University of Florence

International Doctorate in Structural Biology
Cycle XVIII (2003-2006)


# Superoxide Dismutase fALS mutants: insights into molecular mechanisms of aggregation

Ph.D. thesis by

**Elisa Libralesso**

Tutor
Prof. Paola Turano

Coordinator
Prof. Claudio Luchinat

# Contents

# 1. Introduction

## 1.1 Structural biology and non-native protein states

Evidences about connection between familial aggregation and many neurodegenerative disorders were in some cases found years before the molecular genetic or biochemical properties characterizing the disease were known. When the identification of specific mutations in previously unknown genes directed the attention to proteins and pathways typical of the pathogenesis of such diseases, it became clear that only a thoroughly characterization of the aberrant species and the study of their interactions could give the key to find a cure for the diseases themselves.

The number of proteins related to neurodegenerative disorders is quite high; among them are the prion proteins, involved in Creutzfeld– Jacob disease, Bovine spongiform encephalopathy (BSE) Kuru and some others, Huntingtin involved in Huntington disease, α-Synuclein involved in Parkinson disease, $Cu_2$,$Zn_2$-Superoxide Dismutase involved in amyotrophic lateral sclerosis etc[1]. It is quite common that such disorders seem to involve two forms, a so called familial (rare) and a seemingly nonfamilial (common) one, sometimes referred to as “sporadic” or “idiopathic,” despite the increasing evidence of a significant influence of genetic factors also in these cases[2,3].

They share as a common trait the finding of protein deposits distributed in different tissues, from brain like Alzheimer’s disease to central and peripheral nervous system like Neuronal intranuclear inclusion disease, to almost all organs and tissues, like Light chain associated amyloidosis, to upper and lower motor neurons, like ALS[4-8]. In all the cases, relatively unfolded and highly flexible conformations, structurally different from both the molten globule-like and the completely unfolded state, are crucial species in the fibrillation process, which is usually characterized by a first phase of protofibrillation. Many studies evidenced that such pre-fibrillar aggregates appear to be substantially more toxic than mature amyloid fibrils or plaques and their toxicity seems a result of the intrinsic ability of said precursors to impair fundamental cellular processes, via interaction with cellular membranes[9,10].

## 1.2 Aim of the research

The topic of the research carried on during the three years of this Doctorate course was the characterization of non native states of the enzyme Superoxide Dismutase, (SOD hereafter) mainly focused on the study of some mutants related to familial Amyotrophic Lateral Sclerosis.
Despite the increasing amount of evidences about this neurodegenerative disorder, obtained *by in vivo* and *in vitro* studies on the properties of many disease-related mutants[11-17], the pathological mechanism of fALS-associated SOD remains an elusive target. A powerful approach for the understanding of the disorder is the *in vitro* characterization of the involved aberrant proteins, and it has been heavily exploited in the past years by many different research groups, with most of the structural studies on recombinant fALS mutants related to the determination of their X-ray structures[15,18]. In some cases the findings were very interesting, evidencing a packing of the crystal cells that resembled a tubular superstructure[13], but very little is known about the intrinsic mechanisms that lead to the formation of such fibrillar bodies. Since it is becoming clearer that discrete structural instability and aberrant dynamic properties of these mutants are likely to play an essential role in the etiology of the disease[15-17,17,19], and almost no data are available in this sense put aside a complete dynamical characterization of mutant G93A[11], we decided to exploit a similar approach, although widened to include also *in vitro* characterization of the fibrillation step.
Given the need of *in solution* studies of the enzymes, in order to ascertain their dynamical properties, High Field Nuclear Magnetic Resonance (NMR hereafter) was the first technique of choice, since it is well established and many studies have been published, correlating spectral data with structural and dynamic protein characteristics. Moreover, it is favoured by the possibility, offered by expression in systems like *E. coli* or yeast, to easily achieve labelling with NMR optical isotopes (e.g. $^{15}N$, $^{13}C$, $^{2}H$) and provides detailed information about structural and dynamical features of enzymes, since it allows their characterization at atomic level. This is fundamental for the detection of protein regions that could constitute a sort of initiating nucleus for aggregation, usually characterized by fast dynamics, and also for the screening of the polypeptide behaviour in different environments, allowing also the characterization in partially and completely denaturing solution.
The need to characterize also aggregated states of the mutant protein, in order to link the dynamical evidences with the oligomerization process, shifted our attention towards a suitable solution technique for the study of supermolecular compounds, out of the range of molecular weight allowed by standard High Field NMR methods. Fast Field Cycling Relaxometry (FFCR

hereafter) seemed very suitable for the targeted studies and it was first tested on a well know system, namely the folded and unfolded forms of the monomeric non pathogenic variant of SOD (E133QM2) [20]. When applied to the study of soluble aggregation it resulted decidedly suitable, since it reports the water relaxation rates of protein solutions as a function of the applied magnetic field, so allowing the detection of soluble oligomers and the determination of their molecular weights.

A drawback of both NMR and FFCR is their need for quite concentrated samples, that in case of NMR often require an isotopic labelling. Relating to this, the ideal partner for such techniques is Circular Dichroism, that allows a quick, preliminary characterization on the overall folding and stability of a protein, since it employs diluted samples and permits the investigation of the polypeptide stability over a wide range of temperatures, if coupled wit a Peltier unit. This is crucial for the investigation of the unfolding pathway of aberrant proteins, since evidences tell that non-hierarchical, partially folded states of the polypeptidic chain could be triggering the fibrillation process and circular dichroism allows the rough determination of the main structural features of such intermediates, that sometimes are formed in conditions not fully compatible with the requirements of the NMR spectrometry. The drawback of CD is that it provides relatively low resolution structural information, and although it can yield reasonably reliable estimates of the secondary structure content of a protein, in terms of the proportions of α-helix, β-sheet and possibly β-turns, these are overall figures that do not associate regions of the polypeptide with their structural type. This is a result that only thoroughly characterization *via* High Field NMR can achieve.

Among the large number of mutations related to fALS, the criterion of selection for the aberrant proteins to study was oriented so to have an high variability in the involved structural motifs of the enzyme, in order to ascertain the influence of the mutation locus on the dynamic and aggregating properties of the polypeptide. Expression was performed in *E. coli*, for no post translational modification of the expressed protein was needed, and previous expression of the non mutated homologous of the aberrant enzymes gave good results both regarding yield and further purification[21]. As a first approach, the mutations were introduced into the gene encoding *as*SOD (thermostable human $Cu_2Zn_2$-SOD C6A C111S double mutant), since it presents complete structural analogy with the wild type form while being easier to handle thanks to the reversibility of its thermal unfolding pathway, due to the lack of free thiolic side chains[22,23]. Upon choosing S134N as the most promising mutant for structural studies, primarily for its high aggregating tendency and its diminished stability, said mutation was introduced also in the *wt*SOD expressing plasmid, in order to validate the choice of *as*SOD fALS mutants as a

meaningful model for the *in vitro* study of fALS related mutant SODs, and to assess also the free cysteins contribution to the fibrillation process.

Mutants were fully characterized by high field NMR and NMRD spectroscopies, the most appropriate tools to link oligomerization mechanisms to changes in their structural and dynamical features in solution. Being the stability strictly related to fALS mutants aggregating behaviour[24], unfolding studies were also performed, by mean of thermal denaturation both in presence and in absence of 2,2,2-trifluoroethanol (TFE hereafter), a denaturing agent often used to verify the proteins tendency to give rise to amyloid structures[25,26]. In order to better assess the differences between the fALS mutants and native SOD, comparisons were made with the non mutated homologous of the aberrant enzymes, as well as with the engineered monomeric, non-pathogenic form of SOD (mutant E133QM2, discussed in section "an overview on SOD").

The researches carried on during the course of this PhD have been necessarily a team group and although I gave my contribution in each of the involved topics, I primarily focused my attention in the field of protein expression and CD characterization.

## 1.3 An overview on SOD

Copper zinc superoxide dismutase is an essential primary antioxidant,t ubiquitous enzyme, mainly found in the cytosol of eukaryotic cells, although it is present also in their nucleus, peroxisomes, and mitochondrial intermembrane space. It catalyzes the disproportion of superoxide ion to hydrogen peroxide and molecular oxygen, in the cytoplasm[27,28].

$O^{-}{}_2$ + Cu(II)ZnSOD → O2 + Cu(I)ZnSOD  Reaction 1.

$O^{-}{}_2 + 2H^+$ + Cu(I)ZnSOD → $H_2O_2$ + Cu(II)ZnSOD  Reaction 2.

$2O^{-}{}_2 + 2H^+ \rightarrow O_2 + H_2O_2$  Net reaction.

$Cu_2Zn_2$SOD catalyzes this reaction very efficiently, since at physiological pH diffusion can nearly be considered the rate determining step. Activity is almost independent of pH over the range of 5.0 to 9.5 for the holo enzyme[14].

Superoxide is a highly active radical ion, formed in the mitochondria at the end of the respiratory chain, due to partial reduction of $O_2$ to water: a phenomenon that occurs in aerobic organisms in controlled amounts and under physiological conditions[29-31]. It belongs to the so called "Reactive

Oxygen Species" (ROS), that comprise also ions such as the hypochlorite ion, non charged molecules like hydrogen peroxide and radicals like the hydroxyl radical, the most reactive of them all (Halliwell B, Gutteridge JMC: Free Radicals in Biology and Medicine 3rd edition. Oxford: Oxford Science; 1998). Among all these molecular species, the life spans of $OH^{\cdot}$ and $O_2^{-\cdot}$ are the shortest. Although it is believed that $O_2^{-\cdot}$ doesn't directly oxidizes biomolecules such as lipids, proteins, etc, the interaction with metal ions such as iron catalyzes its reduction to $OH^{\cdot}$ through a Fenton reaction. Superoxide reacts also with nitric oxide (NO) to quench physiological activities such as vascular relaxation, generating peroxinitrite ($ONOO^{-}$) that causes oxidative damage too[32 33]. Oxidative stress can be defined as a disturbance in the pro-oxidant–antioxidant balance in favour of the former, leading to potential damage, while oxidative damage is the biomolecular harm caused by direct attack of reactive species during oxidative stress[34]. It can lead to adaptation of the cell or organism by upregulation of defence systems; to cell injury, involving damage to any or all molecular targets such as lipids, DNA, protein, carbohydrate, etc, or in the worst of cases to cell death by apoptosis or necrosis [35] $Cu_2Zn_2$-SOD can thus be regarded as one of the anti-oxidative stress mechanisms the body developed in order to protect itself from highly toxic ROS and their byproducts. Such mechanisms are localized in tissues and inside the cells where ROS are generated. If the amount of ROS exceeds, for any reason, the limit of the defence mechanism of the body, serious disease may be induced.

The comparison of X-ray crystal structures of SOD proteins from many different species, that have been solved mostly in the fully metallated state, show that the structure is highly conserved[36]. The human isoenzyme is composed of two identical subunits, each of them weighting 16-kDa, and characterized by a β-barrel fold, composed of eight antiparallel strands, right-handedly twisted and connected by seven turns and loops numbered in sequence order. I and V are short, β-hairpin connections, while the remaining ones are longer. Strands 1, 2, 3 and 6 form the first sheet, while the second is delimited by strands 4, 5, 7 and 8. Odd loops are located on the opposite side of the barrel with respect to region involved in the subunit-subunit interface, while the even loops are partly located at the intersubunits surface, with loops IV and VI regions involved in the subunit-subunit interactions, together with the N terminus, and the strand 1, and the C terminus and the strand 8. Hydrophobic interactions hold together the two monomers by mean of numerous interactions, specifically main-chain to main-chain hydrogen bonds, water-mediated hydrogen bonds, and hydrophobic contacts. The extension of the interested regions can be invoked to understand the high stability of the dimer towards both thermal and chemical unfolding[37]. The metal binding region, which is completely encased within each single subunit, contains a Zn structural ion and a Cu catalytic ion, responsible for the two-steps catalytic dismutation of the superoxide ion. It is enclosed by loop VII, spanning from residue 121 to 144,

(aka electrostatic loop) and loop IV, called Zn loop since it contains almost all the Zn ligands. Loop IV is the longest one, comprising residues from 49 to 83, and bears also the Cys57, which forms the highly conserved disulfide bridge with Cys146, located in strand 8.

