UNIVERSITÀ DEGLI STUDI FIRENZE

DOTTORATO DI RICERCA
INTERNATIONAL DOCTORATE IN MECHANISTIC AND STRUCTURAL SYSTEMS BIOLOGY

CICLO XXVI

COORDINATORE Prof. Claudio Luchinat

Protein expression strategy for in-cell NMR

Settore Scientifico Disciplinare CHIM/03

Dottorando
Dott. Letizia Barbieri

Tutore
Prof. Lucia Banci

Coordinatore
Prof. Claudio Luchinat

Anni 2011/2013

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

Universiteit Utrecht

CERM Firenze

1 Introduction

1.1 Aim of the research

The purpose of my research activity has been to develop cellular and molecular biology and isotopic labeling strategies to enable optimization of expression of proteins in prokaryotic and eukaryotic cells to characterize them, their interactions and their functional processes in their native environment at atomic level. This could be achieved exploiting in-cell NMR spectroscopy. Protein folding, post translational modifications, such metal ions binding or cysteine oxidation, can be studied in the physiological environment in which the protein of interest operates; also the effects of environmental aspects such as molecular crowding, redox properties, on the properties of the biomolecules and of their functions, can be investigated.

1.2 In-cell NMR Spectroscopy

Atomic resolution methods for the structural characterization of biomolecules are often confined to *in vitro* experimental setups. Macromolecular and, in particular, protein knowledge is largely built on studies of isolated proteins in ideal, diluted solutions. These conditions are vastly different from the physiological environment in cells, which contain complex components and are crowded with macromolecules, their total concentration possibly exceeding 300 g L^{-1}. X-ray crystallography and high resolution electron microscopy are only *in vitro* approaches due to their requirement for pure samples and crystalline or vitrified specimens. Nuclear magnetic resonance (NMR) spectroscopy, the only other method for structural investigations at atomic level, allows for the direct observation of NMR-active nuclei within any NMR-inactive environment and can thus be employed to analyze biomolecules *in vivo* and inside cells. In-cell NMR spectroscopy employs modern methods of isotope labelling and multi-dimensional, isotope-edited correlation experiments to obtain structural information on proteins within living cells. This approach enables high-resolution snapshots of intracellular protein conformation. These structural 'fingerprints' may change upon biological interactions, post-translational modifications or due to structural rearrangements[1].

Several in-cell NMR applications for structural and functional studies of proteins in bacteria have been reported[2-6]. Bacterial in-cell NMR techniques have been successfully employed to analyze protein dynamics[7-8], protein-protein interactions[9], for de-novo resonance assignments[10] and automated structure determinations in crude cell extracts[11]. Also a protein structure determination have been achieved: the structure of the putative heavy-metal binding protein TTHA1718 from *Thermus thermophilus* HB8 overexpressed in *Escherichia Coli* cells was solved[12]. A further application of in-cell NMR spectroscopy is the investigation of the tautomerization and protonation state of histidines in the cellular environment, since this amino acid is frequently found in the active site of enzymes[13].

Why do we wish to extend the applicability of in-cell NMR measurements to eukaryotic cells? Prokaryotic organisms exhibit a limited range of biological activities and many of the cellular processes that define the prevalent topics in modern biological research are absent in bacteria. Post-translational protein modifications serve to regulate biological activities in eukaryotes, but are much less common in prokaryotic organisms. The presence of organelles and the requirement of regulated cellular transport constitutes another characteristic of eukaryotic cell identity. Compartmentalization *per se* results in the creation of sub-cellular environments with different physical and biological properties and little is known about the effects that these compartments exert on a protein's structure or function. Moreover, eukaryotic organisms display a high degree of cellular differentiation, which leads to functional specification and differences in biological activities in neighbouring cell types.

The first in-cell NMR study of isotopically enriched proteins in higher eukaryotic cells was conducted in *Xenopus laevis* oocytes[14]. Their large size (~1 mm diameter) allows microinjection of isotopically labelled proteins, produced in bacteria, into the cytosol. In this system, phosphorylation of folded and intrinsically disordered proteins has been studied and also structures of nucleic acids has been investigated[15-18]. Alternatively, the isotopically labelled proteins could be delivered into the cytoplasm via CPP (cell-penetrating peptides) linked covalently to the proteins. The proteins are subsequently released from CPP by endogenous enzymatic activity or by autonomous reductive cleavage. This technique was successfully apply to different proteins, demonstrating the broad application of this procedure to studying interactions and protein processing. Specifically, the CPP sequence of Tat from HIV-1 was fused to the carboxy terminus of

the human ubiquitin and then incubated with human HeLa cells in the presence of pyrenebutyrate, which mediates the direct translocation of CPP-linked proteins into the cytosol[19].

Another method to insert an isotopically labelled protein into the cytosol was developed by Ogino *et al.* They took advantage of the pore-forming toxin streptolysine O to study the protein thymosin β4 in human 293F cells[20]. One attractive feature of this approach, in contrast to the CPP strategy, is that no modifications to the protein are required for intracellular sample delivery. In-cell NMR spectra are thus uncompromised by CPP tags or by additional amino acids that remain present after intracellular tag removal. One disadvantage is the requirement for mammalian cells that grow in suspension and for highly soluble target proteins. Because the amount of delivered sample depends on passive diffusion through the toxin pore, and not on an active uptake process like the CPP case, the protein concentration gradient across the cell membrane directly relates to the overall transduction efficiency[21].

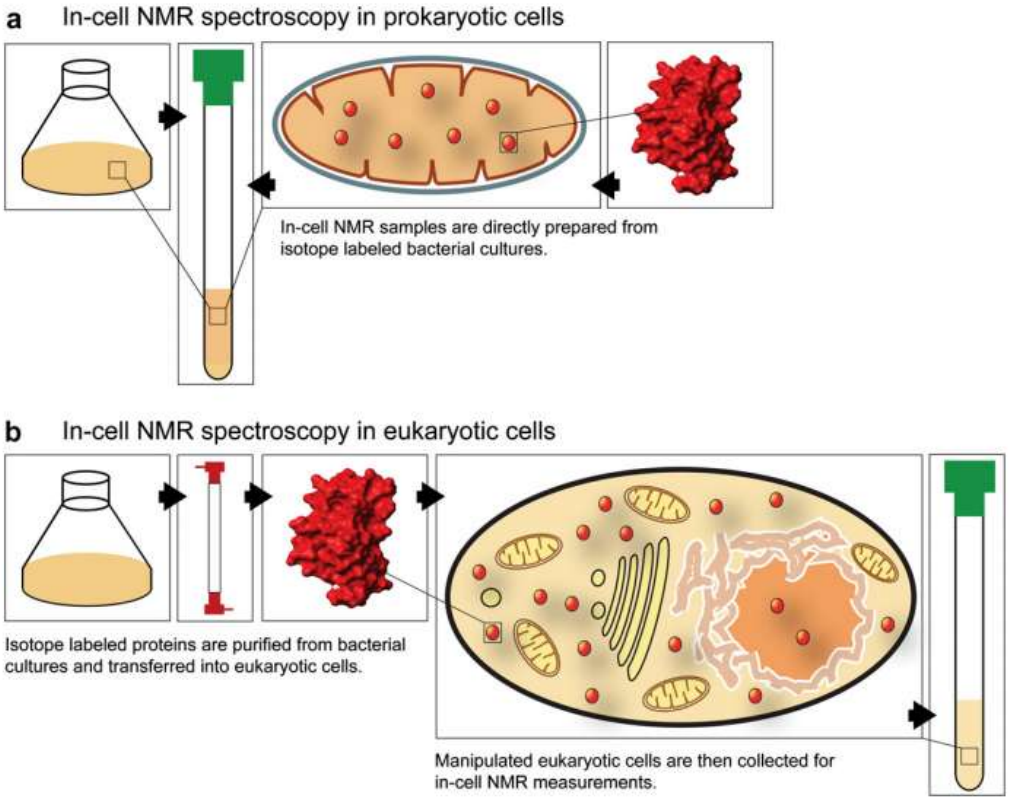


Figure 1: The overall different approach for in-cell sample preparation between prokaryotic (a) and eukaryotic cells (b). Reprinted from: Selenko, Wagner Looking into live cells with in-cell NMR spectroscopy Journal of Structural Biology 158 (2007) 244-253

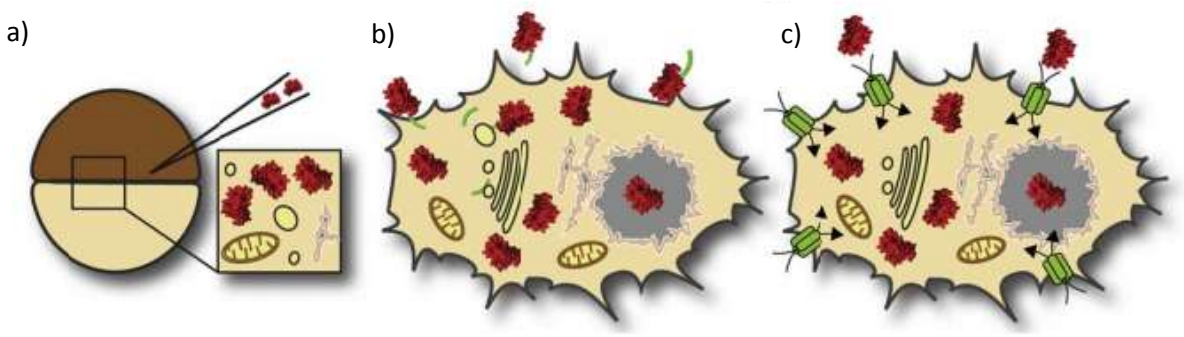


Figure 2: Three methods used up to now to insert proteins into mammalian cells. a) microinjection, b) CPP, c) porins. Reprinted from: Ito and Selenko, Cellular structural biology Current opinion in Structural Biology

The investigation and characterization of proteins in living cells by NMR spectroscopy has to overcame three main difficulties. First, the NMR signals of the molecule of interest must be distinguished from the NMR resonances of all other cellular components. This could be achieved by overexpressing the macromolecule and labelling with the NMR-active isotopes ^{15}N and ^{13}C; ^{19}F labelling has also been employed in some cases. ^{15}N-labeling creates only a small number of background peaks, whereas full ^{13}C-labeling produces strong background signals that make the unambiguous identification of resonance of the macromolecules of interest difficult. The significantly higher background levels in the case of ^{13}C-labeling as compared to ^{15}N-labeling is due to the higher number of C-H groups in the molecules than of N-H groups. Amino acid type selective labelling of proteins can also apply for in-cell NMR experiments. Unfortunately, not all amino acids can be used for selective labelling procedures; good candidates are those at the end of a biosynthetic pathway and do not serve as precursor for other amino acids. Alternatively, in *E. Coli*, specialized auxotrophic strains can be used that have been created for all 20 amino acids. Growth media for NMR labelling in *E. Coli* are simple in their composition, easily prepared and, depending on the type of labelling, relatively cheap. Bacteria will also incorporate isotopes with high efficiency (~98%). Labelling media for eukaryotic cells are sophisticated and expensive. Due to the complexity of most eukaryotic metabolisms, isotope incorporation is typically less than 90%. Induction time for recombinant protein expression are on the order of days rather than hours, which is likely to increase the amount of background labelling artefacts.

Second, the macromolecule must be able to tumble with a sufficient short correlation time. Long rotational correlation times lead to fast relaxation and, therefore, broad peaks. Since the rotational correlation time is proportional to the surrounding viscosity, the intracellular viscosity is an important parameter for the observation of macromolecules inside living cells. The intracellular observation of proteins can, however, become impossible by binding of the proteins to other cellular components; this significantly increase the rotational time of the protein. In particular, binding to large components, such as chaperones and nucleic acids, leads to the disappearance of a protein's resonances due to extensive line broadening. If even larger protein complexes or intracellular aggregates are to be investigated by in-cell NMR experiments, solid-state NMR techniques have to be employed.

Third, the cells have to survive inside the NMR tube at least for the time period of the experiment. The high cellular density can cause problems through oxygen starvation and limiting the amount of available nutrients. If the sensitivity of the selected system (mainly the expression level) is high enough, the NMR spectra can be measured relatively quickly (less than one hour). Modified NMR tubes that allow for continuous exchange of the media can keep cells alive for long times; or bacteria can be kept alive for long periods, encapsulated in low-melting agarose. This solution will not only allow the perfusion of the bacterial sample, it will also solve the problem of cell sedimentation to the bottom of the NMR tube over time. Both problems are more severe with larger and more sensitive eukaryotic cells. For them, special bioreactors were designed, but they are based on 10 mm NMR tubes, and the larger diameter creates additional problems for high resolution NMR spectroscopy, in particular for cryogenic probes[22].

Moreover, the biggest disadvantage of NMR spectroscopy is its inherent low sensitivity. In particular, for the investigation of the behaviour of biological macromolecules in their natural environment, it is important to keep the concentration of the macromolecule of interest as close to their natural level as possible.

1.3 SOD1

Cu,Zn superoxide dismutase (SOD1 here after) is a soluble, 32 kDa homodimeric enzyme, catalyzing the conversion of superoxide radicals to molecular oxygen and hydrogen peroxide, thus providing a defense against oxygen toxicity. SOD1 is one of the three human superoxide dismutases present in mammals: copper-zinc superoxide dismutase (Cu,ZnSOD1), manganese superoxide dismutase (MnSOD or SOD2), and extracellular superoxide dismutase (Cu,ZnSOD3). It was isolated for the first time in 1969 by McCord and Fridovich[23], and it was clear that SOD1 acts as a scavenger of superoxide, through a two-step reaction involving reduction and re-oxidation of the copper ion in its active site:

$Cu^{2+} + O_2^- <-> Cu^+ + O_2$

$Cu^+ + 2H^+ + O_2^- <-> Cu^{2+} + H_2O_2$

Primarily, this reaction occurs in the cytoplasm where SOD1 is highly expressed. Indeed, SOD1 is localized predominantly in the cytosol, but a portion of it is also found in the intermembrane space of mitochondria and in other cellular compartments, including the nucleus and the endoplasmic reticulum[24].

Structurally, human Cu,Zn SOD1 is a homodimer, with subunits of 154 residues. Each subunit is folded in a characteristic eight stranded β-barrel with a Greek key fold. 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. The two subunits of SOD1 are held together by hydrophobic interactions. The contact surface between the subunits 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.

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), it is completely buried in the protein and has a structural role[25]. The Cu ion is coordinated by four histidines (His46, His48, His63, His120). The two metal

ions are bridged by His63. The Cu role in the enzyme is mainly catalytic. The redox properties of copper ion permit the disproportionation of superoxide anion to hydrogen peroxide and dioxygen.

Figure 3: The solution structure of reduced dimeric Cu,Zn SOD1. Reprinted from: Baci et al., The solution structure of reduced dimeric copper zinc superoxide dismutase. The structural effects of dimerization. Eur J Biochem 2002, 269:1905-1915

The insertion of copper into Cu,Zn SOD1 requires a copper metallochaperone called CCS (copper chaperone for superoxide dismutase)[26-28]. CCS folds into three functionally distinct protein domains[29-30]. The N-terminal domain I of CCS has homology to ATX1, another copper chaperone. Surprisingly, a CCS molecule lacking this domain can still insert copper into superoxide dismutase in vivo[31]. It has been therefore proposed that this ATX1-like domain is only needed to maximize CCS function under extreme copper-limiting conditions. The central domain of CCS (domain II) is homologous to Cu,Zn SOD1. This domain physically interacts with SOD1[32] 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 SOD1 in vivo. This domain is highly conserved among CCS molecules from diverse species and includes an invariant CXC motif that can bind copper. Domain III is disordered; 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 SOD1[33]. On the contrary, domain II of CCS is proposed to participate in target recognition rather than in direct transfer of copper.

1.4 ALS-linked SOD1 mutants

Amyotrophic lateral sclerosis (ALS) is a devastating, fatal neurodegenerative disease that specifically targets the motor neurons in the spinal cord and in the brain[34]. It typically has an adult onset, starting first with weakness in arms or legs and proceeding relentlessly to total paralysis and death. Unfortunately, there is no known cure or effective treatment for ALS at present nor the causes of the disease have been identified. 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 165 different mutations in the sod1 gene have been linked to fALS. 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 (ALSod, the Amyotrophic Lateral Sclerosis Online Genetic Database, 2012, http://alsod.iop.kcl.ac.uk/). Mutations are scattered throughout the sequence of the protein. The vast majority of the mutations are genetically dominant.

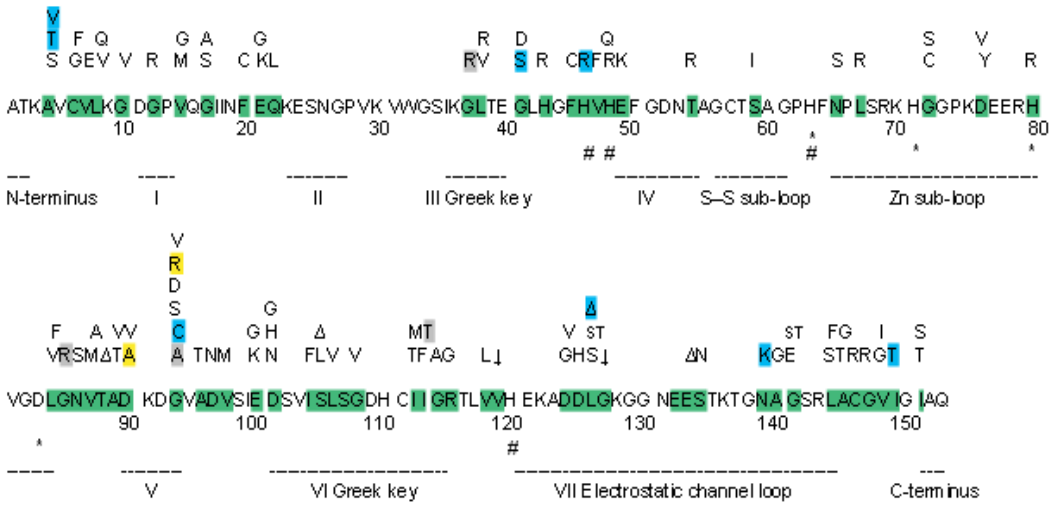


Figure 4: Mutations of SOD1 observed in patients with fALS. Copper-binding residues are indicated by #, zinc-binding residues are indicated by *. ↓ indicates an insertion in the sequence; Δ indicates a deletion; ST indicates a stop codon. Highlights: green, positions mutated in patients with FALS; blue, mutations introduced in cultured cells; yellow: mutations in transgenic rodents; grey: mutations introduced both in cell culture models and transgenic rodents. Reprinted from: Carrì et al., Lessons from models of SOD1.linked familial ALS, Trends in Molecular Medicine

The toxicity of ALS-SOD1 mutants has been linked to the formation of protein aggregates rich in SOD1, observed in tissues from ALS patients, ALS-SOD1 transgenic mice, and in cell culture model systems[35]. The precursors of these protein aggregates are thought to be soluble oligomeric intermediates in the SOD1 aggregation process. These oligomers are believed to be responsible for the toxic gain of function, similar to what occurs in other neurodegenerative diseases[36-39]. The mature form of SOD1 is not prone to aggregation, whereas the intermediate species of the maturation process (i.e. SOD1 lacking metal ions) have been shown to have tendency to oligomerize *in vitro*[40-42]. The oxidation of the two free cysteines of SOD1 (C6 and C111) is responsible of the formation of soluble oligomers[43]. Other mechanisms are possible, in which amyloid-like structures form[44-45] and a number of structurally various aggregation products are possible[46-47].

Therefore, an impaired maturation process would lead to accumulation of immature SOD1 species, which are prone to oligomerization. However, most of the studies addressing the properties of fALS SOD1 mutants and their effects on the maturation process have been performed *in vitro*, so far from physiological conditions.

1.5 Mia40

Mia40 is an oxidoreductase, which exerts its biological role in the intermembrane space of mitochondria (IMS), where it catalyzes oxidative protein folding. All proteins of the IMS are synthesized in the cell cytosol and are then imported into the mitochondria. Most of them do not contain any cleavable sequence and are targeted to the mitochondrial IMS through unidentified import signals. The IMS import pathway differs from the inner membrane or matrix ones, because it does not require ATP or an inner membrane electric potential[48]. Proteins, that should reach IMS, were thought to need to be in an unfolded state in the cytoplasm so to be able to enter mitochondria, thanks to their conformational flexibility that allows them to go through the translocase of the outer membrane (TOM) channel[49]. Once they approach mitochondria and face the IMS, Mia40 interacts with them importing them in the IMS and catalyzes the formation of their internal disulfide bonds, through the intermediate formation of a mixed disulfide bond between the substrate and its catalytic CPC motif[50-53]. Upon interaction with Mia40 and subsequent formation of their disulfide bonds, the substrates become folded and are trapped in the IMS[54-55]. Interestingly,

Mia40 itself obtains its final structure in the IMS upon the formation of two internal disulfide bonds, likely by acting as a substrate of itself[56-57]. Mia40 is then re-oxidized by ALR (augmenter of liver regeneration) protein and Mia40 is capable of forming disulfide intermediates with newly imported substrates proteins[58].

Substrate proteins for Mia40 are IMS proteins, usually less than 20 kDa, containing characteristic cysteine motifs, organized in twin CX_3C, twin CX_9C or CX_2C motifs. Among them there is the mitochondrial copper chaperone Cox17 (a CX_9C substrate), which participates in Cu(I) transfer to cytochrome c oxidase (CcO), and the small Tims (CX_3C substrates), which are chaperones for mitochondrial membrane proteins[59]. Another substrate of Mia40 is ALR protein, which does not contain twin "CX_3C" or twin "CX_9C" motifs. Thus, ALR represents an unusual type of substrate of the Mia40-dependent import pathway[60].

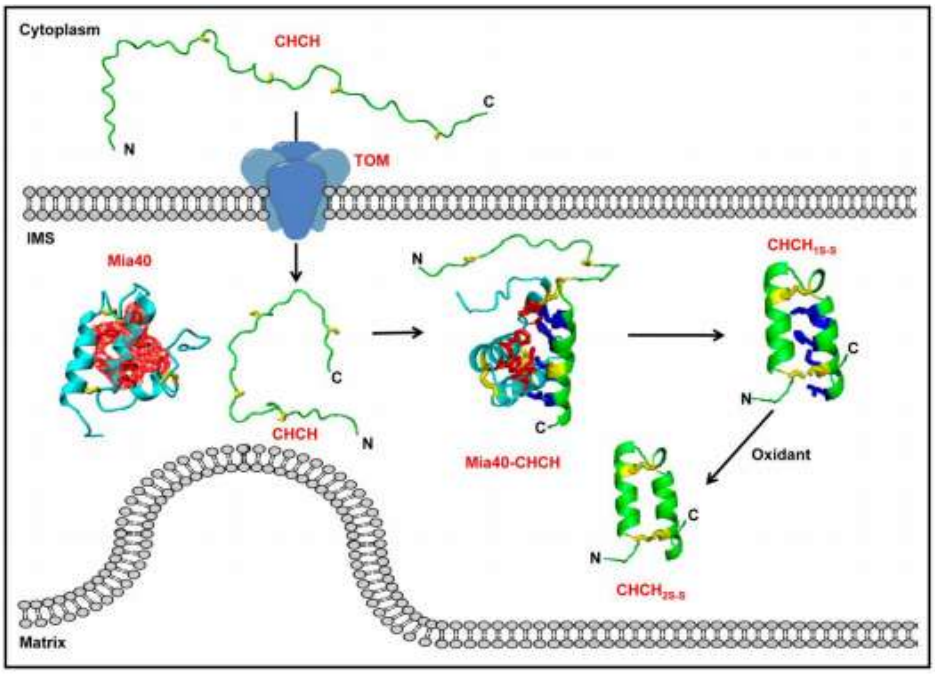


Figure 5: Oxidative protein folding mechanism. Reprinted from: Banci et al. PNAS

Thus, Mia40 cycles between two redox states: the oxidized one ($Mia40_{3S-S}$) containing three intramolecular disulfide bonds, and the partially reduced form where the disulphide bond in the CPC motif is not formed ($Mia40_{2S-S}$).

Structurally, both $Mia40_{3S-S}$ and $Mia40_{2S-S}$ states have a small helical segment in residues 56–59 (helix α1) and two longer helical segments (helix α2, residues 65–77, and helix α3, residues 88–100), whereas the other 80% of residues, essentially located at the N- and C-termini, do not take any secondary structural conformation, a large part of them being highly flexible.

The solution structure of the folded central region of $Mia40_{2S-S}$ consists of a 'core' and a 'lid' on top of it. The core is composed of helices α2 (residues 65–77) and α3 (residues 88–100), which form an antiparallel α-hairpin kept together by two disulfide pairs, Cys64-Cys97 and Cys74-Cys87, juxtaposing the CX9Cmotifs. The lid (residues 41–64) folds onto the core and is structurally rigid, although it does not have defined secondary-structural elements, with the exception of the short helix α1 (residues 56–59). The two residues preceding the lid segment, Pro54 and Cys55, also show some α-helical propensity. However, the disappearance of the NH signal of Cys55 in $Mia40_{2S-S}$ indicates local structural flexibility in the CPC region. In contrast, this NH signal is still detectable in $Mia40_{3S-S}$, suggesting an increased structural rigidity upon disulfide-bond formation. The difference in backbone flexibility between the two redox states of Mia40 may have a role in the catalytic process. Notably, the lid contains conserved hydrophobic residues that interact with conserved aromatic residues located on one side of the α-hairpin core. Indeed, a highly charged region is only present on the α-hairpin face opposite the CPC motif. In summary, the salient structural features of Mia40 are: (i) a high proportion of unfolded segments at the N- and C-termini,(ii) a folded α-hairpin core stabilized by structural disulfide pairings, (iii) a rigid N-terminal lid with an extensive array of hydrophobic interactions with one part of the α-hairpin core and (iv) a solvent-exposed CPC motif ideally placed to be the active site in the oxidation process.

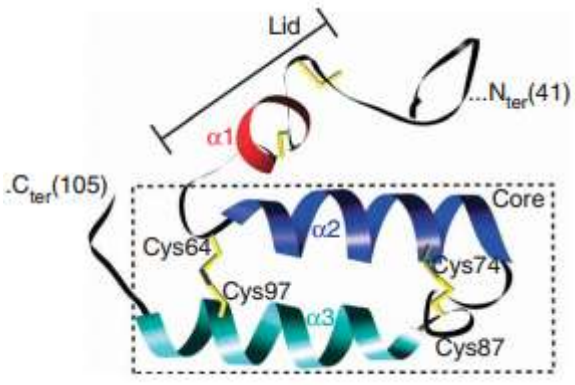


Figure 6: The solution structure of Mia40$_{2S\text{-}S}$. Reprinted from: Banci et al., Mia40 ia an oxidoreductase that catalyzes oxidative protein folding in mitochondria, Nature structural & molecular biology

1.6 Profilin1

Profilin1 is a cytoskeletal protein that is found widely throughout eukaryotes and is essential in all organisms. This small (15 kDa) globular protein binds three types of ligands: actin, phosphoinositides and proteins that contain poly-L-proline[61].

It was among the first actin-binding proteins to be characterized; initial studies suggested that its main role is to sequester actin monomers in a 1:1 complex and then release actin following cell stimulation and an increase in phosphatidylinositol (4,5)-bisphosphate [PtdIns(4,5)P_2] concentration. Profilin1 promote actin polymerization and it is essential in the regulation of cytoskeletal dynamics[62].

The poly-L-proline binding site comprises the N-terminal and C-terminal helices of profilin1, which form the binding clef for poly-L-proline. It is distinct from the actin-binding domain, although both seem to overlap with the phosphoinositides binding region[63].

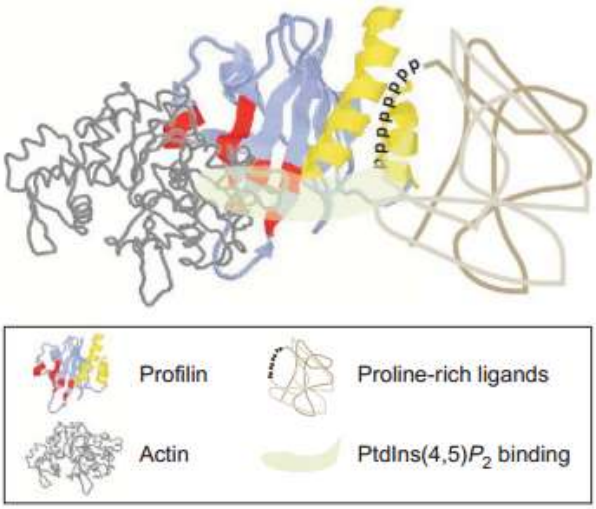


Figure 7: The binding sites for actin (red), poly-L-proline (yellow) and phosphoinositides (light grey). Reprinted from: Witke, The role of profilin complexes in cell mobility and other cellular processes, TRENDS in Cell Biology

Structurally, profilin1 folds into a single compact globular domain. Tracing the profilin1 chain from N- to C-termini, one finds an α-helix that extends from residues W3 through A12 (the two N-terminal residues are disordered). The chain reverses itself in a loop from residues Dl3 to C16 followed by the first two strands of β-sheet: a β-hairpin comprising residues Q17-Y24 and S29-V34. These two strands are connected by an irregular loop. The chain continues through three short stretches of helix: a single turn of $3_{,10}$-helix spanning residues T38-N41; an α-helix spanning residues P44-V51; and an α-helix spanning residues F58-N61. The chain then forms a contiguous stretch of five strands of antiparallel β-sheet: residues L63-L65, Q68-S76, S84-S91, T97-TI05, and T108-KI15. Finally, the chain proceeds to its C-terminus via a long a-helix composed of residues L122-SI37[64].

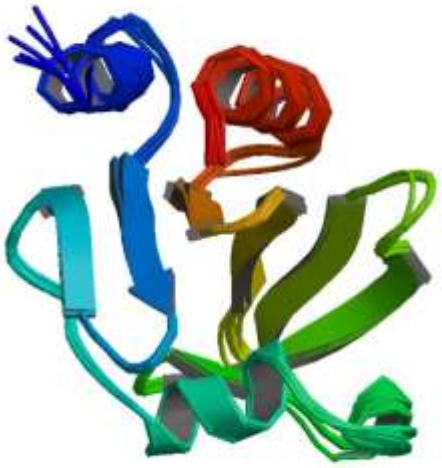


Figure 8: The structure of Profilin1. Reprinted from: PDB

Mutations in the gene of profilin1 have very recently been shown to be the cause of a subgroup of familial amyotrophic lateral sclerosis (ALS). From exome sequencing and a large screen of ALS patients, four mutations have been identified. They are C71G, M114T, E117G and G118V and lie in proximity to actin-binding site. A synthetic mutant H120E profilin1 was utilised as a control, since the mutation is located at a crucial residue for the actin binding. From studies in cultured cells, researchers have found that, while the wild-type profilin1 is present predominantly in the soluble fraction of cells, a considerable proportion of the mutants is detected in the insoluble fraction. Furthermore, several higher molecular mass species are observed, indicative of cytoplasmatic aggregates[65]. Subsequently, another mutant, T109M , has been identified. T109 is positioned close to the actin binding domain and contributes to a cluster together with the other PFN1 ALS mutation. This amino acid is highly conserved among vertebrates and is subject to phosphorylation. Therefore impaired phosphorylation of profilin1 might contribute to motoneuron degeneration[66].

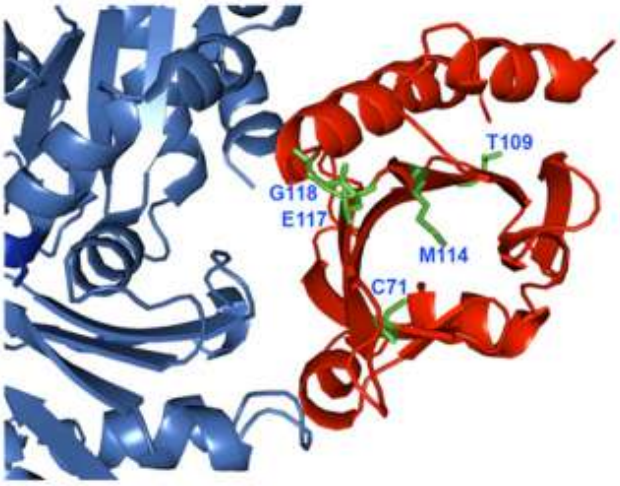


Figure 9: Three dimensional illustration of T109M mutation and the 4 amyotrophic lateral sclerosis mutations. Reprinted from Wu et al. (2012). Blue: actin; red: profilin 1.

2 Results

A new approach to apply in-cell NMR to human proteins in human cells, with respect to those already used by few other researchers, has been developed: proteins has been directly over-expressed in human cells to perform in-cell NMR experiments. Protein expression was obtained by DNA transient transfection mediated by polyethylenimine (PEI), a cationic polymer that binds DNA and enters inside cells via endocytosis and releases it inside the nucleus. To validate our method, proteins with different features were selected, and the high expression level was estimated by SDS-PAGE. These proteins were visible in the ^{1}H NMR spectra above the cellular background, and SOD1, Mia40 and Atox1 could be detected on ^{1}H-^{15}N SOFAST-HMQC spectra on uniformly ^{15}N labelled cell samples. Glutaredoxin1, Thioredoxin1 and Profilin1 were not detected and only became visible upon cell lysis, suggesting that some interactions occur in the cytoplasm which makes the proteins tumbling slower on average, thus broadening the amide crosspeaks beyond detection.

However, before this approach becoming generally applicable, a number of technological limitations has been to overcome: 1) the short lifetime of human cells in the NMR tube requires optimization of experimental conditions to maintain cellular viability and reduce data acquisition time, while increasing measurement sensitivity; 2) expression of isotopically labelled proteins must occur at high enough levels for their detectability in NMR experiments at the currently attainable limits of spectral sensitivity; 3) when biochemical processes of interest require the intervention of other proteins, these need to be co-expressed. Procedures to address and either overcome or reduce the above limitations have been generated and this methodological approach was applied to address few functional processes, including protein folding, metal uptake and disulfide bond formation.

One of the characterized systems is SOD1, an enzyme that is involved in the cellular defence against oxidative stress, catalyzing the dismutation of the superoxide anion into molecular oxygen and hydrogen peroxide. The post-translational modification steps, i.e. zinc binding, homodimer formation, copper uptake and formation of the intrasubunit disulfide bond between cysteine 57 and 146, which occur in SOD1 after its synthesis, have been studied in human HEK293T cells and in *E. Coli* cells.

The hSOD1 gene was transfected in HEK293T cell, and SOD1 was detected by in-cell NMR at closer-to-physiological intracellular concentrations (45 ±10 μM).

In cells grown without supplements of zinc or copper ions and at the highest expression level, two forms of SOD1 were detected in similar amounts: the monomeric, metal-free species (apo-SOD1) and the dimeric species with one Zn^{2+} ion bound to each subunit (E,Zn-SOD1). The observed mixture is due to residual zinc in the culture medium. Cys57 and Cys146 were reduced in both species. The same behaviour in the redox state was found also in *E. Coli* cells: the protein had all its cysteines in the reduced state.

Addition of Zn^{2+} to the culture medium eliminated apo-SOD1 species signals, either in human cells and *E. Coli* cells. Therefore, Zn^{2+} ions are efficiently taken up by cells in culture and bind specifically to the native binding site of dimeric SOD1 in stoichiometric amounts. When reduced apo-SOD1 is exposed to Zn^{2+} ions *in vitro* or in cell lysates, even at sub-stoichiometric concentrations, a mixture of forms is generated that includes apo-SOD1, E,Zn-SOD1 and Zn,ZnSOD1 (in which a second Zn2+ is bound to the copper binding site). Therefore, site-selectivity of Zn^{2+} binding can be achieved only in the physiological cellular context.

Whereas the mechanisms for copper uptake are simple in prokaryotes, they are tightly regulated in eukaryotic cells. In *E.Coli*, 100 μM $CuSO_4$, in addition to $ZnSO_4$, was added in the expression medium 15 minutes before the cells collecting. The Cu(I),Zn-SOD1 was detected by in-cell NMR and ~50% of the SOD1 intrasubunit disulfide bond is formed.

Copper entrance in eukaryotic cells and its delivery to copper proteins require a number of steps, involving specific chaperones responsible of its intracellular trafficking. Copper incorporation of SOD1 in eukaryotes is known to be dependent on the CCS protein.

When cells were co-transfected with hSOD1 and hCCS genes in zinc supplemented medium, SOD1 is in the dimeric zinc-containing species, and the SOD1 intrasubunit disulfide bond is partially oxidized (around 50%). Therefore there is a mechanism of CCS-mediated SOD1 disulfide oxidation which does not rely on the presence of copper.

When cells were incubated with Cu(II), a higher ratio of Cu(I),Zn-SOD1 vs E,Zn-SOD1 is obtained (~1:1) compared with cell samples with basal CCS level, indicating that CCS has also a role in promoting copper incorporation in SOD1. Therefore, copper added as Cu(II)

salt to the culture medium was reduced to Cu(I) and it was stoichiometrically bound to expressed SOD1, forming Cu(I),Zn-SOD1. When the test was repeated at concentration near the physiological ones, the complete formation of oxidized Cu(I),Zn-SOD1 was observed.

The redox state of SOD1 cysteines is affected by both overexpression of CCS and presence of copper. Indeed, when cells overexpressing both SOD1 and CCS are incubated with Cu(II), the intrasubunit disulfide bridge of SOD1 is completely oxidized.

The same maturation steps have been studied also in ALS-linked SOD1 mutants, to investigate if they follow the complete sequence of the post-translational modification steps or if the mutated proteins acquire a misfold. In absence of added metal ions, all mutants were in the apo, partially unfolded state, similar to what was observed for WT SOD1. Zinc binding step was investigated: even with excess Zn^{2+}, several mutants (A4V, I35T, and G85R) did not bind the metal ion, and remained in the apo state. For G37R, G93A and I113T SOD1 mutants, only a small intracellular amount of zinc-containing protein (E,Zn-SOD1) was detected (~15% of the total protein), while the remaining fraction was in the apo state. Conversely, V7E, T54R and V148I SOD1 mutants behaved like WT SOD1: they stoichiometrically bound one Zn^{2+} ion per monomer, and only the folded, homodimeric E,Zn-SOD1 forms were present.

The effect of hCCS in the maturation of SOD1 mutants was assessed by inducing simultaneous expression of each SOD1 mutant and hCCS in media supplemented with zinc and copper, and by analyzing the protein metallation and cysteine redox state. With the exception of G85R mutant, the fully mature form (Cu(I),Zn-$SOD1^{SS}$) was detected for each mutant, indicating that the SOD1 mutants had reached the mature state when both hCCS and copper were available in the cytoplasm in sufficient amount; a fraction of E,Zn-$SOD1^{SS}$ was also present for some mutants (namely V7E, G37R, T54R, G93A, I113T, V148I) as well as for WT SOD1. Notably, for the mutants with impaired intracellular zinc binding, the total amount of zinc-containing protein (i.e. the sum of E,Zn-SOD1 and Cu(I),Zn-SOD1 species) was higher in these cellular conditions than that measured in zinc-supplemented cells with only endogenous levels of CCS. The overexpressed CCS somehow promoted zinc binding to apo-SOD1 mutants, and then subsequently transferred

copper and catalyzed cysteine oxidation, resulting in a higher intracellular amount of mature SOD1. To determine whether copper-free CCS alone could promote zinc binding to apo-SOD1 mutants, we analyzed cell samples overexpressing [^{15}N]Cys-labeled G37R, G93A and I113T SOD1 together with CCS, without supplementing extra copper. In these conditions a mixture of species was generated, which included apo-SOD1^{SH}, E,Zn-SOD1^{SH}, and E,Zn-SOD1^{SS}, showing that Cu-free CCS was actually less efficient than Cu-CCS in increasing the amount of zinc-containing forms for the mutants, and in catalyzing the disulfide bond formation. The ability of Cu-CCS to rescue the zinc binding step in the maturation of a set of fALS mutants suggests that CCS may in fact act as a molecular chaperone for these mutants beyond copper and disulfide transfer.

Another investigated system has been Mia40, that is an oxidoreductase that catalyzes oxidative protein folding in the mitochondrial intermembrane space (IMS). The folding and redox state, prior to mitochondrial import in live human cells, have been observed. Mia40 in the cytoplasm is present in a folded and in a disulfide-oxidized state, while the CPC motif is still reduced.

The dependence of Mia40 redox and folding state on the overexpression of Glutaredoxin1 (Grx1) and Thioredoxin1 (Trx1), two proteins that are involved in the regulation of protein thiol groups in the cellular cytoplasm, was investigated.

In presence of Grx1, the folded conformation of Mia40 was poorly detected, while the crosspeaks of the unfolded parts were still visible in the uniform ^{15}N labelled spectrum. The amount of folded Mia40 was measured in the cell extract by standard addition of pure folded Mia40, using the methyl peak at -0.7 ppm as a marker of the folded state of Mia40. Total Mia40 was determined on the same cell extract by Western-Blot, using the same pure Mia40 for reference. A ratio of 23±3% folded Mia40 over total Mia40 was obtained.

The effect of Trx1 on the Mia40 state was also investigated by inducing co-expression of Trx1 with Mia40. Trx1 affected the folding state of Mia40 to a lesser extent than Grx1 when expressed at similar levels because 50% of cytoplasmic Mia40 was still in the folded, oxidized state. Therefore, the two thiol-regulating proteins have different efficacy in keeping Mia40 reduced, despite being reported to have overall similar functions in the cytoplasm.

2.1

In-cell NMR in *E. Coli* to Monitor Maturation Steps of hSOD1

Banci, L; Barbieri, L; Bertini, I; Cantini, F; Luchinat, E

PLoS ONE 2011 6(8): e23561. doi:10.1371 / journal. pone. 0023561

OPEN ACCESS Freely available online

PLoS one

In-cell NMR in *E. coli* to Monitor Maturation Steps of hSOD1

Lucia Banci[1,2]*, Letizia Barbieri[1,2], Ivano Bertini[1,2]*, Francesca Cantini[1,2], Enrico Luchinat[1,2]

1 Magnetic Resonance Center, University of Florence, Sesto Fiorentino, Italy, **2** Department of Chemistry, University of Florence, Sesto Fiorentino, Italy

Abstract

In-cell NMR allows characterizing the folding state of a protein as well as posttranslational events at molecular level, in the cellular context. Here, the initial maturation steps of human copper, zinc superoxide dismutase 1 are characterized in the *E. coli* cytoplasm by in-cell NMR: from the apo protein, which is partially unfolded, to the zinc binding which causes its final quaternary structure. The protein selectively binds only one zinc ion, whereas *in vitro* also the copper site binds a non-physiological zinc ion. However, no intramolecular disulfide bridge formation occurs, nor copper uptake, suggesting the need of a specific chaperone for those purposes.

Citation: Banci L, Barbieri L, Bertini I, Cantini F, Luchinat E (2011) In-cell NMR in *E. coli* to Monitor Maturation Steps of hSOD1. PLoS ONE 6(8): e23561. doi:10.1371/journal.pone.0023561

Editor: Andreas Hofmann, Griffith University, Australia

Received May 6, 2011; **Accepted** July 20, 2011; **Published** August 24, 2011

Funding: This work was supported by the Italian FIRB PROTEOMICA MIUR contract RBRN07BMCT, and is part of the IRA2 of the Bio-Nuclear Magnetic Resonance project (European Commission's FP7, project number 261863). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: bertini@cerm.unifi.it (IB); banci@cerm.unifi.it (LB)

Introduction

Folding and maturation of proteins characterized by post-translational modifications and formation of quaternary structure is a complex process which progresses through a number of well concerted events. A deep understanding of such processes requires their characterization at molecular level in a cellular context.

In-cell NMR has the unique ability to acquire structural and conformational information of biomolecules in their native cellular environment at atomic level [1,2]. It has been previously shown that the bacterial cytoplasm is a good model of the eukaryotic one, especially to study the effects of molecular crowding on protein folding and non-specific interactions [3], as they have similar pH and redox potential [4–6].

Within this frame, we have characterized by in-cell NMR the wild-type human copper, zinc superoxide dismutase 1 (hSOD1) protein, as well as the initial steps towards its maturation.

hSOD1 is a 32 kDa homodimeric protein involved in the cellular defence against oxidative stress. It is physiologically expressed at relatively high concentrations in human cells, and it exerts its function in the cytoplasm, in the nucleus and in the mitochondrial IMS [7]. In order to reach its mature form, hSOD1 has to incorporate one Zn^{2+} ion and one catalytic Cu^{+} ion per subunit. Additionally, two conserved cysteine residues (Cys 57 and 146) form an intramolecular disulfide bridge during the protein maturation process.

