excess of metals was removed through a PD-10 desalting column. All the NMR samples were prepared in Pi 50mM pH=7.2 buffer. Similar protocol for Cu(I) titration was developed also for the metallation of HCox17.

Polyethylene glycol (PEG)6000/pH, $(NH)_4SO_4$/pH screening using a vapour diffusion methods were performed in order to obtain a Ni(II)-HSco1 crystal suitable for X-Ray diffraction. Different protein concentration (from 30mg/ml to 5 mg/ml) were screened. Crystals of Ni(II)-HSco1 were grew at 20°C from a 0.1 M Tris·HCl/30%PEG6000 solution at pH 8.5. Higher pH, show intensive formation of $Ni(OH)_2$ crystals.

## 2.8 Preliminary protein characterization

Once obtained a pure protein several studies can be done in order to characterize it. First of all, mass spectroscopy analysis is performed in order to verify the protein identity and understand if the sequence has the N-terminal methionine. Solubility and stability of the protein at high concentration generally represent an indication of a good folding. Anyway before proceeding to the preparation of labelled samples the degree of "foldedness" is estimated by 1H-NMR and circular dichroism spectroscopy. The latter technique could be suitable also to evaluate the thermal stability. Size exclusion chromatography coupled with multiangle light scattering is performed in order to determine the aggregation state of the protein in solution. The metal content is analyzed by atomic absorption measurements. Disulfide status could be checked by SDS−PAGE after modification with AMS, or more accurately, by mass spectroscopy after modification with iodoacetamide.

## 2.9 Biophysical characterizations

### - UV fluorescence:

Eukaryotic and prokaryotic cells contain a number of compounds that are fluorescent under UV light excitation. Proteins and peptides, with aromatic amino acids are "intrinsically" fluorescent when excited by UV light. Many enzymatic cofactors, such as FMN, FAD, NAD and porphyrins, which are also intrinsically fluorescent, add to the protein fluorescence. These moieties have a common trait in that they all contain aromatic ring structures that are excited by the UV light. There are also special proteins such as Green Fluorescent Protein, which has an internal serine-tyrosine-glycine sequence that is modified post-translationally and is fluorescent in the visible light region.

The three amino acid residues that are primarily responsible for the inherent fluorescence of proteins are tryptophan, tyrosine and phenylalanine. These residues have distinct absorption and emission wavelengths and differ in the quantum yields (Table.1).

**Table 1. Fluorescent Characteristics of the Aromatic Amino Acids.**

| | Absorption | | Fluorescence | |
|---|---|---|---|---|
| Amino Acid | Wavelength (nm) | Absorbtivity | Wavelength (nm) | Quantum Yield |
| Tryptophan | 280 | 5,600 | 348 | 0.20 |
| Tyrosine | 274 | 1,400 | 303 | 0.14 |
| Phenylalanine | 257 | 200 | 282 | 0.04 |

Tryptophan is much more fluorescent than either tyrosine or phenylalanine. However, the fluorescent properties of tryptophan are solvent dependent. As the polarity of the solvent decreases, the spectrum shifts to shorter wavelengths and increases in intensity. For this reason, tryptophan residues buried in hydrophobic domains of folded proteins exhibit a spectral shift of 10 to 20 nm. Tyrosine can be excited at wavelengths similar to that of tryptophan, but emits at a distinctly different wavelength (303nm). While tyrosine is less fluorescent than tryptophan[14], it can provide significant signal, as it is often present in large numbers in many proteins. Tyrosine fluorescence has been observed to be quenched by the presence of nearby tryptophan moieties via resonance energy transfer, as well as by ionization of its aromatic hydroxyl group. Phenylalanine is very weakly fluorescent[14] and can only be observed in the absence of both tryptophan and tyrosine. Protein fluorescence is generally excited at 280 nm or at longer wavelengths, usually at 295 nm. In the first case, we obtain the excitation of both tryptophan and tyrosine residues but, due to tryptophan's greater absorptivity, the fluorescence spectrum usually resembles that of tryptophan. In the second case, using an excitation wavelength of 295 nm we can obtain a selective excitation of the tryptophan.

The fluorescence of the aromatic residues varies in a somewhat unpredictable manner in various proteins. The quantum yield may be either increased or decreased by the folding. Accordingly, a folded protein can have either greater or less fluorescence than the unfolded form. The intensity of fluorescence is not very informative in itself. The magnitude of intensity, however, can be used as a probe of the perturbation of the folded state.

The fact that protein conformal transitions, corresponding at the transition between different states, like folded and unfolded, oxidated and reduced, are generally characterized by different fluorescence intensities[15] was exploit in order to determine the relative stability of this states under different conditions. Progressive protein unfolding in urea[16], or a disulfide bond red-ox potential[17] are some examples of interesting protein properties that can be monitored in this way.

In the present work, the red-ox potential of WT-HSco1 and P174L-HSco1 mutant was investigated by fluorescence spectroscopy (see Cap.5) in order to verify if the pathogenic mutation affects the red-ox properties of HSco1.

Moreover, proteins can be covalently labeled with various fluorophores, thus producing fluorescent protein conjugates. The emission from these attached tags is called "extrinsic" fluorescence**.** Tagging a protein with fluorescent labels is an important and valuable tool for protein characterization.

In the present work using three different fluoroprobes we performed the following assays:

**1) Thermal-Shift assay (Thermofluor)**: The conceptual basis of this method is that folded and unfolded proteins can be distinguished through exposure to a hydrophobic fluoroprobe, Sypro Orange. This probe has low quantum yield in aqueous solution but is highly fluorescent in non polar environments with low dielectric constants such as hydrophobic sites in proteins. When a protein starts to unfold or melt, the dye binds to exposed hydrophobic parts of the protein, resulting in a significant increase in fluorescence emission **(Fig. 16)**. The fluorescence intensity reaches a maximum and then starts to decrease, probably due to precipitation of the complex of the fluorescent probe and the denatured protein. Thermally induced unfolding is an irreversible unfolding process following a typical two-state model with a sharp transition between the folded and unfolded states, where Tm is defined as the midpoint of temperature of the protein-unfolding transition[18]. For several proteins, melting temperatures, obtained with the thermofluor method have been shown to correlate well with values determined by other biophysical methods such as circular dichroism (CD), turbidity measurements, and differential scanning calorimetry.

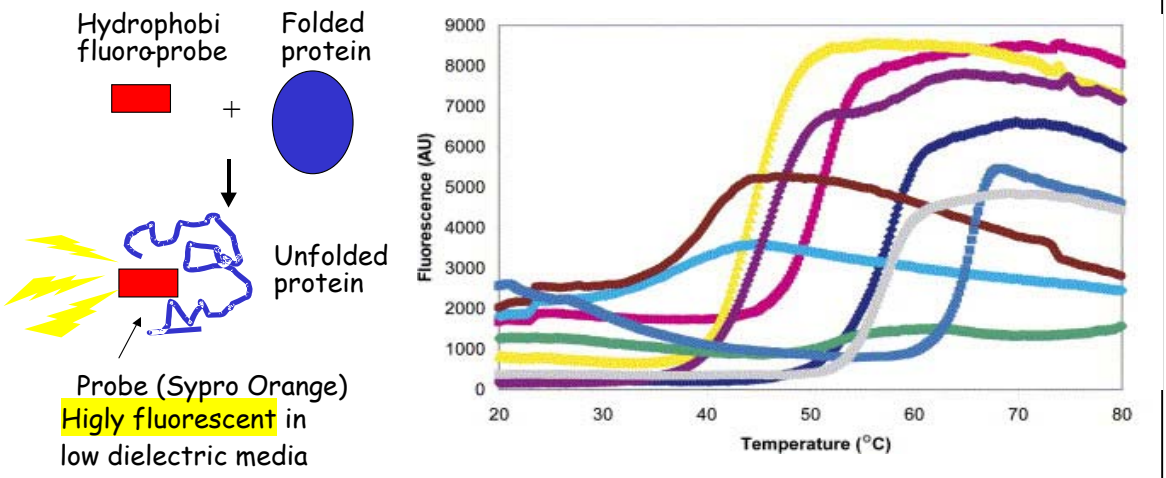



**Fig. 16:** ***Typical Thermal shift assay results.*** (Reprinted from Ericsson UB et.al. Analytical Biochemistry Volume 357, Issue 2 , 15 October 2006, Pages 289-298)

In this work the use of Thermofluor was limited to the Tm measurement of WT-Hsco1 and P174L pathogenic mutant in order to compare their thermal stability (see capther n°5).

**2)ThioflavinT (ThT) fluorescence:** It a commonly used method to monitor fibril formation. This method is particularly attractive since ThT fluoresces only when bound to fibrils. The reaction between the protein and ThT is completed within 1 minute and ThT does not interfere with aggregation. Free ThT has excitation and emission maxima at 350 nm and 450 nm, respectively. However, upon binding to fibrils the excitation and emission change to 450 nm and 485 nm, respectively[21]. This fluorometric technique is really important in the studies of neurodegenerative diseases because it allows the kinetic elucidation of the fibril assembly process. In this work we used ThT probes, coupled with size esclusion chromatography, in order to investigate the tendency of the apo, copper depleted and fully metallated, human SOD1 to form fibrillar aggregates under incubation in condition close to the physiological ones (100uM and 37°C).

**3)Free-thiols quantification:** Estimation of free thiols during aggregation was performed by 4-acetamido 4' maleimidylstilbene-2,2'-disulfonic acid (AMS) modification. This iodoacetamide derivate has high water solubility and is readily conjugated to thiols.

AMS is a stilbene derivate, and shows a typical UV absorption at around 328 nm and emission maximum at 408 nm. Fluorescence measurements were performed in order to monitor the variation of free thiols during the incubation at 37°C of the protein samples. The methods require a preliminary calibration curve for the free cysteine quantification. In our case, this calibration curve, was determined using freshly prepared apo SOD1 WT at different concentrations as a standard. MALDI analysis proved that there is a maximum of two free cysteines per apo SOD1 monomer.

## 2.10 Structural characterization

X-ray crystallography and NMR spectroscopy are the two main techniques that can provide structures of macromolecules at atomic resolution. Both techniques are well established and play a key role in structural biology as a basis for a detailed understanding of molecular functions. Their respective different advantages and disadvantages in terms of sample preparation and data collection and analysis make them complementary in Structual proteomics. X-ray studies usually requires substantial investment of time in order to optimize the crystallization conditions and obtain a crystal with a good diffraction properties This process can take weeks or months, but once a well-diffracting (<2.5Å) crystal is obtained, the structure determination can proceed quite quickly. Whereas X-ray crystallography requires single crystals, NMR measurements are carried out in solution under conditions that can be as close as possible to the physiological state. Some times even if crystal structures are available, additional data in solution would be needed to determine the potential biological function of the protein. NMR is not only capable of solving protein structures to atomic resolution but also has the unique ability to accurately measure the dynamic properties of proteins and can also supply information on protein folding and on intra-, as well as, intermolecular interactions. Furthermore, the analysis through NMR spectroscopy easily allows the characterization under several, different experimental conditions, such as different ionic strength and pH. Generally, a promising protein sample characterized by a good circular dichroism spectra, $^{1}$H-NMR or better from a good $^{1}$H-$^{15}$N HSQC spectra and stable on time (in the mM range) have high probability to be suitable for NMR structural characterization. However, the two majors bottlenecks limiting the NMR in structural biology are represented to the long time request for the data analysis and the drawback size limit of protein. The current is around 35KDa but recent advances in both hardware and experimental design promise to allow the study of much larger proteins[19] .

### 2.10.1 Structure Determination of Proteins with NMR Spectroscopy

NMR spectra of biological macromolecules contain hundreds or even thousands of resonance lines which cannot be resolved in a conventional one-dimensional spectra (1D). Further, the interpretation of NMR data requires correlations between different nuclei, which are implicitly contained in 1D spectra but often difficult to extract. Multidimensional NMR spectra provide both, increased resolution and correlations which are easy to analyse.

The crucial step in increasing the dimensionality of NMR experiments lies in the extension from one to two dimensions. A higher dimensional NMR experiment consists of a combination of two-dimensional (2D) experiments. All 2D NMR experiments use the same basic scheme which consists of the four following, consecutive time periods

**excitation - evolution - mixing – detection**

During the excitation period the spins are prepared in the desired state from which the chemical shifts of the individual nuclei are observed during the evolution period t1. In the mixing period the spins are correlated with each other and the information on the chemical shift of one nucleus ends up on an other nucleus of which the frequency is measured during the detection period t2. A resonance in the 2D spectrum, such as a cross peak, represents a pair of nuclei that suitably interact during the mixing time.

The extension from a 2D to a n-dimensional (nD) NMR experiment consists in the combination of (n-1) two-dimensional experiments which contains only one excitation and one detection period but repeats the evolution and mixing times (n-1) times. A typical nD NMR experiment thus follows the scheme:

**excitation - (evolution - mixing)$_{n-1}$ – detection**

where the bracket repeats (n–1) times. Only during the detection period the signal is physically measured and this period is often referred to as the direct dimension in contrast to the evolution periods which are referred to as indirect dimensions.

The NMR multidimensional measurements almost always use protons ($^{1}$H) and depending on the isotope labelling, $^{13}$C and/or $^{15}$N nuclei. A 3D spectrum can for example be obtained by correlating the amide groups with the α-carbon nuclei attached to $^{15}$N. The

chemical shifts of these carbon nuclei are used to spread the resonances from the 2D plane into a third dimension. The sensitivity obtainable with these types of nuclei greatly varies even if the sample is fully isotope labelled with $^{13}C$ or $^{15}N$. The proton offers the best sensitivity and for this reason constitutes the preferred nucleus for detection of the NMR signal. The other nuclei are usually measured during evolution periods of multidimensional NMR experiments and their information is transferred to protons for detection.
For small proteins (less than 10kDa), it is not required to label the sample with $^{13}C$ or $^{15}N$. In this case the assignment strategy makes use of a combination of 2D homonuclear $^{1}H$ NMR experiments such as COSY/TOCSY, and NOESY/ROESY spectra.

COSY- and TOCSY-type experiments, where COSY stands for *CO*rrelation *S*pectroscop*Y* and TOCSY for *To*tal *C*orrelation *S*pectrosco*py*, correlate different nuclei *via J* coupling[20,21]. In proteins which are isotope labelled with $^{15}N$ and $^{13}C$ *J* couplings between $^{1}H$, $^{15}N$ and $^{13}C$ allow *through-bond* correlations across the peptide bond.

*Through-space* correlations are instead measured via the nuclear Overhauser effect (NOE) and provide the basis for geometric information required to determine the structure of a macromolecule[22]. The NMR method for protein structure determination relies on a dense network of distance constraints derived from NOEs between nearby hydrogen atoms in the protein[22]. NOEs connect pairs of hydrogen atoms separated by less than about 5 Å. In contrast to COSY-type experiments the nuclei involved in the NOE correlation can belong to amino acid residues that may be far apart along the protein sequence but close in space. For molecules with a molecular weight of more than 5 kDa the intensity of an NOE is approximately proportional to $r^{-6}$ and to the molecular weight, where $r$ is the distance between the two interacting spins. NMR experiments which measure the NOE are often referred to as NOESY experiments where NOESY stands for *NOE S*pectroscop*Y* [23].

## *Sequence Resonance Assignment*

For a detailed analysis of the information content of NMR spectra nearly complete assignments of signals in the spectra to individual atoms in the molecule are a requirement. The application of multidimensional NMR spectroscopy allowed the development of general strategies for the assignment of signals in proteins. All procedures use the known protein sequence to connect nuclei of amino acid residues which are neighbours in the sequence. As mentioned before, for unlabelled proteins smaller than 10 kDa the combination of the [$^{1}H$,$^{1}H$]-COSY or TOCSY, used for the sequential assignment, with the

[$^{1}$H, $^{1}$H]-NOESY spectrum allows the assignment of most proton NMR signals to individual protons[23].

For larger proteins extensive signal overlap prevents complete assignments of all $^{1}$H signals in proton spectra. This barrier can be overcome with 3D NMR technique and uniformly $^{13}$C and $^{15}$N labelled proteins. The resonance assignment of single ($^{15}$N or $^{13}$C) labelled proteins using 3D experiments is basically an extension of Wüthrich's strategy which exclusively relies on homonuclear $^{1}$H NMR experiments. With these methods, systems with molecular weights up to approximately 30 kDa can be studied.

In $^{13}$C, $^{15}$N-labeled proteins a sequential assignment strategy is based on through-bond correlations across the peptide-bond between sequential amino acids. This procedure circumvents the use of NOESY spectra already in the assignment step. Most of these correlation experiments use the three types of nuclei $^{1}$H, $^{15}$N, $^{13}$C and are referred to as triple resonance experiments.

The 3D triple resonance experiments exclusively correlate the resonances of the peptide backbone (HN(i), N(i), Cα(i), Hα(i), Cα(i-1), CO(i) and CO(i-1)). **Fig. 17** shows the spin system of the peptide backbone and indicates the size of the coupling constants used for magnetization transfer in double $^{13}$C-, $^{15}$N-labelled proteins.

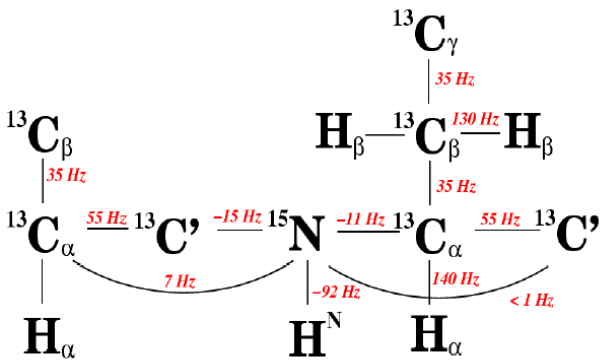



**Fig. 17:** *Spin system of the peptide backbone and the size of the $^{1}J$ and $^{2}J$ coupling constants that are used for magnetization transfer in $^{13}$C-, $^{15}$N-labelled proteins.*

The 3D experiments used to identify the backbone resonances are, usually, HNCA or HNCACB, HN(CO)CA or HN(CO)CACB, HNCO, HN(CA)CO and HNHA[24]. The HNCACB for example, correlates each H-$^{15}$N group with both the intra- and the neighbouring inter-residue Cα and Cβ. These four types of connectivities are discriminated using the HN(CO)CACB experiment, in which only the inter-residue HN-Cα and Cβ couplings are observed.

Similar strategy can be used to assign the other resonances in the other triple resonance spectra.

In the case of proteins with a molecular weight larger than 30 kDa the use of TROSY-type experiments[25] is necessarily. TROSY experiment can reduce the signal loss, which is the direct consequence of the slower correlation tumbling of large molecules which results in faster relaxation and consequently broader lines in the NMR spectrum. TROSY uses constructive interference between different relaxation mechanisms and works best at the highest available magnetic field strengths in the range of 700 to 900 MHz proton resonance frequency. With TROSY the molecular size of proteins accessible for detailed NMR investigations has been extended several fold. The TROSY technique benefits a variety of triple resonance NMR experiments as the 3D HNCA and HNCOCA [26] and the TROSY-based NOESY experiments for the collection of structural constraints are also available [27].

Since the Hα and Cα/β chemical shifts have been assigned, 3D H(C)CH-TOCSY and (H)CCH-TOCSY [28] experiments are then used to link the side chain spin systems to the backbone assignments. These two experiments provide information for the assignment of the side chain protons and of the side chain carbons, respectively.

A complete set of backbone chemical shifts for all Hα, Cα, Cβ and CO resonances can be used to predict the secondary structure of the protein[29]. One technique in particular, the Chemical Shift Index (CSI)[30], has been widely used for the quantitative identification and location of secondary structure in proteins.

The method relies on the fact that the chemical shifts of the different nuclei in the protein backbone are related both to the type of amino acid and to the nature of the secondary structure they are located in. By comparing the actual chemical shift for a nucleus in a specific amino acid with a reference value, it is possible to predict in what secondary structure element the nucleus resides. The reference value that you compare with is the random coil chemical shift for that same nucleus in the same amino acid.

### *Collection of conformational constraints*

In structure calculation, geometric conformational information in the form of distances and/or torsion angles has to be derived from the NMR data. The latter have to be supplemented by information about the covalent structure of the protein, such as the amino acid sequence, bond lengths, bond angles, chiralities, and planar groups, as well as by steric repulsion between non-bonded atom pairs.

Although a variety of NMR parameters contain structural information, the crucial information comes from NOE measurements which provide distance information between pairs of protons. Supplementary constraints can be derived from through bond correlations in the form of dihedral angles[23]. Furthermore, CSI data, provides, as before mentioned, information on the type of secondary structure. Such information can be included in a structure calculation by restricting the local conformation of a residue to the α–helical or β-sheet region of the Ramachandran plot through torsion angle restraints. Hydrogen bonds can also be experimentally detected via through-bond interactions[31] and they can be useful during structure calculations of larger proteins when not enough NOE data are available yet. Finally one other class of conformational restraints originate from residual dipolar couplings in partially aligned or paramagnetic molecules and gives information on angles between covalent bonds and globally defined axes in the molecule, namely those of the magnetic susceptibility tensor [32,33].

With sufficient structural constraint, a folded conformation can be determined at atomic resolution. The result of NMR structure determination is not one model, but a set of similar models, all of which fit the experimentally the structural constraints. The RMSD (root mean square deviation) between these models is used to assess how well the structure calculations have converged. The best structures have backbone RMSD values of less than 1 Å. A final structure is obtained by averaging the models, and then finding the conformation of minimum energy that lies nearest to this average conformation.

## 2.11 Reference list


(1) Venter, J.C.; Adams, M.D.; Myers, E.W.; Li, P.W.; Mural, R.J. (2001) *Science* **291**, 1305-1351.

(2) Lander, E.S.; Linton, L.M.; Birren, B.; Nusbaum, C.; Zody, M.C. (2001) *Nature* **409**, 860-921.

(3) Varshavsky, A. (1996) *Proc.Natl.Acad.Sci.USA* **93**, 12142-1214.

(4) Holz, C.; Hesse, O.; Bolotina, N.; Stahl, U.; Lang, C. (2002) *Protein Expression & Purification*. **25**, 372-378.

(5) Boettner, M.; Prinz, B.; Holz, C.; Stahl, U.; Lang, C.; (2002) *Journal of Biotechnology* **99**,51-62.

(6) Klammt, C.; Schwarz, D.; Lohr, F.; Schneider, B.; Dotsch, V.; Bernhard, F. (2006) *FEBS J.* **273**, 4141-53.

(7) Hammarstrom, M.; Hellgren, N.; Van Den Berg, S.; Berglund, H.; Hard, T. *Protein Sci* (2002) 11, 313–321.

(8) Miroux, B.; Walker J.E. (1996) *J Mol Biol* **260**, 289–298.

(9) Qing, G.; Ma, L.C.; Khorchid, A.; Swapna, G.V.; Mal, T.K.; Takayama, M.M.; Xia, B.; Phadtare, S.; Ke H.; Acton T.; (2004) *Nat Biotechnol* **22,** 877–882.

(10)) De Marco A.; De Marco V. *J Biotechnol* (2004)**109**, 45–52.

(11) Kapust R.B.; Waugh D.S *Protein Sci* (1999) **8**, 1668–1674

(12) Landy, A. (1989) *Ann. Rev. Biochem.* **58**,913.

(13) Woestenenk, E.A.; Hammarstrom, M.; Van den Berg, S.; Hard, T.; Berglund, H. (2004) J Struct Funct Genomics. **5**,217-29.

(14) Principles of Fluorescence Spectroscopy 2nd Edition (1999) Lakowicz, J.R. Editor, Kluwer Academic/Plenum Publishers, New York, New York.

(15) Holmgren, A.*J Biol Chem.* (1972) 247,1992-8

(16) Gasset, M.; Alfonso, C,; Neira, J,L.; Rivas, G. ; Parajes, M.A. Biochem. J. (2002) **361**, 307–315.

(17) Haugstetter, J.; Blicher, T.; Ellgaard, L. (2005) *J Biol Chem.* **280**, 8371-80.

(18) Lo, M.C.; Aulabaugh, A.; Jin, G.; Cowling, R.; Bard, J.; Malamas, M.; Ellestad, G. (2004) *Anal Biochem.* **33**2, 153-9.

(19) Wuthrich, K. (1998) *Nat. Struct. Biol.* **5**, 492–495.

(20) Aue, W.P.; Bartholdi, E.; Ernst, R.R. (1976) *J.Chem.Phys.* **64**, 2229-2235.

(21) Wider, G.; Macura, S.; Kumar, A.; Ernst, R.R.; Wüthrich, K. (1984) *J.Magn.Reson.* **56**, 207-234.

(22) Wider, G. (1998) *Progr.NMR Spectrosc.* **32**, 193-275.

(23) Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley: New York

(24) Kumar, A.; Ernst, R.R.; Wüthrich, K. (1980) *Biochem.Biophys.Res.Commun.* **95**, 1104

(25) Kay, L.E.; Ikura, M.; Tschudin, R.; Bax, A. (1990) *J.Magn.Reson.* **89**, 496-514.

(26) Pervushin, K. (2000) *Q.Rev.Biophys.* **33**, 161-197.

(27) Salzmann, M.; Wider, G.; Pervushin, K.; Senn, H.; Wüthrich, K. (1999) *J.Am.Chem.Soc.* **121**, 844-848.

(28) Pervushin, K.V.; Wider, G.; Riek, R.; Wuthrich, K. (1999) *Proc.Natl.Acad.Sci.USA* **96**, 9607-9612.

(29) Kay, L.E.; Xu, G.Y.; Singer, A.U.; Muhandiram, D.R.; Forman-Kay, J.D. (1993) *J.Magn.Reson.Ser.B* **101**, 333-337.

(30) Wishart, D.S.; Sykes, B.D.; Richards, F.M. (1991) *J.Mol.Biol.* **222**, 311-333.

(31) Wishart, D.S.; Sykes, B.D.; Richards, F.M. (1992) *Biochemistry* **31**, 1647-1651.

(32) Cordier, F.; Grzesiek, S. (1999) *J.Am.Chem.Soc.* **121**, 1601-1602.

[33] Tolman, J.R.; Flanagan, J.M.; Kennedy, M.A.; Prestegard, J.H. (1995) *Proc.Natl.Acad.Sci.USA* **92**, 9279-9283.

# 3

# RESULTS

During the three years of the PhD course, I mainly focused my work on genome browsing, gene cloning, protein expression and purification, preliminary biophysical characterization on copper proteins and NMR sample preparation. Specifically:

A PSI-BLAST search was performed in order to find putative Cox17, Sco1/2, and Cox11 homologs in prokaryotic genomes. Multiple sequence alignments showing sequence similarity within these protein families were then obtained through ClustalW program. The search for sequences similar to Sco1/2 located 102 putative sequences from 69 bacterial and archaeal genomes. All Sco-like sequences have a CxxxC consensus motif. A DxxxD motif is also conserved together with a histidine residue. Multiple Sco-like sequences (up to five) can be found in a single organism. The genomic-context analysis for Sco1 genes supports the role of Sco proteins in copper homeostasis. However, at variance with Cox11, which, when present, is close to genes encoding COX subunits, a dramatic variability is observed in the localization of Sco genes suggesting a multifunctional role of this protein family.

Human Cox17 and two different truncated forms of the human Sco1 gene, both lacking the N-terminal mitochondrial targeting sequence and the single-transmembrane, were cloned. The P174L-HSco1 pathogenic mutant was successfully produced using QuikChange mutagenesis kit. A protocol for expression and purification of unlabelled and labelled ($^{15}N$ and $^{13}C/^{15}N$) samples for each of these proteins was developed.

The degree of folding and thermal stability of the above proteins were evaluate by circular dichroism and thermofloor analysis. Size-exclusion chromatography coupled with multiangle light scattering was performed to verify the aggregation states of the different HSco1 truncated forms. We have succeeded in preparing Cu(I) or Ni(II) derivatives of WT-HSco1, P174L-HSco1 mutant and human Cox17 suitable for NMR structure determination. Cu(II) and Ni(II)HSco1 derivates have been obtained and crystals were produced in the latter case. The metal binding properties, for each of these proteins, were studied using NMR, EPR and UV/VIS spectroscopy. The red-ox potentials of WT-HSco1 and P174L-HSco1 were also measured through fluorescence spectroscopy.

Human SOD1, in which the non-conserved cysteine residues, Cys6 and Cys111, were mutated to Ala and Ser respectively (AS form), were expressed and purified. A protocol for the preparation of fully reduced and demetallated SOD1 forms was also developed. $^{15}N$ labelled apo and metallated SOD1 samples were prepared for NMR analysis. Size-exclusion chromatography was performed to verify the aggregation states of the SOD1

protein under different oxidation and metallation states. Cicular dichroism measuraments were performed to estimate the secondary structure content of the different forms of SOD1.

WT-human SOD1, containing the non-conserved cysteine residues, Cys6 and Cys111 were also expressed and purified. To determine the propensity for oligomerization of the mature, i.e., disulfide-intact, WT human SOD1 under the relatively mild conditions likely to be encountered by the protein *in vivo,* 100uM samples of apo, partially and fully metallated SOD1 were incubated at 37°C. These sample allowed also to investigate the roles of metallation in hindering or promoting the oligomerization. Changes in secondary structure were monitored by CD spectroscopy, and changes in the aggregation state were followed by gel filtration chromatography and ThT binding fluorescence. Variation of free thiols during the incubation, in order to investigate the possible formation of an intermolecular disulfide bonds during the aggregation process, was monitored by 4-acetamido 4' maleimidylstilbene-2,2'-disulfonic acid (AMS). The covalent nature of protein aggregates was confirmed analyzing the persistence of the oligomeric species in presence a chaotropic agents as guanidinium hydrochloride, which disrupts hydrogen bond networks in the secondary and tertiary structure, but not the possible disulfide.

The results and findings related to the above-mentioned proteins are reported in the following articles.

# 3.1

## Ortholog Search of Proteins Involved in Copper Delivery to Cytochrome c Oxidase and Functional Analysis of Paralogs and Gene Neighbors by Genomic Context

Fabio Arnesano, Lucia Banci, Ivano Bertini,* and Manuele Martinelli

*Journal of Proteome Research* (2005), **4**, 63-70.

research articles Journal of proteome research

# Ortholog Search of Proteins Involved in Copper Delivery to Cytochrome *c* Oxidase and Functional Analysis of Paralogs and Gene Neighbors by Genomic Context

**Fabio Arnesano, Lucia Banci, Ivano Bertini,* and Manuele Martinelli**

*Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy*

**Received August 6, 2004**

Cytochrome *c* oxidase (COX) is a multi-subunit enzyme of the mitochondrial respiratory chain. Delivery of metal cofactors to COX is essential for assembly, which represents a long-standing puzzle. The proteins Cox17, Sco1/2, and Cox11 are necessary for copper insertion into $Cu_A$ and $Cu_B$ redox centers of COX in eukaryotes. A genome-wide search in all prokaryotic genomes combined with genomic context reveals that only Sco and Cox11 have orthologs in prokaryotes. However, while Cox11 function is confined to COX assembly, Sco acts as a multifunctional linker connecting a variety of biological processes. Multifunctionality is achieved by gene duplication and paralogs. Neighbor genes of Sco paralogs often encode cuproenzymes and cytochrome *c* domains and, in some cases, Sco is fused to cytochrome *c*. This led us to suggest that cytochrome *c* might be relevant to Sco function and the two proteins might jointly be involved in COX assembly. Sco is also related, in terms of gene neighborhood and phylogenetic occurrence, to a newly detected protein involved in copper trafficking in bacteria and archaea, but with no sequence similarity to the mitochondrial copper chaperone Cox17. By linking the assembly system to the copper uptake system, Sco allows COX to face alternative copper trafficking pathways.

**Keywords:** cytochrome *c* oxidase • enzyme assembly • copper trafficking • genomic-context • paralogs • origin of mitochondria

## Introduction

Cytochrome *c* oxidase (COX) is an enzyme (EC 1.9.3.1) which reduces oxygen to water and generates the proton gradient that drives ATP synthesis. COX is a multi-subunit complex[1,2] which requires a large protein machinery for its assembly.[3] It also contains several metal cofactors, whose insertion and binding in the proper subunit is required to produce the final, active enzyme.[4] It has been shown that the mammalian enzyme is present in a dimeric form.[2]

Mitochondrial COX (aa3-type), present in all the eukaryotic organisms characterized so far, contains two copper centers, designated $Cu_A$ and $Cu_B$. $Cu_A$, present in subunit II of COX (Cox2), is formed by two copper ions bound to two His and two bridging Cys residues of the consensus motif $Hx_nCxExCGx_2$-$Hx_2M$. $Cu_B$, present in subunit I of COX (Cox1), is formed by a copper ion, with a binding motif $Hx_3Yx_{44}HH$, coupled to heme $a_3$, thus forming a binuclear iron/copper center.[2] The $Cu_A$ center receives the electron from cytochrome *c* (cyt *c*) and transfers it to heme *a* and finally to the heme $a_3$−$Cu_B$ center.[5] Three subunits (Cox1, Cox2 and Cox3) are encoded by the mitochondrial genome and the remaining subunits are encoded by the nuclear genome, therefore subunits synthesized in two compartments must be coordinately recruited to assemble in mitochondria.[3]

Many bacteria and archaea contain an aa3-type COX complex with an arrangement similar to that of the mitochondrial enzymes.[6] The prokaryotic aa3-type COX contains identical redox centers and homologous core polypeptides (subunits I and II) to the mitochondrial COX.[1] In some bacterial aa3-type COX, a cyt *c* is fused to the C-terminus of Cox2 (caa3-type COX). In bacterial species, it is often found in addition to, or instead of aa3-type, a cbb3-type COX which, however, lacks the $Cu_A$ center.[7]

It is now well-established that free metal ions are not present in the cytoplasm while they are always bound to proteins, called metallochaperones, which transfer them to the requiring target proteins.[8] In several cases, metal trafficking occurs through a cascade of transfers which involves a series of proteins. Among them there are soluble metal transporters and membrane proteins, like ATPases and permeases, which allow transfer of the metal across membranes, thus controlling metal import/efflux in the cell or its transfer from one cellular compartment to another.[9] While these pathways start to be unraveled for some processes, like copper insertion into multicopper oxidases and superoxide dismutase, little is known on the metal insertion in the COX subunits.

In eukaryotes, it has been found that three proteins, Cox17,

* To whom correspondence should be addressed. Tel: +39-055-4574272. Fax: +39-055-4574271. E-mail: bertini@cerm.unifi.it.

10.1021/pr049862f CCC: $30.25 



Published on Web 01/13/2005

Sco1, and Cox11 are required for copper delivery to COX (reviewed in ref 4). Cox17 is a soluble protein of ~70 amino acids involved in providing copper ions for formation of both $Cu_A$ and $Cu_B$ sites in mitochondria. Cells lacking Cox17 are respiratory deficient, but this defect is complemented by addition of exogenous copper to cells.[10] In Cox17, three out of six conserved cysteines are present in a CCxC sequence motif, essential for Cu(I) binding.[11] Sco1 and Cox11 are involved, together with Cox17, in copper delivery to $Cu_A$ and $Cu_B$, respectively.[12,13] Both Sco1 and Cox11 contain a single transmembrane helix in the N-terminal segment which anchor them to the inner mitochondrial membrane. The specific function of both proteins is not clear. It has been shown that while the lack of Cox17 can be compensated by exogenous copper supply,[10] the lack of Sco1 cannot be and Sco1 deficient cells are not able to form a functionally active COX enzyme.[12] Some eukaryotes contain another protein, highly similar to Sco1, called Sco2, whose role in copper insertion into $Cu_A$ is also not yet fully elucidated.[12] Sco2 is able to restore respiration in Cox17, but not in Sco1 mutants, indicating that Sco1 and Sco2 have overlapping but not identical functions. While the precise role of each of these proteins in copper incorporation remains unclear, recent studies have revealed that inherited mutations in these proteins can result in severe pathology in human infants in association with cytochrome *c* oxidase deficiency.[14,15]

Correlating data on mitochondrial proteins with information about their evolutionary histories might yield insights into the nature and function of eukaryotic cells,[16] aimed at ultimately understanding the molecular bases of mitochondrial disorders, which represent some of the most common metabolic genetic diseases.[17] Given the similarity between mitochondrial and prokaryotic oxygen respiratory chain it is very likely that some of the proteins required for copper delivery to the respiratory complexes are conserved. Indeed, proteins homologous to the eukaryotic COX accessory proteins have been located in some bacterial species even if scattered pieces of information are available. Combining genomic-context analysis (conserved neighborhood, gene fusions, phylogenetic occurrence) with homology-based methods (genome search, structure modeling, correlated mutation analysis) one might be able to predict both the pathway in which a protein operates and its molecular function.[18] Therefore, to have a comprehensive description of proteins involved in copper delivery to COX, we have performed a genome-wide search in prokaryotic organisms for sequences sharing similarity with human Cox17, Sco1/2, and Cox11, and we extended the analysis to genes close to the found proteins. The results shed a new light on the COX assembly system and are discussed in connection with the changing redox and copper metabolisms of the cellular compartment where the soluble COX assembly components are carried or located (i.e., the periplasm in Gram-negative bacteria and the intermembrane space in mitochondria). The evolutionary history of this metalloenzyme is linked to the endosymbiotic origin of mitochondria.

## Procedures

**Sequence Search and Genomic-Context Analysis.** We performed a PSI-BLAST search ($E < 0.01$)[19] in order to find putative Cox17, Sco1/2, and Cox11 homologs in genomic databanks, using the consensus motif to refine the search.[20] Multiple sequence alignments were obtained with the program CLUSTALW.[21] The genomic-context analysis for the selected proteins was performed on fully sequenced genomes and fragments containing gene clusters, which were available at December 2003 in the GenBank (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=Genome). A genomic-context network was built with the program STRING (http://www.bork.embl-heidelberg.de/STRING), which integrates the three types of genomic context (conserved neighborhood, gene fusion and co-occurrence) into a single score function.[22] Assignment of functional categories of genes was derived from the Clusters of Orthologous Groups (COGs) database[23] and automatically made by STRING. When genes could not be assigned to COGs they are referred to as nonsupervised orthologous groups (NOGs). The presence of functional modules[24] was deduced from the network.

**3D Structural Models.** Structural models of various Sco orthologs and paralogs as well as of some cyt *c* domains were generated using the program Modeller-6v2.[25] The input alignments for Modeller were obtained with CLUSTALW.[21] Models of Sco paralogs were created using as template the only experimentally determined structure of a member of this class, Sco from *Bacillus subtilis*.[26] For cyt *c* domains every model was created using as template the structure deposited in the Protein Data Bank with the closest sequence to the target protein. The program MOLMOL[27] was used to analyze the structural models in terms of per-residue solvent accessibility and surface properties (shape, electrostatics).

**Correlated Mutation Analysis.** To obtain the 'interaction index' between two selected proteins (A and B),[28] their sequence alignments were reduced to the set of organisms common to the two proteins, and a virtual concatenated alignment was generated by attaching the sequence of protein A to the sequence of protein B from the same organism. A 'correlation value' was calculated with the program PlotCorr[29] for every pair of positions in the concatenated alignment. The pairs were divided into three sets: two for the intraprotein pairs ($C_{AA}$ and $C_{BB}$; pairs of positions within protein A and within protein B) and one for the interprotein pairs ($C_{AB}$; one position from protein A and one from protein B) and the 'interaction index' was calculated by comparing the distribution of interprotein correlation values with the two distributions of intraprotein correlation values. Interaction indexes > 2.0 correspond mostly to true interactions. This method is called in silico two-hybrid system.[28] Interprotein pairs were used to predict interprotein contacts as previously described.[30] For this analysis, we considered only correlated residue pairs with the highest correlation values (>0.75) and, for each protein, only those residues involved in at least three predicted contacts with the other protein.