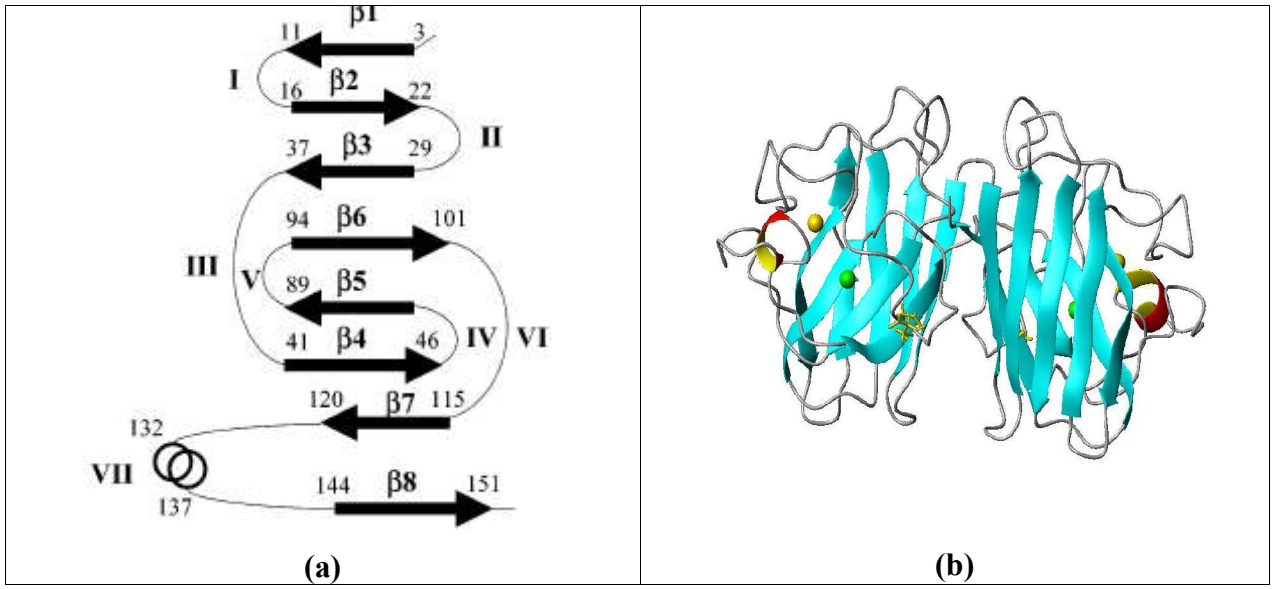



**Figure 1:** SOD **a:** secondary structure; **b:** tertiary structure

The active site is located between the electrostatic and the Zn loops, on the outer wall of the β-barrel and facing strands 5, 6 and 7. It consists of a 15 Å deep and 12 Å long cavity, that narrows to 3 Å in the proximity of the Cu centre in order to enhance the selectivity of the enzyme towards superoxide ion itself. Such channel hosts charged residues, disposed in an alternate fashion, that create an electric field gradient providing the optimal electrostatic potential for the diffusion of the substrate to the reaction site. At the rim of the channel, on the protein surface, lie the two positively charged lysine residues that probably cover a role in attracting anions and guiding them into the crevice. In addition, on the lining of the channel the positively charged side chain of arginine 143 probably interacts with superoxide when it's bound to the copper[38,39].

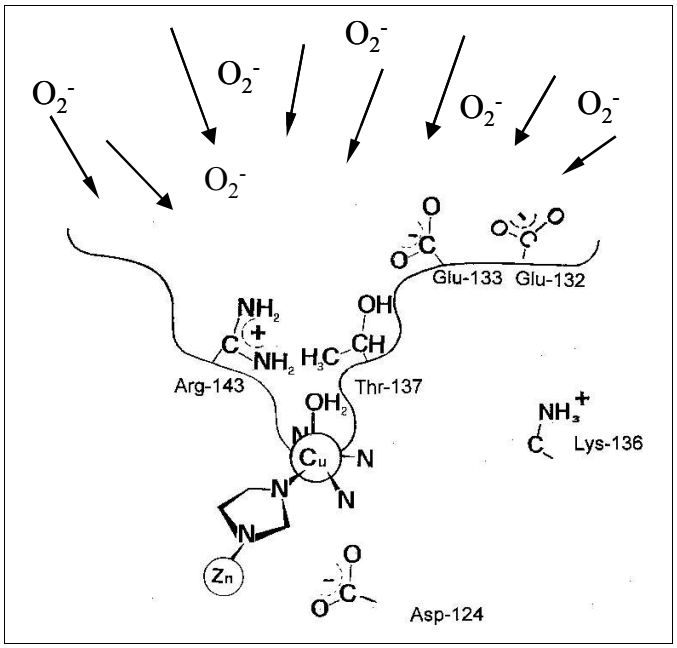



**Figure 2:** Schematic diagram of a cross section of the active site channel in SOD

The Zn (II) ion is completely buried within the protein, and it's coordinated in an approximately tetrahedral geometry by one Aspartic acid (83) and three Histidine residues (80, 71, 63). The oxidized form of the enzyme sees the Cu coordinated by three histidyl side chains, belonging to residues 46, 48 and 120, plus His63 acting as a bridging ligand, and a weakly bound water molecule, with an overall distorted square pyramidal geometry. Upon reduction, the bridging imidazolate ligand is protonated in position $N\varepsilon_2$ and releases the Cu centre, that loses also the water molecule and ends up in a nearly trigonal planar configuration[40,41].

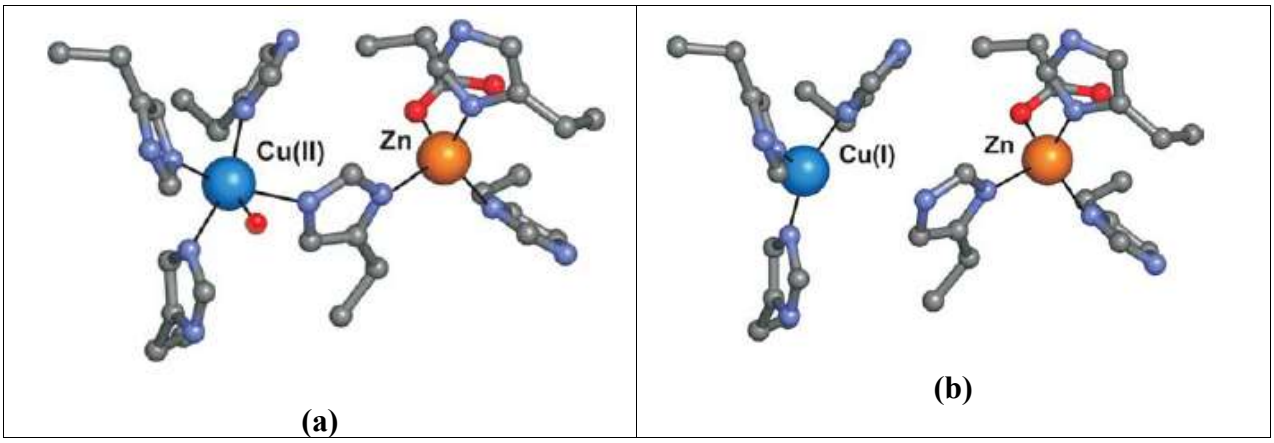



**Figure 3:** Cu and Zn coordination **a:** in the oxidized form; **b:** in the reduced form

Stability of the metal binding region is enhanced by an extensive H-bond which links important portions of the protein to the metal-binding sites, starting from residue 124 (Asp ) that by binding the nonliganding nitrogens of His46 (Cu ligand) and His71 (Zn ligand) helps the active site to maintain its correct orientation. Furthermore, His120 (Cu ligand) is bound to the carbonyl oxygen of Gly141, a residue belonging to the electrostatic loop, while His48 is linked, through hydrogen bond with the carbonyl oxygen of Gly61, to the guanidinium group of the same Arg143 previously named as a catalytically important residue. The latter is also hydrogen bonded to the carbonyl oxygen of Cys57, the one involved in the intrasubunit disulfide bond, which belongs to the loop VI that participates in the formation of the dimer interface[42]. A sequence alignment of SODs from different organisms evidences that 23 out of the 153 amino acids composing each subunit are invariant. They are almost all located in the proximity of the active site, comprising – besides the Cu and Zn ligands - also forementioned Arg 143 and Asp 124, which are involved in the hydrogen bond network that stabilizes the metal binding region, and Pro 66 which stabilizes the overall conformation by rigidifying the zinc loop. Other conserved aminoacids located in crucial regions of the protein are Gly 51 and 114 at the dimeric interface and the previously named Cys 57 and Cys 146, involved in the intramolecular disulfide bridge, which is likely to impose the righteous conformation to a region of the protein also involved in the dimerization. Last but not least, Gly 16, Gly 147, Phe 45 Leu 38 and Leu 106 are likely to contribute in the stabilization of the greek-key β barrel fold[12].

One of the most commonly employed SOD mutant is the C6A, C111S thermostable variant (*as*SOD hereafter), known since the early 90's, that retains the conserved disulfide bond but sees the two free cysteines, a buried one (Cys6) located in a beta-strand, and a solvent accessible one, (Cys111) located in a loop region, mutated respectively to an Ala and a Ser. When expressed and purified from yeast, the mutant enzyme has normal specific activity and its relative resistance to irreversible inactivation of enzymatic activity by heating at 70 degrees is greater with respect to wild type. The mutagenesis of at least one of the free cysteine residues increases thermostability and the double mutant is more stable to irreversible thermal inactivation than either of the single mutants, presumably due to the removal of the reactive thiols[23,43]. *as*SOD structure well superposes on the wild-type one even at high resolution and no significant deviations are observed, safe those at the mutation site, that being local do not have influence on the overall folding[44]. This enzyme isoform reveals to be extremely useful in the field of study of fALS mutants aggregating properties, since it allows to distinguish the direct effects of the FALS mutations from indirect changes resulting from free cysteine mediated oxidation and aggregation. The only drawback could be represented by the fact that C111 has been proposed as an alternative binding residue in the fALS-related H46R SOD mutant[45], but its absence from all the

characterized mammalian SOD, with the only exception of human and chicken, suggests that its role in the biochemistry of the enzyme isn't critical.

The degree of *in vivo* metallation of SOD is mostly unknown and possibly varies depending on the different cell compartments and tissues SOD is found in. Some portion of the SOD protein is normally present in cells in a copper-deficient state, but less evidence can be found about the level of zinc bound to wild-type SOD[46-48]. It was reported that when feeding rats wit a Cu deficient diet, the isolated SOD was partly apo and partly fully metallated, indicating that there may be a connection between the in vivo metallation process of both metal ions[49]. In addition, recent findings demonstrated that in vitro completely immature form of yeast SOD lacking copper, zinc, and the conserved disulfide bond is most efficiently taken up by mitochondria[50].