Apo-hSOD1 has been recently linked to the familial form of amyotrophic lateral sclerosis (fALS), a fatal motor neurodegenerative disease [8–10]. The immature form of hSOD1, i.e. without the metal ions and with a misfolded structure, is believed to play a pivotal role in ALS pathology [11,12]. Therefore folding and metal insertion are important factors to be investigated in the cellular environment.

In this work, we characterized the state of hSOD1, analyzing samples of *E. coli* cells overexpressing hSOD1 protein, in its initial state after expression, and its initial maturation steps through zinc uptake and disulfide bond formation and we determined how this affects the tertiary and quaternary structure of the protein. This study sheds some light on the folding state of the non-mature protein as well as on the process of zinc uptake and protein folding in the cellular environment.

Results

Folding state of apo-hSOD1 in the cytoplasm

$^{1}H,^{15}N$-SOFAST-HMQC spectra [13] were acquired on cell samples overexpressing hSOD1 in a metal-free medium. The spectra were then recorded again on the cleared lysates after cell lysis. The in-cell NMR spectrum shows mainly peaks in the 8.0–8.3 ppm (^{1}H) region. In addition, few more dispersed peaks are visible, at lower S/N ratio (Figure 1 A). When the cells are lysed, still maintaining the sample in anaerobic conditions, a few other dispersed peaks appear, indicating the presence of some structured regions of the protein, while most of the peaks are still in the "unfolded" region. The spectrum of the latter species compares well with that of the monomeric apo form with reduced cysteines, E,E-hSOD1$^{S\text{S}1\text{-}S\text{S}1}$(Figure 1 B) which also shows many peaks in the region of unstructured peptides, overall corresponding to the ones that remain visible in the in-cell spectrum of the protein. This therefore indicates that, once the newly produced protein is in the cytoplasm and in the absence of metal ions, it is in a reduced, metal-free state.

Several interpretations of the in-cell NMR spectrum of apo-hSOD1 are possible: apo-hSOD1 could be completely unfolded in the cytoplasm; in this case all the NH cross-peaks would be overlapped in the "unfolded" region of the spectrum. The protein

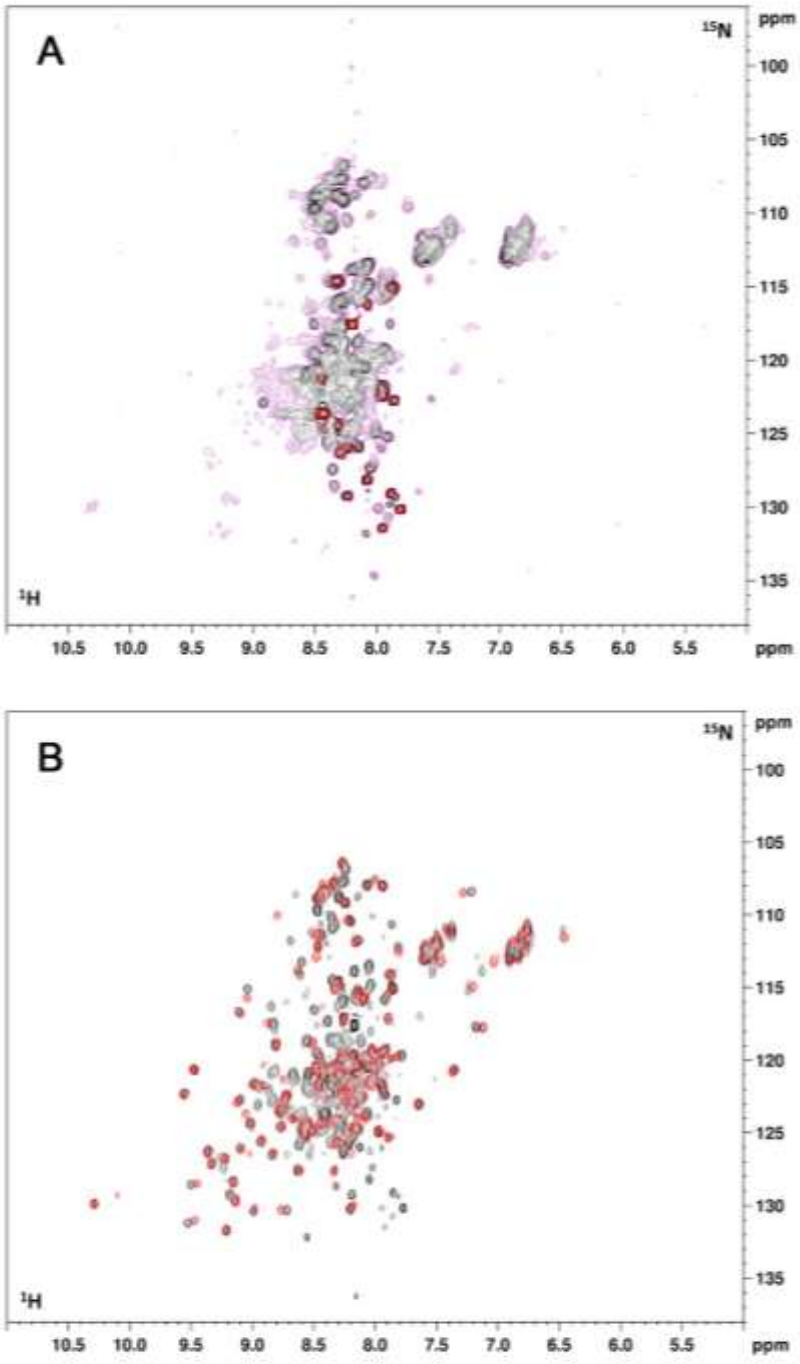


Figure 1. In-cell NMR spectra of apo-hSOD1 and of cell lysate. (A) In-cell ^{1}H-^{15}N SOFAST-HMQC spectrum of *E. coli* cells expressing hSOD1 in defect of Zn(II). Two different thresholds are shown (black, purple). At lower threshold (purple) some weak and broad signals are visible, while only signals in the unfolded region of the spectrum are visible at higher threshold (black). Signals belonging to other cellular components are overlaid in red. These are always present in all in-cell spectra. (B) Overlay of the ^{1}H-^{15}N SOFAST-HMQC spectrum of a cell lysate without addition of Zn(II) (black), and the ^{1}H-^{15}N HSQC of an *in vitro* sample of E,E-hSOD1$^{Se\text{-}Se}$ (red).
doi:10.1371/journal.pone.0023561.g001

could be only partially unfolded, as is the case of *in vitro* E,E-hSOD1$^{SH\text{-}SH}$ [14], and the cross-peaks of the folded region could be lost as a consequence of chemical or conformational exchange phenomena in the NMR timescale. Alternatively, the loss of the signals could be a consequence of interactions with other components of the cellular environment, like membranes, bacterial heat shock proteins or DNA. Finally, the in-cell NMR species could be an oligomer of apo-hSOD1.

To determine whether apo-hSOD1 is completely unfolded in the cytoplasm, the in-cell NMR spectrum was compared to the *in vitro* NMR spectrum of E,E-hSOD1$^{SH\text{-}SH}$ denatured with guanidinium chloride (Figure 2 A). The latter spectrum also shows only signals in the region typical of unfolded proteins, but is somewhat different from the in-cell NMR spectrum. The cross peak of the side-chain NH of Trp 32 for example, which in the folded protein is located in a β-strand of the hSOD1 β-barrel, falls at different chemical shifts in the two spectra (Figure 2 B) and in particular in the cytoplasmic species it has the same chemical shift as in the *in vitro* E,E-hSOD1$^{SH\text{-}SH}$. This indicates that in the cellular environment apo-hSOD1 is not completely unfolded, but the signals of the folded part are lost.

The loss of those signals could be caused by chemical or conformational exchanges phenomena involving the protein amides. In this case, the low S/N ratio of the signals should be increased at different magnetic field strength and temperature. SOFAST-HMQC spectra of cell samples expressing apo-hSOD1

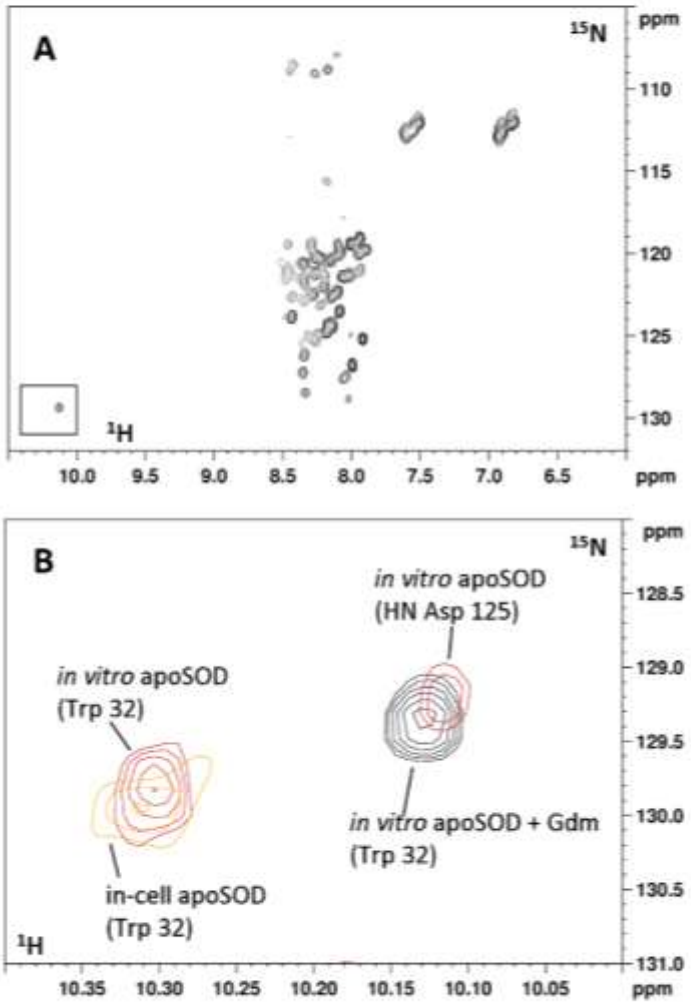


Figure 2. Trp 32 side chain of denatured *in vitro* apo-hSOD1. (A) ^{1}H-^{15}N HSQC of an *in vitro* sample of E,E-hSOD1$^{SH\text{-}SH}$ denatured with 0.5 M guanidinium chloride. (B) Zoom of the (A) spectrum showing the signal of Trp 32 (black). Overlaid are the in-cell NMR spectrum of apo-hSOD1 (orange) and the *in vitro* spectrum of non-denatured E,E-hSOD1$^{SH\text{-}SH}$.
doi:10.1371/journal.pone.0023561.g002

were acquired at 500 MHz at 315 K and at 800 MHz at 288 K and 298 K. In both cases no increase of the S/N ratio of the signals was observed (data not shown), disfavouring this hypothesis.

The possibility of apo-hSOD1 forming aggregates in the cells was also taken into account. Apo-hSOD1 is known to give rise, both *in vivo* and *in vitro*, to soluble oligomeric species, and eventually to fibrils [11,15]. It has been shown that the formation of soluble oligomers of apo-hSOD1 *in vitro* leads to the complete loss of the signals in the HSQC spectrum. This is most true for fibrils, which are insoluble species. Therefore, the presence of several signals in the "unfolded" region of the in-cell NMR spectrum, and the fact that the missing signals are completely recovered after cell lysis, excludes the possibility that oligomers of apo-hSOD1 have formed in the cytoplasm.

Apo-hSOD1 in the cytoplasm could be interacting to some extent with cellular components. Such interactions would lead to the formation of high molecular weight complexes, determining the broadening beyond detection of the signals. In this situation, only the amide signals of the unstructured regions are observable, as they might not interact with the cellular components and therefore move freely with respect to the slow-tumbling complex. Upon cell lysis these interactions are released, the protein becomes free and tumbles faster. The interactions have to be weak enough, in order to be disrupted upon sonication and subsequent dilution of the cell content in M9 buffer (~1:2 dilution).

The line broadening effect could be better explained considering the multiplicity of weak interactions in which apo-hSOD1 is involved. hSOD1 is overexpressed, and is more abundant than any other cellular species (~850 μM in the cytoplasm). Therefore all the different interactions involve some hSOD1 molecules at any given time. In fact we can think of many sub-populations of apo-hSOD1, each experiencing a different chemical environment, thus having different chemical shifts. This "cellular anisotropy" causes an inhomogeneous broadening of different amount for each NH cross-peak. The peaks of the folded part have larger chemical shift dispersion and are thus made invisible by the cellular anisotropy.

Zinc binding properties of cytoplasmic hSOD1

$^{1}H,^{15}N$-SOFAST-HMQC spectra were acquired on cell samples overexpressing hSOD1 in a minimal medium supplied with different amounts of $ZnSO_4$. The concentration of zinc in the medium ranged from 10 μM to 1 mM in the expression medium; this is always in excess with respect to the total amount of hSOD1 expressed. When cells are grown in the presence of extra zinc added in the expression medium, the in-cell NMR spectra of hSOD1 show remarkable differences with respect to those of in-cell hSOD1 expressed without added zinc (Figure 3 A). The appearance of several dispersed peaks, together with the disappearance (or decrease in intensity) of some signals in the 'unfolded' region, indicates that hSOD1 inside the bacterial cells binds zinc when this is added to the culture medium in excess relatively to the total amount of protein expressed. This in-cell NMR spectrum compares very well with that of E,Zn-hSOD1$^{SH\text{-}SH}$ (the species with one zinc ion bound to the zinc binding site), and not with that of Zn,Zn-hSOD1$^{SH\text{-}SH}$ (the non-physiological species with two zinc ions bound to both metal binding sites). Among the various signals, the NH signals from Gly 61 and Thr 135, which are close to the metal binding sites, have chemical shift values which are indicative of the metal binding state of the protein, i.e. of which metal site is occupied by which metal ion. In the in-cell NMR spectra these signals have chemical shifts very close to those observed in the E,Zn-hSOD1$^{SH\text{-}SH}$ *in vitro* spectrum (combined chemical shift difference $^{1}H\text{-}^{15}N$: $\Delta\delta G61 = 0.015$, $\Delta\delta T135 = 0.029$), while are more distant from the corresponding signals of the Zn,Zn-hSOD1$^{SH\text{-}SH}$ form ($\Delta\delta G61 = 0.169$, $\Delta\delta T135 = 0.125$) (Figure 3 B, Figure S1). This is a striking result as it indicates that hSOD1 in the cytoplasm has a higher selectivity than *in vitro* in the binding site mode. Indeed, when sub-stoichiometric amounts up to 1 equivalent of zinc are added to E,E-hSOD1 *in vitro* at physiological conditions (pH around 7), mixtures of E,E-hSOD1, E,Zn-hSOD1 and Zn,Zn-hSOD1 species are formed, while when 2 equivalents of zinc per subunit are added only Zn,Zn-hSOD1 is formed. Instead in the cytoplasm hSOD1 binds zinc only in its native binding site, giving only E,Zn-hSOD1 species, while zinc binding to the copper site does not occur, and Zn,Zn-hSOD1 species is not detected (within the sensitivity of the NMR experiment). Moreover, this effect is seen regardless of the concentration of zinc added to the medium, even at 1 mM.

After the in-cell NMR experiments, the cell samples were washed with metal-free medium in order to remove the external zinc, and NMR spectra of the cleared cell lysates were acquired. Only the species Zn,Zn-hSOD1$^{SH\text{-}SH}$ was detected. Apparently, an excess of zinc is still present inside the cells, which is made available upon cell lysis and binds at the hSOD1 copper binding site.

Cysteine redox state determination

hSOD1, when is either in the cytoplasm or in the cell lysate, without addition of zinc has all its cysteines in the reduced state, as monitored from $^{1}H\text{-}^{15}N$ signals of selectively labelled cysteines.

After cysteine oxidation by air exposure, the spectrum shows the four peaks of the cysteines, with that of Cys 57 only detectable below 298K. Cys 57 and 146 have chemical shifts equal to those previously assigned in the *in vitro* dimeric E,E-hSOD1$^{S\text{-}S}$ species [16] and indicative of their oxidized state. After reduction of the same sample with DTT, the peaks shift back to the chemical shift they have in the cells and in the cell lysate, therefore confirming the reduced state of Cys 57 and 146 in the cytoplasm (Figure 4 A–C). The same analysis, repeated for cells and lysates when hSOD1 is expressed in the presence of Zn(II), shows that also in the case of E,Zn-hSOD1, Cys 57 and 146 in the cytoplasm are reduced (Figure 4 D–F).

The redox state of the cysteines is further confirmed by reaction with AMS performed directly on the cell culture and on the cell lysate after the NMR experiments (Figure S2, Information S1).

Quaternary structure of hSOD1

The spectra of the ^{15}N-cysteine labelled protein provide evidences on the quaternary structure of hSOD1 in the cellular cytoplasm. Indeed disulfide bond formation and protein dimerization are tightly linked processes. *In vitro*, the species E,E-hSOD1$^{SH\text{-}SH}$ is monomeric, while both E,E-hSOD1$^{S\text{-}S}$ and E,Zn-hSOD1$^{SH\text{-}SH}$ are homodimers [14,17]. In the NMR spectra of both *in vitro* ^{15}N-cysteine labelled E,E-hSOD1 and E,Zn-hSOD1, the amide cross-peak of Cys 146 has a large chemical shift difference between the two redox states (Figure 4 B,D). This difference is expected, as Cys 146 is directly involved in the disulfide formation. On the other hand, the cross-peak of Cys 6 has a similar behaviour upon disulfide formation in E,E-hSOD1, but changes little in the case of E,Zn-hSOD1 (Figure 4 C,F). This is because Cys 6 is located close to the interaction surface of the homodimer. Upon E,E-hSOD1 dimerization as a consequence of disulfide bond formation, Cys 6 changes its chemical environment, whereas in E,Zn-hSOD1, which is always dimeric regardless of the oxidation state, it remains unchanged. Therefore the chemical shift of Cys 6 is a marker of the quaternary structure of hSOD1. The chemical shift of Cys 6 of both in-cell E,E-hSOD1$^{SH\text{-}SH}$ and E,

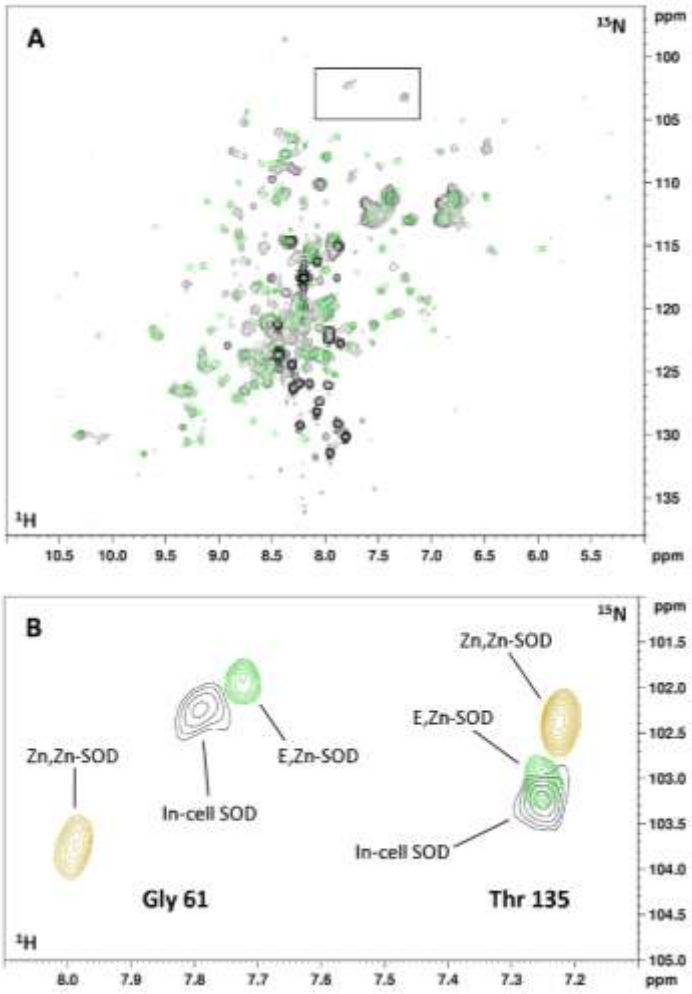


Figure 3. In-cell NMR spectra of hSOD1 with added Zn(II). (A) Overlay of the in-cell ^{1}H-^{15}N SOFAST-HMQC spectrum of *E. coli* cells expressing hSOD1 in presence of Zn(II) (black), and 1H-15N HSQC spectrum of an *in vitro* sample of E,Zn-hSOD1$^{S\text{-}S}$ (green). (B) Zoom of the (A) spectrum showing the peaks of Gly 61 and Thr 135 in-cell (black), *in vitro* E,Zn-hSOD1$^{S\text{-}S}$ (green) and *in vitro* Zn,Zn-hSOD1$^{S\text{-}S}$ (yellow).
doi:10.1371/journal.pone.0023561.g003

Zn-hSOD1$^{SH\text{-}SH}$ species matches that of the corresponding *in vitro* species, thus suggesting that in the cytoplasm E,E-hSOD1$^{SH\text{-}SH}$ is in the monomeric state, while E,Zn-hSOD1$^{SH\text{-}SH}$ is in the dimeric state, confirming that the *in vitro* findings hold true also inside the cell.

Discussion

The process leading a newly synthesized protein to acquire its final, functional state could involve several steps which consist of protein folding to its tertiary and possibly quaternary structure, cofactor binding and post-translational modifications. In most of the cases these steps have been characterized at molecular level only *in vitro*, where protein is isolated in an environment which might be far from the physiological one. In-cell NMR is a quite powerful method to characterize in detail those processes directly in living cells. Indeed, the atomic resolution of the technique, which can be applied to labelled proteins while inside living cells, allows us not only to analyze the folding and organization properties of a protein but also to look at the status of individual residues and the interaction with cofactors.

hSOD1 needs to undergo a number of events after its synthesis, i.e. needs to bind one zinc and one copper ion per molecule, to form a disulfide bond and to dimerize. In this work we have analyzed the protein state directly in the cytoplasm of *E. coli* cells, before it undergoes those maturation steps. Then, by adding zinc to the external medium, we have been able to monitor the changes in tertiary and quaternary structure following the binding of zinc.

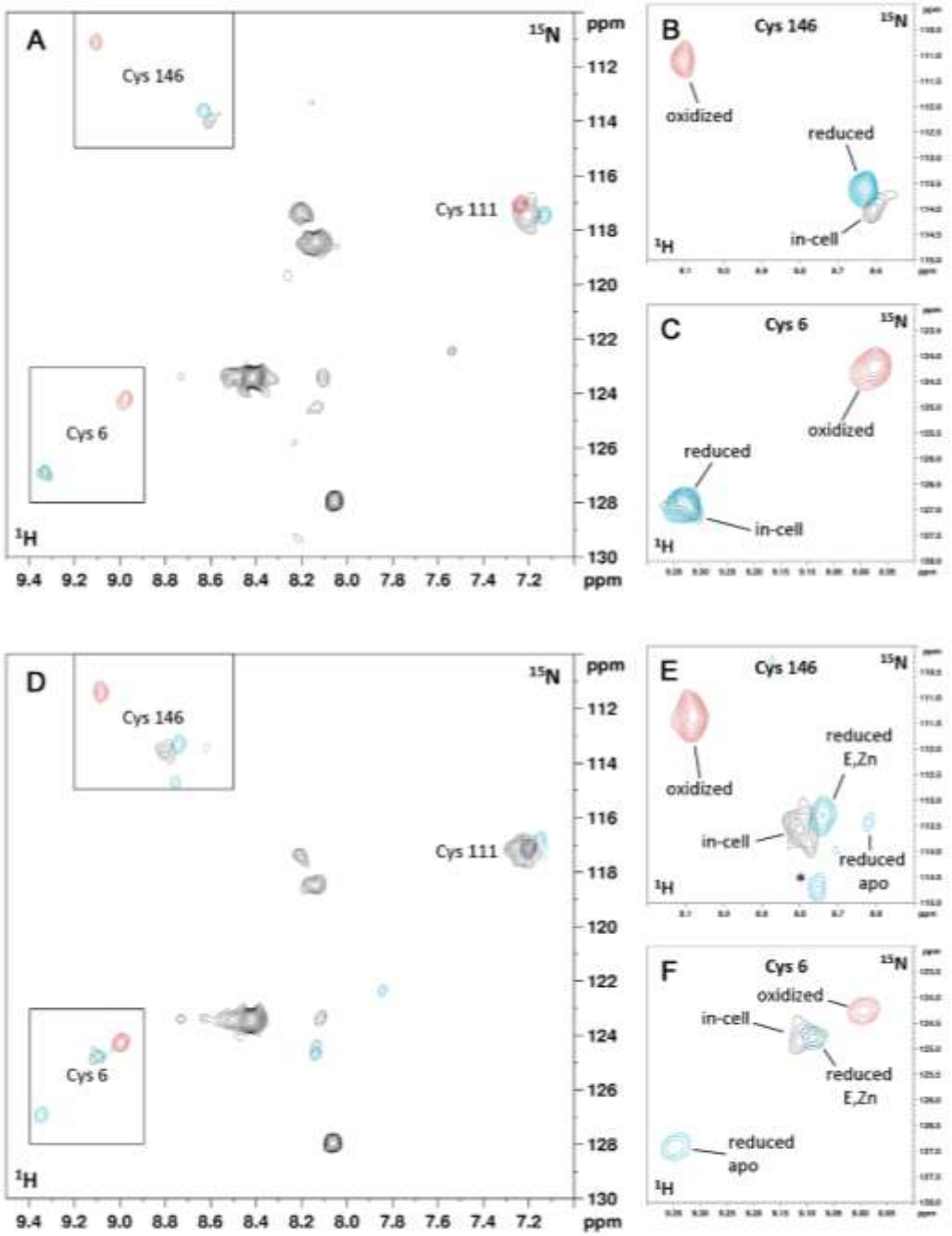


Figure 4. In-cell and *in vitro* NMR spectra of ^{15}N cysteine-labelled E,E-hSOD1 and E,Zn-hSOD1. (A) Overlay of ^{1}H-^{15}N SOFAST-HMQC in-cell spectrum (black) of ^{15}N cysteine-labelled E,E-hSOD1, oxidized E,E-hSOD1 purified from the lysate (red), same sample reduced with DTT (blue). (B,C) Detailed views of (A) showing Cys 146 and Cys 6 amide cross-peaks. (D) Overlay of ^{1}H-^{15}N SOFAST-HMQC in-cell spectrum (black) of ^{15}N cysteine-

labelled E,Zn-hSOD1, oxidized E,Zn-hSOD1 purified from the lysate (red), same sample reduced with DTT (blue). (E,F) Detailed views of (D) showing Cys 146 and Cys 6 amide cross-peaks. When DTT is added to the sample of E,Zn-hSOD1$^{S\text{-}S}$, E,Zn-hSOD1$^{SH\text{-}SH}$ is detected, together with E,E-hSOD1$^{SH\text{-}SH}$ and other species (indicated with an asterisk).
doi:10.1371/journal.pone.0023561.g004

Additionally, we have shown that during this process the cysteines involved in the intramolecular disulfide bridge are in the reduced state. Structural information was obtained by comparing the in-cell NMR spectra with spectra recorded *in vitro* on single, defined protein states.

We have shown that apo-hSOD1 in the cytoplasm is monomeric, reduced and in a partially unstructured state. This is the species which has been suggested and is believed to pass the outer membrane of mitochondria, reaching the IMS where it is trapped upon interaction with the copper chaperone for SOD (CCS) [18]. Interestingly, apo-hSOD1 in the *E. coli* cytoplasm appears to be interacting with different cellular components, giving rise to a distribution of conformations which causes the broadening of part of the NH signals.

A striking result of the present characterization is the in cell selectivity for zinc binding. We have shown that when zinc is added in excess to the culture medium, hSOD1 binds only one equivalent of zinc in its native site, while *in vitro* both metal binding sites are able to bind zinc with comparable affinity. The selective binding of only one zinc ion per subunit occurs only in the intact cells, as when they are lysed and hSOD1 is exposed to an excess of zinc, it binds two equivalents per subunit. Zinc uptake is known to be tightly regulated in the cytoplasm of *E. coli*. While the concentration of free zinc is reported to be less than one atom per cell, the total amount of zinc present is thought to be distributed among other cellular species, either zinc binding proteins or small molecules [19]. It is therefore possible that in this regulated environment zinc binding to hSOD1 becomes highly selective towards the zinc binding site, contrarily to what occurs *in vitro*. Conversely, when the cells are lysed the zinc distribution is not controlled anymore and, if zinc in excess is present, Zn,Zn-hSOD1 species is formed.

To reach its mature state, E,Zn-hSOD1$^{SH\text{-}SH}$ still has to bind a copper ion per monomer and the disulfide bridge has to form. However, in most of Gram-negative bacteria, such as *E. coli*, no copper proteins are known to localize in the cytoplasm, and indeed these organisms have copper efflux pumps to remove copper from the cytoplasm, while no protein is present for cytoplasmic copper uptake [20]. Therefore hSOD1 cannot bind copper after zinc, even if copper is added to the culture medium. Copper binding of course can occur in the cytoplasm of eukaryotic cells, which have systems to regulate cytoplasmic copper intake [21–23], provided the CCS chaperone is present.

Concerning the cysteine redox state, we have shown that after zinc binding the formation of the disulfide bridge still does not occur. This suggests that in the eukaryotic cells the disulfide formation process is also mediated by CCS, and it may be concomitant to the copper loading.

Methods

Cell samples preparation

For uniform ^{15}N labelling, BL21(DE3) Gold *E. coli* strain was used, transformed with a pET28a plasmid containing the WT hSOD1 gene sequence without any additional tag. For selective ^{15}N-cysteine labelling, the auxotroph strain BL21(DE3) CysE was transformed with a pET21 plasmid containing the same WT hSOD1 gene sequence.

Cell samples for in-cell NMR were prepared by adapting a reported protocol [24] (Information S2).

For expression of selective ^{15}N-cysteine labelled WT hSOD1 the samples were prepared as described above, but a M9-based reconstituted medium was used [25], containing ^{15}N-cysteine and the other 19 unlabelled amino acids.

Concentration of hSOD1 inside the cytoplasm after 4 h of overexpression was determined by measuring the absorption of the hSOD1 band on a coomassie-stained SDS-PAGE. A purified hSOD1 sample of known concentration (200 uM) was run on the same gel at different dilutions to provide a calibration curve. hSOD1 is ~850 μM in the cytoplasm, while in the lysate sample hSOD1 is ~350 uM.

Purified hSOD1 samples preparation

Pure WT hSOD1 protein was prepared following an existing protocol [26] (Information S3).

The removal of the metals to obtain E,E-hSOD1$^{S\text{-}S}$ was achieved by dialyzing several times a diluted solution of hSOD1 against 10 mM EDTA in 50 mM acetic acid at pH 3.5. After removal of EDTA E,Zn-hSOD1$^{S\text{-}S}$ and Zn,Zn-hSOD1$^{S\text{-}S}$ were then obtained by adding at pH 5.5 one and two equivalents of $ZnSO_4$, respectively.

To obtain E,E-hSOD1$^{SH\text{-}SH}$ and E,Zn-hSOD1$^{SH\text{-}SH}$, the E,E-hSOD1$^{S\text{-}S}$ and E,Zn-hSOD1$^{S\text{-}S}$ were incubated 1 h at 37°C with 50–60 mM of DTT; 1 mM EDTA was added to the sample of E,E-hSOD1$^{SH\text{-}SH}$ to prevent binding of any metal present in traces. DTT concentration was then brought to 2 mM by dialysis against oxygen-free phosphate buffer. Zn,Zn-hSOD1$^{SH\text{-}SH}$ was obtained by adding $ZnSO_4$ in excess (4 equivalents) to a sample of reduced E,E-hSOD1$^{SH\text{-}SH}$.

The final NMR samples obtained were in 20 mM phosphate buffer at pH 7.5; protein concentration ranged between 0.1 and 0.3 mM (referred to the monomer). All NMR spectra were acquired at Bruker Biospin 600 and 800 MHz spectrometers. The latter is equipped with a cryo-cooled TXI probe.

In-cell NMR experiments

We performed a series of in-cell NMR experiments in which protein expression was induced in M9 medium either without zinc or with increasing amounts of $ZnSO_4$ (from 10 μM up to 1 mM). 2D ^{1}H,^{15}N-SOFAST-HMQC spectra were acquired at 310K. The total acquisition time for each cell sample ranged from 1 to 3 h. After the acquisition, the cells were gently centrifuged and collected to be lysed. The supernatant was then checked in the same experimental conditions, in order to exclude the presence of any signal arising from the protein leaked out of the cells (**Figure S3**). After cell lysis by sonication and centrifugation, 2D ^{1}H,^{15}N-SOFAST-HMQC spectra were acquired on the cleared cell lysate.

The cysteine redox state inside the cells and in the cell lysate was determined via NMR by monitoring the ^{1}H and ^{15}N chemical shifts of the ^{15}N-labelled cysteines. After lysis of the in-cell NMR sample, the protein contained was roughly purified with DEAE anion exchange resin, and left exposed to air for >2 h to allow cysteine oxidation. Finally, *in vitro* reduction of the cysteines was performed with 50 mM DTT.

In vitro NMR experiments

Either 2D ^{1}H,^{15}N-HSQC spectra or both ^{1}H,^{15}N-HSQC and ^{1}H,^{15}N-SOFAST-HMQC spectra were acquired on the *in vitro*

^{15}N-labelled samples of WT hSOD1 at 310K. The linewidth of the HSQC crosspeaks is 10% smaller on average than that of the SOFAST-HMQC crosspeaks, while the crosspeak intensities are comparable, the chemical shifts being the same. Unfolding of an *in vitro* sample of E,E-hSOD1$^{SH\text{-}SH}$ was performed by adding increasing amounts of guanidinium chloride up to 0.5 M, and was monitored with 2D ^{1}H,^{15}N-HSQC spectra acquired at 600 MHz. ^{15}N R_1 and R_2 relaxation measurements [27] were performed at 600 MHz on the ^{15}N labelled E,E-hSOD1$^{SH\text{-}SH}$ sample. The τ_m estimated from R_2/R_1 ratios ($\tau_m = 11.1 \pm 1.5$ µs) confirmed that the protein was in the monomeric state [14].

Supporting Information

Figure S1 Combined Chemical Shift Difference (CCSD) plot of in-cell vs. *in vitro* zinc-bound hSOD1 NMR spectra. CCSD plot of a subset of amide resonances of hSOD1 showing that the in-cell+zinc hSOD1 species is more similar to *in vitro* E,Zn-hSOD$^{SH\text{-}SH}$ compared to *in vitro* Zn,Zn-hSOD1$^{SH\text{-}SH}$. CCSDs between in-cell+zinc hSOD1 and Zn,Zn-hSOD$^{SH\text{-}SH}$ (blue) are higher on average than CCSDs between in-cell+zinc hSOD1 and E,Zn-hSOD$^{SH\text{-}SH}$ (orange). Amide cross-peaks of residues Gly 61 and Thr 135 (marked with an asterisk) are shown in **Figure 3B**. CCSDs were calculated using the formula: $CCSD = \sqrt{\frac{1}{2}(\Delta\delta^{1}H)^{2} + \frac{1}{2}(\Delta\delta^{15}N/5)^{2}}$.
(TIF)

Figure S2 Cysteine redox state determined by reaction of hSOD1 with AMS. Non-reducing SDS-PAGE of AMS reaction performed on cell cultures expressing hSOD1 both in presence and in defect of Zn(II) in the medium (right). AMS reaction on *in vitro* samples (left) of reduced and oxidized hSOD1 is showed as a reference.
(TIF)

Figure S3 Supernatant after centrifugation of the cell sample. ^{1}H-^{15}N SOFAST-HMQC spectrum of the supernatant collected after centrifugation of an in-cell NMR sample of hSOD1. The threshold has been lowered to show the very low S/N ratio of the signals detected.
(TIF)

Information S1 Reaction with AMS.
(DOC)

Information S2 Cell and lysate samples preparation.
(DOC)

Information S3 hSOD1 purification protocol.
(DOC)

Author Contributions

Conceived and designed the experiments: L. Banci IB EL. Performed the experiments: L. Barbieri FC EL. Analyzed the data: L. Barbieri FC EL. Wrote the paper: L. Banci L. Barbieri IB FC EL.

References

1. Reckel S, Hänsel R, Löhr F, Dötsch V (2007) In-cell NMR spectroscopy. Prog NMR Spectrosc 51: 91–101.
2. Selenko P, Wagner G (2007) Looking into live cells with in-cell NMR spectroscopy. J Struct Biol 158: 244–253.
3. Dedmon MM, Patel CN, Young GB, Pielak GJ (2002) FlgM gains structure in living cells. Proc Natl Acad Sci U S A 99: 12681–12684.
4. Riaz D, Beckwith J (2001) Roles of thiol-redox pathways in bacteria. Annu Rev Microbiol 55: 21–48.
5. Wilks JC, Slonczewski JL (2007) pH of the cytoplasm and periplasm of Escherichia coli: rapid measurement by green fluorescent protein fluorimetry. J Bacteriol 189: 5601–5607.
6. Hu J, Dong L, Outten CE (2008) The redox environment in the mitochondrial intermembrane space is maintained separately from the cytosol and matrix. J Biol Chem 283: 29126–29134.
7. Sturtz LA, Diekert K, Jensen LT, Lill R, Culotta VC (2001) A fraction of yeast Cu,Zn-superoxide dismutase and its metallochaperone, CCS, localize to the intermembrane space of mitochondria. A physiological role for SOD1 in guarding against mitochondrial oxidative damage. J Biol Chem 276: 38084–38089.
8. Lindberg MJ, Tibell L, Oliveberg M (2002) Common denominator of Cu/Zn superoxide dismutase mutants associated with amyotrophic lateral sclerosis: Decreased stability of the apo state. Proc Natl Acad Sci USA 99: 16607–16612.
9. Furukawa Y, O'Halloran TV (2005) Amyotrophic Lateral Sclerosis Mutations Have the Greatest Destabilizing Effect on the Apo- and Reduced Form of SOD1, Leading to Unfolding and Oxidative Aggregation. J Biol Chem 280: 17266–17274.
10. Guareschi M, Pesaresi MG, Amori I, Crosio C, Ferri A, et al. (2009) Oligomerization of mutant SOD1 in mitochondria of motoneuronal cells drives mitochondrial damage and cell toxicity. Antioxid Redox Signal 11: 1547–1548.
11. Banci L, Bertini I, Girotto S, Martinelli M, Viezzo M, et al. (2007) Metal-free SOD1 forms amyloid-like oligomers: a possible general mechanism for familial ALS. Proc Natl Acad Sci USA 104: 11263–11267.
12. Furukawa Y, Kaneko K, Yamanaka K, O'Halloran TV, Nukina N (2008) Complete loss of post-translational modifications triggers fibrillar aggregation of SOD1 in the familial form of amyotrophic lateral sclerosis. J Biol Chem 283: 24167–24176.
13. Schanda P, Kupce E, Brutscher B (2005) SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. J Biomol NMR 33: 199–211.
14. Arnesano F, Banci L, Bertini I, Martinelli M, Furukawa Y, et al. (2004) The unusually stable quaternary structure of human SOD1 is controlled by both metal occupancy and disulfide status. J Biol Chem 279: 47998–48003.
15. Chattopadhyay M, Valentine JS (2009) Aggregation of copper-zinc superoxide dismutase in familial and sporadic ALS. Antioxid Redox Signal 11: 1603–1614.
16. Banci L, Bertini I, Boca M, Calderone V, Cantini F, et al. (2009) Structural and dynamic aspects related to oligomerization of apo SOD1 and its mutants. Proc Natl Acad Sci U S A 106: 6980–6985.
17. Banci L, Bertini I, Cantini F, D'Amelio N, Gaggelli E (2006) Human SOD1 Before Harboring the Catalytic Metal: Solution Structure of Copper-Depleted, Disulfide-Reduced Form. J Biol Chem 281: 2333–2337.
18. Kawamata H, Manfredi G (2010) Import, maturation, and function of SOD1 and its copper chaperone CCS in the mitochondrial intermembrane space. Antioxid Redox Signal 13: 1375–1384.
19. Outten CE, O'Halloran TV (2001) Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. Science 292: 2488–2492.
20. Rensing C, Grass G (2003) Escherichia coli mechanisms of copper homeostasis in a changing environment. FEMS Microbiol Rev 27: 197–213.
21. Banci L, Bertini I, Cefaï-Ballouz S, Kozyreva T, Zovo K, et al. (2010) Affinity gradients drive copper to cellular destinations. Nature 465: 645–648.
22. Kim BE, Nevitt T, Thiele DJ (2008) Mechanisms for copper acquisition, distribution and regulation. Nat Chem Biol 4: 176–185.
23. Finney LA, O'Halloran TV (2003) Transition Metal Speciation in the Cell: Insights from the Chemistry of Metal Ion Receptors. Science 300: 931–936.
24. Serber Z, Selenko P, Hansel R, Reckel S, Löhr F, et al. (2006) Investigating macromolecules inside cultured and injected cells by in-cell NMR spectroscopy. Nat Protoc 1: 2701–2709.
25. Cheng H, Wemmer WM, Xia B, Oh BH, Markley JL (1995) Protein expression, selective isotopic labeling, and analysis of hyperfine-shifted NMR signals of Anabaena 7120 vegetative [2Fe-2S]ferredoxin. Arch Biochem Biophys 316: 619–634.
26. Banci L, Benedetto M, Bertini I, Del Conte R, Piccioli M, et al. (1998) Solution structure of reduced monomeric Q133M2 Copper, Zinc Superoxide Dismutase. Why is SOD a dimeric enzyme? Biochemistry 37: 11780–11791.
27. Farrow NA, Muhandiram R, Singer AU, Pascal SM, Kay CM, et al. (1994) Backbone dynamics of a free and phosphopeptide-complexed Src homology 2 domain studied by 15N NMR relaxation. Biochemistry 33: 5984–6003.