## Results

**Sequence Search.** A search in prokaryotic genomes for sequences similar to human Cox17 produced no results. In contrast, the search for Cox11-like sequences located 36 sequences, all containing a CFCF consensus motif, from 36 genomes of Gram-negative bacteria. Cox11-like sequences were found in 21 out of 89 fully sequenced genomes of Gram-negative bacteria, Cox11 is not found in Gram-positive bacteria and archaea, even if these organisms do contain a Cox1 subunit with a classical $Cu_B$ center. The search for sequences similar to Sco1/2 located 102 putative sequences from 69 bacterial and archaeal genomes. Sco-like sequences were found in 38 out of the 131 fully sequenced bacterial genomes, and in 2 out of the 17 fully sequenced archaeal genomes. For 12 prokaryotic complete genomes, which contain an aa3-type COX with a $Cu_A$

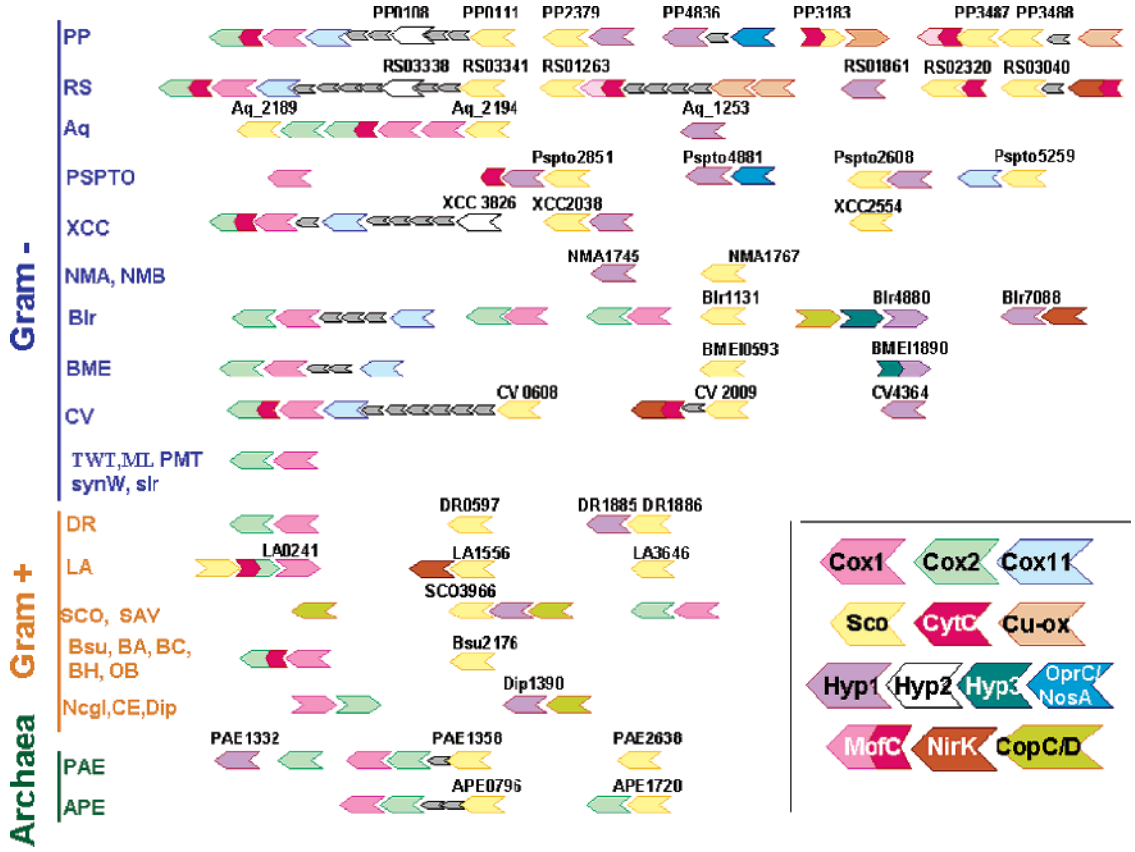



**Figure 1.** Gene neighborhood analysis of cytochrome *c* oxidase accessory proteins. Operon and divergon structures of genes encoding Sco and Cox11 domains and their neighbors are shown. Genes are represented as arrows. The color code is illustrated in the inset. Unrelated genes are shown as gray arrows. For all Sco and most Hyp1 genes is given the numbers they have in the genomes. Gram-negative bacteria are indicated in blue; Gram-positive bacteria are indicated in orange, and Archaea in green. The correspondence of the full species names to the ones used in the figures is as follows: *PP: Pseudomonas putida* KT2440; Bcep: *Burkholderia fungorum*; Rs: *Ralstonia solanacearum*; Aq: *Aquifex aeolicus*; Pspto: *Pseudomonas syringae* pv. tomato str. DC3000; XCC: *Xanthomonas campestris* pv. campestris str. ATCC 33913; NMA: *Neisseria meningitidis* serogroup A strain Z2491; NMB: *Neisseria meningitidis* serogroup B strain MC58; Blr: *Bradyrhizobium japonicum*; BME: *Brucella melitensis*; CV: *Chromobacterium violaceum* ATCC 12472; SynW: *Synechococcus* sp. WH8102; Slr: *Synechocystis* sp. PCC 6803; PMT: *Prochlorococcus marinus* MIT9313; DR: *Deinococcus radiodurans*; LA: *Leptospira interrogans* serovar lai str. 56601 chromosome I; Sco: S*treptomyces coelicolor* A3(2); SAV: *Streptomyces avermitilis* MA-4680; BH: *Bacillus halodurans*; BA: *Bacillus anthracis* A2012; Bsu: *Bacillus subtilis*; BC: *Bacillus cereus* ATCC 14579;OB: *Oceanobacillus iheyensis* HTE83; NCgl: *Corynebacterium glutamicum* ATCC 13032; CE: *Corynebacterium efficiens* YS-314; Dip: *Corynebacterium diphtheriae*; PAE: *Pyrobaculum aerophilum*; APE: *Aeropyrum pernix*.

center, no Sco homologue is present. All Sco-like sequences contain a CxxxC consensus motif. A DxxxD motif is also conserved together with a histidine residue. Sco1-like sequences from Gram-positive bacteria also contain a MxxxM motif, located two residues downstream from CxxxC. Instead, all the eukaryotic Sco1-like sequences share the motif CxxxCxxxxE(D)K(R), i.e., with two adjacent oppositely charged residues after CxxxC. Pairwise residue identities of 20 ± 7% and 37 ± 5% are found over all Sco-like and all Cox11-like sequences, respectively, indicating that Cox11-like sequences are more conserved than Sco-like (see Table 1S and 2S of Supporting Information).

**Genomic-Context Analysis.** We performed a genomic-context analysis in order to predict possible functional associations of Sco and Cox11 homologues with other proteins, based on their coding gene position, phylogenetic occurrence and gene fusions.[31,32] Generally, in prokaryotes, genes which encode the various COX subunits are close each other and are all present or absent together. The conservation of relative gene position derives from the organization of prokaryotic genes into operons which encode proteins involved in the same overall process. This might be then used to extrapolate the findings to species with little or no operon structure, such as eukaryotes, to predict functional relations among genes also for these organisms.[33]

**Context of Cox11 and Sco Genes.** The results on gene neighborhood are summarized in Figure 1 (see Figure 1S of Supporting Information for an extended version). Analyzing the COX operon we found that the Cox11 gene, when present, is close to genes encoding COX subunits. The only exception is represented by *Pseudomonas syringae*, where Cox11 is far from Cox1 and Cox2 genes but it is found adjacent to Sco (see Figure 1). In this peculiar case, Cox2 lacks the ligands of the $Cu_A$ center.

At variance with Cox11, a dramatic variability is observed in the localization of Sco genes. Multiple Sco-like sequences (up to five) can be found in a single organism in different genomic contexts. When five Sco homologues are present (i.e., in *Pseudomonas putida* and *Pseudomonas fluorescens*) one gene is close to caa3-type COX genes, one is close to a protein of unknown function (Hyp1 hereafter, COG2847), one is close to a multicopper oxidase and it is fused to a cyt *c*, and the last

two are adjacent each other and close to another multicopper oxidase. When less than five Sco homologues are present, they can be either close to one of the above-mentioned proteins or to other copper-dependent enzymes, i.e., nitrite reductase or cbb3-type COX. Sco-like sequences of the same organisms can be as different as to have only 15% residue identity.

**Occurrence of Cytochrome *c* Domains.** A Sco gene is found close to COX genes only when a cyt *c* domain is attached to the C-terminus Cox2 (caa3-type COX). This occurs in ten Gram-negative bacteria. Cyts *c* fused to Cox2 are also found in Gram-positive bacteria but, in these organisms, Sco is not close to COX genes. In the case of Sco genes close to multicopper oxidases (i.e., in *Pseudomonas* and *Ralstonia*), a cyt *c* domain with a single CxxCH heme binding motif is present in the same operon, either fused to the periplasmic component of an ABC-type amino acid transporter (and the fused gene is designated MofC,[34] NOG13183) or, remarkably, to Sco itself. In some genomes Sco is found in the vicinity of a gene encoding a copper nitrite reductase (NirK, COG2132), which catalyzes the reduction of nitrite to nitric oxide, a key step in the anaerobic denitrification process, and NirK also contains a cyt *c* fused at its C-terminus. In *Ralstonia*, a Sco gene is found in the cbb3-type COX operon, adjacent to a subunit (FixP, COG2010) containing a *c*-type heme. Also in *Pseudomonas stutzeri*, a Sco homolog, designated ScoP, is close to the operon encoding a cbb3-type COX and it is located downstream of the FnrA gene, which encodes a regulator of the cbb3-type COX gene expression.[35]

**New Potential Copper Transporter.** Some Sco genes are close to a protein of yet unknown function that we call Hyp1 (COG2847). This close neighborhood occurs in a large number of organisms, even when aa3-type COX genes are missing. For instance, Hyp1 is present together with a Sco gene in the pathogens *Neisseria gonorrheae* and *Neisseria meningitidis*, which only have a cbb3-type COX, thus lacking the $Cu_A$ center.[36] We therefore analyzed the sequence and the genomic context of this unknown protein to find possible relationships with known genes.

Hyp1 is a soluble protein mostly occurring in Gram-negative bacteria, and consisting of about 150 amino acids. A Hyp1 gene is also found in a few Gram-positive bacteria where it is characterized by the presence of a single transmembrane segment. No homologues of Hyp1 are found in eukaryotes. All Hyp1 sequences share a conserved H(M)$x_{10}$M$x_{21}$H$x$M consensus motif (Banci L., Bertini I., Ciofi-Baffoni, S., Katsari E., Kubicek K., manuscript in preparation) similar to the Cu(I) binding motif of CopC (COG2372), a well-characterized periplasmic protein involved in copper resistance.[37] CopC is able to selectively bind Cu(I) and Cu(II) at different sites: Cu(II) is bound by two histidines, an aspartate and a glutamate, whereas Cu(I) is bound by a histidine and three methionines. A shift in the redox state causes the copper ion to migrate between the two sites. Thus, CopC acts as a molecular switch that facilitates Cu(II) import to the cytoplasm via the inner membrane protein CopD (COG1276) or Cu(I) export via the outer membrane protein CopB.[37] In several organisms, we found a Hyp1 gene close to a gene encoding CopC fused to CopD (see Figure 1).

Up to three Hyp1 genes can be found in a single organism in different genomic contexts. Hyp1 can be found either close to a Sco gene or to a copper-transporting outer membrane channel protein (NosA or OprC, COG1629)[38] of the family of TonB receptors, which are mostly involved in siderophore-iron uptake.[39] NosA is known to be involved in the copper delivery pathway for the $Cu_A$ site of the nitrous oxide reductase,[40] which is the terminal oxidoreductase of a respiratory process that generates $N_2$ from $NO_3^-$. The observed regulatory responses indicate that the outer membrane protein, NosA, functions in anaerobic metabolism, and since it is repressed by Cu, its role seems to be limited to conditions in which the Cu supply is low.[40] In some organisms, a Hyp1 gene is fused to a membrane protein of unknown function (Hyp3 hereafter, COG4549), while in other organisms Hyp1 and Hyp3 are encoded by two separated genes which form a cluster with the fusion gene encoding CopCD (COG2372 and COG1276) (see Figure 1). In this case, CopC lacks the Cu(I) binding motif due to internal sequence deletion of the Met-rich region, but maintains all four Cu(II) ligands. It is therefore not unlikely that, in a hypothetical interaction between CopC and Hyp1, a shift in the redox state may cause the copper ion to migrate between the Cu(II) binding site of CopC to the Met-rich region of Hyp1.

The genomic-context analysis strongly suggests a role of Hyp1 in copper trafficking in prokaryotes.

**Genomic-Context Network.** A genomic-context network was obtained with the program STRING,[22] which integrates information on conserved neighborhood, gene fusion and co-occurrence (Figure 2). The network shows that the orthologous group of Cox11 (COG3175) is densely linked to the subcluster (or functional module) of proteins involved in (c)aa3-type COX assembly, and it appears to have a univocal function in this process. Cox11 is linked, among others, to Surf1 (COG3346). This latter protein is conserved in eukaryotes and mutations in its gene are observed in patients with Leigh syndrome and COX deficiency.[41] Surf1 and Cox11 are both linked to a protein of unknown function (Hyp2, NOG10163), which is found in the COX operon (Figure 1). Similarly to Sco, Hyp2 proteins have a conserved CxxxC motif and a thioredoxin-fold, as predicted by threading methods, but lack the DxxxD motif and the histidine (sequence identity to Sco = 7±3%). Surf1 is also linked to Cox10 and Cox15, which are involved in the synthesis of heme *a* before its insertion into Cox1.[15,42]

Within a genomic-context network orthologous groups that connect separate subclusters tend to be multifunctional and/or to play a role in different processes.[24] The multifunctionality does not necessarily reside in the individual proteins of the orthologous group, but can be achieved by gene duplication, leading to different functional associations and assignment to multiple subclusters.[24] This is the case of Sco (COG1999), Hyp1 (COG2847), and cyt *c* (COG2010), which act as linkers connecting different functional modules. In particular, the subclusters of (c)aa3-type and cbb3-type COXs are connected through a single linker, i.e., cyt *c* (COG2010), while (c)aa3-type COX and the copper uptake subclusters are linked through a two-linker-connection involving the orthologous groups of Sco and Hyp1. Importantly, the linker proteins are more conserved than those in nonlinkers and mutations in their sequences have a significantly higher effect on biological processes.[43]

**3D Structural Models.** Structural models were built for all five Sco paralogs of *Pseudomonas putida* and for the neighbor cyt *c* domains, the models of yeast and human Sco1 and Sco2[26] and of human mitochondrial cyt *c*[44] being available as well as the experimental structure of cyt *c* from yeast.[45] All the three

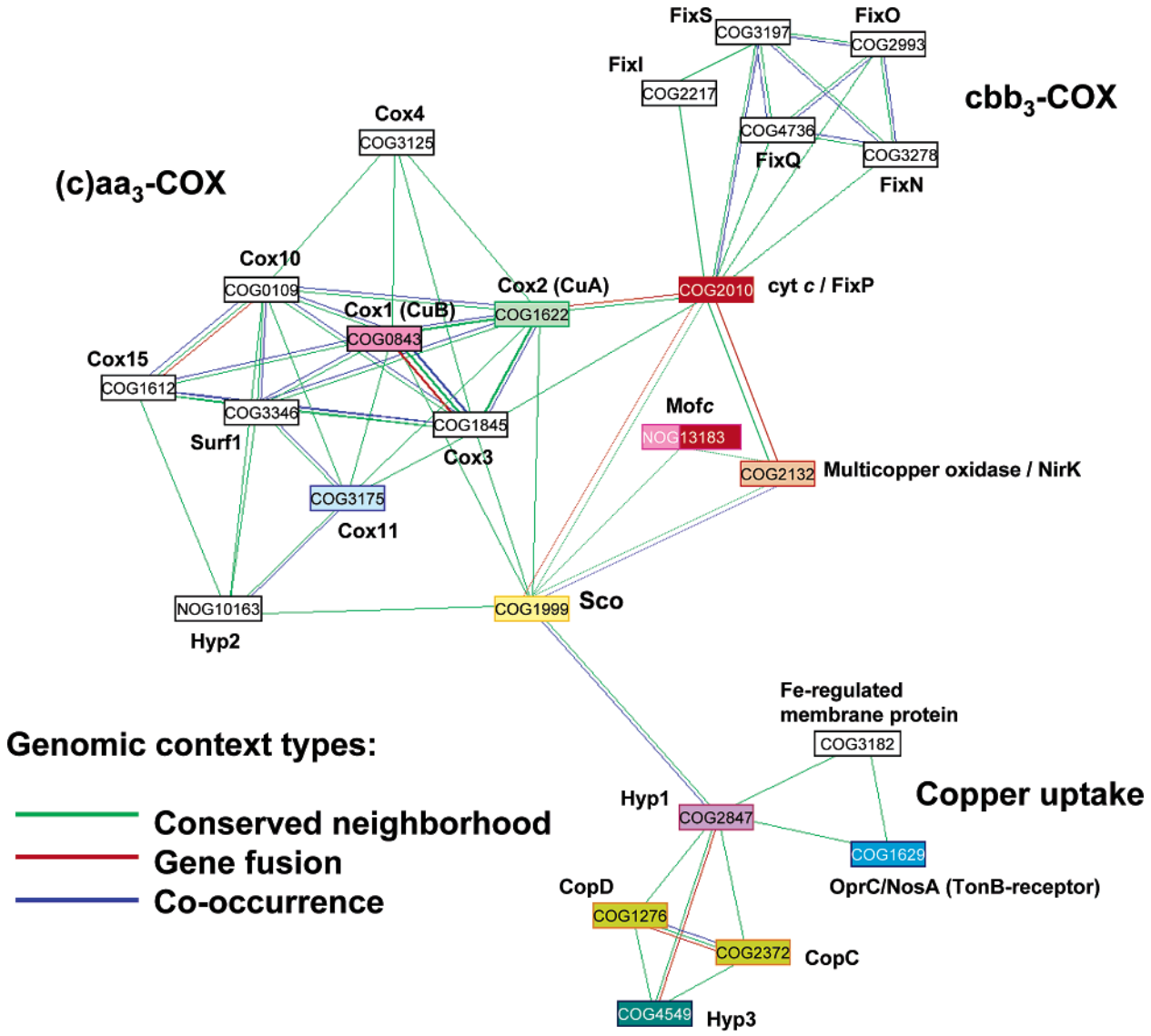



**Figure 2.** Functional modules in a genomic-context network obtained with the program STRING.[22] Shown are the orthologous groups linked via genomic context to COG1999 (Sco) either directly or via another COG (network depth = 2). For each COG the gene name is indicated. Color coding of genes is the same as that reported in Figure 1. The three types of context evidence (gene order, gene fusion and co-occurrence) are indicated by separate lines in the network (full lines correspond to a combined association score > 0.4 and dashed lines to a score <0.4 and >0.3). The three subclusters ((c) aa3-COX, cbb3-COX and copper uptake) are connected to each other through either one orthologous group (COG2010) or one link (between COG1999 and COG2847).

cyt *c* domains present in the genome of *Bacillus subtilis* were also modeled, including the cyt *c* attached to the C-terminus of Cox2. The electrostatic potential surfaces of Sco paralogs and cyt *c* domains of four different organisms, i.e., two eukaryotes, a Gram-negative and a Gram-positive bacterium, are compared in Figure 3. The CxxxC motif is exposed to the solvent in all Sco models but the electrostatic surface in the proximity of the motif is variable. In the human and yeast proteins, as well as in Sco from *Bacillus subtilis*, the region around the CxxxC motif is largely negative and surrounded by a ring of positive charges. On the contrary, among the five Sco paralogs from *Pseudomonas putida*, a large negative patch is present only in the Sco paralog fused to cyt *c* and close to a multicopper oxidase, while in the other Sco paralogs this negative region is reduced, at various extent, due to the presence of scattered neutral or positively charged residues. In particular, the electrostatic potential surface of Sco paralogs encoded by two adjacent genes in *Pseudomonas putida* are remarkably different.

Human and yeast mitochondrial cyt *c* show surface complementarity with both Sco1 and Sco2 from the corresponding organisms in proximity of the exposed heme edge, where a number of lysine and arginine residues form a positive patch. The cyt *c* domain attached to the C-terminus of Cox2 in *Bacillus subtilis* also shows a positive region surrounded by a ring of negative charges on the side of the exposed heme edge. Also in this system the electrostatic surface of cyt *c* is complementary to that of Sco, around the CxxxC motif. In contrast, in the other two cyt *c* domains of *Bacillus subtilis* the face corresponding to the exposed heme edge is largely neutral. Electrostatic surface complementarity is also present between some Sco paralogs and neighbor cyt *c* domains of *Pseudomonas putida*. This is more evident in the case of fused Sco and cyt *c* (see Figure 3).

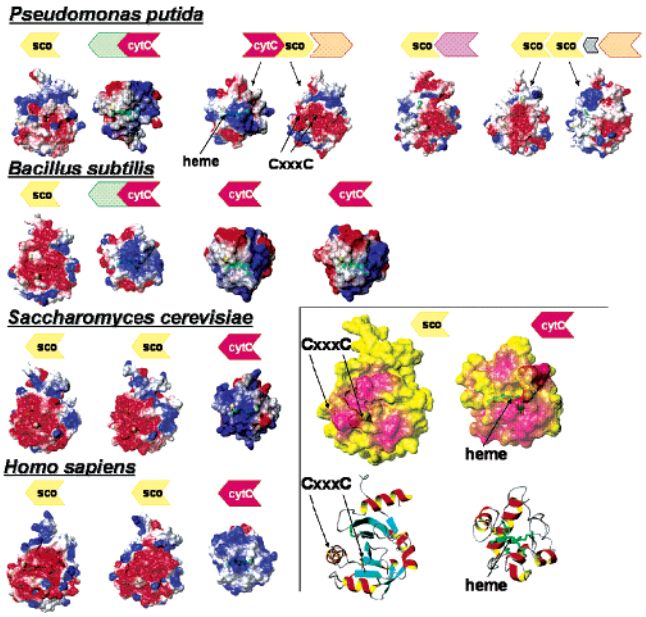



**Figure 3.** Electrostatic surface potential of structural models of Sco1 paralogs and cyt *c* domains of four different organisms. The positively, negatively charged and neutral amino acids are represented in blue, red and white, respectively. The genomic context is indicated with arrows above each protein. Color coding of genes is the same as that reported in Figure 1. The inset in the bottom right shows the mapping of correlated mutations (see text for details) on the surfaces of Sco1 and cyt *c* from *Bacillus subtilis*. Correlated residues are shown in magenta. A yellow surface indicates no correlations. A ribbon representation of the two proteins is also shown. The cysteines of the CxxxC motif of Sco1 and of the CxxC motif of cyt *c* are shown in yellow, the heme cofactor of cyt *c* and the conserved His of Sco1 are shown in green. The orientation of all Sco1 and cyt *c* structures is chosen to allow to see the face where lie the CxxxC motif of Sco1 and the exposed heme edge of cyt *c*, respectively.

**Correlated Mutation Analysis.** Correlated mutation pattern can be defined as the tendency of residues to be conserved or to mutate in tandem between (sets of) sequences. Correlated mutations have been suggested to be related to protein−protein interactions.[30] Therefore, we analyzed them to predict potential interactions of Sco. For this analysis we selected all Sco sequences close to caa3-type COX genes or close to Hyp1 genes. As mentioned above, in caa3-type COX a cyt *c* domain is attached to the C-terminus of the Cox2 subunit, whose N-terminal domain contains the $Cu_A$ center. We analyzed the correlated mutation pattern of the following protein/domain pairs: Sco/cyt *c*, $Cu_A$/cyt *c*, Sco/$Cu_A$ (Figure 2S of Supporting Information) and Sco/Hyp1 (Figure 3S of Supporting Information). On the basis of interaction indexes calculated from correlated mutations (see Methods), one might speculate a possible interaction for the pairs Sco/cyt *c* and Sco1/Hyp1. The values of interaction indexes are 4.2 and 3.9 for the pairs Sco/cyt *c* and Sco/Hyp1, respectively, which are equal or even larger than those found for the $Cu_A$/cyt *c* and Sco/$Cu_A$ pairs (2.8 and 3.8, respectively), for which experimental evidence of interaction was reported.[46,47]

The information from correlated mutation analysis might also suggests an evolutionary compensation between pairs of positions which are possibly in physical proximity. Therefore, it might also highlight potential interaction sites.[30] For the Sco/cyt *c* pair, correlated mutations were mapped on the structure of Sco from *Bacillus subtilis* and of the cyt *c* domain attached to the caa3-type COX from the same organism (see inset of Figure 3). Remarkably, correlated residues are clustered around the CxxxC motif of Sco and around the exposed heme edge of cyt *c*. These are the two regions which also show complementary electrostatic surfaces in the structural models of Sco paralogs and cyt *c* domains. For the Sco/Hyp1 pair the mapping of correlated mutations on the protein structure of Sco indicates that correlated residues are again clustered around the CxxxC motif, whereas in Hyp1 they are close to methionines of the consensus motif, and indicate that this region could be part of the surface interacting with Sco.

For the $Cu_A$/cyt *c* and Sco/$Cu_A$ pairs a direct interaction has been experimentally proved.[46] Our correlated mutation analysis indicates that correlated residues are close to the electron-transfer sites of $Cu_A$ and cyt *c* and to the conserved CxxxC motif of Sco.

## Discussion

**Possible Roles of Cox11 and Sco.** In the absence of direct experimental data, some clues on the properties of Cox11, Sco and related proteins, and their involvement in copper delivery to COX, can be gained from their genomic context, including operon structures and conserved domain fusions in prokaryotes.[32]

Cox11 has been shown to be involved in the insertion of $Cu_B$ in *Rhodobacter sphaeroides*[13] and, indeed, we found Cox11 genes in the aa3-type COX operon in a large number of Gram-negative bacteria. No Cox11 genes were found in the proximity of quinol oxidases, which contain only the $Cu_B$ site and use the lipid-soluble quinol as electron carrier. This indicates that the role of Cox11 is confined to a function in copper delivery to $Cu_B$ in aa3-type COX, whereas alternative copper transport systems could be involved in copper delivery to quinol oxidases. The absence of Cox11 in Gram-positive bacteria and archaea also suggests a different mechanism of copper incorporation in $Cu_B$ in these organisms as well as in some Gram-negative bacteria. For example, it is known that a copper transporting P-type ATPase, called FixI, is involved in copper delivery to the $Cu_B$ site of cbb3-type COX in *Rhizobia*.[48]

The function of Sco is more complex and less univocal. Multiple Sco-like sequences can be found in a single organism. These might be all derived from a gene duplication event and can be considered paralog proteins. In support of this, in some bacteria two Sco genes are found in adjacent positions and in eukaryotes two paralog proteins, Sco1 and Sco2, exist. The variety in number and localization of Sco genes suggests that Sco paralogs can be involved in other functions besides copper insertion into $Cu_A$. Indeed, some Sco genes are neighbors to copper-dependent enzymes others than COX. Furthermore, the variability of the electrostatic surface features of the 3D structural models of Sco paralogs (Figure 3) may suggest their involvement in different partnerships.

**Connection between Redox and Copper Homeostasis.** In this analysis, a correlation emerges between Sco and cyt *c*. Cyt *c*, which acts as an electron donor for the $Cu_A$ center in Cox2 and for other copper-dependent enzymes. The structure of Sco has a thioredoxin fold and a thiol-disulfide oxidoreductase function has been proposed for this protein.[26] Therefore, Sco might be involved, in addition to its well documented role in copper transfer to $Cu_A$,[4] in the reduction of disulfide bonds of $Cu_A$ prior to copper insertion. Also, the two cysteines of the CxxxC motif of Sco should remain in a reduced state, despite

the oxidizing environment, where it is located. Thioredoxins are usually associated to transmembrane electron transporters involved in a disulfide bond reduction cascade which carries electrons (reducing equivalents) from the cytoplasm to the periplasm, which is a strongly oxidizing environment.[49] As we did not find any correlation between Sco and transmembrane electron transporters, a candidate molecule for reducing the disulfide of the CxxxC motif might be cyt *c*. In eukaryotes, cyt *c* is required not only for electron transfer but also for COX assembly through a still unknown mechanism. Indeed, in mitochondria lacking the folded and mature (heme-containing) form of cyt *c*, the COX subunits are not properly assembled.[50,51] It is therefore likely that cyt *c* and Sco are jointly involved in enzyme assembly.

Some Sco paralogs in bacteria and archaea are related, both in terms of gene neighborhood and phylogenetic occurrence, to Hyp1 (COG2847), a conserved bacterial protein which possesses a Met-rich motif potentially involved in copper binding. Several periplasmic proteins involved in copper homeostasis have been identified so far (e.g., CopC and CueO), which adopt a Cu(I)-thioether coordination chemistry involving Met-rich motifs.[9] The periplasm is the cell compartment of Gram-negative bacteria where COX and other copper enzymes acquire their metal cofactors. Under normal conditions, the copper concentration in the periplasm is not limiting.[9] However, under conditions of low copper supply, a more efficient copper uptake mechanism might be activated. Hyp1 might be expressed only in oxygen and copper limiting conditions, as it has been demonstrated for one of its neighbor genes, NosA, which encodes an outer membrane channel responsible for copper uptake in the periplasm.[38,40] As there are no prokaryotic homologues of the mitochondrial copper chaperone Cox17, is proposed that Hyp1 takes the role of Cox17 in bacteria and archaea, probably trafficking Cu(I) ions only in copper starving conditions and/or anaerobiosis. Hyp1 has been expressed and structurally characterized in our lab and its metal binding properties have been determined, demonstrating that Hyp1 is indeed a Cu(I) binding protein (Banci L., Bertini I., Ciofi-Baffoni, S., Katsari E., Kubicek K., manuscript in preparation). In the genomic-context network, we identified Sco and Hyp1 as linkers, which are responsible for networking two functional modules: one involved in (c)aa3-type COX assembly and one in copper uptake (see Figure 2). Therefore, the association between Sco and Hyp1 may represent the event at the origin of copper insertion into subunit II of COX.

A potential interaction of Sco with Hyp1 and cyt *c* is consistent with the values of the interaction index calculated from correlated mutations. In addition, 3D structural models of Sco paralogs and neighbor cyt *c* domains from different organisms highlight an electrostatic surface complementarity between the two proteins around their active sites, that are the regions where correlated mutations are also clustered (Figure 3).

From the present analysis of protein sequences and of their genomic context, we can summarize that Sco could play a role in delivering copper to a variety of different enzymes whose activity depends on copper, and/or to play a role as a thioreductase, keeping the residues that coordinate copper in those enzymes in a reduced state for copper delivery.

Thanks to its dual nature and the paralogs, in response to variations in oxygen levels and copper availability occurred during the evolution,[52] Sco may have led to the incorporation of a copper cofactor into subunit II of COX, possibly being recruited among thioredoxins responsible for heme incorporation into apocytochromes.[53] In support of this, a possible evolutionary relationship exist between subunit II of caa3-type COX and the dihaemic subunit (called FixP) of cbb3-type COX.[7] This is based on the observation that the cyt *c* domain fused at the C-terminus of Cox2 is quite similar to the second cyt *c* domain of FixP. Furthermore, the binding site of the first c-type heme of FixP seems to be part of the binding site of $Cu_A$: the axial ligands of the heme, the histidine and the methionine, are ligands of the copper center.[7]

## Conclusions

**From Bacteria to Mitochondria.** It is well established that mitochondria originated from ancient invasion of Gram-negative α-proteobacteria (endosymbiont) into an archaea-type or an eukaryotic host. In losing their autonomy, endosymbionts elaborated mechanisms for organelle biogenesis and metabolite exchange, thus acquiring many host-derived properties.[54] A fundamental step in this process was the adaptation to the metal uptake mechanisms of the host. As a consequence the functional link between Sco and the potential copper transporter Hyp1 was lost and a new link was established between Sco and the mitochondrial copper chaperone Cox17, this latter protein being only present in eukaryotes, including the protists *Plasmodium falciparum* and *Chlamidomonas reinhardtii*. On the other hand, Hyp1 is conserved among the α-proteobacteria (see Figure 1S of Supporting Information) and occurs in many other prokaryotes, including the *Neisseria* and *Vibrio* pathogens, but it lacks a homologue among the eukaryotes. Therefore Hyp1 represents a potential drug target.

In the lifestyle transition from free-living versus obligate intracellular, the adaptation to the copper transport and metabolism of the host was possible thanks to the pivotal role of the multifunctional protein Sco, while keeping unvaried the core components of the COX assembly module inherited from Gram-negative α-proteobacteria (see Figure 2). This fits with the observation that most of the ancestral bacterial genes present in the mitochondrial genome are involved in bioenergetic and translational processes, while novel genes recruited from the host nuclear genome are primarily involved in transport and regulatory functions.[55,56]

Gene duplications are regarded as an efficient engine that enables rapid responses to alterations in the environmental conditions.[57] After gene duplication, multiple partnerships of a single ancestral gene may become separately allocated among paralogs by acquisition of functional specialization. The presence of two Sco paralogs in eukaryotes (i.e., Sco1 and Sco2) can be rationalized in the light of the genomic-context analysis of prokaryotic paralogs: one Sco gene may preferentially interact with subunit II of COX thus favoring COX assembly, while the second gene may assist the metallochaperone Cox17 which is responsible for copper recruitment in the inter membrane space of mitochondria. In this scenario, it is possible that the two eukaryotic Sco paralogs interact to promote copper insertion into $Cu_A$. Similar conclusions for human Sco genes were reached using an experimental approach.[58]

To conclude, we found that Cox11 is highly networked within the COX assembly module and likely fulfills a univocal function in COX assembly, while Sco represents a multifunctional linker or adaptor which allows the COX enzyme to interface with alternative redox and copper metabolisms.

**Acknowledgment.** This work has been supported by the European Commission (contracts HPRI-CT-2001-50028 and

QLG2-CT-2002-00988). The Italian MURST COFIN03 is acknowledged for financing.

**Supporting Information Available:** Two tables reporting a list of Cox11- and Sco-like genes identified through the BLAST searches. One figure with results of gene neighnorhood analysis showing operon and divergon structures of genes encoding Cox11, Sco and their neighbors. Two figures showing the correlated mutation pattern of the domain pairs Sco/cyt *c*, $Cu_A$/cyt *c*, Sco/$Cu_A$, and of Sco/Hyp1. This material is available free of charge at http://pubs.acs.org.

## References


(1) Iwata, S.; Ostermeier, C.; Ludwig, B.; Michel, H. *Nature* **1995**, *376*, 660−669.
(2) Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yoshikawa, S. *Science* **1995**, *269*, 1069−1074.
(3) Poyton, R. O. *Nat. Genetic.* **1998**, *20*, 316−317.
(4) Carr, H. S.; Winge, D. R. *Acc. Chem. Res.* **2003**, *36*(5), 309−316.
(5) Ramirez, B. E.; Malmström, B. G.; Winkler, J. R.; Gray, H. B. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 11949−11951.
(6) Garcia-Horsman, J. A.; Barquera, B.; Rumbley, J.; Ma, J.; Gennis, R. B. *J. Bacteriol.* **1994**, *176*, 5587−5600.
(7) Pereira, M. M.; Santana, M.; Teixeira, M. *Biochim. Biophys. Acta* **2001**, *1505*(2−3), 185−208.
(8) O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057−25060.
(9) Finney, L. A.; O'Halloran, T. V. *Science* **2003**, *300*, 931−936.
(10) Glerum, D. M.; Shtanko, A.; Tzagoloff, A. *J. Biol. Chem.* **1996**, *271*, 14504−14509.
(11) Heaton, D.; Nittis, T.; Srinivasan, C.; Winge, D. R. *J. Biol. Chem.* **2000**, *275*, 37582−37587.
(12) Glerum, D. M.; Shtanko, A.; Tzagoloff, A. *J. Biol. Chem.* **1996**, *271*, 20531−20535.
(13) Hiser, L.; Di Valentin, M.; Hamer, A. G.; Hosler, J. P. *J. Biol. Chem.* **2000**, *275*, 619−623.
(14) Shoubridge, E. A. *Hum. Mol. Genet.* **2001**, *10*, 2277−2284.
(15) Barrientos, A.; Barros, M. H.; Valnot, I.; Rotig, A.; Rustin, P.; Tzagoloff, A. *Gene* **2002**, *286*, 53−63.
(16) Karlberg, E. O.; Andersson, S. G. *Nat. Rev. Genet.* **2003**, *4*, 391−397.
(17) Wallace, D. C. *Science* **1999**, *283*, 1482−1488.
(18) Huynen, M.; Snel, B.; Lathe, W.; Bork, P. *Curr. Opin. Struct. Biol.* **2000**, *10*, 366−370.
(19) Altschul, S. F.; Madden, T. L.; Schaeffer, A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J. *Nucl. Acids Res.* **1997**, *25*(17), 3389−3402.
(20) Zhang, Z.; Schaffer, A. A.; Miller, W.; Madden, T. L.; Lipman, D. J.; Koonin, E. V.; Altschul, S. F. *Nucl. Acids Res.* **1998**, *26*(17), 3986−3990.
(21) Thompson, J. D.; Higgins, D. G.; Gibson, T. J. *Nucl. Acids Res.* **1994**, *22*(22), 4673−4680.
(22) von Mering, C.; Huynen, M.; Jaeggi, D.; Schmidt, S.; Bork, P.; Snel, B. *Nucl. Acids Res.* **2003**, *31*, 258−261.
(23) Tatusov, R. L.; Natale, D. A.; Garkavtsev, I. V.; Tatusova, T. A.; Shankavaram, U. T.; Rao, B. S.; Kiryutin, B.; Galperin, M. Y.; Fedorova, R. D.; Koonin, E. V. *Nucl. Acids Res.* **2001**, *29*, 22−28.
(24) Snel, B.; Bork, P.; Huynen, M. A. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 5890−5895.
(25) Sali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*(3), 779−815.
(26) Balatri, E.; Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S. *Structure* **2003**, *11*, 1431−1443.
(27) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graph.* **1996**, *14*, 51−55.
(28) Pazos, F.; Valencia, A. *Proteins* **2002**, *47*, 219−227.
(29) Pazos, F.; Olmea, O.; Valencia, A. *CABIOS* **1997**, *13*, 319−321.
(30) Pazos, F.; Helmer-Citterich, M.; Ausiello, G.; Valencia, A. *J. Mol. Biol.* **1997**, *271*, 511−523.
(31) Marcotte, E. M. *Curr. Opin. Struct. Biol.* **2000**, *10*, 359−365.
(32) Galperin, M. Y.; Koonin, E. V. *Nat. Biotechnol.* **2000**, *18*, 609−613.
(33) Huynen, M.; Snel, B.; von Mering, C.; Bork, P. *Curr. Opin. Cell Biol.* **2003**, *2*(15), 191−198.
(34) De Vrind, J.; De Groot, A.; Brouwers, G. J.; Tommassen, J.; De Vrind-De Jong, E. *Mol. Microbiol.* **2003**, *47*, 993−1006.
(35) Cuypers, H.; Zumft, W. G. *J. Bacteriol.* **1993**, *175*, 7236−7346.
(36) Seib, K. L.; Jennings, M. P.; McEwan, A. G. *FEBS Lett.* **2003**, *546*, 411−415.
(37) Arnesano, F.; Banci, L.; Bertini, I.; Mangani, S.; Thompsett, A. R. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 3814−3819.
(38) Yoneyama, H.; Nakae, T. *Microbiology* **1996**, *142*, 2137−2144.
(39) Ferguson, A. D.; Deisenhofer, J. *Cell (Cambridge, Mass)* **2004**, *116*, 15−24.
(40) Wunsch, P.; Herb, M.; Wieland, H.; Schiek, U. M.; Zumft, W. G. *J. Bacteriol.* **2003**, *185*, 887−896.
(41) Zhu, Z.; Yao, J.; Johns, T.; Fu, K.; De Bie, I.; Macmillan: C.; Cuthbert, A. P.; Newbold, R. F.; Wang, J.; Chevrette, M.; Brown, G. K.; Brown, R. M.; Shoubridge, E. A. *Nat. Genet.* **1998**, *20*, 337−343.
(42) Antonicka, H.; Leary, S. C.; Guercin, G. H.; Agar, J. N.; Horvath, R.; Kennaway, N. G.; Harding, C. O.; Jaksch, M.; Shoubridge, E. A. *Hum. Mol. Genet.* **2003**, *12*, 2693−2702.
(43) Winzeler, E. A.; Shoemaker, D. D.; Astromoff, A.; Liang, H.; Anderson, K.; Andre, B.; Bangham, R.; Benito, R.; Boeke, J. D.; Bussey, H.; Chu, A. M.; Connelly, C.; Davis, K.; Dietrich, F.; Dow, S. W.; El Bakkoury, M.; Foury, F.; Friend, S. H.; Gentalen, E.; Giaever, G.; Hegemann, J. H.; Jones, T.; Laub, M.; Liao, H.; Davis, R. W. *Science* **1999**, *285*, 901−906.
(44) Banci, L.; Bertini, I.; Rosato, A.; Varani, G. *J. Biol. Inorg. Chem.* **1999**, *4*, 824−837.
(45) Banci, L.; Bertini, I.; Bren, K. L.; Gray, H. B.; Sompornpisut, P.; Turano, P. *Biochemistry* **1997**, *36*, 8992−9001.
(46) Zhen, Y.; Hoganson, C. W.; Babcock, G. T.; Ferguson-Miller, S. *J. Biol. Chem.* **1999**, *274*, 38032−38041.
(47) Lode, A.; Kuschel, M.; Paret, C.; Rodel, G. *FEBS Lett.* **2000**, *485*(1), 19−24.
(48) Koch, H. G.; Winterstein, C.; Saribas, A. S.; Alben, J. O.; Daldal, F. *J. Mol. Biol.* **2000**, *297*, 49−65.
(49) Katzen, F.; Beckwith, J. *Cell* **2000**, *103*, 769−779.
(50) Pearce, D. A.; Sherman, F. *J. Biol. Chem.* **1995**, *270*, 20879−20882.
(51) Barrientos, A.; Pierre, D.; Lee, J.; Tzagoloff, A. *J. Biol. Chem.* **2003**, *278*, 8881−8887.
(52) Frausto da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of the Elements: The Inorganic Chemistry of Life*; University Press: New York, Oxford; 2001.
(53) Thony-Meyer, L. *Microbiol. Mol. Biol. Rev.* **1997**, *61*, 337−376.
(54) Dyall, S. D.; Brown, M. T.; Johnson, P. J. *Science* **2004**, *304*(5668), 253−257.
(55) Karlberg, O.; Canback, B.; Kurland, C. G.; Andersson, S. G. *Yeast* **2000**, *17*, 170−187.
(56) Gabaldon, T.; Huynen, M. A. *Science* **2003**, *301*, 609.
(57) Boussau, B.; Karlberg, E. O.; Frank, A. C.; Legault, B. A.; Andersson, S. G. *Proc. Natl. Acad. Sci. U.S.A.* **2004**.
(58) Leary, S. C.; Kaufman, B. A.; Pellecchia, G.; Guercin, G. H.; Mattman, A.; Jaksch, M.; Shoubridge, E. A. *Hum. Mol. Genet.* **2004**, *13*, 1839−1848.


PR049862F

# 3.2

## A hint for the function of human Sco1 from different structures

Lucia Banci, Ivano Bertini, Vito Calderone, Simone Ciofi-Baffoni, Stefano Mangani, Manuele Martinelli, Peep Palumaa, and Shenlin Wang

*Proc Natl Acad Sci U S A.* (2006) **103**, 8595-600.

# A hint for the function of human Sco1 from different structures

**Lucia Banci*, Ivano Bertini*†, Vito Calderone*, Simone Ciofi-Baffoni*, Stefano Mangani*‡, Manuele Martinelli*, Peep Palumaa§, and Shenlin Wang***

*Magnetic Resonance Center and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Florence, Italy; ‡Department of Chemistry, Università degli Studi di Siena, Via Aldo Moro 1, 53100 Siena, Italy; and §Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia

Edited by Joan Selverstone Valentine, University of California, Los Angeles, CA, and approved April 12, 2006 (received for review February 17, 2006)

**The solution structures of apo, Cu(I), and Ni(II) human Sco1 have been determined. The protein passes from an open and conformationally mobile state to a closed and rigid conformation upon metal binding as shown by electrospray ionization MS and NMR data. The metal ligands of Cu(I) are two Cys residues of the CPXXCP motif and a His residue. The latter is suitably located to coordinate the metal anchored by the two Cys residues. The coordination sphere of Ni(II) in solution is completed by another ligand, possibly Asp. Crystals of the Ni(II) derivative were also obtained with the Ni(II) ion bound to the same His residue and to the two oxidized Cys residues of the CPXXCP motif. We propose that the various structures solved here represent the various states of the protein in its functional cycle and that the metal can be bound to the oxidized protein at a certain stage. Although it now seems reasonable that Sco1, which is characterized by a thioredoxin fold, has evolved to bind a metal atom via the di-Cys motif to act as a copper chaperone, the oxidized form of the nickel-bound protein suggests that it may also maintain the thioredoxin function.**

cytochrome *c* oxidase | NMR | x-ray | assembly factor

Sco is a family of proteins ubiquitous to all kingdoms of life. Ortholog and paralog genome browsing has shown that one or more representative of this class are present in most bacterial and eukaryotic genomes (1, 2). In the bacterial operons, Sco proteins often are associated with copper enzymes, suggesting that they are involved in the maturation or functioning of such enzymes (1). Eukaryotic genomes contain two paralogs, Sco1 and Sco2 (3, 4), both involved in copper-dependent assembly of cytochrome *c* oxidase (CcO) (5). CcO contains two functional copper ions located in the binuclear $Cu_A$ site and one located in the binuclear $Cu_B$-heme $a_3$ site (6). CcO is a multimeric enzyme complex embedded in the inner mitochondrial membrane of all eukaryotes and in the plasma membrane of prokaryotes, and it functions as terminal enzyme of the respiratory chain (7). Sco1 was first suggested to be involved in copper ion delivery to the CcO complex based on the observation that, in the presence of high copper concentrations, overexpression of either Sco1 or the homologous Sco2 can restore the CcO activity of *Saccharomyces cerevisiae* strains lacking the gene of the mitochondrial copper chaperone Cox17 (8). Yeast strains lacking Sco1 also are respiratory deficient, and an excess of copper and/or overexpression of either Cox17 or Sco2 cannot compensate for the Sco1-associated CcO deficiency (8). The absolute requirement of Sco1 in the activation of CcO indicates that Sco1 functions downstream from Cox17 in the delivery of copper to CcO. Indeed, *in vitro* Sco1 can receive copper from the copper chaperone Cox17 (9).