Since its early characterization, it has been evident that the loss of metal ions, especially the structural zinc, significantly decreases the otherwise high thermodynamic stability of SOD[51]. To study the stability of wild-type SOD in dependence of its metal content, many differently metallated forms of SOD have been studied during the years, aiming to assess the individual contribution of metals to the protein dynamical behaviour[52-55]

In order to structurally characterize such SOD variants by NMR, a monomeric analog was initially engineered by mutating two residues at the dimer interface (Phe50 and Gly51) with Glu's. This substitution disrupts the hydrophobic interactions at the interface, yielding as a final result the protein monomerization. Such enzyme showed reduced activity, probably due to the distortion of the electrostatic loop, which forms one side of the active site channel and is fundamental in determining the optimal electrostatic gradient to drive the superoxide anions to the catalytic site[56]. For this reason, further engineering of the polypeptide was performed by replacing Glu133 with Gln[21], a mutation which showed to increase enzymatic activity also in dimeric SOD[57]. Despite the structural changes induced on the active site coordination geometry by such mutation, the enzymatic activity results partially restored, with an increased affinity of the polypeptide for anions, including superoxide. This analog (E133QM2 hereafter) was the first monomeric SOD showing substantial activity. Although the remodelling and exposure of the dimer interface to solvent was accompanied by the distortion of the active site geometry, with an increases in the axiality of the copper chromophore and of the Cu-$OH_2$ distance, this mutant has demonstrated a good, relatively low-molecular weight model suitable for accurate High Field NMR[53,54,58,58,59]. Characterization of the apo E133QM2 SOD confirmed that the removal of the Zn ion increases the mobility of the electrostatic and zinc loops, that appear largely unstructured in this variant[59], while they resulted well ordered in the monomeric fully metallated form. The high mobility of these structural elements isn't devoid of implications for the overall subunit fold, since the latter relies significantly on an extensive network of H-bonds between the two loops

and several residues in the $\beta$-barrel, as already observed in the dimeric wild-type protein. Conversely, the Cu-free monomeric SOD undergoes smallest changes, involving mainly structure and mobility of copper binding site residues.

Subsequently, extensive use of triple resonances NMR experiments coupled with deuteration schemes, have been employed to successfully afford the assignment and the structure and dynamics characterization of the dimeric protein (MW 32,000). Taking advantage of the extensive available assignment of the WT dimeric SOD[53], the structural characterization of the fALS mutants performed during this PhD could be performed without using deuterated samples

## 1.4 An overview on ALS pathogenesis

The first establishment of ALS as a distinct pathology was operated in 1874 by Jean Martin Charcot, who first noted the characteristics of ALS and established it as a distinct condition, but it wasn't until 1993 that mutations in $Cu_2Zn_2$SOD were associated with familial ALS[60] (http://dev.nsta.org/evwebs/2150/history.htm).

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig's disease or Motor Neuron Disease (MND) is a progressive, fatal neurological disease belonging to a class of disorders called motor neuron diseases. Every year, the number of newly diagnosed individuals with ALS is 1-3/100,000, with a prevalence not much higher given the short life expectation[61]. The ethnic incidence is similar world wide, with exception for the South Pacific area where it is higher, while men are about one-and-a-half times more likely to develop the disease than women.10% of ALS stricken people have at least one other affected family member and are said to have familial ALS (fALS hereafter), that can be categorized by mode-of-inheritance and subcategorized by specific gene or chromosomal locus[62]. About 20% of fALS cases are linked to a genetic defect on chromosome 21, in locus 21q22.1 in the gene SOD1 encoding $Cu_2Zn_2$-SOD.

SOD 1 is a 12 kb gene, with five exons and four introns. More than 100 mutations have been catalogued throughout all five exons, gathered in 64 different loci[63]. They are predominantly missense, even if frameshifts, deletions, insertions or truncations have been reported, such as V118 -> Stop 122 that removes 32 a.a. including the catalytically important Arg143, the Cys146 involved in the disulfide bridge and the eight strand wholly[61,64]. All of them are dominant, except D90A that can be recessive in certain families [65,66].

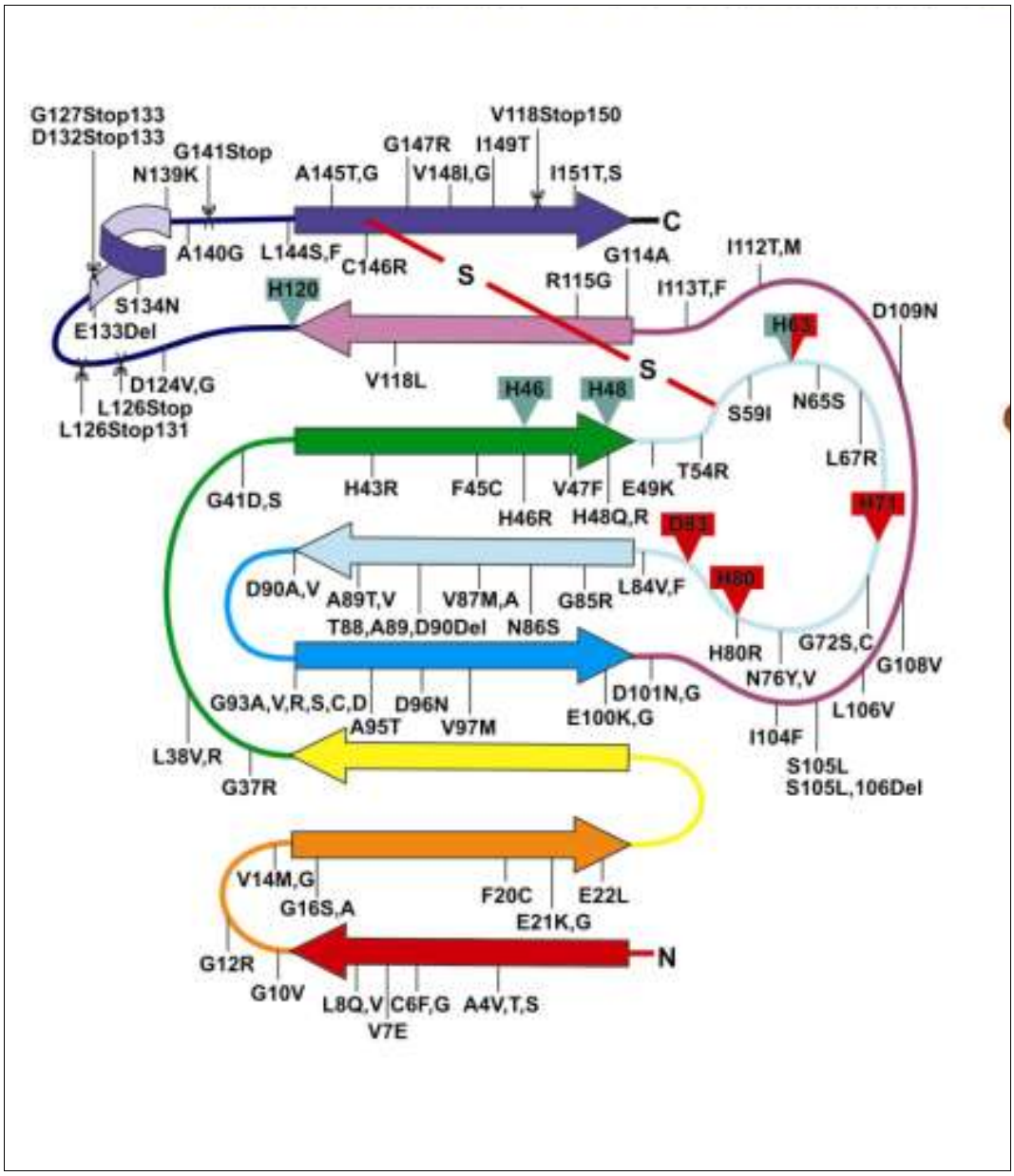



**Figure 4:** Secondary structural representation of SOD showing the locations of fALS-associated mutations

ALS strikes in mid-life and the average age of onset in individuals with no known family history is 56 years, while for individuals with more than one family member affected (familial ALS or fALS), it is 46 years. Its pathogenesis implies the gradual degeneration of upper and lower motor neuron(UMN and LMN hereafter), where intracellular accumulations and perikaryal inclusions of neurofilaments and cytoplasmic inclusions such as Bunina bodies and Lewy body–like inclusions are found[2]. Lower motor neurons are located in the spinal cord and innervate striated muscle; symptoms of their defeat include weakness, muscle wasting, hyporeflexia, muscle cramps, and fasciculations. Upper motor neurons are located in the motor cortex of the frontal lobe and their axons descend through the bulbar great corticofugal tracts to the brainstem and

spinal cord, influencing patterned activity of the LMN. In the case of UMN death, we will have hyperreflexia, extensor plantar response, increased muscle tone, and weakness in a topographical representation. In dependence of the class of motor neuron stricken at first, the early stage of the disease can be characterized by different clinical aspects. On the contrary, the subsequent course is common to almost all patients, that become paralyzed and often require ventilation and surgery to provide a new opening in the stomach. Sphincter control, sensory function, intellectual ability, and skin integrity are usually preserved, even if about 5% of the cases present cognitive impairment and dementia. Progression in fALS may be significantly shorter or longer than in sporadic ALS (sALS), where the average life expectation is of about three years. Usually death occurs for loss of respiratory function[67].

sALS etiology is unknown but consists probably of a multifactorial combination of oxidative stress, glutamate excitotoxicity, mitochondrial dysfunction, inflammation, and apoptosis; some evidence suggests that also the immune system may be involved[68]. Environmental causes of ALS have been studied for many years, since local increased incidence of ALS have showed up, for example, in Gulf War veterans[69] or Italian soccer players [70]. A quite unique case is Guamanian ALS, where a dietary neurotoxin (an amino acid, the BMAA found in the seed of the cycad *Cyas cirinalis,* a tropical plant that was a major dietary component during the 50s) seems to be the primary cause of the disease [71].

Treatment is just palliative, up to now, and the only currently FDA-approved drug is Riluzole. Its mechanism of action is thought to be glutamate inhibition. Clinical trials have shown marginal slowing of disease progression in some but not all individuals[72].

## 15. Insights into fALS molecular biology

The initial hypothesis that the disease could be due to lack of SOD activity was quickly refuted when a transgenic mice line expressing the mutant G93A developed progressive motor neuron degeneration despite high SOD measured activity[73] while in contrast SOD1 knockout mice had normal lifespan and didn't develop motor neuron disease[74]. This finding was extended by more mice lines that expressed G37R[75] and G85R[76]. In these cases the disorder onset was observed despite the unchanged SOD activity.

These evidences, and the autosomal dominant nature of SOD1-associated fALS[65], suggested that the mutant SOD-mediated toxicity arises from a gain of functions, not fully delineated but related either to the tendency of these variants to give rise to aggregates or to oxidative damage caused by aberrant catalytic activity[77,78].

As reported above, about 20% of fALS cases are related to mutations in the gene SOD1, responsible for $Cu_2Zn_2$SOD expression. They are gathered in 64 positions, not clustered in preferential regions of the protein, but scattered all along both the primary and the tertiary structure. This is probably the most unusual feature of fALS, since it should be expected that in dependence on the area of the enzyme involved in mutagenesis, different functions of the protein would be disrupted so causing different disease symptoms. Extensive characterization of some of the fALS mutants, in fact, shows that they can be divided in two main pools, wild type-like (WTL hereafter) and metal binding region (MBR hereafter) mutants, characterized by distinct features with respect to spectroscopic properties, metal content, stability and activity in scavenging the superoxide ion[16,17]. Enzymes belonging to the first subgroup (e.g. G37R or G93A,R,C,) where mutations are located in "safer" regions of the tertiary structure, are given of highly similar three-dimensional crystal packing with respect to wild type $Cu_2Zn_2$SOD. For example, crystal structures of metal-bound WTL mutants A4V and I113T reveal both minor differences in the dimer interface region, if compared with *wt*SOD, but aside this they are not significantly different in any other aspect[79]. G37 represents a peculiarity in this sense, since its X-ray structure evidences that its two subunits have different properties, namely one of them retains the Cu(II) four-coordinate geometry with histidine ligands 46, 48, 63, and 120, while the other sees the histidinato bridge (His 63) broken, leaving the Cu in trigonal coordination. Despite this major change, the refined structure shows essentially no gross deviations in backbone positions relative to the wild-type and thermostable mutant human SOD protein coordinates and no aberrant superstructure upon crystallization[15]. For this reason, it has been hypothesized that those mutant proteins may adopt a toxic form either prior to folding and metallation or subsequent to covalent modification or partial degradation. On the opposite, the MBR mutants (e.g. H46R, H48Q,R, and S134N) with amino acid substitutions that interest Cu binding residues or the electrostatic and zinc loops, diminish the polypeptide affinity for metals and activity, yielding an enzyme with altered SOD activities and spectroscopic properties, that easily crystallizes in filament-like arrays with abnormal contacts between the single protein dimers, although the dimer structure undergoes only local changes. A good example of this are the MBR fALS mutants apo-H46R and $Zn_2$-ΔH46R, and S134N [13], that evidenced the presence of higher order structures in the crystal packing, very similar in some cases to amyloid fibres described for other proteins and resembling a pore if their helical structure is extended to the width of a lipid bilayer membrane. This pores could be responsible for permeabilizing membranes causing cell death and are probably triggered by disordered electrostatic and zinc loops, usually seen in the metal-deficient SOD structures, despite a full complement of zinc in the $Zn_2$ H46R zinc site and in the fully metallated S134N. The final effect is an alternate conformations of the loops, that allows

non native interactions between dimers,. This kind of packing is not happening in those SOD structures in which the zinc and electrostatic loops are quite rigid and properly ordered, since such an arrangement is one of the examples of "negative design" that prevents oligomerization in wild-type SOD[13]. Being not present in other SOD structures, this dimer-dimer contact was termed a gain-of-interaction.