Figure S1: Combined Chemical Shift Difference (CCSD) plot of in-cell vs. *in vitro* zinc-bound hSOD1 NMR spectra

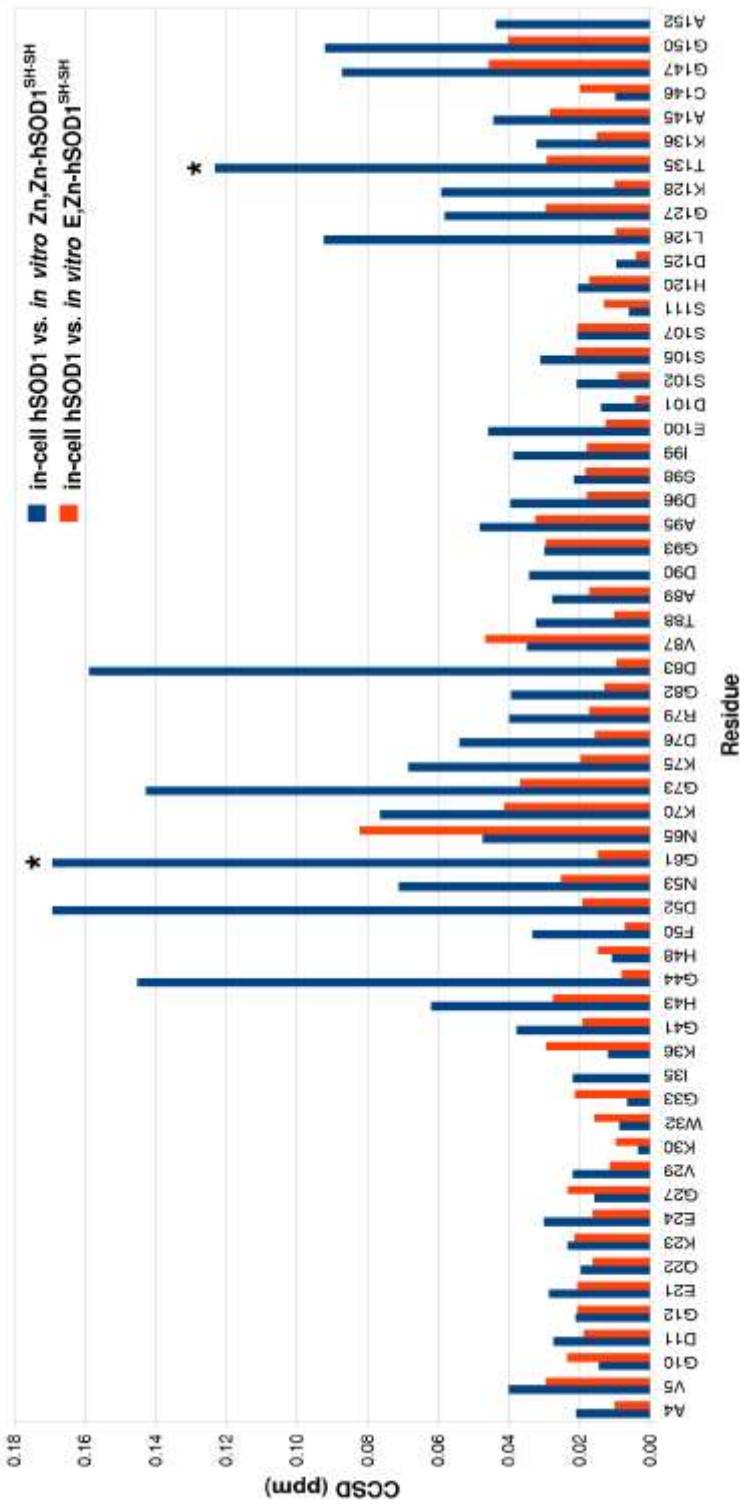


Figure S2: Cysteine redox state determined by reaction of hSOD1 with AMS

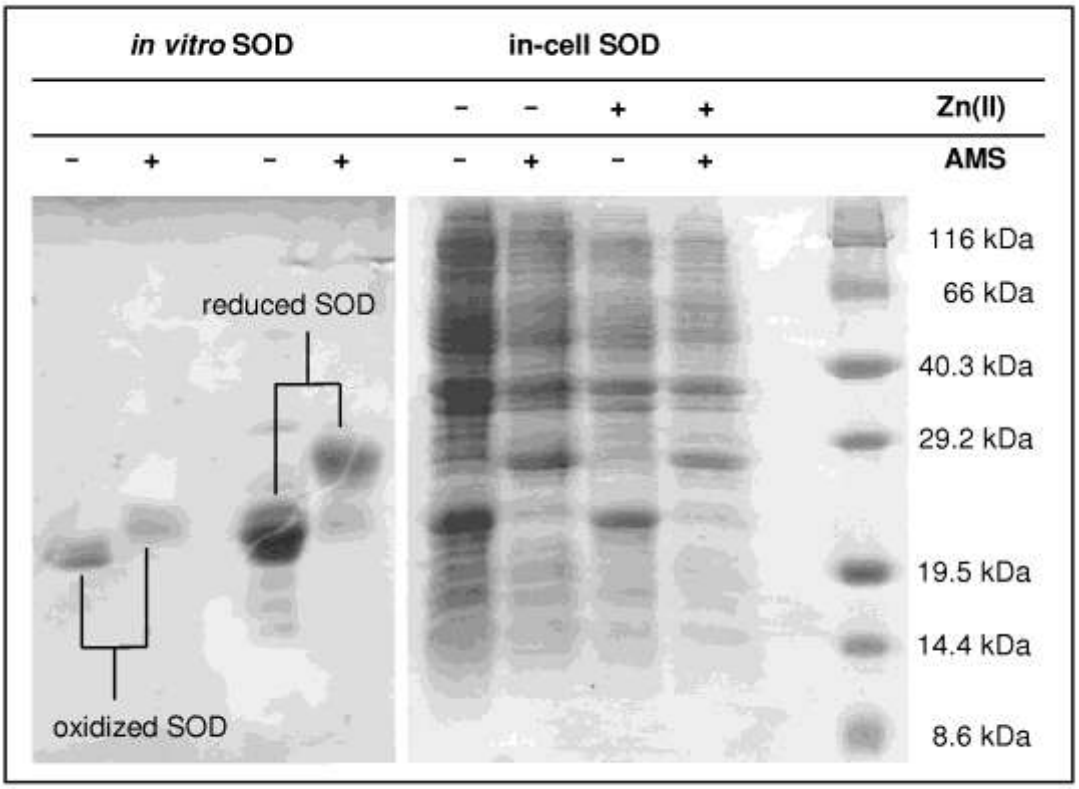
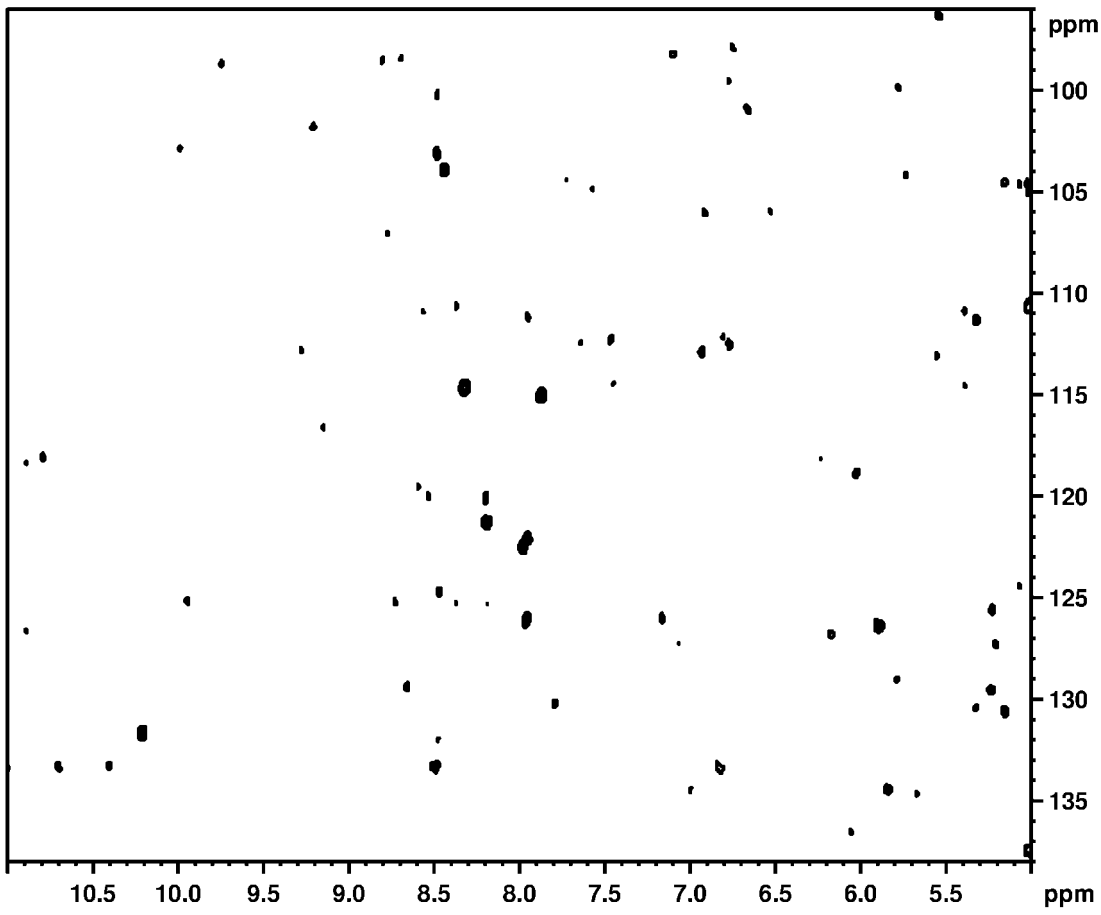


Figure S3: Supernatant after centrifugation of the cell sample



Supporting Information S1. Reaction with AMS

Reaction with 4-acetamido-4'-maleimidylstilbene-2,2'-disulfonic acid (AMS) was performed directly on cell samples in oxygen-free conditions [1]. 1 mL of cell culture was precipitated with 10% trichloroacetic acid (TCA), washed with 80 µL acetone and re-suspended in 100 µL 100 mM Tris pH 7 + 2% SDS. 10 µL of the mixture obtained was incubated 1 h at 37°C with 20 mM AMS, and finally run on a non-reducing SDS-PAGE. The same reaction was performed on *in vitro* samples of $hSOD1^{S\text{-}S}$ and $hSOD1^{SH\text{-}SH}$. 20 µL of an *in vitro* protein sample (0.2 mM) were precipitated with 10% TCA, washed with 40 µL acetone and re-suspended in 50 µL 100 mM Tris pH 7 + 2% SDS. 10 µL of the mixture obtained was incubated 1 h at 37°C with 20 mM AMS.

References:

1. Kobayashi T, Ito K (1999) Respiratory chain strongly oxidizes the CXXC motif of DsbB in the Escherichia coli disulfide bond formation pathway. EMBO J. 18: 1192-1198.

Supporting Information S2. Cell and lysate samples preparation

Cell samples preparation

The M9 buffer used in all cell cultures was checked through ICP-AES for traces of zinc. The concentration of Zn^{2+} was below the detection limit of 0.04 µM.

Cell samples for in-cell NMR were prepared as follows: a cell culture was grown overnight at 30°C in 35 mL of LB medium. After gentle centrifugation (3000 g) for 20 minutes, the cells were re-suspended in 50 mL of M9 minimal medium [M9 buffer (7 g/L K_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, pH 7.4), 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 mg/L biotin, 1 mg/L thiamine, antibiotic] containing 1 g/L $(^{15}NH_4)_2SO_4$ and 3 g/L of unlabelled glucose to obtain an OD_{600} of ~1.6. After 10 min recovery time, overexpression was induced with 0.5 mM IPTG, and carried out at 30°C for 4 h. The cells were washed once with 50 mL of metal-free M9 buffer in order to remove nutrients, metal ions and any excreted by-product, and they were harvested through gentle centrifugation. The pellet was then re-suspended in metal-free M9 buffer until 500 µL of a ~50% v./v. cell slurry were obtained. 50 µL of D_2O were added, and the final volume was put in a 5 mm NMR tube.

Cell lysates preparation

Cleared cell lysates for in-cell NMR experiments were prepared as follows: after removal of the supernatant to be checked by NMR, the cell pellet was re-suspended in an equal volume of metal-free M9 buffer. The cells were then lysed by ultrasonication. Then the lysate was centrifuged at 18000 g for 20 minutes, the supernatant was collected and its volume was brought to 500 µL with M9 buffer. 50 µL of D_2O were then added. The final dilution of the cytoplasm in M9 buffer is around 1:2.

Supporting Information S3. hSOD1 purification protocol

Pure hSOD1 protein was prepared as follows: a cell culture (BL21(DE3) Gold (Stratagene), transformed with a pET28a plasmid containing the WT hSOD1 gene) was grown overnight at 30°C in 750 mL LB, harvested and re-suspended in 2.25 L ^{15}N-labelled M9 medium. After 4 h from induction with 0.5 mM IPTG at 30°C the cells were harvested and re-suspended in 20 mM Tris, pH 8 buffer for lysis. The cleared lysate was loaded on an anion exchange column (DEAE Sepharose Fast Flow resin, GE Healthcare) for a first purification of hSOD1 by elution with NaCl gradient. The collected fractions containing hSOD1 (checked by SDS-PAGE) were further purified by gel filtration (Superdex75 16/60 column, GE Healthcare) in 20 mM Tris, 100 mM NaCl, pH 8 buffer. Fractions containing pure hSOD1 were collected. $ZnSO_4$ was added to the protein solution to increase hSOD1 stability, and 1 mM DTT was added in all buffers to prevent protein aggregation through disulfide bridges.

2.2

Atomic-resolution monitoring of protein maturation in live human cells by NMR

Banci, L; Barbieri, L; Bertini, I; Luchinat, E; Secci, E; Zhao, Y; Aricescu A.R

Nat. Chem. Biol.9, 297–299

nature chemical biology

BRIEF COMMUNICATION

PUBLISHED ONLINE: 3 MARCH 2013 | DOI: 10.1038/NCHEMBIO.1202

Atomic-resolution monitoring of protein maturation in live human cells by NMR

Lucia Banci[1,2]*, Letizia Barbieri[1], Ivano Bertini[1,2,4], Enrico Luchinat[1], Erica Secci[1], Yuguang Zhao[3] & A Radu Aricescu[3]*

We use NMR directly in live human cells to describe the complete post-translational maturation process of human superoxide dismutase 1 (SOD1). We follow, at atomic resolution, zinc binding, homodimer formation and copper uptake, and discover that copper chaperone for SOD1 oxidizes the SOD1 intrasubunit disulfide bond through both copper-dependent and copper-independent mechanisms. Our approach represents a new strategy for structural investigation of endogenously expressed proteins in a physiological (cellular) environment.

Functional understanding of cellular processes requires a detailed characterization of molecular players and their structural and dynamic properties and networks of interactions. A biomolecule should, ideally, be characterized within its cellular milieu to match the physiological environment, including pH, redox potential, viscosity and the presence of all relevant interaction partners. A new approach to structural biology is therefore needed to explore the cellular context using atomic resolution techniques. In principle, in-cell NMR[1–5] represents an ideal method for monitoring protein structure during functional processes and in 'close-to-physiological' conditions. However, in practice, one must overcome a number of technological limitations in order to (i) express (or coexpress) isotopically labeled proteins in cells derived from a suitable organism—human proteins, for example, should be endogenously synthesized and studied in human cells; (ii) establish experimental conditions for maintaining cellular viability inside the NMR tube and (iii) reduce data-acquisition time while increasing measurement sensitivity. We sought to address these challenges and directly monitor the steps involved in post-translational modifications of a model protein, SOD1, within live human cells. Furthermore, we aimed to define the role of copper chaperone for SOD1 (CCS), proposed to be a major contributor to SOD1 maturation[6]. A correct understanding of this process is important given the fundamental role of SOD1 in cellular defense against oxidative stress[7]. Impaired SOD1 maturation has been linked to disease states, including the onset of amyotrophic lateral sclerosis[8,9].

We induced transient expression of human SOD1 in human embryonic kidney (HEK293T) cells[10] and modulated expression levels by transfecting different amounts of cDNA (see Online Methods). Under our culture conditions, the endogenous concentration of SOD1 in HEK293T cells was 10 ± 2 μM (s.d.) (**Supplementary Results**, **Supplementary Fig. 1**). SOD1 concentrations of up to 40 μM have been reported in the cytoplasm of mammalian cells[11]. After recombinant expression, the highest observed intracellular concentration of monomeric SOD1 was 360 ± 30 μM. However, we also detected SOD1 by in-cell NMR at closer-to-physiological intracellular concentrations (45 ± 10 μM) (**Supplementary Fig. 2**). We assessed cellular distribution of SOD1 in isolated nuclear, cytoplasmic and mitochondrial fractions. Irrespective of the total protein abundance, the majority of SOD1 was present in the cytoplasm; <1% was localized in the mitochondrial fraction (**Supplementary Fig. 3**), a lower proportion than previously reported[12].

In cells grown without supplements of zinc or copper ions, we detected two forms of SOD1, in similar amounts: the monomeric, metal-free species (apo-SOD1) and the dimeric species with one Zn^{2+} ion bound to each subunit (E,Zn-SOD1) (**Supplementary Fig. 4a**). To analyze the NMR data, we took advantage of previous *in vitro* spectra and backbone assignments[13,14]. The observed mixture is probably due to residual zinc in the culture medium (~4.5 μM and ~1 μM in medium supplemented with 10% and 2% FBS, respectively; see Online Methods). Cys57 and Cys146, which form an intrasubunit disulfide bond in the mature enzyme, were reduced in both species (**Fig. 1a**). We further confirmed that only these two species were present by NMR analysis of cell extracts (**Supplementary Fig. 4b**). The ^{1}H-^{15}N spectrum of apo-SOD1 was consistent with that of its equivalent produced in *Escherichia coli* cells without addition of metal ions[15], as only the cross-peaks of the less-structured parts were detected (**Supplementary Fig. 4a**). Addition of Zn^{2+} to the culture medium eliminated apo-SOD1 species signals (**Fig. 1b**). The uniformly ^{15}N-labeled cell sample yielded a well-resolved spectrum of E,Zn-SOD1 (**Fig. 1c**), which was improved by removal of background signals arising from nonselective labeling of cellular components (**Supplementary Fig. 5**). Therefore, Zn^{2+} ions are efficiently taken up by cells in culture and bind specifically to the native binding site of dimeric SOD1 in stoichiometric amounts, as further confirmed by ^{1}H-^{15}N cross-peak analysis (**Supplementary Fig. 6**). When reduced apo-SOD1 is exposed to Zn^{2+} ions *in vitro* or in cell lysates, even at sub-stoichiometric concentrations, a mixture of forms is generated that includes apo-SOD1, E,Zn-SOD1 and Zn,Zn-SOD1 (in which a second Zn^{2+} is bound to the copper binding site)[15]. Therefore, site-selectivity of Zn^{2+} binding can be achieved only in the physiological cellular context.

Whereas the mechanisms for copper uptake and excretion are simple in prokaryotes, they are tightly regulated in eukaryotic cells[16]. Accordingly, copper added as Cu(II) salt to the *E. coli* cell culture medium was reduced to Cu(I) and, only in this redox state, bound readily and stoichiometrically to recombinantly expressed SOD1, forming Cu(I),Zn-SOD1 (**Supplementary Fig. 7a,b**). Notably, Cu(I) added to *E. coli* cells either as an acetonitrile or glutathione complex did not become available to E,Zn-SOD1 (**Supplementary Fig. 7c**). In-cell NMR spectra also showed that in *E. coli*, a sizable fraction of the SOD1 intrasubunit disulfide bond is formed (~50%) (**Supplementary Fig. 7d**). In eukaryotic cells, unlike bacteria, copper entrance and delivery to copper-binding proteins requires a number of steps involving specific chaperones responsible for its intracellular trafficking[17,18]. Accordingly, only ~25% of the recombinant SOD1 protein incorporated copper (in the Cu(I) state) in human cells cultured in the presence of Cu(II) (**Fig. 2a**).

[1]CERM, Magnetic Resonance Center, University of Florence, Florence, Italy. [2]Department of Chemistry, University of Florence, Florence, Italy. [3]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK. [4]Deceased. *e-mail: banci@cerm.unifi.it or radu@strubi.ox.ac.uk.

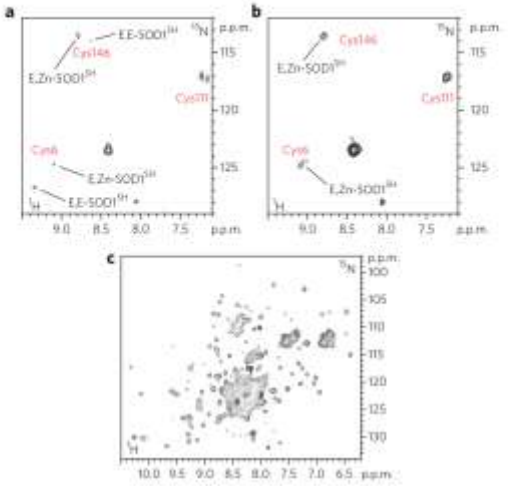


Figure 1 | Zn(II) added to the culture medium promotes binding of one Zn^{2+} ion per apo-SOD1 subunit in the cytoplasm. (**a**,**b**) ^1H-^{15}N SOFAST HMQC spectra acquired on human cells expressing [^{15}N]cysteine-labeled SOD1 in the absence of metals (**a**) and with Zn(II) added to the culture medium (**b**). Assigned cysteine residues are indicated in red. When two species of SOD1 are present, labels indicate the species to which each cross-peak belongs (E,E-SOD1[ox], reduced apo-SOD1; E,Zn-SOD1[red], reduced SOD1 containing one Zn^{2+} ion per subunit). Unlabeled cross-peaks correspond to cellular background signals. (**c**) ^1H-^{15}N SOFAST HMQC acquired on human cells expressing uniformly ^{15}N-labeled SOD1 in Zn(II)-supplemented medium. Comparison between the chemical shifts of the in-cell spectrum with the backbone assignments of SOD1 in different metallation and redox states *in vitro*[13,14] allowed the determination of the species present in the cytoplasm (**Supplementary Fig. 6**).

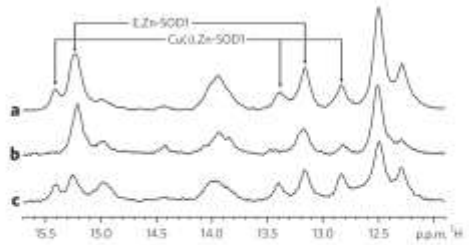


Figure 2 | Cu(II) addition to the culture medium induces Cu(I) binding to a fraction of cytoplasmic SOD1. (**a**,**b**) Histidine regions of ^1H NMR spectra acquired on human cells expressing unlabeled SOD1 in Zn(II)-supplemented medium with (**a**) and without (**b**) incubation with Cu(II). (**c**) ^1H NMR spectrum of human cells coexpressing SOD1 and CCS in Zn(II)-supplemented medium after incubation with Cu(II). Histidine protons unambiguously assigned to Cu(I),Zn-SOD1 species are indicated.

The remaining SOD1 fraction contained only one Zn^{2+} ion per subunit, as observed from the histidine proton signals in the 1D ^1H NMR spectrum (**Fig. 2a–c**). Additionally, the spectra of the [^{15}N]cysteine-labeled protein showed only ~20% intrasubunit disulfide bond formation in SOD1 (**Supplementary Fig. 8**).

Copper incorporation of SOD1 in eukaryotes is dependent on the CCS protein[15,18]. Although basal expression of the human gene *CCS* does occur during normal cell growth, it is likely that the amount of SOD1 produced in our experimental setup was too high for complete copper insertion through the CCS-dependent pathway. However, a CCS-independent copper-insertion pathway has also been reported for human SOD1 (refs. 21,22) and might have contributed to the observed partial formation of Cu(I),Zn-SOD1 (**Fig. 2a**). Simultaneous overexpression and isotopic enrichment of both SOD1 and CCS caused a reduction in overall SOD1 expression (**Supplementary Fig. 2**). Nevertheless, SOD1 signals were readily identifiable (**Fig. 3a**,**b**), and interference from CCS was minimal owing to its greater molecular mass. Intracellular protein concentration ranged from 70 ± 10 μM to 45 ± 10 μM for SOD1 and 50 ± 10 μM to 15 ± 3 μM for CCS (**Supplementary Fig. 2**). Coexpression of SOD1 and CCS in zinc-supplemented medium resulted in dimeric, zinc-containing SOD1 species. We observed no difference in the SOD1 metal content with respect to the cells expressing basal levels of CCS (**Supplementary Fig. 9**). However, in spectra recorded from cells coexpressing the two proteins, we observed partial formation (~50%) of the SOD1 intrasubunit disulfide bond (**Fig. 3a**). This result is consistent with a mechanism of CCS-mediated SOD1 disulfide oxidation that does not require copper insertion into SOD1.

When we added Cu(II) to cells coexpressing SOD1 and CCS we observed a higher ratio of Cu(I),Zn-SOD1 to E,Zn-SOD1 (~1:1; **Fig. 2c**) than in cells expressing basal levels of CCS (**Fig. 2a**), indicating that CCS promotes the incorporation of copper into SOD1. The redox state of Cys57 and Cys146 in SOD1 is affected by both CCS overexpression and copper presence. Indeed, when we incubated cells overexpressing SOD1 and CCS with Cu(II), we observed complete formation of the SOD1 intrasubunit disulfide bond (**Fig. 3b**). We attempted to explore the above mechanism at SOD1 and CSS concentrations that were lower and as close as possible to reported physiological levels[11]. We observed complete disulfide bond formation in E,Zn-SOD1 at concentrations of 45 ± 10 μM SOD1 and 15 ± 3 μM CCS (**Supplementary Fig. 10a**,**b**), and additional incubation with Cu(II) resulted in the complete formation of oxidized Cu(I),Zn-SOD1 (**Supplementary Fig. 10c**,**d**). These results suggest that recombinant protein abundance affects the relative amounts of species but not the sequence of maturation events. We speculate that lower concentrations of SOD1, undetectable by NMR, would result in complete formation of mature SOD1 even with endogenous levels of CCS and copper. In such conditions, however, it would not be possible to obtain information on the intermediate maturation steps. When only SOD1 is overexpressed, the components of its maturation pathway are not abundant enough to complete the process, allowing intermediate SOD1 maturation states to be detected. Furthermore, by selectively increasing single components (for example, CCS or copper), we can recover specific steps of SOD1 maturation and detect different SOD1 states. This 'knock-in' approach can help in understanding which components are necessary for each step of the process.

We have observed the key steps of SOD1 maturation by in-cell NMR (**Fig. 3c**). Specifically, we found that (i) apo-SOD1 is largely unfolded and monomeric in the HEK293T cytoplasm, (ii) zinc uptake occurs without the need of a chaperone, (iii) copper uptake and formation of the Cys57-Cys146 disulfide bond occur partially in cells exposed to Cu(II) (which is reduced to Cu(I)), and (iv) copper loading and complete oxidation of Cys57 and Cys146 are achieved in cells coexpressing SOD1 and CCS. Our experiments also reveal that, in a physiological context, CCS is able to oxidize Cys57 and Cys146 in the absence of copper binding to SOD1. This finding is consistent with the previously reported effect of CCS in promoting SOD1 disulfide bond formation[23], and demonstrates that cysteine oxidation can occur *in vivo* independently of copper transfer and, thus, differs from the mechanism observed *in vitro*[24].

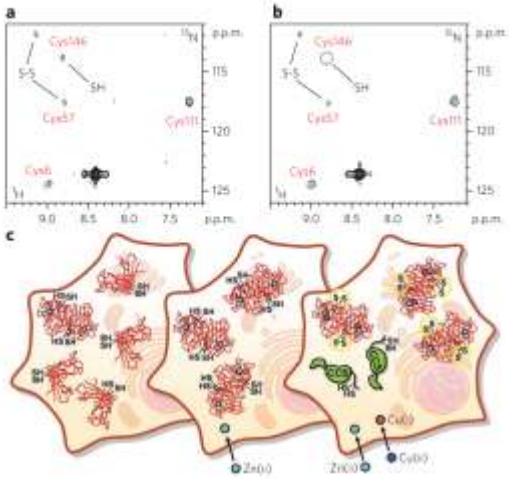


Figure 3 | The redox state of SOD1 is influenced by both copper binding and the presence of CCS. (**a**,**b**) ^{1}H-^{15}N SOFAST HMQC spectra acquired on human cells coexpressing [^{15}N]cysteine-labeled SOD1 and CCS in Zn(II)-supplemented medium before (**a**) and after (**b**) incubation with Cu(I). Assigned cysteine residues are indicated in red. When two species of SOD1 are present, labels indicate the cysteine redox state of each species. Unlabeled cross-peaks are cellular background signals. (**c**) Summary illustration of SOD1 maturation steps. Left, in cells with no additional metals, SOD1 is present mainly in the monomeric, partially unfolded apo form. A fraction of SOD1 binds the zinc present in the expression medium. Middle, SOD1 quantitatively binds one Zn^{2+} ion per monomer and dimerizes upon addition of Zn(II) (cyan) to the expression medium; the cysteines involved in the intrasubunit disulfide bond are completely reduced. Right, a fraction of SOD1 binds Cu(I) (orange) when Zn(II) and Cu(I) (blue) are added to the expression medium and CCS is coexpressed; the disulfide bond is completely formed (yellow circles). S-S, oxidized cysteines forming a disulfide bond; SH, reduced cysteines.

To establish whether the approach described here is applicable to proteins other than SOD1, we selected four human proteins (mitochondrial intermembrane space import and assembly protein 40 (Mia40), copper transport protein Atox1, glutaredoxin-1 (Grx1) and thioredoxin (Trx)) in which properties such as protein folding, binding of metal ions and redox potential of cysteine residues differ from those of SOD1. Using the protocol we established for SOD1 and CCS, we found that all of these targets were highly expressed and visible in the ^{1}H NMR spectra above the cellular background (**Supplementary Fig. 11**). Mia40 and Atox1 were detectable in ^{1}H-^{15}N spectra on uniformly ^{15}N-labeled cell samples (**Supplementary Fig. 12a**,**b**). Grx1 and Trx became visible only upon cell lysis, suggesting that some interaction in the cytoplasm makes the protein tumbling slower on average, thus broadening the amide cross-peaks beyond detection (**Supplementary Fig. 12c**,**d**). Such proteins may be successfully characterized through different NMR techniques (such as solid-state MAS NMR for slow-tumbling proteins)[15].

Successful NMR analysis of endogenously expressed proteins in live mammalian cells requires efficient cDNA transfection, relatively high protein expression levels, the application of different labeling strategies and the maintenance of cell integrity during measurements. In this study, we have addressed all of these aspects and followed the sequential, physiological order of the events in the SOD1 post-translational modification process to obtain information that cannot be retrieved *in vitro*. To our knowledge, this is the first time a complete protein-maturation process has been followed in atomic detail in a living cell. This strategy is potentially applicable for many other protein targets and, thus, opens the way for a broad range of in-cell structural studies of proteins at the molecular level.

Received 26 September 2012; accepted 31 January 2013; published online 3 March 2013

Methods

Methods and any associated references are available in the online version of the paper.

References

1. Reckel, S., Hänsel, R., Löhr, F. & Dötsch, V. *Prog. Nucl. Magn. Reson. Spectrosc.* **51**, 91–101 (2007).
2. Burz, D.S. & Shekhtman, A. *PLoS ONE* **3**, e2571 (2008).
3. Inomata, K. *et al. Nature* **458**, 106–109 (2009).
4. Ogino, S. *et al. J. Am. Chem. Soc.* **131**, 10834–10835 (2009).
5. Selenko, P. *et al. Nat. Struct. Mol. Biol.* **15**, 321–329 (2008).
6. Culotta, V.C., Yang, M. & O'Halloran, T.V. *Biochim. Biophys. Acta* **1763**, 747–758 (2006).
7. Michiels, C., Raes, M., Toussaint, O. & Remacle, J. *Free Radic. Biol. Med.* **17**, 235–248 (1994).
8. Lindberg, M.J., Tibell, L. & Oliveberg, M. *Proc. Natl. Acad. Sci. USA* **99**, 16607–16612 (2002).
9. Banci, L. *et al. Proc. Natl. Acad. Sci. USA* **104**, 11263–11267 (2007).
10. Ariescu, A.R., Lu, W. & Jones, E.Y. *Acta Crystallogr. D Biol. Crystallogr.* **62**, 1243–1250 (2006).
11. Chang, L.Y., Slot, J.W., Geuze, H.J. & Crapo, J.D. *J. Cell Biol.* **107**, 2169–2179 (1988).
12. Sturtz, L.A., Diekert, K., Jensen, L.T., Lill, R. & Culotta, V.C. *J. Biol. Chem.* **276**, 38084–38089 (2001).
13. Banci, L., Bertini, I., Cramaro, F., Del Conte, R. & Viezzoli, M.S. *Biochemistry* **42**, 9543–9553 (2003).
14. Banci, L., Bertini, I., Cantini, F., D'Amelio, N. & Gaggelli, E. *J. Biol. Chem.* **281**, 2333–2337 (2006).
15. Banci, L., Barbieri, L., Bertini, I., Cantini, F. & Luchinat, E. *PLoS ONE* **6**, e23561 (2011).
16. Kim, B.E., Nevitt, T. & Thiele, D.J. *Nat. Chem. Biol.* **4**, 176–185 (2008).
17. Puig, S. & Thiele, D.J. *Curr. Opin. Chem. Biol.* **6**, 171–180 (2002).
18. Banci, L. *et al. Nature* **465**, 645–648 (2010).
19. Schmidt, P.J. *et al. J. Biol. Chem.* **274**, 23719–23725 (1999).
20. Furukawa, Y., Torres, A.S. & O'Halloran, T.V. *EMBO J.* **23**, 2872–2881 (2004).
21. Carroll, M.C. *et al. Proc. Natl. Acad. Sci. USA* **101**, 5964–5969 (2004).
22. Leitch, J.M., Yick, P.J. & Culotta, V.C. *J. Biol. Chem.* **284**, 24679–24683 (2009).
23. Proescher, J.B., Son, M., Elliott, J.L. & Culotta, V.C. *Hum. Mol. Genet.* **17**, 1728–1737 (2008).
24. Banci, L. *et al. Proc. Natl. Acad. Sci. USA* **109**, 13555–13560 (2012).
25. Reckel, S., Lopez, J.J., Löhr, F., Glaubitz, C. & Dötsch, V. *ChemBioChem* **13**, 534–537 (2012).

Acknowledgments

We would like to thank E.Y. Jones and D.I. Stuart for critically reading the manuscript and providing advice. This work was supported by the Access to Research Infrastructures activities in the Seventh Framework Programme of the European Commission (Bio-NMR contract 261863 and P-CUBE contract 227764), by the Italian grant program MIUR-PRIN 2009 'Biologia strutturale meccanicistica: avanzamenti metodologici e biologici' and by the European Integrated Structural Biology Infrastructure (INSTRUCT), which is part of the European Strategy Forum on Research Infrastructures (ESFRI) and supported by national member subscriptions. Specifically, we thank the INSTRUCT Core Centres CERM (Italy) and the University of Oxford. A.R.A. was supported by a UK Medical Research Council Career Development Award fellowship.

Author contributions

L. Banci, I.B. and A.R.A. conceived the work; L. Banci, L. Barbieri, E.L. and A.R.A. designed the experiments; L. Barbieri and Y.Z. grew and transfected the human cells; L. Barbieri cloned the genes and produced the cell samples of SOD1, CCS, Atox1, Grx1 and Trx; E.S. produced the cell sample of Mia40; L. Barbieri and E.L. produced and analyzed the *E. coli* cell samples; E.L. performed the NMR experiments and analyzed the data; L. Banci, L. Barbieri, I.B., E.L. and A.R.A. wrote the paper.

Competing financial interests

The authors declare no competing financial interests.

Additional information

Supplementary information is available in the online version of the paper. Reprints and permissions information is available online at http://www.nature.com/reprints/index.html. Correspondence and requests for materials should be addressed to L. Banci or A.R.A.

ONLINE METHODS

Constructs. cDNAs encoding full-length human SOD1 (amino acids 1–154, GenBank: NP_000445.1), CCS (amino acids 1–274, GenBank: NP_005116.1), Mia40 (amino acids 1–142, GenBank: NP_001091972.1), Atox1 (amino acids 1–68, GenBank: NP_004036.1), glutaredoxin-1 (amino acids 1–106, GenBank: NP_001112362.1) and thioredoxin (amino acids 1–105, GenBank: NP_003320.2) were amplified by PCR and cloned into the pHLsec[16] vector between EcoRI and XhoI restriction enzyme sites to generate the mammalian expression plasmids. All clones were verified by DNA sequencing.

Cell culture and transfection. HEK293T cells were maintained in DMEM (high glucose, D6546, Sigma) supplemented with L-glutamine, antibiotics (penicillin and streptomycin) and 10% FBS (Gibco) in uncoated 75-cm^2 plastic flasks and incubated at 310 K, 5% CO_2 in a humidified atmosphere. Cells were transiently transfected with the pHLsec plasmid containing the hSOD1 cDNA using polyethylenimine (PEI), as described elsewhere[18]. Different DNA/PEI ratios were tested for maximizing protein expression (PEI was kept constant at 50 µg per flask), and an optimal ratio of 1:2 was found (25 µg DNA, 50 µg PEI per flask. For coexpression of SOD1 and CCS, cells were transfected with plasmids containing hSOD1 and hCCS constructs in different amounts and ratios. The highest expression of both proteins was obtained by transfection with a ratio of 1:1:2 (hSOD1/hCCS/PEI), thus doubling the total DNA amount. Lower expression levels of SOD1 were obtained by transfecting cells with a 1:4 hSOD1/PEI ratio. To decrease the expression levels of both proteins, cells were transfected with a 1:1:4 hSOD1/hCCS/PEI ratio. PEI was kept constant at 50 µg per flask. Different times of SOD1 expression were tested (1, 2, 3 and 6 d), and the highest amount of protein was reached after 2 d (48 h) of expression. During protein expression, cells were incubated at 310 K in 75-cm^2 flasks. Commercial DMEM medium was used for unlabeled in-cell NMR samples; BioExpress6000 medium (CIL) was used for uniform ^{15}N-labeling; for selective [^{15}N]cysteine labeling, a reconstituted medium was prepared following the DMEM (Sigma)-reported composition, in which [^{15}N]cysteine was added together with all the other unlabeled components. Expression medium was supplemented with 2% FBS. Zn(II) was supplemented as $ZnSO_4$, which was added to the expression medium to a final concentration of 10 µM immediately after transfection. Basal zinc concentration was calculated by considering 45 µM zinc present in the FBS, as reported by Sigma-Aldrich Media Expert (http://www.sigmaaldrich.com/life-science/cell-culture/learning-center/media-expert.html). Cu(II) was supplemented as $CuCl_2$, added to expression medium to a final concentration of 100 µM after 48 h of protein expression and incubated for 24 h. Protein expression was monitored by comparing the protein band intensities in the cell extracts with bands of *in vitro* samples of known concentration run on Coomassie-stained SDS-PAGE.

Human cell samples for in-cell nuclear magnetic resonance. Samples for in-cell NMR were prepared following a reported protocol[5] with some variations: HEK293T cells from a 75-cm^2 culture flask were detached with trypsin-EDTA 0.05% (Gibco) and resuspended in 20 mL DMEM containing 10% FBS to inactivate trypsin. Cells were gently centrifuged (800 × *g*), resuspended in 10 mL PBS, washed once with PBS and resuspended in one cell-pellet volume of DMEM medium supplemented with 90 mM glucose, 16 mM HEPES buffer and 20% D_2O (to obtain 10% D_2O in the final NMR sample). The cell suspension was transferred to a 3-mm Shigemi NMR tube; the glass plunger was not used. Cells were allowed to settle at the bottom of the tube, thus filling up the active coil volume. The supernatant was kept during the NMR experiments to obtain good field homogeneity. After the experiments, cells were resuspended in the supernatant, removed from the NMR tube and spun down again to collect the medium for the protein-leakage test (**Supplementary Fig. 13**). Cells were lysed by the freeze-thaw method after being suspended in one pellet volume of PBS buffer supplemented with 0.5 mM EDTA and 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF). The lysate was centrifuged at 16,000 × *g* for 30 min at 4 °C, and the cleared cell extract was collected for NMR and SDS-PAGE analysis.

Nuclear magnetic resonance experiments. NMR experiments were acquired using a 950-MHz Bruker Avance III spectrometer equipped with a CP TCI CryoProbe. 1D ^{1}H and 2D ^{1}H-^{15}N SOFAST HMQC[26] spectra were acquired at 305 K. The total acquisition time for each cell sample ranged from 1 to 2 h. The supernatant of each cell sample was checked in the same experimental conditions to exclude the presence of any signal arising from protein that had leaked out of the cells. In the above experimental conditions, very low protein signal was detected in the external medium (<10% of the signal in cells). The same NMR spectra were also acquired on the cell extracts. Cell viability before and after NMR experiments was assessed by trypan blue staining[27]. Cell viability remained above 90%, as damaged cells ranged from 3% before the experiments to 8% after the experiments.

***E. coli* cell samples.** Samples of *E. coli* cells expressing human SOD1 were prepared as previously described[15]. For copper incorporation experiments, after 4 h of SOD1 expression, cells were incubated with either 100 µM Cu(II)SO_4, Cu(I)–acetonitrile complex or Cu(I)–glutathione complex for 15 min. Cells were then washed once with M9 buffer and collected for NMR sample preparation. NMR spectra on *E. coli* cell samples were acquired at 305 K using an 800-MHz Bruker Biospin spectrometer equipped with a TXI CryoProbe.

26. Schanda, P. & Brutscher, B. *J. Am. Chem. Soc.* **127**, 8014–8015 (2005).
27. Freshney, R. *Culture of Animal Cells: A Manual of Basic Technique*, 2nd edn. (Wiley-Liss, 1987).

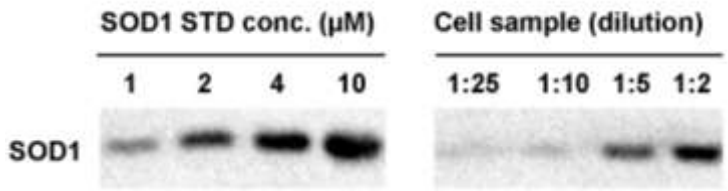
doi:10.1038/nchembio.1202

Supplementary Information

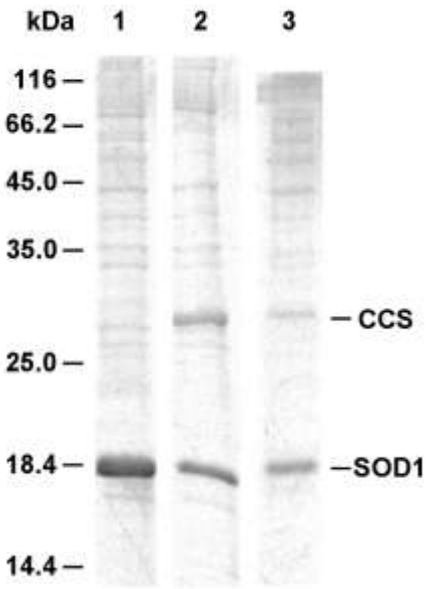
Atomic-resolution monitoring of protein maturation in live human cells

Lucia Banci*, Letizia Barbieri, Ivano Bertini, Enrico Luchinat, Erica Secci, Yuguang Zhao, A. Radu Aricescu*

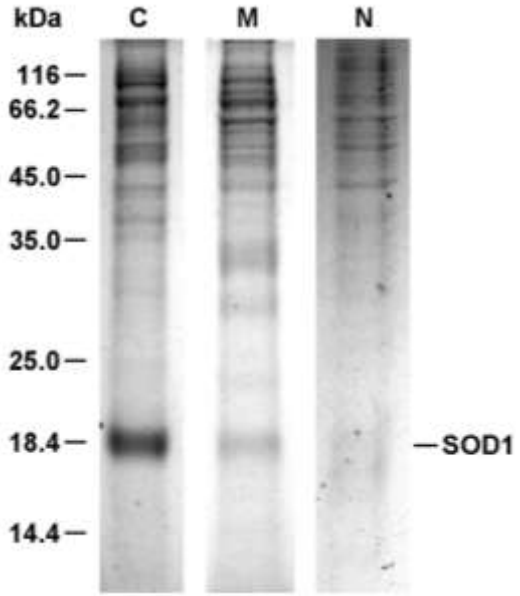
Contains Supplementary Figures 1 to 14.



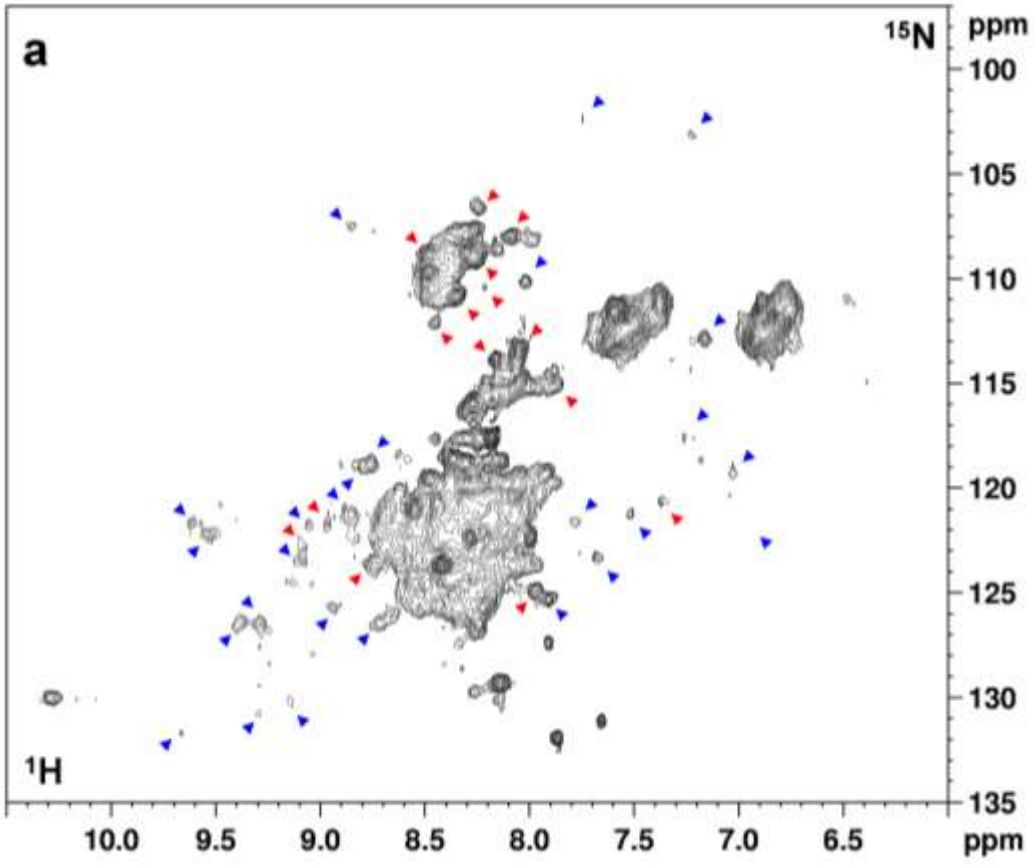
Supplementary Figure 1. Endogenous SOD1 levels in untransfected HEK293T cells were measured by Western Blot analysis. The endogenous amount of SOD1 was estimated in untransfected total cell lysates by densitometric analysis of band intensities, in comparison with different dilutions of pure SOD1. SOD1 was detected using a rabbit anti-human SOD1 polyclonal primary antibody (BioVision) diluted 1:100 (2μg/ml) and a goat anti-rabbit IgG-peroxidase secondary antibody for detection (Sigma), diluted at 1:80000.