Human Sco1 (HSco1) is a 301-residue polypeptide anchored through a single helix to the inner mitochondrial membrane of eukaryotes (10). The functional part of the Sco1 protein is composed of a single soluble domain, located in C-terminal region, whereas the N terminus contains a mitochondrial-targeting sequence followed by a transmembrane helix (11). The structure of the soluble domain was first resolved by NMR for apoSco1 from *Bacillus subtilis* (12). The structure revealed a potential metal binding site constituted by two Cys residues present in a conserved motif CPXXCP and a fully conserved His residue, in agreement with earlier extended x-ray absorption fine structure investigations of yeast Cu(I)Sco1 (13). Besides extended x-ray absorption fine structure data many other spectroscopic data on the human, yeast, and *B. subtilis* Sco proteins have confirmed that the protein binds Cu(I), Cu(II), and other metal ions (14–17). The fold of apoSco1 protein, which contains four $\alpha$-helices and seven $\beta$-strands organized in two $\beta$-sheets (12), is atypical for a metal chaperone given that it resembles the fold of thioredoxins, which are enzymes specialized for the reduction of protein disulfides (18). Also, the x-ray structures of the apo forms of the *B. subtilis* Sco1 and HSco1 proteins have been determined (16, 19). Some crystals of *B. subtilis* apoSco1 contained Sco1 with S—S bonds, thus supporting a role of the protein in redox processes. In addition, it has been recently suggested on the basis of the extreme sensitivity of the yeast *sco1*-null mutant to hydrogen peroxide that HSco1 might function as a mitochondrial redox signaling molecule (19). Thus, the specific role of Sco1 in maturation of CcO, either as a copper chaperone or connected with redox processes, is still elusive.

To date there are no structures available for any metallated forms of Sco1 proteins. The structure of the metal adducts is, however, crucial for understanding the mechanism of Sco-mediated copper insertion into CcO. We have succeeded in preparing a human Cu(I)Sco1 derivative [hereafter referred to as Cu(I)HSco1] *in vitro* and have determined its solution structure through NMR. We also have investigated the structure of the Ni(II) derivative of HSco1 as a model for the binding of bipositive metal ions like Cu(II). We would like to stress that NMR solution structure determination of metalloproteins is a difficult challenge as far as the protein metal-binding mode is concerned because NMR does not provide direct information on protein–metal interaction. In contrast, it is also common that the crystallization procedures might provide derivatives different than the physiological ones. Still, these derivatives may be significant as models of transient species and for the mechanism of action. Indeed, the x-ray structure of human Ni(II)Sco1 [hereafter referred to as Ni(II)HSco1], also reported here, displays a completely different metal binding with respect to the solution structure, being the metal bound to the S—S bond of the oxidized HSco1. Structural information for different conformers of the same protein is extremely valuable because individual structures often mimic transient species, which enable elucidation of the mechanism of protein action.

Conflict of interest statement: No conflicts declared.

This paper was submitted directly (Track II) to the PNAS office.

Abbreviaions: CcO, cytochrome *c* oxidase; ESI, electrospray ionization; HSco1, human Sco1; HSQC, heteronuclear single quantum correlation.

Data deposition: The atomic coordinates, structural restraints, chemical shifts, and structural factors have been deposited in the Protein Data Bank, www.pdb.org [PDB ID codes 2GT5 and 2GVP for apoSco1; 2GQM and 2GT6 for Cu(I)Sco1; and 2GQK, 2GQL, and 2GGT for Ni(II)Sco1].

†To whom correspondence should be addressed. E-mail: bertini@cerm.unifi.it.

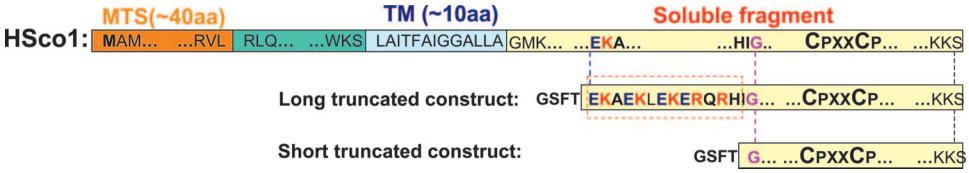



**Fig. 1.** Diagram of the protein sequence and cloned constructs of HSco1. The positions of the mitochondrial target sequence (MTS), transmembrane helix (TM), and the soluble fragment of HSco1 are orange, blue, and yellow, respectively. The N terminus of HSco1 protruding into the mitochondrial matrix is green. The soluble fragments used in the present study are named long and short truncated constructs. The truncates lack only the mitochondrial target sequence and transmembrane helix. The essential CPXXCP motif, the tobacco etch virus protease recognition site (GSFT), and the 14-aa segment (red dotted box) at the N terminus side are indicated in bold. The positively and negatively charged amino acids are represented in blue and red, respectively.

Here, we propose that the Cu(I)HSco1 species act as chaperones for copper ion delivery and that the specific oxidation of metal-ligating thiolates of HSco1 could have a role in the transfer of copper ions to CcO.

## Results

Two truncated forms of the HSco1 gene have been engineered, both lacking the N-terminal mitochondrial targeting sequence and the single-transmembrane helix (Fig. 1). The two truncated forms differ by a 14-aa segment at the N-terminal side. This 14-aa segment is rich in positively and negatively charged residues and is predicted to be unstructured. Size-exclusion chromatography equipped with multiangle light scattering shows that the shorter apoHSco1 construct eluted in fractions corresponding to a monomeric state for the protein, whereas the longer construct eluted as a dimer (see Fig. 6, which is published as supporting information on the PNAS web site). This observation suggests that the residues separating the transmembrane helix from the folded domain are essential to promote dimerization. *In vivo* studies also have shown that this N-terminal region is crucial for yeast Sco1 function and cannot be replaced, even by its Sco2 counterpart (20).

For the structural and electrospray ionization (ESI)-MS characterization, the short construct of 170 aa (corresponding to residue segment 132–301) of the HSco1 gene was used. Additionally, this construct contains an additional 4-aa segment (GSFT), corresponding to the tobacco etch virus protease recognition site at the N terminus (Fig. 1), thus producing a final construct of 174 aa whose identity was confirmed by ESI-MS analysis. The ESI-MS spectrum of the oxidized apoWT-HSco1 showed two main peaks corresponding to +9 and +10 ions, and the deconvolution of the spectrum gave the expected molecular mass of 19,741.3 Da. Incubation of oxidized HSco1 with 1 mM DTT increased molecular mass by 2.3 Da (molecular mass = 19,743.9 Da), which indicates that the disulfide bridge in HSco1 could be reduced easily by 1 mM DTT.

**Reconstitution of Fully Reduced HSco1 with Metal Ions.** We are able to prepare HSco1 containing Cu(I) ions at 1:1, a ratio suitable for structural studies. Mass spectra of fully reduced HSco1 reconstituted at pH 7.5 with increasing concentrations of Cu(I) ions, are presented in Fig. 7, which is published as supporting information on the PNAS web site. Addition of one equivalent of Cu(I) ions to reduced HSco1 generates a new peak in MS spectrum, which corresponds to $Cu_1$HSco1 (Fig. 7). Addition of further equivalents of Cu(I) ions leads to a slight additional increase of the $Cu_1$HSco1 peak but does not induce metalloforms with higher metal stoichiometry (Fig. 7). We also showed that oxidized HSco1 does not form complex with Cu(II) ions in an ESI-MS experiment (Fig. 7), which indicates that such a complex, even if present in solution, is weak and dissociates during the ESI process. Accordingly, addition of Ni(II)$Cl_2$ to oxidized $^{15}$N HSco1 does not affect the $^1$H–$^{15}$N heteronuclear single quantum correlation (HSQC) spectrum. However, we have prepared the Cu(II) and Ni(II)HSco1 derivatives of reduced HSco1 in a 1:1 metal/protein ratio. The ultraviolet/visible (UV/VIS) and EPR spectra of Cu(II)HSco1 are identical to those recently reported (15). Similar UV/VIS spectra also were reported for bacterial Sco homologues (12, 14, 17) and Cu(II) nitrosocyanin (21, 22). In addition, the UV/VIS spectrum of Ni(II)HSco1 (Fig. 8, which is published as supporting information on the PNAS web site) is similar to that reported for the *B. subtilis* and *Rhodobacter sphaeroides* homologues (14, 17), with two weak bands at 380 and 540 nm, respectively, correlating with the two low-energy bands of the Cu(II)HSco1 spectrum although shifted to lower energy (Fig. 8). The intense thiolate-Ni(II) charge transfer band (23) is also red-shifted at 304 nm with respect to the Cu(II)HSco1 spectrum (Fig. 8).

The aggregation state and the conformational properties of the Cu(I) and Ni(II)HSco1 forms have been investigated by multiple techniques, including size-exclusion chromatography, ESI-MS, and NMR. In NMR, $^{15}$N relaxation rates are modulated by the correlation time for the protein tumbling ($\tau_m$), which is directly related to the molecular weight of the protein, thus monitoring its aggregation state (24). The correlation times of Cu(I)HSco1 and Ni(II)HSco1 proteins (at millimolar concentrations) are 14.5 ± 1.1 ns and 15.6 ± 1.2, respectively, as expected for a protein of this size in a monomeric state. These values also are similar to the $\tau_m$ of apoHSco1 (13.8 ± 1.6 ns), which reorients in solution as a monomeric protein as shown by size-exclusion chromatography and multiangle light scattering experiments. ESI-MS experiments conducted at a 1.8 $\mu$M concentration of protein did not detect any higher aggregates for apoHSco1 or $Cu_1$HSco1. ESI-MS spectra and, especially, the charge-state distribution of ions also can yield information about the conformational states of proteins under a variety of conditions (25). Upon the binding of Cu(I), the charge state distribution of $Cu_1$HSco1 species shifted toward ions with lower charges (+9 and +8), which indicates that binding of metal induces a conformation change of the protein to a more compact state (Fig. 7).

**Cu(I), Ni(II), and apoHSco1 Solution Structures.** The solution structures of Cu(I)HSco1 and Ni(II)HSco1 (Fig. 2) were determined by using distance and angle restraints as obtained from 2D and 3D heteronuclear NMR experiments (Table 1, which is published as supporting information on the PNAS web site). The overall fold of Cu(I)HSco1 and Ni(II)HSco1 structures is the same as that of the crystal structure of apoHSco1 (19) and contains four $\alpha$-helices and nine $\beta$-strands organized into the thioredoxin fold. The solution structure of apoHSco1 also displays the same global thioredoxin fold. However, the $\beta$-hairpin present in the extended, solvent-exposed loop-8 region does not form anymore (Fig. 2).

Cu(I) is coordinated by the two Cys residues of the CPXXCP conserved motif, shared by the third loop and helix $\alpha$1, and by the conserved His-260 (Fig. 2), located in the seventh $\beta$-strand, as shown by $^2J_{NH}$ coupling-based $^1$H–$^{15}$N HSQC experiments

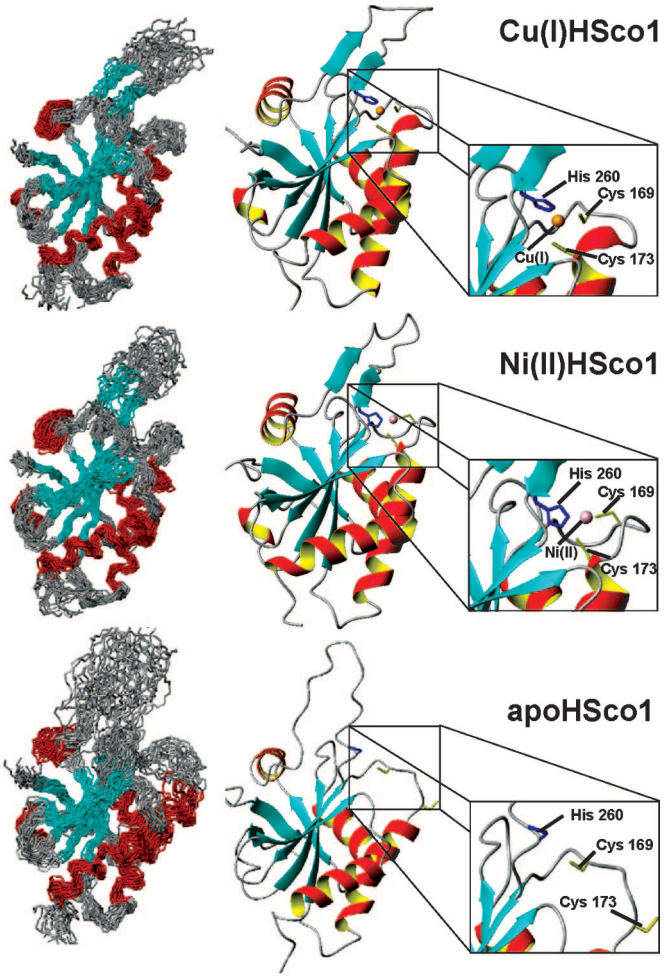



**Fig. 2.** Solution structures of human Cu(I), Ni(II), and apoSco1. (*Left*) The superimposition of 20 structures of Cu(I), Ni(II) and apoHSco1 are shown. $\alpha$-helices and $\beta$-strands are colored in red and cyan, respectively. (*Right*) The average structures of the lowest energy ensemble are shown. The metal-binding residues Cys-169, Cys-173, and His-260 are shown in yellow and blue, respectively. Cu(I) and Ni(II) ions are depicted as orange and pink spheres, respectively.

(Fig. 9, which is published as supporting information on the PNAS web site). From these experiments, it appears that, in Cu(I)HSco1, His-260 acquires a preferential conformation where $N^{\delta 1}$ is protonated and $N^{\varepsilon 2}$ is coordinated to the metal ion. The involvement in the metal binding of residues from two different regions of the protein produces a compact state of protein with respect to the apo form, in agreement with observations in ESI-MS experiments. Backbone NH resonances of three regions comprising residues 166–180, 202–204, and 244–264 are indeed not detected in the $^{1}H$–$^{15}N$ HSQC spectrum of apoHSco1, although they are present in both Cu(I) and Ni(II)HSco1 spectra. These three regions comprise the CPXXCP metal-binding motif and the surrounding loops 5 and 8, the latter containing the third ligand, His-260. The inability to detect the backbone NH signals listed above is because of their fast exchange with the bulk solvent or because of the presence of multiple backbone conformations in the metal-binding area of apoHSco1, whereas the metal binding is able to "freeze" these regions in a more rigid conformation. In particular, the large conformational variability of the long loop 8 observed in the apoHSco1 solution structure (no long-range nuclear Overhauser effects are detected in loop 8) indicates that backbone structural

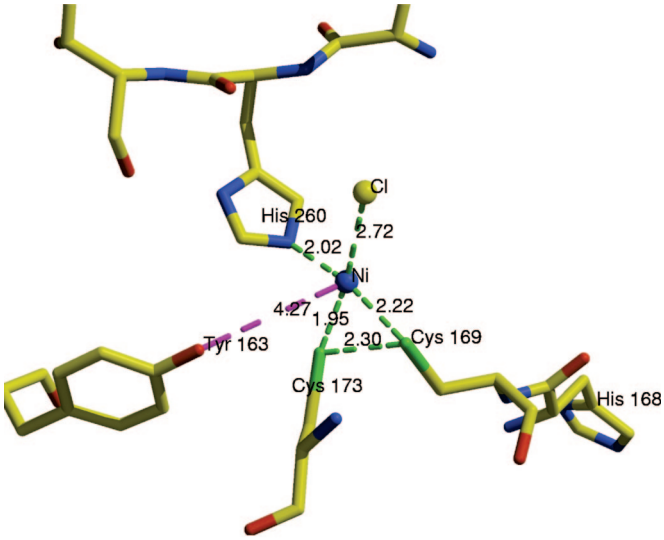



**Fig. 3.** The distorted square planar coordination of nickel as present in the x-ray structure of oxidized Ni(II)HSco1. Bonding distances between nickel and the coordinating moieties are shown in green along with the distance between the two Cys residues. The distance between the nickel ion and the OH oxygen of Tyr-163 is also shown in magenta, and it is too large to be considered a bonding interaction.

changes are necessary to locate the metal ligand His-260 in the vicinity of the other two ligands, Cys-169 and Cys-173 (Fig. 2). Analysis of backbone dynamics (Fig. 10, which is published as supporting information on the PNAS web site) also is in agreement with the latter picture and demonstrates that both metallated forms of HSco1 do not display extensive motions on both milli- to microsecond and/or nano- to picosecond time scales, with the exception of the C and N termini.

When reduced, HSco1 binds a divalent cation, as Ni(II), and the metal is still bound by the two Cys residues of the CPXXCP motif and by His-260 through the $N^{\varepsilon 2}$ atom (Fig. 2), as confirmed by $^{2}J$ $^{1}H$–$^{15}N$ HSQC NMR experiments. Because Ni(II) is expected to be at least four-coordinated, it is feasible that a forth ligand is completing its coordination sphere. This ligand could be a water molecule or a residue donated by the protein. The solution structure of Ni(II)HSco1 shows that two acidic groups (Asp-171 and Asp-259) could complete the Ni(II) coordination sphere, but our data do not allow discrimination between these two possibilities.

**Ni(II)HSco1 Crystal Structure.** Crystals of the Ni(II)HSco1 derivative also were obtained in aerobic conditions. The overall structure of the Ni(II)HSco1 complex is essentially superimposable with that of apoHSco1 [Protein Data Bank (PDB) ID code 1WP0] (19), which was used as the model in the molecular replacement. The main exception is the solvent exposed a region involving residues 240–260 (loop 8 and the $\beta$-hairpin), ending with the metal-binding His-260. This loop indeed acquires a more ordered conformation as a consequence of the metal binding, according to the behavior in solution; this greater stability is confirmed by the good quality of the electron density map in that region for both molecules in the asymmetric unit, which is indeed better than that observed in the apoHSco1 crystal structure (19). A further confirmation is the significantly lower temperature factors of the atoms belonging to the above-mentioned loop in the structure of Ni(II)HSco1 with respect to those of the crystal structure of apoHSco1 (19).

The coordination sphere of Ni(II) in the crystal structure of Ni(II)HSco1 is quite odd and unexpected. In this structure, the

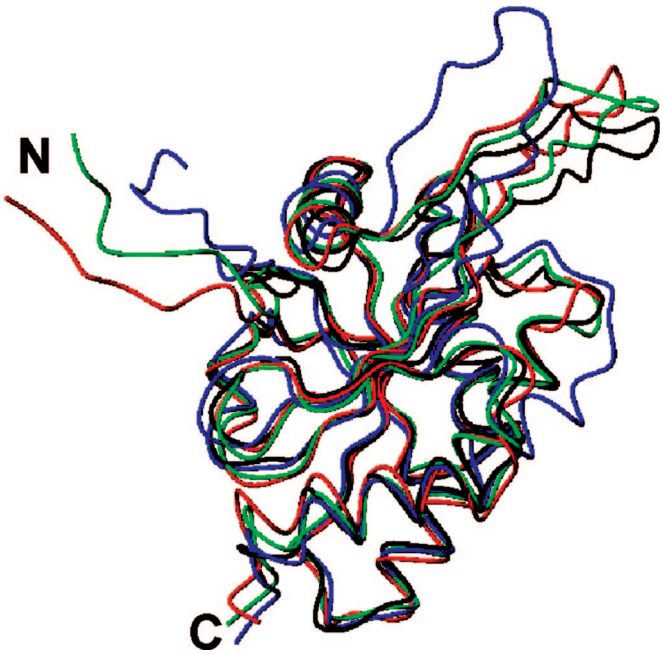



**Fig. 4.** Overlay of the backbone of apo (blue), Cu(I) (green), and Ni(II) (red) HSco1 solution structures and of Ni(II) (black) HSco1 x-ray structure.

two metal-binding Cys residues are oxidized and form a disulfide bond (Fig. 3; see also Fig. 11, which is published as supporting information on the PNAS web site) and therefore not capable of binding the Ni(II) ion as thiolates. Still, the metal ion remains in contact with the S—S bridge with a Ni—S distance of 2.0–2.2 Å, suggesting the formation of bonds with the available lone pairs of sulfur atoms (Fig. 3). The coordination sphere of Ni(II) is completed by His-260 ($N^{\varepsilon 2}$—Ni, 2.03–2.45 Å), in agreement with the solution structure of Ni(II)HSco1, and a water molecule or more likely an anion such as $Cl^-$ arranged in a distorted square planar geometry.

We have been able to trace one case only in the PDB of a nickel ion coordinated to Cys residues (PDB ID code 1FRF) that are at bonding distance; in this case, the metal-binding site is dinuclear (iron and nickel) and is made of four Cys residues pointing toward two metal ions that are at a distance of 3.2 Å from each other. Cys-75 and Cys-546 are at interaction distance (2.4 Å), whereas Cys-72 and Cys-543 are farther apart (3.0 Å) (26). The distances between nickel and the four sulfur atoms are as follows: Cys-72–Ni, 2.15 Å; Cys-543–Ni, 2.11 Å; Cys-75–Ni, 1.61 Å; and Cys-546–Ni, 2.44 Å.

## Discussion

The solution and crystal structures of the metal derivatives of HSco1 are completely superimposable along the entire amino acid sequence (Fig. 4) (backbone rms deviation to the new structure within 0.8 Å). These structures also are very similar to the solution structure of apoHSco1 with the exception of loop 8, which displays a different backbone conformation in apoHSco1, positioning the imidazole ring of His-260 at ≈10 Å from the sulfur atoms of the metal binding Cys residues (Fig. 4). In addition, helix $\alpha 2$, which encompasses the CPXXCP metal-binding site at its N terminus, is tilted in the apo solution structure with respect to the structure of the metallated HSco1 (Fig. 4). From the NMR structures, it appears that the apo form is highly disordered around the metal binding site, sampling more open conformations than in the metallated forms (Fig. 2). This observation also agrees with the results of conformation analysis by ESI-MS. In the x-ray structure of apoHSco1 (19), however, the protein is frozen in a specific conformation that does not reflect the real condition in solution. Therefore, metal binding is accompanied by a relatively large, albeit localized, effect on the protein structure, mainly involving loop 8: From an open conformation with local disorder, the structure converts into a well defined, compact state in which a metal ion is bound. In particular, the presence of the His ligand, suitably located in loop 8 to coordinate both divalent and monovalent metal ions, is important to modulate the order and disorder state of loop 8 observed in the metallated and apo forms, respectively. Also taking into account that disordered regions in protein structure often are engaged in protein–protein interactions (27), one may speculate that loop 8 modulates association–dissociation of HSco1 with its partner, the Cu(I) chaperone Cox17. For example, it is possible that, once Cu(I)Cox17 interacts transiently with apoHSco1 and donates its copper cargo to HSco1, loop 8 structurally rearranges and allows His binding and concomitant formation of the compact Cu(I)HSco1 structure, which might not exchange copper with Cox17. The formation of the stable, compact Cu(I)HSco1 state could thus constitute the important driving force of the copper transfer from Cox17 to HSco1.

**Biological Context.** The debate on whether HSco1 is a metalloprotein or a thioredoxin can significantly be advanced in light of the structural results presented here. HSco1 forms 1:1 complexes with the Cu(I) and Ni(II) ions by exploiting the same metal binding ligands, which confirms that HSco1 is suitable for binding both monovalent and divalent metal ions (15). Reduced HSco1 also can bind one equivalent of Cu(II); however, the reconstituted Cu(II)HSco1 complex shows two different coordination environments with different populations (15). Similar results on Cu(II)Sco1 complexes also were obtained for the *B. subtilis* and yeast proteins (12, 15). It is therefore reasonable to assign a copper chaperone role to HSco1, the metal ion being coordinated by two Cys residues and one His residue. A similar metal-binding site also is found in another copper chaperone, i.e., the ATX1 from *Synecocystis* (28). Similarly to the latter system, the metal-donating and the metal-receiving coordination sites are different, thus overcoming the condition that "donor" and "recipient" protein partners in metal transfer processes need metal-binding sites similar in structure, as recently suggested (29). The presence of three ligands, one of each being a "flexible" His residue, also makes the site suitable for the binding of divalent metal ions. Indeed, Cu(II), Ni(II), and, presumably, Zn(II) (14) can bind at the same site of HSco1. In the case of a divalent metals, the coordination can be completed by an additional exogenous ligand, e.g., $H_2O$, or by a protein carboxylate. The latter hypothesis is supported by the observation that the affinity of HSco1 for the Cu(II) ion is reduced if Asp-259 is mutated (15).

The fold of HSco1 is similar to that of redox-active proteins like thioredoxins and peroxiredoxins, with the metal-binding Cys residues located at the same positions as the conserved catalytic Cys residues in thioredoxins. This feature became apparent when the first structure of a Sco1 homolog was solved (12). Therefore, a thioredoxin fold has evolved as a metal chaperone to bind the metal atom via the di-Cys motif, and one may speculate that the thioredoxin function is still maintained. Indeed, apoHSco1 can be easily oxidized to form S—S bonds. In this research, we also have shown that, in the oxidized form, the protein has low affinity for metal ions because their addition did not affect the $^1H$–$^{15}N$ HSQC spectrum of oxidized apoHSco1, and no metal adduct has been detected by ESI-MS experiments. In the PDB, only one example of a metal ion bound to an oxidized S—S bond is reported (26). Therefore, it is feasible that the present crystal structure of the nickel derivative of oxidized HSco1 represents the transient copper-delivery complex, which might exist just before the copper is transferred to the $Cu_A$ site of the COXII subunit. Indeed, biochemical and genetic studies on yeast Sco1

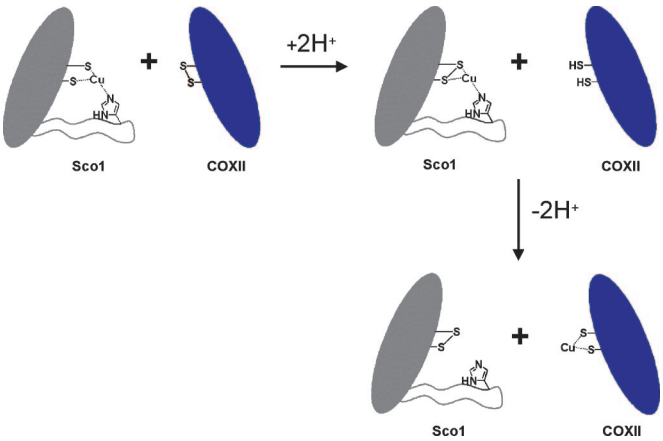



**Fig. 5.** Proposed mechanism for copper transfer from HSco1 to the COXII subunit of CcO. This model implies that HSco1 may form a transient species characterized by an oxidized S—S moiety still able to interact with a copper ion. This state might exist just before the copper is transferred to the $Cu_A$ site of CcO. In the metal transfer mechanism, HSco1 also can work as a thioredoxin in the reduction of the Cys residues in the $Cu_A$ site of COXII. At this stage, the oxidization state of copper during the metal transfer cannot be assessed. The $Cu_A$ site of the COXII subunit can indeed accept both Cu(I) and Cu(II) ions from HSco1. Broken lines indicate the metal coordination bonds.

demonstrated its ability for direct interaction with the COXII subunit (30, 31). Because numerous proteins in the mitochondrial intermembrane space have disulfide bonds (32–34), the two Cys residues of the $Cu_A$ site of COXII may also require reduction before metal transfer. The same requirement applies to the bacteria in which the COXII subunit is exposed to oxidizing periplasmatic or extracelluar environments. Therefore, it may be argued that Sco1 participates, along with metal transfer, in the reduction of the $Cu_A$ site of COXII and its metal-transfer mechanism might include the following steps: (*i*) CuHSco1 interacts with the oxidized COXII and reduces the Cys residues of the $Cu_A$ site as a thioredoxin, and (*ii*) the oxidized HSco1 protein transfers Cu to the reduced $Cu_A$ site (Fig. 5). The oxidized apoHSco1 could be reduced before the next metal transfer and its reduction can be performed by another thioredoxin-like protein, such as HSco2, which is known to play a key role in $Cu_A$ formation (35), or by cytochrome *c*, an electron-transfer protein that is found in the same operon of Sco1 homologs in some bacteria or is fused to some Sco1 paralog proteins (1). The copper delivery process therefore seems to involve several proteins in a complex molecular mechanism that requires further investigation. Interestingly, a dual functional role in the assembly mechanism of the copper enzyme superoxide dismutase has been similarly proposed for its copper chaperone CCS, which is, along with the copper donation, involved in the formation of an intramolecular disulfide essential for superoxide dismutase activity (36).

**Concluding Remarks.** We have succeeded in preparing Cu(I) and Ni(II) derivatives of HSco1 and determined their solution structures. The solution structure of apoHSco1 also has been determined. The structures confirm that the metal ions are bound by two Cys residues and one His residue, and they show the transition from a locally disordered apo protein to a compact metallated form, as confirmed by ESI-MS studies. We also obtained crystals of Ni(II)HSco1, which suggest the binding of the metal ion to the oxidized form of HSco1. This species may represent a transition state of the copper transfer from HSco1 to the $Cu_A$ site of COXII. This transient species may present a missing link that integrates the metal transfer and thioredoxin functions already proposed for this fascinating protein. A similar mechanism of copper transfer also has been previously suggested for the bacterial Sco1 homologue PrrC from *Rhodobacter sphaeroides* (17).

## Methods

**Protein Preparation and Characterization.** Long and short soluble domains of HSco1 (lacking the first 351 and 393 bp, respectively, corresponding to residues 1–117 and 1–131) were amplified by PCR, cloned into the Gateway Entry vector pENTR/tobacco etch virus/D-TOPO (Invitrogen), and subcloned into pETG-30A (European Molecular Biology Laboratory Protein Expression and Purification Facility) by Gateway LR reaction to generate N-terminal, His-GST fused proteins. The proteins were expressed in *Escherichia coli* BL21-Gold(DE3) cells (Stratagene), which were grown in LB and minimal medium [$(^{15}NH_4)_2SO_4$ and/or [$^{13}C$]-glucose] for the production of labeled samples. Purification was performed by using a HiTrap chelating HP column (Amersham Pharmacia Biosciences) charged with Zn(II). His-GST tag was cleaved with AcTEV, and separated from the C-terminal domain with a second purification step. After this purification, the protein preparations showed a single component by SDS/PAGE with <5% of copper bound to the protein, as checked through ESI-MS spectrometry. DTT was added to the apoprotein in a 10 mM concentration to reduce the Cys residues of the CPXXCP motif before metal reconstitution. The Cu(I), Cu(II), Ni(II) metallated forms were obtained by addition of stoichiometric amounts of the metal ions {as $[Cu(I)(CH_3CN)_4]PF_6$, $CuSO_4$, and $NiCl_2$} to diluted protein solutions in 50 mM phosphate buffer at pH 7.2, followed by protein concentration under nitrogen atmosphere. The metal content was finally determined by inductively coupled plasma MS.

Electronic spectra on the metal derivatives were recorded on a Cary 50 spectrophotometer (Varian). EPR spectra on Cu(II)HSco1 were recorded at 180 K on an Elexsys E500 spectrometer (Bruker) equipped with a X-band microwave bridge (microwave frequency, 9.45 GHz) and an ER 4131 VT unit for temperature control. To investigate the aggregation state of HSco1, 0.5–1 mM protein samples were run on a Superdex75 HR-10/30 size-exclusion column on an AKTA-FPLC system (Amersham Pharmacia Biosciences) connected with a multiangle light scattering (DAWN-EOS, Wyatt Technologies, Santa Barbara, CA) coupled with quasielastic light-scattering detectors.

Before ESI-MS experiments, purified apoHSco1 protein was brought into 50 mM ammonium acetate buffer (pH 7.5) by using a HiPrep26/10 desalting column (Amersham Pharmacia Biosciences). In MS experiments, 1.8 $\mu$M protein samples were infused by a syringe pump at 15 $\mu$l/min into an Ettan API ESI-TOF mass spectrometer (Amersham Pharmacia Biosciences). Mass spectra were recorded during 2–3 min at a capillary exit voltage of 150 V. apoHSco1 was reduced by addition of 0.5 or 1.0 mM DTT at 25°C. Reconstitution of apoHSco1 with copper was conducted as follows. First, Cu(II) acetate was dissolved at 150 $\mu$M concentration in argon-saturated 50 mM ammonium acetate, pH 7.5, and Cu(II) was reduced to Cu(I) by addition of 0.5 mM DTT. Different equivalents of freshly prepared Cu(I)DTT complex were added to the apoHSco1 samples (protein concentration, 1.8 $\mu$M), the mixture was incubated for 1 min at 25°C, and ESI-MS spectra were recorded as described above.

**Solution Structures Determination.** NMR spectral assignment and structure determination were obtained through the experiments listed in Table 1, which also indicates the magnetic fields at which they were collected. Overall, the resonances of 95% of carbon atoms, 97% of nitrogen atoms, and 90% of protons were assigned in Cu(I)HSco1 and Ni(II)HSco1 (deposited in the Protein Data Bank database). In the case of apo form, the resonances of 85% of carbon atoms, 80% of nitrogen atoms, and 80% of proton

atoms were assigned. The $^{1}H$, $^{13}C$, and $^{15}N$ resonance assignments of the apo, Cu(I), and Ni(II)HSco1 forms are reported, respectively, in Tables 2–4, which are published as supporting information on the PNAS web site. The His ring protons were assigned through a $^{1}H$–$^{15}N$ HSQC experiment tailored to the detection of $^{2}J$ $^{1}H$–$^{15}N$ couplings and from the analysis of the $^{13}C$-NOESY-HSQC spectra. For all His residues, all of the nonexchangeable protons were assigned in Cu(I)HSco1 and Ni(II)HSco1. The exchangeable proton of the metal-binding ligand, His-260, also was detected in both metallated forms. $R_1$ and $R_2$ $^{15}N$ relaxation rates and $^{1}H$–$^{15}N$ nuclear Overhauser effects (with and without $^{1}H$ saturation) (Table 1) were measured at 298K on Avance 500 and 600 MHz Bruker spectrometers and then analyzed by using a standard procedure (37). After conversion of the NMR data in structural constraints [3,035, 2,776 and 2,066 meaningful proton–proton distances, together with 85 $\psi$ and 83 $\phi$ angle constraints for Cu(I)HSco1, Ni(II)HSco1, and apoHSco1, respectively], the structures were calculated using the program DYANA (38). The best 30 structures of the DYANA family were then subjected to restrained energy minimization with AMBER 8.0 (39). The force-field parameters for the metal ions were adapted from similar systems (40, 41). The statistical analysis of the restrained energy minimization family of apoHSco1, Cu(I)HSco1, and Ni(II)HSco1 structures are reported, respectively, in Tables 5, 6, and 7, which are published as supporting information on the PNAS web site. The programs PROCHECK and PROCHECK-NMR (42, 43) were used in the evaluation of the quality of the structures. More than 90% of residues were located in the allowed regions of the Ramachandran plot.

**Crystallization, Data Collection, and Structure Solution.** Crystals of Ni(II)HSco1 grew at 20°C from a 0.1 M Tris·HCl/30% polyethylene glycol 6000 solution at pH 8.5 by the vapor diffusion technique. The final protein concentration was ≈10 mg/ml. The data set was collected by using synchrotron radiation at beamline ID-29 (European Synchrotron Radiation Facility, Grenoble, France) at 100 K, with the crystal cryocooled, in the presence of 10–15% of ethylen glycol. The Ni(II)HSco1 crystal diffracted up to 2.5-Å resolution and belongs to space group $P2_12_12_1$ ($a$ = 51.46 Å, $b$ = 52.44 Å, $c$ = 136.41 Å) with two molecules in the asymmetric unit, a solvent content of 47.1%, and a mosaicity of 0.7°. The structure was solved by using the molecular replacement technique, with the structure of the apoHSco1 (PDB ID code 1WP0) as starting model.

The Ramachandran plot of the refined model shows that 97.5% of residues are in allowed regions of the plot, 2.5% of residues are in generously allowed, and no residues are in disallowed regions. Table 8, which is published as supporting information on the PNAS web site, reports the data collection and refinement statistics.

We thank Dr. Rannar Sillard (Karolinska Institute, Stockholm) for providing access to ESI-MS instruments. This work was supported by European Community "Structural Proteomics in Europe" Grant QLG2-CT-2002-00988, by Marie Curie Host Fellowship MEST-CT-2004-504391 for early stage research training ("NMR in Inorganic Structural Biology"), Estonian Science Foundation Grant 5635, and by Ente Cassa Risparmio di Firenze.

1. Arnesano, F., Banci, L., Bertini, I. & Martinelli, M. (2005) *J. Proteome Res.* **4,** 63–70.
2. Chinenov, Y. V. (2000) *J. Mol. Med.* **78,** 239–242.
3. Schulze, M. & Rodel, G. (1989) *Mol. Gen. Genet.* **216,** 37–43.
4. Smits, P. H., de Hann, M., Maat, C. & Grivell, L. A. (1994) *Yeast* **10,** 75–80.
5. Carr, H. S. & Winge, D. R. (2003) *Acc. Chem. Res.* **36,** 309–316.
6. Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. & Yoshikawa, S. (1996) *Science* **272,** 1136–1144.
7. Khalimonchuk, O. & Rodel, G. (2005) *Mitochondrion* **5,** 363–388.
8. Glerum, D. M., Shtanko, A. & Tzagoloff, A. (1996) *J. Biol. Chem.* **271,** 20531–20535.
9. Horng, Y. C., Cobine, P. A., Maxfield, A. B., Carr, H. S. & Winge, D. R. (2004) *J. Biol. Chem.* **279,** 35334–35340.
10. Buchwald, P., Krummeck, G. & Rodel, G. (1991) *Mol. Gen. Genet.* **229,** 413–420.
11. Beers, J., Glerum, D. M. & Tzagoloff, A. (2002) *J. Biol. Chem.* **277,** 22185–22190.
12. Balatri, E., Banci, L., Bertini, I., Cantini, F. & Ciofi-Baffoni, S. (2003) *Structure (London)* **11,** 1431–1443.
13. Nittis, T., George, G. N. & Winge, D. R. (2001) *J. Biol. Chem.* **276,** 42520–42526.
14. Andruzzi, L., Nakano, M., Nilges, M. J. & Blackburn, N. J. (2005) *J. Am. Chem. Soc.* **127,** 16548–16558.
15. Horng, Y. C., Leary, S. C., Cobine, P. A., Young, F. B., George, G. N., Shoubridge, E. A. & Winge, D. R. (2005) *J. Biol. Chem.* **280,** 34113–34122.
16. Ye, Q., Imriskova-Sosova, I., Hill, B. C. & Jia, Z. (2005) *Biochemistry* **44,** 2934–2942.
17. McEwan, A. G., Lewin, A., Davy, S. L., Boetzel, R., Leech, A., Walker, D., Wood, T. & Moore, G. R. (2002) *FEBS Lett.* **518,** 10–16.
18. Arner, E. S. & Holmgren, A. (2000) *Eur. J. Biochem.* **267,** 6102–6109.
19. Williams, J. C., Sue, C., Banting, G. S., Yang, H., Glerum, D. M., Hendrickson, W. A. & Schon, E. A. (2005) *J. Biol. Chem.* **280,** 15202–15211.
20. Lode, A., Paret, C. & Rodel, G. (2002) *Yeast* **19,** 909–922.
21. Arciero, D. M., Pierce, B. S., Hendrich, M. P. & Hooper, A. B. (2002) *Biochemistry* **41,** 1703–1709.
22. Basumallick, L., Sarangi, R., DeBeer George, S., Elmore, B., Hooper, A. B., Hedman, B., Hodgson, K. O. & Solomon, E. I. (2005) *J. Am. Chem. Soc.* **127,** 3531–3544.
23. Maroney, M. J., Choudhury, S. B., Bryngelson, P. A., Mirza, S. A. & Sherrod, M. J. (1996) *Inorg. Chem.* **35,** 1073–1076.
24. Bruschweiler, R. (2003) *Curr. Opin. Struct. Biol.* **13,** 175–183.
25. Eyles, S. J. & Kalatshov, I. A. (2004) *Methods* **34,** 88–99.
26. Rousset, M., Montet, Y., Guigliarelli, B., Forget, N., Asso, M., Bertrand, P., Fontecilla-Camps, J. C. & Hatchikian, E. C. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 11625–11630.
27. Dyson, H. J. & Wright, P. E. (2005) *Nat. Rev. Mol. Cell. Biol.* **6,** 197–208.
28. Banci, L., Bertini, I., Borrelly, G. P. M., Ciofi-Baffoni, S., Robinson, N. J. & Su, X. C. (2004) *J. Biol. Chem.* **279,** 27502–27510.
29. Rosenzweig, A. C. (2001) *Acc. Chem. Res.* **34,** 119–128.
30. Lode, A., Kuschel, M., Paret, C. & Rodel, G. (2000) *FEBS Lett.* **485,** 19–24.
31. Dickinson, E. K., Adams, D. L., Schon, E. A. & Glerum, D. M. (2000) *J. Biol. Chem.* **275,** 26780–26785.
32. Hofmann, S., Rothbauer, U., Muhlenbein, N., Baiker, K., Hell, K. & Bauer, M. F. (2005) *J. Mol. Biol.* **353,** 517–528.
33. Field, L. S., Furukawa, Y., O'Halloran, T. V. & Culotta, V. C. (2003) *J. Biol. Chem.* **278,** 28052–28059.
34. Sturtz, L. A., Diekert, K., Jensen, L. T., Lill, R. & Culotta, V. C. (2001) *J. Biol. Chem.* **276,** 38084–38089.
35. Leary, S. C., Kaufman, B. A., Pellecchia, G., Guercin, G. H., Mattman, A., Jaksch, M. & Shoubridge, E. A. (2004) *Hum. Mol. Genet.* **13,** 1839–1848.
36. Furukawa, Y., Torres, A. S. & O'Halloran, T. V. (2004) *EMBO J.* **23,** 2872–2881.
37. Peng, J. W. & Wagner, G. (1992) *J. Magn. Reson.* **98,** 308–332.
38. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298.
39. Case, D. A., Darden, T. A., Cheatham, T. E., Simmerling, C. L., Wang, J., Duke, R. E., Luo, R., Merz, K. M., Wang, B., Pearlman, D. A., *et al.* (2004) AMBER (Univ. of California, San Francisco), Version 8.0.
40. Poger, D., Fuchs, H. J. R., Nedev, H., Ferrand, M. & Crouzy, S. (2005) *FEBS Lett.* **579,** 5287–5292.
41. Banci, L., Bertini, I., Bruni, B., Carloni, P., Luchinat, C., Mangani, S., Orioli, P. L., Piccioli, M., Rypniewski, W. & Wilson, K. (1994) *Biochem. Biophys. Res. Commun.* **202,** 1088–1095.
42. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486.
43. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26,** 283–291.