Probably the most challenging feature of such mutants is their common relation with the same neurodegenerative disorder, characterized by the same symptoms, although each of them shows different aberrant features not shared by *as*SOD. What's even more perplexing is that there is a correlation between life expectations and different mutations, but aminoacid substitution likely to cause highly disruptive effects in the overall enzyme architecture, like in MBR mutants do not correspond to the shortest life spans, as could be expected, and vice versa (e.g patients with the WTL mutation A4V have a life expectation of less than 1 year after the disease onset, while patients with mutation H46R, that involves one of the Cu ligand Histidines, have about 12 years of survival) [80,81]. Activity data are not available for all of the mutants; G37R demonstrated normal SOD enzymatic activity [82], while H48Q exerted a novel superoxide-dependent peroxidase behaviour and a very low SOD activity, namely about 1% of the wt enzyme[83], despite binding a similar amount of copper with respect to *wt*SOD. S134N showed an activity 3-5 fold lower with respect to *wt*SOD, even after accounting for the lower copper content found when the mutant is expressed in an Sf21 insect cell/baculovirus system[16]

# 2.Experimental part

## 2.1 PCR mutagenesis, protein expression and purification

**State of the art**

At the moment, several host systems are available for protein expression, including bacteria, yeast, plants, filamentous fungi, insect or mammalian cells grown in culture and transgenic animals. The choice of one instead of another is strongly dependent upon the specific requirements for the final product, since it will affect not only the protein expression, but also the subsequent purification. The latter depends upon the location of production within the host; for example, bacteria may secrete the protein into the growth media, transport it to the periplasmic space or store it as insoluble inclusion bodies within the cytoplasm. In each of the cases, the isolation will be performed in a different way. Other specifications to be considered are the need for proper metallation, in the case of metalloproteins, or of post-translational modifications (e.g. *E. coli* yields not properly N-acetylated proteins), fundamental to preserve the full activity or some inherent structural properties of some enzymes, since they can't be accomplished in every host. Often the first expression system of choice is *E. coli*, especially when small, soluble and non toxic proteins are the expression target, since it is fairly easy and fast to grow with respect for example to insect or mammaliam cells.

After choosing the organism of interest, the vector plasmid for expression must also be carefully designed, taking into consideration the specific requirements of the application since it will influence the behaviour of the target protein. A key factor that can greatly simplify isolation, purification and detection of polypeptides is the use of fusion protein vectors, containing a tag of known size and biological function that can enhance, for example, solubility or affinity for some specific purification systems. In some cases the protein yield can also be increased.

The functional analysis of a protein can often require the generation of a number of single amino acid variants of the wild type gene, in order to gain information on the importance of individual residues in the encoded polypeptide, or when attempting to assess whether a particular missense mutation found in a known disease gene is pathogenic, for the characterization of the disease-related mutant proteins. When an expressing, fully functional plasmid is already available for the

wild type enzyme, the easiest way to obtain a gene encoding the mutant/s is to perform PCR mutagenesis directly on the construct, instead of creating a new one which requires time consuming cloning operations. For this purpose, the first step is the design of mutagenic oligonucleotide primers, whose sequence is complementary to the gene sequence in the to-be-mutated region, with a single difference at the intended mutation site where the bases complementary to the desired mutant nucleotide, rather than the original ones, must be inserted. The mutagenic primers are then allowed to prime new DNA synthesis to create complementary full-length sequences containing the desired mutation. The methylated, nonmutated parental DNA is then digested with a suitable endonuclease such as DpnI, and the newly formed, mutated plasmid is used to transform cells suitable to propagate it. The mutagenesis product isn't absolutely unambiguous, since the reaction can yield a small fraction of randomly mutated or non mutated plasmids; for this reason, the screening via sequencing of constructs isolated from different colonies is needed, in order to assess the actual presence of the desired mutation.

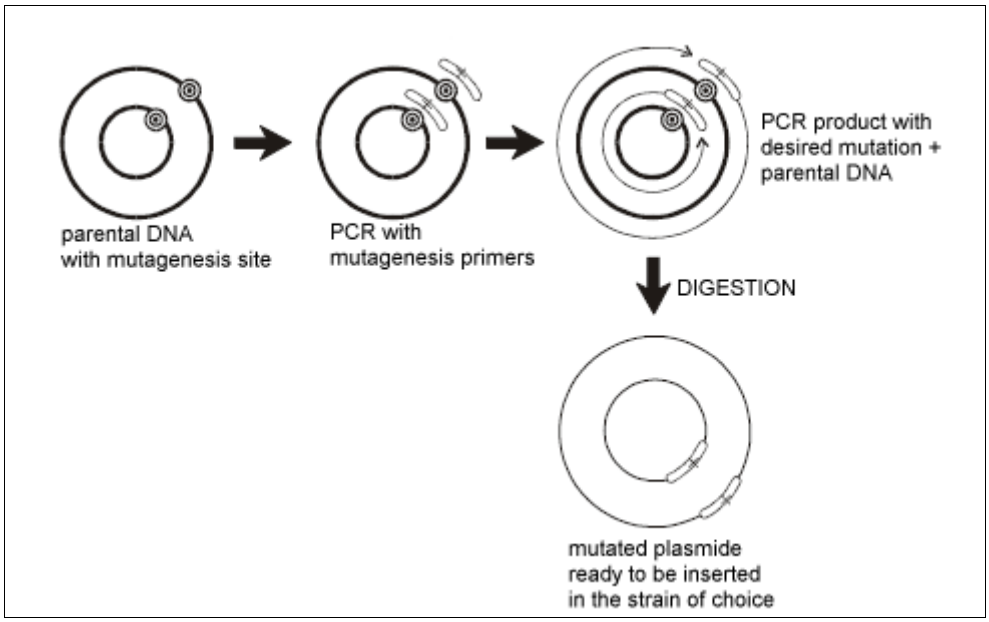



**Figure 5:** Main steps of PCR mutagenesis

In dependence of the spectroscopical technique of choice, the protein expression is performed in differently composed media. In fact, when large amounts of proteins must be isolated for the sake of techniques that do not require isotopic labelling, the colture is usually performed in a so-called rich or complex medium. Complex media contain water soluble extracts of plant or animal tissue

(e.g., enzymatically digested animal proteins such as peptone and tryptone), and for this reason are rich in nutrients and minerals, assuring a fast bacterial growth and an high expression level. Their exact composition is unknown and this can impair the coltures reproducibility. Chemically defined (or minimal) media are composed of pure ingredients in measured concentrations, dissolved in milliQ water; this way the exact chemical composition of the medium is known, allowing high reproducibility of protein yields and grade and type of interferents. Typically, this class of media is composed of a buffering agent (e. g. phosphate buffer) due to maintain the colture pH around physiological values, a carbon and energy source like a simple sugar (glucose) that can be replaced with glycerol, and an inorganic nitrogen source, usually an ammonium inorganic salt. In dependence of the bacterial strain and of the expressed proteins various mineral salts can be added and, if necessary, growth factors such as purified amino acids, vitamins, purines and pyrimidines. Chemically defined media are easier to isotopically label, simply using $^{15}N$ and $^{13}C$ enriched nitrogen and carbon sources in its composition, even if isotopically enriched complex media are also commercially available.

Uniform labelling of proteins is particularly important when dealing with 2D and 3D NMR. $^{15}N$ is very convenient because of the strategic locations of nitrogens in the backbone, and the absence of homonuclear couplings due to the intervening carbons[84]. The replace of all the carbons with $^{13}C$ is also possible, although in this case the spectroscopy has to deal with the network of couplings among bonded carbons. Partial or total deuteration (perdeuteration) attenuates the dipolar couplings among protons, reducing the relaxation rates of NMR-active nuclei, since the gyromagnetic ratio of $^{2}H$ is 6.5 times smaller than $^{1}H$. This last kind of enrichment is the most difficult to achieve since incorporation of $^{2}H$ reduces the growth rate of organisms up to 50%, and decreases protein production as a consequence of the isotopic effect. For this reason, deuteration can require a gradual adaptation of bacteria to increasing concentration of $D_2O$.

Cell lysis is the first step in protein purification and the technique chosen for the disruption of cells must be compatible with the amount of material to be processed and the intended downstream applications. Many techniques are available, spanning from physical to detergent-based methods. Among the first class are mechanical disruption**,** that relies on the use of rotating blades to grind and disperse large amounts of complex tissue; liquid-based homogenization, such as the french press, where cells are lysed by forcing the suspension through a narrow space, thereby shearing the cell membranes, and sonication, which uses pulsed, high frequency sound waves to agitate and lyse cells, bacteria and spores. Mechanical energy from the probe initiates the formation of microscopic vapour bubbles that form momentarily and implode, causing shock waves to radiate through a sample. To prevent excessive heating, ultrasonic treatment is applied

in multiple short bursts to a sample immersed in an ice bath. Gentler ways of disrupting cells are freeze-thaw lysis, detergent lysis, enzymatic lysis and osmotic lysis. All of them rely on a precise physical-chemical process, e.g. for the freeze-thaw protocol, freezing a cell suspension in a dry ice/ethanol bath and then thawing the material at room temperature causes cells to swell and ultimately break as ice crystals form during the freezing process and then contract during thawing, while in enzymatic or detergent lysis the membrane breakage is achieved by enzymatic removal of the cell wall with an enzyme specific for the type of cell or by addition of micellar compounds that solubilize the cells walls, freeing the content. Lastly, osmotic shock is a very gentle method, well suited for applications in which the protein is expressed in the periplasm. The big advantage of this technique is that all the interferents coming from the cytoplasmic region are eliminated, since the inner membrane is not broken, preventing also the leakage of genomic DNA and of a good fraction of the proteinases. Osmotic shock is usually achieved by equilibrating the cells in a hypotonic solution; breakage can be boosted if bacteria are equilibrated first in an hypertonic solution, such as 20% sucrose.

The purification of a protein is a crucial matter and the key to successful polypeptide isolation is to select the most appropriate techniques and combine them in a logical way to maximise yield and minimise the number of steps required, avoiding the need for intermediate conditioning passages. Purification generally exploits proteins intrinsic properties to separate them one from the other. The choice of method depends on a number of variables, such as chemical-physical properties of the target protein, its location and form within the cell, vector construct, and intended applications for the expressed protein. Two traditional classes of methods are differential precipitation (by ethanol, ammonium sulphate or heat), or chromatography.

In laboratory scale, salting out of proteins, particularly by use of ammonium sulphate, is one of the best known and used methods for purifying and concentrating enzymes. It is based on the increase in the ionic strength of the solution, that causes a reduction in the repulsive effect of alike charges between identical molecules of a protein and also diminishes the forces holding the solvation shell around the molecules themselves. When such interactions are sufficiently weakened, the protein will precipitate, with hydrophobic polypeptides precipitating at lower salt concentrations than hydrophilic ones. Ammonium sulphate isn't the only salt that can be used, but it is widely employed thanks to its high solubility, cheapness, lack of toxicity, and its stabilizing effect on some polypeptides. When performing salting out, the trick is to ensure that only the target protein precipitates. This can be achieved by using an appropriate ammonium sulphate concentration, determined by experiment or from literature references. The precipitation reaction should be performed at low temperatures to prevent denaturation of the target protein. The ammonium sulphate must be added slowly and mixing is important to distribute the salt and

prevent inclusion of unwanted proteins in the precipitate. After precipitation, the pellet is separated by centrifugation, then dissolved in buffer for further purification or analysis.