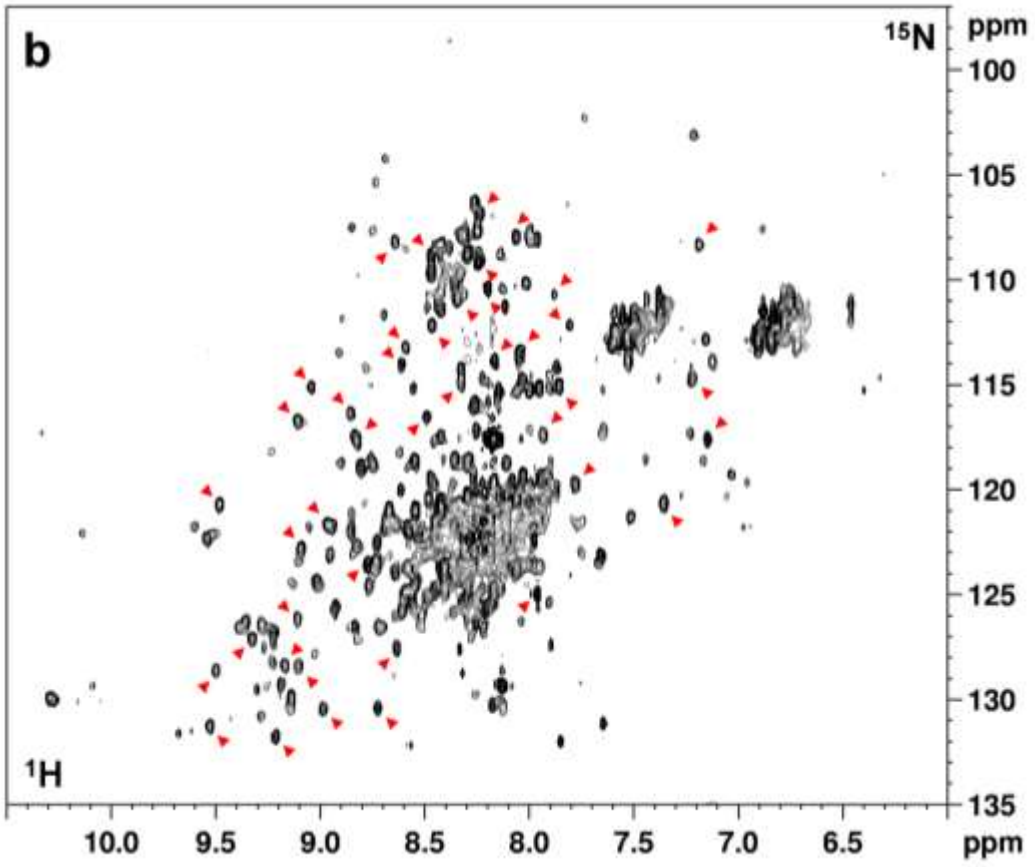


Supplementary Figure 2. Overexpression of SOD1 and CCS in human cells. Coomassie-stained SDS-PAGE of cell extracts from NMR samples of human cells expressing SOD1 (1:2 hSOD1:PEI ratio, lane 1), co-expressing SOD1 and CCS (1:1:2 hSOD1:hCCS:PEI ratio, lane 2) in Zn(II)-supplemented medium. The expression level was reduced by lowering the amount of DNA (1:1:4 hSOD1:hCCS:PEI ratio, lane 3). Cell extracts were obtained by freeze/thaw lysis followed by centrifugation. Protein bands and reference molecular weights are indicated. Protein concentration was estimated by comparing band intensities with serial dilutions of a pure SOD1 sample at known concentration, ran in the same gel (see Supplementary Fig. 14a,b).

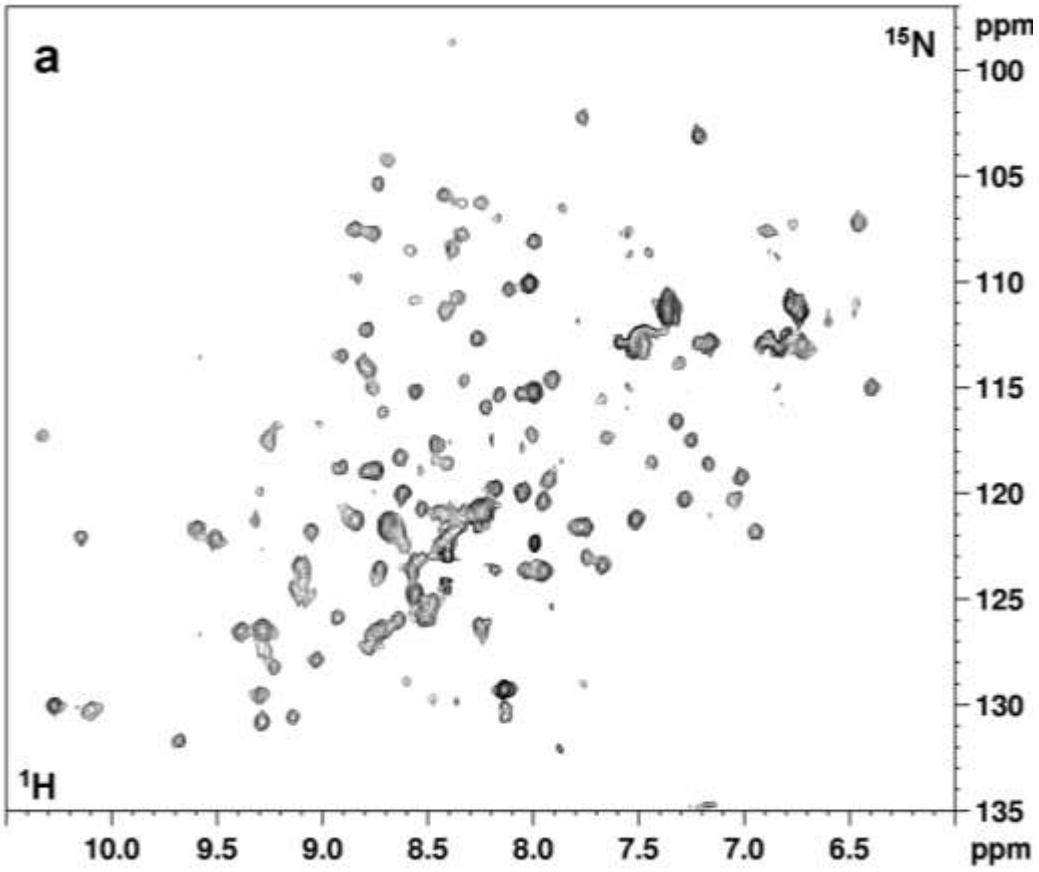


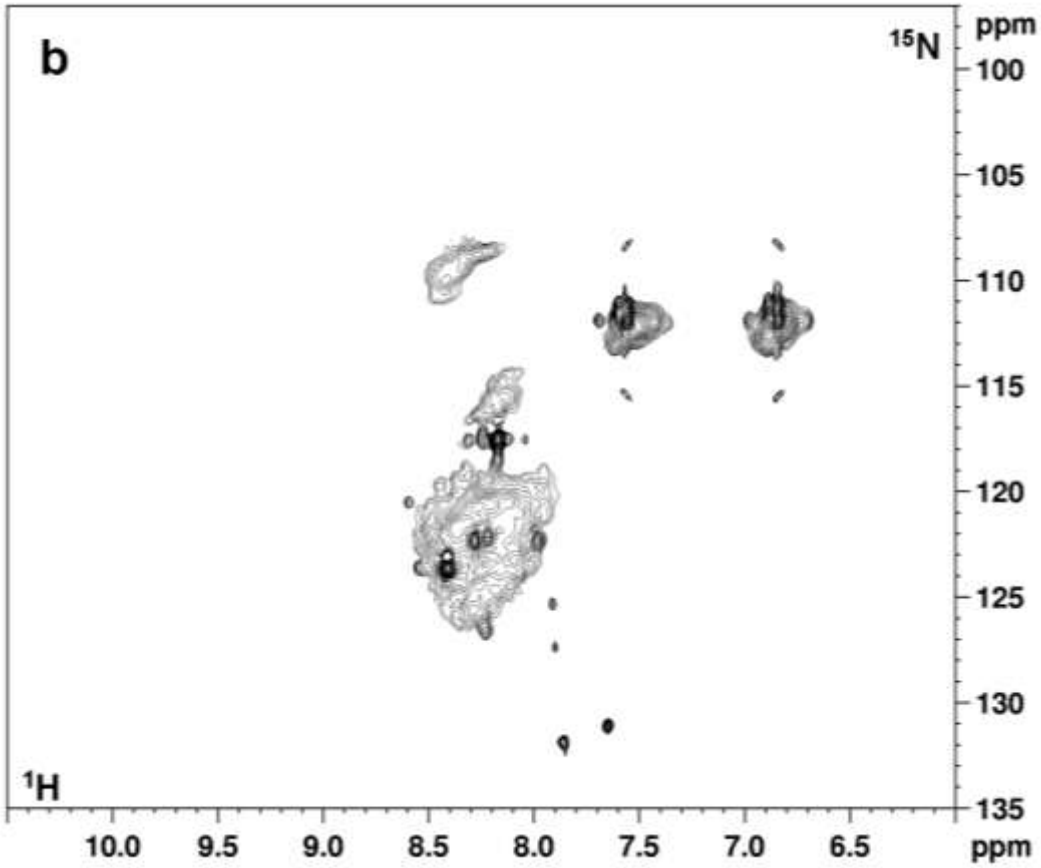
Supplementary Figure 3. Cellular fractionation shows that SOD1 is present mainly in the cytoplasm. Coomassie-stained SDS-PAGE of sub-cellular fractions obtained from human cells expressing SOD1 using a mitochondria isolation kit for cultured cells (Thermo Scientific). The nuclear fraction was obtained by washing once the pellet obtained after cell rupture, and resuspending it in PBS buffer. C = cytoplasm; M = mitochondria; N = nuclei. Relative dilutions are 1:16 (C), 1:1 (M), 1:10 (N) respectively (see Supplementary Fig. 14c,d). SOD1 concentration was estimated via SDS-PAGE by comparing each fraction at different dilutions with a dilution series of pure SOD1.


a
^{15}N
ppm
100
105
110
115
120
125
130
135
^{1}H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

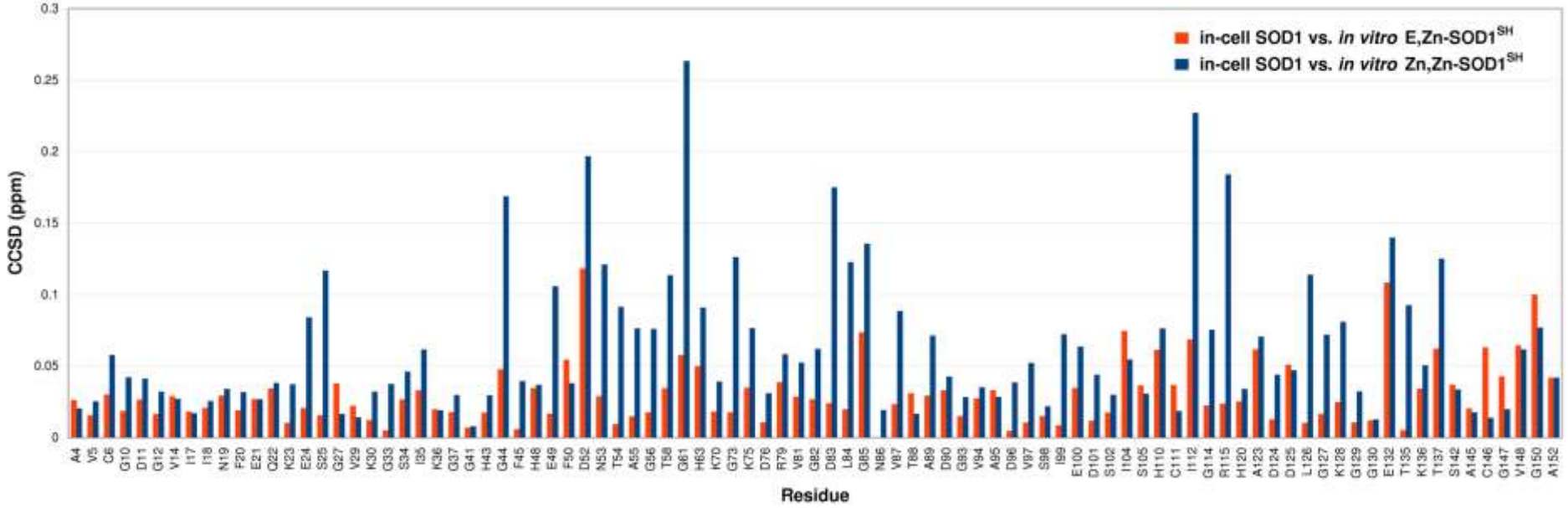

b
^{15}N
ppm
100
105
110
115
120
125
130
135
^{1}H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

Supplementary Figure 4. SOD1 is present both in the metal-depleted form (apo-SOD1^{SH}) and in the zinc-containing form (E,Zn-SOD1^{SH}) in cells without metal supplementation of tissue culture media. ^{1}H-^{15}N SOFAST HMQC spectra were acquired: **a**, on human cells expressing uniformly ^{15}N-labelled SOD1; **b**, on the corresponding cell extract. In the in-cell NMR spectrum (**a**) only the crosspeaks of the unfolded region of apo-SOD1^{SH} are detected, (at 8.0-8.5 ppm ^{1}H, above the cellular background signals) while most of the peaks of the folded region are broadened beyond detection. Crosspeaks unambiguously assigned to apo-SOD1^{SH} and E,Zn-SOD1^{SH} are labelled with red and blue arrows, respectively. Most of the crosspeaks of apo-SOD1^{SH} are detected in the cell extract (**b**), together with those of E,Zn-SOD1^{SH}, while no crosspeaks from other species are detected. Crosspeaks unambiguously assigned to apo-SOD1^{SH} are labelled with red arrows.

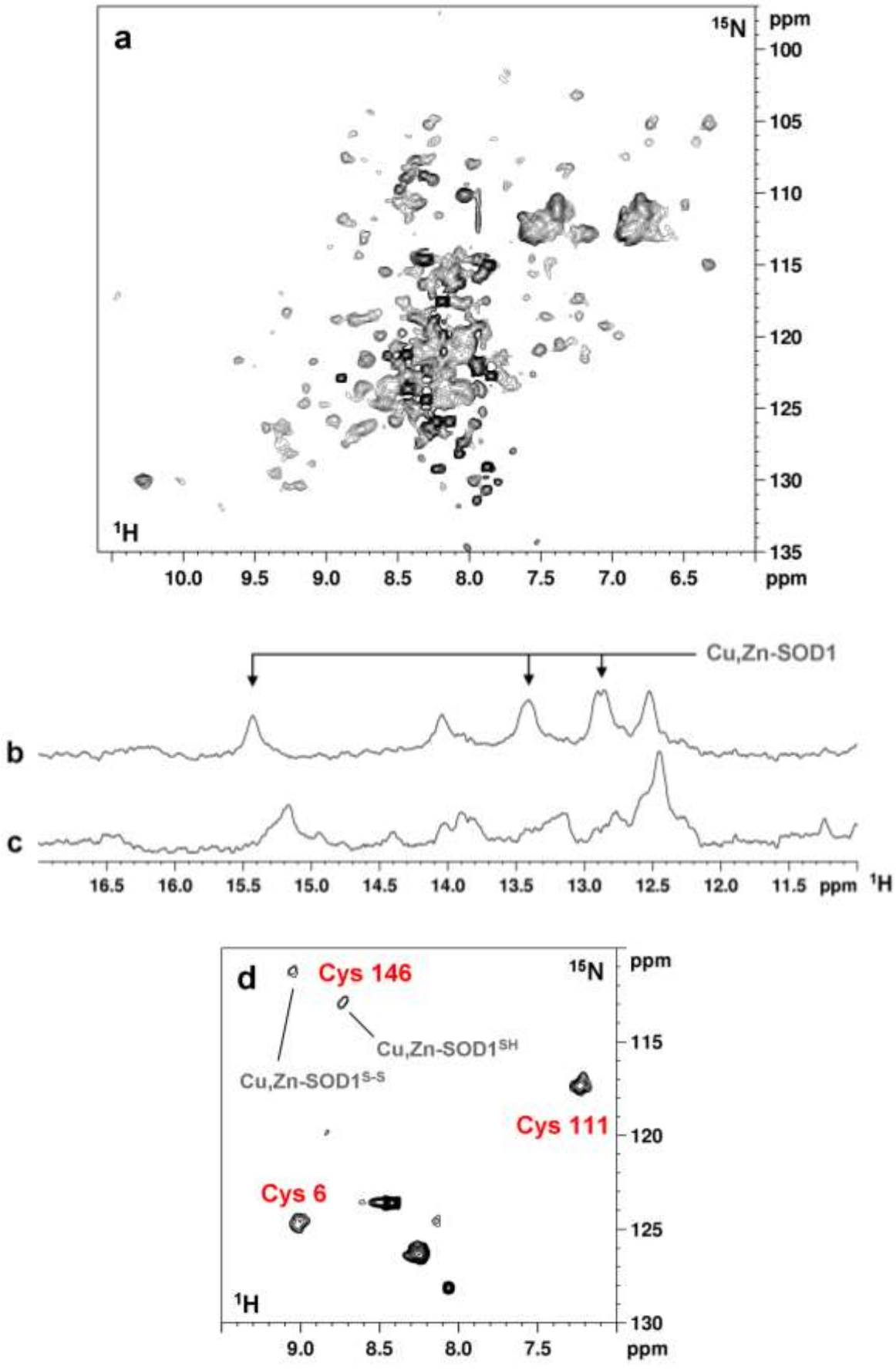

a
^{15}N
ppm
100
105
110
115
120
125
130
135
^{1}H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm


b
^{15}N
ppm
100
105
110
115
120
125
130
135
^{1}H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm

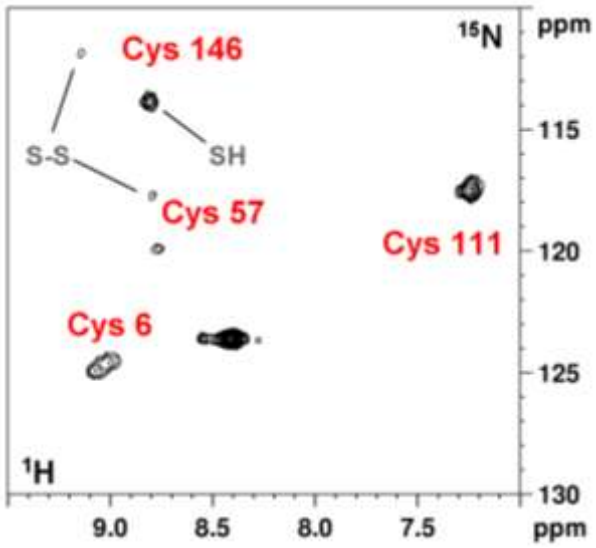
Supplementary Figure 5. Spectral crowding is reduced by subtracting the background signals arising from the non-selective labelling of cellular components. a, ^{1}H-^{15}N correlation spectrum of cytoplasmic E,Zn-SOD1^{SH} free of background signals obtained by subtracting to the ^{1}H-^{15}N SOFAST-HMQC spectrum shown in Figure 1c the spectrum **b**, of a cell sample transfected with the empty pHLsec vector, containing only signals from cellular components. The two spectra were acquired in the same experimental conditions.


in-cell SOD1 vs. *in vitro* E,Zn-SOD1^{SH}
in-cell SOD1 vs. *in vitro* Zn,Zn-SOD1^{SH}
CCSD (ppm)
0.3
0.25
0.2
0.15
0.1
0.05
0
Residue
A4 V5 C6 G10 D11 G12 V14 I17 I18 N19 F20 E21 Q22 K23 E24 S25 G27 V29 K30 G33 S34 I35 K36 G37 G41 H43 G44 F45 H48 E49 F50 D52 N53 T54 A55 G56 T58 G61 H63 K70 G73 K75 D76 R79 V81 G82 D83 L84 G85 N86 V87 T88 A89 D90 G93 V94 A95 D96 V97 S98 I99 E100 D101 S102 I104 S105 H110 C111 I112 G114 R115 H120 A123 D124 D125 L126 G127 K128 G129 G130 E132 T135 K136 T137 S142 A145 C146 G147 V148 G150 A152

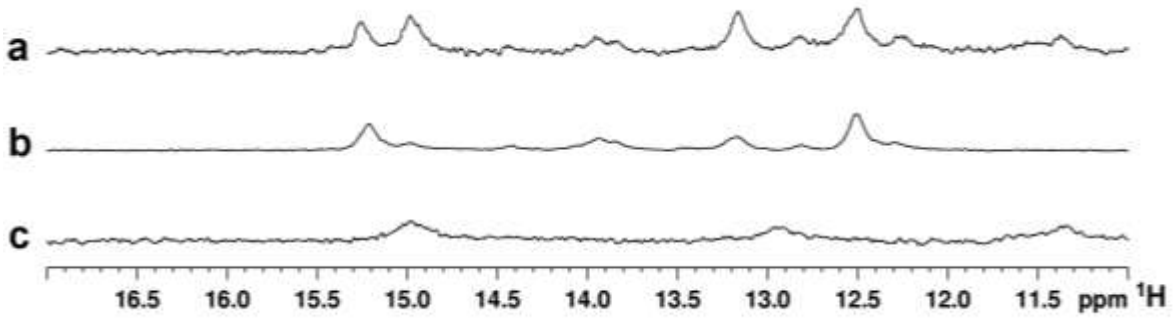
Supplementary Figure 6. Combined Chemical Shift Difference (CCSD) plot of in-cell SOD1 vs. in vitro E,Zn-SOD1 and Zn,Zn-SOD1. CCSD plot of a subset of ^{1}H-^{15}N resonances of SOD1 showing that the cytoplasmic zinc-containing SOD1 species corresponds to *in vitro* E,Zn-SOD1^{SH}. CCSDs between cytoplasmic SOD1 and Zn,Zn-SOD1^{SH} (blue) are higher on average than CCSDs between cytoplasmic SOD1 and E,Zn-SOD1^{SH} (orange). CCSDs were calculated using the formula: $CCSD = \sqrt{\frac{1}{2}\left(\Delta\delta^{1}H\right)^{2} + \frac{1}{2}\left(\Delta\delta^{15}N/5\right)^{2}}$.


a
^{15}N
ppm
100
105
110
115
120
125
130
135
^{1}H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm
Cu,Zn-SOD1
b
c
16.5
16.0
15.5
15.0
14.5
14.0
13.5
13.0
12.5
12.0
11.5
ppm
^{1}H
d
Cys 146
^{15}N
ppm
Cu,Zn-SOD1^{SH}
Cu,Zn-SOD1^{S-S}
115
Cys 111
120
Cys 6
125
^{1}H
130
9.0
8.5
8.0
7.5
ppm

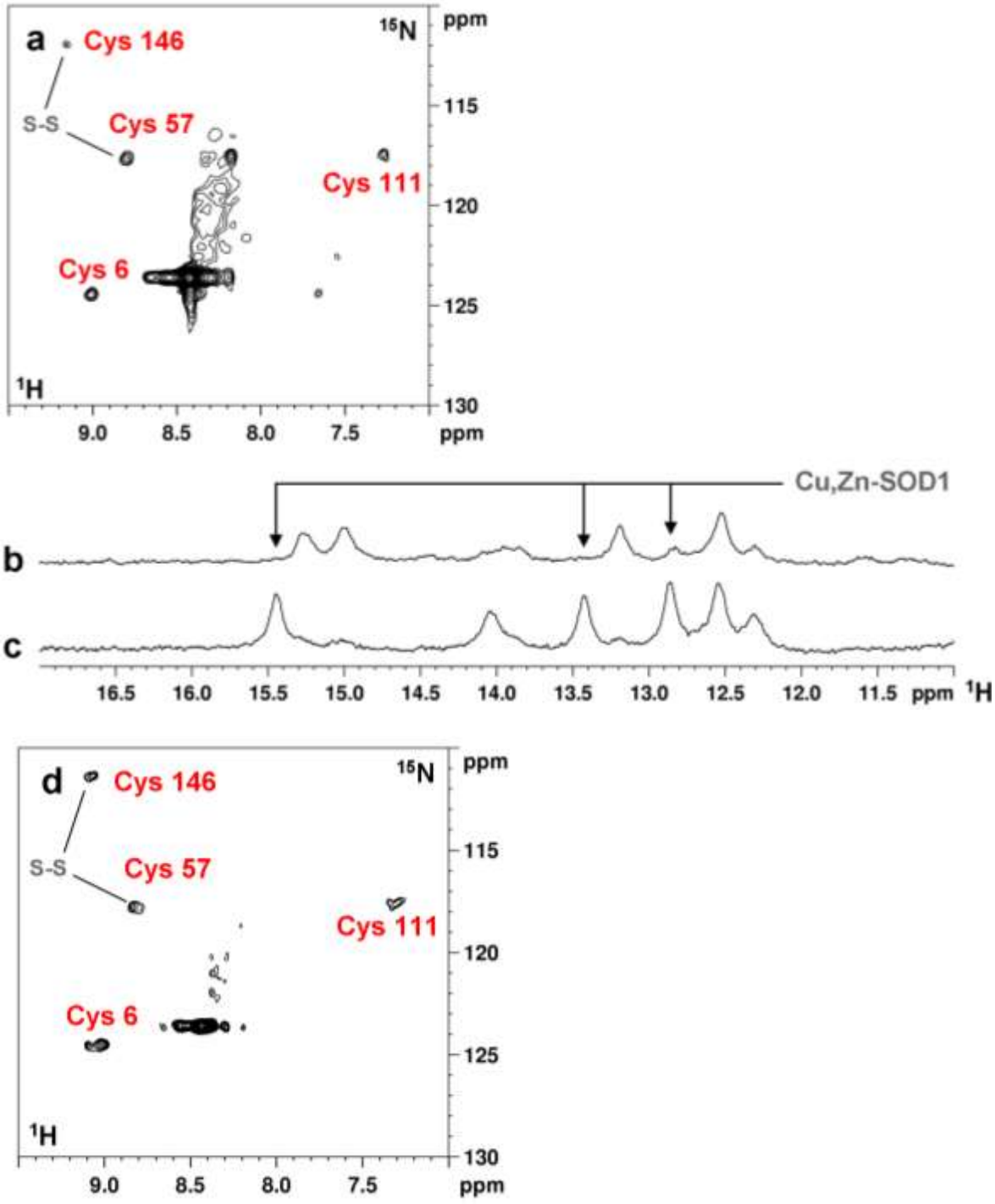
Supplementary Figure 7. Cu(II) addition to *E. coli* cells induces formation of Cu(I),Zn-SOD1 and partial disulfide oxidation. a, ^{1}H-^{15}N SOFAST HMQC spectrum of *E. coli* cells expressing uniformly ^{15}N-labelled SOD1 incubated with Cu(II); **b**, ^{1}H histidine NMR spectrum of *E. coli* cells expressing SOD1 incubated with Cu(II); **c**, NMR spectrum of *E. coli* cells expressing SOD1 incubated with Cu(I)-acetonitrile complex; **d**, ^{1}H-^{15}N SOFAST HMQC spectrum of *E. coli* cells expressing ^{15}N-cysteine labelled SOD1 incubated with Cu(II). Assigned cysteine residues in **d** are indicated in red. When two species of SOD1 are present, labels indicate the disulfide redox state of each species. Unlabelled crosspeaks are cellular background signals.



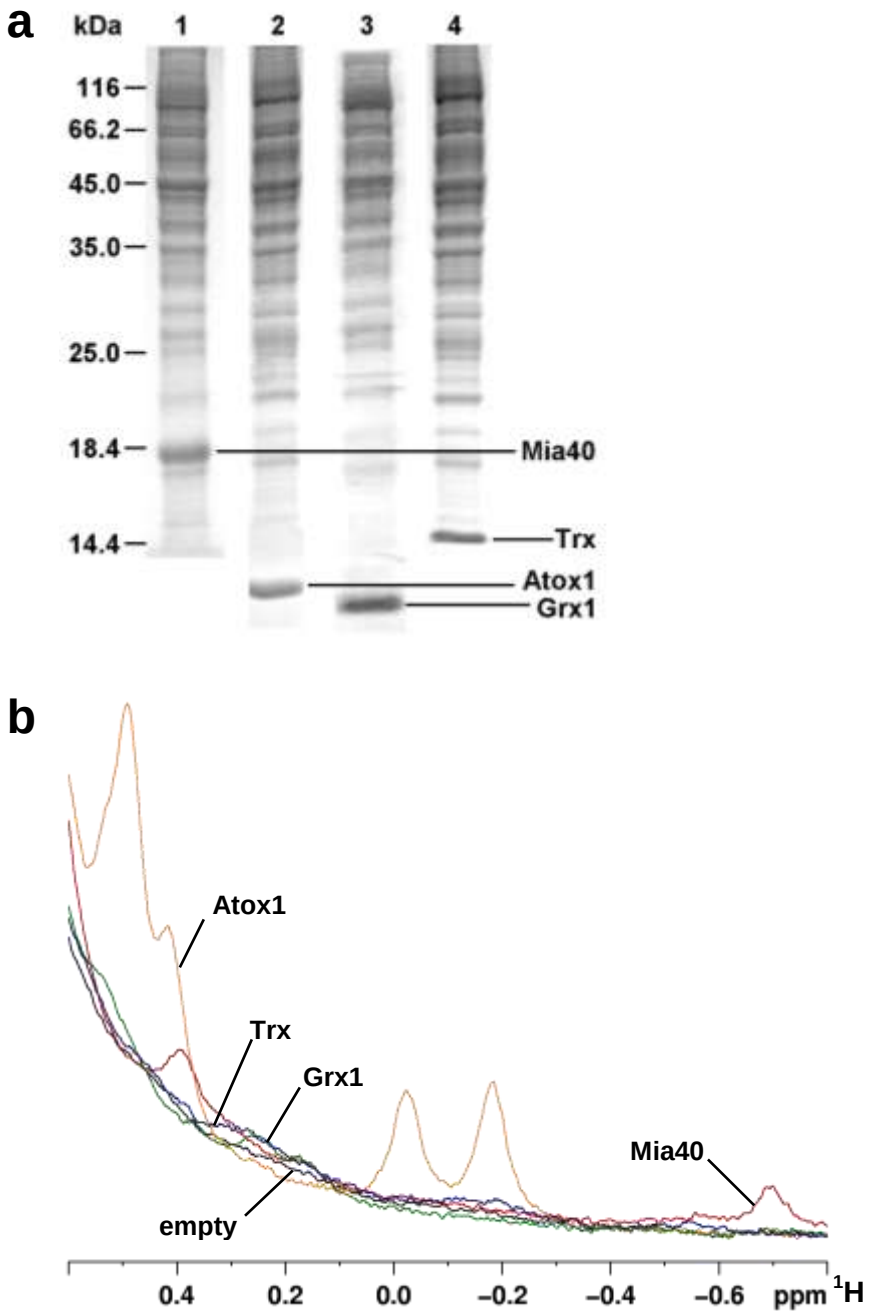
Supplementary Figure 8. The redox state of SOD1 is influenced by copper binding. ^{1}H-^{15}N SOFAST HMQC spectra were acquired on human cells expressing ^{15}N-cysteine labelled SOD1 in Zn(II)-supplemented medium, after incubation with Cu(II). Assigned cysteine residues are indicated in red. When two species of SOD1 are present, labels indicate the disulfide redox state of each species. Unlabelled crosspeaks are cellular background signals.



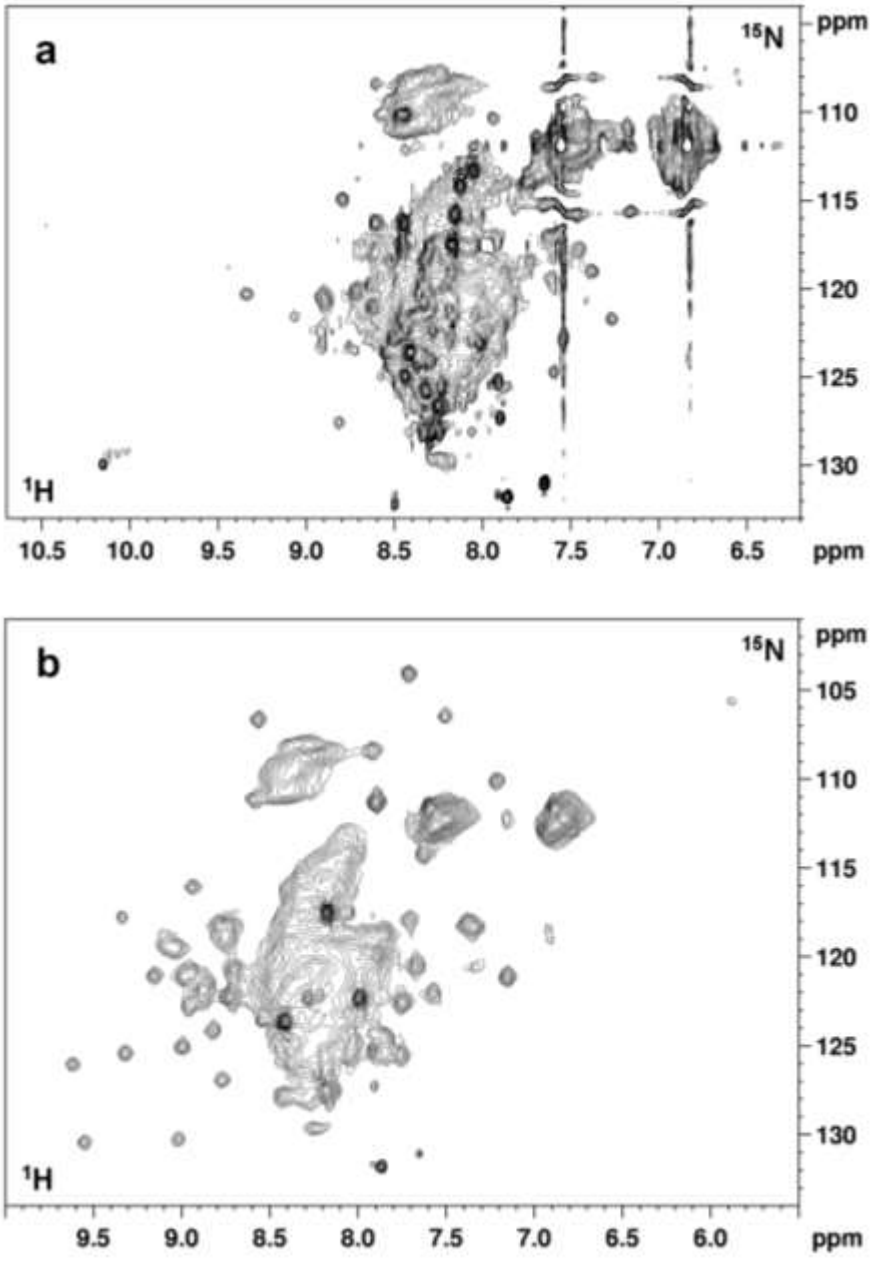
Supplementary Figure 9. Histidine region of ^{1}H NMR spectra acquired on human cells co-expressing SOD1 and CCS. a, cells co-expressing unlabelled SOD1 and CCS; **b**, expressing unlabelled SOD1 (scaled down to match the intensity of SOD1 peaks in **a**); **c**, expressing unlabelled CCS. Cell samples were transfected in Zn(II)-supplemented medium.

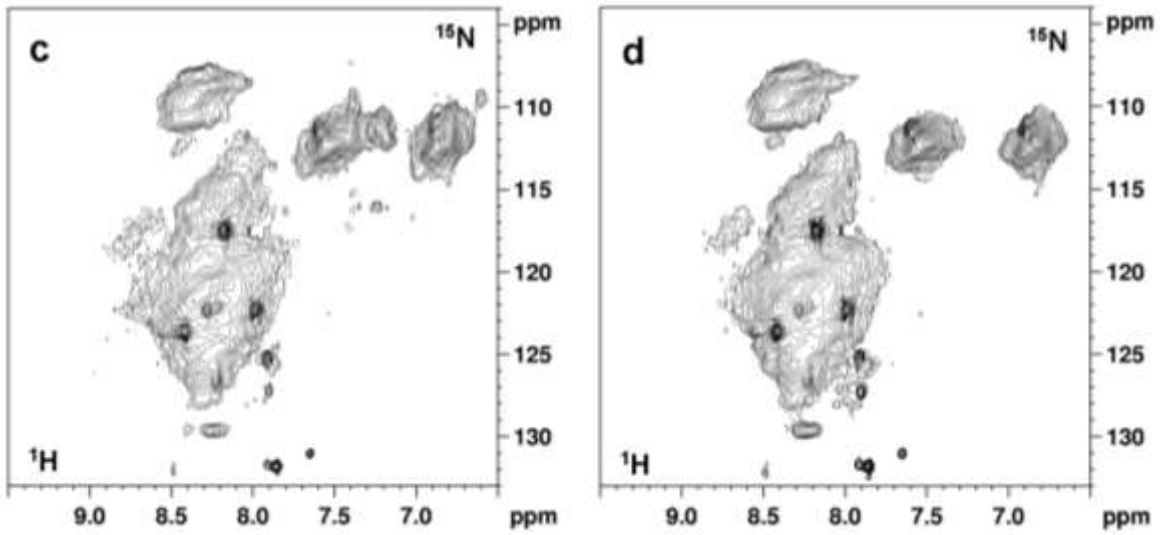

a
Cys 146
S-S
Cys 57
Cys 111
Cys 6
15N
ppm
115
120
125
130
1H
9.0
8.5
8.0
7.5
ppm
Cu,Zn-SOD1
b
c
16.5
16.0
15.5
15.0
14.5
14.0
13.5
13.0
12.5
12.0
11.5
ppm
1H
d
Cys 146
S-S
Cys 57
Cys 111
Cys 6
15N
ppm
115
120
125
130
1H
9.0
8.5
8.0
7.5
ppm

Supplementary Figure 10. When SOD1 is expressed at lower amounts, CCS catalyzes the complete formation of the disulfide bond, and incubation with copper results in the complete formation of Cu,Zn-SOD1. ^{1}H-^{15}N SOFAST HMQC spectra and ^{1}H NMR spectra were acquired on human cells expressing 45 ± 10 µM SOD1 and 15 ± 3 µM CCS : **a**, ^{1}H-^{15}N SOFAST HMQC acquired on cells co-expressing ^{15}N-cysteine labelled SOD1 and CCS in Zn(II)-supplemented medium; **b**, histidine region of ^{1}H NMR spectrum acquired on the same cells; **c**, histidine region of ^{1}H NMR spectrum acquired on cells co-expressing ^{15}N-cysteine labelled SOD1 and CCS in Zn(II)-supplemented medium, after incubation with Cu(II) **d**, ^{1}H-^{15}N SOFAST HMQC acquired on the same cells. Assigned cysteine residues are indicated in red. In both samples the cysteine crosspeaks of reduced SOD1 are not detected (**a**,**d**). Histidine protons unambiguously assigned to Cu(I),Zn-SOD1 species are indicated (**b**,**c**). Complete copper incorporation in SOD1 is observed. Protein concentration was estimated by comparing band intensities with serial dilutions of a pure SOD1 sample at known concentration, ran in the same gel.