# SUPPLEMENTARY MATERIAL

## A hint for the function of human Sco1 from different structures

Lucia Banci, Ivano Bertini, Vito Calderone, Simone Ciofi-Baffoni, Stefano Mangani, Manuele Martinelli, Peep Palumaa, and Shenlin Wang

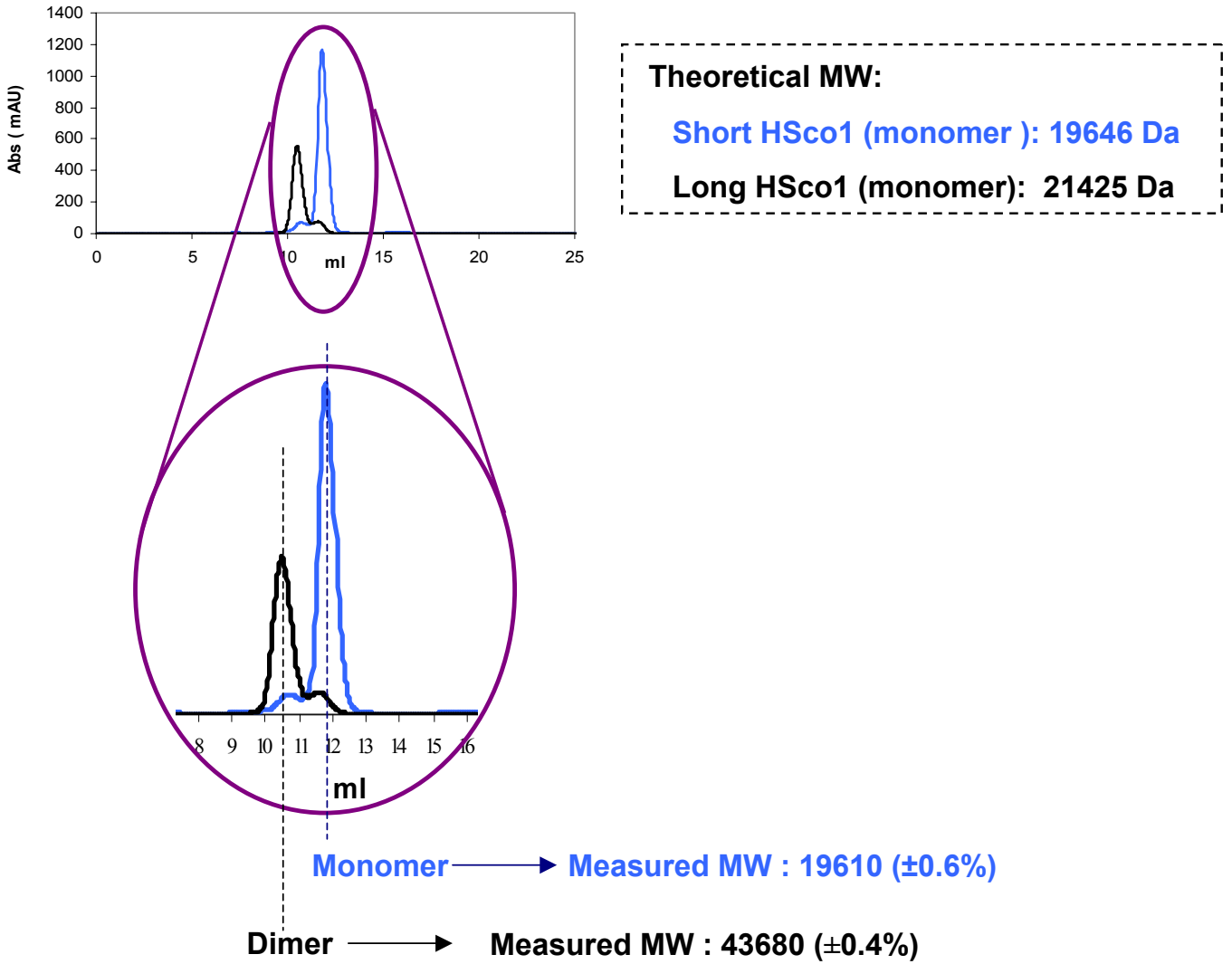



**Fig. 6 Aggregation state of HSco1 domains monitored by size exclusion chromatography coupled with multi angle light scattering.**

Long and short truncated constructs of HSco1 eluted mainly as a single peak with a molecular weight (MW) corresponding to a dimeric (black) and monomeric (blue) state, respectively.

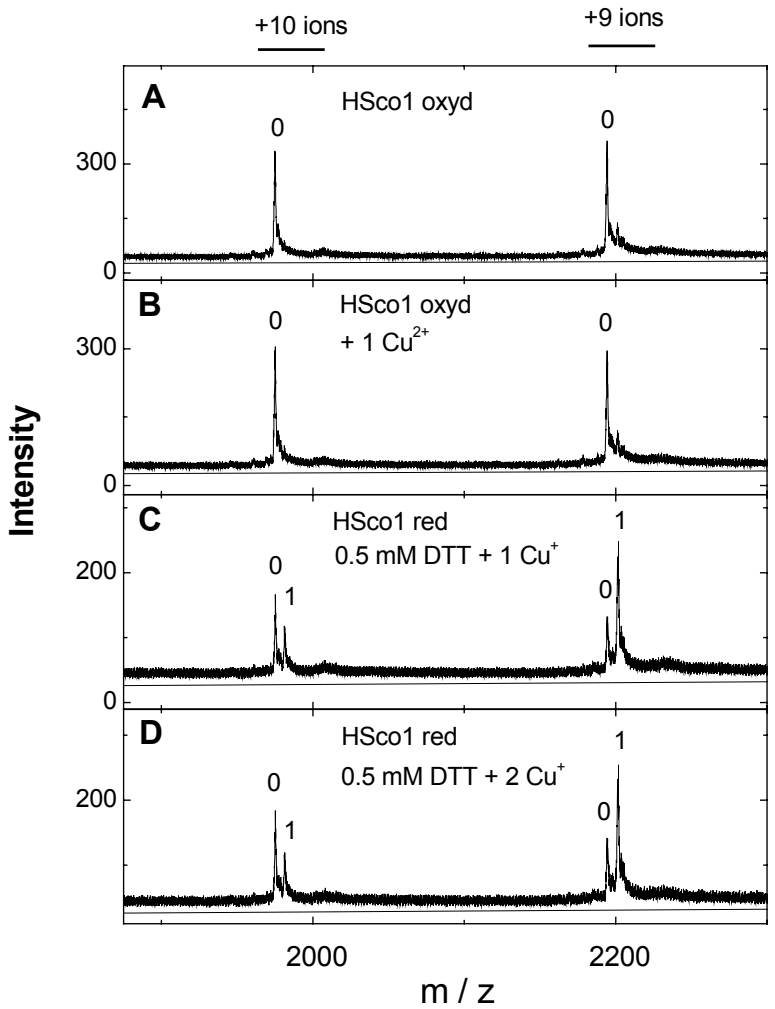



**Fig. 7 Binding of copper ions to HScо1.**

ESI-TOF mass spectra of HScо1 (1.8 μM) reconstituted with Cu(II)acetate or Cu(I)DTT complex in 50 mM ammonium acetate, pH 7.5. A - HScо1; B - HScо1 + 1 equivalent of Cu(II)acetate; C - HScо1 + 1 equivalent of Cu(I) in the presence of 0.5 mM DTT; D - HScо1 + 2 equivalents of Cu(I) in in the presence of 0.5 mM DTT. Charge state +9 and +10 ions are presented and numbers on the peaks denote the metal stoichiometry of the complex.

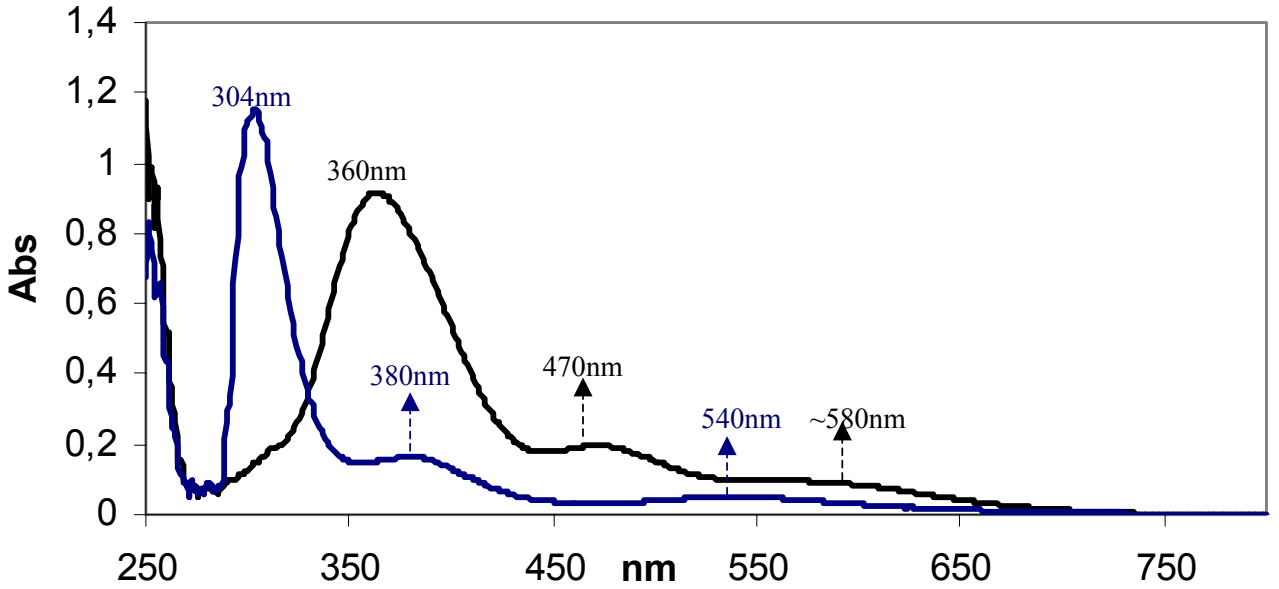



**Fig. 8 Electronic absorption spectra of Cu(II) (black) and Ni(II) (blu) HSco1.**
The spectra were obtained at 298K in 50 mM phosphate (pH = 7.2). Both spectra were baseline corrected by subtracting the spectrum of apoHSco1.

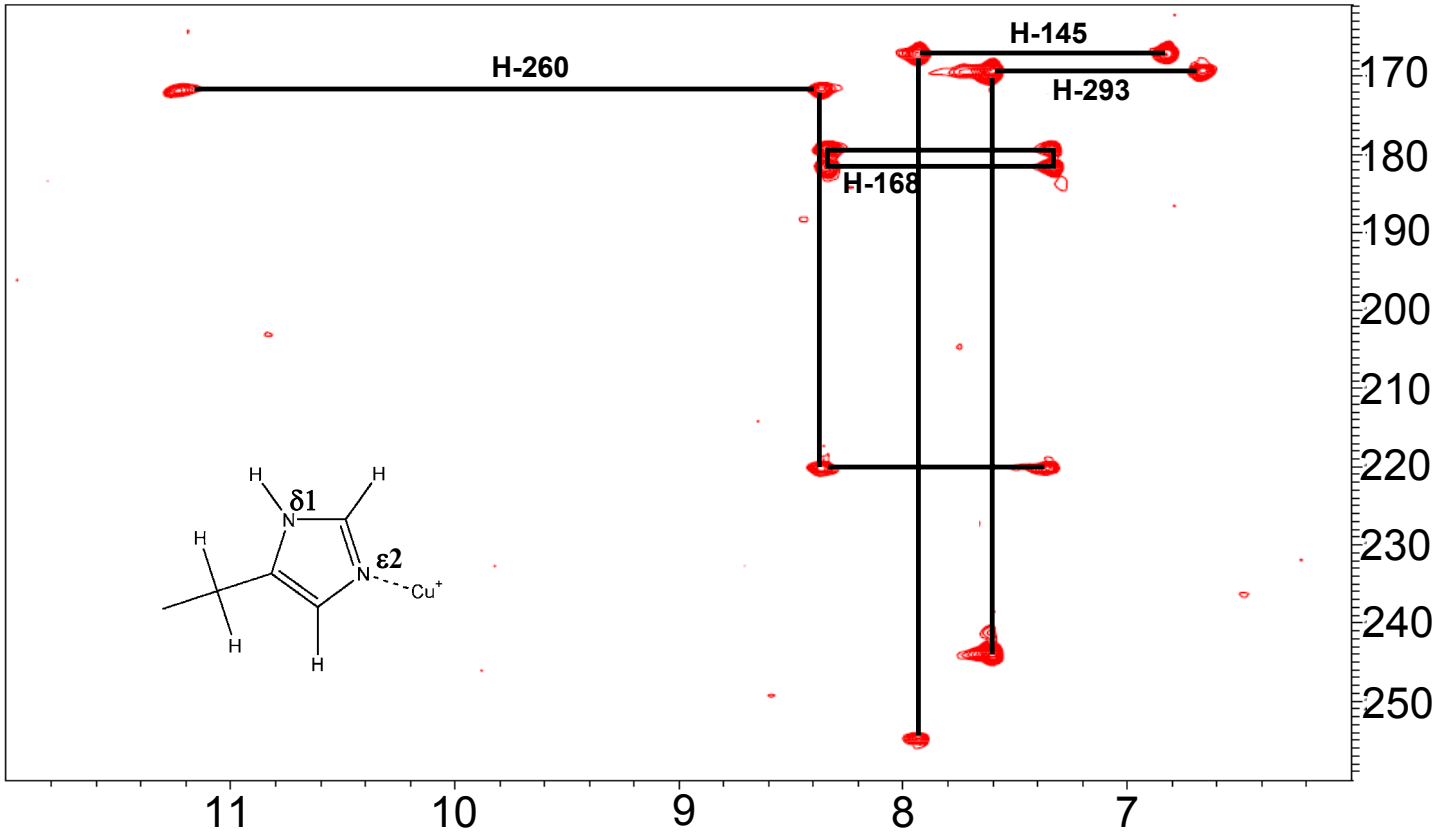



**Fig. 9 $^{1}$H-$^{15}$N HSQC spectra optimized for the detection of $^{2}$J NH of histidine rings, recorded on Cu(I)HSco1.**
The spectra were collected at 298 K and pH 7.2 on a spectrometer equipped with cryo-probe and operating at 500 MHz. The tautomeric species of His ring of copper(I) bound form is also depicted.

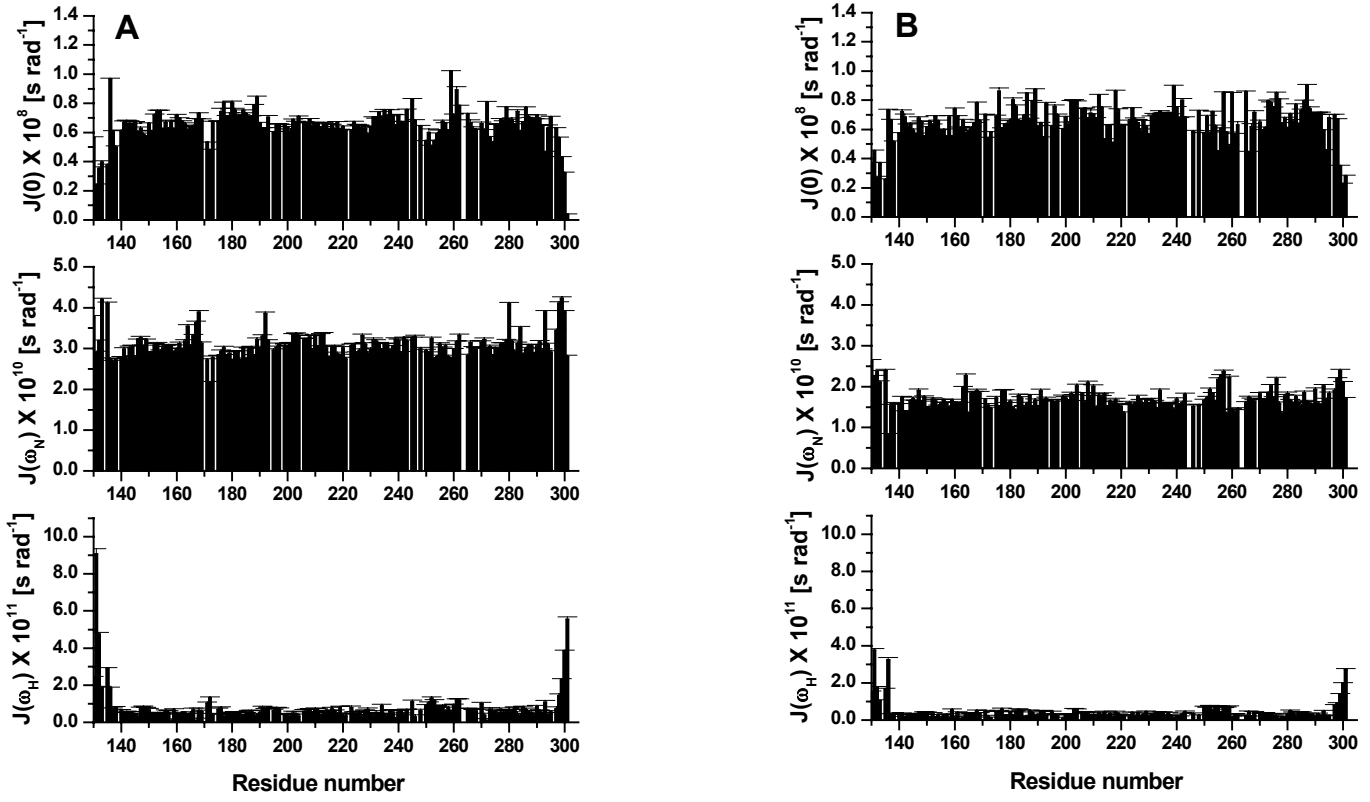



**Fig. 10 Spectral density functions $J(\omega_H)$, $J(\omega_N)$ and $J(0)$ versus the residue numbers as obtained from the $^{15}N$ relaxation data measured at 500 MHz and 600 MHz for Cu(I) (A) and Ni(II)HSco1 (B), respectively.**

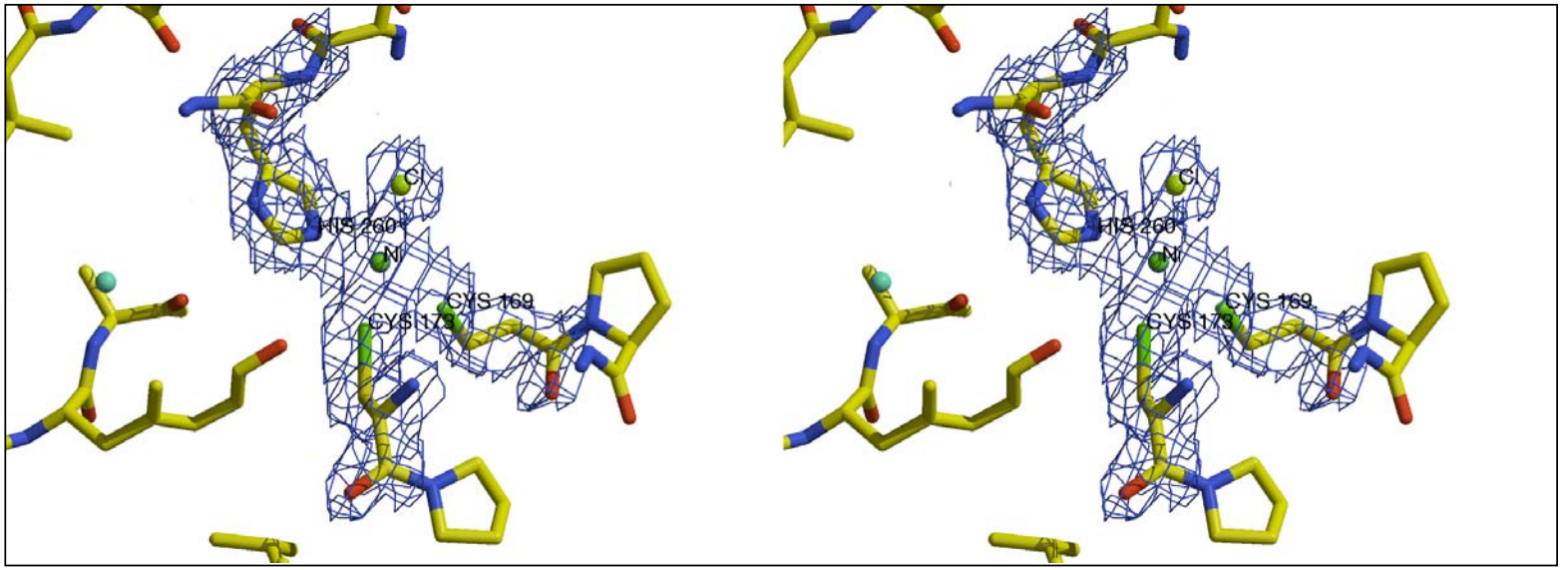



**Fig. 11 Stereo view of the electron density map (2Fo-Fc contoured at 1 sigma level) around the residues involved in nickel binding.**

# 3.3

## Human Sco1 functional studies and pathological implications of P174L mutant

Lucia Banci, Ivano Bertini, Simone Ciofi-Baffoni, Iliana Leontari, Manuele Martinelli, Peep Palumaa, Rannar Sillard, Shenlin Wang.

***(Proc Natl Acad Sci U S A*. In Press)**

# Human Sco1 functional studies and pathological implications of the P174L mutant

Lucia Banci*†, Ivano Bertini*‡, Simone Ciofi-Baffoni*, Iliana Leontari*, Manuele Martinelli*, Peep Palumaa§, Rannar Sillard§¶, and Shenlin Wang*

*Magnetic Resonance Center (CERM) and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy; †FiorGen Foundation, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy; ¶Department of Medical Biochemistry and Biophysics, Karolinska Institutet, SE-171 77 Stockholm, Sweden; and §Department of Gene Technology, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia

Edited by Joan Selverstone Valentine, University of California, Los Angeles, CA, and approved October 30, 2006 (received for review July 21, 2006)

**The pathogenic mutant (P174L) of human Sco1 produces respiratory chain deficiency associated with cytochrome *c* oxidase (CcO) assembly defects. The solution structure of the mutant in its Cu(I) form shows that Leu-174 prevents the formation of a well packed hydrophobic region around the metal-binding site and causes a reduction of the affinity of copper(I) for the protein. $K_D$ values for Cu(I)WT-HSco1 and Cu(I)P174L-HSco1 are $\approx 10^{-17}$ and $\approx 10^{-13}$, respectively. The reduction potentials of the two apo proteins are similar, but slower reduction/oxidation rates are found for the mutant with respect to the WT. The mitochondrial metallochaperone in the partially oxidized $Cu_1(I)Cox17_{2S\text{-}S}$ form, at variance with the fully reduced $Cu_4(I)Cox17$, interacts transiently with both WT-HSco1 and the mutant, forming the Cox17/Cu(I)/HSco1 complex, but copper is efficiently transferred only in the case of WT protein. $Cu_1(I)Cox17_{2S\text{-}S}$ indeed has an affinity for copper(I) ($K_D \approx 10^{-15}$) higher than that of the P174L-HSco1 mutant but lower than that of WT-HSco1. We propose that HSco1 mutation, altering the structure around the metal-binding site, affects both copper(I) binding and redox properties of the protein, thus impairing the efficiency of copper transfer to CcO. The pathogenic mutation therefore could (*i*) lessen the Sco1 affinity for copper(I) and hence copper supply for CcO or (*ii*) decrease the efficiency of reduction of CcO thiols involved in copper binding, or both effects could be produced by the mutation.**

cytochrome *c* oxidase | mass spectrometry | NMR | copper chaperone | respiratory chain deficiency

Cytochrome *c* oxidase (CcO) is the terminal enzyme of the respiratory chain, embedded in the inner mitochondrial membrane of all eukaryotes and in the plasma membrane of many prokaryotes. In eukaryotes, a large number of nuclear genes are required for the proper assembly and function of the CcO complex, among which at least six proteins (Cox17, Cox23, Cox19, Sco1, Sco2, and Cox11) are involved in the delivery and insertion of copper into the binuclear $Cu_A$ and $Cu_B$-heme $a_3$ sites of CcO (1–3). Sco1 and Sco2, which are anchored to the inner mitochondrial membrane through a single transmembrane helix, are specifically implicated in the assembly of the $Cu_A$ site of CcO (4–7). Sco1 and Sco2 are both metal-binding proteins capable to bind either copper(I) or copper(II) in the same metal-binding site (8, 9); mutations of the ligands that abrogate their copper-binding ability drastically compromise CcO activity (9, 10). The mitochondrial copper chaperone Cox17 is capable of donating copper(I) to Sco1 (11), implying that it might transfer copper to Sco1, which, in turn, inserts it into the $Cu_A$ site. According to this mechanism, Sco1 was shown to interact with the $Cu_A$-containing COXII subunit of CcO (6).

Important insights on the function of this class of proteins also recently were obtained from a genome-wide search for all available prokaryotic genomes (12). This search showed the presence of several Sco paralogs adjacent to copper and/or redox proteins, thus suggesting that Sco proteins can have multifunctional properties involved in different physiological processes.

The structure of Sco proteins is characterized by a thioredoxin fold (13–15), and, from the first structural characterization (13), it was suggested that the protein also may fulfill a redox function. Recently, we have solved the solution structures of the demetallated, the copper(I), and the nickel(II) derivatives of human Sco1 (HSco1 hereafter) as well as the crystal structure of a nickel(II) derivative (16). The structural analysis showed that Cu(I) is coordinated by ligands located in two distant protein regions: Cys-169 and Cys-173 of the CXXXC conserved motif in loop 3 and His-260 in loop 8. In the structure of the apo protein, the latter loop is highly disordered and mobile in solution (16), suggesting that the metal binding is able to "freeze" this loop in a more rigid conformation, thus representing an important step in the metal-binding process of HSco1. In the crystal structure of oxidized Ni(II)HSco1, the two metal-binding cysteines are oxidized and form a disulfide bond, still with the metal ion interacting with the S-S moiety. This metal environment showed that the protein can have a peculiar mode of interaction with a metal ion, suggesting that cysteine oxidation of HSco1 also might play a role in the metal release to $Cu_A$ site of CcO (16). These structural data further support our initial suggestion that Sco1 can act both as a copper chaperone and a thioredoxin (13).

Human *SCO1* and *SCO2* genes can experience pathogenic mutations that produce respiratory chain deficiency associated with CcO assembly defects (17). The missense mutation in human Sco1 gene of a proline into a leucine, P174L, is associated with a fatal neonatal hepatopathy when the second allele also is nonfunctional (18). This finding implies that the pathology is caused by a loss of normal Sco1 function rather than gain of some aberrant action of the mutant protein. This proline, adjacent to the conserved CXXXC domain of HSco1, is completely conserved in eukaryotes. Introduction of the P174L mutant chimeric HSco1 protein in Sco1-null yeast mutants impaired CcO assembly and induced loss of CcO activity (19). WT protein and P174L mutant chimera are present in comparable concentrations in mitochondrial extracts, thus indicating that the pathogenicity of the mutant does not result from its instability but rather from an impaired function (19). During the preparation of our work, a paper on the same P174L-HSco1 mutant came out indicating that the mutation compromises Cox17-dependent metalation

Author contributions: L.B., I.B., S.C.-B., and P.P. designed research; S.C.-B., I.L., M.M., P.P., and S.W. performed research; R.S. contributed new reagents/analytic tools; S.C.-B., M.M., P.P., and S.W. analyzed data; and L.B., I.B., S.C.-B., and P.P. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Abbreviations: CcO, cytochrome *c* oxidase; HSQC, heteronuclear single quantum correlation; ESI, electrospray ionization; GSH, reduced glutathione; GSSG, oxidized glutathione.

Data deposition: The atomic coordinates and structural restraints for human Cu(I)P174LSco1 structure have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2HRN).

‡To whom correspondence should be addressed. E-mail: bertini@cerm.unifi.it.

This article contains supporting information online at www.pnas.org/cgi/content/full/0606189103/DC1.

without impairing copper binding (20). In the present research, we have explored at the molecular level the effect of the mutation on the Sco1 function by solving the solution structure of the P174L-HSco1 mutant, characterizing the metal-binding and redox properties of the P174L-HSco1 mutant with respect to WT HSco1 (WT-HSco1 hereafter), and investigating the role of Cox17 in the copper transfer to both WT and mutant HSco1 proteins. Our work demonstrates that the P174L mutation alters hydrophobic interactions around the metal-binding site, thus affecting both the copper-binding and the redox properties of HSco1 and, therefore, determining a reduction of the copper transfer efficiency from Cox17 to HSco1. We also learned that the partially oxidized $Cu_1(I)Cox17_{2S\text{-}S}$ and not the fully reduced $Cu_4(I)Cox17$ are the preferential species delivering copper to HSco1.

## Results

**Structural Properties of the Pathogenic Mutant P174L-HSco1.** ApoP174L-HSco1 at millimolar concentration is a well folded protein as it appears from the $^1H$-$^{15}N$ heteronuclear single quantum correlation (HSQC) spectrum [see supporting information (SI) Fig. 5], which also is similar to that of apoWT-HSco1. Accordingly, the mutant maintains the same fold of WT-HSco1, featuring indeed the same content of $\beta$-strands (21%) and $\alpha$-helices (26%) as derived from circular dichroism spectra. Fluorimetric thermal unfolding assay (SI Fig. 5) shows that the thermal stability of apoP174L-HSco1 also is the same as that of apoWT-HSco1, both having a melting temperature of ≈54°C. These data show that the pathogenic nature of the mutation does not originate from fold and/or stability differences of apoP174L-HSco1 with respect to the WT protein. This conclusion also is consistent with *in vivo* results, which showed that comparable concentrations of either WT or P174L mutant protein are present in mitochondrial extracts (19).

ApoP174L-HSco1, as apoWT-HSco1, is able to bind Cu(I) ions, when the latter is provided as Cu(I) acetonitrile complex. The metalation of both proteins, followed through $^1H$-$^{15}N$ HSQC spectra in the presence of 1 mM DTT (SI Fig. 6), however, shows meaningful differences. Indeed, although 1 eq of Cu(I) fully metallates WT-HSco1, complete metalation of P174L-HSco1 is not achieved even when 3 eq of Cu(I) are added, as ≈20% of P174L-HSco1 is still in the apo form (SI Fig. 6). In the latter experiment, Cu(I) thus competes between the mutant and the DTT chelator, allowing us to determine the dissociation constant ($K_D$) of the Cu(I)P174L-HSco1 complex. Following the relative ratio of apo and copper(I) forms, as obtained from the $^1H$-$^{15}N$ HSQC spectra of the mixture during the titration steps [the interaction between the protein and Cu(I)DTT complex is slow on the chemical-shift time scale], $K_D$ of the Cu(I)P174L-HSco1 complex is estimated to be $3.2 \pm 0.6 \times 10^{-13}$ M. This value is much higher than the $K_D \approx 10^{-17}$ of the WT protein (see *Transfer of Copper from Cox17 to HSco1*). When copper(I) is added in the absence of DTT, 1.5 eq are enough to fully metallate the mutant.

Copper binding to WT-HSco1 induces major conformational changes to the protein, which passes from an open state in the apo form to a closed, compact one upon copper binding (16). This change clearly is evident from the solution structures of these two states of WT-HSco1 and also can readily be assessed from the charge state analysis of electrospray ionization (ESI)-MS spectra. Conformational analysis by ESI-MS is based on observation that, during the ESI, tightly compact proteins acquire only a narrow range of low-charge values, whereas open conformations produce a more heterogeneous and highly charged population (21). Charge-state distribution of Cu(I)P174L-HSco1 ($Int_{+9}/Int_{+10} = 1.6$; Fig. 1) is similar to that of apoP174L-HSco1 ($Int_{+9}/Int_{+10} = 1.7$) and apoWT-HSco1 ($Int_{+9}/Int_{+10} = 1.4$), indicating that all these systems have a more

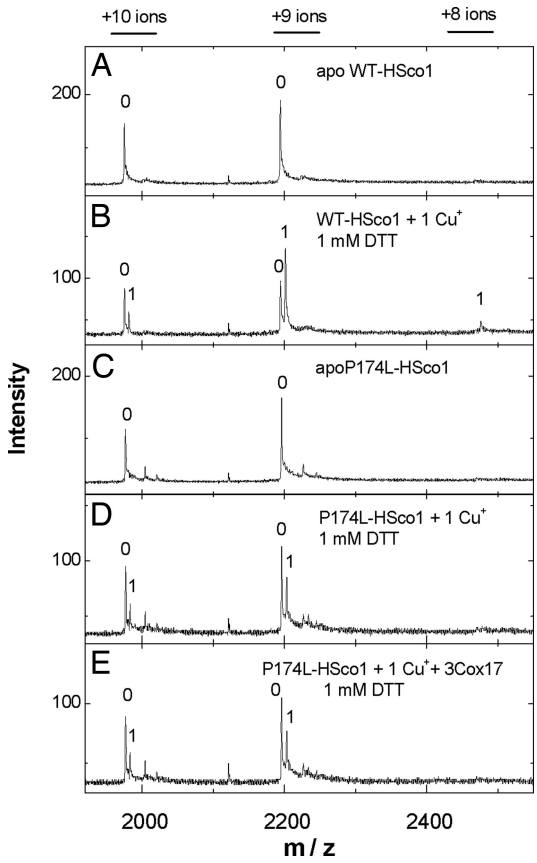



**Fig. 1.** Binding of copper ions to human Sco1 and mutant P174LSco1. ESI-TOF mass spectra of WT-HSco1 (0.8 $\mu$M) and P174L-HSco1 (1.2 $\mu$M) reconstituted with Cu(I)DTT complex in 50 mM ammonium acetate, pH 7.5. (*A*) apoWT-HSco1. (*B*) WT-HSco1 + 1 eq of Cu(I) in the presence of 1.0 mM DTT. (*C*) apoP174L-HSco1. (*D*) P174L-HSco1 + 1 eq of Cu(I) in the presence of 1.0 mM DTT. (*E*) P174L-HSco1 + 1 eq of Cu(I) + 3 eq of $Cox17_{2S\text{-}S}$ in the presence of 1.0 mM DTT. Charge state +8, +9, and +10 ions are presented and numbers on the peaks denote the metal stoichiometry of the complex.

open conformation than Cu(I)WT-HSco1, in which the lower +9 state was more populated ($Int_{+9}/Int_{+10} = 3.8$; Fig. 1). These data therefore show that Cu(I) binding to apoP174L-HSco1 is unable to induce the compact conformation as it occurs in the WT protein.

In the $^1H$-$^{15}N$ HSQC spectrum of the fully metallated P174L-HSco1 ($Cu_1(I)$P174L-HSco1), 34 residues show a splitting in two or, in a few cases, three NH cross-peaks, indicating the occurrence of structural heterogeneity, which is absent in $Cu_1(I)$WT-HSco1. The two major species are present in a ratio of 60:40. The relative intensity of the resonances of the two species is temperature-dependent, indicating that there is an equilibrium between the two. The solution structure of $Cu_1(I)$P174L-HSco1 (determined from the more abundant species) shows the same thioredoxin-like fold of the WT protein (Fig. 2) with an average backbone rmsd of 1.30 Å. The only significant structural difference is in loop 8, which contains the His copper ligand. In the mutant, this loop does not have the short $\beta$-hairpin (Fig. 2), similarly to what occurs in the apoWT-HSco1 solution structure but at variance with the $Cu_1(I)$WT-HSco1 structure (16). Because no apo form of the mutant is present in the NMR sample (see *Methods*), the latter less defined secondary structure could be the consequence of the structural heterogeneity in this region, which in turn could induce a weaker copper-affinity in the mutant. Several residues with double conformations are located in helix $\alpha$2, which is in contact with the metal-binding CXXXC

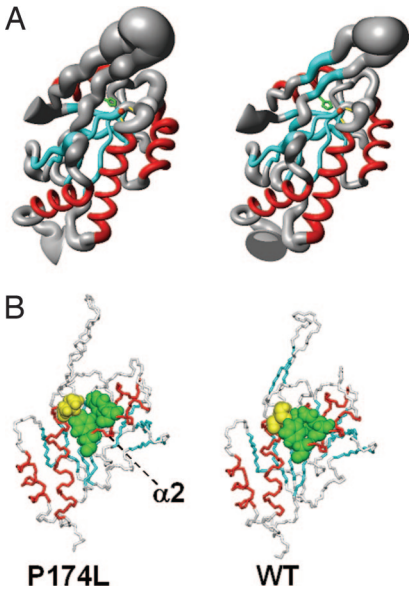



**Fig. 2.** Solution structure of $Cu_1$(I)P174L-HSco1 mutant. (*A*) Backbone atoms are represented as a tube with variable radius, proportional to the backbone rmsd value of each residue for $Cu_1$(I)P174L-HSco1 and $Cu_1$(I)WT-HSco1. The side chains of Cys-169, Cys-173, His-260, and the copper(I) ions are shown in yellow, green, and orange, respectively. The secondary structure elements also are indicated: $\beta$-strands are in cyan and $\alpha$-helices in red. (*B*) Side-chain packing involving Pro-174 or Leu-174 (in yellow) and Tyr-216 and Phe-220 (in green) is shown on the WT and P174L mean minimized solution structures. The mutant protein was copper-saturated resulting in a 1:1 copper:protein complex.

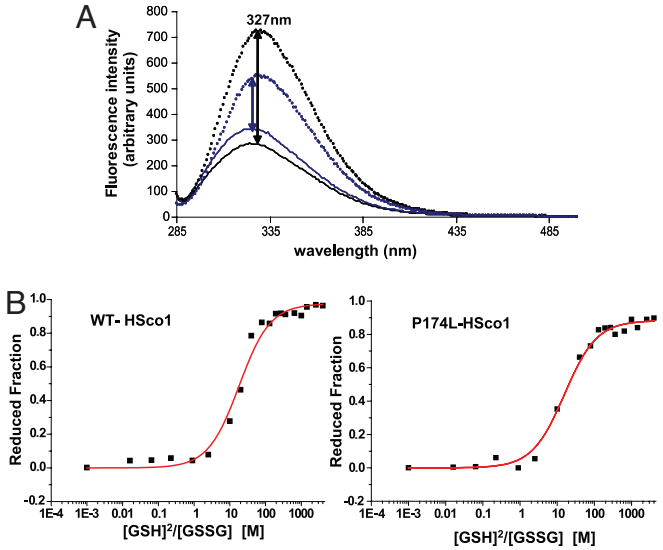



**Fig. 3.** Redox properties of WT-HSco1 and P174L-HSco1. (*A*) Fluorescence emission spectra of 5 $\mu$M WT-HSco1 (black) and P174L-HSco1 (blue) recorded under native conditions of the oxidized (50 mM phosphate buffer, pH 7.0, 0.01 mM GSSG; dotted lines) and the reduced (50 mM phosphate buffer, pH 7.0, 200 mM GSH; solid lines) protein after excitation at 280 nm. (*B*) Redox equilibrium of WT-HSco1 and P174L-HSco1 with different $[GSH]^2$/GSSG ratios as followed by measuring fluorescence emission at 327 nm. After nonlinear regression, values of $K_{eq}$ = 17.70 ± 2.23 mM and of 15.66 ± 1.86 mM (correlation coefficient: 0.986 and 0.996) were obtained for the WT-HSco1 and P174L-HSco1/glutathione systems, corresponding to a redox potentials of −0.277 ± 0.030 V and −0.275 ± 0.028 V, respectively, by using the glutathione standard potential of −0.240 V at pH 7.0 and 298 K.

region through hydrophobic interactions involving Leu-174 or Pro-174 and Leu-177 (helix $\alpha$1), and Tyr-216 and Phe-220 (helix $\alpha$2). Although Pro-174 has extensive hydrophobic contacts with both Tyr-216 and Phe-220 in $Cu_1$(I)WT-HSco1, thus forming a compact hydrophobic patch, in the mutant, the bulkier Leu side chain prevents the optimal packing with the two aromatic side chains (Fig. 2), which indeed display double conformation. Leu-174 essentially interacts only with Phe-220 (Fig. 2) and is more solvent-exposed (25%) than Pro-174 is (12%). The lack of a well organized hydrophobic core on this side of the metal-binding region also causes the presence of a double conformation for His-260, located in loop 8. Two patterns for the aromatic ring of this His are observed in the $^2J_{NH}$ coupling-based $^1H$-$^{15}N$ HSQC spectra (SI Fig. 7). One pattern, belonging to the minor species, fully matches the shifts of His-260 in $Cu_1$(I)WT-HSco1, whereas the other has different shifts but still is indicative of copper(I) coordination through N$\varepsilon$2 as in the WT protein. In the major species, His-260 lacks NOEs contacts with the CXXXC metal-binding region that are present in the WT form. A decreased number of NOEs with respect to WT protein also is observed in the metal-binding loop 3 and in the surrounding loops 5, 7, and 8, likely reflecting the effect of the observed conformational disorder. From this structural investigation, we thus can conclude that the mutation (*i*) perturbs the hydrophobic interactions and (*ii*) induces conformational heterogeneity around the copper-binding site, with both effects overall reducing the affinity of the protein for copper(I).