This technique can be applied in reverse mode, so to have the greatest part of the interferents precipitated while the target protein remains in solution. In this case, such method can be successfully coupled with hydrophobic Interaction Chromatography (HIC hereafter), since protein binding to HIC adsorbents is promoted by moderately high concentrations of anti-chaotropic salts, which also have a stabilizing influence on protein structure. This chromatography, in fact, is related to the differential absorption of polypeptides on columns stacked with polystyrene beads, functionalyzed by mean of bound hydrophobic groups. The strength of interaction depends on the number of hydrophobic amino acids externally positioned on a protein; elution is achieved by a linear or stepwise decrease in the concentration of salt in the adsorption buffer. Recoveries are often very satisfactory.

Another powerful chromatographic technique is Ion Exchange Chromatography (IEX hereafter), that separates the proteins on the basis of a reversible interaction between the polypeptide and a specific, charged ligand attached to a chromatographic matrix. The sample is applied in conditions that favour specific binding, namely low ionic strength of the solution and carefully calibrated pH, in order to enhance interaction with the target of interest. The unbound material is washed away, then the bound protein is recovered by changing conditions to those favouring desorption. This can be achieved specifically, using a competitive ligand, or non specifically by changing pH or ionic strength of the eluent. Poorly binding proteins will be eluted first at the lower salt concentration, while strongly bound ones at higher concentration. The ability to incorporate different buffers in IEX allows proteins to adsorb on either anion or cation exchangers, respectively for positively or negatively charged polypeptides. The most common exchanger is an anionic one; since the majority of proteins carry an overall net negative charge at environmental conditions.

In case of peptide fusion tags, a variety of affinity purification methods are available. The protein tag specifically interacts with an affinity resin (e.g chelating resins loaded with Ni or Zn for $His_6$ tags) and binds the tagged protein whereas other polypeptides are not bound. This allows a simple purification by releasing the tagged protein with a suitable competing buffer. In this case, usually an intermediate purification step includes the cleavage of part or all of the fusion tag with proteases such as enterokinase, Factor Xa, thrombin, TEV in dependence of the cleavage site. The fusion protein is then separated, usually by size exclusion chromatography provided that the difference in size with the target polypeptide is big enough.

Gel filtration is the simplest and mildest of all the chromatographic techniques. Unlikely to IEX or HIC, the molecules do not bind to the resin, but are allowed to travel through a medium given

of controlled porosity, so the buffer composition does not play any role in the resolution. Consequently, a significant advantage of this technique is that conditions can be varied to suit the type of sample, allowing the use of buffers with high or low ionic strength and added of essential ions or cofactors, detergents, urea or other additives if needed for the protein stability. Size exclusion can be performed on a "rough" level, separating the components of a sample in major groups to remove for example high or low molecular weight contaminants or to exchange buffers, while high resolution fractionation of biomolecules allows to isolate one or more components of a protein mixture, to separate monomers from aggregates and last but not least to determine molecular weights or to perform a molecular weight distribution analysis, provided that suitable standards are available.

**Results**

In our specific case, the plasmid pPCuZnSODIIᵠ, containing the gene for the expression of *as*SOD, was already available, as well as an already mutated one for the monomeric SOD. Both of them were derivatives of pBR322, with HindIII and SalI as restriction sites, as well as a leader sequence for the periplasmic expression. Mutations G37R, H48Q, and S134N were introduced in pPCuZnSODIIᵠ by PCR mutagenesis, using the QuikChange® mutagenesis kit (Stratagene, California). Analogously, a plasmid encoding *wt*SOD was obtained by double backmutation on pPCuZnSODIIᵠ, and the mutation S134N was inserted also in this one. All the resulting constructs were sequenced to confirm the mutations.

To achieve high yields in protein production, careful screening for favourable expression conditions was performed, since even a single amino acid mutation can yield a differently expressed protein. Many small sized coltures were set, with high variability of expression parameters such as media composition, inducing agent concentration, length of induction time and special attention to trace metal supplements. Since the first expression tests gave good results with *E. coli* Topp1 strain, used also for the expression of *as*SOD in our laboratories, no screening of different strains was performed. Tests with different $pO_2$ were also performed, when the process was scaled-up from flasks to fermenters.

SOD and its variants were expressed both in complex and chemically defined media. Fermentation in rich media was performed in 2xYT medium, supplemented with 1 mM $CuSO_4$ at the induction, achieved by addition of 1mM isopropyl-beta-D-thiogalactopyranoside (IPTG hereafter). Single ($^{15}N$) and doubly ($^{13}C$, $^{15}N$) isotopically labelled, minimal media were also employed, whose formulation arises from a modification of the M9 medium. They were

composed of a phosphate buffer, added with 2 g $L^{-1}$ of glucose, 0.6 g $L^{-1}$ of ammonium sulphate, and is enriched by B vitamins and trace metals. 0.25 mM $CuSO_4$, was added at the induction and its concentration increased up to 1mM 1 hour before the cells harvesting. Also in this case the expression of SOD was induced by addition of 1mM IPTG, but the induction time was shortened from overnight, in the case of rich medium, to 6 hours. In the case of fALS mutants, the yield lowered sharply upon scaling up the coltures from small to big flasks. This was to be ascribed to insufficient aeration of the medium, being more difficult, when working with 2L flasks, to get turbulent motions in all the colture volume so to allow an uniform and effective oxygen diffusion. For this reason, in the optimized procedure 150 mL of colture for each 2 L flask were set and this restored yield to reasonable levels. Finally, in the case of *wt*S134N mutant, $CuSO_4$ was not added since it triggered the precipitation of the enzyme in inclusion bodies even if added just one hour before harvesting the cells. The fermentations were performed in flask, and in some cases in fermenters, too. The yields were very different, in dependence of the mutation, the medium and the method of choice.

| | flask | | fermenter | |
|---|---|---|---|---|
| mutant | 2xYT | M9 | 2xYT | M9 |
| G37R | 40 | 16 | - | 19 |
| H48Q | 43 | 18 | - | - |
| *as*S134N | 53 | 25 | - | 36 |
| *wt*S134N | 28 | 10 | - | - |

**Table 1:** fALS mutants yields expressed in mg of purified protein obtained per litre of colture

The cells were disrupted by osmotic shock, since the plasmids encoding SOD, as well as all its derivatives, contain a leader sequence for the periplasmic migration. The cells were first equilibrated in a hypertonic 20% sucrose solution, buffered at pH 8 by 20 mM Tris then centrifuged and equilibrate in mQ water, that broke the external membrane by osmotic pressure. This method was not applied in the case of *wt*SOD and *wt*S134N mutant, since their migration in the periplasmic zone was inhibited, probably by aggregation in the bacterial host. In that case, the protein was recovered by sonication of cells, using a DTT added buffer, in order to keep reduced the free cysteins and inhibit aggregation to some extent.

SOD and its variants were purified mainly by cromatografic techniques, employing an Akta FPLC system (Amersham Biosciences) that allowed fast and highly automated purification steps.

*as*SOD, E133QM2 SOD and all the *as* fALS mutants, in 20 mM Tris buffer pH 8, were satisfactorily isolated with two IEX steps. The first lysate rough purification was achieved by mean of a 20 mL Hi Load Sepharose Q FF (Amersham Biosciences) strong anion exchange column, using a linearly increasing salt gradient towards a Tris buffer containing 150 mM NaCl. The enzymes were typically eluted between 60 and 100 mM NaCl. The fractions containing less pure SODs, identified by SDS-PAGE, were typically concentrated, desalted by mean of disposable Pd-10 columns and applied onto a 6 mL Resource Q-Sepharose analytical column (Amersham Biosciences). The elution was performed in the same way, yielding the target enzyme as pure fractions. *wt*SOD, and its derivative S134N, gave some problems during the first attempts of purification by mean of IEX, probably due to their tendency to aggregate, that caused aberrant surface charge distribution whose ultimate result was the finding of the non purified enzymes all along the elution gradient. In the case of unlabelled polypeptides a first, rough purification was achieved by ammonium sulphate precipitation of most of the interferents, eased by the fact that SODs are highly soluble in ammonium sulphate solution (up to about 70% of saturation). The first chromatographic step consisted of an HIC performed on two-in-series 5 mL Phenyl-Sepharose 6 Fast Flow high sub Hi Trap columns (Amersham Biosciences), employing a sodium phosphate pH 7.0 buffer, containing 2.0 M $(NH_4)_2SO_4$, 150 mM NaCl, 0.1 mM EDTA, and 0.25 mM dithiothreitol (DTT) to avoid cysteines triggered aggregation. Proteins were eluted with a linearly decreasing salt gradient toward a low salt phosphate buffer and released with good specificity between 1.6 and 1.1 M ammonium sulphate. Again, the non pure fractions were desalted and purified on an analytical IEX column in the same way as *as*SOD and its mutants, but adding 1 mM DTT to all the buffers.

Since the $^{15}N$ uniformly labelled wtS134N mutant showed lower solubility in ammonium sulphate with respect to the unlabelled one, and a loading buffer with 6 M NaCl hadn't ionic strength high enough to allow the protein binding to HIC columns, this specific enzyme was purified by anionic exchange chromatography using the same protocol as the *as*SOD mutants. In this case the elution buffers were added of 2 mM DTT, in order to keep the cysteins reduced to the maximum possible extent.

## 2.2 Metals content determination by ICP

**State of the art**

Inductively Coupled Plasma (ICP) is an analytical technique used for the detection of trace metals and allows multi-elemental, simultaneous analysis of most elements in the periodic table, with excellent sensitivity (ng/L or ppt) and high sample throughput. The analytical principle used is optical emission spectroscopy, triggered by plasma ionization. The liquid sample is introduced in the prechamber, usually by mean of a peristaltic pump, then nebulized and entrained in the flow of plasma support gas, which is typically Argon. The analytes are excited into a state of radiated light emission by the plasma vapour and the emitted radiation is converted to an electrical signal that can be measured quantitatively by resolving the light into its component radiation (almost always by means of a diffraction grating). The light intensity is measured, at the specific wavelength for each element line, with a CCD (*charge coupled device*, basically an array of tiny, light-sensitive diodes) which converts it to electrical signals. The electron signal intensity is then compared to previously measured intensities of standard solutions of the target elements, and a concentration is computed. This spectroscopic technique is very suitable for the proteins metal content determination, since it requires very small amounts of sample, and is not perturbed by the polypeptidic matrix.