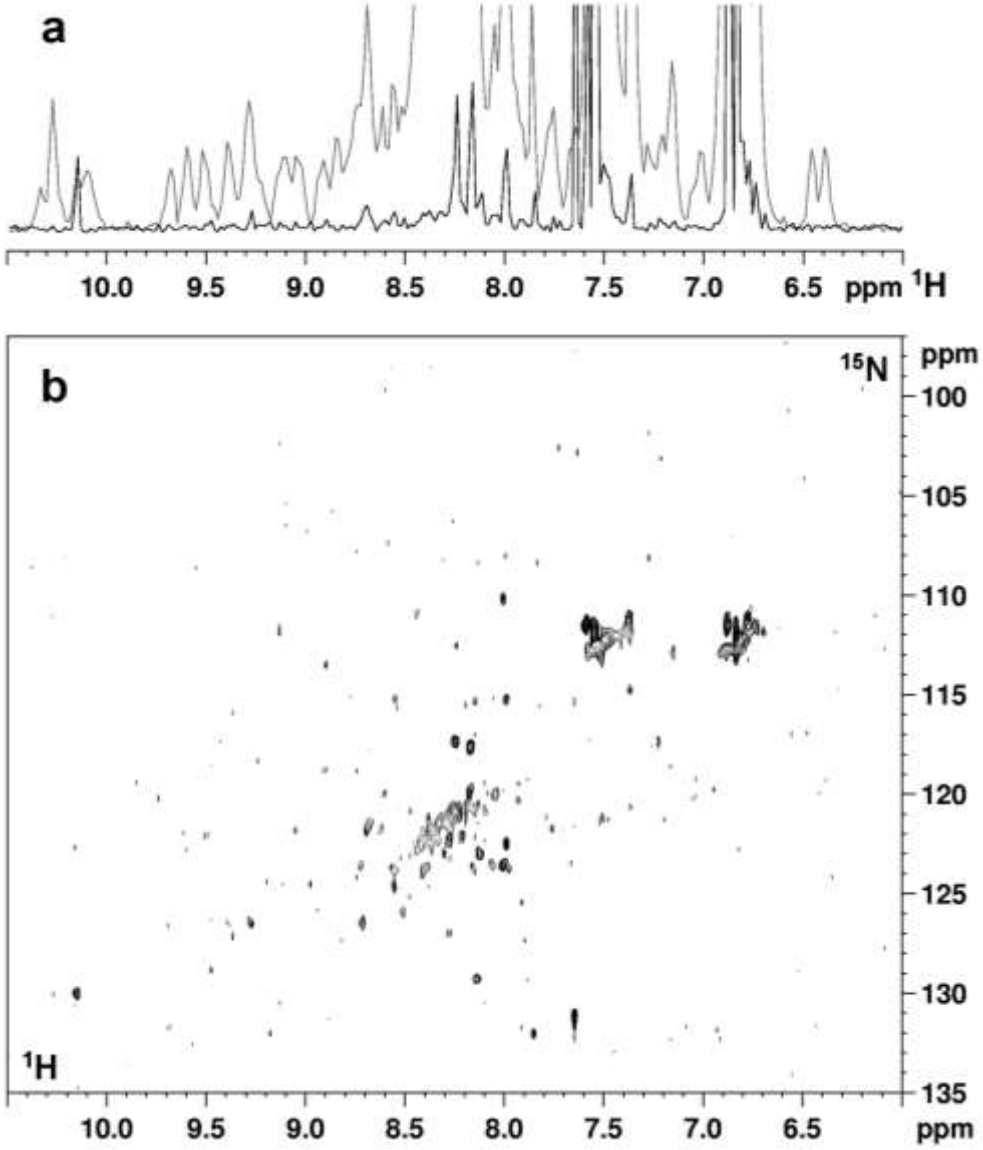


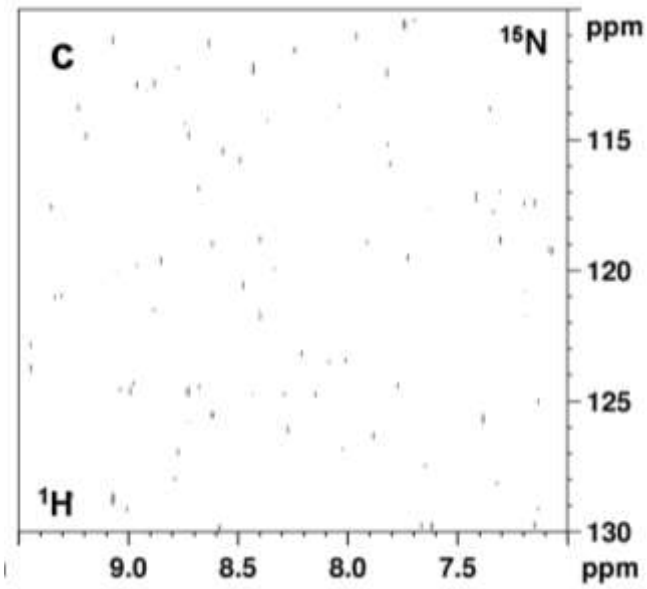
Supplementary Figure 11. Overexpression of non-hSOD1 protein targets in HEK293T cells. a, Coomassie-stained SDS-PAGE of total cell extracts showing the overexpression of Mia40 (lane 1), Atox1 (lane 2), glutaredoxin-1 (Grx1, lane 3) and thioredoxin (Trx, lane 4), 48 h after transfection (see Supplementary Fig. 14e-g); **b**, aliphatic region of ^{1}H NMR spectra of cells expressing Mia40 (red), Atox1 (yellow), Grx1 (green), Trx (blue) superimposed with cells transfected with empty vector (black).


a
^{15}N
ppm
110
115
120
125
130
^{1}H
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm
b
^{15}N
ppm
105
110
115
120
125
130
^{1}H
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
ppm

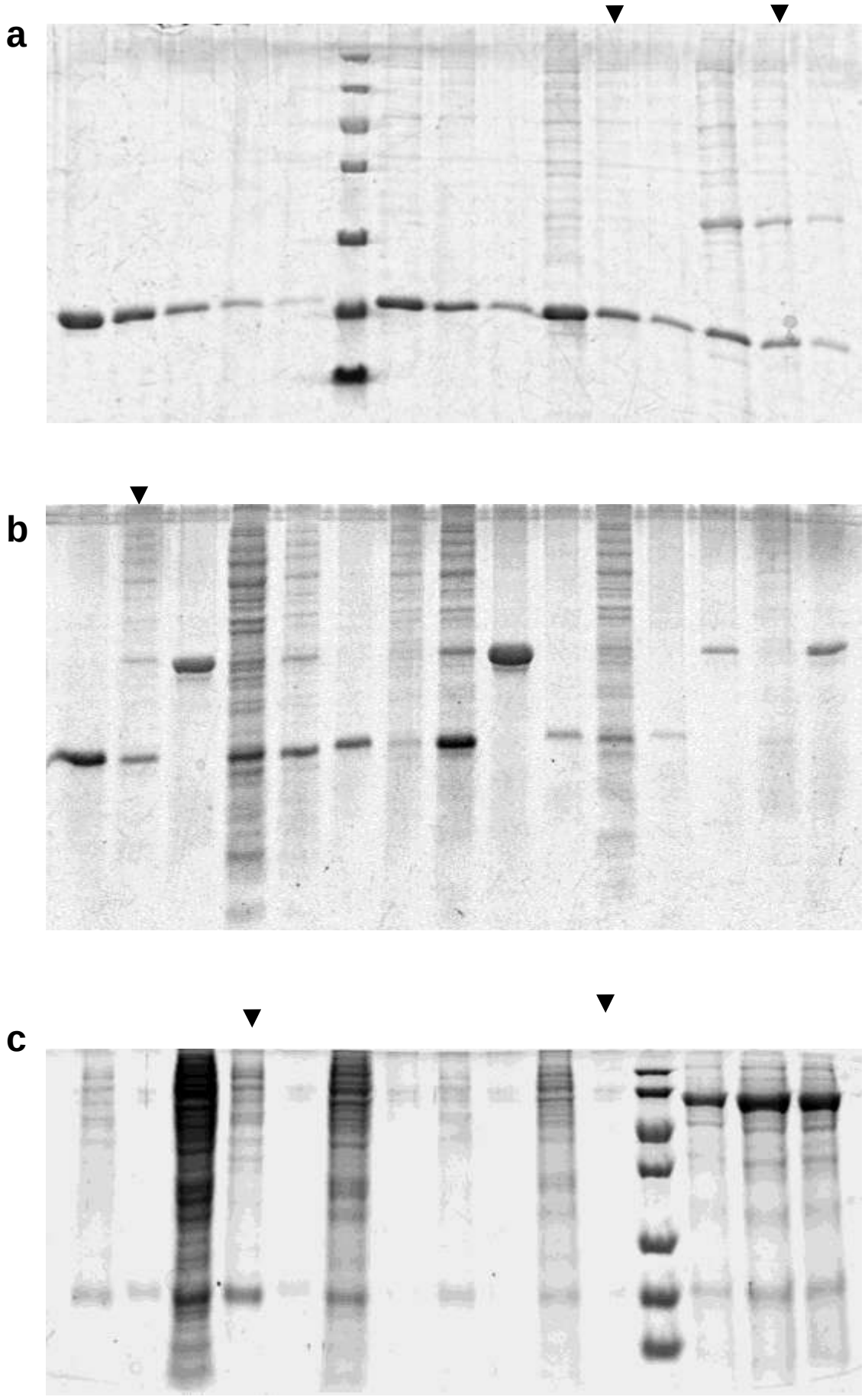


Supplementary Figure 12. In-cell NMR detection of non-hSOD1 protein targets. a, ^{1}H-^{15}N SOFAST HMQC spectrum of HEK293T cells expressing U-^{15}N labelled Mia40; **b**, ^{1}H-^{15}N SOFAST HMQC spectrum of HEK293T cells expressing U-^{15}N labelled Atox1; **c**, ^{1}H-^{15}N SOFAST HMQC spectrum of HEK293T cells expressing U-^{15}N labelled glutaredoxin-1; **d**, ^{1}H-^{15}N SOFAST HMQC spectrum of HEK293T cells expressing U-^{15}N labelled thioredoxin. Mia40 (**a**) and Atox1 (**b**) are clearly detected above the cellular background, while glutaredoxin-1 (**c**) and thioredoxin (**d**) are not detected due to slow tumbling in the cytoplasm.

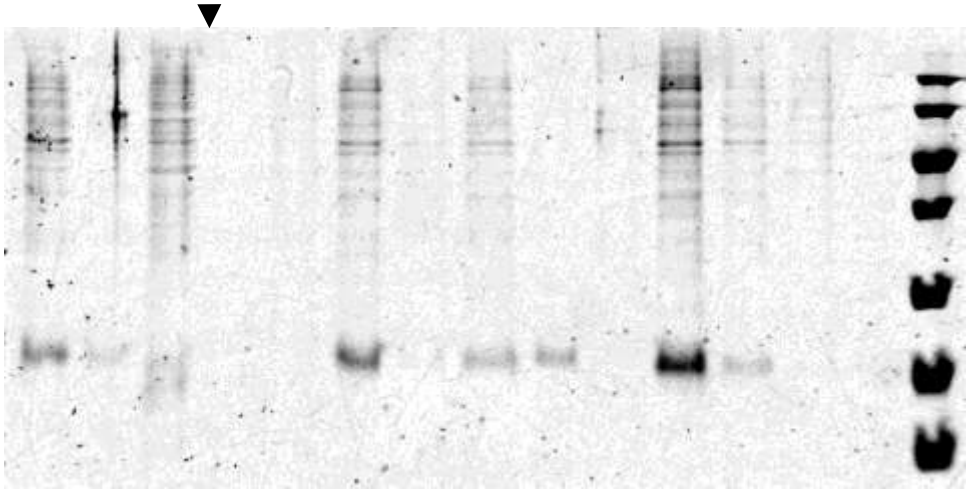

a
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm
1H
b
15N
ppm
100
105
110
115
120
125
130
135
1H
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
ppm



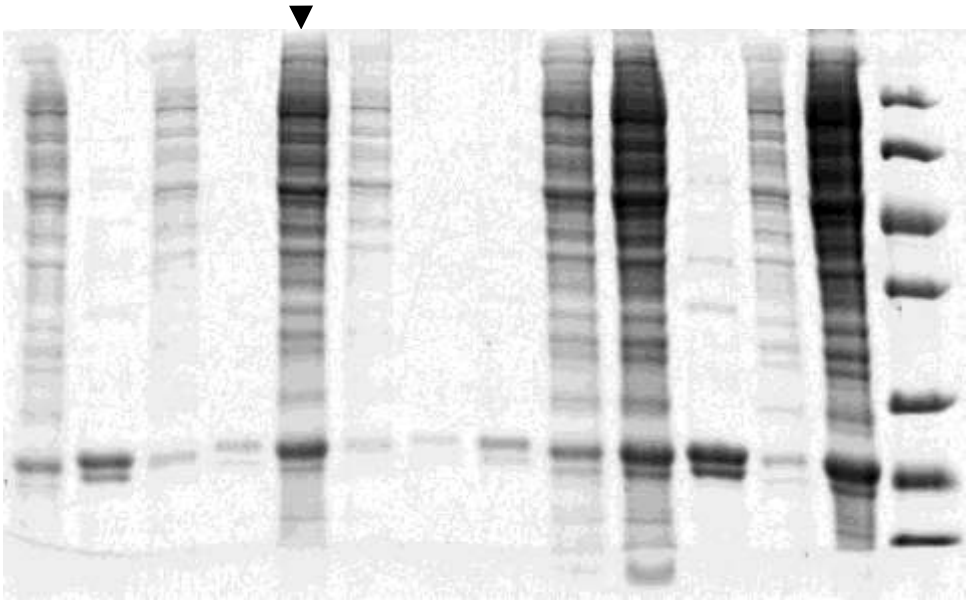
Supplementary Figure 13. NMR spectra of the cell sample supernatants confirm the absence of protein leaked out of the cells. ^{1}H-^{15}N SOFAST HMQC spectra were acquired on the supernatants obtained from each cell sample, after the actual in-cell NMR experiments, to check whether protein leakage has occurred. **a**, ^{1}H projection of the ^{1}H-^{15}N SOFAST HMQC supernatant spectrum (shown in **b**) of cells containing U-^{15}N E,Zn-SOD1 (black line) compared to the ^{1}H projection of the in-cell NMR spectrum (grey line); **b**, 2D ^{1}H-^{15}N SOFAST HMQC supernatant spectrum; **c**, 2D ^{1}H-^{15}N SOFAST HMQC supernatant spectrum of a cell sample containing ^{15}N-cysteine labelled SOD1 and CCS.


a
b
c

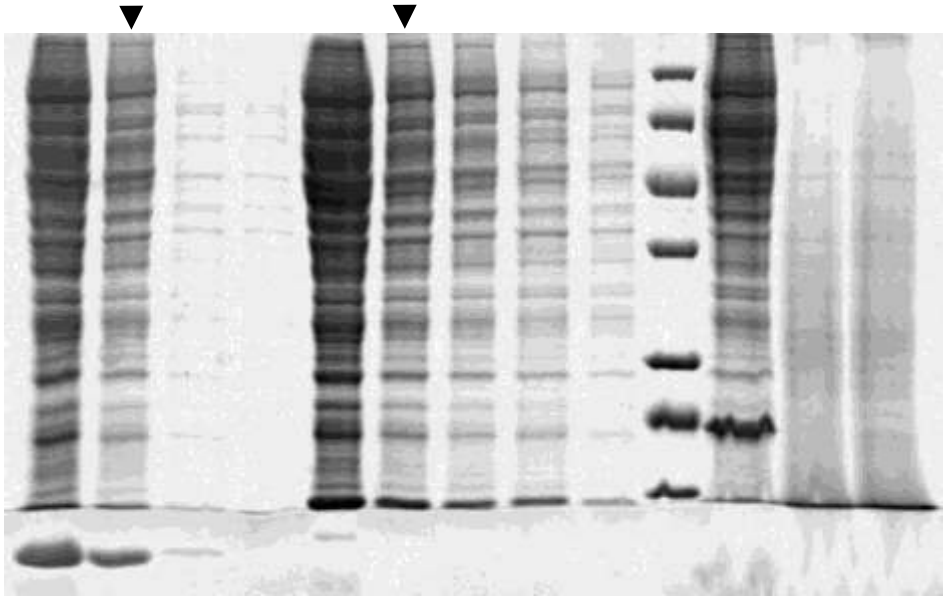
d

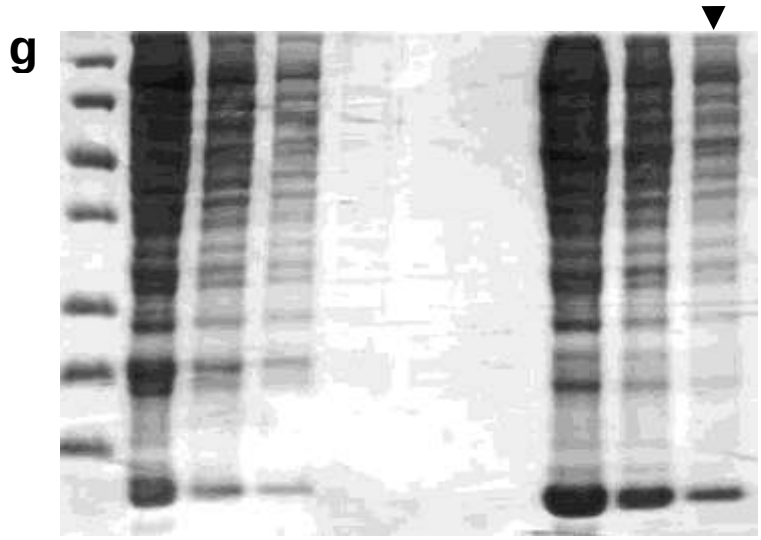


e



f





Supplementary Figure 14. Uncut SDS-PAGEs. a, SDS-PAGE of Supplementary Fig. 2 (lanes 1,2); **b**, SDS-PAGE of Supplementary Fig. 3 (lane 3); **c**, SDS-PAGE of Supplementary Fig. 3 (lanes C,M); **d**, SDS-PAGE of Supplementary Fig. 3 (lane N); **e**, SDS-PAGE of Supplementary Fig. 11 (lane 1); **f**, SDS-PAGE of Supplementary Fig. 11 (lanes 2,4); **g**, SDS-PAGE of Supplementary Fig. 11 (lane 3). Cut lanes are indicated with arrows.

2.3

In living cells folding and metal binding of SOD1 fALS mutants is impaired

Banci, L; Barbieri, L; Cantini, F; Kozyreva, T; Luchinat, E; Rubino, J

In preparation

In living cells folding and metal binding of SOD1 fALS mutants is impaired

Lucia Banci[1,2,*], Letizia Barbieri[1], Francesca Cantini[1,2], Tatiana Kozyreva[3], Enrico Luchinat[1,4], Jeffrey T. Rubino[1]

Affiliations:

[1]CERM, Magnetic Resonance Center, University of Florence, Via Luigi Sacconi 6, 50019, Sesto Fiorentino, Florence, Italy.

[2]Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019, Sesto Fiorentino, Florence, Italy.

[3]Giotto Biotech S.r.l., Via Madonna del Piano 6, 50019, Sesto Fiorentino, Florence, Italy.

[4]Department of Biomedical, Clinical and Experimental Sciences, University of Florence, Viale Morgagni 50, 50134, Florence, Italy.

* To whom correspondence should be addressed. L.B. (banci@cerm.unifi.it)

Introduction

An atomic level approach is required to characterize physiological processes and their alteration, in order to understand the molecular cause of their malfunction in pathological states and to develop treatments and therapeutic protocols. Ideally, these processes should be studied in the "natural" environment, i.e. within a live cellular context. For that purpose, in-cell NMR represents a promising technique (1, 2). We have recently developed an in-cell NMR approach aimed at obtaining information on the folding and maturation process of proteins directly in live human cells (3, 4). This in-human-cell NMR approach was applied to monitor maturation events for human proteins such as Mia40 (4) and superoxide dismutase 1 (SOD1) (3). The latter is a stable dimer binding a copper and a zinc ion per monomer, and forming an intramolecular disulfide bond. We showed that intracellular apoSOD1 selectively binds one zinc ion per monomer, and this step is required for the protein to interact with its specific chaperone CCS (5), which then transfers copper to SOD1 and catalyzes the disulfide bond formation (6-8).

SOD1 is strongly implicated in the onset of amyotrophic lateral sclerosis (ALS), an adult-onset neurodegenerative disease characterized by death of motor neurons in the brain and spinal cord. While the majority of ALS cases are sporadic (sALS) with no known etiology, the remaining cases are genetic in origin and classified as 'familial' (fALS). In 1993, it was discovered that about 20 % of fALS cases could be related to mutations in the SOD1 gene (9). 165 unique fALS-linked SOD1 mutations have been identified so far, which are scattered throughout SOD1 amino acid sequence (ALSod, the Amyotrophic Lateral Sclerosis Online Genetic Database, 2012, http://alsod.iop.kcl.ac.uk/). The pathogenic role of SOD1 in ALS has been linked to the formation of protein aggregates rich in SOD1, observed both in the spinal cords of patients with ALS (10-12) and of transgenic mice expressing human forms of the protein (13-16), with accumulation primarily in the late stage of the disease. The precursors of these protein aggregates are believed to be soluble oligomeric intermediates of the SOD1 aggregation process. These oligomers are thought to be responsible for the toxic gain of function, similar to what has been proposed for other neurodegenerative diseases (17-20). Generation of soluble oligomers was found to occur through oxidation of the two free cysteines of SOD1 (C6 and C111) (21), as well as through other possible mechanisms (22), which form amyloid-like structures (23, 24) among what is likely a number of structurally various aggregation products (25, 26).

The mature form of SOD1 is not prone to aggregation, whereas the intermediate species of the maturation process (i.e. SOD1 lacking metal ions) have been shown to have tendency to oligomerize *in vitro* (21, 27-29). Therefore, an impaired maturation process would lead to accumulation of immature SOD1 species, which are prone to oligomerization. Recent studies reported the occurrence of a higher propensity for deficient protein maturation *in vivo* or in cell culture models for some fALS mutants (30, 31). However, most of the studies addressing the properties of fALS SOD1 mutants and their effects on the

maturation process have been performed *in vitro*, i.e. far from physiological conditions (32-34).

In this work, the maturation process of a set of fALS mutants was studied through the in-human-cell NMR approach, already successfully applied to wild-type hSOD1 (3). With this technique, proteins are overexpressed and isotopically labeled in cultured human cells, and information on the intracellular folding, cysteine oxidation and metallation state of the protein, is obtained by observing them directly in living cells by NMR. The maturation levels of fALS mutants in different cellular conditions (e.g. different levels of copper and/or zinc and co-expression of human CCS) were compared to those of wild-type (WT) hSOD1. The CCS-dependent maturation of the same set of fALS mutants was also examined through *in vitro* NMR experiments following the same procedure applied to WT hSOD1 (8). These two approaches (in-cell and *in vitro* NMR) allowed us to monitor each step of the fALS mutants maturation process, providing atomic-level information on how each mutation affects the ability of intracellular hSOD1 to reach its mature, stable state.

Results

A number of fALS-linked SOD1 mutations distributed over the protein structure was investigated (**Figure 1**). We selected the mutants A4V, V7E, G37R, T54R, G85R, G93A, I113T and V148I, featuring various properties as summarized in Table 1. We also selected a mutation (G85R) that destabilizes the zinc-binding site (35), and a SOD1 mutation not currently linked to fALS (I35T), but which is reported to promote SOD1 aggregation (23) (Table 1). With the exception of G85R, all selected mutations occur far from the metal binding sites.

Table 1. Properties of the hSOD1 mutations studied in the present work.

Mutations	
A4V	Side chain size variation, this mutation appears to have a uniformly aggressive disease course with a mean of survival of 1 year after onset.
V7E	Dimer interface and charge variation to negative residue
I35T	Non-fALS mutation located on SOD1 β3 strand, the only one devoid of known fALS mutations; mutation from polar

	hydrophobic amino acid to non-polar hydrophilic amino acid.
G37R	Charge variation to positive residue
T54R	Dimer interface. Residue near the disulfide bond and charge variation to positive residue
G85R	Charge variation to positive residue. The mutation is located in the proximity of the zinc-binding site.
G93A	G93A mutation occurs in a region of hSOD1 in which hydrogen bonding interactions are thought to facilitate the "plug" end of the β-barrel.
I113T	Dimer interface, change to hydrophilic amino acid
V148I	Dimer interface.

Impaired zinc binding causes mutant apo-SOD1 to accumulate in the cytoplasm

SOD1 mutants were overexpressed in HEK293T cells, transfecting them with the same amount of DNA for all the mutants, and the expression level of each mutant was compared to that of WT SOD1 in the same conditions. The amounts of soluble mutant proteins in the cell extracts ranged from 36 ± 8 μM (G85R SOD1) to 162 ± 8 μM (T54R SOD1), compared to 150 μM WT SOD1. For each mutant the total amount of intracellular SOD1 was the same as that of the soluble SOD1, indicating that no significant formation of insoluble aggregates occurred to any of the mutants in these cellular conditions (**Figure S1**). In absence of added metal ions, all mutants appear to be in a predominately unfolded apo state, similar to what was observed for WT SOD1 expressed in the same conditions, where the in-cell NMR spectra show mainly peaks in the 8.0-8.3 ppm (^{1}H) region (**Figure S2**). Zinc binding was investigated by inducing protein expression in cells supplemented with $ZnSO_4$. Even with excess zinc, most of the fALS mutants were observed in predominately unfolded apo forms (**Figure 2**). For A4V, I35T, G85R, no zinc forms were observed at all, while for G37R, G93A, and I113T only a small intracellular amount of zinc containing protein (E,Zn-SOD1) was observed (**Figure S3**). We have previously described the folded E,Zn form of WT SOD1 when the protein is expressed in the presence of zinc (3). In the current study, a similar behavior was observed only for V7E, T54R and V148I SOD1 mutants: they stoichiometrically bound one Zn^{2+} ion per monomer, and only the folded, homodimeric E,Zn-SOD1 forms were present (**Figure 3**). The absence or low amounts of E,Zn-SOD1 observed in fALS-linked mutants suggests a possible detrimental effect of these mutations for zinc binding. To more thoroughly characterize the ability for these mutants to bind zinc, they were also expressed and purified from *E. coli* for analysis

in vitro (**Figure S4**). E,Zn forms of the G37R, G93A, and I113T were obtained with only minimal amounts of the predominately unfolded apo state observed for G93A and G37R. The E,Zn form of A4V was also observed, however this mutant appeared to mostly be in the predominately unfolded apo state. No metallated form of G85R was obtained and was also predominately unfolded, however some dispersed peaks suggested a more structured apo form was also present. In preparation for titration with Zn^{2+} G37R, G93A, I113T, and AV4 were demetallated and subsequently reduced to their biologically relevant apo-reduced forms (apo-$SOD1^{SH}$) to directly observe zinc binding. In each case, signals for both the predominately unfolded apo and a more structured apo froms were observed, however in different ratios, with I113T appearing to have the highest ratio of the more structured apo form (Figure S5). Subsequent titration with Zn^{2+} revealed that only the more structured apo form has the ability to bind zinc, as only chemical shifts associated with the more structured form were perturbed upon addition of the metal while the chemical shifts associated with the predominately unfolded form remained unchanged (Figure S5). G85R was also reduced and subsequently titrated with Zn^{2+}, and although was initially isolated in the apo form, was also observed to bind zinc (Figure S5).

Using both in-cell and *in vitro* NMR techniques we identified a predominately unfolded form of SOD1 that is unable to bind zinc. To further analyze this species we conducted analytical gel filtration experiments on several *in vitro* samples of fALS mutants (Figure S6). A4V, G37R, G85R, and G93A all appeared to have peaks associated with what has previously been identified as soluble oligomers (21, 23). Consistently, all of these mutants showed considerably high ratios of the predominately unfolded to folded species both in-cell and *in vitro* by NMR, in contrast to T54R, which appeared completely folded by NMR, and did not exhibit peaks associated with soluble oligomers in the analytical gel filtration experiment.

CCS rescues the impaired zinc-binding step of SOD1 mutants

The effect of hCCS in the maturation of SOD1 mutants in living cells was assessed by inducing simultaneous expression of each SOD1 mutant with hCCS in either U-^{15}N-labeled or [^{15}N]Cys-labeled medium supplemented with zinc and copper, and by analyzing the protein metallation and cysteine redox state. With the exception of G85R mutant, the fully mature form (Cu(I),Zn-$SOD1^{SS}$) was detected for each mutant (**Figures 4, 5**), while no unmetallated forms were detected (**Figure S7**), indicating that the SOD1 mutants had reached the mature state when both hCCS and copper were available in the cytoplasm in sufficient amount; a fraction of E,Zn-$SOD1^{SS}$ was also present for some mutants (namely V7E, G37R, T54R, G93A, I113T, V148I) as well as for WT SOD1 (**Figure 4**). Notably, for the mutants where the soluble oligomer was detected, the total amount of zinc-containing protein (i.e. the sum of E,Zn-SOD1 and Cu(I),Zn-SOD1 species) was higher in these cellular conditions than that measured in zinc-supplemented cells with only endogenous levels of CCS. Fully mature SOD1 was not observed when mutants were

coexpressed with hCCS in the absence of supplemented copper (**Figure S8**). In these conditions a mixture of species was generated, which included apo-SOD1^{SH}, E,Zn-SOD1^{SH}, and E,Zn-SOD1^{SS}, indicating that copper-free hCCS did fully restore the zinc-binding step of SOD1 mutants. Together, these results demonstrate that copper is required with the coexpression of hCCS to promote mutant SOD1 maturation and prevent the formation of the soluble oligomeric species.

In the in-cell NMR experiments, SOD1 and hCCS were co-expressed in the presence of copper and oxygen, therefore the two steps of CCS-dependent maturation – copper transfer and disulfide bond oxidation – could not be separated. In order to examine these steps separately, and determine if fALS mutations affect the kinetics of either process, we characterized the interaction of Cu(I)-hCCS with WT and mutants forms of E,Zn-hSOD1^{SH} by *in vitro* NMR experiments. Incubation of WT and fALS mutants V7E, G37R, T54R, G85R, G93A, I113T and V148I with Cu(I)-hCCS anaerobically resulted in the formation of the Cu,Zn-disulfide reduced forms of hSOD1**,** with 100% copper transfer for all the fALS mutants, and an apparent rate similar to that of the WT protein (**Figure 6)**. Once fully copper loaded, exposure to air produced the fully mature Cu,Zn-hSOD1^{SS} form, with rates comparable to that of WT SOD1 (**Figure 7**). Only the T54R mutant protein resulted in a considerably slower oxidation of the C57-C146 disulfide bond, although its oxidation was faster when copper was transferred by CCS than when Cu(I)-acetonitrile was used as the source of copper ,in the absence of CCS. For mutants where the soluble oligomeric state was observed together with E,Zn-hSOD1^{SH} (G37R, G85R, and G93A) signals associated with the soluble oligomer were unaffected by titration with Cu(I)hCCS (for example G85R, Figure S9).

Discussion

We analyzed the maturation process of a set of fALS-linked SOD1 mutants in their physiological environment, by comparing the distribution of species produced by the cells in different metal abundance, with respect to that of WT SOD1 in the same conditions, and we investigated the role of CCS in promoting mutant SOD1 maturation. Many of the SOD1 mutants analyzed here appeared to form soluble oligomers when overexpressed in cells with only endogenous levels of hCCS, or when coexpressed with hCCS in the absence of copper. Conversely, the soluble oligomers were not observed for the same mutants when hCCS was coexpressed with the mutant SOD1 in the presence of copper, where, in fact, only mature SOD1 forms were detected. This demonstrates that the soluble oligomeric species do not form when the SOD1 CCS-dependent maturation mechanism is intact and can readily convert immature E,Zn-hSOD1 forms to fully mature protein. Under the current conditions, the absence of Cu(I)-hCCS results in the accumulation of monomeric species, which, particularly for the more unstable fALS mutant proteins, likely results in unfolding, as predicted by the proposed three-state folding model of SOD1 (2U

⇆ 2M ⇆ N_2). In this model the properly folded monomer (M) is in equilibrium with an unfolded monomeric state (U) in addition to the native dimer (N_2) (36, 37). While the properly folded monomer readily binds zinc, the unfolded monomer has a significantly reduced affinity for the metal (38-40). ALS mutations destabilize the folded monomer and shift the equilibrium to favor the unfolded state (41-43), which has been suggested to be the starting material for the potentially toxic oligomeric species (33, 41), observed here for the first time in the living cellular environment by NMR.

Identification of the in-cell soluble oligomers was aided by use of *in vitro* experiments using the same fALS mutants. HSQC spectra for the same mutants using both techniques reveal strikingly similar peak patterns, suggesting they are in fact the same species. Moreover, both seem to be equally unaffected by the presence of excess zinc: supplementation of zinc in the media of cells expressing fALS mutants with only endogenous hCCS, in the case of in-cell NMR experiments, did not appear to affect the peak signals associated with the soluble oligomers, nor, in the case of the *in vitro* experiments, did direct titration of zinc into fALS mutant samples affect the signals associated with the soluble oligomers. Together these observations highly suggest that the predominately unfolded states observed in fALS mutant samples by both in-cell and *in vitro* NMR are in fact the same species. Furthermore, based on the results of analytical gel filtration on *in vitro* apo-SOD1 samples, we have concluded that the predominately unfolded species is most likely in a soluble oligomeric state.

In experiments with CCS *in vitro*, with the exception of T54R, the mutations examined did not inhibit the interaction of E,Zn-SOD1^{SH} with Cu(I)CCS, and the observed rates of copper transfer and disulfide formation for these mutants were similar to those of the WT protein. This behavior suggests that the mutations selected did not significantly alter the dimer interface, despite their relative proximity, as the mutants were still able to interact with hCCS and proceed with normal CCS-dependent maturation. However, the T54R mutation may result in the loss of a potentially important intermolecular hydrogen bond and introduces a potentially destabilizing positive charge in the vicinity of the disulfide bond, which may explain the reduced rate of disulfide formation (see Figure S10 for a detailed analysis).

As previously mentioned, the G85R mutation is located near and destabilizes the zinc binding site. In fact, only unmetallated forms of G85R were observed in the in-cell NMR samples and when the mutant protein was expressed in *E. coli* and purified for *in vitro* experiments. In the in-cell NMR experiment where the G85R mutant was coexpressed with hCCS in the presence of copper, no mature protein was detected. This is in agreement with previous *in vitro* work where we also observed that WT apo-SOD1 is unable to interact with hCCS, and that zinc acquisition is required before CCS-dependent maturation can occur (8). In the current work, with high enough zinc concentrations *in vitro*, we were able to metallate the apo-reduced form of G85R, and monitor its progression through CCS-dependent maturation (Figure S9).

Figures:

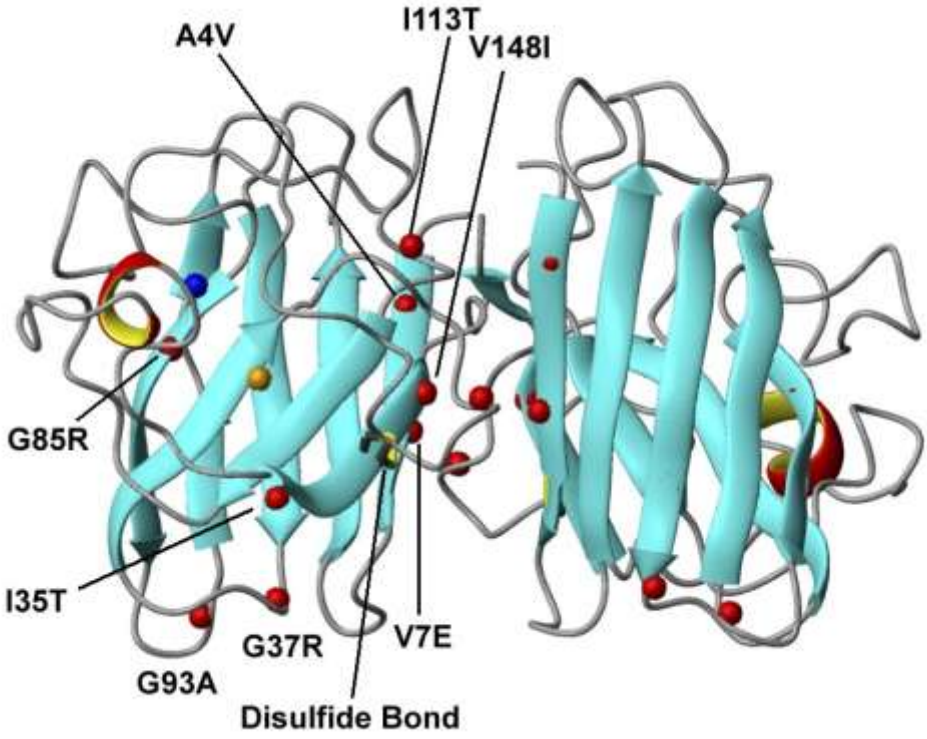


Figure 1: Location of selected fALS mutations of SOD1 (PDB code: 1L3N) examined, copper and zinc are shown as yellow and blue spheres, respectively. The side chains of Cys57 and Cys146 are also shown as yellow sticks.

.

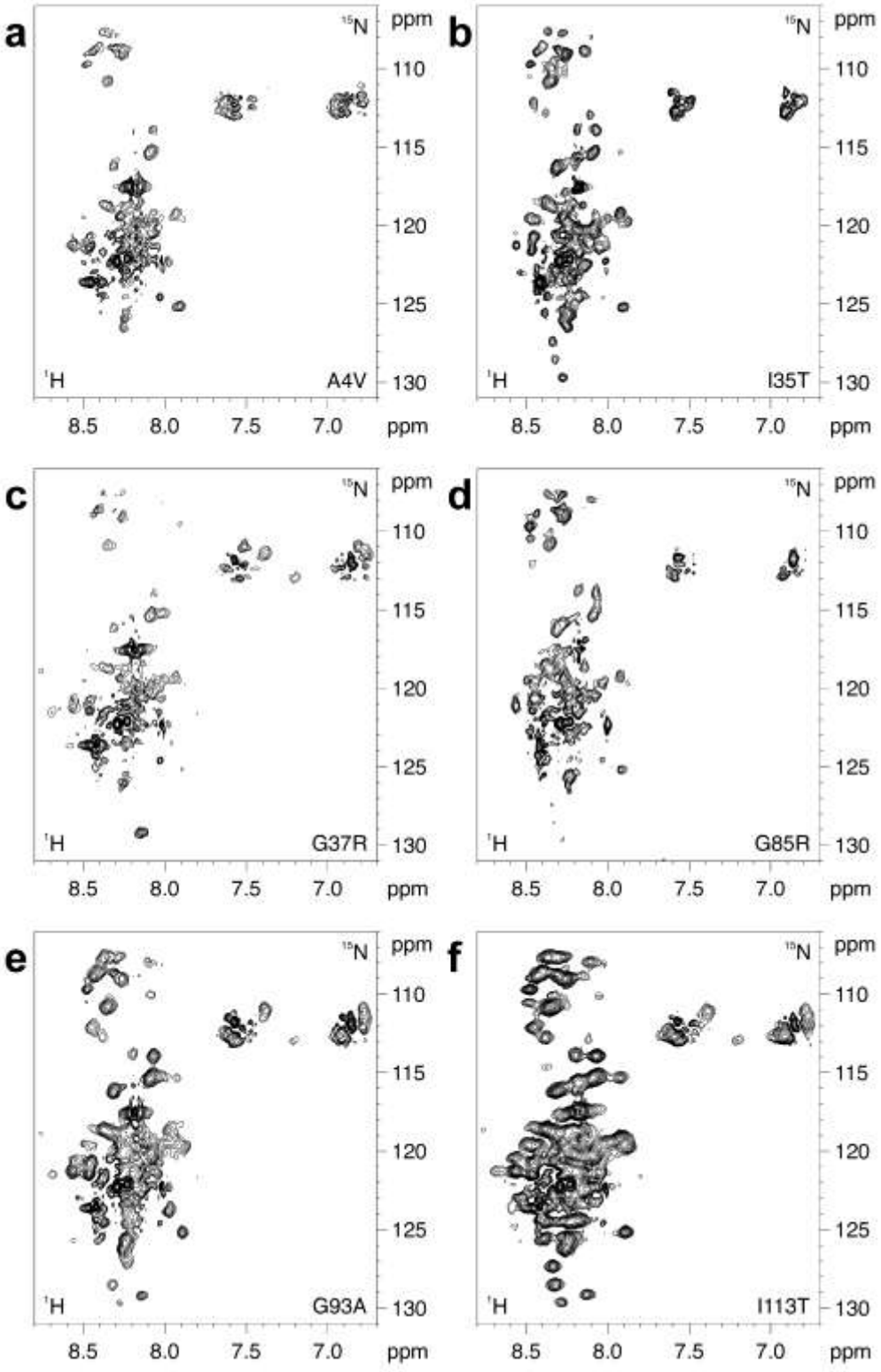


Figure 2: A predominately unfolded apo form was observed for some fALS SOD1 mutants, even with excess zinc in the cell culture. ^{1}H-^{15}N SOFAST HMQC spectra acquired on human cell samples expressing uniformly ^{15}N-labeled SOD1 mutants in Zn(II)-supplemented medium.

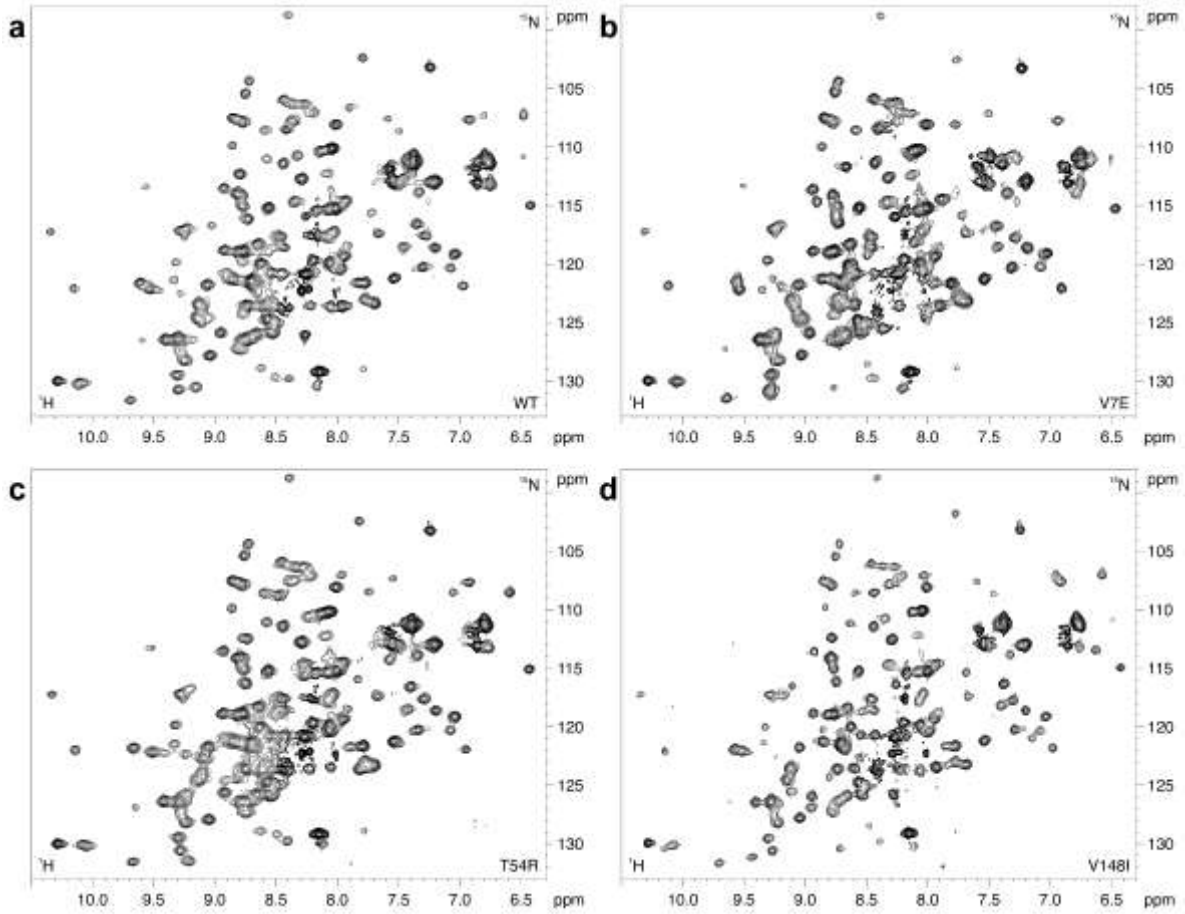


Figure 3: Some SOD1 mutants behaved like WT SOD1, and stoichiometrically bound one Zn^{2+} ion per monomer, when present in excess in the cell culture. ^{1}H-^{15}N SOFAST HMQC spectra acquired on human cell samples expressing uniformly ^{15}N-labeled WT SOD1 and mutants in Zn(II)-supplemented medium.

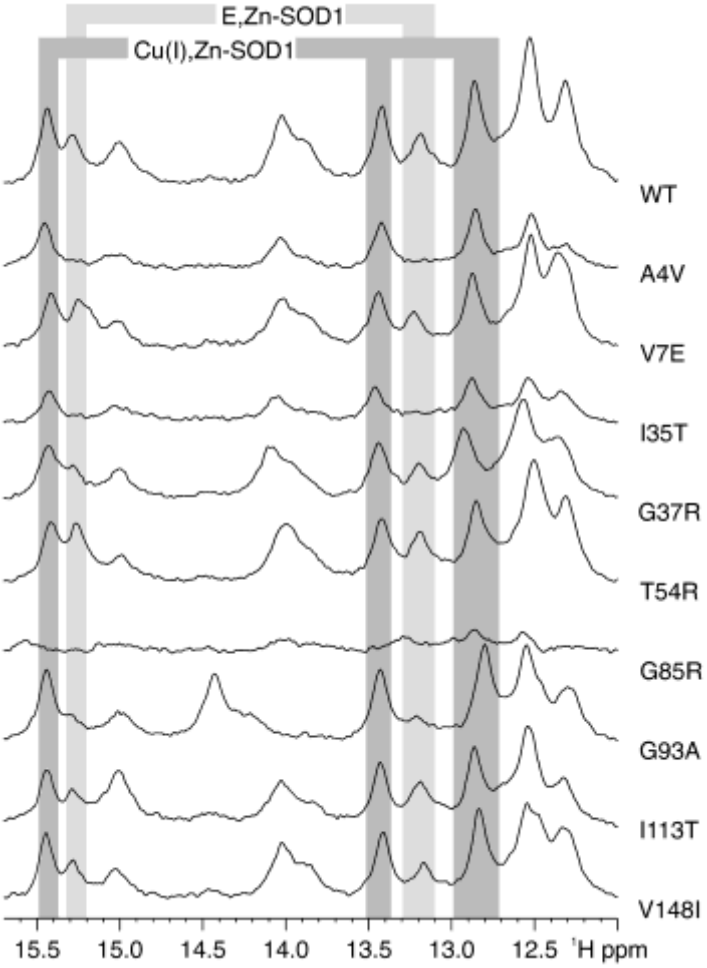


Figure 4: In excess of zinc and copper, co-expressed CCS allowed the formation of Cu,Zn-SOD1 for all mutants, with the exception of G85R. A small amount of E,Zn-SOD1 was still present for some mutants.

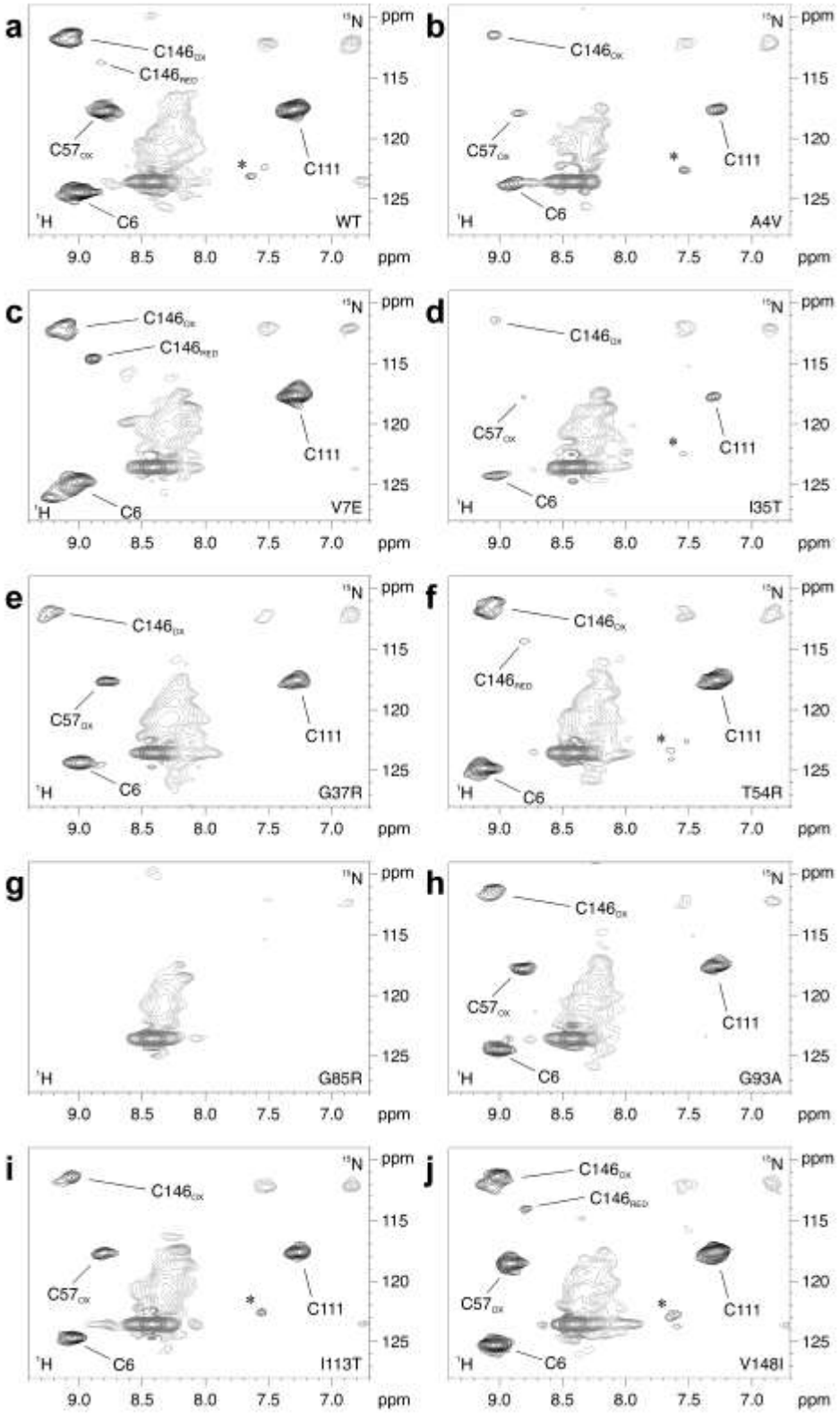


Figure 5: Co-expression of CCS in presence of both zinc and copper catalyzed the formation of the mutant SOD1 intramolecular disulfide bond.