**Redox Properties of the Pathogenic Mutant P174L-HSco1.** Based on its thioredoxin fold, it has been suggested that the two Cys residues of the CXXXC conserved motif in Sco protein family can be involved in redox reactions (13–15), perhaps in the reduction of the oxidized cysteines in $Cu_A$ site of CcO (16). Oxidized and reduced forms of apoHSco1 are characterized by different fluorescence spectra (Fig. 3). Trp-159 and, to a lesser extent, Tyr-163 are the residues principally contributing to the difference of florescence intensity upon change of the oxidation state of the protein (see *SI Text* for details), as assessed by comparing the fluorescence spectra of W159F, W159F/Y163F, and Y216F mutants with that of WT-HSco1 (SI Fig. 8). Tyr-163 is close to the CXXXC motif, whereas Trp-159 is far but still belonging to the same $\beta$-strand of Tyr-163. From the difference in fluorescence of the two redox states of the protein, the reduction potential of the two cysteines for WT- and P174L-HSco1 in their apo forms can be determined. The measured reduction potentials at pH 7.2 are −0.277 ± 0.030 V and −0.275 ± 0.028 V for WT-HSco1 and P174L-HSco1, respectively (Fig. 3), falling in the standard state redox potentials range of the thioredoxin family of thiol-disulfide oxidoreductases (−124 to −270 mV) (22), consistently with a possible function of HSco1 as a thioredoxin. The similar reduction potentials of WT-HSco1 and P174L-HSco1 also indicate that the pathogenic mutation does not significantly affect the thermodynamic equilibrium between reduced and oxidized HSco1. However, the kinetic rates of disulfide bond formation and reduction are significantly different in WT-HSco1 and P174L-HSco1 forms. Air oxidation of cysteines in apoWT-HSco1 has an half-life time of ≈4 h. Oxidation of apoP174L-HSco1 is a much slower process, because ≈70% of the reduced form still is present after 4 h of air exposure (SI Fig. 9). Reduction of the disulfide bond in oxidized apoHSco1, by 1 mM DTT, also has different rates in the two proteins, being much faster in WT-HSco1 than in the mutant. In the former, the reduction of the disulfide bond is completed already after 1 min (SI Fig. 10); however, after the same period, only 20% of oxidized apoP174L-HSco1 is reduced, and, after 5 min, a reduction level of only 55% is reached (SI Fig. 10). In summary, P174L-HSco1 mutant has a similar redox potential as compared with WT protein but significantly slower reduction and oxidation rates.

**Transfer of Copper from Cox17 to HSco1.** After the characterization of the structural and redox properties of the P174L-HSco1 mutant, we also addressed and analyzed the mechanism of copper acquisition from the copper chaperone Cox17, which also could be affected by the mutation. For this analyses, we have used porcine Cox17, which differs from human Cox17 only in the first four amino acid residues. Cox17 could exist in three different oxidation states, which have different metal-binding properties (23). In oxidative environment, a completely oxidized form with three disulfide bonds ($Cox17_{3S\text{-}S}$ hereafter) is present that is unable to bind copper. In the presence of 1 mM DTT, Cox17 is partially oxidized with two disulfide bonds ($Cox17_{2S\text{-}S}$) and two reduced Cys residues, which bind one Cu(I) ion [$Cu_1(I)Cox17_{2S\text{-}S}$ hereafter]. In more reducing conditions, Cox17 exists in the fully reduced state, which binds cooperatively four Cu(I) ions forming a tetracopper-thiolate cluster [$Cu_4(I)Cox17$ hereafter] (23). In principle, both metallated forms of Cox17 could transfer metals to HSco1 protein.

To test which form of Cox17 can transfer Cu(I) to WT-HSco1, we have produced a mixture of $Cu_4(I)Cox17$ and $Cu_1(I)Cox17_{2S\text{-}S}$ forms and mixed it to reduced apoWT-HSco1. ESI-MS data of the Cox17 mixture showed metalation of apoWT-HSco1 with one Cu(I) ion and the concomitant decrease of the peak of $Cu_1(I)Cox17_{2S\text{-}S}$, whereas that of $Cu_4(I)Cox17$ remains unaltered (SI Fig. 11). This finding indicates that $Cu_1(I)Cox17_{2S\text{-}S}$ preferentially transfers Cu(I) to apoWT-HSco1. All of the subsequent NMR and ESI-MS experiments thus have been performed with the $Cu_1(I)Cox17_{2S\text{-}S}$ form.

Competition experiments for Cu(I) between the $Cox17_{2S\text{-}S}$ protein and the Cu(I) chelator DTT were carried out with ESI-MS technique as described in ref. 23. From these data, it results that $Cox17_{2S\text{-}S}$ bind one Cu(I) ion tightly with a $K_D$ value of $6.4 \pm 0.6 \times 10^{-15}$ M (SI Fig. 12).

When $^{15}N$-labeled apoWT-HSco1 is titrated with increasing amounts of unlabelled $Cu_1(I)Cox17_{2S\text{-}S}$ in the presence of 1 mM DTT, the intensity of some of its NH signals decreases in the $^1H$-$^{15}N$ HSQC spectra, and, concomitantly, NH signals of the Cu(I)WT-HSco1 species appeared with increasing intensities along the titration (SI Fig. 13). No additional signals from a possible (transiently populated) intermediate could be detected at any point of the titration. ApoWT-HSco1 reaches 90% of its metalation upon addition of 1 eq of $Cu_1(I)Cox17_{2S\text{-}S}$, indicating a lower limit of the equilibrium constant of $10^2$. Considering the latter value and the $K_D$ value of $Cu_1(I)Cox17_{2S\text{-}S}$ ($K_D = 6.4 \pm 0.6 \times 10^{-15}$ M), $K_D$ of Cu(I)WT-HSco1 is estimated to be $\approx 10^{-17}$ or lower, thus WT protein displaying a higher affinity for Cu(I) compared with the mutant ($K_D = 3.2 \pm 0.6 \times 10^{-13}$ M). The rotational correlation time ($13.2 \pm 0.9$ ns) of WT-HSco1 in the 1:1 mixture with $Cu_1(I)Cox17_{2S\text{-}S}$ is not higher than that of isolated apo or Cu(I)WT-HSco1 ($13.8 \pm 1.6$ and $14.5 \pm 1.1$ ns, respectively), indicating that no complex is present at detectable concentrations. The copper(I) transfer process is slow on the chemical-shift time scale, setting a lower limit for the equilibration rate of $\approx 10^2\ s^{-1}$ [determined by the smallest chemical-shift difference between apoWT-HSco1 and Cu(I)WT-HSco1 that can be detected, i.e., $\approx$0.1 ppm at 800 MHz]. The comparison of the $^1H$-$^{15}N$ HSQC spectra of the Cu(I)WT-HSco1 metalated via $Cu_1(I)Cox17_{2S\text{-}S}$ with that of Cu(I)WT-HSco1 metallated via Cu(I) acetonitrile complex (SI Fig. 14) shows that a few NH resonances (residues 164, 243, 245, 258, and 260–262) close to the metal-binding site are missing, broader, or having two conformational states in the former spectrum, indicating that Cox17 is transiently interacting with this region of WT-HSco1 producing conformational exchange processes of these NH resonances.

ESI-MS spectrometry confirms the NMR data, showing that, by adding an excess of $apoCox17_{2S\text{-}S}$, the apoWT-HSco1/Cu(I)WT-HSco1 ratio does not change (SI Fig. 15 *C* versus *B*), indicating that Cu(I) ion is bound to WT-HSco1 with higher affinity than to $Cu_1(I)Cox17_{2S\text{-}S}$. According to the different metal-binding affinities, from ESI-MS experiments, we also established that even 10 mM DTT cannot extract Cu(I) from Cu(I)WT-HSco1, whereas already 5 mM DTT extracts $\approx$40% of copper from $Cu_1(I)Cox17_{2S\text{-}S}$. Therefore, $Cox17_{2S\text{-}S}$ does not compete with WT-HSco1 for Cu(I) ions, but vice versa, it contributes to metalation of WT-HSco1 protein as it follows from the fact that $Cox17_{2S\text{-}S}$ shifts the metalation equilibrium of HSco1 toward the Cu(I)WT-HSco1 form more than the Cu(I)DTT complex does (SI Fig. 15 *A* and *B*). These data are in agreement with the copper content of human Sco1, determined from the cytoplasm of yeast cells overexpressing human Sco1 in the presence and absence of co-overexpression of human Cox17 (9).

ESI-MS spectra also indicate that, in the absence of metal ion, WT-HSco1 and Cox17 proteins do not form any complex. However, in the presence of 1 eq of Cu(I) there are two minor twin-peaks in the ESI-MS spectra (SI Fig. 15*D*) whose deconvolution yields molecular masses of 26,497 and 26,562 Da, which correspond to WT-HSco1/$Cu_1(I)$/$Cox17_{2S\text{-}S}$ and WT-HSco1/$Cu_2(I)$/$Cox17_{2S\text{-}S}$ complexes (theoretical $M_r$ 26,499.8 and 26,563.4 Da, respectively). These peaks have very low intensity, indicating that these complexes are transient and not highly populated states, in agreement with NMR results obtained at millimolar concentration. However, the relative intensity of these minor peaks increases at higher concentration of proteins, indicating that an equilibrium between free proteins and protein complexes is present. Moreover, at substoichiometric concentrations of Cu(I) ions, WT-HSco1/$Cu_1(I)$/$Cox17_{2S\text{-}S}$ is prevalent, whereas at higher Cu(I) concentrations, the WT-HSco1/$Cu_2(I)$/$Cox17_{2S\text{-}S}$ form prevails. This is another example of formation of a metal-mediated protein–protein complex, also observed in the case of other copper chaperones (24–26). As a whole, the ESI-MS results indicate that WT-HSco1/$Cu_1(I)$/$Cox17_{2S\text{-}S}$ is the transient intermediate involved in metal transfer from $Cu_1(I)Cox17_{2S\text{-}S}$ to WT-HSco1. At higher metal concentrations, WT-HSco1/$Cu_2(I)$/$Cox17_{2S\text{-}S}$ complex also exists, which suggests that the WT-HSco1/$Cox17_{2S\text{-}S}$ complex can accommodate two Cu(I) ions.

When $Cu_1(I)Cox17_{2S\text{-}S}$ is added to apoP174L-HSco1, differently from what happens for WT-HSco1, no copper transfer occurs at 1:1 $Cu_1(I)Cox17_{2S\text{-}S}$/P174L-HSco1 ratio, as judged by the $^1H$-$^{15}N$ HSQC spectrum of the mixture. By increasing the $Cu_1(I)Cox17_{2S\text{-}S}$/P174L-HSco1 ratio up to 2:1, copper transfer occurs slowly on the chemical-shift time scale, similarly to the WT protein, but only $\approx$20% of P174L-HSco1 is metallated, accordingly to the higher affinity of $Cox17_{2S\text{-}S}$ for copper(I) [$K_D = 6.4 \pm 0.6 \times 10^{-15}$ M for $Cu_1(I)Cox17_{2S\text{-}S}$; $K_D = 3.2 \pm 0.6 \times 10^{-13}$ M for Cu(I)P174L-HSco1]. ESI-MS spectra also show that the addition of $Cox17_{2S\text{-}S}$ up to 3 eq does not improve the metalation of P174L-HSco1 (Fig. 1 *E* versus *D*), which is different from what occurs with the WT protein. Addition of $Cox17_{2S\text{-}S}$ also does not change the charge-state distribution of Cu(I)P174L-HSco1 (SI Fig. 16), which remains similar to that of apoP174L-HSco1 (Fig. 1*E*). Similarly to WT-HSco1, no protein–protein complexes are observed in ESI-MS spectra in the absence of metal, and a minor twin peak, corresponding to P174L-HSco1/$Cu_1(I)$/$Cox17_{2S\text{-}S}$ and P174L-HSco1/$Cu_2(I)$/$Cox17_{2S\text{-}S}$ complexes, is detected in ESI-MS spectrum in presence of Cu(I) ions (SI Fig. 16), which indicates that transient protein–protein complexes with 1 and 2 Cu(I) ions also do exist in case of the mutant protein. All of the data indicate that, even if Cox17 is still capable of interacting and exchanging copper(I) with the P174L-HSco1 mutant, the latter is not efficiently metallated into a compact copper(I) form, at variance with what occurs in WT protein.

## Discussion

Copper incorporation into CcO is biologically crucial; it is, however, an extremely complex process, which is tightly regulated and requires a large array of proteins, each of them with

a specific, in most of the cases, nonreplaceable role. The comprehension and description of these processes are still in their infancy, and only now knowledge about the structure and functional role of the assisting proteins has begun to become available. Many lines of evidence show that copper insertion into $Cu_A$ site of COXII subunit is not a simple, passive copper transfer from Cox17 to Sco1/2 proteins and finally to COXII, but redox reactions are interlinked with the process of copper transfer (12, 15, 16, 27).

In this work, we have shown that the P174L mutation in HSco1 significantly affects the redox and metal-binding properties of this protein. Structurally, the mutation produces the lack of well organized hydrophobic contacts and a structural heterogeneity in the vicinity of the metal-binding region. The local structural changes induced by the point mutation decrease the copper(I) affinity of the mutant, thus negatively affecting the cometallochaperone function of HSco1 in the $Cu_A$ assembly. The mutation also determines that the oxidation and reduction rates of the disulphide bond of the mutant are much slower than those of the WT protein. It is likely that introduction of the bulky Leu-174 disturbs the mutual orientation of cysteines and slows down the local conformational changes necessary for disulfide bond formation and disruption. The thioredoxin role of HSco1, proposed for maintaining the receiving $Cu_A$ site in the suitable reduced state (15, 16), therefore is significantly perturbed because the mutation could slow down the reduction of the copper-binding cysteines in $Cu_A$ site of CcO. Overall, these alterations induced by the mutation on copper-binding and redox properties of HSco1 can sizably decrease the efficiency of the copper transfer from Cu(I)HSco1 to the $Cu_A$ site of CcO.

Our results also provide a detailed understanding of the metalation process of HSco1 by Cox17 and show how the mutation at position 174 of HSco1 affects this process. ESI-MS data demonstrate that $Cu_1(I)Cox17_{2S\text{-}S}$, but not $Cu_4(I)Cox17$, is capable of transferring Cu(I) to apoWT-HSco1 and that minor amounts of the metal-bridged species, WT-HSco1/$Cu_1(I)$/$Cox17_{2S\text{-}S}$ and WT-HSco1/$Cu_2(I)$/$Cox17_{2S\text{-}S}$, do occur, which are probably the transient intermediates involved in the metal transfer from $Cu_1(I)Cox17_{2S\text{-}S}$ to apoWT-HSco1. Such transient interaction with $Cox17_{2S\text{-}S}$ assists WT-HSco1 to pass from the open to the closed state, which binds Cu(I) ions extremely strongly (Fig. 4). The mutant P174L-HSco1 behaves differently: the conformational disorder experienced by Cu(I)P174L-HSco1 indeed is associated with a weakly bound metal (Fig. 4), and $Cox17_{2S\text{-}S}$ is not able to assist the mutant P174L-HSco1 to form the more compact state as it happens with WT-HSco1.

In conclusion, the P174L mutation in HSco1 decreases the efficiency of both metal-binding and redox properties of HSco1, which both are potentially involved in HSco1-mediated transfer of the copper and electrons from Cox17 to the $Cu_A$ site of COXII (16).

## Methods

**Protein Expression and Characterization.** The pathogenic mutation at the gene position C520T, corresponding to P174L amino acid change, was made by using QuikChange mutagenesis kit (Stratagene, La Jolla, CA). WT-HSco1 protein and P174L-HSco1 pathogenic mutant were produced in *E. coli* BL21-Gold(DE3) (Stratagene) following an already reported protocol (16). Porcine Cox17 was isolated from porcine intestine as fully oxidized apo protein ($Cox17_{3S\text{-}S}$) and lyophilized (23).

The oxidation state of the cysteines of the WT- and P174L-HSco1 proteins was evaluated through the selective reaction of the free thiol groups with 4-acetamido-4-maleimidylstilbene-2,2-disulfonic acid, which adds ≈500 Da per reactive thiol to the total mass, thus shifting the mobility of the protein on a SDS/PAGE denaturing gel. The metal content was determined by inductively coupled plasma mass spectrometry. Far-UV CD spectra (185–260 nm) were recorded on a Jasco J-810 spectropolarimeter on 10 $\mu$M protein solutions in 20 mM phosphate buffer, pH 7.2, at 298 K.

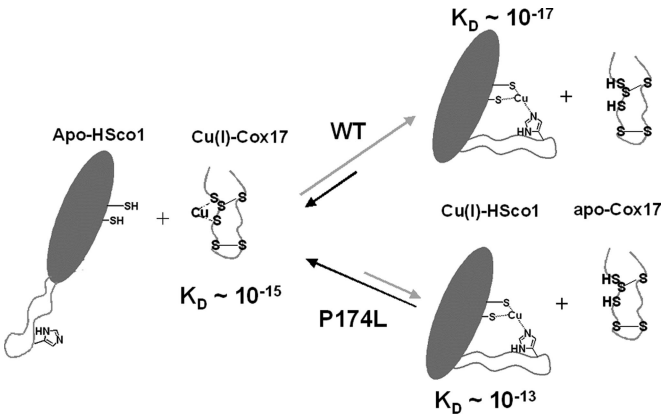



**Fig. 4.** The effect of the P174L-Sco1 mutation in the copper transfer with Cox17. The pathogenic human Sco1 mutation P174L, next to the copper-binding motif, determines a reduction of the efficiency of the copper transfer from the mitochondrial metallochaperone Cox17 to Sco1.

The melting temperatures of WT- and P174L-HSco1 proteins in the apo state were determined through the fluorescence-based thermal-shift assay (28). Fluorescence intensity was measured at different temperatures with excitation/emission wavelengths of 490 and 575 nm, respectively.

Redox equilibration between WT- or P174L-HSco1 and glutathione was followed through the change of fluorescence monitored on 5 $\mu$M concentrations of protein samples previously incubated for at least 6 h under an $N_2$ atmosphere, with different reduced/oxidized glutathione (GSH/GSSG) ratios (0.01 mM GSSG and varying concentrations of GSH, 0.1–200 mM) in 50 mM phosphate buffer, pH 7.2. The collection and data analysis was performed by following a protocol previously described for thioredoxins (29).

**Preparation of Cox17, HSco1, and P174L-HSco1 Samples for MS Measurements.** For MS experiments, purified WT- and P174L-HSco1 proteins containing one disulfide bond were brought into 50 mM ammonium acetate, pH 7.5, by using HiPrep26/10 desalting column (Amersham Biosciences, Uppsala, Sweden). Lyophilized $Cox17_{3S\text{-}S}$ was dissolved in argon-saturated 50 mM ammonium acetate, pH 7.5, at 75 $\mu$M concentration, and this stock solution was used for further experiments. $Cox17_{2S\text{-}S}$ was produced by reduction of $Cox17_{3S\text{-}S}$ with 1 mM DTT at 298 K (incubation time 2 min), which leads to reduction of one most labile disulfide bond in $Cox17_{3S\text{-}S}$. Fully reduced Cox17 was prepared by incubation of $Cox17_{3S\text{-}S}$ with 1 mM DTT at 45°C for 100 min.

**MALDI-TOF and ESI-MS Protein Characterization.** In MALDI-TOF experiments, reduction of oxidized apoWT-HSco1 and apoP174L-HSco1 (2 $\mu$M protein concentration) by 1 mM DTT was followed in 50 mM ammonium acetate buffer, pH 7.5, by stopping the reduction with 5 mM iodoacetamide. Protein adducts were identified by MALDI-TOF MS on a Voyager STR instrument (Applied Biosystems, Foster City, CA).

In ESI-MS experiments, 0.8–4.8 $\mu$M concentrations of protein samples or mixtures were infused by a syringe pump at 15 $\mu$l/min into an Ettan API ESI-TOF mass spectrometer (Amersham Biosciences). Mass spectra were recorded during 2 to 3 min at a capillary exit voltage of 150 V. HSco1 and P174L-HSco1 were reduced with 1.0 mM DTT at 298 K. ESI-MS spectrum of P174L-HSco1 also exposed two minor peaks corresponding to incorrectly cleaved protein forms, which did not disturb measurements. Reconstitution of HSco1 and Cox17 forms with

copper was conducted as follows. First, Cu(II) acetate was dissolved at 150 $\mu$M concentration in argon-saturated 50 mM ammonium acetate, pH 7.5, and Cu(II) was reduced to Cu(I) by addition of 1.0 mM DTT. Different equivalents of freshly prepared Cu(I)DTT complex were added to HSco1 or Cox17 and incubated in buffer with 1.0 mM DTT, and the mixture was incubated for additional 1 min at 25°C, and ESI-MS spectra were recorded as described above. In Cox17/HSco1 titration experiments, $Cu_1(I)Cox17_{2S\text{-}S}$ was prepared by addition of Cu(I) to $Cox17_{2S\text{-}S}$, which subsequently was mixed with a sample of HSco1 protein. All buffers contained 1 mM DTT. The dissociation constant of $Cu_1(I)Cox17_{2S\text{-}S}$ complex has been calculated as previously reported (23).

**NMR Protein Characterization.** The fully copper-loaded state of the $Cu_1(I)$P174L-HSco1 NMR sample used for structure determination has been produced by adding 1.5 eq of $[Cu(I)(CH_3CN)_4]PF_6$ complex to reduced apo protein in 50 mM phosphate buffer at pH 7.2 and in the absence of DTT. NMR spectral assignment and structure determination of $Cu_1(I)$P174L-HSco1 were obtained through the experiments listed in SI Table 1. The $^1H$, $^{13}C$, and $^{15}N$ resonance assignments of the $Cu_1(I)$P174L-HSco1 are reported in SI Table 2. After conversion of the NMR data in structural restraints, the structure was calculated by using the program DYANA. The relative intensity of the split resonances was taken into account when converting NOE volumes into distances. The best 30 conformers of the DYANA family were subjected to restrained energy minimization (REM) with AMBER 8.0, by using the metal force field already applied for WT-HSco1 (16). The statistical analysis of the REM family of $Cu_1(I)$P174L-HSco1 structures is reported in SI Table 3. The programs PROCHECK, PROCHECK-NMR (30, 31), and WHATIF (32) were used in the evaluation of the quality of the structures. More than 90% of residues were located in the allowed regions of the Ramachandran plot.

Titrations of $^{15}N$-labeled apoHSco1 (0.1 mM) and apoP174L-HSco1 (0.1 mM) with unlabelled $Cu_1(I)Cox17_{2S\text{-}S}$ or $[Cu(I)(CH_3CN)_4]PF_6$ were performed with NMR spectroscopy following the $^1H$-$^{15}N$ spectral changes in HSQC spectra upon addition of increasing amounts of the unlabelled protein partner or copper(I) complex in the presence of 1 mM DTT. Aliquots were added in a Coy chamber under nitrogen atmosphere at 298 K. The metalation state of WT-HSco1 and P174L-HSco1 proteins was monitored along the NMR titration through a few residues that are next to the copper(I) binding motif CXXXC and therefore are experiencing chemical shifts sensitive to the metalation state. In addition, their NHs are not overlapped in the $^1H$-$^{15}N$ HSQC maps, thus they are easily integrated during the titration steps to estimate the relative population of apo and copper(I) loaded forms, which allows the estimation of the metal affinity. The dissociation constant $K_D$ of Cu(I)P174L-HSco1 is obtained according to the following scheme:

$$\text{Cu(I)P174L-HScol} + \text{DTT} \Leftrightarrow \text{apoP174L-HScol} + \text{Cu(I)DTT}. \quad [1]$$

The conditional dissociation constant for the Cu(I)DTT complex, necessary for the above $K_D$ estimation, is $6.31 \times 10^{-12}$ M at pH = 7.4 and $T$ = 298 K (33).

$R_1$ and $R_2$ $^{15}N$ relaxation rates on 1:1 $Cu_1(I)Cox17$/WT-HSco1 mixture were measured at 298 K on a Bruker (Billerica, MA) Avance 500 spectrometer and then analyzed by using a standard procedure (34). An estimate of the overall rotational correlation time was derived from the measured $R_2/R_1$ ratio.

This work was supported by Marie Curie Host fellowships for early stage research training ("NMR in Inorganic Structural Biology" Fellowship MEST-CT-2004-504391), the European Community (EU-NMR Contract 026145), Estonian Science Foundation Grant 5635, and a grant from Ente Cassa di Risparmio di Firenze (to FiorGen Foundation).

1. Carr HS, Winge DR (2003) *Acc Chem Res* 36:309–316.
2. Nobrega MP, Bandeira SCB, Beers J, Tzagoloff A (2002) *J Biol Chem* 277:40206–40211.
3. Barros MH, Johnson A, Tzagoloff A (2004) *J Biol Chem* 279:31943–31947.
4. Glerum DM, Shtanko A, Tzagoloff A (1996) *J Biol Chem* 271:20531–20535.
5. Beers J, Glerum DM, Tzagoloff A (2002) *J Biol Chem* 277:22185–22190.
6. Lode A, Kuschel M, Paret C, Rodel G (2000) *FEBS Lett* 485:19–24.
7. Mattatall NR, Jazairi J, Hill BC (2000) *J Biol Chem* 275:28802–28809.
8. Andruzzi L, Nakano M, Nilges MJ, Blackburn NJ (2005) *J Am Chem Soc* 127:16548–16558.
9. Horng YC, Leary SC, Cobine PA, Young FB, George GN, Shoubridge EA, Winge DR (2005) *J Biol Chem* 280:34113–34122.
10. Nittis T, George GN, Winge DR (2001) *J Biol Chem* 276:42520–42526.
11. Horng YC, Cobine PA, Maxfield AB, Carr HS, Winge DR (2004) *J Biol Chem* 279:35334–35340.
12. Arnesano F, Banci L, Bertini I, Martinelli M (2005) *J Proteome Res* 4:63–70.
13. Balatri E, Banci L, Bertini I, Cantini F, Ciofi-Baffoni S (2003) *Structure (London)* 11:1431–1443.
14. Williams JC, Sue C, Banting GS, Yang H, Glerum DM, Hendrickson WA, Schon EA (2005) *J Biol Chem* 280:15202–15211.
15. Ye Q, Imriskova-Sosova I, Hill BC, Jia Z (2005) *Biochemistry* 44:2934–2942.
16. Banci L, Bertini I, Calderone V, Ciofi-Baffoni S, Mangani S, Martinelli M, Palumaa P, Wang S (2006) *Proc Natl Acad Sci USA* 103:8595–8600.
17. Shoubridge EA (2001) *Am J Med Genet* 106:46–52.
18. Valnot I, Osmond S, Gigarel N, Mehaye B, Amiel J, Cormier-Daire V, Munnich A, Bonnefont JP, Rustin P, Rotig A (2000) *Am J Hum Genet* 67:1104–1109.
19. Paret C, Lode A, Krause-Buchholz U, Rodel G (2000) *Biochem Biophys Res Commun* 279:341–347.
20. Cobine PA, Pierrel F, Leary SC, Sasarman F, Horng YC, Shoubridge EA, Winge DR (2006) *J Biol Chem* 281:12270–12276.
21. Eyles SJ, Kalatshov IA (2004) *Methods* 34:88–99.
22. Aslund F, Berndt KD, Holmgren A (1997) *J Biol Chem* 272:30780–30786.
23. Palumaa P, Kangur L, Voronova A, Sillard R (2004) *Biochem J* 382:307–314.
24. van Dongen EM, Klomp LW, Merkx M (2004) *Biochem Biophys Res Commun* 323:789–795.
25. Banci L, Bertini I, Cantini F, Felli IC, Gonnelli L, Hadjiliadis N, Pierattelli R, Rosato A, Voulgaris P (2006) *Nat Chem Biol* 2:367–368.
26. Strausak D, Howie MK, Firth SD, Schlicksupp A, Pipkorn R, Multhaup G, Mercer JF (2003) *J Biol Chem* 278:20821–20827.
27. Imriskova-Sosova I, Andrews D, Yam K, Davidson D, Yachnin Y, Hill BC (2006) *Biochemistry* 44:16949–16956.
28. Lo MC, Aulabaugh A, Jin G, Cowling, R, Bard J, Malamas M, Ellestad G (2004) *Anal Biochem* 332:153–159.
29. Haugstetter J, Blicher T, Ellgaar L (2005) *J Biol Chem* 280:8371–8380.
30. Laskowski RA, Rullmann JAC, MacArthur MW, Kaptein R, Thornton JM (1996) *J Biomol NMR* 8:477–486.
31. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) *J Appl Crystallogr* 26:283–291.
32. Vriend G (1990) *J Mol Graphics* 8:52–56.
33. Krezel A, Lesniak W, Jezowska-Bojczuc M, Mlynarz P, Brasun J, Kozlowski H, Bal W (2001) *J Inorg Biochem* 84:77–88.
34. Peng JW, Wagner G (1992) *J Magn Reson* 98:308–332.

# SUPPLEMENTARY MATERIAL

## Human Sco1 functional studies and pathological implications of P174L mutant

Lucia Banci, Ivano Bertini, Simone Ciofi-Baffoni, Iliana Leontari,
Manuele Martinelli, Peep Palumaa, Rannar Sillard, Shenlin Wang

**Fluorescence spectral properties of W159F, W159F/Y163F and Y216F Sco1 mutants**

Fluorescence spectra (**Fig. 4** of the Supplementary Material) of apoWT-HSco1 and of W159F, W159F/Y163F and Y216F apoHSco1 mutants (5 μM) were carried out at 298K on Cary 50 Eclipse Spectrophotometer supplied with a Single cell Peltier thermostatted cell holder. After excitation at either 280 nm and 295 nm an emission scan was recorded between 200 nm and 500 nm. The fully reduced protein was obtained incubating the sample for at least 6 h in the presence of 1 mM DTT or 200 mM GSH, while the completely oxidized protein was obtained incubating the sample for at least 6 h in the presence of 0.1 mM GSSG. To avoid air oxidation, the solutions were degassed and subsequently flushed with nitrogen and the sample were manipulated under an $N_2$ atmosphere. Fluorescence spectra of oxidized and reduced HSco1 were measured in 50 mM sodium phosphate, pH 7.0.

**Determination of the Redox Potential of the Active-site Cysteines of Sco1**

The fluorescence properties of HSco1 and the reduction of its active-site disulfide by GSH were used to measure the equilibrium constant between HSco1 and glutathione. After excitation at 280 nm, the fluorescence intensity at 327 nm, for WT-HSco1 and P174L-HSco1, and at 302, for W159F-HSco1, (the wavelengths at which the maximal difference between the emission of oxidized and reduced protein was observed, see **Fig. 4** of the Supplementary Material) was recorded for 60 second and averaged. The redox equilibrium of HSco1 with glutathione is given by Eq. 1 and Eq. 2.

$$\mathbf{HSco1^{SH} + GSSG \rightleftharpoons HSco1^{SS} + 2GSH} \quad \textbf{(Eq. 1)}$$

$$\mathbf{K_{eq} = \frac{[HSco1^{SS}][GSH]^2}{[HSco1^{SH}][GSSG]}} \quad \textbf{(Eq. 2)}$$

Oxidized HSco1 (5 μM or 20 μM) was incubated, under anaerobic conditions, in 50 mM sodium phosphate, pH 7.0 at 298K containing different concentrations of GSH (0.1 mM to 200 mM). 10

uM GSSG was added to all the reaction mixtures. The relative amount of reduced HSco1 (R) at equilibrium was measured by the intrinsic HSco1 fluorescence (F). R was calculated according to Eq. 3, where F is the measured fluorescence intensity (at 327 nm or 303 nm), and $F_{red}$ and $F_{ox}$ are the fluorescence intensities of completely reduced or oxidized Sco1, respectively.

$$R=\frac{F - F_{ox}}{F_{red}-F_{ox}} \qquad \textbf{(Eq. 3)}$$

$F_{red}$ , for the completely reduced protein, was measured at 1 mM DTT or 200mM GSH, and $F_{ox}$ for the completely oxidized protein was measured at 0.1 mM GSSG.

By fitting the data shown in **Fig. 4** and **Fig. 4** of the Supplementary Material by non-linear regression, according to Eq. 4, an equilibrium constant for the reaction shown in Eq. 1 was determined.

$$\mathbf{R=\frac{F - F_{ox}}{F_{ox} - F_{red}} = \frac{[GSH]^2/[GSSG]}{Keq + [GSH]^2/[GSSG]}} \qquad \textbf{(Eq. 4)}$$

The standard redox potential of WT-HSco1, P174L-HSco1 and W159F-HSco1 mutants active-site cysteines was calculated from the Nernst equation (Eq. 5) using the glutathione standard potential of -0.240 V at 298K and pH 7.0.

$$\mathbf{E'_{0(HSco1)} = E'_{0(GSH/GSSG)} - 0.059/2 \log\ Keq} \qquad \textbf{(Eq. 5)}$$

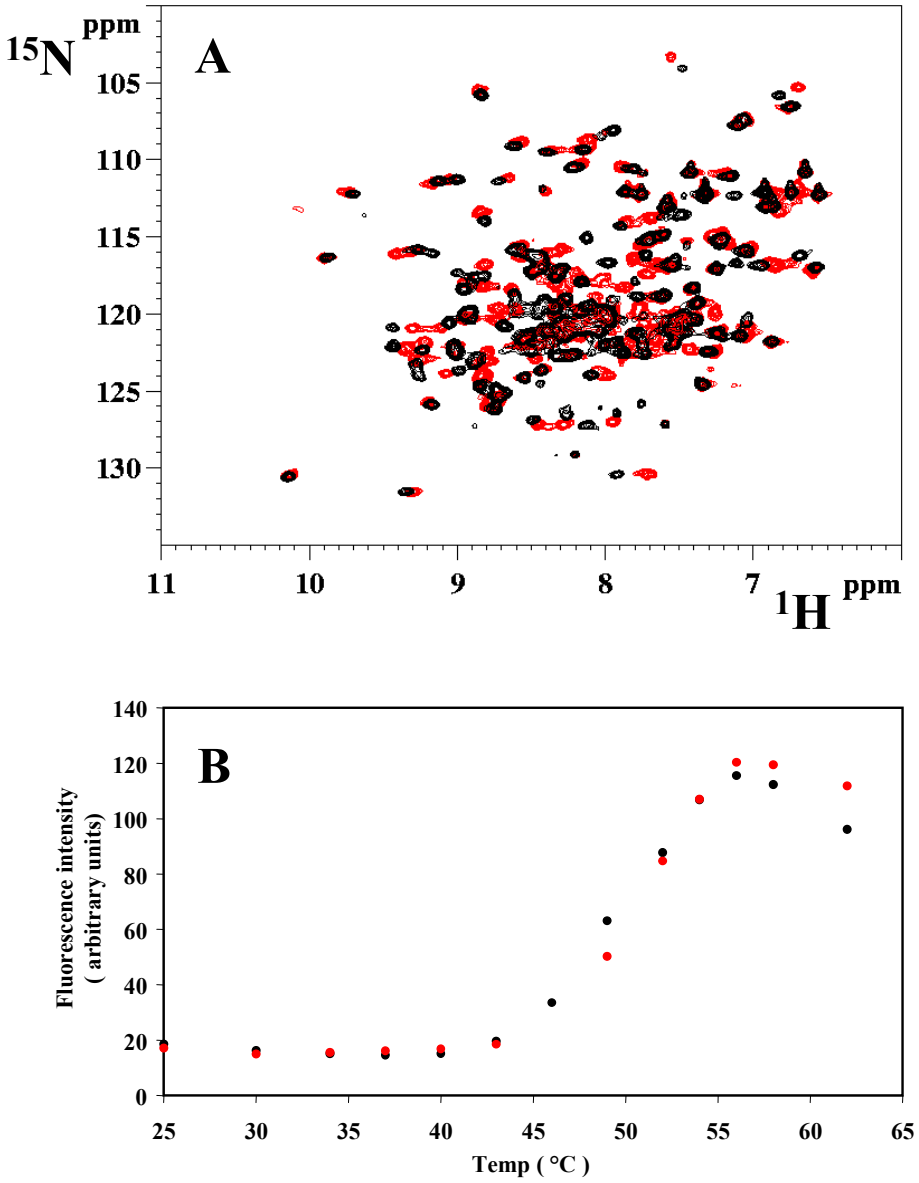



**Fig. 5.** ***Folding and thermal stability properties of WT-HSco1 and P174L-HSco1 mutant***. (A) Superimposition of $^{1}$H-$^{15}$N HSQC spectra of WT-HSco1 (black) and P174L-HSco1 mutant (red) in 50 mM phosphate pH 7.2 at 298K; (B) Fluorimetric thermal unfolding assay of WT-HSco1 (•) and P174L-HSco1 mutant (•) in 20 mM phosphate pH 7.2 containing SYPRO orange dye.

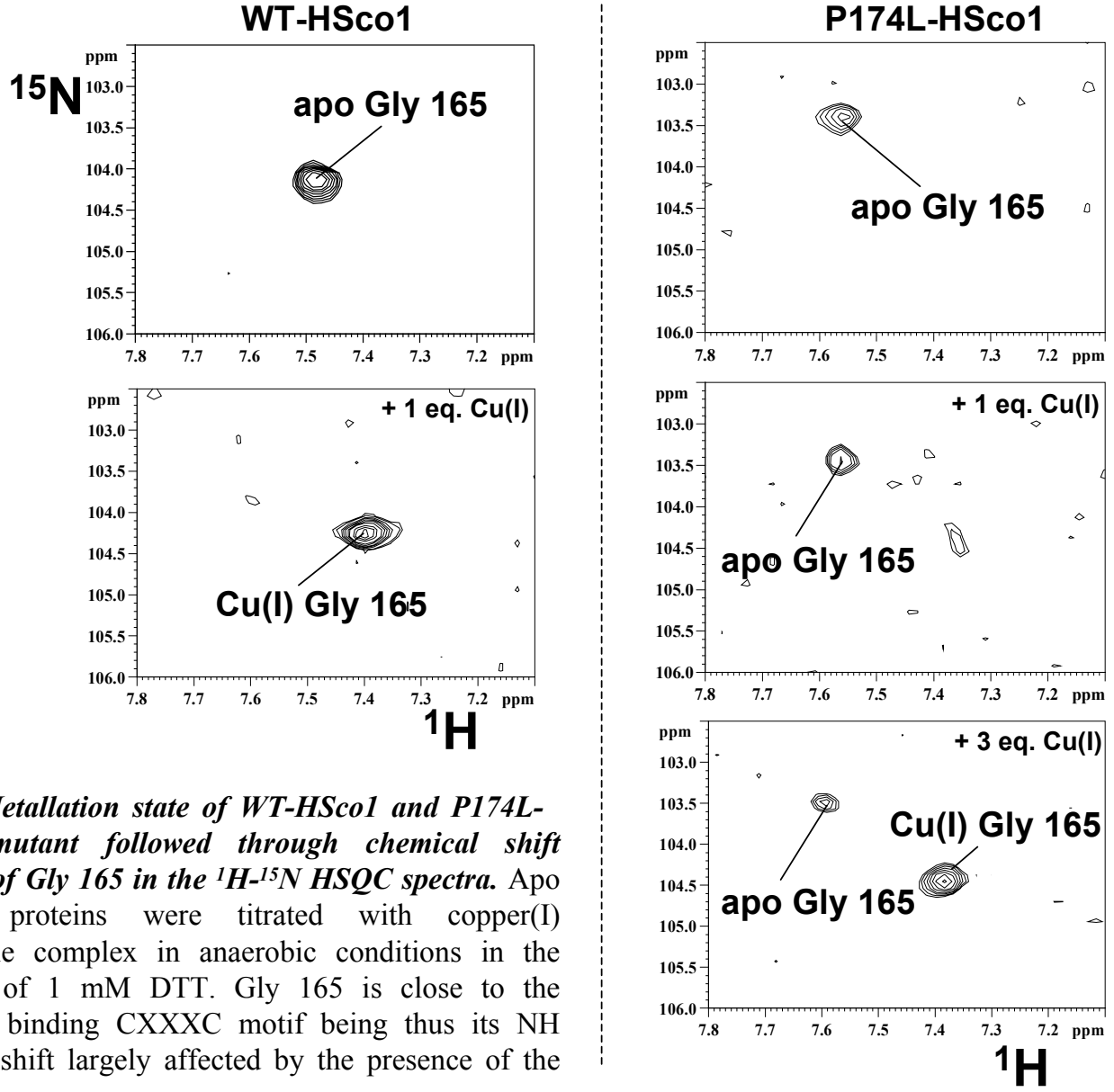



**Fig. 6.** ***Metallation state of WT-HSco1 and P174L-HSco1 mutant followed through chemical shift changes of Gly 165 in the $^{1}H$-$^{15}N$ HSQC spectra.*** Apo reduced proteins were titrated with copper(I) acetonitrile complex in anaerobic conditions in the presence of 1 mM DTT. Gly 165 is close to the copper(I) binding CXXXC motif being thus its NH chemical shift largely affected by the presence of the metal ion.

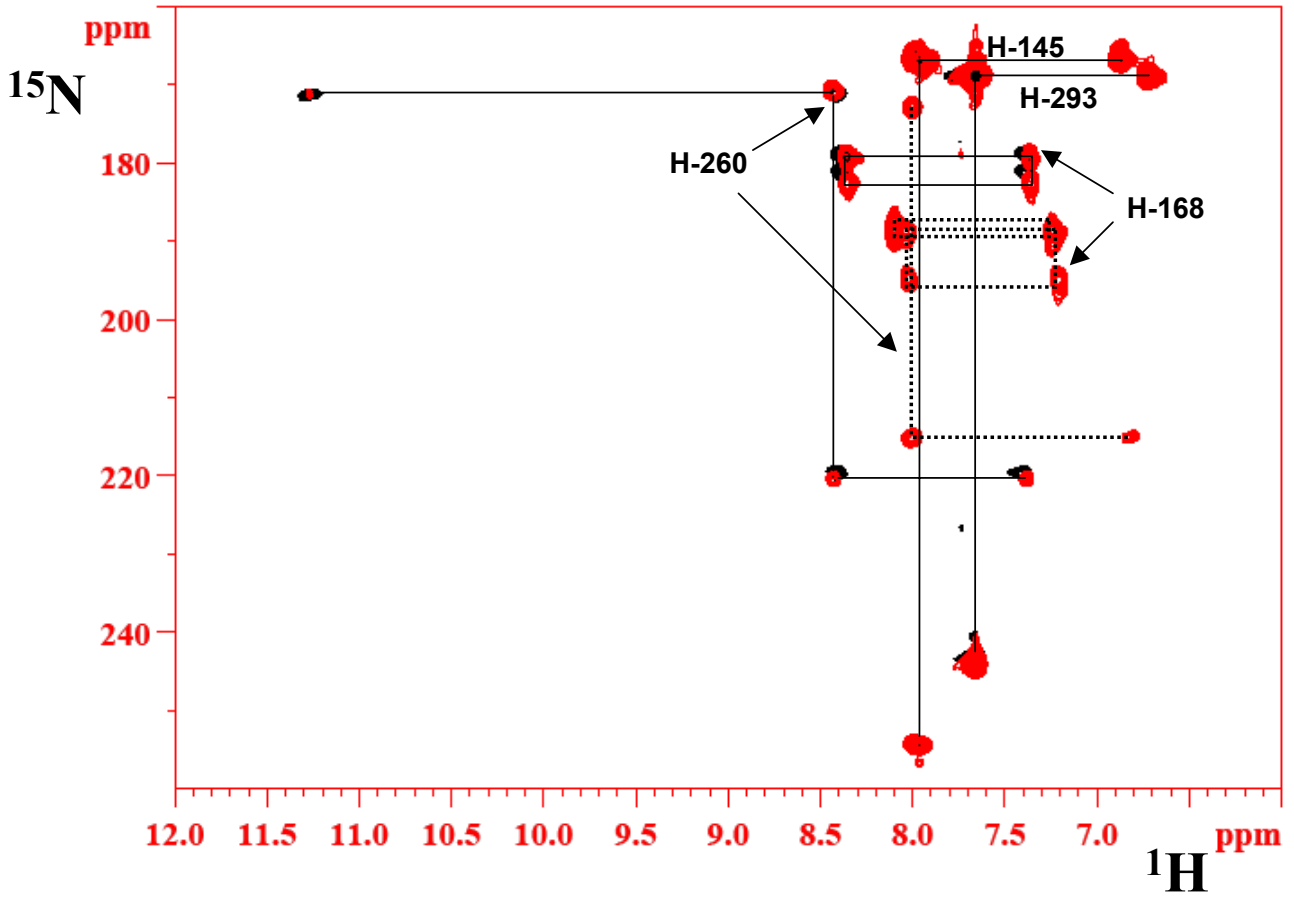



**Fig. 7.** ***$^{1}$H-$^{15}$N HSQC spectra optimized for the detection of $^{2}$J-NH of histidine rings, recorded on Cu(I)P174L-HSco1 (red) and Cu(I)WT-HSco1 (black).***
The spectra were collected at 298 K and pH 7.2 on a spectrometer equipped with cryo-probe and operating at 500 MHz. His 260 and His 168 show the presence of two and three patterns, respectively, one of those matching with those of the WT protein (black straight line).