**Results**

Metal content analysis was performed of fALS mutants G37R, obtained from complex medium, and *as*S134N, obtained both from complex and chemically defined media. This step was absolutely crucial, since some aberrant properties of *as*SOD and its mutants could depend on the grade of the enzymes metallation and usually these enzymes are isolated with incorrect metallation from baculovirus systems. What we found is that the mutant *as*S134N is isolated from *E. coli* with a non aberrant metals content, namely a zinc and a copper ion per subunit, while G37R is given of a more abundant content of Zn, which is about three fold the Cu one.

| | Cu equivalents | Zn equivalents |
|---|---|---|
| S134N from M9. | 2.1 | 2.2 |
| S134N from 2xYT | 1.8 | 2.1 |
| G37R | 0.8 | 2.5 |

**Table 2:** Metal content of some of the isolated mutants

The findings for S134N are very interesting, since in the employed expression system SOD was not coexpressed with any Cu chaperone. For this reason the chance that the heterologous expression could lead to a non native metallation state, even if the mutants affinities for its ions were unchanged, was quite high, especially because *E. Coli* lacks the bacterial isoenzyme.
The correct metallation state of S134N suggests also that the diminished affinity of mutant G37R for Cu is driven by structural aberration more than by technical issues related to the expression system

## 2.3 EPR

Electron Paramagnetic Resonance (EPR), often called Electron Spin Resonance (ESR), is the process of resonant absorption of microwave radiation (GHz range) by paramagnetic ions or molecules, with at least one unpaired electron spin and in the presence of a static magnetic field. In a typical experiment the sample is placed in a resonant cavity which has a high quality factor or high Q. At a fixed microwave frequency, the magnetic field $B_o$ is varied until resonance occurs at the value $B_e$ given by

$B_e = h\nu = g\beta B_o$

In this equation g is the "gfactor" of the unpaired electron and β is the unit magnetic moment of the spinning electron (Bohr magneton). Free radicals have values that are fairly close to
g = 2.0023 which characterizes a free electron, while some transition ions, such as cobalt ($Co^{2+}$) and copper ($Cu^{2+}$), have gvalues which differ significantly form that theoretical value.
The resonant absorption is not infinitely narrow, since unpaired electrons do not only interact with the externally applied magnetic field, but also with the magnetic fields in their neighbourhood. By observing spectral line width and intensity, it is possible to obtain information about the spin environment. Electron spin exchange between identical and non-identical molecules, chemical exchange between the paramagnetic molecule and its environment, and the interaction of nearby molecules having unpaired spins are some examples of environmental effects which can influence line width and intensity in the EPR spectrum. Moreover, an observed spectrum can split in several lines referred to as hyperfine structure, arising from the electrons interacting with surrounding nuclear spins. This last property becomes very useful when investigating biological molecules that contain transition metal ions in their structure, since it can be exploited to characterize their coordination environment in dependence of its intrinsic geometry.
Advanced EPR has represented a tool for the study of the electronic structure of the SOD copper site since a long time ago, by comparison of the structure of *wt*SOD Cu centre with the ones of

some carefully chosen mutants in positions 124, 125 and 133, aiming for a deeper understanding of the relationship between point mutations in the protein and variations in the Cu binding site geometry. The EPR spectrum of the native enzyme is rhombic at every temperature and characterized by a resolved hyperfine structure in the $g_{\parallel}$ region[36], while in the case of mutants 124, 125 and 133, a more or less markedly rhombic distortion usually appears[85].

**Results**

Also in the case of our fALS mutants, an increase in rhombicity was witnessed, when passing from *as*SOD to *as* and *wtS*134N mutants to G37R mutant. This well fits in the overall picture and could come of help to explain the lessened affinity of the latter for the Cu ion. Such distortion is evidenced also in the CD spectrum recorded in the visible region, that will be discussed later.

| mutant | $g_{\parallel}$ | $g_{\perp}$ | $A_{//}$ ($10^{-4}$ $cm^{-1}$) |
|---|---|---|---|
| G37R | 2.27 | 2.08 | 139 |
| *as*S134N | 2.26 | 2.06 | 154 |
| *wt*S134N | 2.28 | 2.06 | |
| *wt*SOD | 2.27 | 2.05 | 154 |

**Table 3:** Main parameters of EPR spectra of mutants G37R, *as*S134N, *wt*S134N compared with SOD

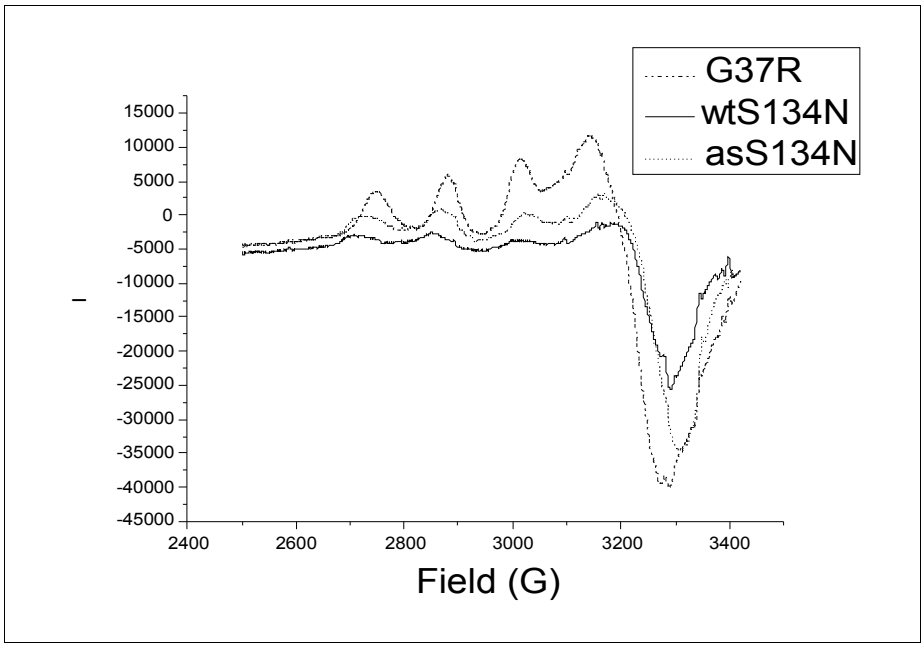



**Figure 6:** EPR spectra of mutants G37R, *as*S134N, *wt*S134N

## 2.4 Fast Field Cycling Relaxometry

**State of the art**

Fast Field Cycling Relaxometry is a NMR technique used for the determination of longitudinal relaxation times of the solvent ($T_1$) over a wide interval of magnetic fields, ranging from about $10^{-6}$ to about 1 Tesla. The range boundaries are set by sheer technical issues, being the lower limit affected by the local fields, while the upper limit is mainly determined by technical choices and compromises.

This interval is very wide if compared with the 0.1 T-20T range covered by standard NMR, not considering the impractical technical issues of studying $T_1$ dispersion curve with an array of standard magnets. On the other hand, FFC relaxometry requires a specialized system, given of much lesser resolution if compared with most NMR spectrometers.

The basic scheme of an experiment can be divided into three phases. At first the sample is polarised in a high field $B_p$ for the time needed to achieve saturation of the nuclear magnetization, then the magnetic field is switched to a lower value $B_r$ for a time $t_r$, during which the magnetization is allowed to relax towards a new equilibrium value. Eventually, the magnetic field is increased again and the equilibrium magnetization is measured, by applying a 90° pulse followed by acquisition.

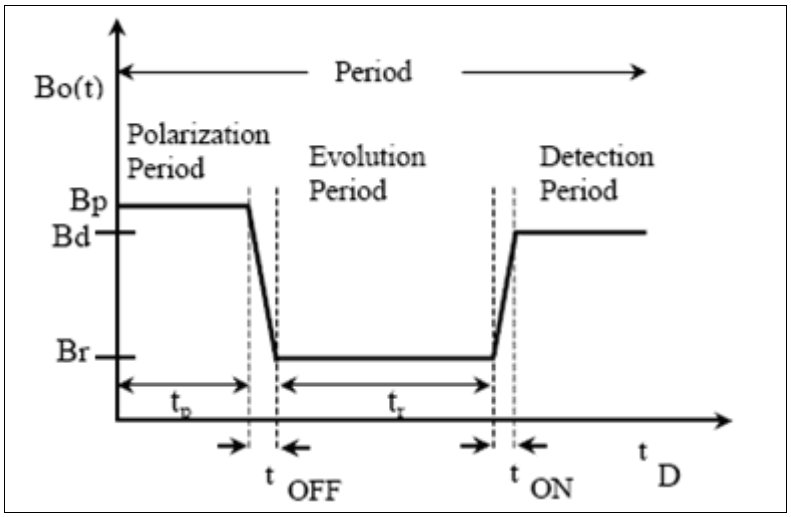



**Figure 7:** A typical field cycle in FFCR

Although FFCR can be applied to several research fields, spanning from dynamics of liquid polymers to proton quantum tunnelling investigation, a useful biological application is the characterization of the hydration of proteins in solution.

When recording the relaxation profile of an aqueous, concentrated protein solution, the observed proton relaxation is given by the sum of the contributions of all sample protons They can be categorized into four main groups, namely bulk water protons, protons of exchanging water molecules interacting with the protein and exchangeable and non-exchangeable protein protons. Given the lack of sensitivity and resolution of the technique, the only appreciable contribution to relaxation profiles in diluted protein samples comes from bulk water protons, although information about protein and bulk water exchanging protons can be indirectly inferred from the comparison of profiles registered in the absence and in the presence of protein in solution.

The water $^{1}$H NMRD profile for a protein solution is composed by a low field plateau, an inflection point, also called dispersion, and a high field limit value different from zero. The first two components constitute the so called dispersive part of the spectrum, which arises from contributions of long-lived water molecules and exchangeable protein protons, whose relaxation is modulated by the rotational correlation time of the protein. On the contrary, the non-dispersive plateau found at high fields is due to short-lived water molecules, whose dipolar interactions are modulated by their own short residence times, and can thus be treated as a constant value.

In complex systems such as a protein aqueous solution, NMR correlation functions often do not decay exponentially due to local mobility of side chains such as lysines and arginines, to strongly bound water molecules or to partial aggregation. For this reason the Lorentzian theoretical shape of a FFCR profile is usually stretched to some extent and the relaxation rates profile is invariably well described as a superposition of Lorentzian processes, each with a characteristic time and activation energy. Phenomenological “Cole and Cole” expressions[86], further modified by Hallenga and Koenig[87], demonstrated their validity when applied for the fitting of FFCR profiles[88]. Also based on the multi-lorentzian principle, a Model Free approach was developed by Halle et al[89], that yields multiple values for the three crucial parameters i.e. $_{c}$ the rotational correlation time, corresponding to the middle point of the NMRD profile dispersion; , defined in terms of the integral of dispersion and representing the overall strength of the interaction, hence being proportional to the water exposed protein surface and (usually subtracted of 0.3 $s^{-1}$ which represents the bulk water relaxivity), the non dispersive portion of the profile, which accounts for the contributions of short lived and near diffusing water molecules. Another fitting approach consists in the simulation of the relaxivity profile by mean of the implemented version of the CORMA program[90,91] which is able to calculate the complete relaxation matrix, including long-lived water protons and bulk water protons contributions, besides the protein protons

contribution. Such an approach is applicable only provided that a 3D protein structure, comprising also the position of long-lived water molecules, has been solved.

**Results**

The CORMA method was heavily employed for the study of GdmCl denaturation titration of monomeric mutant E133QM2 SOD from a relaxivity point of view. The task was possible since detailed characterization of mobility and hydrodynamic radius of its native and denatured states has been performed[20,53]. The profile of the unfolded protein shows similar features to those of the native one: the relaxation rate is constant at low fields, until a dispersion occurs. The general trend of the three main parameters $\alpha$, $\beta$, $\tau_c$ upon titration indicates an increase in the solvent exposed protein surface in the presence of denaturant, evidenced by the lengthening of the reorientational time and the relatively small increase in the $\alpha$ value. Such values are consistent with a transition from a folded to a molten globular structure more than to a randomly coiled peptide, which is consistent with the findings of high resolution NMR studies.

In the fALS mutants field of study, NMRD has demonstrated its potential in characterizing the aggregation tendency of the enzymes in solution. In this case the Model Free approach was exploited, since it provides an estimate of correlation time for the protein tumbling and so for its molecular weight. The fitting of the profiles for freshly prepared enzymes yields a single $\tau_c$ value for every mutant, with values in good agreement with *as*SOD one, confirming that the enzymes are in their dimeric state After ten days of incubation at room temperature, the profiles of G37R and *as*SOD were unvaried, while *as*S134N showed two distinct flexes, that after fitting with Model-Free provided two independent correlation times, one as an average value for the large molecular weight species and one for the dimeric protein. The higher $\tau_c$ value is consistent with oligomeric aggregates composed of 20 – 30 subunits.

| $\tau_c$ values (ns) | G37R | S134N | *as*SOD |
|---|---|---|---|
| freshly prepared | 19.4 | 20.0 | 20.5 |
| after 10 dd. @ RT | 19.6 | 19.0 / 548 | 19.3 |

**Table 4:** Comparison of FFCR parameters for mutants G37R and *as*S134N with *as*SOD

## 2.5 Circular Dichroism

**State of the art**

Circular Dichroism (CD) is a phenomenon occurring when asymmetrical molecules interact with circularly polarized light, thus absorbing left and right hand circularly polarized light with different absorption coefficients. Linearly polarized light can be viewed as a superposition of two opposite circularly polarized electromagnetic radiations of equal amplitude and phase and can be split in left and right polarized light by passing through a quarter wave retarder such as a Pockel cell. In this case, the projection of the components will yield a circle. When such polarized electromagnetic radiation interacts with an optically active sample, i.e. given of different absorbance for the two components, the amplitude of the stronger absorbed one will be smaller than that of the less absorbed one, thus yielding an ellipse as projection instead of a circle. The occurrence of ellipticity is accompanied by rotation of the polarization plane (i.e. the axes of the dichroic ellipse) by a small angle *a*, that occurs when the phases for the 2 circular components become different due to the difference in the two refractive indexes. The change of optical rotation with wavelength is called optical rotary dispersion, ORD and can be related to Circular Dichroism by a Kronig-Kramers transformation.