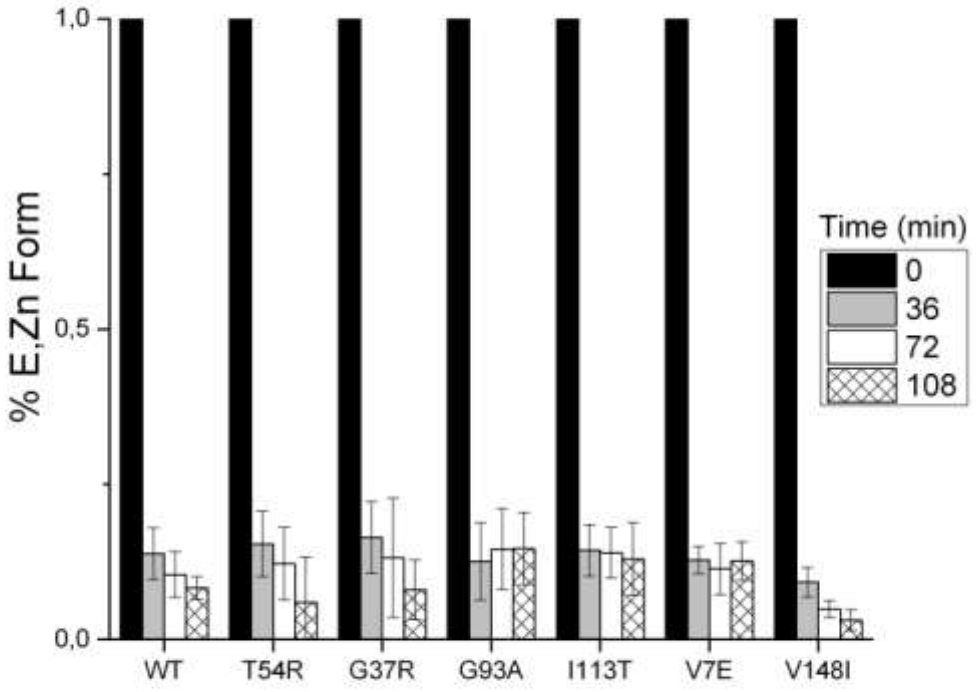


Figure 6: % E,Zn hSOD1 form before (0) and 36, 72, and 108 minutes after addition of Cu(I)hCCS. Conversion of E,Zn-to-Cu,Zn form was followed by monitoring the amide signals for residues G73, N139, N140, D124, and L126, as they are a located in proximity to the catalytic metal-binding site. Bewteen 3-5 of these residues were used for each experiment. The intenisty of the peaks assocaited with the E,Zn-hSOD1 form for were used to determine percent of the E,Zn form, and then averaged. Error bars indicat one standard deviation from the mean. (Data for G85R mutant not included, see Figure S9).

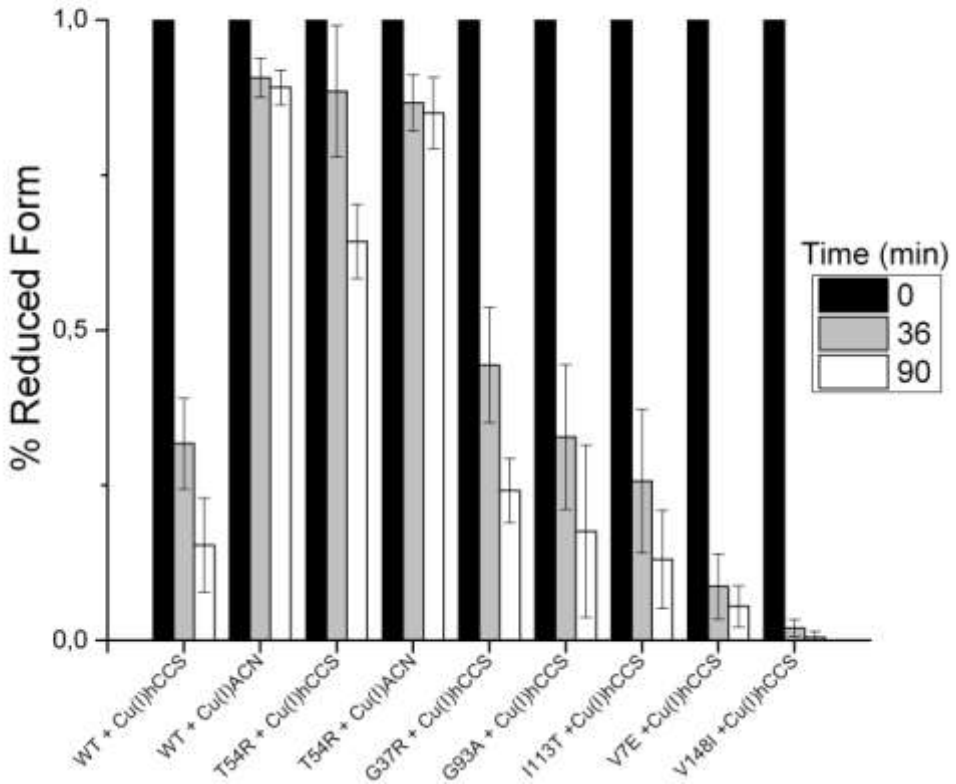


Figure 7: % Cu,Zn-hSOD1 reduced form before (0) and 36, and 90 minutes after exposure to air for 10 minutes. Copper forms were generated by addition of Cu(I)hCCS or Cu(I)acetonitrile, Cu(I)ACN, as labeled. Conversion of the reduced to the oxidized form of the protein was followed by monitoring the amide signals for G141, G61, and C146; G141 and G61 are located near the disulfide bond formed between C57 and C146. The percent of the reduced form of the protein was determined by the ratio of the intensity of the peaks associated with the reduced and oxidized form, for each of the 3 residues, and averaged. Error bars indicate one standard deviation from the mean. (Data for G85R mutant not included, see Figure S9).

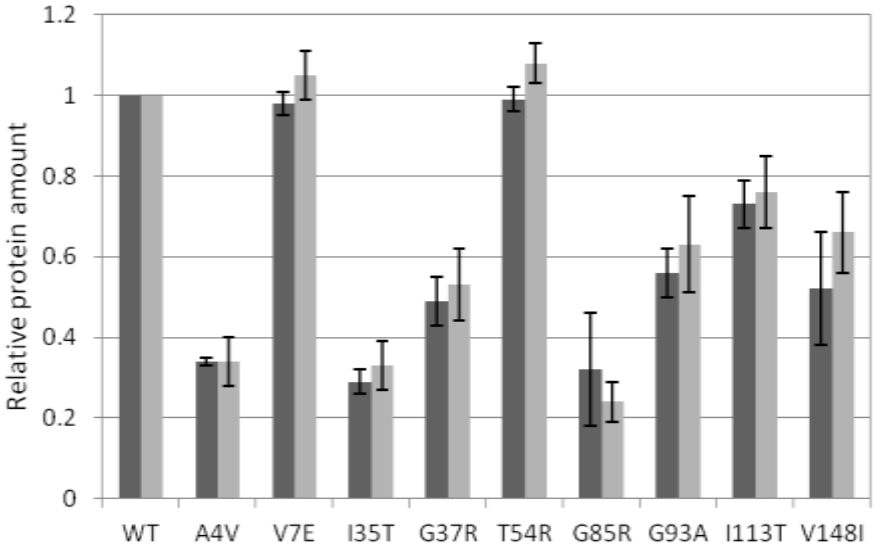


Figure S1: Amount of overexpressed SOD1 mutants in whole cells (dark gray) and in cell extracts (light gray) from cells grown in zinc-supplemented medium, relative to WT SOD1 overexpressed in the same conditions.

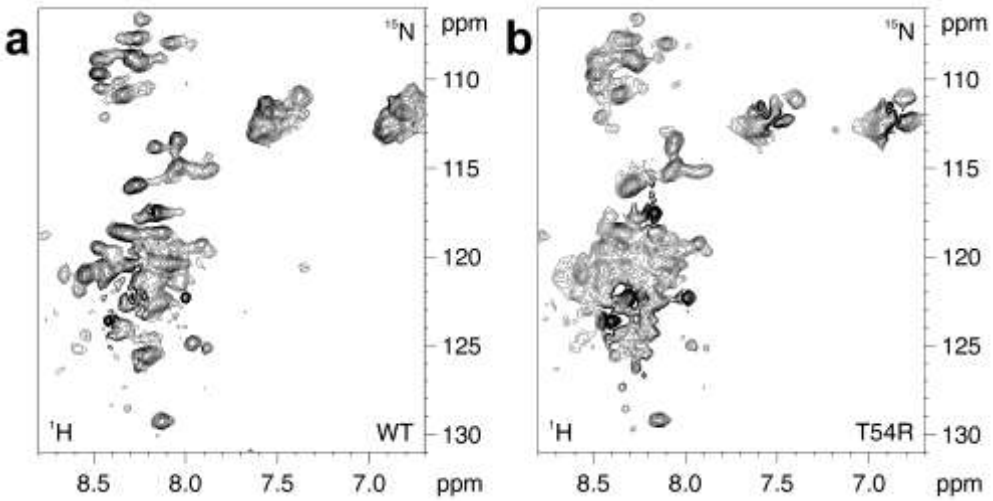


Figure S2: In absence of added metal ions, SOD1 mutants were mostly in the apo state, which is partially unfolded, similar to WT SOD1.

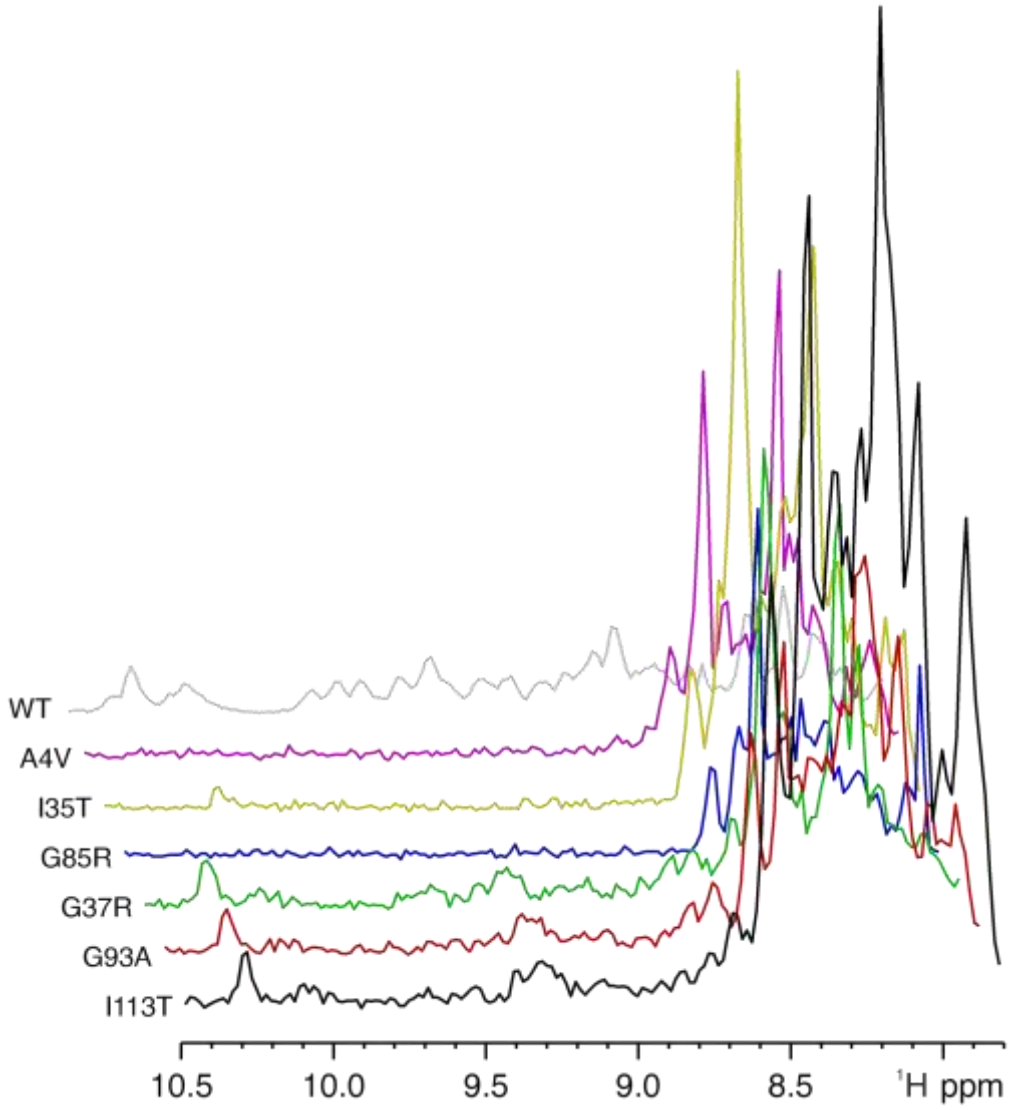


Figure S3: Projections over the 15N dimension of the in-cell NMR spectra of figure 2, showing that small amounts of E,Zn-SOD1 species are detected in cells expressing G37R, G93A and I113T SOD1 mutants. The in-cell NMR spectrum of WT E,Zn-SOD1 (scaled at 0.125x) is shown in gray as reference.

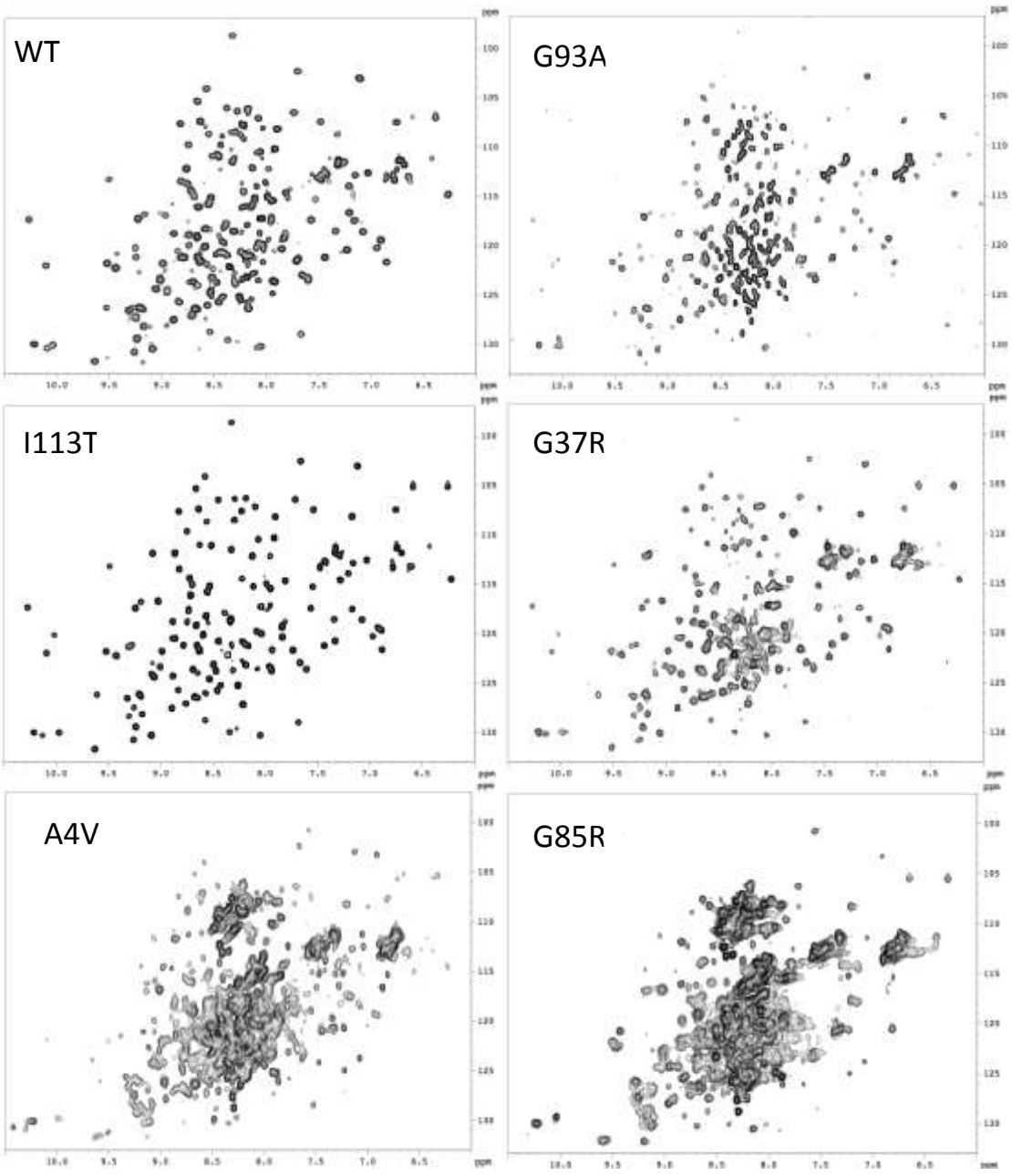


Figure S4: HSQC spectra of WT SOD1 and fALS mutants after purification indicate that in each case the E,Zn-hSOD1^{SS} form was isolated, with the exception of G85R, where the apo-oxidized form was isolated.

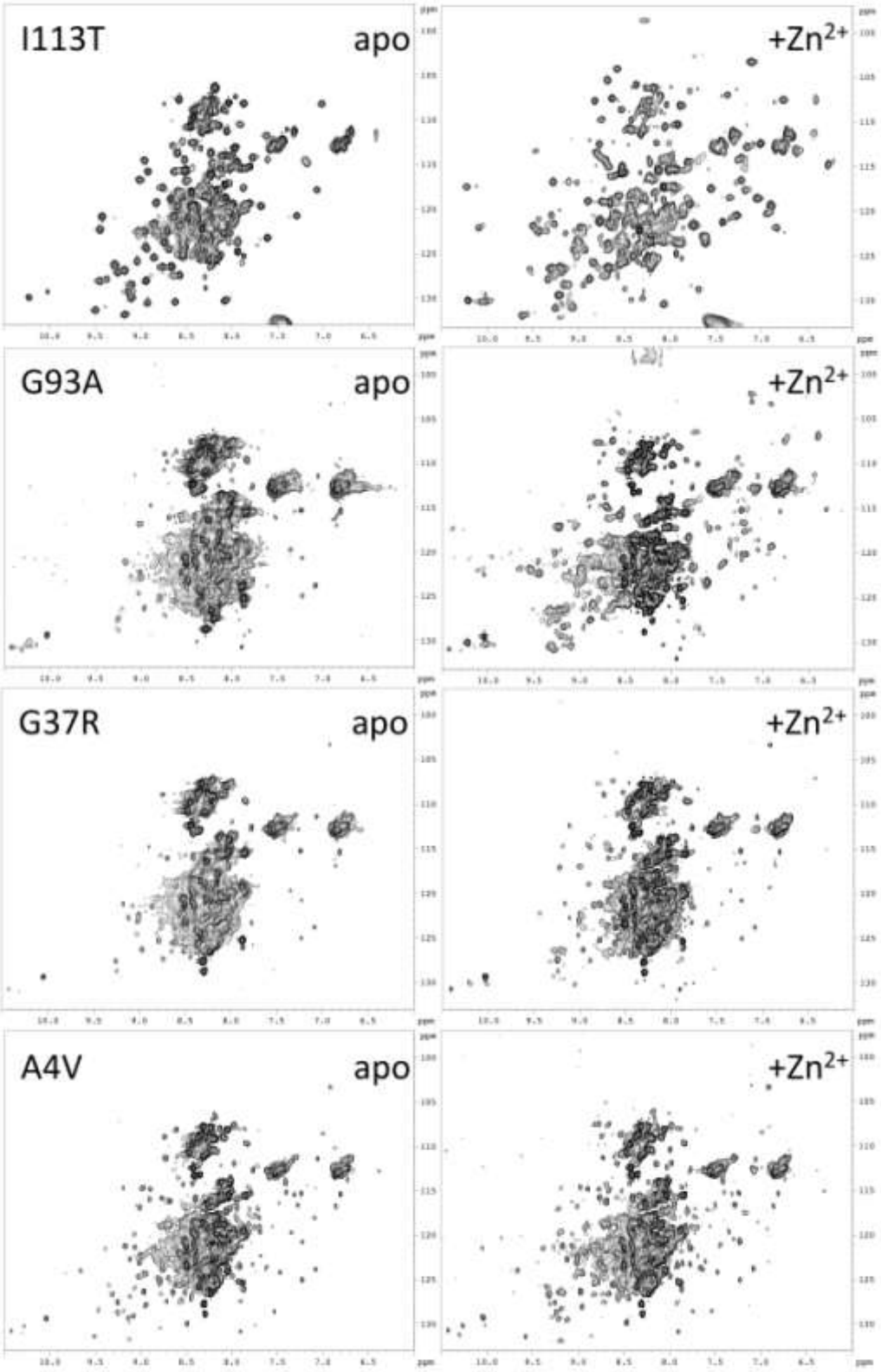


Figure S5: HSQC spectra of fALS mutants after demetallation and reduction (left side) and titrated with Zn^{2+} (right side).

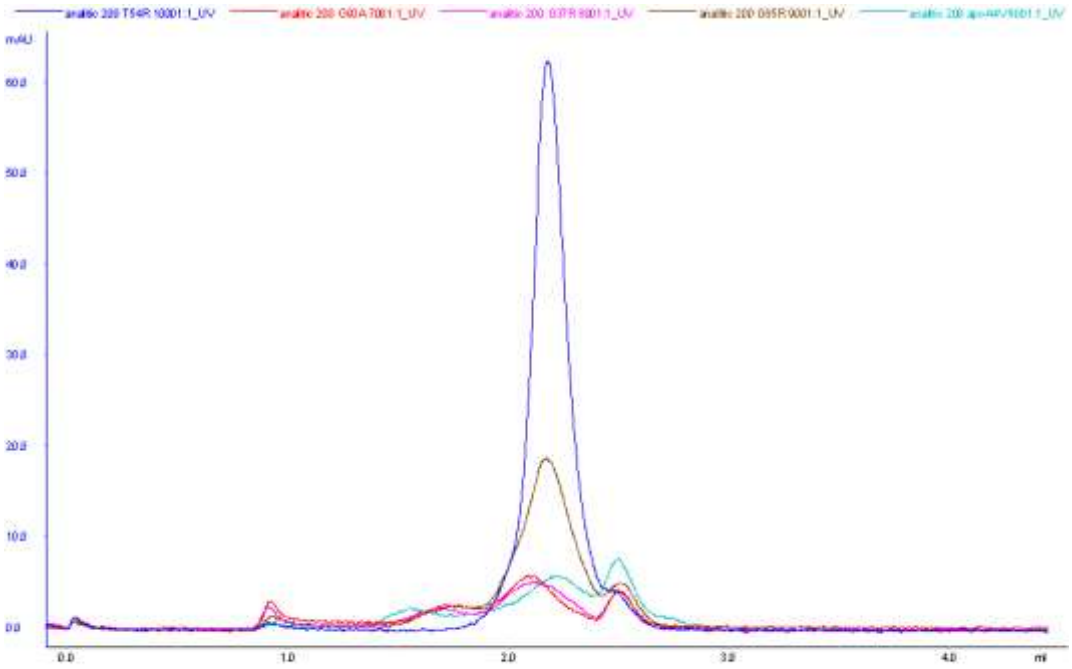


Figure S6: Analytical size exclusion gel filtration of fALS apo reduced mutants: T54R (blue), G93A (red), G37R (pink),G85R (green), A4V (grey).

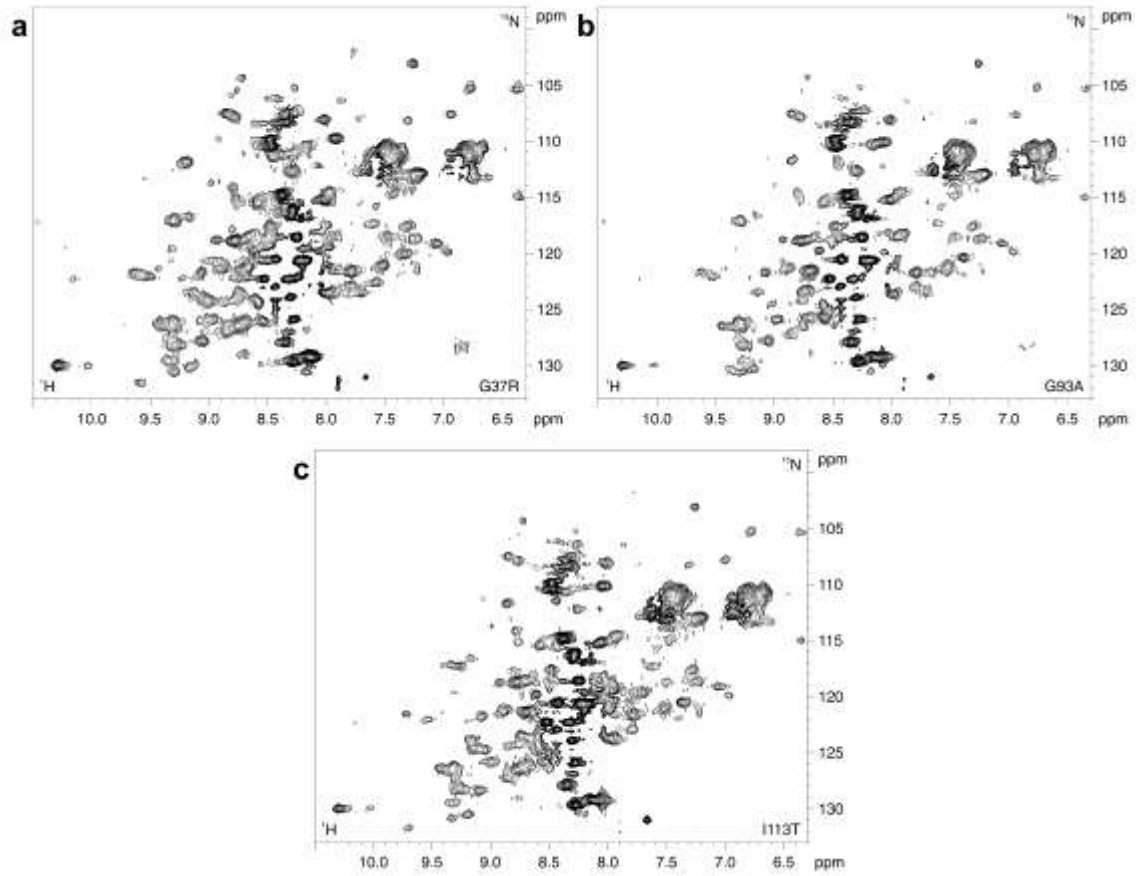


Figure S7: In presence of zinc and copper, co-expression of CCS allowed the formation of the mature form of SOD1 mutants (Cu,Zn-SOD1^{SS}), and no apo-form was detected.

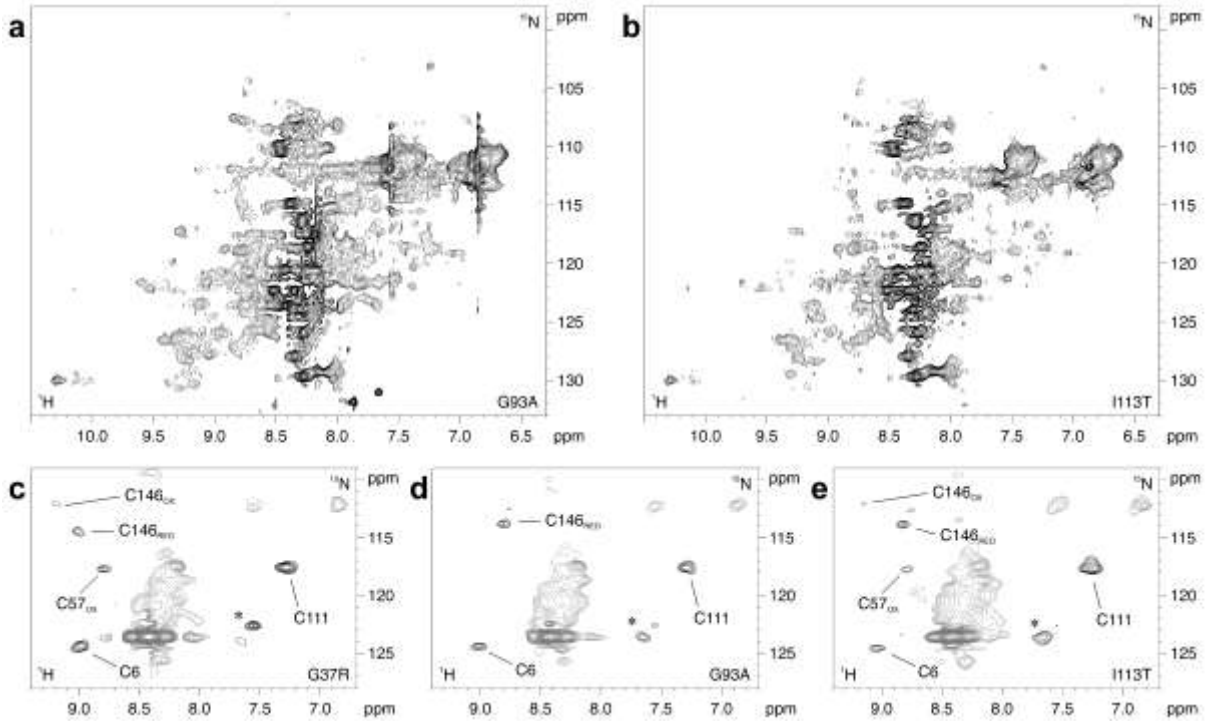


Figure S8: Cu-free CCS was less efficient than Cu-loaded CCS in increasing the amount of Zn-containing mutant SOD1 and in catalyzing disulfide bond formation, and a mixture of species was generated.

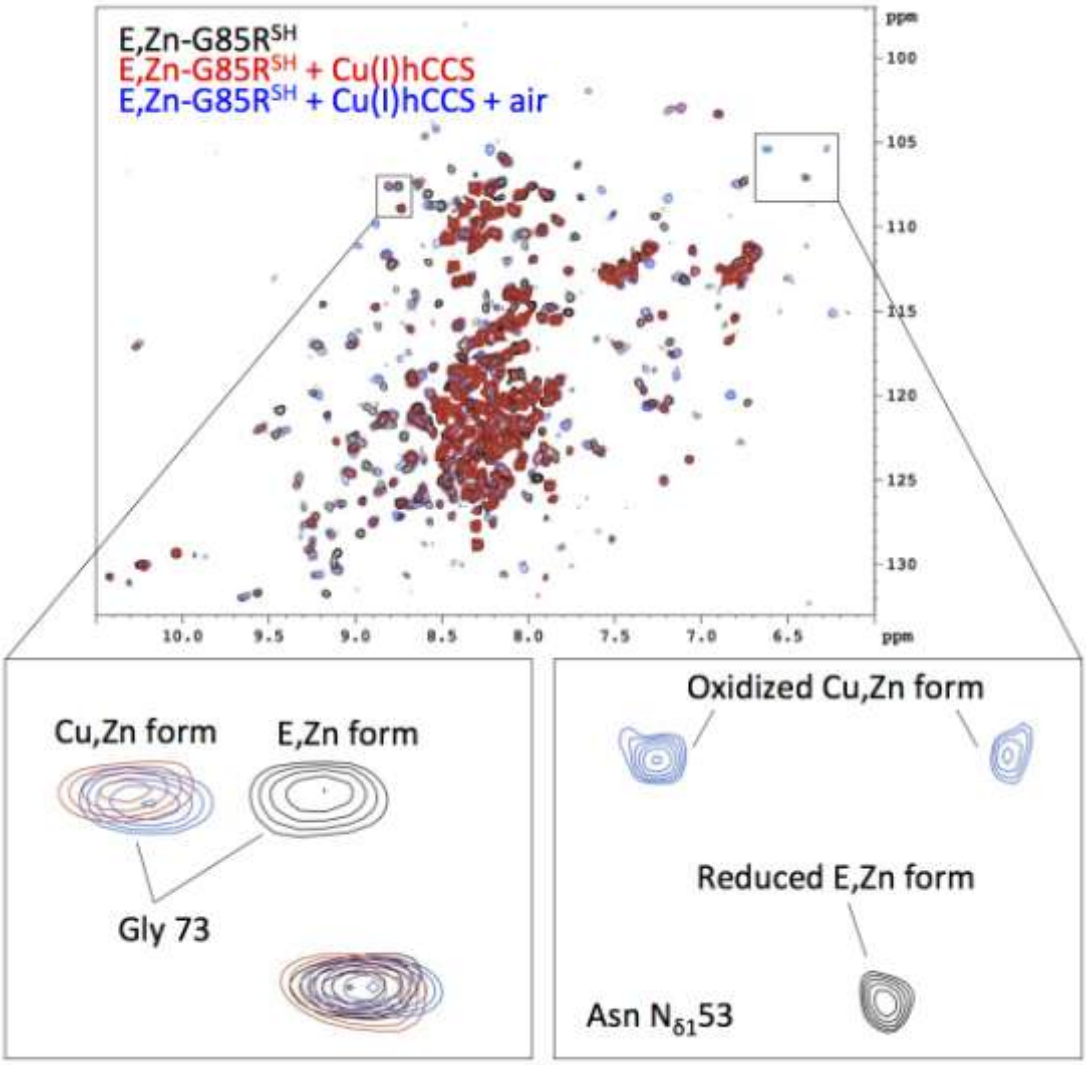


Figure S9: Overlay of HSQC spectra of E,Zn-G85R^{SH} (black) after addition of Cu(I)hCCS (red) and exposure to air (blue). Selected regions highlight the amide signal of residue Gly-73, which is demonstrative of copper binding because it is located in proximity to the catalytic metal-binding site, and the chemical shifts of side chain signals of Asn-53, which are affected by both copper binding and oxidation. Asn-53 is located in loop 4 of hSOD1 and faces the C57-C146 disulfide on its formation. Due to the low intensity of the folded G85R signals, in comparison to those associated with the soluble oligomer, we were unable to analyze the spectra of the G85R titration with Cu(I)CCS in the same fashion as the other mutants.

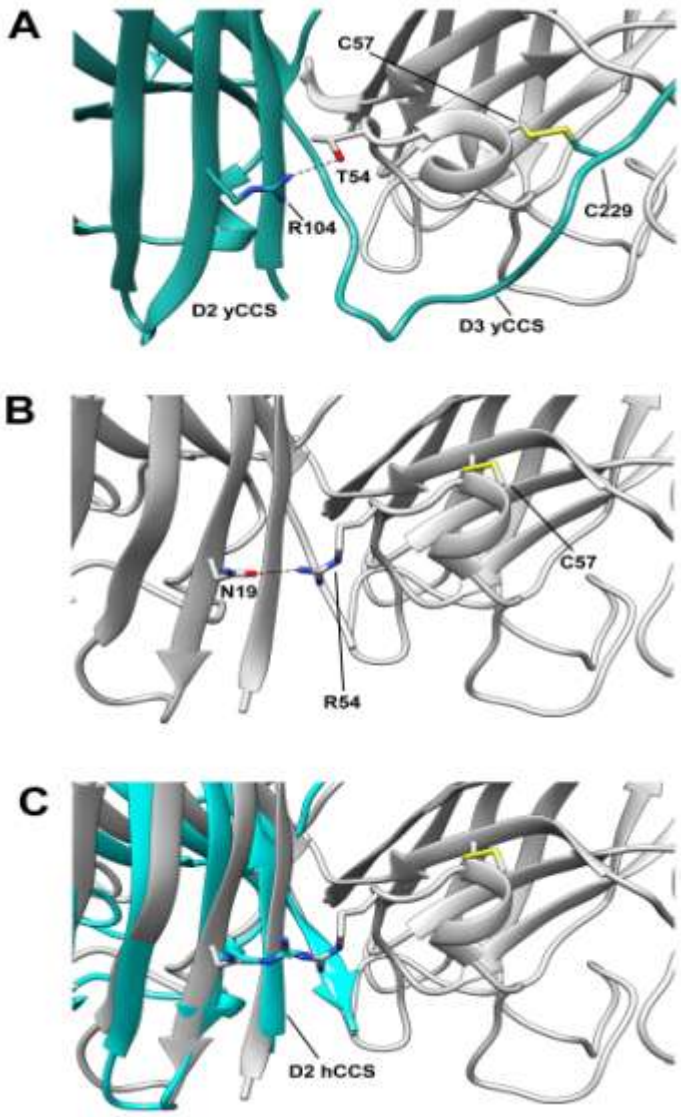


Figure S10: Crystal structure interfaces of (A) yCCS (cyan) and ySOD (grey) (PDB: 1JK9), (B) dimeric apoT54R^{SS} (PDB: 3ECW), and (C) dimeric apoT54R^{SS} (grey) superimposed with domain 2 of hCCS (cyan) (PDB: 1DO5)

The slower rate of oxidation observed for T54R SOD1 may be the result of the loss of a potential intermolecular hydrogen bond between T54 of hSOD1 and R104 of hCCS, as observed in the yeast proteins (**a**) (44). This hydrogen bond may stabilize the region around C57 of hSOD1 optimizing it for the disulfide bond transfer from the CXC motif of domain 3 of hCCS to hSOD1. In the crystal structure of dimeric T54R apo-hSOD1^{SS}, R54 from one subunit forms a novel hydrogen bond with N19 from the other subunit (**b**) (45). Superimposition of the structure of domain 2 of hCCS (46) over one monomer of T54R apo-hSOD1^{SS} brings R104 of hCCS in proximity of R54 of the other hSOD1 monomer, suggesting that the T-to-R mutation, in addition to eliminating the potential hydrogen bond, further destabilizes the region due to the proximity of the positively charged arginine side chains (**c**).

References:

1. Banci L, Barbieri L, Bertini I, Cantini F, & Luchinat E (2011) In-cell NMR in E. coli to Monitor Maturation Steps of hSOD1. *PLoS One* 6(8).

2. Selenko P & Wagner G (2007) Looking into live cells with in-cell NMR spectroscopy. *Journal of Structural Biology* 158(2):244-253.

3. Banci L, *et al.* (2013) Atomic-resolution monitoring of protein maturation in live human cells by NMR. *Nat. Chem. Biol.* 9(5):297-+.

4. Banci L, Barbieri L, Luchinat E, & Secci E (2013) Visualization of Redox-Controlled Protein Fold in Living Cells. *Chem. Biol.* 20(6):747-752.

5. Culotta VC, *et al.* (1997) The copper chaperone for superoxide dismutase. *J. Biol. Chem.* 272(38):23469-23472.

6. Furukawa Y, Torres AS, & O'Halloran TV (2004) Oxygen-induced maturation of SOD1: a key role for disulfide formation by the copper chaperone CCS. *EMBO J.* 23(14):2872-2881.

7. Carroll MC, *et al.* (2004) Mechanisms for activating Cu- and Zn-containing superoxide dismutase in the absence of the CCS Cu chaperone. *Proc. Natl. Acad. Sci. USA* 101(16):5964-5969.

8. Banci L, *et al.* (2012) Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS). *Proc. Natl. Acad. Sci. USA* 109(34):13555-13560.

9. Rosen DR, *et al.* (1993) Mutations in Cu/Zn superoxide-dismutase gene are associated with familial amyotrophic-lateral-sclerosis. *Nature* 362(6415):59-62.

10. Shibata N, *et al.* (1996) Intense superoxide dismutase-1 immunoreactivity in intracytoplasmic hyaline inclusions of familial amyotrophic lateral sclerosis with posterior column involvement. *Journal of Neuropathology and Experimental Neurology* 55(4):481-490.

11. Watanabe M, *et al.* (2001) Histological evidence of protein aggregation in mutant SOD1 transgenic mice and in amyotrophic lateral sclerosis neural tissues. *Neurobiol. Dis.* 8(6):933-941.

12. Jonsson PA, *et al.* (2004) Minute quantities of misfolded mutant superoxide dismutase-1 cause amyotrophic lateral sclerosis. *Brain* 127:73-88.

13. Bruijn LI, *et al.* (1997) ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Neuron* 18(2):327-338.

14. Bruijn LI, *et al.* (1998) Aggregation and motor neuron toxicity of an ALS-linked SOD1 mutant independent from wild-type SOD1. *Science* 281(5384):1851-1854.

15. Johnston JA, Dalton MJ, Gurney ME, & Kopito RR (2000) Formation of high molecular weight complexes of mutant Cu,Zn-superoxide dismutase in a mouse model for familial amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. USA* 97(23):12571-12576.

16. Wang J, Xu G, & Borchelt DR (2002) High molecular weight complexes of mutant superoxide dismutase 1: Age-dependent and tissue-specific accumulation. *Neurobiol. Dis.* 9(2):139-148.

17. Mulligan VK & Chakrabartty A (2013) Protein misfolding in the late-onset neurodegenerative diseases: Common themes and the unique case of amyotrophic lateral sclerosis. *Proteins: Struct., Funct., Bioinf.*:n/a-n/a.

18. Ross CA & Poirier MA (2004) Protein aggregation and neurodegenerative disease. *Nature Medicine* 10(7):S10-S17.

19. Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. *Annu. Rev. Biochem,* Annual Review of Biochemistry), Vol 75, pp 333-366.

20. Bucciantini M, *et al.* (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. *Nature* 416(6880):507-511.

21. Banci L, *et al.* (2007) Metal-free superoxide dismutase forms soluble oligomers under physiological conditions: a possible general mechanism for familial ALS. *Proc. Natl. Acad. Sci. USA* 104(27):11263-11267.

22. Toichi K, Yamanaka K, & Furukawa Y (2013) Disulfide Scrambling Describes the Oligomer Formation of Superoxide Dismutase (SOD1) Proteins in the Familial Form of Amyotrophic Lateral Sclerosis. *J. Biol. Chem.* 288(7):4970-4980.

23. Banci L, *et al.* (2008) SOD1 and Amyotrophic Lateral Sclerosis: Mutations and Oligomerization. *PLoS One* 3(2).

24. Wang J, *et al.* (2002) Fibrillar inclusions and motor neuron degeneration in transgenic mice expressing superoxide dismutase 1 with a disrupted copper-binding site. *Neurobiol. Dis.* 10(2):128-138.

25. Kerman A, *et al.* (2010) Amyotrophic lateral sclerosis is a non-amyloid disease in which extensive misfolding of SOD1 is unique to the familial form. *Acta Neuropathologica* 119(3):335-344.

26. Wang J, *et al.* (2005) Somatodendritic accumulation of misfolded SOD1-L126Z in motor neurons mediates degeneration: alpha B-crystallin modulates aggregation. *Hum. Mol. Genet.* 14(16):2335-2347.

27. Lindberg MJ, Tibell L, & Oliveberg M (2002) Common denominator of Cu/Zn superoxide dismutase mutants associated with amyotrophic lateral sclerosis: Decreased stability of the apo state. *Proc. Natl. Acad. Sci. USA* 99(26):16607-16612.

28. Furukawa Y, Kaneko K, Yamanaka K, O'Halloran TV, & Nukina N (2008) Complete loss of post-translational modifications triggers fibrillar aggregation of SOD1 in the familial form of amyotrophic lateral sclerosis. *J. Biol. Chem.* 283(35):24167-24176.