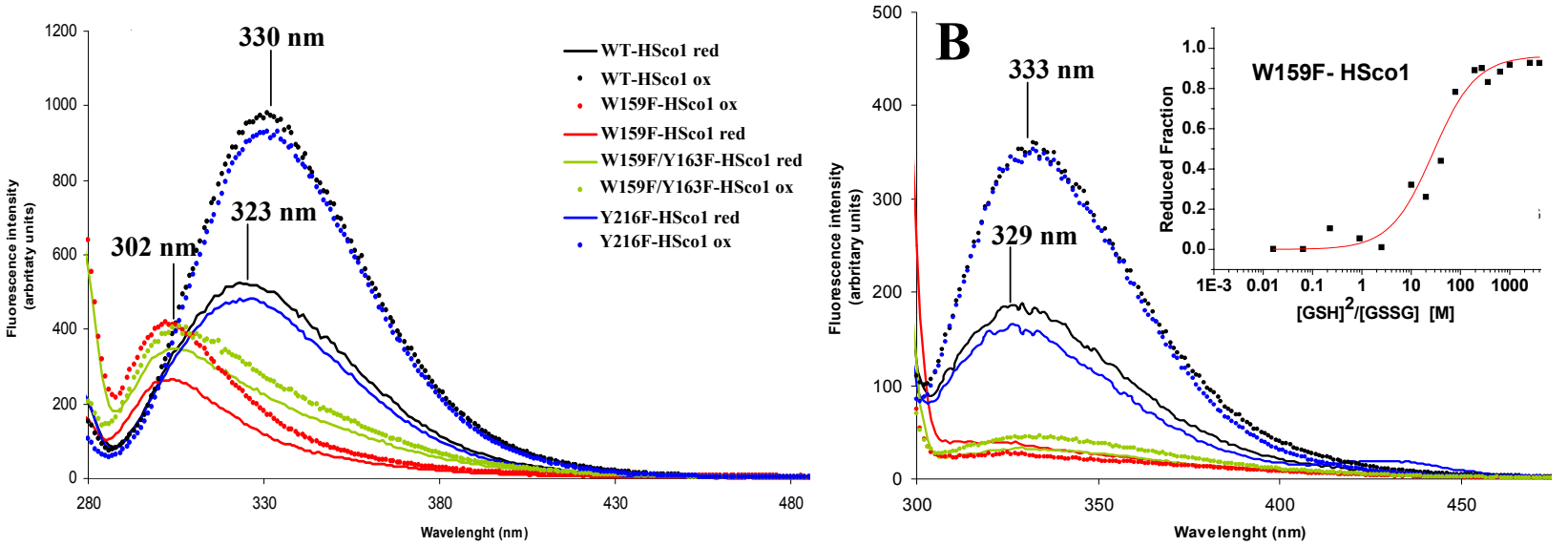



**Fig. 8.** ***Fluorescence emission spectra of W159F, W159F/Y163F and Y216F mutants of HSco1 compared with that of wild-type, recorded under native conditions of the oxidized (50 mM phosphate buffer, pH 7.0, 0.01 mM GSSG; dotted line) and the reduced (50 mM phosphate buffer, pH 7.0, 200 mM GSH; solid line) protein after excitation at 280 (A) and 295 nm (B). In the inset, the redox equilibrium of W159F-HSco1 with different [GSH]$^2$/GSSG ratios as followed by measuring fluorescence emission at 302 nm is shown.***

No significant difference of fluorescence emission is observed in the spectra of the double W159F/Y163F mutant between the oxidized and reduced forms, indicating that W159 and Y163 residues are the main responsible of the change in fluorescence intensity upon Cys reduction. On the contrary, Y216F mutant does not show a meaningful contribution to the fluorescence intensity upon change of Cys redox state. The change in fluorescence intensity at 327 nm, after excitation at 280 nm, used for measuring the reduction potentials of WT and P174L pathogenic mutant, is mainly due to W159 residue. The reduction potential of W159F-HSco1 measured by monitoring the emission at 303 nm (-0.282 ± 0.030 V) is the same of that obtained for the wild-type protein measured at 327 nm, indicating that both W159 and Y163 are suitable to observe the same reduction process of the two cysteine residues.

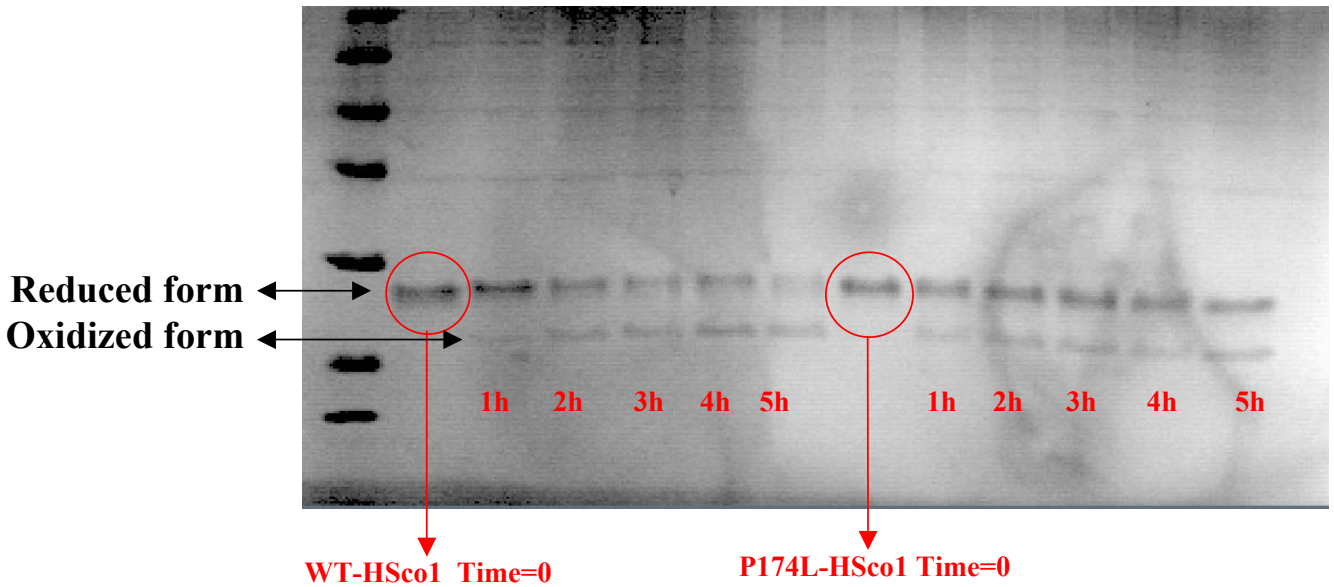



**Fig. 9.** ***Disuplhide oxidation rate of apoWT-HSco1 and apoP174L-HSco1 analyzed by SDS-page.*** The disulfide status is readily evaluated by the selective reaction of the free thiol groups with 4-acetamido-4-maleimidylstilbene-2,2-disulfonic acid, which adds ca.500 Da per reactive thiol and thus shifts the mobility of the protein on a SDS page gel.

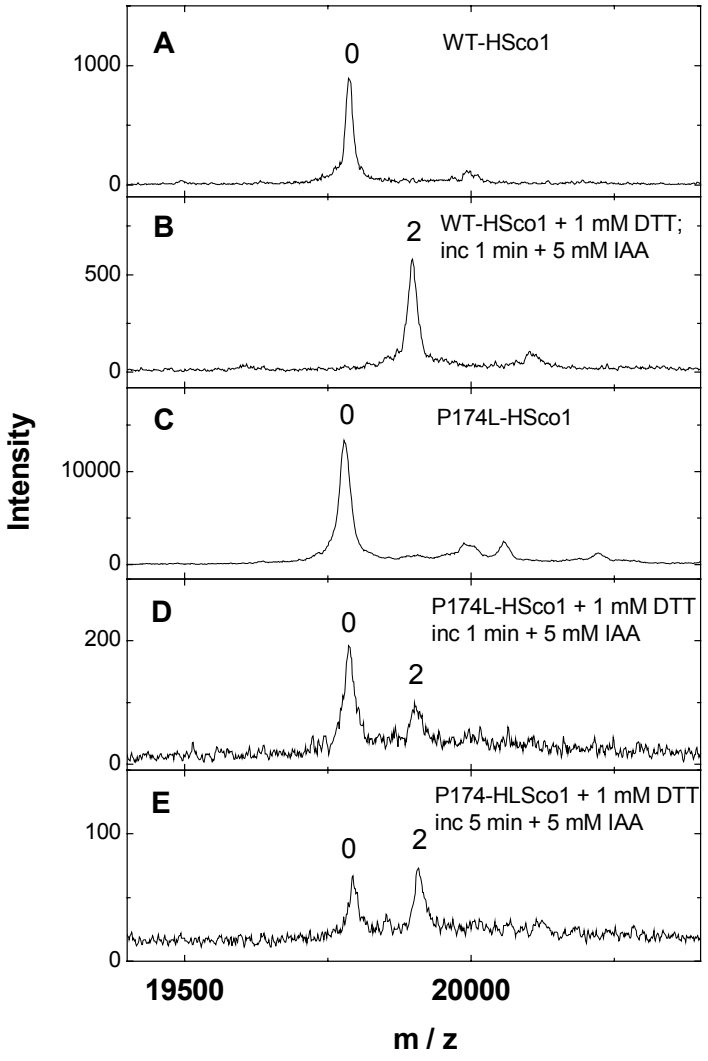



**Fig. 10.** ***Reduction of WT-H Sco1 and P174L-HSco1 mutant with DTT.*** MALDI-TOF mass spectra of WT-HSco1 (2.0 μM) and P174L-HSco1 (2.0 μM) reduced with 1 mM DTT at 25° C in 50 mM ammonium acetate, pH 7.5. Reduction was stopped with addition of 5 mM iodoacetamide. (A) WT-HSco1; (B) WT-HSco1 + 1.0 mM DTT incubated for 1 min; (C) P174L-HSco1; (D) P174L-HSco1 + 1.0 mM DTT incubated for 1 min; (E) P174L-HSco1+ 1.0 mM DTT incubated for 5 min; Numbers on the peaks denote the number of carboxyamidomethyl groups in the protein.

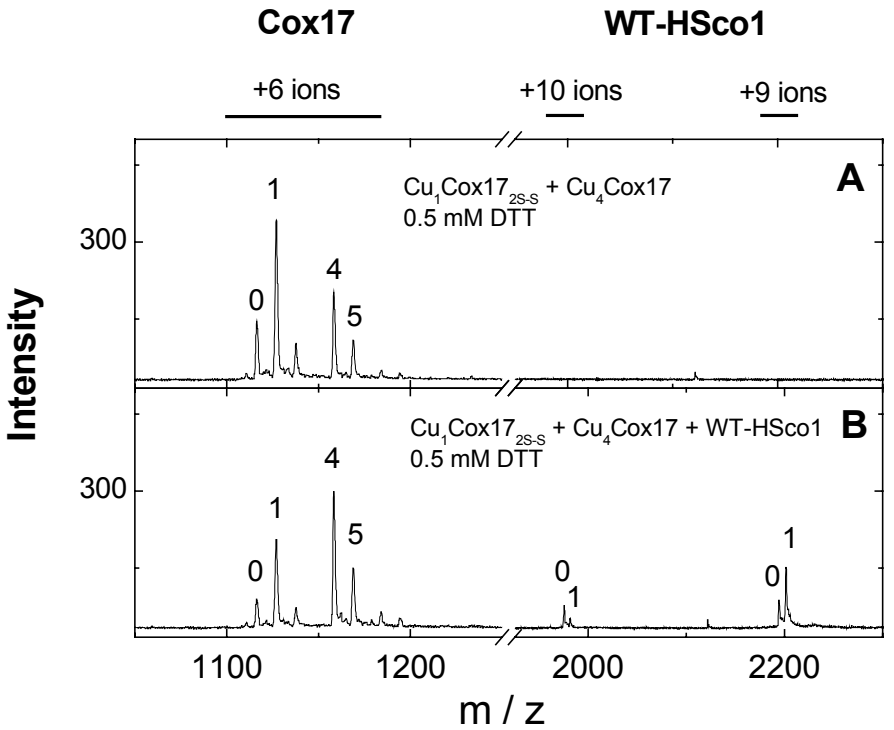



**Fig. 11.** ***Interaction of human WT-Sco1 with different metalloforms of Cox17.*** (A) ESI-TOF mass spectrum of the mixture of $Cox17_{3S\text{-}S}$, $Cu_1Cox17_{2S\text{-}S}$, $Cu_4Cox17$ and $Cu_5Cox17$ (total concentration of Cox17 6.0 μM) in 50 mM ammonium acetate, pH 7.5, 0.5 mM DTT. (B) effect of WT-HSco1 (1.6 μM) on the mixture of $Cox17_{3S\text{-}S}$, $Cu_1Cox17_{2S\text{-}S}$, $Cu_4Cox17$ and $Cu_5Cox17$ (total concentration of Cox17 6 μM) in 50 mM ammonium acetate, pH 7.5, 0.5 mM DTT. Charge state +6 ions are presented for Cox17 and +9 and +10 ions are presented for WT-HSco1. Numbers on the peaks denote the metal stoichiometry of the complex

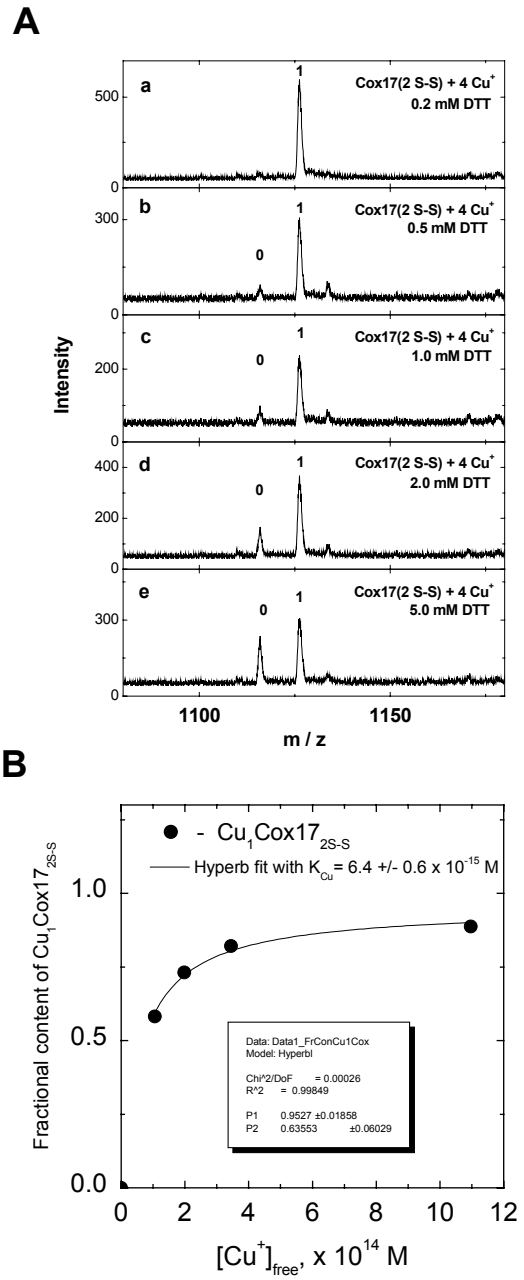



**Fig.12.** ***Binding of Cu(I) ions to Cox17 in the presence of DTT.*** A. ESI mass spectra of Cox17 (1.2 μM) reconstituted with 4 equivalents of Cu(I) ions in the presence of various concentrations of DTT in 20 mM ammonium acetate, pH 7.5, 25ºC. (a) 0.2 mM DTT; (b) 0.5 mM DTT; (c) 1.0 mM DTT; (d) 2 mM DTT; (e) 5 mM DTT. Charge state +6 ions are presented and numbers on the peaks denote the metal stoichiometry of the complex. B. Determination of the dissociation constant for $Cu_1Cox17_{2S-S}$ complex. Dependence of the fractional content of $Cu_1Cox17_{2S-S}$ ($Y = I_{Cu1Cox17} / (I_{Cox17} + I_{Cu1Cox172S-S})$ from the concentration of free Cu(I) ions in metal competition experiment. Solid line shows the fitted curve for hyperbola with $K_{Cu} = 6.4 \times 10^{-15}$ M.

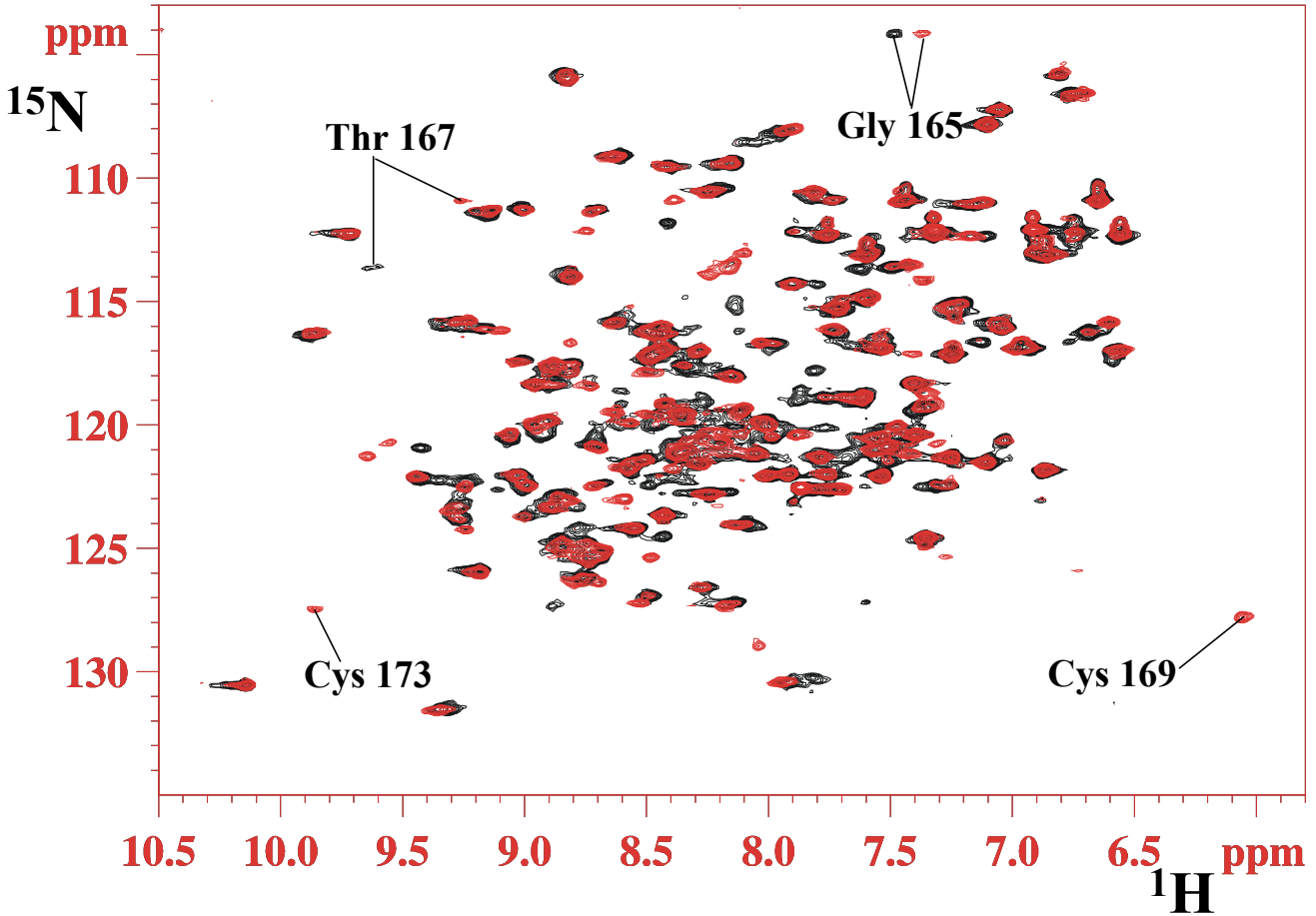



**Fig. 13.** ***$Cu_1$(I)Cox17$_{2S-S}$/WT-HSco1 interaction.*** Superimposition of $^{1}$H-$^{15}$N HSQC spectra of isolated apoWT-HSco1 (black) and in the presence of 1 equivalent of $Cu_1$(I)Cox17$_{2S-S}$ (red) in 50 mM phosphate pH 7.2, 1 mM DTT at 298K. The Cys ligands with surrounding residues are indicated.

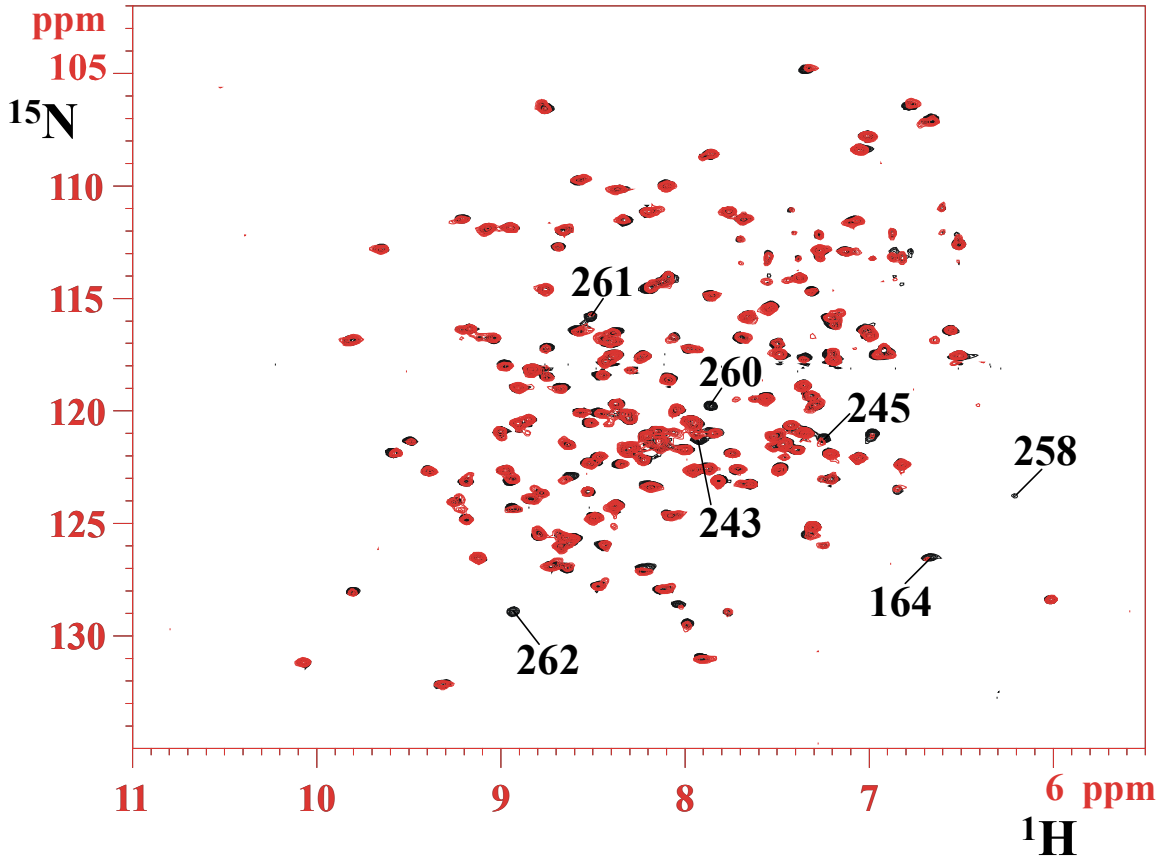



**Fig. 14**. ***The effect of Cu $_1$Cox17$_{2S\text{-}S}$ on the WT-HSco1 metallation state***. Superimposition of $^{1}$H-$^{15}$N HSQC spectra of Cu(I)WT-HSco1 metallated *via* Cu(I) acetonitrile complex (black) and *via* $Cu_1Cox17_{2S\text{-}S}$ (red) in 50 mM phosphate pH 7.2, 1 mM DTT at 298K.

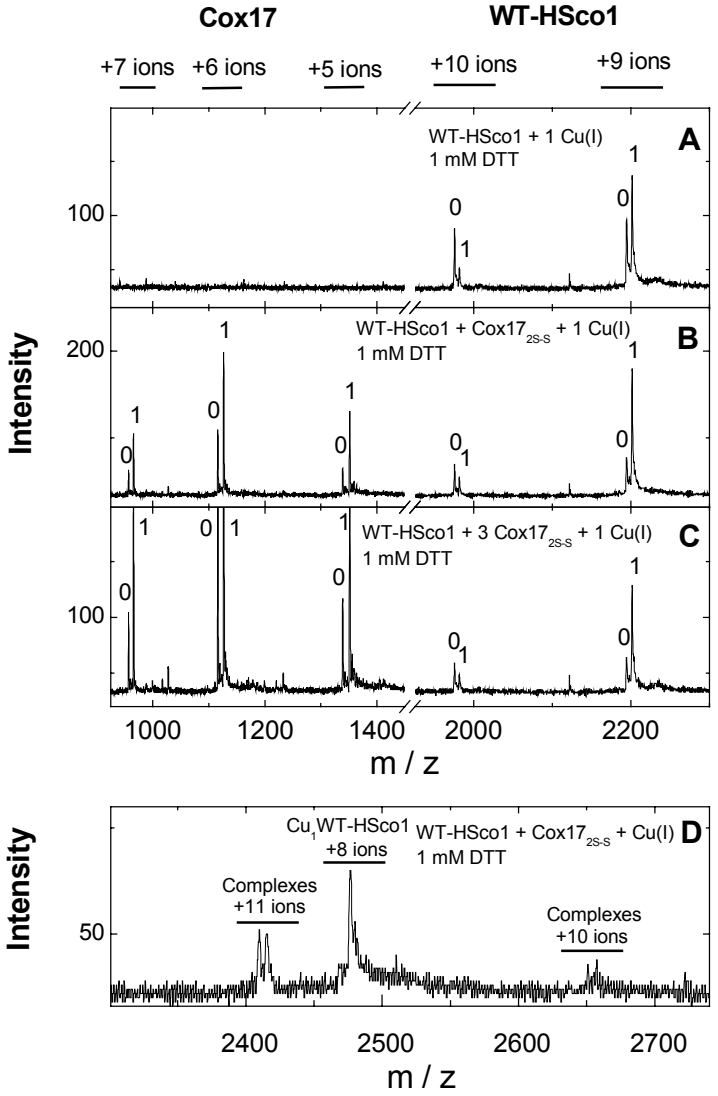



**Fig. 15.** ***The effect of Cox17 to the metallation of human Sco1 with Cu(I) ions.*** ESI-TOF mass spectra of WT-HSco1 reconstituted with Cu(I)DTT complex or $Cu_1$(I)$Cox17_{2S-S}$ in 50 mM ammonium acetate, pH 7.5. (A) WT-HSco1 (0.8 μM) + 1 equivalent of Cu(I) in the presence of 1.0 mM DTT; (B) WT-HSco1 (0.8 μM) + 1 equivalent of Cu(I) + 1 equivalent of $Cox17_{2S-S}$ in the presence of 1.0 mM DTT; (C) WT-HSco1 (0.8 μM) + 1 equivalent of Cu(I) + 3 equivalent of $Cox17_{2S-S}$ in the presence of 1.0 mM DTT; charge state +9 and +10 ions are presented for WT-HSco1 and +5, +6 and +7 ions for $Cox17_{2S-S}$ and numbers on the peaks denote the metal stoichiometry of the complex. (D) Determination of protein-protein complexes in the mixture of $Cox17_{2S-S}$ (2.4 μM) + Cu(I) (2.4 μM) + WT-HSco1 (2.4 μM) in the presence of 1.0 mM DTT.

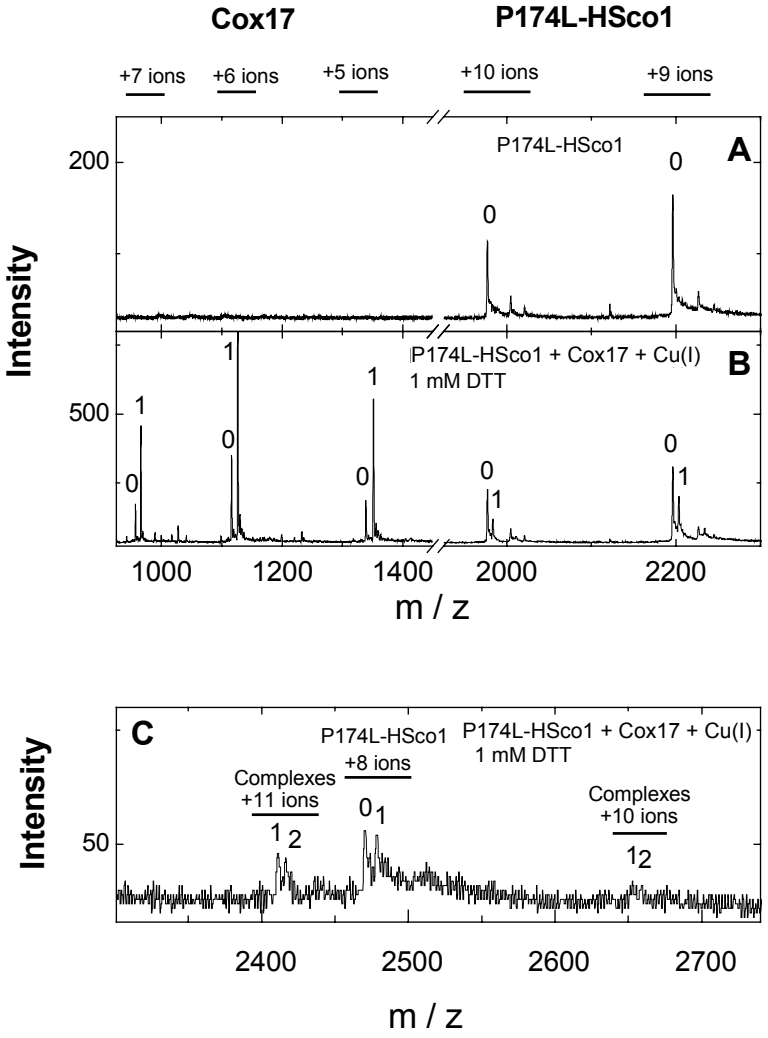



**Fig. 16.** ***Interaction of Cox17 to the metallation of human P174L-HSco1 with Cu(I) ions.*** ESI-TOF mass spectra of P174L-HSco1 in the presence of Cu(I) ions and $Cox17_{2S\text{-}S}$ in 50 mM ammonium acetate, pH 7.5. (A) P174L-HSco1 (1.6 μM); (B) P174L-HSco1 (7.2 μM) + Cu(I) (4.8 μM) + $Cox17_{2S\text{-}S}$ (7.2 μM) in the presence of 1.0 mM DTT; charge state +9 and +10 ions are presented for P174L-HSco1 and +5, +6 and +7 ions for $Cox17_{2S\text{-}S}$ and numbers on the peaks denote the metal stoichiometry of the complex. (C) Peaks corresponding to protein-protein complexes in the mixture of P174L-HSco1 (7.2 μM) + Cu(I) (4.8 μM) + $Cox17_{2S\text{-}S}$ (7.2 μM) in the presence of 1.0 mM DTT. Charge state +10 and +11 ions are denoted for P174L-HSco1-$Cox17_{2S\text{-}S}$ complexes and numbers on the peaks denote the metal stoichiometry in the complex.

# 3.4

## The Unusually Stable Quaternary Structure of Human Cu,Zn-Superoxide Dismutase 1 Is Controlled by Both Metal Occupancy and Disulfide Status

Fabio Arnesano, Lucia Banci, Ivano Bertini, Manuele Martinelli,
Yoshiaki Furukawa,and Thomas V. O'Halloran

*(J Biol Chem.* 2004 Nov 12;279(46):47998-8003)

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 279, No. 46, Issue of November 12, pp. 47998–48003, 2004
Printed in U.S.A.


# The Unusually Stable Quaternary Structure of Human Cu,Zn-Superoxide Dismutase 1 Is Controlled by Both Metal Occupancy and Disulfide Status*[S]

Received for publication, June 1, 2004, and in revised form, August 9, 2004
Published, JBC Papers in Press, August 23, 2004, DOI 10.1074/jbc.M406021200

**Fabio Arnesano‡, Lucia Banci‡, Ivano Bertini‡§, Manuele Martinelli‡, Yoshiaki Furukawa¶‖, and Thomas V. O'Halloran¶****

*From the ‡Magnetic Resonance Center CERM and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy and ¶Departments of Chemistry and Biochemistry, Molecular Biology, and Cell Biology, Northwestern University, Evanston, Illinois 60208*

**The eukaryotic copper,zinc superoxide dismutases are remarkably stable dimeric proteins that maintain an intrasubunit disulfide bond in the reducing environment of the cytosol and are active under a variety of stringent denaturing conditions. The structural interplay of conserved disulfide bond and metal-site occupancy in human copper,zinc superoxide dismutase (hSOD1) is of increasing interest as these post-translational modifications are known to dramatically alter the catalytic chemistry, the subcellular localization, and the susceptibility of the protein to aggregation. Using biophysical methods, we find no significant change in the gross secondary or tertiary structure of the demetallated form upon reduction of the disulfide. Interestingly, reduction does lead to a dramatic change in the quaternary structure, decreasing the monomer-to-dimer equilibrium constant by at least four orders of magnitude. This reduced form of hSOD1 is monomeric, even at concentrations well above the physiological range. Either the addition of Zn(II) or the formation of the disulfide leads to a shift in equilibrium that favors the dimeric species, even at low protein concentrations (*i.e.* micromolar range). We conclude that only the most immature form of hSOD1, *i.e.* one without any post-translational modifications, favors the monomeric state under physiological conditions. This finding provides a basis for understanding the selectivity of mitochondrial SOD1 import and may be relevant to the toxic properties of mutant forms of hSOD1 that can cause the familial form of amyotrophic lateral sclerosis.**

Eukaryotic copper,zinc superoxide dismutase (SOD1)[1] catalyzes the dismutation of superoxide radical to oxygen and hydrogen peroxide and is a 32-kDa homodimeric enzyme found predominantly in the cytosol (1). SOD1 is one of the most thermally stable enzymes known in mesophilic organisms. Dismutase activity declines at 80 °C with a corresponding melting temperature, $T_m$, above 90 °C (2). The protein is stable in the presence of strong denaturants, and the activity is observed in 4% SDS or 10 M urea (3). Structural properties of SOD1 that contribute to this extreme thermochemical stability are thought to include an eight-stranded $\beta$-barrel motif, hydrophobic interactions associated with dimerization, coordinate covalent bonds, and an intrasubunit disulfide bond between highly conserved pair of cysteines, namely $Cys^{57}$ and $Cys^{146}$ in the human form. Whereas the dimerization can contribute to the structural stability through the reduction of its mobility (4), the roles of the disulfide bond in the SOD1 function and/or structure are only now beginning to emerge. Inspection of the SOD1 structure reveals that the loop containing $Cys^{57}$ can influence the conformation of the catalytically important residue, $Arg^{143}$, through a hydrogen-bonding network (5). Portions of this loop contribute to the dimer interface (6), leading to the possibility that the disulfide bond influences the protein dimerization and thereby the SOD1 quaternary structure.

To attain the correctly folded quaternary structure and become enzymatically active, several post-translational modifications need to occur in SOD1 such as the acquisition of copper and zinc ions, formation of the disulfide bond, and dimerization. Whereas the mechanism by which SOD1 acquires Zn(II) is not fully understood, several aspects of the copper insertion by the copper chaperone for SOD1 (CCS) are well established (7–12). More recently, Furukawa *et al.* (13) have shown that the intrasubunit disulfide bond is correctly introduced in yeast SOD1 by the copper-bound form of yeast CCS. However, given that the cytosol is a strongly reducing environment due to the high GSH/GSSG ratio (100:1–1000:1) (14), the disulfide formation is an unfavorable process. Those results suggest that the immature disulfide-reduced hSOD1 is a more important species in the cytosolic environment than previously thought.

* This work was supported in part by the European Commission (Contract QLG2-CT-2002-00988), by the Italian MURST Project COFIN03, and by National Institutes of Health Grants GM 54111 (to T. V. O.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains Supplemental Fig. S1.

‖ Supported by Japan Society for the Promotion of Science Postdoctoral Fellowship for Research Abroad.

§ To whom correspondence may be addressed: CERM and Dept. of Chemistry, University of Florence, Via L. Sacconi 6, Sesto Fiorentino 50019, Italy. Tel.: 39-055-45-4272; Fax: 39-055-4574271; E-mail: bertini@cerm.unifi.it.

** To whom correspondence may be addressed: Dept. of Chemistry, Northwestern University, 2145 Sheridan Rd., Evanston, Illinois 60208. Tel.: 847-491-5060; Fax: 847-491-7713; E-mail: t-ohalloran@northwestern.edu.

[1] The abbreviations used are: SOD1, Cu,Zn-superoxide dismutase; hSOD1, human SOD1; ySOD1, yeast SOD1; E,E-hSOD1$^{SH}$, fully reduced and demetallated hSOD1; E,Zn-hSOD1$^{SH}$, fully reduced Zn-loaded hSOD1; E,E-hSOD1$^{SS}$, oxidized and demetallated hSOD1; E,Zn-hSOD1$^{SS}$, oxidized and Zn-loaded hSOD1; Q133M2SOD1, human SOD1 with the mutations F50E/G51E/E133Q; CCS, copper chaperone for SOD1; AMS, 4-acetamide-4′maleimidylstilbene-2,2′-disulfonic acid; WT, wild type; HSQC, heteronuclear single quantum coherence; CD, circular dichroism; fALS, familial form of amyotrophic lateral sclerosis.

Field *et al.* (15) have also recently shown that uptake of the SOD1 molecule into the intermembrane space of the mitochondria is dependent on the status of the disulfide bond. The reduced form of SOD1 is imported through the mitochondrial outer membrane, but the disulfide-bonded apo-SOD1, the Zn(II)-loaded SOD1, and the holo-form or fully mature form of SOD1 are not readily transferred from the cytosol into the intermembrane space of the mitochondria. The effects of disulfide reduction on the SOD1 structure are therefore relevant to our understanding of the intracellular localization and stability of the SOD1 molecule. In this study we show that, even after removal of both copper and zinc ions from the active and mature form of hSOD1, the dimeric state still persists; however, upon reduction of the disulfide bond, the protein can readily dissociate to the monomer form. Zn(II) addition to the reduced apo-hSOD1 restores the dimeric state, indicating that only the most immature form of hSOD1 before any post-translational modifications favors the monomeric state. These results provide a molecular basis for understanding factors that control the SOD1 monomer-dimer equilibrum in the cytosol and have direct relevance to models for the toxic gain of function mutations in SOD1 that are associated with familial amyotrophic lateral sclerosis (fALS).

## EXPERIMENTAL PROCEDURES

*Sample Preparation*—hSOD1 was expressed in the *Escherichia coli* TOPP1 (Stratagene) or BL21(DE3) strain. The mutations were performed using a QuikChange™ site-directed mutagenesis kit (Stratagene). The $^{15}$N-labeled protein in which the non-conserved cysteine residues, Cys$^{6}$ and Cys$^{111}$, were mutated to Ser was obtained by growing the cells in the M9 minimal medium with $N^{15}$-$NH_4Cl$ following a reported procedure (16), whereas LB medium was used for the non-labeled protein. The cells were grown at 37 °C until $A_{600}$ = 0.6 and induced with 1.0 mM isopropyl 1-thio-β-D-galactopyranoside for 6 h. The protein was isolated and purified according to previously published protocols (16). Fully reduced and demetallated hSOD1 (E,E-hSOD1$^{SH}$) was prepared by treating the isolated protein with dithiothreitol at 37 °C for 1 h in an anaerobic chamber to reduce the disulfide bond (13). The protein solution was then acidified using 0.4% trifluoroacetic acid, and organic solvents (15% $CH_3CN$, 10% $CH_3OH$) were included to remove the metal ions. The protein solution was purified using reverse-phase high pressure liquid chromatography through a 300-Å C18 Jupiter column (Phenomenex) equilibrated with 0.1% trifluoroacetic acid in water. The fractions containing hSOD1 were eluted with a linear gradient of 0.1% trifluoroacetic acid in $CH_3CN$ and lyophilized. The metal content of E,E-hSOD1 was checked by inductively coupled plasma atomic emission spectrometry using a Thermo Jarrell Ash Atomscan Model 25 Sequential inductively coupled spectrometer, and zinc and copper ions were <10 nM in the 2 μM protein sample. Zinc reconstitution was obtained as previously described (17). Protein reduction and metallation were carried out under a nitrogen atmosphere in an anaerobic chamber.

*Thiol-Disulfide Reduction Assay*—The thiol-disulfide status of purified hSOD1 was determined by chemical modification with the thiol-specific reagent 4-acetamide-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS) (Molecular Probes, Inc.) (13). AMS conjugation results in a ~1 kDa increase in the molecular mass of hSOD1 as visualized by non-reducing SDS-PAGE and Coomassie Blue staining. 3 μg of the SOD1 protein that is dissolved in 10 μl of the buffer was mixed with 2.5 μl of 100 mM AMS and 2.5 μl of 10% SDS. The reaction mixture was incubated at 37 °C for an hour in an anaerobic chamber, and then the Laemmli buffer without any reducing agent was added. After boiling at 95 °C for 2 min, the sample was loaded on SDS-PAGE gel.