Asymmetry occurs in molecules when they lack a centre plane or an integer number of axis of symmetry (e.g. quaternary carbons with four different substituents) or when the arrangement of inherently symmetric molecules in space is asymmetric due to restricted rotation or highly ordered arrays. In proteins the major optically active groups are the amide bonds of the peptide backbone, typically disposed in highly ordered arrays such as α-helices or β-pleated sheets. In dependence of the orientation of the peptide bonds in the arrays, given by the symmetry of its disposition, optical transitions are differently split by exciton splitting, thus yielding characteristic spectral profiles for each of the three basic secondary structures of a polypeptide chain. A protein consisting of these elements will therefore display a spectrum that can be deconvoluted into the tree individual contributions, although a major drawback of this approach is that there’s no standard reference CD spectra available for "pure" secondary structures. The use of synthetic homopolypeptides in order to obtain reference spectra is, in general, a poor model for the secondary structures found in proteins since, for example, the CD of an alpha helix has been shown to be length dependent while no good example of a homopolypeptide has been found for the beta sheet structure.

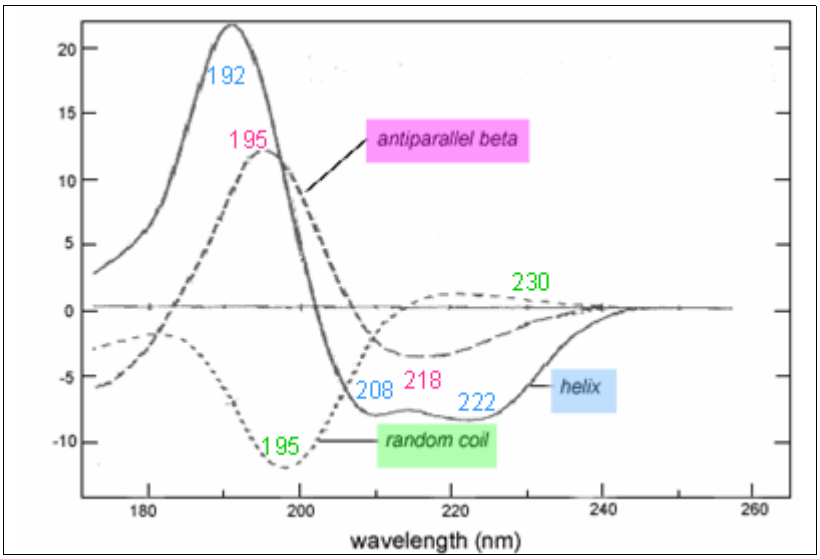



**Figure 8:** CD spectra of "pure" secondary structure

To bypass this problem, several mathematical methods have been developed, all of them relying on the assumption that the spectrum of a protein can be represented by a linear combination of the spectra of its secondary structural elements, plus a noise term which includes the contribution of aromatic chromophores

$\theta_\lambda = \Sigma \varepsilon_i S_i$ + noise

The main difference between one another is that some methods use selection procedures so that only proteins with spectral characteristics that are similar to those of the protein to be evaluated are used as standards. In this case, the main advantage is that no assumptions are made on the CD form of the individual secondary structural elements, and it assures that all the structural characteristics of the analyzed protein are represented in the set of reference spectra. In this way, irregularities of secondary structure and length dependencies should be taken into account Aside the simplest linear regression method, Convex Constraint Algorithm (CCA)[92], Ridge Regression (CONTIN)[93], Singular Value Decomposition (SVD)[94] the Self Consistent Method (SELCON)[95] and Variable Selection (VARSLC)[96] are some of the developed approaches.

Neural Networks (e. g. CDNN or K2D)[97,98] are very interesting software for the secondary structure assignment. They are artificial intelligence programs which can detect patterns and correlations in data. Usually they are first trained using a set of known proteins, so that the input of the CD at each wavelength results in the output of the correct secondary structure, then used to

analyze unknown proteins. Also tertiary structure class of globular proteins can be determined by CD, using the CLUSTER method,[99] which claims 100% accuracy for predicting all alpha, alpha/beta, and denatured proteins; 85% for alpha + beta; and 75% for all beta proteins. In this case, CD spectra of native proteins and denatured samples in the range 190–236 nm, at 2-nm intervals, are used to construct a 24-dimensional hyperspace corresponding to the ellipticity values at the 24 wavelengths. Proteins belonging to different tertiary structure classes (αα, , α+ , α/ , and denatured) clusters in this hyperspace and the hyperplanes equations separating the different groups are used to assign any new CD spectrum to one of the five tertiary structure classes.

Protein chromophores are not constituted only by the backbone peptide bonds. In different zones of the electromagnetic spectrum also aromatic sidechains, prosthetic groups and metals can give rise to absorption.

In the range 240 - 290 nm, aromatic amino acids absorption can be employed as a characteristic fingerprint of the native structure, while the appearance of the weak, unstructured band around 260 nm can help in the determination of disulphide bonds dihedral angles. In the near UV-VIS region, the absorption of cofactors, such as flavins, haem, pyridoxal 5X-phosphate, yields information about the groups environment. Finally, in the visible region, the absorption of coordinates metals can be used for the characterization of the metal environment geometry. In this case the signals can be fitted by a series of Gaussian lineshapes and provided that an angular overlap analysis is available for the system, the assignment of d-d and LMCT bands is possible.

**Results**

The CD spectrum of holo SOD has been previously reported and its general features agrees with the fact that native SOD is predominantly a β-sheet structure, as it is evidenced by the single, negative broad band centred at 280 nm[100].

The secondary structure content of all the mutants expressed during this PhD curse: E133QM2 SOD (monomeric non pathogenic one) and fALS related ones G37R, H48Q, and *as* and *wt* S134N was investigated in the range 190-250 nm. Results show that, as expected, at room temperature the secondary structure elements of all of them are fully conserved.

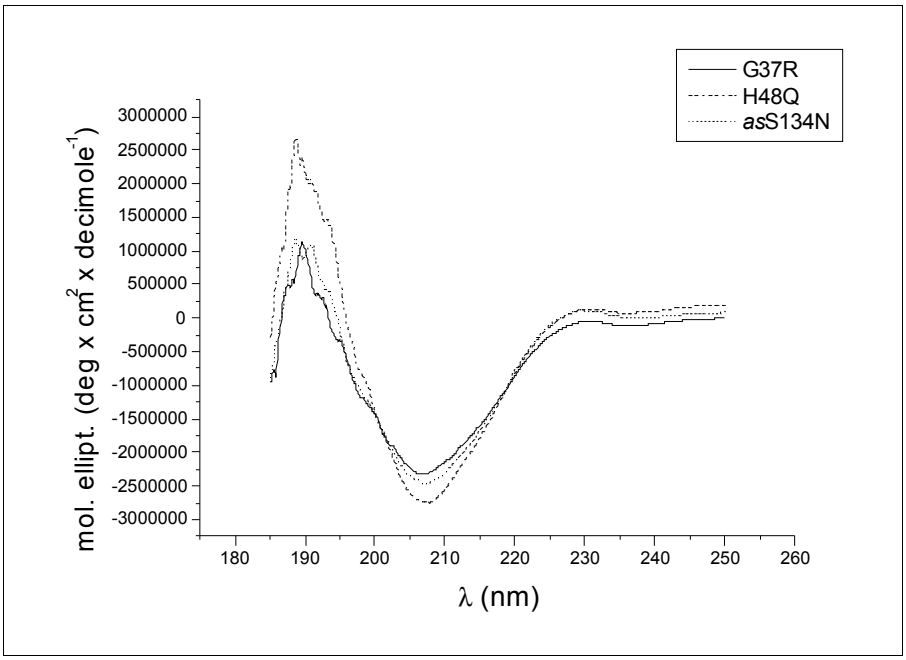



**Figure: 9:** CD spectra of mutants G37R H48D and asS134N in the UV region

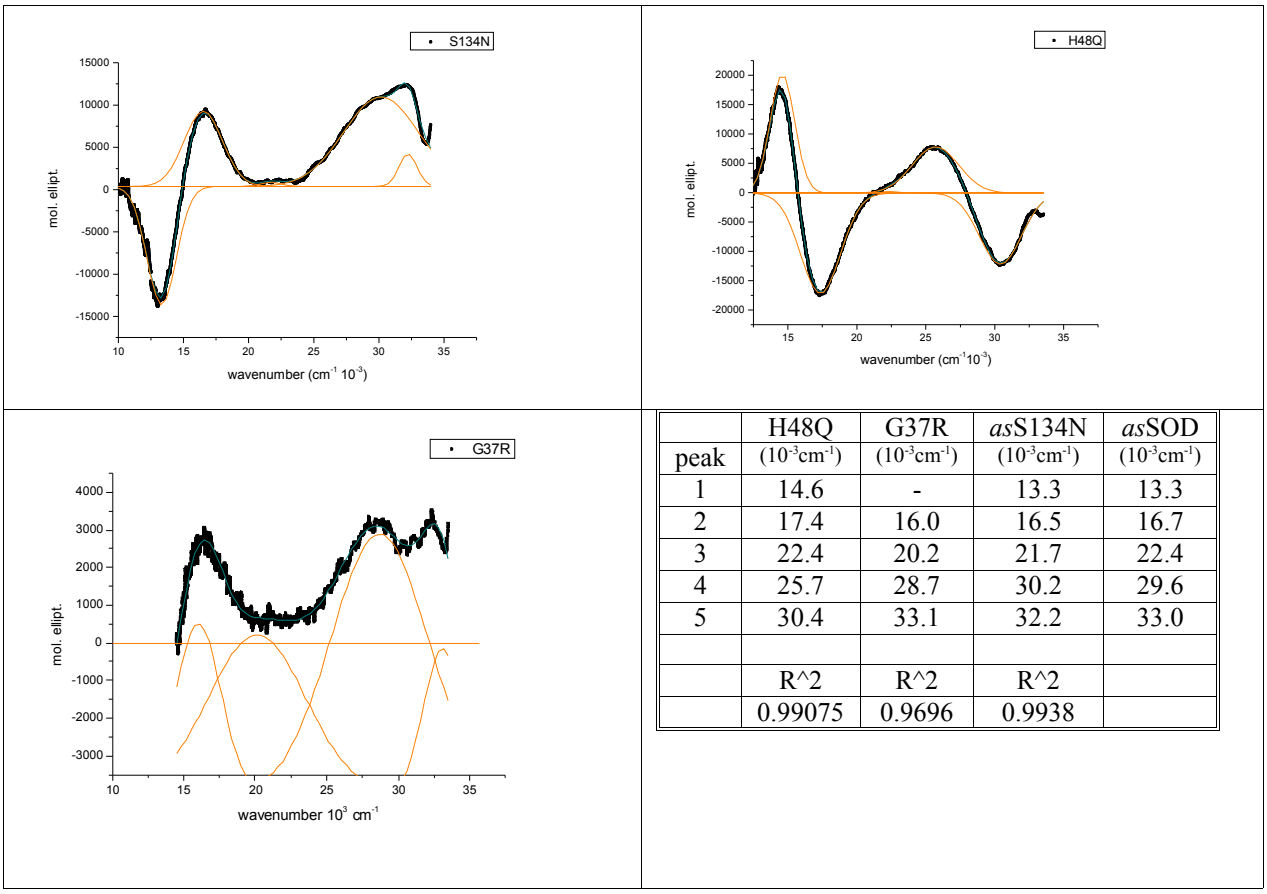
To make preliminary tests on the monomeric mutant of SOD, thermal unfolding and unfolding with Trifluoroethanol (TFE) and Guanidinium Chloride (GdmCl) were tested by mean of CD, in order to choose the most significant and less measurement-interfering conditions. Spectra were recorded in a range of wavelenghts from 185 to 250 nm, which covers all the relevant structural features, although at high GdmCl concentration the part of the spectra above 200 nm was covered by artefacts caused by the high OD of the sample. As already reported for dimeric SOD, the CD spectra progressively decreases in magnitude as GdmCl concentration increases[100]. To obtain the titration curve for the denaturation with GdmCl, the residual secondary structure content at 218 nm (the minimum for pure b-sheet spectrum) was plotted against GdmCl concentration. The sigmoidal fit of data shows the middle point of titration at 2.5 M GdmCl and complete unfolding of SOD for GdmCl concentrations higher than 3.5 M, which well agrees with previous data obtained by High Field NMR techniques [20].