29. Durer ZAO, *et al.* (2009) Loss of Metal Ions, Disulfide Reduction and Mutations Related to Familial ALS Promote Formation of Amyloid-Like Aggregates from Superoxide Dismutase. *PLoS One* 4(3).

30. Hayward LJ, *et al.* (2002) Decreased metallation and activity in subsets of mutant superoxide dismutases associated with familial amyotrophic lateral sclerosis. *J. Biol. Chem.* 277(18):15923-15931.

31. Lelie HL, *et al.* (2011) Copper and zinc metallation status of copper-zinc superoxide dismutase from amyotrophic lateral sclerosis transgenic mice. *J. Biol. Chem.* 286(4):2795-2806.

32. Hough MA, *et al.* (2004) Dimer destabilization in superoxide dismutase may result in disease-causing properties: Structures of motor neuron disease mutants. *Proc. Natl. Acad. Sci. USA* 101(16):5976-5981.

33. Svensson A-KE, *et al.* (2010) Metal-free ALS variants of dimeric human Cu,Zn-superoxide dismutase have enhanced populations of monomeric species. *PLoS One* 5(4).

34. Tiwari A & Hayward LJ (2003) Familial amyotrophic lateral sclerosis mutants of copper/zinc superoxide dismutase are susceptible to disulfide reduction. *J. Biol. Chem.* 278(8):5984-5992.

35. Cao XH, *et al.* (2008) Structures of the G85R variant of SOD1 in familial amyotrophic lateral sclerosis. *J. Biol. Chem.* 283(23):16169-16177.

36. Lindberg MJ, Normark J, Holmgren A, & Oliveberg M (2004) Folding of human superoxide dismutase: Disulfide reduction prevents dimerization and produces marginally stable monomers. *Proc. Natl. Acad. Sci. USA* 101(45):15893-15898.

37. Rumfeldt JAO, Stathopulos PB, Chakrabarrty A, Lepock JR, & Meiering EM (2006) Mechanism and thermodynamics of guanidinium chloride-induced denaturation of ALS-associated mutant Cu,Zn superoxide dismutases. *J. Mol. Biol.* 355(1):106-123.

38. Kayatekin C, Zitzewitz JA, & Matthews CR (2008) Zinc Binding Modulates the Entire Folding Free Energy Surface of Human Cu,Zn Superoxide Dismutase. *J. Mol. Biol.* 384(2):540-555.

39. Nordlund A, *et al.* (2009) Functional features cause misfolding of the ALS-provoking enzyme SOD1. *Proc. Natl. Acad. Sci. USA* 106(24):9667-9672.

40. Mulligan VK, Kerman A, Ho S, & Chakrabartty A (2008) Denaturational Stress Induces Formation of Zinc-Deficient Monomers of Cu,Zn Superoxide Dismutase: Implications for Pathogenesis in Amyotrophic Lateral Sclerosis. *J. Mol. Biol.* 383(2):424-436.

41. Lindberg MJ, Bystrom R, Boknas N, Andersen PM, & Oliveberg M (2005) Systematically perturbed folding patterns of amyotrophic lateral sclerosis (ALS)-associated SOD1 mutants. *Proc. Natl. Acad. Sci. USA* 102(28):9754-9759.

42. Vassall KA, Stathopulos PB, Rumfeldt JAO, Lepock JR, & Meiering EM (2006) Equilibrium thermodynamic analysis of amyotrophic lateral sclerosis-associated mutant apo Cu,Zn superoxide dismutases. *Biochemistry* 45(23):7366-7379.

43. Kayatekin C, Zitzewitz JA, & Matthews CR (2010) Disulfide-Reduced ALS Variants of Cu, Zn Superoxide Dismutase Exhibit Increased Populations of Unfolded Species. *J. Mol. Biol.* 398(2):320-331.

44. Lamb AL, Torres AS, O'Halloran TV, & Rosenzweig AC (2001) Heterodimeric structure of superoxide dismutase in complex with its metallochaperone. *Nat. Struct. Biol.* 8(9):751-755.

45. Banci L, *et al.* (2009) Structural and dynamic aspects related to oligomerization of apo SOD1 and its mutants. *Proc. Natl. Acad. Sci. USA* 106(17):6980-6985.

46. Lamb AL, Wernimont AK, Pufahl RA, O'Halloran TV, & Rosenzweig AC (2000) Crystal structure of the second domain of the human copper chaperone for superoxide dismutase. *Biochemistry* 39(7):1589-1595.

2.4

Visualization of the protein fold and redox state in living cells as a function of other proteins

Banci, L; Barbieri, L; Luchinat, E; Secci, E

Chem.Biol., 20, 747-752, 2013

Chemistry & Biology
Brief Communication

Visualization of Redox-Controlled Protein Fold in Living Cells

Lucia Banci,[1,2,*] Letizia Barbieri,[1] Enrico Luchinat,[1,3] and Erica Secci[1]
[1]Magnetic Resonance Center - CERM, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy
[2]Department of Chemistry, University of Florence, Via della Lastruccia 3, 50019 Sesto Fiorentino, Florence, Italy
[3]Department of Biomedical, Clinical and Experimental Sciences, University of Florence, Viale Morgagni 50, 50134 Florence, Italy
*Correspondence: banci@cerm.unifi.it
http://dx.doi.org/10.1016/j.chembiol.2013.05.007

SUMMARY

Most mitochondrial proteins are encoded by nuclear DNA, synthesized in the cytoplasm, and imported into mitochondria. Several proteins of the intermembrane space (IMS) are imported and localized through an oxidative process, being folded through the formation of structural disulfide bonds catalyzed by Mia40, and trapped in the IMS. To be imported, these proteins need to be reduced and unfolded; however, no structural information in situ exists on these proteins in the cytoplasm. In humans, Mia40 undergoes the same mechanism, although its folding state in the cytoplasm is unknown. We provide atomic-level details on the Mia40 folding state in the human cell cytoplasm through in-cell nuclear magnetic resonance. Overexpressed cytoplasmic Mia40 is folded, and its folding state depends on the glutaredoxin 1 (Grx1) and thioredoxin 1 (Trx1) systems. Specifically, increased Grx1 levels keep most Mia40 unfolded, while Trx1 is less effective.

INTRODUCTION

The majority of the human proteins are produced by nuclear DNA and released in the cytoplasm and/or in the endoplasmic reticulum (ER). These proteins then need to complete their folding and maturation process, which could involve several steps, from cofactor binding to cysteine oxidation to other posttranslational modifications. Furthermore, if the protein destiny is a cellular compartment other than the cytoplasm, some or all maturation and folding steps may take place in the final cellular localization. This is particularly true for a large share of proteins present in the intermembrane space (IMS) of mitochondria, but which do not feature any target sequence for this organelle (Banci et al., 2009b; Longen et al., 2009; Neupert and Herrmann, 2007). They were proposed to be in an unfolded state in the cytoplasm and thus can enter mitochondria, thanks to their conformational flexibility that allows them to go through the translocase of the outer membrane (TOM) channel (Neupert and Herrmann, 2007). Once they have entered the IMS, they fold to their native form, thus being blocked in a defined, more rigid conformation, which prevents them from crossing back to the outer membrane (Chacinska et al., 2004; Lu et al., 2004; Mesecke et al., 2005). It is therefore evident that the folding state of a protein depends on the cellular compartment where the protein is located and on its properties. It has been shown that the import of some of these IMS proteins is in fact modulated by cytosolic thiol-disulfide regulation systems (Durigon et al., 2012). However no detailed, atomic resolution study has confirmed such findings.

In this work we show, by exploiting in-cell nuclear magnetic resonance (in-cell NMR; Banci et al., 2013; Inomata et al., 2009; Ogino et al., 2009; Reckel et al., 2007; Selenko et al., 2008), that indeed the properties of the cell compartment in terms of redox-regulating components influence the folding state of one of such proteins, whose oxidation state is compartment-dependent. We have specifically characterized Mia40, a hub protein for the mitochondrial protein import process. Mia40 is an oxidoreductase that catalyzes in the IMS the formation of internal disulfide bonds on its protein substrates through the intermediate formation of a mixed disulfide bond between the substrate and its catalytic CPC motif (Banci et al., 2009a; Chacinska et al., 2004; Grumbt et al., 2007; Naoé et al., 2004). Together, these proteins constitute the disulfide relay system of the IMS (Mesecke et al., 2005; Tokatlidis, 2005). Upon interaction with Mia40 and subsequent formation of their disulfide bonds, the substrates of Mia40 become folded and are trapped in the IMS (Banci et al., 2010; Gabriel et al., 2007). Interestingly, Mia40 itself obtains its final structure in the IMS upon the formation of two internal disulfide bonds, likely by acting as a substrate of itself (Chacinska et al., 2008; Grumbt et al., 2007). Like the other substrates of the disulfide relay system, Mia40 has to cross the outer mitochondrial membrane in an unfolded, reduced state. Glutaredoxin 1 (Grx1) and thioredoxin 1 (Trx1) are cytoplasmic oxidoreductases involved in the regulation of protein thiol groups and in the cellular defense against oxidative stress (Holmgren, 1989; Meyer et al., 2009). Trx1 was recently shown to be responsible for facilitating the mitochondrial import of the small Tim proteins (Durigon et al., 2012). It can be hypothesized that other small proteins of the IMS sharing the same import mechanism, including Mia40, are regulated by such thiol-regulating proteins.

We show here that Mia40, even in the reducing environment of the cytoplasm, is largely in the oxidized, folded state when it is overexpressed, thus indicating that the high cytoplasmic level of reduced glutathione is not sufficient alone to maintain the reduced state of Mia40. Co-expression of Grx1 (Meyer et al., 2009; Sagemark et al., 2007) keeps Mia40 mostly in the unfolded, reduced state, while co-expression of Trx1, which has a similar role in keeping protein thiols reduced in the

cytoplasm (Holmgren, 1979; Meyer et al., 2009), affects the oxidation state of Mia40 to a lesser extent. Additionally, Grx1 does not catalyze the reduction of Mia40 in the presence of reducing agents in vitro, implying some effect of the cytoplasmic environment that is not reproduced in vitro. Overall, these results indicate that Mia40, and likely other mitochondrial proteins that share the same structural motif and import pathway, tend to reach the oxidized, folded conformation in the absence of specific proteins, even in the reducing environment of the cytoplasm, and therefore require intervention of the cytoplasmic thiol-disulfide regulation mechanisms, especially those of Grx1, to reach the outer mitochondrial membrane in the reduced, import-competent state.

RESULTS

Transient expression of human Mia40, either alone or together with Grx1 or Trx1, was induced in human embryonic kidney (HEK293T) cells, and the relative expression levels were modulated by transfecting different amounts of DNA for each gene. The expression levels of each protein were determined with western blot analysis of the cell extracts (Figure 1). The concentration of overexpressed Mia40 remained constant in most of the samples, around 60 ± 10 μM. Grx1 was co-expressed in various amounts, ranging from 20 ± 10 μM to 100 ± 30 μM. Trx1 was co-expressed in two amounts: 35 ± 4 μM and 70 ± 15 μM. We assessed the cellular distribution of the overexpressed proteins in isolated cytosolic and mitochondrial fractions. Irrespective of the total protein abundance, >96% of Mia40 and >99% of either Grx1 or Trx1 were present in the cytosolic fraction (Figure S1A available online). Endogenous Mia40 remained undetected (<20 nM, data not shown), while endogenous levels of Grx1 were 60 ± 20 nM, ~10% of which was found in mitochondria, and 5 ± 2 μM of Trx1 was localized in the cytoplasm only, as measured on extracts of untransfected cells (Figure S1B).

Mia40, overexpressed alone in the cytoplasm of human cells, took a folded conformation, as monitored through ^{1}H-^{15}N and ^{1}H NMR spectra, which corresponds to the functional state of Mia40 that is normally found in the IMS (Figure 2A), and is the same conformation as that of oxidized Mia40 in vitro ($Mia40^{2S\text{-}S}$; Figure S2A). The central region of the protein is stably folded in two α helices while the N- and C-terminal regions are intrinsically unstructured. The cross-peaks of the folded part were well dispersed in the spectrum, while the peaks of the unstructured regions appeared as intense overlapped signals, which fell in the central part of the spectrum. Analysis of the methyl region of the ^{1}H NMR spectrum, which provides more sensitivity to assess the relative amount of folded Mia40 (Figure 2B), indicated that the folded protein was ~85% of the total protein amount (Figure 4). When it was co-expressed with various amounts of Grx1, a large fraction of Mia40 was present in the reduced, unfolded state. This effect was confirmed by a decrease of the methyl signal and also reflected in the ^{1}H-^{15}N NMR spectra, in which the cross-peaks of the folded part of Mia40 were barely detected as expected, while the cross-peaks of the unfolded parts were still visible, both in intact cells (Figures 2C and 2D) and in cell extracts (Figures 3C and 3D). In the presence of increasing amounts of Grx1, oxidized Mia40 ($Mia40^{2S\text{-}S}$) in the cytoplasm decreased from ~85% to ~25% of the total

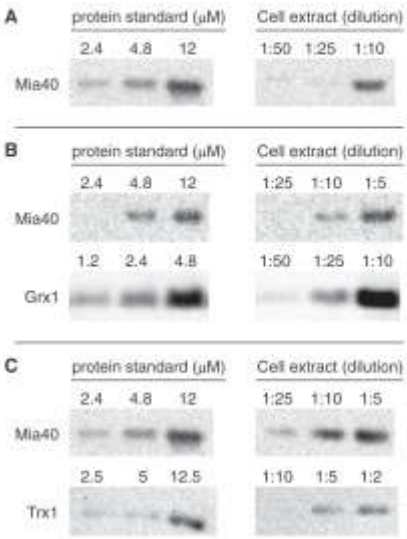


Figure 1. The Total Amount of Mia40, Grx1, and Trx1 in the Cell Extracts Is Measured with Western Blot Analysis

Samples of cell extracts were blotted at increasing dilutions together with samples of pure Mia40, Grx1 and Trx1 at known concentrations.

(A) Extract from cells expressing Mia40 alone.

(B) Extract from cells co-expressing Mia40 and Grx1 (1:0.25:2 *hMia40:hGrx1*:PEI ratio).

(C) Extract from cells co-expressing Mia40 and Trx1 (1:1:2 *hMia40:hTrx1*:PEI ratio) in selenium-supplemented medium. A calibration curve was obtained for each blot from the intensities of the standard Mia40 samples.

The concentration of Mia40 in the cell extracts was calculated for each dilution from the calibration curve and averaged.

See also Figure S1.

Mia40 (Figure 4). Co-expressed Grx1 was in the fully reduced and active state (Figure S1C), consistent with its reported redox potential (E_0 = −230 mV, compared to −290 mV for the GSH/GSSG couple in the cytoplasm; Sagemark et al., 2007). Grx1 is invisible in the in-cell NMR spectra because its signals are broadened beyond detection as a consequence of its slow tumbling rate, which is likely due to interactions with the cellular environment (a similar effect has been reported for several proteins overexpressed in *E. coli* cytoplasm, including wild-type ubiquitin [Sakai et al., 2006], cytochrome c [Crowley et al., 2011], *Escherichia coli* thioredoxin, and FKBP [Reckel et al., 2012]). This behavior was confirmed by the empty ^{1}H-^{15}N in-cell NMR spectra when Grx1 alone was expressed (Figure S3A), whereas Grx1 signals became visible upon cell lysis (Figure S3B). The NMR properties of cytoplasmic Grx1 therefore allowed us to obtain ^{1}H-^{15}N NMR spectra of Mia40 free of interference from Grx1 signals.

The effect of Trx1 on the Mia40 state was also investigated by inducing co-expression of Trx1 with Mia40. Similar to Grx1, Trx1 in the human cell cytoplasm is not detectable with NMR, consistent with what was reported in *E. coli* (Reckel et al., 2012; Figure S3C). The intracellular activity of Trx1 depends on the

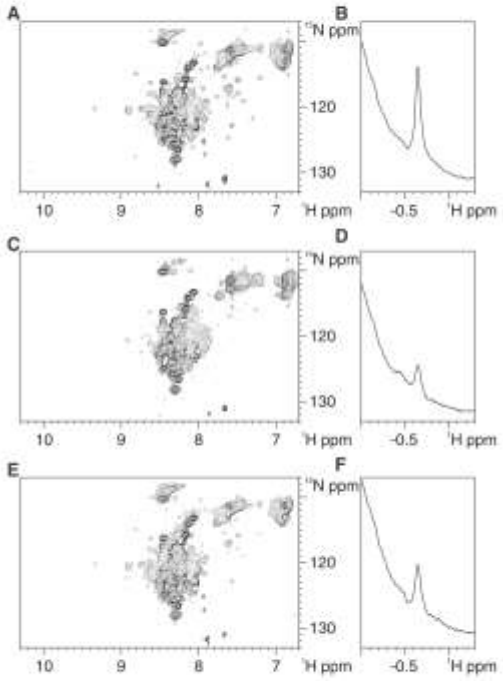


Figure 2. The Folding State of Mia40 in the Cytoplasm Depends on the Presence of Different Redox-Regulating Proteins

NMR spectra were acquired on human cells expressing uniformly ^{15}N-labeled Mia40 in the cytoplasm.

(A and B) Spectra of cells expressing Mia40.

(C and D) Spectra of cells co-expressing Mia40 and Grx1.

(E and F) Spectra of cells co-expressing Mia40 and Trx1 in selenium-supplemented medium.

(A, C, and E) ^{1}H-^{15}N SOFAST HMQC spectra. The strong, overlapped cross-peaks between 8.0 and 8.5 ppm (^{1}H) correspond to the unfolded regions, while the weaker, dispersed cross-peaks belong to the residues of the folded region of Mia40 (see also Figure S2).

(B, D, and F) Aliphatic region of ^{1}H spectra showing the ^{1}Hγ peak of Ile 53, which is a marker of the folded conformation of Mia40 and falls in a cellular background-free region.

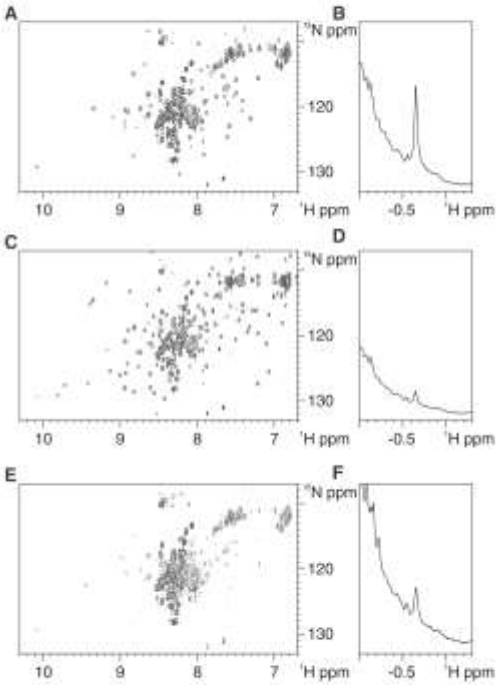


Figure 3. NMR Spectra Acquired on the Cell Extracts Corresponding to the Samples in Figure 2

NMR spectra were acquired on the cell extract of samples expressing uniformly ^{15}N-labeled Mia40. The relative amounts of the two Mia40 folding states remained unchanged upon cell lysis.

(A and B) Spectra of the cell extract containing Mia40.

(C and D) Spectra of the cell extract containing Mia40 and Grx1.

(E and F) Spectra of the cell extract containing Mia40 and Trx1 in selenium-supplemented medium.

Upon cell lysis, both Grx1 and Trx1 became visible in the ^{1}H-^{15}N spectra (C and E; see also Figure S3). Glutathionylation of unfolded Mia40 was excluded with mass spectrometry (see also Figure S4).

availability of selenium, which is required for the activation of thioredoxin reductase (TrxR; Ueno et al., 2007). Indeed, Trx1 co-expression in cells grown with basal amounts of selenium did not affect the folding state of Mia40 (data not shown). To ensure the complete activation of TrxR, and consequently active Trx1, sodium selenite was supplemented to cells co-expressing Trx1 and Mia40. In these conditions, co-expressed Trx1 was fully reduced (Figure S1C) and active, consistent with the reported Trx1 redox state at endogenous levels (Watson et al., 2003). Co-expression of active Trx1 affected the folding state of Mia40 to a lesser extent than Grx1 when expressed at similar levels because ~50% of cytoplasmic Mia40 was still in the folded, oxidized state (Figures 2E, 2F, 3E, and 3F). Therefore, the two thiol-regulating proteins have different efficacy in keeping Mia40 reduced, despite being reported to have overall similar functions in the cytoplasm (Figure 4).

In contrast to what was observed in the cells, fully reduced Grx1 had no effect on the Mia40 redox state in vitro. U-^{15}N Mia40$^{2S\text{-}S}$ was incubated in reduction conditions, either in the presence of 20 mM dithiothreitol (DTT) or 20 mM reduced glutathione (GSH), with increasing concentrations of unlabeled, fully reduced Grx1, and each step was monitored with NMR. No change in the ^{1}H-^{15}N NMR spectrum occurred upon addition of up to 2 eq of Grx1 to Mia40$^{2S\text{-}S}$ and 48 hr incubation, thus excluding a direct mechanism of reduction of Mia40 by Grx1. In a control experiment, U-^{15}N Mia40$^{2S\text{-}S}$ (Figure S2A) was completely reduced by heat denaturation at 95°C in buffer containing either 20 mM DTT or 20 mM GSH. No protein degradation occurred. The protein remained reduced in both reductants when cooled down at 25°C (Figure S2C). In these conditions,

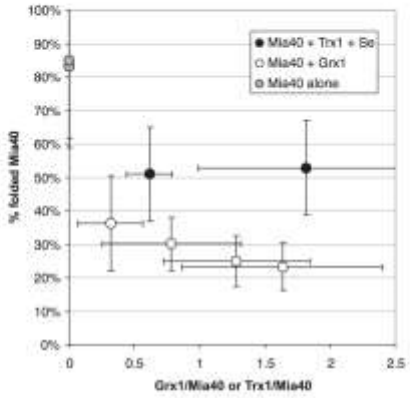


Figure 4. Effect of Grx1 and Trx1 on the Mia40 Folding State in the Cytoplasm

The ratio of folded Mia40 over total Mia40 was determined in cell samples with varying relative amounts of co-expressed Grx1 or Trx1. The amount of each protein was measured on cell extracts with western blot analysis. The amount of folded Mia40 was measured on the cell extracts by NMR. When only Mia40 was expressed, it was largely found in the folded state in the cytoplasm (gray circles). Co-expression of Grx1 caused a remarkable decrease of folded Mia40, which also occurred at substoichiometric amounts relative to Mia40 (white circles). Co-expression of Trx1 in the presence of selenium affected the folding state of Mia40 to a lesser extent (black circles).

Error bars represent SDs; n = 3.

the ^{1}H-^{15}N cross-peaks of the α-helical region disappeared, analogous to what was observed in the cytoplasm in the presence of Grx1. The alkylation reaction with 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS) depicted with SDS-PAGE confirmed the complete reduction of all cysteines (Figures S2B and S2D). Upon removal of the reducing agent and exposure to air, Mia40 rapidly reverted back to the folded, oxidized conformation. To test whether Grx1 could bind fully reduced Mia40 and prevent its oxidation, 2 eq of Grx1 was added to a sample of U-^{15}N labeled, unfolded Mia40 in presence of 20 mM GSH. Upon exposure to air, Mia40 rapidly folded, thus indicating that Grx1 did not protect reduced Mia40 from oxidation. Grx1-catalyzed glutathionylation of reduced Mia40, either in vitro in presence of GSH or in the cytoplasm, was excluded by mass spectrometry analysis of the in vitro Mia40 samples and of Mia40 isolated from cell extracts (Figure S4).

DISCUSSION

The current model of Mia40 maturation pathway requires that the protein, which is natively expressed in the cytoplasm from nuclear mRNA, crosses the outer mitochondrial membrane through the TOM channel in a reduced and unfolded conformation. No Mia40 has been reported to reside in the cytoplasm, except for the time required to translocate to the mitochondria (Hofmann et al., 2005). However, when Mia40 was overexpressed, it did not translocate quantitatively to the mitochondria, and close to all of the Mia40 remained in the cytoplasm. This effect could be explained with the limited availability of the import machinery on the outer mitochondrial membrane (i.e., the TOM channel), which determines an upper limit to the import level of Mia40. Consequently, unimported Mia40 accumulates in the cytoplasm. Our results show that in these conditions, cytoplasmic Mia40 reached the folded state, which is thus a thermodynamically favored conformation in the cytoplasm. Therefore, in physiologic conditions there is the need for a mechanism to keep Mia40 reduced until it reaches the outer mitochondrial membrane because the amount and ratio of cytoplasmic GSH are apparently not sufficient for that purpose.

Overexpression of Grx1 allowed most Mia40 to remain in the reduced, unfolded state. This effect also occurred with substoichiometric amounts of Grx1, consistent with a catalytic role of Grx1. Conversely, when oxidized Mia40 was incubated in vitro with Grx1 in the presence of GSH, it did not change its redox state and the reduction of the structural disulfide bonds was only possible upon heat denaturation in reducing conditions. Moreover, Grx1 did not prevent oxidation of heat-denatured Mia40 when added in the presence of GSH and molecular oxygen, and neither partially oxidized Mia40 nor glutathionylated Mia40 were formed. The lack of a direct interaction of Grx1 with Mia40 suggests that the effect of Grx1 on the oxidation state of intracellular Mia40 is mediated by some other component of the cytoplasm.

Overexpression of Trx1 resulted in a weaker effect because a sizable fraction of Mia40 reached the folded state. As with Grx1, the effect occurred also with a substoichiometric amount of Trx1, therefore suggesting that the effect of Grx1, and to a lesser extent that of Trx1, is not the consequence of a generic increase in reducing power in the cytoplasm, which would be concentration-dependent.

Overall, the data show that cytoplasmic redox-regulation systems, Grx1 in particular, have a specific catalytic—although likely indirect—role in keeping a sizable fraction of Mia40 unfolded in the cytoplasm, implying a link between these systems and the Mia40 maturation pathway.

SIGNIFICANCE

We characterized the folding state of Mia40 in the cytoplasm, obtaining atomic-level information in living human cells with NMR. We found that the folded form of Mia40, which cannot be imported into mitochondria, is thermodynamically favored in the cytoplasm. We also showed that the folding of Mia40 is controlled by the cytoplasmic Grx1 and Trx1 redox-regulation systems, Grx1 having a stronger effect. These results show the general relevance of atomic resolution studies performed in living cells, which provide a way to describe cellular physiologic processes such as the redox-controlled protein folding, and to understand how other involved pathways can affect and regulate such processes.

EXPERIMENTAL PROCEDURES

Overexpression of Mia40, Grx1, and Trx1 in human cells was performed by following a previously established protocol (Ariescu et al., 2006; Banci et al., 2013). Briefly, the cDNA sequences encoding Mia40 (amino acids 1–142, GenBank accession number: NP_001091972.1), Grx1 (amino acids

1–106, GenBank accession number: NP_001112362.1), and Trx1 (amino acids 1–105, GenBank accession number: NP_003320.2) were amplified with PCR and subcloned into the pHLsec vector (Aricescu et al., 2006; Banci et al., 2013) between EcoRI and XhoI restriction enzyme sites. The clones were verified by gene sequencing. Transient transfection was obtained by treating the cells with a DNA:polyethylenimine mixture. Cells were collected after 48 hr for in-cell NMR and western blot analysis. Intracellular distribution was assessed by separating the cytoplasmic and the mitochondrial fractions from cell extracts using a mitochondria isolation kit for cultured cells (Thermo Scientific).

Grx1 was co-expressed with Mia40, either unlabeled or with uniform ^{15}N labeling. Different ratios of DNA were chosen to obtain different protein ratios: 1:0.75:2; 1:0.5:2; 1:0.25:2; and 1:0.125:2 *hMia40:hGrx1*:PEI. Trx1 was co-expressed with Mia40 with the following ratios of DNA: 1:1:2 and 0.5:1:2 *hMia40:hTrx1*:PEI. To ensure complete activation of Trx1, sodium selenite (kindly provided by Prof. A. Arcangeli lab, University of Florence) was supplemented to the cell culture media to a final concentration of 100 nM, starting 24 hr before transfection.

The amount of folded Mia40 in the cell extracts was measured with NMR through the standard addition of pure, folded Mia40 at known concentration. The ^1H resonance at −0.7 ppm of Ile 53 $H\gamma$ was used as a marker of the folded conformation of Mia40. Total Mia40 was determined on the same cell extracts by western blot analysis, by using the same pure Mia40 sample at increasing dilutions as a reference; Grx1 and Trx1 concentrations in the cell extracts were determined as above by using samples of purified Grx1 and Trx1 as references. Mia40 was stained with a rabbit polyclonal anti-Mia40 antibody (Abcam: ab87033, diluted to 0.5 μg/ml); Grx1 with a rabbit polyclonal anti-glutaredoxin 1 antibody (Abcam: ab45953, 1.0 μg/ml); Trx1 with a rabbit polyclonal anti-thioredoxin/TRX antibody (Abcam: ab26320, 0.4 μg/ml). Goat anti-rabbit IgG (whole molecule)-peroxidase secondary antibody (Sigma: A0545) was used for detection, diluted at 1:160,000 (for Mia40) or 1:80,000 (for Grx1 and Trx1). For checking the purity of the cytoplasmic and mitochondrial fractions, antibodies against a cytoplasmic marker (rabbit polyclonal anti-GAPDH antibody, Abcam: ab9485) and a mitochondrial marker (rabbit polyclonal anti-COX IV, Abcam: ab16056) were used.

U-^{15}N-labeled Mia40 for in vitro experiments and unlabeled Mia40 for NMR quantifications were produced as previously described (Banci et al., 2009b). Glutaredoxin 1 for in vitro interaction with Mia40 was produced as follows: a pTH34 vector containing the human Grx1 gene (N-term fused with His-tag and TEV recognition site) was transformed in *E. coli* BL21(DE3) gold competent cells. Cells were grown at 37°C in minimal medium to an optical density 0.6 and then induced with 0.5 mM IPTG for 16 hr at 25°C. Glutaredoxin 1 was purified by affinity chromatography using a nickel-chelating HisTrap (GE Healthcare) column. After digestion with AcTEV protease (Invitrogen) O/N at 25°C, the protein was separated from the affinity tag in a HisTrap column. The sample buffer was then exchanged with 50 mM potassium phosphate, 0.5 mM EDTA, pH = 7. Pure thioredoxin 1 for western blot analysis was prepared from HEK293T cells overexpressing Trx1. Cells were lysed in 20 mM potassium phosphate buffer (pH 7) and the cleared extract was incubated at 70°C for 4 min. After centrifugation (30 min at 16,000 × *g*), Trx1 was purified by anionic exchange chromatography using a HiTrap DEAE FF (Amersham Biosciences) column applying a linear gradient of potassium phosphate buffer (10–100 mM, pH 7).

Mia40 for mass spectrometry analysis was isolated from cell extracts by ionic exchange chromatography using a HiTrap DEAE FF (Amersham Biosciences) column applying a linear gradient of NaCl (0–500 mM NaCl in 20 mM TRIS buffer, pH 8). For both in vitro and cell extracts samples, the buffer was exchanged with 0.1% trifluoroacetic acid/ 50% acetonitrile by Zip Tip_{C18} tips (MILLIPORE). Mass spectrometry analysis was performed at a Bruker Ultraflex III MALDI TOF/TOF instrument, using an α-cyano-4-hydroxycinnamic acid matrix.

For thiol alkylation reaction with 4-acetamido-4′-maleimidylstilbene-2,2′-disulfonic acid (AMS), either cell or protein samples were precipitated with 10% trichloroacetic acid, washed with acetone, and resuspended in 100 mM Tris pH 7 + 2% SDS. The mixtures were then incubated 1 hr at 37°C with 20 mM AMS, and run on a nonreducing SDS-PAGE.

NMR experiments were acquired with a 950 MHz Bruker Avance III spectrometer equipped with a CP TCI CryoProbe. 1D ^1H and 2D $^1H,^{15}N$-SOFAST-HMQC (Schanda and Brutscher, 2005) spectra were acquired at 305K. The total acquisition time for each cell sample ranged from 1 to 2 hr. The supernatant of each cell sample was checked for protein leakage in the same experimental conditions. The same NMR spectra were also acquired on the cell extracts. Cell viability before and after NMR experiments was assessed with trypan blue staining (Freshney, 1987). Cell viability remained above 90% because damaged cells ranged from 3% before the experiments to 8% after the experiments. 2D $^1H,^{15}N$-SOFAST-HMQC NMR spectra of cell samples and extracts were processed with Bruker Topspin 3.1 software by subtracting a spectrum of untransfected cells/cell extract acquired within the same experimental conditions, thereby eliminating the interference of signals arising from unspecific ^{15}N labeling.

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures and can be found with this article online at http://dx.doi.org/10.1016/j.chembiol.2013.05.007.

ACKNOWLEDGMENTS

This work was supported by the Programmi di Ricerca di Rilevante Interesse Nazionale (PRIN; 2009FAKHZT_001 "Biologia strutturale meccanicistica: avanzamenti metodologici e biologici"), the Italian "Rete Nazionale per lo studio della Proteomica Umana (Italian Human ProteomeNet)" FIRB PROTEOMICA MIUR -RBRN07BMCT, and Ente Cassa di Risparmio di Firenze ("Biologia Strutturale Integrata").

Received: January 5, 2013
Revised: May 8, 2013
Accepted: May 17, 2013
Published: June 20, 2013

REFERENCES

Aricescu, A.R., Lu, W., and Jones, E.Y. (2006). A time- and cost-efficient system for high-level protein production in mammalian cells. Acta Crystallogr. D Biol. Crystallogr. *62*, 1243–1250.

Banci, L., Bertini, I., Cefaro, C., Ciofi-Baffoni, S., Gallo, A., Martinelli, M., Sideris, D.P., Katrakili, N., and Tokatlidis, K. (2009a). MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria. Nat. Struct. Mol. Biol. *16*, 198–206.

Banci, L., Bertini, I., Ciofi-Baffoni, S., and Tokatlidis, K. (2009b). The coiled coil-helix-coiled coil-helix proteins may be redox proteins. FEBS Lett. *583*, 1699–1702.

Banci, L., Bertini, I., Cefaro, C., Cenacchi, L., Ciofi-Baffoni, S., Felli, I.C., Gallo, A., Gonnelli, L., Luchinat, E., Sideris, D.P., and Tokatlidis, K. (2010). Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import. Proc. Natl. Acad. Sci. USA *107*, 20190–20195.

Banci, L., Barbieri, L., Bertini, I., Luchinat, E., Secci, E., Zhao, Y., and Aricescu, A.R. (2013). Atomic-resolution monitoring of protein maturation in live human cells by NMR. Nat. Chem. Biol. *9*, 297–299. http://dx.doi.org/10.1038/nchembio.1202.

Chacinska, A., Pfannschmidt, S., Wiedemann, N., Kozjak, V., Sanjuan Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., and Pfanner, N. (2004). Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins. EMBO J. *23*, 3735–3746.

Chacinska, A., Guiard, B., Müller, J.M., Schulze-Specking, A., Gabriel, K., Kutik, S., and Pfanner, N. (2008). Mitochondrial biogenesis, switching the sorting pathway of the intermembrane space receptor Mia40. J. Biol. Chem. *283*, 29723–29729.

Crowley, P.B., Chow, E., and Papkovskaia, T. (2011). Protein interactions in the Escherichia coli cytosol: an impediment to in-cell NMR spectroscopy. ChemBioChem *12*, 1043–1048.

Dungen, R., Wang, Q., Ceh Pavia, E., Grant, C.M., and Lu, H. (2012). Cytosolic thioredoxin system facilitates the import of mitochondrial small Tim proteins. EMBO Rep. *13*, 916–922.

Freshney, R. (1987). Culture of Animal Cells: A Manual of Basic Technique (New York: Alan R. Liss).

Gabriel, K., Milenkovic, D., Chacinska, A., Müller, J., Guiard, B., Pfanner, N., and Meisinger, C. (2007). Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. J. Mol. Biol. *365*, 612–620.

Grumbt, B., Stroobant, V., Terziyska, N., Israel, L., and Hell, K. (2007). Functional characterization of Mia40p, the central component of the disulfide relay system of the mitochondrial intermembrane space. J. Biol. Chem. *282*, 37461–37470.

Hofmann, S., Rothbauer, U., Mühlenbein, N., Baiker, K., Hell, K., and Bauer, M.F. (2005). Functional and mutational characterization of human MIA40 acting during import into the mitochondrial intermembrane space. J. Mol. Biol. *353*, 517–528.

Holmgren, A. (1979). Reduction of disulfides by thioredoxin. Exceptional reactivity of insulin and suggested functions of thioredoxin in mechanism of hormone action. J. Biol. Chem. *254*, 9113–9119.

Holmgren, A. (1989). Thioredoxin and glutaredoxin systems. J. Biol. Chem. *264*, 13963–13966.

Inomata, K., Ohno, A., Tochio, H., Isogai, S., Tenno, T., Nakase, I., Takeuchi, T., Futaki, S., Ito, Y., Hiroaki, H., and Shirakawa, M. (2009). High-resolution multi-dimensional NMR spectroscopy of proteins in human cells. Nature *458*, 106–109.

Longen, S., Bien, M., Bihlmaier, K., Kloeppel, C., Kauff, F., Hammermeister, M., Westermann, B., Herrmann, J.M., and Riemer, J. (2009). Systematic analysis of the twin cx(9)c protein family. J. Mol. Biol. *393*, 356–368.

Lu, H., Allen, S., Wardleworth, L., Savory, P., and Tokatlidis, K. (2004). Functional TIM10 chaperone assembly is redox-regulated in vivo. J. Biol. Chem. *279*, 18952–18958.

Mesecke, N., Terziyska, N., Kozany, C., Baumann, F., Neupert, W., Hell, K., and Herrmann, J.M. (2005). A disulfide relay system in the intermembrane space of mitochondria that mediates protein import. Cell *121*, 1059–1069.

Meyer, Y., Buchanan, B.B., Vignols, F., and Reichheld, J.P. (2009). Thioredoxins and glutaredoxins: unifying elements in redox biology. Annu. Rev. Genet. *43*, 335–367.

Naoé, M., Ohwa, Y., Ishikawa, D., Ohshima, C., Nishikawa, S.I., Yamamoto, H., and Endo, T. (2004). Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space. J. Biol. Chem. *279*, 47815–47821.

Neupert, W., and Herrmann, J.M. (2007). Translocation of proteins into mitochondria. Annu. Rev. Biochem. *76*, 723–749.

Ogino, S., Kubo, S., Umemoto, R., Huang, S., Nishida, N., and Shimada, I. (2009). Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, streptolysin O. J. Am. Chem. Soc. *131*, 10834–10835.

Reckel, S., Hänsel, R., Löhr, F., and Dötsch, V. (2007). In-cell NMR spectroscopy. Prog. Nucl. Magn. Reson. Spectrosc. *51*, 91–101.

Reckel, S., Lopez, J.J., Löhr, F., Glaubitz, C., and Dötsch, V. (2012). In-cell solid-state NMR as a tool to study proteins in large complexes. ChemBioChem *13*, 534–537.

Sagemark, J., Elgán, T.H., Bürglin, T.R., Johansson, C., Holmgren, A., and Berndt, K.D. (2007). Redox properties and evolution of human glutaredoxins. Proteins *68*, 879–892.

Sakai, T., Tochio, H., Tenno, T., Ito, Y., Kokubo, T., Hiroaki, H., and Shirakawa, M. (2006). In-cell NMR spectroscopy of proteins inside Xenopus laevis oocytes. J. Biomol. NMR *36*, 179–188.

Schanda, P., and Brutscher, B. (2005). Very fast two-dimensional NMR spectroscopy for real-time investigation of dynamic events in proteins on the time scale of seconds. J. Am. Chem. Soc. *127*, 8014–8015.

Selenko, P., Frueh, D.P., Elsaesser, S.J., Haas, W., Gygi, S.P., and Wagner, G. (2008). In situ observation of protein phosphorylation by high-resolution NMR spectroscopy. Nat. Struct. Mol. Biol. *15*, 321–329.

Tokatlidis, K. (2005). A disulfide relay system in mitochondria. Cell *121*, 965–967.

Ueno, H., Kajihara, H., Nakamura, H., Yodoi, J., and Nakamuro, K. (2007). Contribution of thioredoxin reductase to T-cell mitogenesis and NF-kappaB DNA-binding promoted by selenite. Antioxid. Redox Signal. *9*, 115–121.

Watson, W.H., Pohl, J., Montfort, W.R., Stuchlik, O., Reed, M.S., Powis, G., and Jones, D.P. (2003). Redox potential of human thioredoxin 1 and identification of a second dithiol/disulfide motif. J. Biol. Chem. *278*, 33408–33415.

Supplemental Information

Visualization of redox-controlled protein fold in living cells

Lucia Banci, Letizia Barbieri, Enrico Luchinat, Erica Secci

Supplementary Figures:

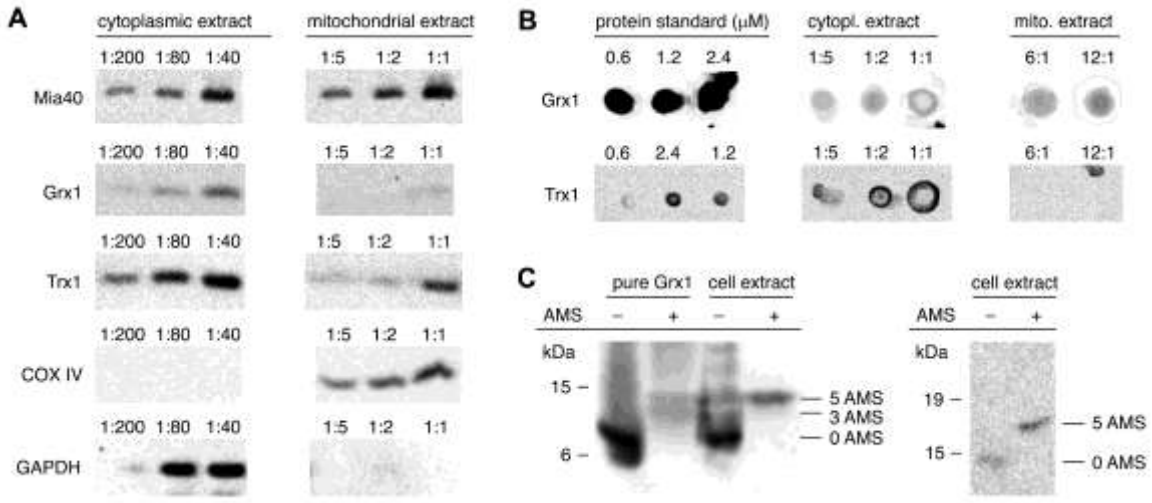


Figure S1, related to Figure 1: Overexpressed Mia40, Grx1 and Trx1 are mostly localized in the cytoplasm; Grx1 and Trx1 are fully reduced. (A) Cytoplasmic and mitochondrial fractions of cells were blotted at increasing dilutions to assess the intracellular distribution of overexpressed Mia40, Grx1 and Trx1. Dilutions are expressed relative to the most concentrated mitochondrial extract. The subcellular fractions were obtained using a mitochondria isolation kit for cultured cells (Thermo Scientific). The fraction purity was checked with a mitochondrial and a cytoplasmic marker (COX IV and GAPDH, respectively). (B) Endogenous levels and subcellular localization of Grx1 and Trx1 were measured by dot blot analysis of cytoplasmic and mitochondrial extracts. Pure samples at known concentration were blotted as a reference. (C) AMS thiol alkylation reaction on a cell extract containing overexpressed Grx1 was detected by Western Blot, showing that all Grx1 cysteines are reduced. Purified Grx1 was run as a reference. The same reaction was performed on a cell extract containing Trx1 overexpressed in presence of fully active TrxR, showing that all Trx1 cysteines are reduced.