*CD Spectroscopy*—Far-UV CD spectra (185–260 nm) of hSOD1 were recorded on JASCO J-810 spectropolarimeter. A cell with a path length of 1 mm was used for the measurement, and the parameters were set as follows: bandwidth, 2 nm; step resolution, 0.1 nm; scan speed, 50 nm/min; and response time: 2 s. Each spectrum was obtained as the average of five scans. The protein concentration was typically around 20 μM. Prior to the calculation of the mean residue molar ellipticity, all of the spectra were corrected by subtracting the contributions from the buffer. Spectra were then smoothed using adjacent averaging or Fast Fourier transform filter. Quantitative estimations of the secondary structure contents were made using the DICROPROT software package (18).

*NMR Spectroscopy*—Data were collected on Bruker Avance 500 spectrometer, operating at a proton nominal frequency of 500.13 MHz. A triple resonance Cryoprobe equipped with pulsed field gradients along the *z*-axis was used. The two-dimensional $^{1}$H-$^{15}$N HSQC spectra and relaxation experiments were acquired on 0.5 mM samples of $^{15}$N-labeled E,E- and E,Zn-hSOD1$^{SH}$ in 20 mM sodium phosphate buffer (pH 7.0). The $^{15}$N backbone longitudinal and transverse relaxation rates, $R_1$ and $R_2$, were measured as previously described (19). The value of reorientational correlation time $\tau_m$ was estimated from the $R_2/R_1$ ratio with the program Quadric_diffusion (20). All of the spectra were collected at 298 K, processed using the standard Bruker software (XWINNMR). All of the NMR samples were prepared under nitrogen atmosphere in a glove box where they were loaded into 5-mm quartz NMR tubes capped with latex serum caps.

*Gel Filtration Chromatography*—200 μl of 30 μM hSOD1 protein was loaded on Superose 12 HR 10/30 (Amersham Biosciences) at 4 °C. The column was preequilibrated with 50 mM potassium phosphate, pH 7.5, and the flow rate was 1.0 ml/min. To prevent the possible air-oxidation of the thiol groups, 1 mM dithiothreitol was added in the above buffer for the gel filtration analysis of E,E- and E,Zn-hSOD1$^{SH}$. For the experiments using E,E-hSOD1$^{SH/SS}$, 0.1 mM EDTA was included in the buffer. The chromatogram was obtained by monitoring the absorbance at 215 nm. The calibration of the column for the estimation of molecular weight was performed using 200 μl of 0.25 g/liter immunoglobulin G, bovine serum albumin, ovalbumin, carbonic anhydrase, horse heart skeletal myoglobin, *E. coli* thioredoxin, and aprotinin as protein standards.

## RESULTS

Chemical reduction of the disulfide in hSOD1 with dithiothreitol followed by acidification to remove bound metal ions yields the fully reduced and demetallated E,E-hSOD1$^{SH}$ sample as confirmed by AMS modification and inductively coupled plasma atomic emission spectrometry analysis. In the NMR and CD studies, two of the four Cys residues in hSOD1 have been mutated to Ser (*i.e.* C6S/C111S) to avoid the possible oxidation of the free thiol groups and it has been reported that these Cys residues, Cys$^{6}$ and Cys$^{111}$, which are not involved in disulfide formation, have little effects on the SOD1 activity and structure (21). When the disulfide bond is intact, previous studies have shown that SOD1 is mainly comprised of β sheets and has little α-helical structure (19). To examine the possible structural changes upon disulfide reduction, CD spectroscopy was used to probe the secondary structure.

*Disulfide Reduction Has a Little Effect on the SOD1 Secondary Structure*—As seen in Fig. 1, the CD spectrum of the most immature form, E,E-hSOD1$^{SH}$, exhibits a negative peak at 207 nm, indicating that E,E-hSOD1$^{SH}$ is predominantly comprised of the β sheets (22). The absence of a strong band at 222 nm in the spectrum indicates low α-helical content (22). We attempted to see whether any major changes in the secondary structure upon the disulfide formation could be detected by using E,E-form of the monomeric hSOD1 mutant, E,E-Q133M2SOD1$^{SS}$. However, both E,E-hSOD1$^{SH}$ and E,E-Q133M2SOD1$^{SS}$ give CD signals similar to that of the matured form of the enzyme, Cu,Zn-hSOD1$^{SS}$ (Fig. 1). The fitting of CD data (18, 23) suggests that these forms of hSOD1 have similar secondary structure content as reported in Table I. Disulfide reduction does not significantly alter the secondary structure, suggesting that several features of the β barrel-folding pattern are acquired before any post-translational modifications. We also examined the effects of the disulfide reduction on the tertiary and quaternary structure of hSOD1 by NMR spectroscopy.

*Disulfide Reduction and Zinc Removal Disrupt SOD1 Quaternary Structure*—The two-dimensional $^{1}$H-$^{15}$N HSQC spectrum of E,E-hSOD1$^{SH}$ is shown in Fig. 2*A* (*red contours*). Although several signals are present in a spectral region typical of unfolded polypeptides (between 8 and 8.5 ppm in the $^{1}$H

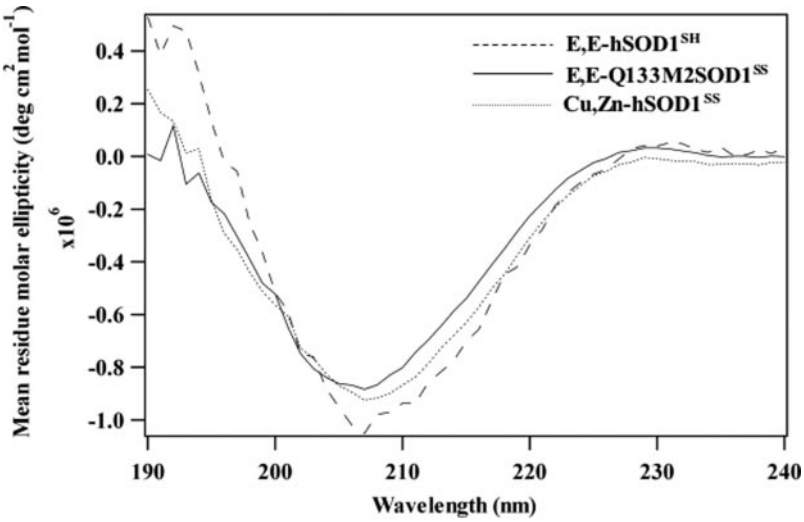



FIG. 1. **The CD spectra of E,E-hSOD1$^{SH}$(C6S/C111S) (*broken curve*), E,E-Q133M2SOD1$^{SS}$ (*solid curve*), and Cu,Zn-hSOD1$^{SS}$(C6S/C111S) (*dotted curve*) in 20 mM sodium phosphate buffer, pH 7.0.** 20 $\mu$M protein was used for the measurements.

TABLE I
*Secondary structure content of various forms of hSOD1 obtained from the fitting of far-UV CD spectra by the DICROPROT software package (18) using the least square method of Chen et al. (23)*

| | $\alpha$ Helix | $\beta$ Sheet | Random coil |
|---|---|---|---|
| | | % | |
| E,E-hSOD1$^{SH}$ | 8 | 50 | 42 |
| E,E-Q133M2SOD1$^{SS}$ | 4 | 46 | 50 |
| Cu,Zn-hSOD1$^{SS}$ | 6 | 43 | 51 |

dimension), this spectrum is remarkably similar to that of E,E-Q133M2SOD1$^{SS}$ (Fig. 2*B*) for which the NMR signals have been assigned (19). This finding suggests that the tertiary structure of E,E-hSOD1$^{SH}$ also is similar to that of E,E-Q133M2SOD1$^{SS}$, *i.e.* the $\beta$ barrel structure is maintained, and loops IV (residues 45–85) and VII (residues 121–142) are severely disordered (19). Because loop IV contains three His ligands involved in Zn(II) binding and also involves a disulfide-bond residue, namely Cys$^{57}$, it is likely that zinc removal and disulfide reduction can destabilize and increase the structural mobility of loop IV. Likewise, it has been reported that loop VII, which produces an optimal electrostatic field for uptake of the superoxide substrate, becomes well organized upon the zinc binding (19, 24). Therefore, even after demetallation and reduction of the disulfide bond, most of the protein tertiary structure is retained with the exception of the disorder of some loop regions.

However, when zinc ion is added to the E,E-hSOD1$^{SH}$ sample, a dramatic change in the $^{1}$H-$^{15}$N HSQC spectrum is observed (Fig. 2*A*, *blue contours*). Compared with the E,E state, E,Zn-hSOD1$^{SH}$ shows a larger signal dispersion in the spectrum, which resembles that of the matured protein, Cu,Zn-hSOD1$^{SS}$ (Fig. 2*C*). Monomeric E,Zn-Q133M2SOD1$^{SS}$ shows the intermediate features in the HSQC spectrum between monomeric E,E-hSOD1$^{SH}$ and dimeric Cu,Zn-hSOD1$^{SS}$ (see Supplemental Fig. S1). The $^{1}$H and $^{15}$N chemical shift differences between the three zinc-bound forms of the protein are shown in Fig. 3 as the weighted average chemical shift differences, $\Delta_{avg}$(HN) (*i.e.*$\{[(\Delta H)^2 + (\Delta N/5)^2]/2\}^{1/2}$, where $\Delta$H and $\Delta$N are chemical shift differences for $^{1}$H and $^{15}$N, respectively). Differences between E,Zn-hSOD1$^{SH}$ and E,Zn-Q133M2SOD1$^{SS}$ (Fig. 3*A*) are due to disulfide reduction and/or dimerization, differences between E,Zn-hSOD1$^{SH}$ and Cu,Zn-hSOD1$^{SS}$ (Fig. 3*B*) are due to disulfide reduction and/or copper binding, and differences between E,Zn-Q133M2SOD1$^{SS}$ and Cu,Zn-hSOD1$^{SS}$ (Fig. 3*C*) are due to dimerization and/or copper binding. In all of the cases, significant chemical shift differences are present in loops IV and VII, whereas those for the residues at the N- and C-terminal regions are relatively small. The disulfide formation/copper binding in the zinc-bound proteins affects loops IV and VII more significantly than the N- and C-terminal regions, which are at the dimer interface. Cross-peaks of some of these residues at the dimer interface, which are outside the crowded regions and well resolved in $^{1}$H-$^{15}$N HSQC spectra, are indicated with *arrows* in Fig. 2, *A–D*.

The overall correlation time for molecular reorientation ($\tau_m$) provides insights into the hSOD1 quaternary structure. The $\tau_m$ values were estimated from the averaged values of $^{15}$N $R_1$ and $R_2$, which are obtained for 69 and 72 backbone NH resonances of E,E- and E,Zn-hSOD1$^{SH}$, respectively (Table II). The overall correlation time is highly sensitive to the protein size (25), and the dimerization is expected to increase $\tau_m$. Estimated from the $R_2/R_1$ ratios, $\tau_m$ is significantly larger in E,Zn-hSOD1$^{SH}$ (20.6 $\pm$ 0.9 ns) than in E,E-hSOD1$^{SH}$ (10.3 $\pm$ 0.4 ns). In particular, the $\tau_m$ value of E,E-hSOD1$^{SH}$ is very similar to that of monomeric E,E-Q133M2SOD1$^{SS}$ (10.1 ns) (19) and E,Zn-Q133M2SOD1$^{SS}$ (8.4 $\pm$ 0.3 ns) (24), whereas E,Zn-hSOD1$^{SH}$ exhibits a $\tau_m$ value similar to that found for dimeric Cu,Zn-hSOD1$^{SS}$ (25.3 ns) (4). These results suggest that E,E-hSOD1$^{SH}$ is monomeric and that Zn(II) addition to the reduced protein can lead to the dimerization. This result is quite surprising, because previous studies have shown that SOD1 dimer is very stable, even after complete demetallation (26). To directly examine the effects of the disulfide reduction on the monomer-dimer equilibrium at physiological concentrations of protein, we employed gel filtration chromatography.

*Monomer-Dimer Equilibrium of E,E- and E,Zn-hSOD1$^{SH/SS}$*—Fig. 4*A* compares the gel filtration chromatograms between the wild-type E,E- and E,Zn-hSOD1$^{SH}$ proteins. The E,E-form of reduced hSOD1(WT) favors the monomeric state at the concentration of ~30 $\mu$M (*broken curve* in Fig. 4*A*). The physiological concentration of SOD1 in the cell has been estimated to be ~10 $\mu$M in yeast by quantitative Western blot (8) and ~100 $\mu$M in the cytosol of cultured hepatocytes by quantitative immunocytochemistry (27). After E,E-hSOD1$^{SH}$(WT) is anaerobically incubated with an equimolar amount of zinc ion for an hour at 37 °C, the resultant protein, E,Zn-hSOD1$^{SH}$(WT), strongly favors the dimeric state (*solid curve* in Fig. 4*A*). The preference for the monomeric state in E,E-hSOD1$^{SH}$(WT) can be also confirmed by using a hSOD1 mutant in which all of the Cys residues are changed to Ser, C6S/C57S/C111S/C146S (*i.e.* the C$^{4}$S form). The protein conformation of reduced hSOD1 can be modeled by this quadruple mutant, which cannot undergo any type of thiol oxidation. Consistent with the WT data, the E,E-form of this C$^{4}$S

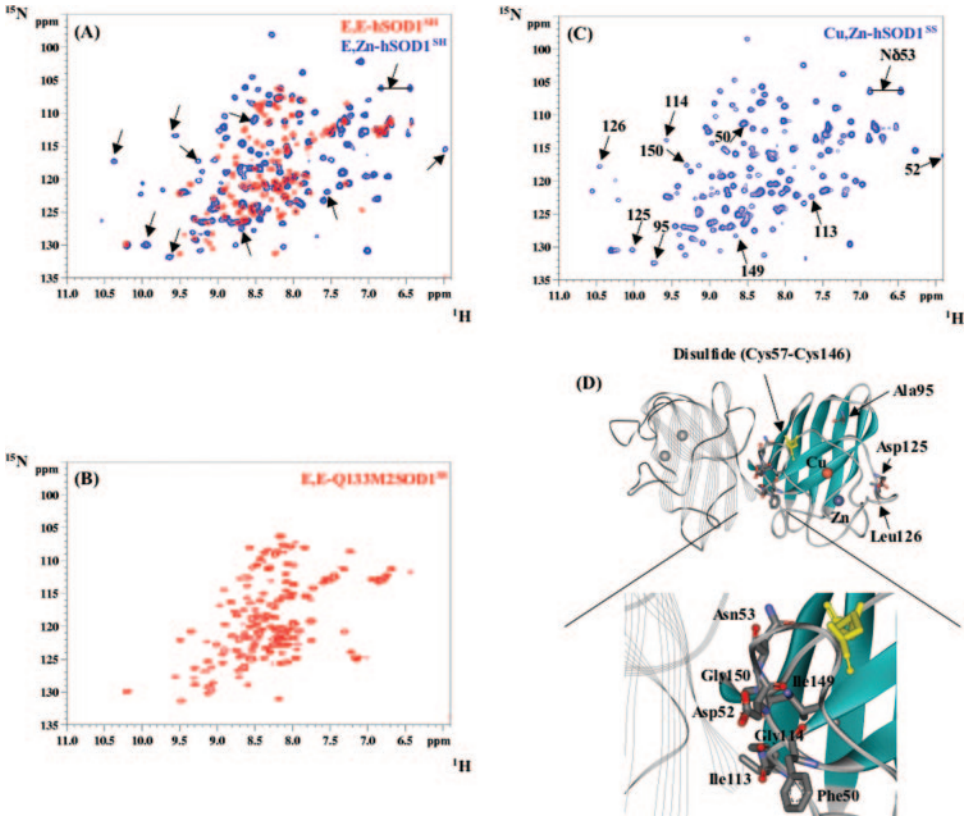



FIG. 2. **The two-dimensional $^{1}H$-$^{15}N$ HSQC spectra of E,E-hSOD1$^{SH}$(C6S/C111S) (*A*, *red contours*) and E,Zn-hSOD1$^{SH}$(C6S/C111S) (*blue contours*), monomeric E,E-Q133M2SOD1$^{SS}$ (*B*), and dimeric Cu,Zn-hSOD1$^{SS}$(C6S/C111S) (*C*).** Some cross-peaks of residues at the dimer interface and in the metal binding region are indicated by *arrows* and are shown in *D*.

mutant favors the monomeric state and turns to the dimeric state upon the addition of the zinc ion (Fig. 4*B*).

To investigate the effects of the conserved disulfide bond (Cys$^{57}$-Cys$^{146}$) on monomer-dimer equilibrium, the non-conserved cysteine residues, Cys$^{6}$ and Cys$^{111}$, were mutated to Ser. As seen in the WT protein, the E,E-form of the C6S/C111S mutant favors the monomeric state when the disulfide bond is reduced and the addition of Zn ion can dimerize the protein (Fig. 4*C*). These results show that the monomer-dimer equilibrium is not affected by the non-conserved cysteine residues. In contrast, when the conserved Cys residues are oxidized to form the intramolecular disulfide bond, the E,E-hSOD1$^{SS}$(C6S/C111S) form elutes at the peak position corresponding to the dimeric state (*broken curve* in Fig. 4*D*). This result suggests that, in the absence of any metal ions, the SOD1 monomerization is promoted by reduction of the canonical disulfide. The addition of Zn(II) ion to E,E-hSOD1$^{SS}$(C6S/C111S) does not further change the elution profile, and E,Zn-hSOD1$^{SS}$(C6S/C111S) still favors the dimeric state (*solid curve* in Fig. 4*D*). Therefore, Zn(II) removal alone cannot monomerize the hSOD1 protein unless the disulfide is reduced. We conclude that the nascent or folded form of the hSOD1 polypeptide favors the monomer state until it undergoes the first of several post-translational modifications in the cell.

## DISCUSSION

The subunits of copper,zinc superoxide dismutase are not linked by covalent bonds but are nonetheless unusually resistant to dissociation. Strong interaction is observed under extreme denaturing conditions such as 8–10 M urea, 7 M guanidine HCl, or SDS (3). Accordingly, it is surprising that the E,E-hSOD1$^{SH}$ protein is monomeric, even without any detergents present. In fact, we find that the E,E-hSOD1$^{SH}$ form is dominantly monomeric even at the high concentrations (500 $\mu$M) used in the NMR experiments. When we assume that ~10% total 500 $\mu$M E,E-hSOD1$^{SH}$ is in the dimeric state, the upper limit for the association constant is estimated as $K_a < 2 \times 10^2$ M$^{-1}$. At the other extreme, E,E-hSOD1$^{SS}$ still favors the dimeric state, even at the low concentrations used (30 $\mu$M) in the gel filtration experiments. Based on a conservative estimate of the absorbance in the gel filtration, <10% of the total E,E-hSOD1$^{SS}$ exists as the monomer in this condition, leading to an estimate of the lower limit of $K_a > 3 \times 10^6$ M$^{-1}$. Thus, although disulfide formation does not significantly change the secondary structure of hSOD1 (Fig. 1), it clearly favors dimerization by at least four orders of magnitude. The observations here show that the disulfide bond plays a quite significant but not clearly exclusive role in determining the quaternary structure of the human form of SOD1.

Zinc acquisition by the E,E-hSOD1$^{SH}$ state seems to have as profound effect on the monomer-dimer equilibrium as disulfide formation, *i.e.* both E,E-hSOD1$^{SS}$ and E,Zn-hSOD1$^{SH}$ favor the dimeric state, even when protein concentration is as low as 10 $\mu$M. Zinc binding has been shown to reduce the mobility of the loop IV (residues 48–85, *colored* with *purple* in Fig. 5) (19, 24), which contains the important amino acid residues for zinc ligation, *i.e.* His$^{63}$, His$^{71}$, and His$^{80}$ (*colored* with *green* in Fig. 5). Because loop IV in SOD1 is adjacent to the interface between the subunits, such a structurization of loop IV upon the binding of the Zn ion is also implicated to play important roles in the SOD1 dimerization (19, 24). The results here show that, even when the disulfide is reduced, Zn(II) binding alone is enough to stabilize a conformation of the protein that favors dimer formation. Likewise, it is interesting to note that loop IV is linked to the $\beta$ sheet unit via the disulfide formation between Cys$^{57}$ and Cys$^{146}$. Whereas a structural determination of several SOD1 microstates is currently in progress, reduction of the

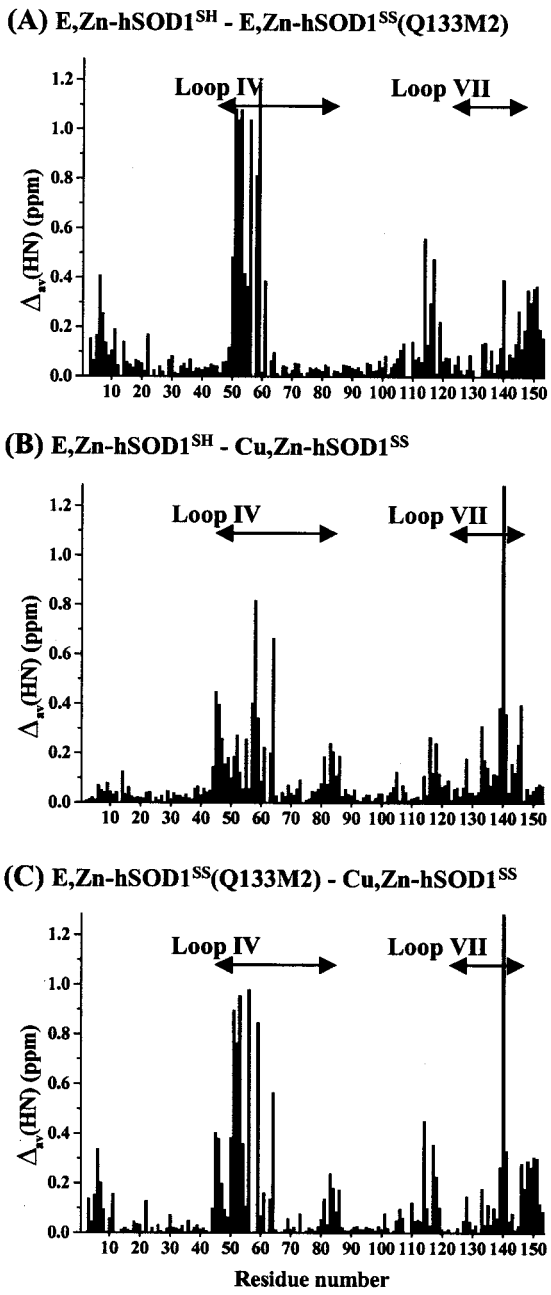



FIG. 3. **$^{1}$H and $^{15}$N amide chemical shift differences between E,Zn-hSOD1$^{SH}$ and E,Zn-Q133M2SOD1$^{SS}$ (*A*), between E,Zn-hSOD1$^{SH}$ and Cu,Zn-hSOD1$^{SS}$ (*B*), and between E,Zn-Q133M2SOD1$^{SS}$ and Cu,Zn-hSOD1$^{SS}$ (*C*).** The weighted average chemical shift differences $\Delta_{avg}$(HN) are shown (see "Results"). The extents of loops IV and VII are indicated with *arrows*.

TABLE II
*Average $^{15}$N $R_1$ and $R_2$ values and the overall correlation time, $\tau_m$*

| | Average $R_1$ | Average $R_2$ | $\tau_m$ |
|---|---|---|---|
| | $s^{-1}$ | | ns |
| E,E-hSOD1$^{SH}$ | 1.79 ± 0.17[a] | 13.8 ± 2.0[a] | 10.3 ± 0.4 |
| E,Zn-hSOD1$^{SH}$ | 1.01 ± 0.10[a] | 29.2 ± 3.4[a] | 20.6 ± 0.9 |
| E,E-Q133M2SOD1$^{SS}$ | 1.24 ± 0.09[b] | 12.6 ± 2.7[b] | 10.1 ± 0.3[c] |
| E,Zn-Q133M2SOD1$^{SS}$ | 1.36 ± 0.29[b] | 13.4 ± 0.9[b] | 8.4 ± 0.3[d] |
| Cu,Zn-hSOD1$^{SS}$ | 0.60 ± 0.11[b] | 33.9 ± 4.1[b] | 25.3 ± 1.3[e] |

[a] Measured at 500 MHz.
[b] Measured at 600 MHz.
[c] Ref. 19.
[d] Ref. 24.
[e] Ref. 4.

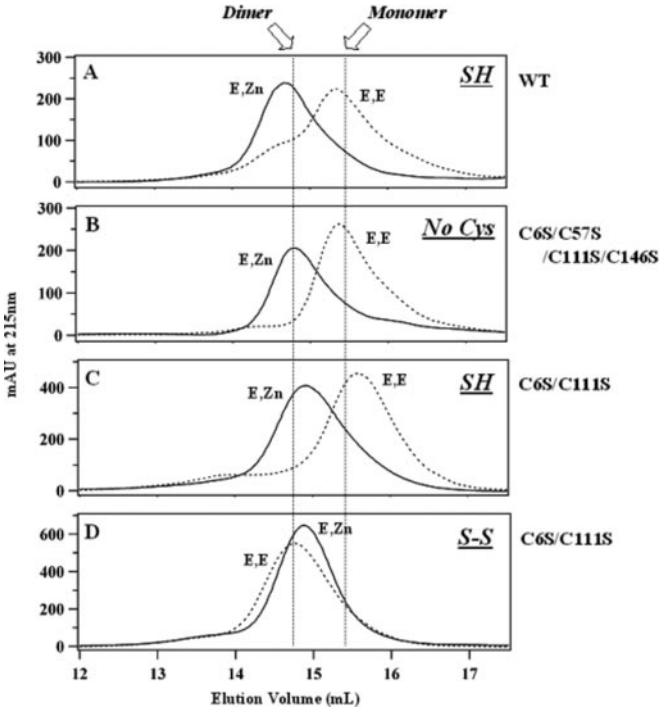



FIG. 4. **The chromatograms of the gel filtration experiments using hSOD1$^{SH}$(WT) (*A*), hSOD1(C6S/C57S/C111S/C146S) (*B*), hSOD1$^{SH}$(C6S/C111S) (*C*), and hSOD1$^{SS}$(C6S/C111S) (*D*).** The *solid* and *broken curves* represent E,Zn- and E,E-forms of each protein, respectively. The elution profiles were obtained by using Superose 12 (Amersham Biosciences) column and monitoring the absorbance change at 215 nm. The flow rate was 1 ml/min.

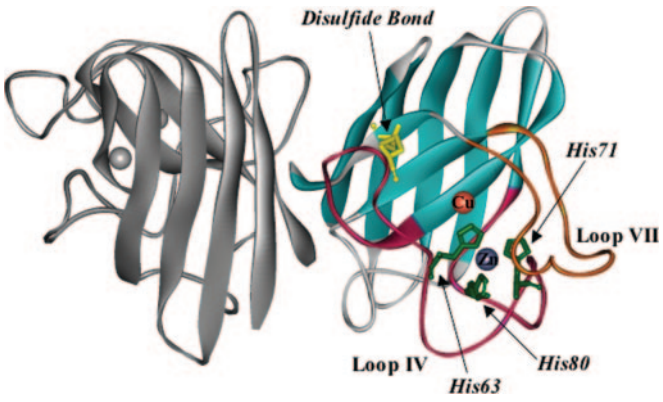



FIG. 5. **The structure of E,Zn-hSOD1$^{SS}$ (Protein Data Bank code 1HL4).** The disulfide bond and the zinc ligands are shown as *yellow* and *green sticks*, respectively. Copper and zinc ions are indicated in the sphere style. Loop IV (residues 48–85) is colored *purple*, whereas loop VII (residue 121–142) is colored *orange*.

disulfide bond would release and disorder loop IV, resulting in the increase of its conformational flexibility. Such a disordered structure of loop IV may obstruct the interaction between the subunits. Therefore, both disulfide formation and Zn(II) binding can add structure to the protein conformation around loop IV, which would promote the interaction between the subunits.

Important roles of the disulfide bond in the monomer-dimer equilibrium have also been reported in the yeast SOD1 (ySOD1) and are in accordance with this study on human SOD1. However, Furukawa *et al.* (13) have found that reduction of the disulfide in E,Zn-ySOD1$^{SS}$ (dimer) leads to conformation changes that favor the monomeric state under the same conditions. Thus, there is an interesting difference in the behavior of the yeast and human proteins. The E,Zn-form of reduced human protein favors the dimeric state, whereas the

same state of the yeast protein favors the monomer state. Although the structural characterization of E,Zn-ySOD1$^{SH}$ and E,Zn-hSOD1$^{SH}$ is necessary, we speculate that this difference may be attributed to two proline residues, Pro$^{142}$ and Pro$^{144}$, near the intrasubunit disulfide bond that are present in the yeast but not in the human protein. Culotta and co-workers (28) have recently shown that these residues play a key role in the CCS-independent activation pathways, which differ for the yeast and human enzyme. The trans-configuration of these proline residues would limit the ySOD1 conformation, especially around the disulfide bond and loop IV to reduce the interaction between the subunits. Given that the cytosol can provide the strongly reducing conditions (100–1000:1 of the GSH/GSSG ratio), the monomeric form of this disulfide-reduced SOD1 protein appears to be more physiologically relevant than has been appreciated to date.

In the yeast system, the most immature form, *i.e.* E,E-ySOD1$^{SH}$, is the only one that is efficiently taken up from the cytosol into the intermembrane space of mitochondria. Mitochondrial retention of SOD1 is dependent upon its activation by yeast CCS inside the intermembrane space of mitochondria (15); however, once the Zn(II) ion is incorporated or the disulfide bond is introduced in E,E-ySOD1$^{SH}$, mitochondrial import of the SOD1 protein is significantly inhibited (15). This selectivity for SOD1 mitochondrial import could be explained by our current results. The completely demetallated and disulfide-reduced form has a smaller size than any other forms. Furthermore, it is expected to be easier to unfold and thread its way through machinery in the mitochondrial outer membrane. Dimerization that accompanies metallation and disulfide formation may prevent the SOD1 dimer from crossing the mitochondrial membrane.

Perturbation of the SOD1 quaternary structure by disulfide reduction may be relevant to the etiology of fALS, which has been associated with a number of mitochondrial pathologies in fALS patients (29). Point mutations in the human SOD1 can cause 20% of total fALS, which is a fatal and late-onset neurodegenerative disorder (30, 31). The fALS-associated mutants do not necessarily lose the SOD1 activity but gain some new activities to cause the disease, such as peroxidase activity or adventitious protein aggregation (29). It has been suggested that the apoform of the fALS mutant exhibits decreased stability, which has some correlations with disease duration (32). Furthermore, it has been proposed that protein monomerization plays a role in formation of misfolded intermediates, leading to protein aggregation (33). Because the conserved disulfide bond in SOD1 is adjacent to the dimer interface (Fig. 4) (34), we suspect that SOD1 monomerization can increase the exposure of these Cys residues. A thiol group in the Cys residue is in general susceptible to the oxidative modification, which plays an important role in the protein aggregation in some neurodegenerative disease (35). Recently, it has been shown that exposed Cys residues in the SOD1 monomer can be modified by oxidative stress, leading to disulfide-linked multimerization of SOD1 (13). These SOD1 multimers could be involved in protein aggregation and the pathology of amyotrophic lateral sclerosis. Interestingly, increased susceptibility to disulfide reduction has been observed in some fALS mutants (36); therefore, protein monomerization caused by the disulfide reduction and demetallation might be an important process in causing the fALS diseases.

*Acknowledgments*—We thank A. Herrnreiter for preparation of several protein samples. We also thank the O'Halloran group for the helpful discussions.

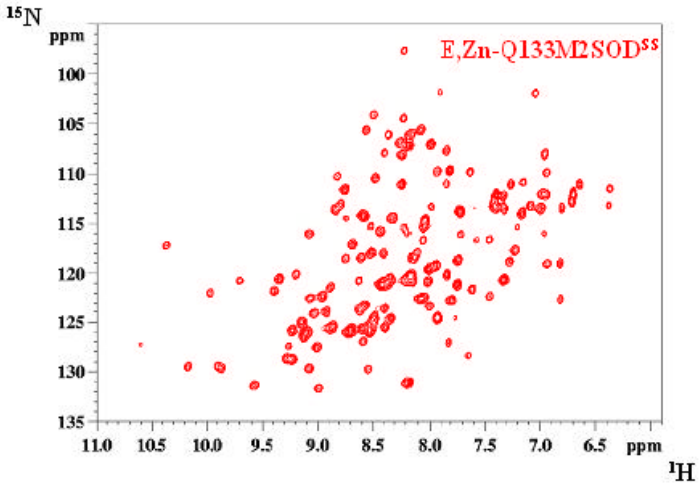
REFERENCES

1. McCord, J. M., and Fridovich, I. (1969) *J. Biol. Chem.* **244,** 6049–6055
2. Roe, J. A., Butler, A., Scholler, D. M., Valentine, J. S., Marky, L., and Breslauer, K. J. (1988) *Biochemistry* **27,** 950–958
3. Forman, H. J., and Fridovich, I. (1973) *J. Biol. Chem.* **248,** 2645–2649
4. Banci, L., Bertini, I., Cramaro, F., Del Conte, R., Rosato, A., and Viezzoli, M. S. (2000) *Biochemistry* **39,** 9108–9118
5. Fisher, C. L., Cabelli, D. E., Tainer, J. A., Hallewell, R. A., and Getzoff, E. D. (1994) *Proteins* **19,** 24–34
6. Bertini, I., Mangani, S., and Viezzoli, M. S. (1998) in *Advanced Inorganic Chemistry* (Sykes, A. G., ed) pp. 127–250, Academic Press, San Diego
7. O'Halloran, T. V., and Culotta, V. C. (2000) *J. Biol. Chem.* **275,** 25057–25060
8. Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C., and O'Halloran, T. V. (1999) *Science* **284,** 805–808
9. Rae, T. D., Torres, A. S., Pufahl, R. A., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 5166–5176
10. Lamb, A. L., Torres, A. S., O'Halloran, T. V., and Rosenzweig, A. C. (2000) *Biochemistry* **39,** 14720–14727
11. Culotta, V. C., Klomp, L. W. J., Strain, J., Casareno, R. L. B., Krems, B., and Gitlin, J. D. (1997) *J. Biol. Chem.* **272,** 23469–23472
12. Eisses, J. F., Stasser, J. P., Ralle, M., Kaplan, J. H., and Blackburn, N. J. (2000) *Biochemistry* **39,** 7337–7342
13. Furukawa, Y., Torres, A. S., and O'Halloran, T. V. (2004) *EMBO J.* **23,** 2872–2881
14. Hwang, C., Sinskey, A. J., and Lodish, H. F. (1992) *Science* **257,** 1496–1502
15. Field, L. S., Furukawa, Y., O'Halloran, T. V., and Culotta, V. C. (2003) *J. Biol. Chem.* **278,** 28052–28059
16. Banci, L., Benedetto, M., Bertini, I., Del Conte, R., Piccioli, M., and Viezzoli, M. S. (1998) *Biochemistry* **37,** 11780–11791
17. Beem, K. M., Rich, W. E., and Rajagopalan, K. V. (1974) *J. Biol. Chem.* **249,** 7298–7305
18. Deleage, G., and Geourjon, C. (1993) *Comput. Appl. Biosci.* **9,** 197–199
19. Banci, L., Bertini, I., Cramaro, F., Del Conte, R., and Viezzoli, M. S. (2003) *Biochemistry* **42,** 9543–9553
20. Bruschweiler, R., Liao, X., and Wright, P. E. (1995) *Science* **268,** 886–889
21. Parge, H. E., Hallewell, R. A., and Tainer, J. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6109–6113
22. Greenfield, N., and Fasman, G. D. (1969) *Biochemistry* **8,** 4108–4116
23. Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) *Biochemistry* **13,** 3350–3359
24. Banci, L., Bertini, I., Cantini, F., D'Onofrio, M., and Viezzoli, M. S. (2002) *Protein Sci.* **11,** 2479–2492
25. Einstein, A. (1956) *Investigations on the Theory of the Brownian Movement*, Dover, New York
26. Hartz, J. W., and Deutsch, H. F. (1972) *J. Biol. Chem.* **247,** 7043–7050
27. Chang, L. Y., Slot, J. W., Geuze, H. J., and Crapo, J. D. (1988) *J. Cell Biol.* **107,** 2169–2179
28. Carroll, M. C., Girouard, J. B., Ulloa, J. L., Subramaniam, J. R., Wong, P. C., Valentine, J. S., and Culotta, V. C. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5964–5969
29. Julien, J. P. (2001) *Cell* **104,** 581–591
30. Deng, H. X., Hentati, A., Tainer, J. A., Iqbal, Z., Cayabyab, A., Hung, W. Y., Getzoff, E. D., Hu, P., Herzfeldt, B., Roos, R. P., *et al*. (1993) *Science* **261,** 1047–1051
31. Gurney, M. E., Pu, H., Chiu, A. Y., Dal Canto, M. C., Polchow, C. Y., Alexander, D. D., Caliendo, J., Hentati, A., Kwon, Y. W., Deng, H. X., Chen, W., Zhai, P., Sufit, R. L., and Siddique, T. (1994) *Science* **264,** 1772–1775
32. Lindberg, M. J., Tibell, L., and Oliveberg, M. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 16607–16612
33. Rakhit, R., Crow, J. P., Lepock, J. R., Kondejewski, L. H., Cashman, N. R., and Chakrabartty, A. (2004) *J. Biol. Chem.* **279,** 15499–15504
34. Bordo, D., Djinovic, K., and Bolognesi, M. (1994) *J. Mol. Biol.* **238,** 366–386
35. Lee, S., and Eisenberg, D. (2003) *Nat. Struct. Biol.* **10,** 725–730
36. Tiwari, A., and Hayward, L. J. (2003) *J. Biol. Chem.* **278,** 5984–5992

# SUPPLEMENTARY MATERIAL

## The Unusually Stable Quaternary Structure of Human Cu,Zn-Superoxide Dismutase 1 Is Controlled by Both Metal Occupancy and Disulfide Status

Fabio Arnesano, Lucia Banci, Ivano Bertini, Manuele Martinelli,
Yoshiaki Furukawa,and Thomas V. O'Halloran




# Figure S1

**Figure S1** The two-dimensional $^{1}H$-$^{15}N$ HSQC spectra of monomeric E,Zn-Q133M2SOD1$^{SS}$.

# 3.5

## Metal-free SOD1 forms amyloid-like oligomers: a possible general mechanism forfamilial ALS.

Lucia Banci, Ivano Bertini, Stefania Girotto, Manuele Martinelli,
Joan Selvestrone Valentine, Miguela Vierù, Julian Whitelegge

**Submitted**

## Abstract

It is now well established that in SOD1-related FALS disease, aberrant protein oligomerization of the SOD1 mutants is strongly implicated in disease pathology. Here we have found that, even WT SOD1 when lacking both its metal ions, forms large protein aggregates at physiological conditions, i.e. 37 ºC, pH 7 and 100 μM protein concentration. It is here suggested that Cys 6 and Cys 111, which are reduced in the mature active form of SOD1, are implicated, in the absence of metal ions, in formation of disulfide bonds resulting in stable, soluble oligomers which reach a MW of about 800 kDa The covalent nature of the oligomer is confirmed by its disruption upon addition of reducing agents which break the disulfide bonds thus leading to SOD1 monomers. These oligomers are further stabilized by interactions among beta sheets in amyloid-type aggregates. The latter property is evidenced by the interaction of the aggregate with ThT, an amyloid-specific dye and by its disruption upon addition of the chaotropic agent guanidinuim hydrochloride. Our results suggest a general, unifying picture of the process of SOD1 aggregation which can be operative when the SOD1 metalation process is altered or SOD1 fails to maintain its metallated state. Even if we cannot exclude that other mechanisms are operative for SOD1 aggregation, the one here proposed has the strength of rationalize a common event, i.e. protein aggregation, as caused by a quite diverse and large number of SOD1 point mutations. If proved *in vivo*, it could be a starting point for designing therapeutic approaches to FALS and possibly to ALS.

Over 100 different variants of human copper-zinc superoxide dismutase (CuZnSOD, SOD1) have been identified and linked to the neurodegenerative disease familial amyotrophic lateral sclerosis (FALS) by a gain-of-function mechanism (1, 2). Although the exact cellular site and mechanism of the toxicity are unknown, aberrant SOD1 protein oligomerization has been strongly implicated in disease causation (3, 4). Several recent publications have presented compelling evidence that abnormal disulfide-crosslinking of ALS-mutant SOD1 plays a role in this oligomerization, and disulfide-linked SOD1 multimers have been detected in neural tissues of SOD1-ALS transgenic mice that are presumed to be components of higher molecular weight species or intermediates in their formation (4-7).

Wild type (WT) human SOD1 is an exceptionally stable protein, and, although some of the ALS-mutant SOD1 proteins are severely destabilized by their mutations, others retain much of the stability of WT SOD1 (5). Moreover, in the fully demetallated states, some of the ALS-mutant SOD1 proteins are actually more stable than apo WT SOD1 (8). Wild type human SOD1 contains 4 cysteines, Cys 6, 57, 111, and 146. *In vivo,* in the mature protein, Cys 57 and Cys 146 are normally linked by a disulfide bond, a characteristic that is highly unusual for an intracellular protein in the highly reducing environment normally encountered inside the cell. Because of the high stability of the WT and many of the ALS-mutant SOD1 proteins in their mature, disulfide-intact forms, it has generally been assumed that the abnormal disulfide crosslinks, that are associated with SOD1 aggregation, result from oxidative damage to the immature, disulfide-reduced forms of the protein in which all four cysteines are reduced and that either Cys 57, Cys 146, or both are involved (7, 9).

Abnormal protein inclusions have been observed in neural tissue of most of the SOD1-ALS transgenic mice and, in several cases, it has been shown that these inclusions are amyloid-like, based on their observed reactivities with a fluorescent dye frequently

used to visualize the amyloid inclusions found in Alzheimers brains (10). In experiments designed to mimic the conditions that might lead to such inclusions, ALS-mutant SOD1 proteins have been induced to oligomerize *in vitro* to form aggregates that bind the amyloid-specific dye thioflavin T (ThT). However, the conditions used in each case have been far from physiological, i.e., either very low pH (11) or extensive metal-catalyzed oxidation (12), suggesting that oligomerization occurred only if the SOD1 protein was substantially damaged or unfolded.

The current study was undertaken to determine the propensity for oligomerization of mature, i.e., disulfide-intact, WT human SOD1 under the relatively mild conditions likely to be encountered by the protein *in vivo* and to determine the roles of metallation in hindering or promoting such oligomerization. We report here our finding that WT human SOD1 apoprotein, with its intrasubunit disulfide bond intact, forms soluble, ThT-positive, high molecular weight oligomeric assemblies upon incubation in solution at conditions very close to the physiological ones, i.e. at pH 7.0, 37 ºC, and physiological concentration; that these assemblies are remarkably stable, persisting in the soluble state for months; and that metallation with both zinc and copper or even with zinc alone totally suppresses this behavior. These results indicate that human SOD1 proteins, wild type or mutant, when they are metal-free, are highly aggregation-prone under relatively mild conditions, even when the intrasubunit disulfide bond is intact. These results suggest that the gain of toxic function of SOD1 in FALS, and possibly even in sporadic ALS (SALS), may be due to the inability of WT and mutant forms to achieve or to maintain the metallated state that is required to protect against oligomerization *in vivo*.

SOD1 has been estimated to be present in cells at concentrations of approximately 100 μM (13-15). In order to be as close as possible to the physiological situation experienced by the protein *in vivo*, we therefore prepared 100 μM protein solutions at pH 7.0 and incubated them at 37 ºC. Changes in secondary structure were monitored by CD

spectroscopy, and changes in the state of aggregation were followed by gel filtration chromatography and ThT fluorescence. Solutions of fully metallated Cu,Zn-SOD1 (100 μM, pH 7.0) were found to be unchanged after incubation at 37 ºC for more than one month, and the same results were obtained for the copper-free derivative, E,Zn -SOD1. On the contrary, the behavior of the completely metal-depleted protein (apo form) was dramatically different. Incubation in the presence of ThT showed a progressive increase in fluorescence, with a fast process in the first 230 hours of incubation which slows down after one months and then reaches a plateau in the following four months **(Fig. 1)**. No turbidity was detected along the entire time period.

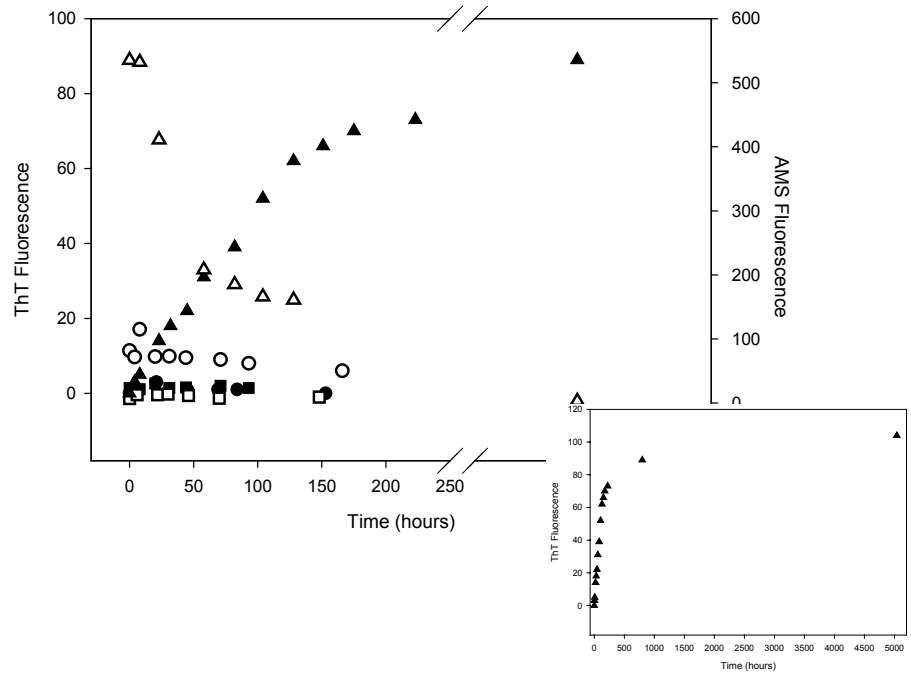



**Fig. 1 Amyloid-like oligomer formation and free cysteines disappearance.**
Binding of thioflavin T to amyloid structures is accompanied by an increase in fluorescence. Thioflavin T does not bind to unaggregated protein. ThT binding to fully metallated SOD1 (●); apo SOD1 (▲ ), copper deprived SOD1 (○), apo AS SOD1 (□), and fully metallated AS SOD1 (■). Fluorescence of AMS bound to the free cysteines of the apo SOD1 WT sample (△).during its incubation at 37 ºC. Inset: ThT binding fluorescence of the apo SOD1 protein during seven months of incubation.

The detection of ThT-binding fluorescence is a clear indication of the formation of amyloid-like oligomeric assemblies (16). The rate of ThT fluorescence enhancement was found to be strongly dependent on temperature; the higher the temperature the faster the rate of aggregation. Different conditions of pH also affected the reaction: all other things being equal, the rate doubled at pH 5 relative to pH 7.

The heterogeneous population of states of SOD1, after incubation at 37 ºC for a variable period of time, was fractioned through gel chromatography, and the fractions were analyzed by multiangle light scattering **(Fig.2)**. The high molecular weight broad band originates from a mixture of species having an average molecular weight that increases as the time of incubation increases (Table.2 in the supplementary material). The maximum molecular weight value observed for a sample incubated for seven months was estimated to be about 800 kDa; this value suggests that this aggregate is formed by fifty SOD1 monomers.

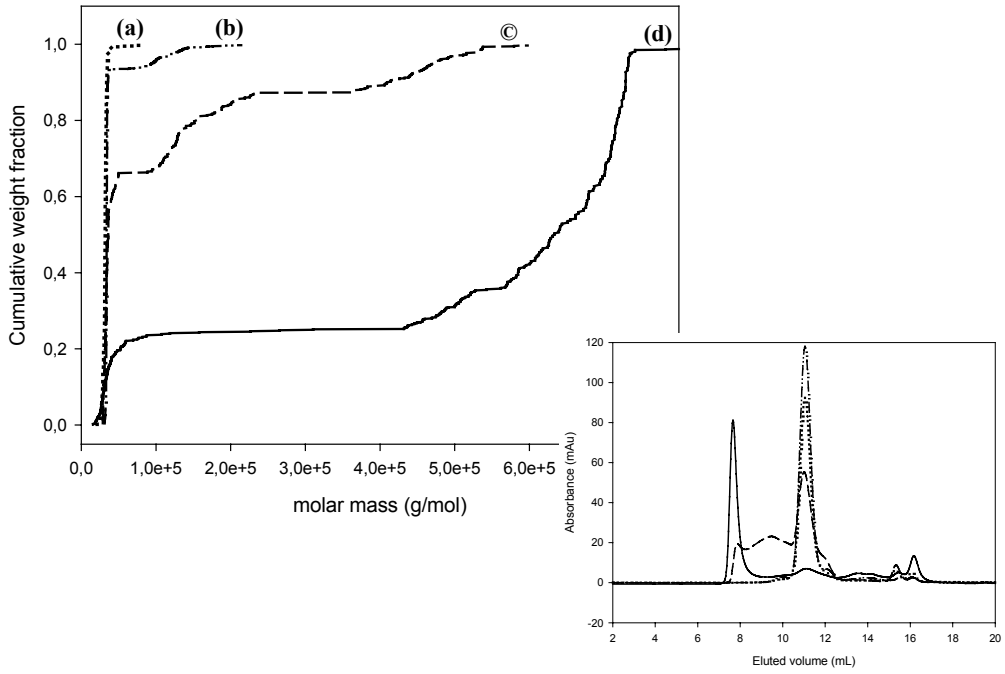



**Fig.2:Light Scattering analysis** Cumulative weight fraction *vs* molar mass reported for the same sample after different periods of incubation: before incubation (a), 8 hours (b), 2.5 days (c) and 7 months (d).

When apo WT SOD1 was incubated in the presence of a reducing agent, either (tris(2-carboxyethyl)phosphine hydrochloride (TCEP) or dithiothreitol (DTT), no enhanced fluorescence was observed in the presence of ThT (Fig.3A), and only the monomeric state was observed through gel filtration (Fig.3B), as is expected for disulfide-reduced apo SOD1 (17). The absence of high molecular weight species in reducing conditions, therefore, shows that the SOD1 oligomeric assemblies involve disulfide bonds

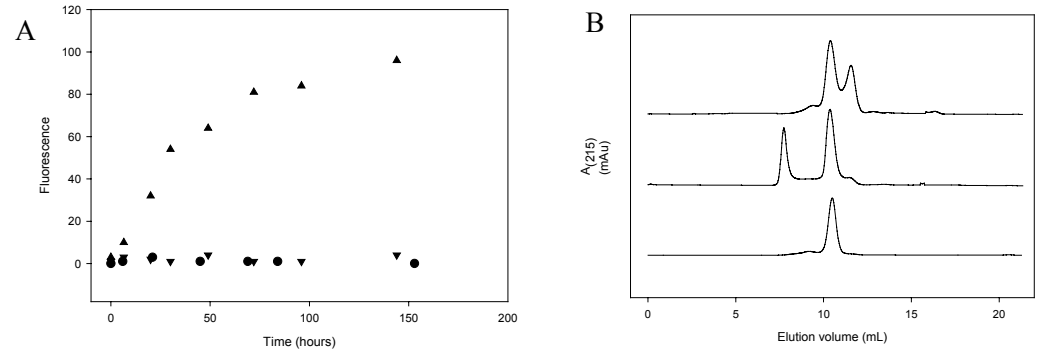



**Fig.3 TCEP addition analysis** (A) ThT binding to fully metallated SOD1 (●); apo SOD1 (▲), and apo SOD1 after the addition of TCEP 1mM (▼). (B) Gel filtration chromatogram showing the monomer, dimer and aggregate composition of fully metallated SOD1 (**a**); apo SOD1 (**b**), and apo SOD1 after the addition of TCEP 1mM (**c**) on a Superdex 75 HR 10/30 column. All samples were analyzed about 50 hours after incubation started.

Freshly prepared, oxidized SOD1 contains two free cysteines (Cys 6 and Cys 111) per SOD1 monomer. Free cysteines content in apo SOD1 (monitored by AMS fluorescence on AMS-modified cysteine samples) decreased simultaneously with protein aggregation (monitored by ThT fluorescence) **(Fig.1)**. The two processes have the same time course: aggregation and free cysteine disappearance are initially fast, successively reaching almost a plateau after nine/ten days. When the chaotropic agent guanidinium hydrochloride (GdnHCl), which disrupts hydrogen bond networks that stabilize secondary and tertiary structure is added to a soluble aggregate of apo WT SOD1, the ThT-binding fluorescence is quenched in 15 minutes. Still, gel filtration of the resulting solution shows that high molecular weight species remain present. This behavior

indicates that the oligomeric assemblies of apo SOD1 are both formed by covalent disulfide bonds probably involving the two free cysteines (Cys6 and Cys111) and stabilized by hydrogen bond interactions, similar to those found in amyloid. Since guanidinium disrupts hydrogen bond networks, the loss of ThT-binding ability is the result of the disruption of the amyloid-like tertiary structure of the assemblies, but the SOD1 oligomeric state due to the covalent S-S bonds between cysteines of the SOD1 subunits, persists.

Based on all these experimental observations, we can hypothesize a process of aggregation where initially a fraction of apo SOD1 is in the monomeric state as a consequence of quaternary structure destabilitation due to lack of metal ions as already described in literature (15, 19). Large molecular weight aggregates are then assembled through disulfide bonds formation and further stabilized by non covalent interactions between beta sheets (Fig. 4).

From all these data it appears that SOD1, when lacking both its metal ions but still maintaining the disulfide bond intact, has a dramatic tendency to form high molecular weight soluble aggregates in conditions very close to the physiological ones. This process occurs *in vitro* as well as *in vivo*, in *E.coli* cells, both for WT SOD1 and its mutants, provided Cys6 and Cys111 are present. A failure to either activate and/or maintain the proper metallation of SOD1 makes the two free cysteines prone to be oxidized and form inter-disulfide bonds, therefore initiating the aggregation process. A large number of ALS-related SOD1 mutants have a decreased metal affinity and indeed some fail even to be metal reconstituted *in vitro* (20). Also for WT, the protein is produced in an immature form which then needs several steps to reach the final active form. Among the various posttranslational modifications SOD1 should undergo, metallation is a complex process which involves several processes and interactions with other proteins. Metal uptake might be altered by misregulation of any of these steps by external factors. Therefore, when

some SOD1 is present in the immature, demetalated state, the aggregation process can start giving rise to amyloid-like soluble oligomers. We have also to take into account that SOD1 is present in two different cell compartments, i.e. cytoplasm and mitochondria, and it independently acquires metal ions in the two cell compartments. These two cell compartments have quite different redox properties which can further modulate the aggregation process *in vivo*. This is also consistent with a recent suggestion of Wang et al. (4); they report that in *in vivo* experiments non-native intermolecular disulfide bonds help stabilize the aggregates and that the redox state of the cell may play a role in the aggregation process.

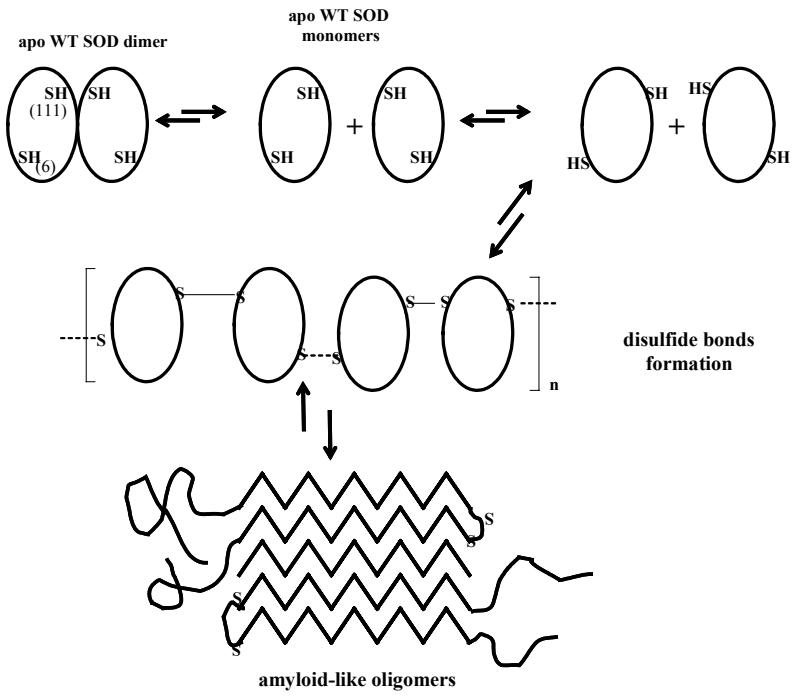



**Fig. 4 Proposed mechanism for SOD1 aggregation.** Possible mechanism for *in vivo* formation of amyloid-like oligomers which occur when the apo WT SOD1 protein is kept at 37 ºC for a period of time. Dimeric apo SOD1 is in equilibrium with the monomeric species which are known to be quite mobile. This internal dynamics could make accessible locally disordered conformations when the two free cysteines (Cys6 and Cys111) become solvent exposed. In this state, they are prone to form In this , they are prone to form disulfide bonds and therefore oligomeric assemblies. The latters arrange in amyloid-like structures through hydrogen bonds.

This mechanism of SOD1 aggregation can be quite general since it does not depend on a specific mutation and could provide a unified picture of the general process of aggregation observed for SOD1 mutants, though other pathways cannot be excluded. The amyloid-like soluble oligomers observed under mild conditions appear to us to be likely candidates for the toxic species that cause SOD1-linked ALS. The knowledge of the factors determining SOD1 aggregation is likely to be the starting point to develop therapeutic strategies having the goal of finding exogenously applied agents that maximize the formation of the mature, fully metalated SOD1 form or interfere with the aggregation process.

*Acknowledgments*

This work was supported by the European Community: Understanding Protein Misfolding and Aggregation by NMR (UPMAN) (n° LSHG-CT-2004-512052 (1/11/04-31/10/07)), by Marie Curie Host Fellowships for early stage research training (n°MEST-CT-2004-504391 " NMR in Inorganic Structural Biology "). It was also supported by NIH grants DK46828 and NS049134 to JSV.

References

1. Valentine, Doucette & Potter (2005) *Annu Rev Biochem* **74**, 563-593.
2. Bruijn, L. I., Miller, T. M. & Cleveland, D. W. (2004) *Annu. Rev. Neurosci.* **27**, 723-749.
3. Jonsson, P. A., Ernhill, K., Andersen, P. M., Bergemalm, D., Brannstrom, T., Gredal, O., Nilsson, P. & Marklund, S. L. (2004) *Brain* **127**, 73-88.
4. Wang, J., Xu, G. & Borchelt, D. R. (2006) *J. Neurochem.* **96**, 1277-1288.
5. Jonsson, P. A., Graffmo, K. S., Andersen, P. M., Brannstrom, T., Lindberg, M., Oliveberg, M. & Marklund, S. (2006) *Brain* **129**, 451-464.
6. Deng, H. X., Shi, Y., Furukawa, Y., Zhai, H., Fu, R., Liu, E., Gorrie, G. H., Khan, M. S., Hung, W.-Y., Bigio, E. H. *et al.* (2006) *Proc. Natl. Acad. Sci. U. S. A* **103**, 7142-7147.
7. Furukawa, Y., Fu, R., Deng, H. X., Siddique, T. & O'Halloran, T. V. (2006) *Proc. Natl. Acad. Sci. U. S. A* **103**, 7148-7153.
8. Rodriguez, J. A., Shaw, B. F., Durazo, A., Sohn, S. H., Doucette, P. A., Nersissian, A. M., Faull, K. F., Eggers, D. K., Tiwari, A., Hayward, L. J. *et al.* (2005) *Proc. Natl. Acad. Sci. U. S. A* **102**, 10516-10521.
9. Furukawa, Y. & O'Halloran, T. V. (2005) *J. Biol. Chem.* **280**, 17266-17274.

10. Wang, J., Xu, G. L., Gonzales, V., Coonfield, M., Fromholt, D., Copeland, N. G., Jenkins, N. A. & Borchelt, D. R. (2002) *Neurobiol. Dis.* **10**, 128-138.
11. DiDonato, M., Craig, L., Huff, M. E., Thayer, M. M., Cardoso, R. M., Kassmann, C. J., Lo, T. P., Bruns, C. K., Powers, E. T., Kelly, J. W. *et al.* (2003) *J. Mol. Biol.* **332**, 601-615.
12. Rakhit, R., Cunnigham, P., Furtos-Matei, A., Dahan, S., Qi, X. F., Crow, J. P., Cashman, N. R., Kondejewski, L. H. & Chakrabartty, A. (2002) *J. Biol Chem.* **277**, 47551-47556.
13. Kurobe, N., Suzuki, F., Okajima, K. & Kato, K. (1990) *Clin. Chim. Acta* **187**, 11-20.
14. Lindenau, J., Noack, H., Possel, H., Asayama, K. & Wolf, G. (2000) *Glia.* **29**, 25-34.
15. Rakhit, R., Crow, J. P., Lepock, J. R., Kondejewski, L. H., Cashman, N. R. & Chakrabartty, A. (2004) *J. Biol Chem.* **279**, 15499-15504.
16. Krebs, M. R., Bromley, E. H. & Donald, A. M. (2005) *J. Struct. Biol.* **149**, 30-37.
17. Arnesano, F., Banci, L., Bertini, I., Martinelli, M., Furukawa, Y. & O'Halloran, T. V. (2004) *J. Biol. Chem.* **279**, 47998-48003.
18. Banci, L., Bertini, I., Cabelli, D. E., Hallewell, R. A., Tung, J. W. & Viezzoli, M. S. (1991) *Eur. J. Biochem.* **196**, 123-128.
19. Lindberg, M. J., Normark, J., Holmgren, A. & Oliveberg, M. (2004) *Proc. Natl. Acad. Sci. U. S. A* **101**, 15893-15898.
20. Hayward, L. J., Rodriguez, J. A., Kim, J. W., Tiwari, A., Goto, J. J., Cabelli, D. E., Valentine, J. S. & Brown, R. H. J. (2002) *J. Biol. Chem.*

**Materials and Methods**

*Sample Preparation*— Wild type (WT) SOD1 expressed in yeast, was purified in the fully metallated form (1). $N^{\alpha}$-acetylation (alanine), as expected for human SOD1, was confirmed by mass spectroscopy analysis. WT SOD1 was also expressed in the *Escherichia coli* BL21(DE3) strain. The protein was obtained by growing the cells in the LB medium. The protein was isolated by osmotic shock in a 20mM Tris, 5mM dithiothreitol (DTT) buffer at pH 8. After incubation for 30 minutes at 37 ºC, the protein was centrifuged at 40000 rpm for 20 minutes. Supernatant was purified following a reported procedure (2) modified by the addition of 1 mM DTT to each chromatographic buffer. The protein obtained with this procedure contained substoichiometric amounts of the metal ions (Table S1). Mutations were performed using a QuikChange™ site-directed mutagenesis kit (Stratagene).

The demetallated (apo) protein was prepared according to previously published protocols (3). Zinc reconstitution was carried out as previously described (4). Metal content of the various forms of SOD1 was checked by inductively coupled plasma mass

spectrometry (ICP-MS) using a Thermo Jarrell Ash Atomscan Model 25 Sequential inductively coupled spectrometer (Table S1).

*Spectroscopic characterization*—Protein samples were 100 μM in SOD1 concentration (as dimer) in 20mM phosphate buffer at pH 7. The protein was incubated at 37 °C to mimic physiological conditions. Optical and fluorescence spectroscopies, coupled with gel filtration chromatography, were used to monitor the formation of oligomeric species at these sample conditions.

Fluorescence was followed with Thioflavin T, (ThT) probe, which specifically binds to amyloid-like structures (6). Free ThT has excitation and emission maxima at 350 nm and 450, respectively. However, upon binding to amyloid-like oligomers, the excitation and emission λ changes to 450 and 485 nm, respectively. 54μl aliquots of sample were added to 646 μl of a 215 μM ThT solution in a 20 mM phosphate buffer at pH 7. The solution fluorescence emission was measured, over time of incubation,with a Cary 50 Eclipse Spectrophotometer supplied with a Single cell Peltier thermostatted cell holder regulated at 37 ºC. The background fluorescence spectrum of the buffer was subtracted. The excitation wavelength was 446 nm (bandwidth 10 nm) and the emission was recorded at 480 nm (bandwidth 10 nm). Fluorescence intensity at 483 nm was plotted against time of incubation.

Turbidity was measured at 400 nm to detect possible formation of insoluble precipitate. Solution turbidity was measured as apparent absorbance at 400 nm using a Cary UV-visible spectrophotometer. Experiments were performed by diluting 120 μl of the incubation SOD1 stock into 280 μl of 20mM phosphate buffer at pH 7. A 1 cm quartz cuvette was used. Instrumental detection limit was 0.001 at 400 nm.

*Monitoring SOD1 Aggregation by Gel Filtration and Light Scattering*—100 μl aliquots of the incubated protein at 37 ºC were periodically taken and analyzed by gel

filtration on Superdex 75 HR 10/30 (Amersham Biosciences) at room temperature. The column was preequilibrated with 20 mM potassium phosphate, pH 7.0, and the flow rate was 0.6 ml/min. The chromatogram, which monitors the species formed during incubation, was obtained by monitoring the absorbance at 215 nm. The Superdex 75 HR 10/30 column was connected to a light scattering spectrometer. The online multiangle light scattering (MALS) detector (DAWN EOS, Wyatt Technology, Santa Barbara, CA) and differential refractive index (DRI) detector (Optilab DRI, Wyatt Technology) setup was used to measure the light scattered as a function of angle and absolute protein concentration of fractions eluting from the size-exclusion chromatography column. The Zimm approximation was used in the Astra software (Wyatt Technology) to estimate molar mass. Data were fit using a first-order polynomial. The analysis was performed for each one of the 100 μl aliquots periodically taken from the incubation batch so as to monitor the increase in molecular weight of the soluble species formed during aggregation.

*Free-thiol quantification*—Estimation of free thiols during aggregation was performed by 4-acetamido 4' maleimidylstilbene-2,2'-disulfonic acid (AMS) modification. This iodoacetamide derivate has high water solubility and is readily conjugated to free thiols. 800 μl of reaction mixture in phosphate buffer 20 mM pH 7 containing proteins 7.5 μM, AMS 250 μM, SDS 1% was incubated for 30 minutes at 37 °C in order to complete the reaction. The excess of AMS was removed by dialysis. Reacted aliquots were taken from an apo WT SOD1 sample (100 μM) along the incubation at 37 ºC. AMS is a stilbene derivate, and shows a typical UV absorption at around 328 nm and emission maximum at 408 nm. Fluorescence measurements (excitation 322 nm, emission 406 nm, excitation/emission slits 10) were performed in order to monitor the variation of free thiols during the aggregation process. The calibration curve for the free cysteine quantification was prepared using freshly prepared apo SOD1 WT at different concentrations as a

standard. MALDI analysis proved that there is a maximum of two free cysteines per apo SOD1 monomer.

Aliquots taken from the incubation stock used for the AMS test, were simultaneously reacted with ThT, and the fluorescence of the solution was measured to monitor sample aggregation.

Materials and Methods References


1. Banci, L., Bertini, I., Luchinat, C. & Hallewell, R. A. (1988) *J. Am. Chem. Soc.* **110**, 3629-3633.
2. Banci, L., Benedetto, M., Bertini, I., Del Conte, R., Piccioli, M. & Viezzoli, M. S. (1998) *Biochemistry* **37**, 11780-11791.
3. McCord, J. M. & Fridovich, I. (1969) *J. Biol. Chem.* **244**, 6049-6055.
4. Beem, K. M., Rich, W. E. & Rajagopalan, K. V. (1974) *J. Biol. Chem.* **249**, 7298-7305.
5. Deleage, G. & Geourjon, C. (1993) *Comp. Appl. Biosc.* **9**, 197-199.
6. Krebs, M. R., Bromley, E. H. & Donald, A. M. (2005) *J. Struct. Biol.* **149**, 30-37.
7. Wang, J., Xu, G. & Borchelt, D. R. (2006) *J. Neurochem.* **96**, 1277-1288.
8. Whitelegge, J. P., le Coutre, J. L. J. C., Engel, C. K., Prive, G. G., Faull, K. F. & Kaback, H. R. (1999) *Proc. Natl. Acad. Sci. U. S. A* **96**, 10695-10698.
9. Whitelegge, J. P., Gundersen, C. B. & Faull, K. F. (1998) *Protein Sci* **7**, 1423-1430.

# 4

# GENERAL DISCUSSION AND PERSPECTIVE

S*tructural biology* provides a detailed analysis of the structure of biological macromolecules at atomic level, with the ultimate goal of investigating the relationship between structure and function. Metalloproteins, which comprise a significant share of the proteome, contain biologically relevant metal ions which can have one or more accessible oxidation states. The investigation of mechanisms how metals are transported and inserted into diverse protein locations is an emerging field in the scientific world, which is unravelling only in the most recent years. Therefore structural characterization of metalloproteins is an important aspect of post-genomic research. Out activity is integrated into this framework and intends to contribute to the coverage of metalloprotein structures and to pursue functional studies aiming to elucidate the role of metal cofactors, their redox properties, and implications in protein-protein interactions.

Intriguingly, several metalloproteins are involved in diseases, and often they contain copper[(2)]. Copper is a naturally occurring redox metal essential for the activity of many enzymes, but it is highly reactive and toxic at high concentrations. Among the factors required to achieve copper ion homeostasis are the metallochaperones[(3)], that specifically traffic copper to cuproenzymes. Different copper trafficking pathways are indeed present in the cell, i.e. copper delivery to cytosolic superoxide dismutase, to cytochrome *c* oxidase (CcO) and to the multicopper oxidase Fet3. During the PhD course my attention was focused on three copper proteins (Cu,Zn SOD, Sco1 and Cox17) belonging to the first two pathways.

Copper incorporation into CcO is biologically crucial; it is, however, an extremely complex process, which is tightly regulated and which requires a large array of proteins, each of them with a specific, in most of the cases, non-replaceable role. The knowledge of these processes is still at an infant state. Therefore, to have a comprehensive description of proteins involved in copper delivery to CcO, we have performed a genome-wide search in prokaryotic organisms for sequences sharing similarity with human Cox17, Sco1/2, and Cox11, and we extended the analysis to genes close to the found proteins.

We found that, while Cox11 is highly networked within the CcO assembly module and likely fulfills an univocal function of copper delivery to $Cu_B$, the variety of number and localization for Sco genes suggesting that, the related proteins, could have multifunctional properties and be involved in different physiological processes. In this analysis a strong correlation was observed for Sco genes with both red-ox and copper homeostasis. The structure of Sco has indeed a thioredoxin fold and a thiol-disulfide oxidoreductase function

has been proposed for this protein[4]. Therefore, Sco might be involved, in addition to its well documented role in copper transfer to $Cu_A$ center[5], in the reduction of disulfide bonds of its protein patner Cox2 prior to copper insertion. The presence of two Sco paralogs in eukaryotes (i.e., Sco1 and Sco2) can be also rationalized in the light of the genomic-context analysis of prokaryotic paralogs: one Sco gene may preferentially interact with Cox2 subunit of CcO, thus favoring CcO assembly, while the second gene may assist the metallochaperone Cox17, which is responsible for copper recruitment in the intermembrane space of mitochondria. In this scenario, it is possible that the two eukaryotic Sco paralogs interact to promote copper insertion into $Cu_A$. Similar conclusions for human Sco genes were reached using an experimental approach[6]. Many other lines of evidence support the hypothesis that copper insertion into $Cu_A$ site of Cox2 subunit, is not a simple, passive copper transfer from Cox17 to Sco1/2 proteins and finally to Cox2, but that several redox reactions are probabily interlinked with the process of copper transfer

The structure of the metal adducts and the protein-protein interaction analysis are crucial for better understanding the mechanism of Sco-mediated copper insertion into CcO. Therefore, we cloned and expressed the WT-HSco1 and P174L-HSco1 pathogenic mutant. We have succeeded in preparing Cu(I) and Ni(II) derivatives of HSco1 and determined their solution structures. The solution structure of apoHSco1 has been also determined. The structures indicate that the metal ion is bound by two cysteines and one histidine. They also show the transition from a locally disordered apo protein to a compact metallated form, as confirmed also by ESI-MS studies. We obtained crystals of Ni(II)HSco1, and the relative X-ray structure suggest the binding of the metal ion to the oxidized form of HSco1. The species, described from the above mentioned structure, may be represent a transition state of the copper transfer from HSco1 to the $Cu_A$ site of Cox2. It may represent a missing link, which integrates the metal transfer and thioredoxin functions already proposed for this fascinating protein.

The solution structure of P174L-HSco1 mutant showed that the mutation significantly affects the metal binding properties of HSco1 protein. Structurally, the mutation produces the lack of well-organized hydrophobic contacts and a structural heterogenity in the vicinity of the metal binding region. The local structural changes induced by the point mutation decrease the copper(I) affinity of the mutant, thus negatively affecting the co-metallochaperone function of HSco1. ESI-MS and NMR interaction studies between HSco1 and HCox17 demonstrate that $Cu_1$(I)-$HCox17_{2S\text{-}S}$, is capable of transfering efficiently Cu(I) to apoWT-HSco1, but not to apoP174L-HSco1. Morever the oxidation

and reduction rates of the mutant disulfide bond are much slower than those of the wild-type protein. The thioredoxin role of HSco1, proposed for maintaining the receiving $Cu_A$ site in the suitable reduced state, is, in this case, significantly perturbed since the mutation could slow down the reduction of the copper binding cysteines in the $Cu_A$ site of CcO. Overall, these alterations, induced by the mutation, on copper binding and redox properties of HSco1 can substantially decrease the efficiency of the copper transfer from $Cu_1$(I)HSco1 to the $Cu_A$ site of CcO. All these data support the multifunctional role of HSco1, which may act at the same time, as copper chaperone and thioredoxin in the process of copper insertion in the $Cu_A$ site.

The functional role of the paralog HSco2 protein is not yet defined as well as the redox function of the Sco protein family. Therefore, further studies are necessary to fully elucidate the assembly machinery of CcO. The next step will be the investigation of structural and biochemical properties of HSco2. Moreover, nothing is known at the moment about how the two cysteines of CXXXC conserved motif of HSco1/2 are reduced prior to copper transfer from Cox17. Therefore interaction studies between HSco1/HSco2 and HCox17 in both reducing and oxidizing environment will be performed.

The structure-function paradigm claims that a specific function of a protein is determined by its unique and rigid three-dimensional (3D) structure. Thus, following the biosynthesis of a protein in the ribosome, it should fold to be functional. This idea represents one of the cornerstones of modern biology. This concept is confirmed by numerous cases describing polypeptide chains that loose the capability of gaining a proper 3D structure(i.e. became misfolded) due to the effect of environmental factors or because of genetic defects (mutations). Consequences of such misfolding are well known and represent loss of function, aggregation, development of conformational disorders and cell death.

ALS may be classified as a “conformational” disorders. In this class of disorders, protein alteration, through an abnormal folding pathway, generally results in protein aggregation and high-molecular weight complexes formation (**Fig. 19**). Our target is the understanding of the relationships between structural perturbations of the ubiquitous SOD1 and the neurodegenerative disorder ALS.

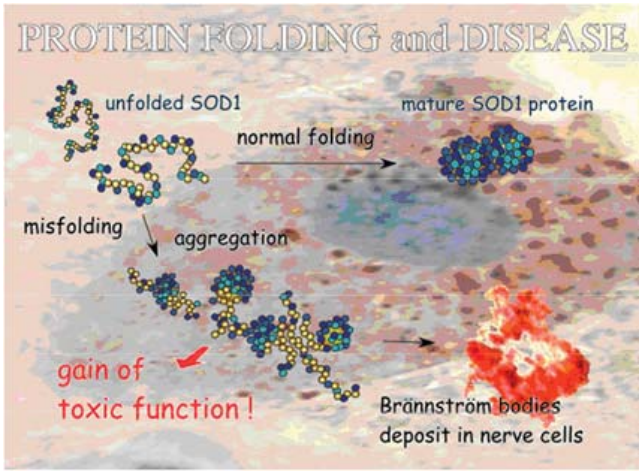



**Fig. 19:** *Example of the two possible different pathways, normal and aberrant, to which the immature SOD1 may be subjected.*

Over 100 different variants of human copper-zinc superoxide dismutase (CuZn-SOD1) have been identified and linked to the neurodegenerative disease familial amyotrophic lateral sclerosis (fALS). Recent studies of a large number of ALS-mutant Cu,Zn-superoxide dismutase (SOD1) proteins have revealed that there is great diversity in the biophysical properties of these proteins. Nowadays, no common denominator of SOD1 pathogenic mutant are able to explain their aggregation propensity with respect to WT-SOD1. Some of the ALS-mutant SOD1 proteins, in both apo and/or fully metallated form, are severely destabilized by their mutations, others retain much of the stability of WT SOD1[7,8]. What is in known, is that holo-form of WT-SOD1 (Cu,Zn-SOD1) is an exceptionally stable protein to the extent that its melting temperature, Tm, is around 90°C[9]; however, removing a copper ion from holo-SOD1(E,Zn-SOD1) decreases the Tm to ~70°C, and further removal of the Zn ion from E,Zn-SOD1 (E,E-SOD1) results in ~60°C of Tm[10]. SOD1 has to undergo several post-translational modifications before reaching its mature form. The protein, requires insertion of zinc and copper atoms, followed by the formation of a conserved S-S bond between Cys-57 and Cys-146 (human numbering), which makes the protein fully active. Given that these modifications generally increase the protein structural stability, it is well expected that most immature, the apo and disulfide-reduced SOD1 (E,E-SOD1$^{SH}$), and the other possible intermediate forms (for example: E,Zn-SOD$^{SH}$ or E,E-SOD$^{SS}$), are more susceptible to destabilization and aggregation.

Using biophysical methods, we investigated the difference in the secondary, tertiary and quaternary structure of both oxidated and reduced SOD1 in all the possible different metallation states. From this investigation we show that, even after removal of both copper and zinc ions from the active and mature form of human SOD1, the dimeric state still

persists; however, upon reduction of the disulfide bond, the protein can readily dissociate to the monomer form. Zn(II) addition to the reduced apo-SOD1 restores the dimeric state, indicating that only the most immature form of SOD1, before any post-translational modifications, favors the monomeric state, even if no major changes in the secondary structure were detected.

It has been proposed that protein monomerization plays a role in formation of misfolded intermediates, leading to protein aggregation[11]. In conclusion, the observed perturbation of the SOD1 quaternary structure by disulfide reduction and demetallation may be relevant to the etiology of both sporadic and familiar ALS.

In the light of this first interesting result we proceeded to determine the propensity for oligomerization of mature, i.e., disulfide-intact, WT human SOD1 under relatively mild conditions likely to be encountered in the protein "*in vivo"* and to determine the role of metallation in hindering or promoting such oligomerization.

We found that demetallation of SOD1 induces the formation of soluble high molecular weight aggregates upon incubation at 37 °C, while the fully metallated WT SOD1 retains, under the same conditions, the dimeric state. Lack of aggregate formation for the metallated protein shows that the metallation state is a critical issue in the aggregation process. Formation of intermolecular disulfide bonds, probably involving the Cys6 and Cys111, is an essential factor for aggregation, as proved by absence of aggregation under reducing conditions. Disulfide bonds involvement in amyloid fibrils formation has previously been reported for other neurodegenerative diseases[12,13].

The increase fluorescence intensity upon ThT binding to the apo-SOD1 suggest the formation of amyloid-like oligomeric assemblies. The finding of this potentially pathological, soluble protofibrillar structure in WT SOD1 might have relevance with respect to ALS, pointing at a common behaviour with many other neurodegenerative disorders. In fact a general pathological feature of many neurological disorders is the presence of inclusions bodies and other visible protein aggregates, which might represent the end stage of a molecular cascade of several steps. Earlier steps in this cascade are more directly tied to pathogenesis than the aggregates themselves[14], and are the result of the pathogenic species elimination induced by the cells. The formation of the protofibrillar specie in apo WT SOD1 suggests a common explanation for both, familiar and sporadic ALS. Incorrect and/or incomplete SOD1 metallation starts the onset of pathogenic events, and the mutations could be only a time modulation of the disease onset and development. This hypothesis is consistent with the fact that most of SOD1 pathogenic mutants have a

decreased affinity for one or both the native metal ions[15,16]. This would also explain the widespread of fALS mutations, with no evident correlation with structural properties.

In conclusion our data suggest a possible general aggregation mechanism common for both sporadic and familiar ALS. Recently In has been shown that a significant fraction of the insoluble SOD1 aggregates in spinal cord of the ALS-model transgenic mice contain multimers cross-linked via intermolecular disulfide bonds. At this point, in order to validate our model, it will be necessary to test, for the presence of disulfide-linked multimers, the spinal cord from of transgenic mice expressing SOD1 which lacks the two not conserved cysteine residues, Cys-6 and Cys-111 that we suggest to be involved in the aggregate formation.

Another point to address is how SOD1 incorrect metallation can be active also *in vivo*. ALS afflicted individuals express SOD1 from birth in all tissues[17], even though the pathology in this disease is largely limited to motor neurons. Neurofilaments-L (NF-L) are one of the most abundant proteins in motor neurons; their presence in aggregates is a pathologic hallmarks of both sporadic and familial ALS[18]; they have a strong affinity for zinc[16] and form aggregates that are positive to the thioflavin T binding assay, in the presence of copper and $H_2O_2$[19]. Therefore, under specific conditions, they could remove zinc and/or copper from both WT and mutant SOD1 explaining, in this way, the restricted localization of SOD1 inclusions in motor neurons. This hypothesis is supported by the fact that the absence of neurofilaments slows the disease onset in mice expressing a specific fALS mutant[20].

# Reference list


(1) Trojanowski, J.Q.; Lee, V.M; (2000) Ann.N.Y.Acad.Sci. **924**, 62-67.

(2) Strausak, D.; Mercer, J.F.; Hermann, H.D.; Stremmel, W.; Multhaup, G. (2001) *Brain Res.Bull.* **55**,175-185

(3) Pufahl, R.; Singer, C.P.; Peariso, K.L.; Lin, S.J.; Schmidt, P.J.; Fahrni, C.J.; Culotta, V.; Penner-Hahn, J.E.; O'Halloran, T.V. (1997) *Science* **278**, 853-856.

(4) Balatri, E.; Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S. (2003) Structure **11**, 1431-1443.

(5) Carr, H.S.; Winge , D.R. (2003) *Acc.Chem.Res* **36**, 309-316.

(6) Leary, S. C.; Kaufman, B. A.; Pellecchia, G.; Guercin, G. H.; Mattman, A.; Jaksch, M.; Shoubridge, E. A. (2004) Hum. Mol. Genet **13**, 1839-1848.

(7) Jonsson, P.A.; Graffmo, K.S.; Andersen, P.M.; Brannstrom, T.; Lindberg, M.; Oliveberg, M.; Marklund, S. (2006) *Brain* **129**, 451-464.

(8) Rodriguez, J.A.; Shaw, B.F.; Durazo, A.; Sohn, S.H.; Doucette, P.A.; Nersissian, A.M.; Faull, K.F.; Eggers, D.K.; Tiwari, A.; Hayward, L.J. (2005) *Proc. Natl. Acad. Sci. U. S. A* **102**, 10516-10521.

(9) Forman, H.J.; Fridovich, I. *J Biol Chem* (1973) 248, 2645–2649.

(10) Roe, J.A.; Butler, A.; Scholler, D.M.; Valentine, J.S.; Marky, L.; Breslauer, K.J. (1988) *Biochemistry* **27**, 950–958.

(11) Lindberg, M.J.; Tibell, L.; Oliveberg, M. (2002) *Proc. Natl Acad Sci. U. S. A.* **99**, 16607–16612.

(12) Welker, E.; Wedemeyer, W.J.; Scheraga, H.A. (2001) *Proc. Natl. Acad. Sci. U. S. A* **98**, 4334-4336.

(13) Chen, Y. & Dokholyan, N. V. (2005*) J. Mol. Biol.* **354**, 473-482.

(14) Ross, C. A.; Poirier, M. A. (2006) *Nat. Med.* **10**, S10-17

(15) Lynch, S.M.; Boswell, S.A.; Colon, W. (2004) *Biochemistry* **43**, 16525-16531.

(16) Crow, J. P.; Sampson, J.B.; Zhuang, Y.; Thomson, J.A.; Beckman, J.S. (1997) *J. Neurochem.* **69**, 1936-1944.

(17) Lindberg, M.J.; Bystrom, R.; Boknas, N.; Andersen, P.M.; Oliveberg, M. (2005) *Proc. Natl. Acad. Sci. U. S. A* **102**, 9754-9759.

(18) Chou, S.M.; Wang, H. S.; Komai, K. (1996) *J. Chem. Neuroanat.* **10**, 249-258.

(19) Kim, N.H.; Kang, J.H. (2003) *J Biochem Mol Biol.* **36**, 488-92.

(20) Williamson, T.L.; Bruijn, L.I.; Zhu, Q.; Anderson, K.L.; Anderson, S.D.; Julien, J.P.; Cleveland, D.W. (1998) *Proc. Natl. Acad. Sci. U. S. A* **95**, 9631-9636.