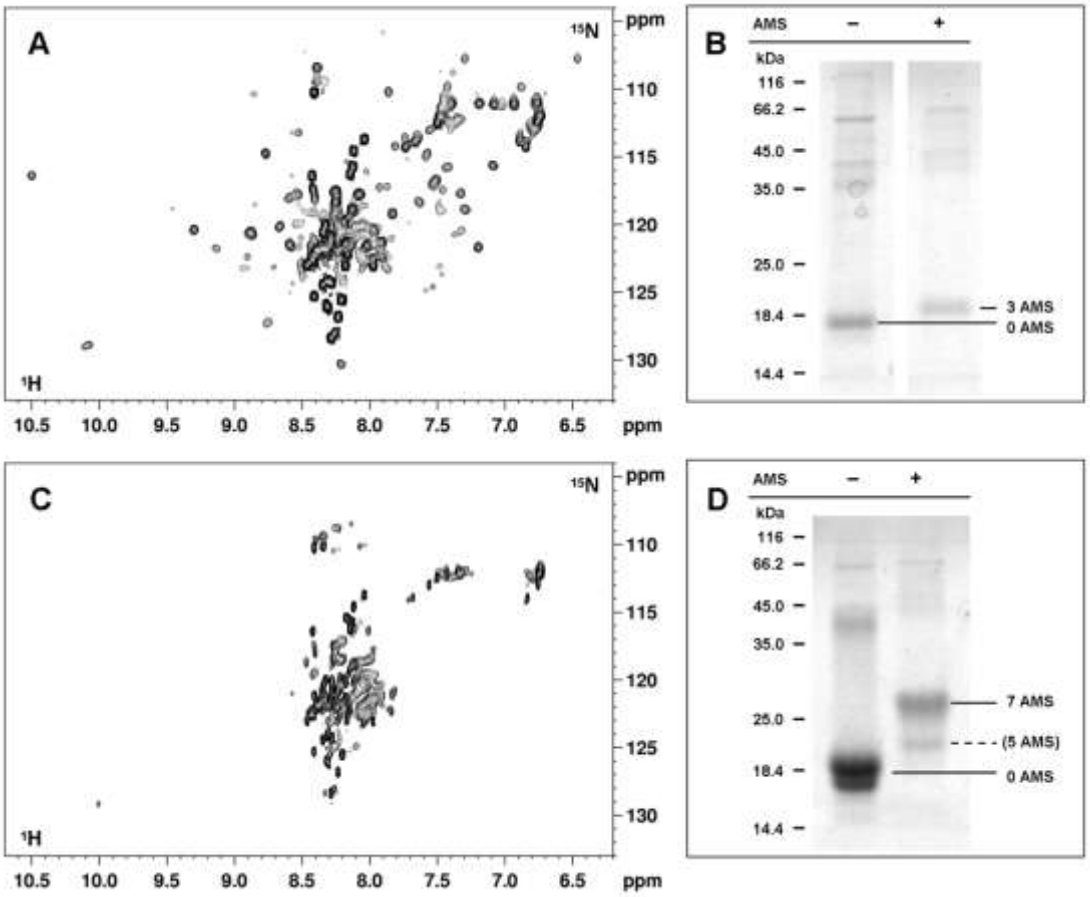


Figure S2, related to Figure 2: Oxidized, folded Mia40 is reduced *in vitro* upon heat denaturation in reducing conditions. ^{1}H-^{15}N SOFAST HMQC spectra were acquired on pure U-^{15}N Mia40$^{2S\text{-}S}$ (A) before and (C) after denaturation at 95°C in presence of 10 mM GSH. After denaturation, the crosspeaks of the N- and C-terminal unfolded segments of Mia40 are still visible (with sharper lines), while those corresponding to the core segment of oxidized Mia40 disappear. AMS reaction was performed on Mia40 samples before and after denaturation, which were run on Coomassie-stained SDS-PAGE (B, D), confirming the reduction of the two structural disulfide bonds.

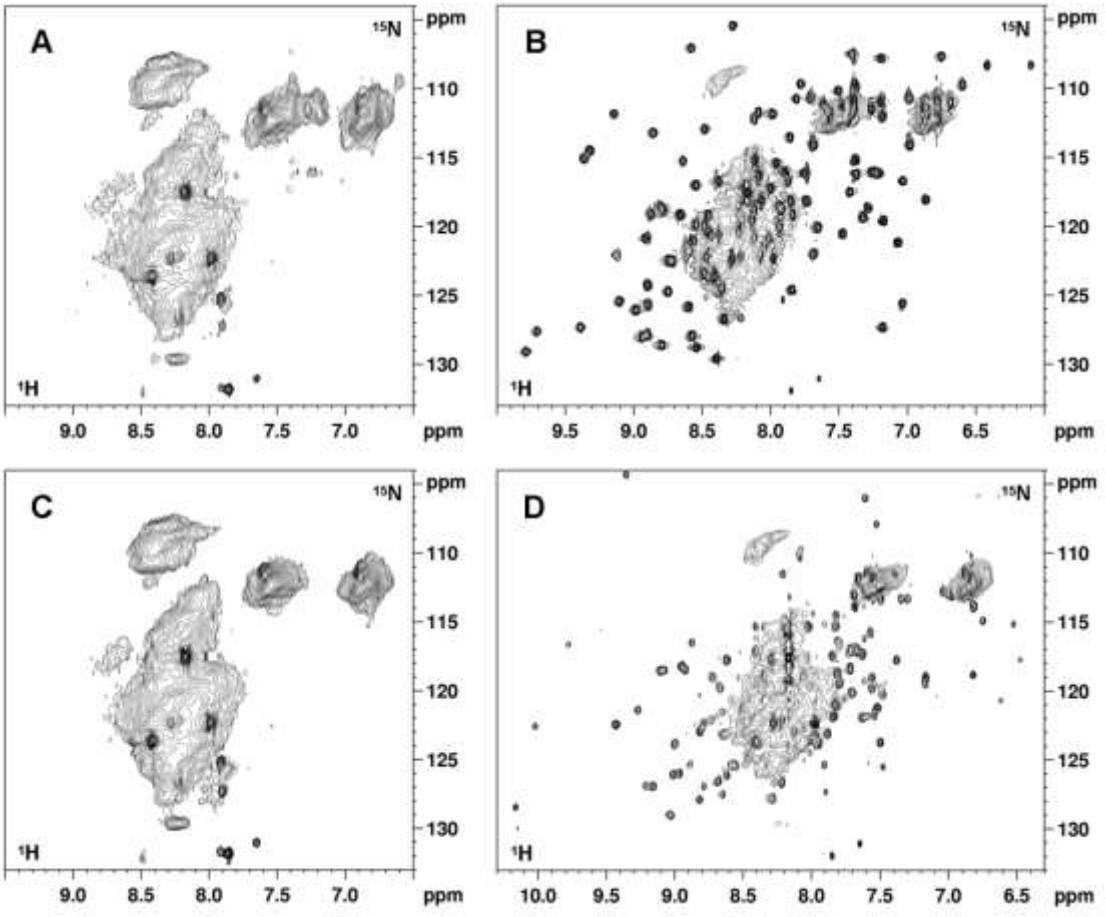


Figure S3, related to Figure 3: The amide signals of cytoplasmic Grx1 and Trx1 are broadened beyond detection. ^{1}H-^{15}N SOFAST-HMQC spectra were acquired on samples of human cells expressing (A) Grx1 and (C**)** Trx1, and on the corresponding cell extracts (B, D). Only the cellular background signals were visible in the spectra of intact cells, where the amide crosspeaks of Grx1 and Trx1 are broadened beyond detection. Upon cell lysis, the crosspeaks of both proteins were detected.

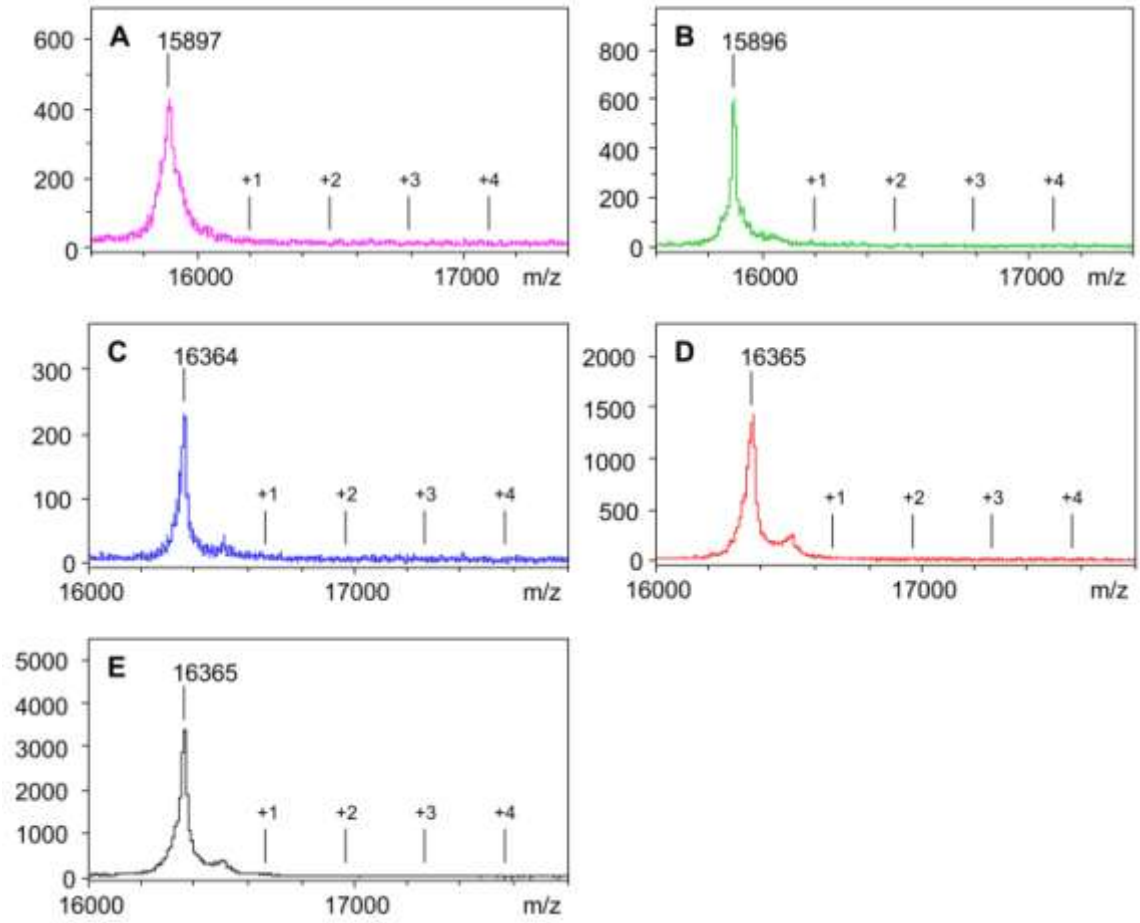


Figure S4, related to Figure 3: Mass spectrometry analysis shows absence of glutathionylated Mia40. MALDI TOF/TOF analysis of Mia40 isolated from cell extracts (native Mia40) and Mia40 *in vitro* samples (GSFT-Mia40). (A) Extract of cells overexpressing Mia40 alone; (B) extract of cells co-expressing Mia40 and Grx1; (C) *in vitro* Mia40 heat denatured in 20 mM GSH (phosphate buffer, pH 7), after removal of GSH; (D) same sample as (C), incubated with 2 eq. of Grx1 in 20 mM GSH, after removal of GSH. Points in the spectra corresponding to increasingly glutathionylated forms of Mia40 are indicated (+305.3 Da for each GSH molecule).

2.5 Profilin1 results:

The synthetic gene encoded for profilin1 protein was purchased from *eurofins* company. It was cloned in the pHLsec vector between the restriction enzymes Eco RI and XhoI. ^{1}H-^{15}N SOFAST-HMQC spectra on uniformly ^{15}N labelled cells over-expressing profilin1 were acquired. Protein signals were not detected and only became visible upon cell lysis, suggesting that some interactions occur in the cytoplasm which makes the proteins tumbling slower on average, thus broadening the amide crosspeaks beyond detection. The same behaviour was observed for the mutant H120E. The other ALS mutants (C71G, M114T and G118V) did not reach an enough protein expression level in the soluble fraction and none NMR spectra were acquired.

Since this protein interacts with phosphoinositides inside cells, the residues in this binding domain should be mutated in order to minimize this interactions and to try to make it visible by in-cell NMR experiments. Mutations (K70N, R89A, K91E, R136A and R137D) were introduced in the gene of the wild-type protein and test expression was positive for all constructs.

The wild-type and the mutated gene was cloned in some pET vectors in order to express them in *E. Coli* cells, to purify the proteins and to study their behaviour *in vitro*. pET vectors were chosen to obtain or the native or the histidine tagged proteins to facilitate the purification. For all the constructs the exploited restriction enzymes were NdeI and XhoI. The purified sample was produced as follow: a massive culture of the histidine tagged protein was grown in minimal medium M9 at 37°C for 8h (for WT and H120E) or at 25°C O/N (for M114T and G118V). The cleared lysate was loaded in a nickel-chelating HisTrap column for an affinity chromatography. After digestion with Factor Xa protease O/N at 25°C, the protein was separated from the affinity tag in a HisTrap column. The purity of the sample was checked with SDS-PAGE.

3 Methodological Aspects

3.1 Cell lines

Human cells, adherent HEK (Human Embryonic Kidney) 293T cells, were maintained in DMEM medium supplemented with l-glutamine, antibiotics (penicillin/streptomycin) and 10% foetal bovine serum (FBS) and were incubated at 310 K, 5% CO_2 in a humidified atmosphere. HEK293T cells have some advantages like to be easy-handling, have robust growth rate, excellent transfectability, high capacity for recombinant protein expression and low-cost media requirements. They are available from all major cell banks and support repeated passages (twice a week for at least four months) without losing the above properties.

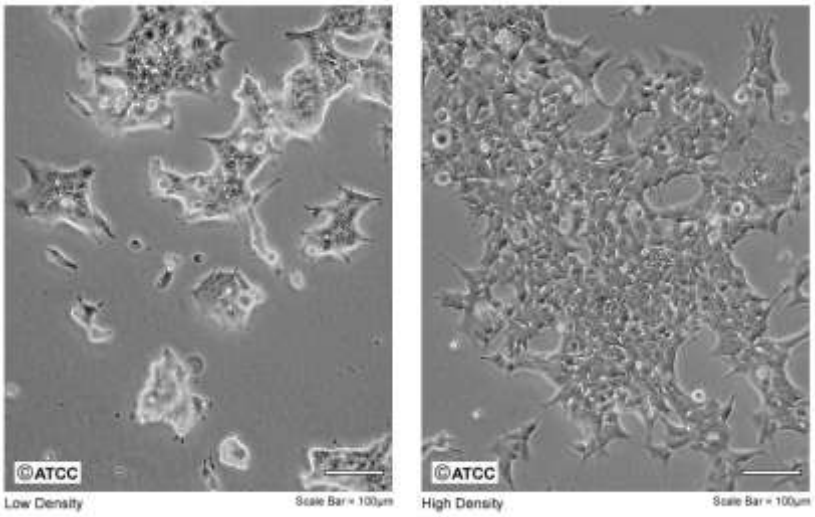


Figure 10: HEK293T cells. Reprinted from ATCC web site

3.2 Gene cloning

The genes of all proteins were cloned in the pHLsec plasmid between the restriction enzymes Eco RI and XhoI. The pHLsec plasmid contains a secretion signal sequence for the expressed protein secretion in the medium. The choice of these two particular restriction enzymes allows the elimination of this signal sequence and thus the cytoplasmatic protein expression. Mutated proteins were obtained by site-directed mutagenesis. All the DNA clones were confirmed by DNA sequencing.

DNA was purified using the Endotoxin-Free Plasmid Maxi Kit (Macherey-Nagel). High-quality DNA is essential for successful transfection and only samples with an OD_{260}/OD_{280} ratio of 1.8 or higher are suitable. The DNA samples must be sterile; therefore, care should be taken to wash the DNA precipitates properly with 70% ethanol before dissolving them in sterile water.

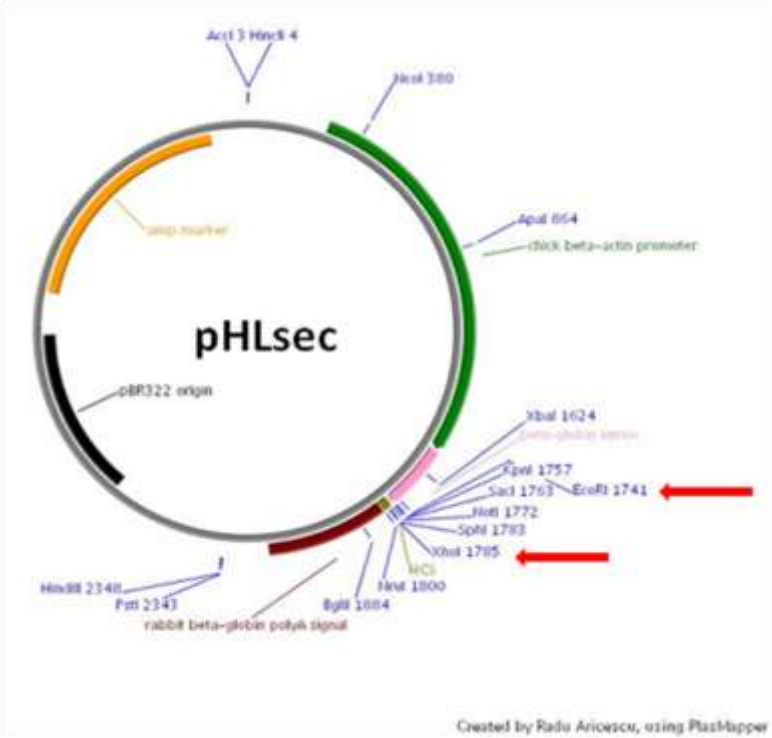


Figure 11: pHLsec vector

3.3 DNA transient transfection in human cells

DNA transient transfection differentiate itself from stable transfection because the DNA does not integrate into the host genome. A number of delivery reagents are available. Polyethylenimine (PEI) is a very affordable and highly efficient transfection reagent and there is now a significant number of publications reporting its use (Kichler, 2004; Demeneix & Behr, 2005). A large selection of PEI forms are available, varying in molecular weight and branching. The employed PEI is '25 kDa branched', which was found to be most effective in transfecting various cells lines (Durocheret al.2002). Stock solutions are made in water, first at 100 mg ml^{-1}(PEI is an extremely viscous liquid). Once

the solution is homogeneous, it is further diluted to 1 mg ml^{-1}, the pH adjusted to 7, filter sterilized and aliquoted. PEI is a polycationic reagent that binds the negatively charged DNA and promotes the binding of the complex to anionic proteoglycans that are present on the cell surface. Through endocytosis, the PEI/DNA complex enters into the cells, reach the nucleus where the DNA is released.

Optimal transfection is achieved when adherent cells reach about 90% confluency. For a standard 75 cm^2 flask, 25 µg of plasmid DNA are required. The DNA is added to 2,5 mL of serum-free medium and is followed by the addition of 50 µg of PEI in serum-free medium and gently mixed. The solution is incubated for 20 min at room temperature to allow DNA/PEI complex formation. During complex formation, media from the flask to be transfected should be changed, lowering the serum concentration to 2%. Finally, the DNA/PEI complex is added to the flask, which is then briefly rotated to allow mixing and the cells are placed in the incubator. For co-expression of two proteins, the optimal DNA ratio between the two construct was chosen.

3.4 Protein expression

During protein expression, cells are incubated at 310 K in 75 cm^2 flasks. DMEM medium is used for unlabelled in-cell NMR samples; BioExpress6000 medium (CIL) is used for uniform ^{15}N labelling; for selective ^{15}N-cysteine labelling, a reconstituted medium is prepared following the DMEM reported composition, in which ^{15}N-cysteine is added together with all the other unlabelled components.

When metals are needed, Zn(II) is supplemented as $ZnSO_4$, which is added to the expression medium to a final concentration of 10 µM immediately after transfection. Cu(II) is supplemented as $CuCl_2$, added to a final concentration of 100 µM after 48 hours of protein expression, and incubated for 24 hours. To ensure complete activation of Trx1, sodium selenite is supplemented to the cell culture medium to a final concentration of 100 nM, starting 24 hr before transfection.

3.5 NMR sample preparation

Samples for in-cell NMR were prepared following a reported protocol with some variations: HEK293T cells from a 75 cm^2 culture flask were detached with trypsin-EDTA 0,05% and resuspended in 20 mL DMEM containing 10% FBS to inactivate trypsin. Cells were gently centrifuged (800 g), resuspended in 10 mL PBS, washed once with PBS and resuspended in one cell pellet volume of DMEM medium supplemented with 90 mM glucose, 16 mM HEPES buffer, 10% D_2O. The cell suspension was transferred to a 3 mm Shigemi NMR tube; the glass plunger was not used.

Cleared cell lysates for in-cell NMR experiments were prepared as follows: after the experiments, cells were resuspended in the supernatant, removed from the NMR tube and spun down again to collect the medium for the protein-leakage test. Cells were lysed by the freeze-thaw method after being suspended in one pellet volume of PBS buffer supplemented with 0.5 mM EDTA and 4-(2-aminoethyl)-benzenesulfonyl fluoride hydrochloride (AEBSF). The lysate was centrifuged at 16,000 × g for 60 min at 4 °C, and the cleared cell extract was collected for NMR and SDS-PAGE analysis.

3.6 NMR experiments

NMR experiments on human HEK293T cells were acquired using a 950-MHz Bruker Avance III spectrometer equipped with a CP TCI CryoProbe. 1D ^{1}H and 2D ^{1}H-^{15}N SOFAST HMQC spectra were acquired at 305 K. The total acquisition time for each cell sample ranged from 1 to 2 h.

NMR spectra on *E. Coli* cell samples were acquired at 305 K using an 800-MHz Bruker Biospin spectrometer equipped with a TXI CryoProbe. The total acquisition time for each cell sample ranged from 1 to 4 h.

- 1D ^{1}H

The aliphatic region of ^{1}H NMR spectra (-1 to 1 p.p.m.) were used to check the protein over-expression over the cellular background, obtained with cells transfected with empty

vector. This spectrum region is free from cellular background signals, hence the methyl groups of the proteins are easily detectable.

For the study of the metallation state of hSOD1 protein, a 1D ^{1}H NMR spectrum with a larger spectral window (0-16 p.p.m.) was utilised to allow the detection of histidines signals, since their resonance frequencies are shifted at higher p.p.m., due to the metal binding.

- 2D ^{1}H-^{15}N SOFAST-HMQC

Fast data acquisition in SOFAST-HMQC is realized by using very short inter-scan delays. SOFAST-HMQC combines the advantages of a small number of radio-frequency pulses, Ernst-angle excitation, and longitudinal relaxation optimization to obtain an increased signal to noise ratio for high repetition rates of the experiment.

The main features of SOFAST-HMQC are the following: (i) the HMQC-type H–X transfer steps require only few radio-frequency pulses which limits signal loss due to B1-field inhomogeneities and pulse imperfections. A reduced number of radio-frequency pulses become important if the experiment is performed on a cryogenic probe, where B1-field inhomogeneities are more pronounced. (ii) The band-selective ^{1}H pulses reduce the effective spin-lattice relaxation times (T1) of the observed ^{1}H spins. The presence of a large number of non-perturbed ^{1}H spins, interacting with the observed ^{1}H via dipolar interactions (NOE effect), significantly reduces longitudinal relaxation times and the equilibrium spin polarization is more quickly restored. (iii) The adjustable flip angle of the ^{1}H excitation pulse allows further enhancement of the available steady-state magnetization for a given recycle delay The performance of SOFAST-HMQC critically depends on the choice of the pulse shapes for the band-selective excitation and refocusing pulses on the ^{1}H channel. [Schanda and Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds, Journal of Biomolecular NMR (2005) 33: 199–211]

3.7 Assessment of cell viability

The viability of the cells before and after each in-cell NMR experiment is checked by trypan blue staining. Live cells with intact cell membranes are not coloured by the dye. Since cells are very selective in the compounds that pass through the membrane, in a viable cell trypan blue is not absorbed; however, it traverses the membrane in a dead cell. Hence, dead cells are shown as a distinctive blue colour. The number of stained cells was counted under phase-contrast microscope and the cell viability always remained above 90%.

3.8 Western Blotting

Western blot is an analytical technique used to detect and analyse specific proteins in a given sample of cells or cellular extract. The sample is subjected to gel electrophoresis for protein separation and the proteins are then immobilized on a nitrocellulose membrane following electrophoretic transfer from the gel. Non-protein binding areas on the membrane are blocked to prevent non-specific binding of antibodies. The membrane is incubated with a primary antibody that specifically binds to the protein of interest. Unbound antibodies are removed by washing and a secondary antibody conjugated to an enzyme, a fluorophore or an isotope is used for detection. The detected signal from the protein:antibody:antibody complex is proportional to the amount of protein on the membrane. The most commonly used method for detection is chemiluminescence, based on secondary antibodies conjugated with horseradish peroxidase enzyme. On the addition of a peroxide-based reagent, the enzyme catalyses the oxidation of luminol resulting in the emission of light. The light signal can be captured either using a charge-coupled device (CCD) camera-based imager or by exposure to X-ray film. Quantitative protein analysis was achieved by densitometry using ImageJ program.

3.9 *E. Coli* samples preparation

Cell samples for in-cell NMR were prepared as follows. For uniform ^{15}N labelling, BL21(DE3) Gold *E. Coli* strain was used, transformed with a pET28a plasmid containing the WT hSOD1 gene sequence without any additional tag. For selective ^{15}N-cysteine

labelling, the auxotroph strain BL21(DE3) CysE was transformed with a pET21 plasmid containing the same WT hSOD1 gene sequence. A cell culture was grown overnight at 30°C in 35 mL of LB medium. After gentle centrifugation (3000 g) for 20 minutes, the cells were re-suspended in 50 mL of M9 minimal medium [M9 buffer (7 g/L K_2HPO_4, 3 g/L KH_2PO_4, 0.5 g/L NaCl, pH 7.4), 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 1 mg/L biotin, 1 mg/L thiamine, antibiotic] containing 1 g/L $(^{15}NH_4)_2SO_4$ and 3 g/L of unlabelled glucose to obtain an OD_{600} of ~1.6. For expression of selective ^{15}N-cysteine labelled WT hSOD1 the samples were prepared as described above, but a M9-based reconstituted medium was used, containing ^{15}N-cysteine and the other 19 unlabelled amino acids. After 10 min recovery time, over-expression was induced with 0.5 mM IPTG, and carried out at 30°C for 4 h. The cells were washed once with 50 mL of metal-free M9 buffer in order to remove nutrients, metal ions and any excreted by-product, and they were harvested through gentle centrifugation. The pellet was then re-suspended in metal-free M9 buffer until 500 µL of a ~50% v./v. cell slurry were obtained. 50 µL of D_2O were added, and the final volume was put in a 5 mm NMR tube.

Cleared cell lysates for in-cell NMR experiments were prepared as follows: after removal of the supernatant to be checked by NMR, the cell pellet was re-suspended in an equal volume of metal-free M9 buffer. The cells were then lysed by ultrasonication. Then the lysate was centrifuged at 16000 x g for 20 minutes, the supernatant was collected and its volume was brought to 500 µL with M9 buffer. 50 µL of D_2O were then added. The final dilution of the cytoplasm in M9 buffer is around 1:2.

3.10 Protein purification

- hSOD1

  Pure hSOD1 protein was prepared as follows: a cell culture (BL21(DE3) Gold (Stratagene), transformed with a pET28a plasmid containing the WT hSOD1 gene) was grown overnight at 30°C in 750 mL LB, harvested and re-suspended in 2.25 L ^{15}N-labelled M9 medium. After 4 h from induction with 0.5 mM IPTG at 30°C the cells were harvested and re-suspended in 20 mM Tris, pH 8 buffer for lysis. The cleared lysate was loaded on an anion exchange column (DEAE Sepharose Fast Flow

resin, GE Healthcare) for a first purification of hSOD1 by elution with NaCl gradient. The collected fractions containing hSOD1 (checked by SDS-PAGE) were further purified by gel filtration (Superdex75 16/60 column, GE Healthcare) in 20 mM Tris, 100 mM NaCl, pH 8 buffer. Fractions containing pure hSOD1 were collected. $ZnSO_4$ was added to the protein solution to increase hSOD1 stability, and 1 mM DTT was added in all buffers to prevent protein aggregation through disulfide bridges. The removal of the metals to obtain E,E-hSOD1$^{S\text{-}S}$ was achieved by dialyzing several times a diluted solution of hSOD1 against 10 mM EDTA in 50 mM acetic acid at pH 3.5. After removal of EDTA E,Zn-hSOD1$^{S\text{-}S}$ and Zn,Zn-hSOD1$^{S\text{-}S}$ were then obtained by adding at pH 5.5 one and two equivalents of $ZnSO_4$, respectively. To obtain E,E-hSOD1$^{SH\text{-}SH}$ and E,Zn-hSOD1$^{SH\text{-}SH}$, the E,E-hSOD1$^{S\text{-}S}$ and E,Zn-hSOD1$^{S\text{-}S}$ were incubated 1 h at 37°C with 50–60 mM of DTT; 1 mM EDTA was added to the sample of E,E-hSOD1$^{SH\text{-}SH}$ to prevent binding of any metal present in traces. DTT concentration was then brought to 2 mM by dialysis against oxygen-free phosphate buffer. Zn,Zn-hSOD1$^{SH\text{-}SH}$ was obtained by adding $ZnSO_4$ in excess (4 equivalents) to a sample of reduced E,E-hSOD1$^{SH\text{-}SH}$. The final NMR samples obtained were in 20 mM phosphate buffer at pH 7.5.

- hMia40

Pure hMia40 protein was prepared as follows: a cell culture of BL21(DE3) Gold (Stratagene) transformed with a pDEST-His-MBP (Addgene) plasmid containing the hMia40 gene, was grown overnight at 30°C in 15 mL LB, harvested and re-suspended in 1.5 L of ^{15}N-labelled M9 medium. Protein expression was induced with 0.7 mM IPTG for 16h at 25°C. Cells were harvested and re-suspended in 20 mM Tris, 500 mM NaCl, 5 mM Imidazole, pH 8 buffer for the lysis by ultrasonication. The cleared lysate was then processed for the purification. It was performed using a HiTrap chelating HP column (Amersham Pharmacia Biosciences) charged with Ni(II). The protein was eluted with the 20 mM Tris, 500 mM NaCl, 300 mM Imidazole, pH 8, buffer. His-MBP tag was cleaved with AcTEV protease (Invitrogen) O/N at 25°C, and separated from the N-terminal His-MBP domain with a second purification step. A gel filtration using a HiLoad 26/60 Superdex 75 (Amersham Pharmacia Bioscience) column was

further required to obtain a high purity (>95%) protein sample. The sample buffer was then exchanged with 50 mM potassium phosphate, 0.5 mM EDTA, 1 mM DTT, pH=7 to perform NMR experiments.

- hGlutaredoxin1

hGlutaredoxin1 for *in vitro* interaction with hMia40 was produced as follows: a pTH34 vector containing the human Grx1 gene (N-term fused with His-tag and TEV recognition site) was transformed in *E. coli* BL21(DE3) Gold competent cells. A cell culture was grown overnight at 30°C in 10 mL LB, harvested and re-suspended in 1L of minimal medium M9. When $O.D._{600}$ reached 0.6, protein expression was induced with 0.5 mM IPTG for 16 h at 25 °C. hGlutaredoxin1 was purified by affinity chromatography using a nickel chelating HisTrap (GE Healthcare) column. The fractions containing pure hGrx1 (checked by SDS-PAGE) were digested with AcTEV protease (Invitrogen) O/N at 25 °C. The protein was separated from the affinity tag in a HisTrap column. The sample buffer was then exchanged with 50 mM potassium phosphate, 0.5 mM EDTA, pH=7 to perform NMR experiments.

4 Conclusions

In conclusion, in this research project, cellular and biological strategies were developed in order to optimize protein expression in prokaryotic and eukaryotic cells. Functional processes, such as protein folding, metal-uptake and cysteine oxidation, were investigated through in-cell NMR spectroscopy, that is a technique that allows the characterization of biological molecules, their interactions and their functional processes in their native environment at atomic level. To obtain human proteins over-expression in human cells, a new approach with respect to those already used by few other researchers, was developed. Procedures to address and either overcome or reduce the observed technological limitations were created. In particular, the functional processes involving the maturation steps of SOD1 protein and of ALS-linked SOD1 mutants and the folding of Mia40 prior the import in the inter-membrane space of mitochondria, also when Glutaredoxin1 and Thioredoxin1 were co-expressed, were studied.

In-cell NMR spectroscopy has demonstrated to be a useful and suitable technique to study functional processes involving proteins in their physiological environment and should be applied to other proteins implicated in human diseases.

5 References

[1]Selenko P, Wagner G, Looking into live cells with in-cell NMR spectroscopy, Journal of Structural Biology 158 (2007) 244-253

[2]Dedmon M.M., Patel, C.N., Young, G.B., Pielak, G.J., 2002, FlgM gains structure in living cells, PNAS 99, 12681-12684

[3]Hubbard J.A., MacLachlan, L.K., King, G.W., Jones, J.J., Fosberry, A.P., 2003, Nuclear magnetic resonance spectroscopy reveals the functional state of the signalling protein CheY *in vivo* in *Escherichia coli*. Mol. Microbiol. 49, 1191-1200

[4]McNulty, B.C., Young, G.B., Pielak, G.J., 2006, Macromolecular crowding in the *Escherichia coli* periplasm maintains alpha-synuclein disorder, J. Mol. Biol. 355, 893-897

[5]Serber, Z., Straub, W., Corsini, L., Nomura, A.M., Shimba, N., Craik, C.S., Ortiz de Montellano, P., 2004; Dötsch V., 2004, Methyl groups as probes for proteins and complexes in in-cell NMR experiments, J. Am. Chem. Soc. 126, 7119-7125

[6]Wieruszeski, J.M., Bohin, A., Bohin, J.P., Lippens, G., 2001, In vivo detection of the cyclic osmoregulated periplasmic glucan of *Ralstonia solanacearum* by high-resolution magic angle spinning NMR, J. Magn. Reson., 151, 118-123

[7]Bryant, J.E., 2006, In-cell protein dynamics, Mol. Biosyst. 2, 406-410

[8]Bryant, J.E., Lecomte, J.T., Lee, A.L., Young, G.B., Pielak, G.J., 2005 Protein dynamics in living cells, Biochemistry 44, 9275-9279

[9]Burz, D.S., Dutta, K., Cowburn, D., Shekhtman, A., 2006, Mapping structural interactions using in-cell NMR spectroscopy (STINT-NMR) , Nat. Methods 3, 91-93

[10]Reardon, P.N., Spicer, L.D, 2005, Multidimensional NMR spectroscopy for protein characterization and assignment inside cells, J. Am. Chem. Soc. 127, 10848-10849

[11]Etezady-Esfarjani, T., Herrmann, T., Horst, R., Wuthrich, K., 2006, Automated protein NMR structure determination in crude cell-extract, J. Biomol. NMR 34, 3-11

[12]Sakakibara, D., Sasaki, A., Ikeya, T., Shirakawa, M., Güntert, P., Ito, Y., 2009, Protein structure determination in living cells by in-cell NMR spectroscopy, Nature Letters 5;458(7234):102-5

[13]Reckel S., Löhr F., Dötsch V., In-cell NMR Spectroscopy ChemBioChem 2005, 6, 1601-1606

[14]Selenko, Serber, Gadea, et al 2006 Quantitative NMR analysis of the protein G B1 domain in *Xenopus laevis* egg extracts and intact oocytes. PNAS 103, 11904-11909

[15]Selenko, P., Frueh, D.P., Elsaesser, S.J., Haas, W., Gygi, S.P., Wagner, G., 2008, *In situ* observation of protein phosphorylation by high-resolution NMR spectroscopy, Nat. Struct. Mol. Biol. 15, 321-329

[16]Bodart, J.F., Wieruszeski, J.M., Amniai, L., Leroy, A., Landrieu, I., Rousseau-Lescuyer, A., Vilain, J.P., Lippens, G., 2008, NMR observation of Tau in *Xenopus* oocytes, J. Magn. Reson. 192, 252-257

[17]Pielak et al., Protein Nuclear Magnetic Resonance under Physiological Conditions Biochemistry 2009, 48, 226-234

[18]Trantìrek et al., Investigation of quadruplex structure under physiological conditions using In-cell NMR, Top Curr Chem (2012)

[19]Shirakawa et al., High-resolution multi-dimensional NMR spectroscopy of proteins in human cells 2009 Nature Letters

[20]Ogino, Kubo, Umemoto et al., 2009 Observation of NMR signals from proteins introduced into living mammalian cells by reversible membrane permeabilization using a pore-forming toxin, Streptolysin O. Journal of the American Chemical Society 131, 10834

[21]Ito and Selenko, Cellular structural biology, Current opinion in structural biology 2010, 20:640-648

[22]Reckel, Hänsel, Löhr, Dötsch, In-cell spectroscopy, Progress in Nuclear Resonance Spectroscopy 51 (2007) 91-101

[23]McCord, J.M., Fridovich, I., "Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein)," Journal of Biological Chemistry, vol. 244, no. 22, pp. 6049–6055, 1969

[24]Kawamata, H., Manfredi, G., Import, Maturation, and Function of SOD1 and Its Copper Chaperone CCS in the Mitochondrial Intermembrane Space, Antioxid Redox Signal. 2010 Nov 1;13(9):1375-84

[25]Tainer, J.A.; Getzoff, E.D.; Richardson, J.S.; Richardson, D.C. (1983) Nature 306, 284-287

[26]Culotta, V.C.; Klomp, L.W.; Strain, J.; Casareno, R.L.; Krems, B.; Gitlin, J.D. (1997) J.Biol.Chem. 272, 23469-23472

[27]Portnoy, M.E.; Schmidt, P.J.; Rogers, R.S.; Culotta, V.C.; (2001) Mol.Genet.Genomics 265, 873-882

[28]Gamonet, F., Lauquin, G.J.M. (1998) Eur.J.Biochem. 251, 716-723

[29]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

[30]Hall, L.T., Sanchez, R.J., Valentine, J.S., Hart, P.J. et al., (2000) Biochemistry 39, 3611-3623

[31]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

[33]Casareno, R.L.; Waggoner, D.J.; Gitlin, J.D. (1998) J.Biol.Chem. 273, 23625-23628

[33]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

[34]Rowland, L.P.; Shneider, N.A. (2001) N. Engl. J. Med. 344, 1688–700

[35]Bruijn, L.I.; Miller, T.M.; Cleveland D.W (2004) Annu. Rev. Neurosci. 27, 723–49

[36]Mulligan VK & Chakrabartty A (2013) Protein misfolding in the late-onset neurodegenerative diseases: Common themes and the unique case of amyotrophic lateral sclerosis. Proteins: Struct., Funct., Bioinf: Vol 81,8,1285-1303.

[37]Ross CA & Poirier MA (2004) Protein aggregation and neurodegenerative disease. Nature Medicine 10(7):S10-S17

[38]Chiti F & Dobson CM (2006) Protein misfolding, functional amyloid, and human disease. Annu. Rev. Biochem, Annual Review of Biochemistry), Vol 75, pp 333-366

[39]Bucciantini M, *et al.* (2002) Inherent toxicity of aggregates implies a common mechanism for protein misfolding diseases. Nature 416(6880):507-511

[40]Lindberg MJ, Tibell L, Oliveberg M (2002) Common denominator of Cu/Zn superoxide dismutase mutants associated with amyotrophic lateral sclerosis: Decreased stability of the apo state. Proc. Natl. Acad. Sci. USA 99(26):16607-16612

[41]Furukawa Y, Kaneko K, Yamanaka K, O'Halloran TV, & Nukina N (2008) Complete loss of post-translational modifications triggers fibrillar aggregation of SOD1 in the familial form of amyotrophic lateral sclerosis. J. Biol. Chem. 283(35):24167-24176

[42]Durer ZAO, *et al.* (2009) Loss of Metal Ions, Disulfide Reduction and Mutations Related to Familial ALS Promote Formation of Amyloid-Like Aggregates from Superoxide Dismutase. PLoS One 4(3)

[43]Banci L, *et al.* (2007) Metal-free superoxide dismutase forms soluble oligomers under physiological conditions: a possible general mechanism for familial ALS. Proc. Natl. Acad. Sci. USA 104(27):11263-11267

[44]Banci L, *et al.* (2008) SOD1 and Amyotrophic Lateral Sclerosis: Mutations and Oligomerization. PLoS One 3(2)

[45]Wang J, *et al.* (2002) Fibrillar inclusions and motor neuron degeneration in transgenic mice expressing superoxide dismutase 1 with a disrupted copper-binding site. Neurobiol. Dis. 10(2):128-138

[46]Kerman A, *et al.* (2010) Amyotrophic lateral sclerosis is a non-amyloid disease in which extensive misfolding of SOD1 is unique to the familial form. Acta Neuropathologica 119(3):335-344

[47]Wang J, *et al.* (2005) Somatodendritic accumulation of misfolded SOD1-L126Z in motor neurons mediates degeneration: alpha B-crystallin modulates aggregation. Hum. Mol. Genet. 14(16):2335-2347

[48]Tokatlidis, K., 2005, Cell. 121(7):965-7

[49]Neupert, W., Herrmann, J.M., 2007, Translocation of proteins into mitochondria, Annu Rev Biochem. 2007;76:723-49.

[50] Sideris, D.P., Petrakis, N., Katrakili, N., Mikropoulou, D., Gallo, A., Ciofi-Baffoni, S., Banci, L., Bertini, I., Tokatlidis, K., 2009, A novel intermembrane space-targeting signal docks cysteines onto Mia40 during mitochondrial oxidative folding, J Cell Biol. 28;187(7):1007-22

[51]Chacinska, A., Pfannschmidt,S., Wiedemann, N., Kozjak, V., Sanjuán Szklarz, L.K., Schulze-Specking, A., Truscott, K.N., Guiard, B., Meisinger, C., Pfanner, N., 2004 Essential role of Mia40 in import and assembly of mitochondrial intermembrane space proteins, EMBO J. 23:3735–3746

[52]Terziyska, N., Grumbt, B., Bien, M., Neupert, W., Herrmann, J.M., Hell, K., 2007, The sulfhydryl oxidase Erv1 is a substrate of the Mia40-dependent protein translocation pathway, FEBS Lett. 2007 Mar 20;581(6):1098-102

[53]Naoe´, M., Ohwa, Y., Ishikawa, D., Ohshima, C., Nishikawa, S., Yamamoto, H., Endo, T., 2004, Identification of Tim40 that mediates protein sorting to the mitochondrial intermembrane space, J. Biol. Chem. 279:47815–47821

[54]Banci, L., Bertini, I., Cefaro, C., Cenacchi, L., Ciofi-Baffoni, S., Felli, I.C., Gallo, A., Gonnelli, L., Luchinat, E., Sideris, D., Tokatlidis, K., 2010, Molecular chaperone function of Mia40 triggers consecutive induced folding steps of the substrate in mitochondrial protein import, Proc Natl Acad Sci U S A. 23;107(47):20190-5

[55]Gabriel, K., Milenkovic, D., Chacinska, A., Müller, J., Guiard, B., Pfanner, N., Meisinger, C., 2007, Novel mitochondrial intermembrane space proteins as substrates of the MIA import pathway. J. Mol. Biol. 365:612–620

[56]Müller, J.M., Milenkovic, D., Guiard, B., Pfanner, N., Chacinska. A., 2008, Precursor oxidation by Mia40 and Erv1 promotes vectorial transport of proteins into the mitochondrial intermembrane space, Mol. Biol. Cell. 19:226–236

[57]Grumbt et al., 2007, Functional characterization of Mia40p, the central component of the disulfide relay system of the mitochondrial intermembrane space, J Biol Chem 282:37461–37470

[58]Hell, K., 2007, The Erv1–Mia40 disulfide relay system in the intermembrane space of mitochondria, Biochimica et Biophysica Acta 1783 (2008) 601–609

[59] Banci, L., et al., 2009, MIA40 is an oxidoreductase that catalyzes oxidative protein folding in mitochondria, Nat Struct Mol Biol16:198–206

[60]Terziyska, N., Grumbt, B., Bien, M., Neupert, W., Herrmann, J.M., Hell, K., 2007, FEBS Letters 581 1098–1102

[61]Oakley et al. NMR Investigation of the binding between human profilin1 and inositol 1,4,5-triphosphate, the soluble headgroup of phosphatidylinositol 4,5-bisphosphate, Biochemistry (2008) 47,51

[62]Paavilainen, V.O; Bertling, E; Falck, S; Lappalainen, P; Regulation of cytoskeletal dynamics by actin-monomer-binding proteins, Trends in Cell Biology, Volume 14, Issue 7, July 2004, Pages 386–394

[63]Witke, The role of profilin complexes in cell mobility and other cellular processes, TRENDS in Cell Biology

[64]Metzler et al., 1995, Refined solution structure of human profilin1, Protein Science, 4, 450-459

[65]Landers et al., 2012, Mutations in the profilin1 gene cause familial amyotrophic lateral sclerosis, Nature Letter

[66]Weishaupt et al., 2012, A novel phosphorylation site mutation in profilin1 revealed in a large screen of US, Nordic, and German amyotrophic lateral sclerosis/frontotemporal dementia cohorts, Neurobiology of Aging