Also the relative stability of mutants was tested, by mean of a double 2,2,2-trifluoroethanol (TFE)/temperature titration, in the ranges 0 – 25% TFE and 25 – 60/80°C. There is increasing evidence that soluble aggregation do not develop directly from the native conformations of proteins but from precursors which are only partially folded[1] (Booth, D. R.et al. (1997) Nature 385, 787–793.), for this reason we decided to use such fluorinated alcohol, known to stabilize partially unfolded states of proteins, in order to check whether there were non hierarchical

folding intermediates in the enzymes unfolding pathways, and if such forms were also characterized by aggregation. For this kind of investigation we chose to use both *as isolated* and demetallated (apo) enzymes, in order to check the differences in stability upon metal sequestration. It was found, in fact, that both *as*SOD and some fALS mutants suffer a loss of stability when in their apo form[16,17,50,59,101], and this could imply that not completely mature forms of SOD are involved in the aggregation process.

As expected, the studies performed on our mutants evidenced that all the metallated proteins show little or no denaturation upon addition of low amounts of TFE followed by increasing of the temperature, while apo forms are much more prone to unfolding. When performing double TFE-temperature titration on the holo enzymes, mutant S134N shows the biggest destabilization with respect to *as*SOD and it is the only one giving net helices formation even in its metallated form, although also *as*SOD and mutant G37R show a certain α-propensity. In fact, their final titration spectra are quite similar to *as*S134N unfolding intermediate, especially at high TFE concentrations. Unfolding titration performed on the apo enzymes evidence that while the rise of net helical structure is not a common feature for the metallated mutants, in the range of temperature and TFE covered, all the apo protein, *as*SOD included, present helical propensity to a different extent, being mutant H48Q the less affected. Apo mutant G37R is given of an unfolding pathway very different from the other apo enzymes, seemingly with a smaller content in random coil. Apo *as*SOD and apo *as*S134N both suffer a loss of secondary structure even at room temperature for high TFE concentrations, indicating that the loss of Zn and Cu ions destabilizes them to an heavier extent with respect to H48Q. Their unfolding pathway and relative stability look very similar one each other, with *as*S134N showing sharper transitions. In the end, it must to be reported that for TFE concentrations higher than 10%, the aggregation of apo and holo S134N and apo H48Q was appreciable by eye at the end of the temperature increasing slope.

CD was employed also to characterize the Cu environment of all the expressed fALS mutants, in the wavelength range 290-1000 nm.




| peak | H48Q ($10^{-3}cm^{-1}$) | G37R ($10^{-3}cm^{-1}$) | *as*S134N ($10^{-3}cm^{-1}$) | *as*SOD ($10^{-3}cm^{-1}$) |
|---|---|---|---|---|
| 1 | 14.6 | - | 13.3 | 13.3 |
| 2 | 17.4 | 16.0 | 16.5 | 16.7 |
| 3 | 22.4 | 20.2 | 21.7 | 22.4 |
| 4 | 25.7 | 28.7 | 30.2 | 29.6 |
| 5 | 30.4 | 33.1 | 32.2 | 33.0 |
| | | | | |
| | R^2 | R^2 | R^2 | |
| | 0.99075 | 0.9696 | 0.9938 | |

**Figure 10:** CD spectra of the Cu binding region for mutants G37R H48Q S134N, and deconvolutions

The spectra of G37R, S134N and H48Q were fitted using Gaussian functions.

Up till now the bands for the S134N and G37R mutants were identified, since the spectral profiles are similar to *as*SOD one[85] and were fitted with the same number of Gaussians (actually the lowest band was not identified for G37R, due to the poor s/n ratio in the range 10000 -15000 $cm^{-1}$). Specifically, the two higher resonances are attributed to d-d transitions, while the lowest are ligand to metal charge transfer bands (LMCT). The small intermediate one could be either a further d-d transition or the lowest LMCT arising from the bridging imidazolato ligand. Just by rough comparison of the profiles of G37R and *as*S134N with *as*SOD one, it can be seen that there's a certain similarity between them, constituting an indication that these mutants roughly share the same ions binding geometry as *as*SOD, although the shift of the band resonances suggest a certain grade of distortion already evidenced by EPR. On the contrary, for the mutant H48Q, where one of the ligand histidines is mutated to a glutamine, the spectrum is completely different and so it is the deconvolution. The assignment of the fitting has not been performed.

## 2.6 High Field NMR

**State of the art**

Nuclear magnetic resonance (NMR) spectroscopy is the only bio-physical method which can provide structures at atomic resolution of biological molecules, such as proteins, in solution.
Importantly, biomolecular NMR spectrocopy do not only provide information about structural features of a macromolecule, helping also in the characterization of conformational dynamics and exchange processeses of biomolecules, at timescales ranging from picoseconds to seconds. Another key application of this technique is the determination of ligand binding and the mapping of interaction surfaces of protein/ligand complexes.
Nuclear magnetic resonance is a phenomenon which occurs when the nuclei of atoms given of non zero nuclear spin are immersed in a static magnetic field and exposed to a second oscillating magnetic field. The nuclei of naturally occurring atomic isotopes that constitute biological molecules either have a nuclear spin at natural abundance ($^{1}H$, $^{31}P$) or naturally less-abundant isotopes with a nuclear spin ($^{13}C$, $^{15}N$) can be incorporated into biomolecules by isotope labelling. Said spins give rise to a nuclear magnetic moment, which defines the basic resonance frequency for the nucleus of interest at a given magnetic field, (e.g. 600/150 MHz for $^{1}H/^{13}C$ at 14.1 Tesla.) while the exact resonance frequency depends on the chemical environment of each spin. Differences from the basic resonance, expressed in ppm, are called chemical shifts: their assignment is the first step to solve the structure of a polypeptide. Spins can interact through chemical bonds via J-coupling, providing information about dihedral angles, thus individuating the peptide backbone and side chain conformations, or through-space via nuclear Overhauser effect (NOE), which is a result of cross-relaxation between different spins. NOEs are typically only observed between protons which are separated by less than 5-6 Å and they yield the most important structural information derived from NMR, since their values depend on the spatial disposition of nuclei themselves. Recently, new NMR parameters, such as residual dipolar couplings (RDCs) and cross-correlated relaxation effects (CCRs) have been shown to provide distance independent projection angles for bond vectors (e.g. N−H and $C^{\alpha}$−$H^{\alpha}$ bonds in proteins).
Although NMR has a wide range of applications, for example solid state NMR methods for studying membrane proteins, crystallized or immobilized in lipid bilayers, or Magnetic Resonance Imaging (MRI) for in vivo imaging of human tissue, the predominant application of nuclear magnetic resonance in chemical and biological research is NMR in solution.

In this field, several different pieces of information can be achieved. Some examples regard 3D structure of proteins, nucleic acids, sugars and their complexes in solution, dynamics and mobility, as well as chemical exchange, conformational exchange and hydrogen exchange rates, very useful to determine the strength of hydrogen bonding. Also kinetics can be successfully investigated, such as enzyme mechanism and chemical reactions, and folding and unfolding studies, especially if NMR is coupled to stopped-flow techniques. NMR is very efficient in mapping interactions with other molecules, e.g. protein/protein, protein/nucleic acid, protein/ligand or nucleic acid/ligand interactions. Titration of a polypeptide with an interacting ligand induces changes in the NMR spectrum of the macromolecule, regarding the signals of atoms near the binding site. This chemical shift mapping, therefore, allows to localize the regions involved in the adduct formation and can be also used to determine dissociation constants ($K_d$).

Standard sets of experiments are used for the determination of a protein structure and its characterization; usually 1D experiments are not employed, although they can come of hand when investigating specific regions of the protein (e. g. the histidyl side chains, whose resonances fall in a clear region at high chemical shift), since they are way too crowded given the high number of non equivalent protons in a macromolecule.

By the introduction of additional spectral dimensions these profiles are simplified, allowing to resolve all the signals. If the spectrum is homonuclear, the additional dimensions will report signals of the same isotope (usually $^{1}H$), while $^{15}N$ and/or $^{13}C$ are usually the nuclei of choice in the heteronuclear NMR spectroscopy of proteins. Since the natural abundance of $^{15}N$ and $^{13}C$ is very low and their gyromagnetic ratio is markedly lower than the one of protons, usually the two strategies for increasing the low sensitivity of these nuclei are isotopic enrichment of proteins and enhancement of the signal to noise ratio by the use of inverse NMR experiments where the magnetization is tranferred from protons to the heteronuclei. A 2D experiment is usually composed of four steps. The first is nuclei preparation, where the magnetization is labelled with the chemical shift of the first nucleus, followed by an indirect evolution period of spins free precession. After this there will be a mixing time, where the magnetization is transferred from the first nucleus to a second one; mixing sequences exploit two different mechanisms for magnetization transfer: scalar coupling (J coupling) or dipolar interaction (NOE). Data are acquired at the end of the experiment; during acquisition the magnetization is labelled with the chemical shift of the second nucleus.. Two dimensional Fourier Transform yields the 2D spectrum with two frequency axes.

Classical example of homonuclear 2D pulse sequences are COSY and TOCSY, where the magnetization transfer is accomplished by mean of scalar coupling, and NOESY. The NOESY experiment exploits dipolar interaction of spins and is crucial for the determination of protein

structure, since it correlates all protons which are close enough, not only in the amino acid sequence, but also in space due to tertiary folding. The intensity of the NOE is in first approximation proportional to $1/r^6$, with r being the distance between the protons.

The most important heteronuclear NMR experiment is the $^{1}H$ - $^{15}N$ HSQC (heteronuclear single quantum correlation) which correlates the nitrogen atom of an NH group with the directly attached proton(s).

When analyzing relatively large proteins, it can happen that even $^{1}H$-$^{1}H$ 2D spectra, like NOESY or TOCSY, are too crowded to resolve part of the signals. To overcome this problem it's possible to further spread out the signals in a third dimension, so that they are distributed in a cube instead of a plane, by extending NOESY or TOCSY with a HSQC step. The resulting 3D experiment will be called NOESY-HSQC or TOCSY-HSQC.

Triple resonance experiments can also be performed, resulting in the correlation of three different nuclei ($^{1}H$, $^{13}C$, $^{15}N$) so to accomplish the sequential assignment of larger proteins They can obviously be performed only on doubly labelled ($^{13}C$, $^{15}N$) proteins and their names descend from the ordered names of all nuclei involved in the magnetization transfer during the experiment, with brackets around the nuclei used only for transfer, whose frequencies are not detected (e.g. HNCO and HN(CA)CO).

Also mobility is of high relevance for the characterization of a polypeptide in terms of dynamical behaviour, which is intimately connected with structural changes and molecular flexibility. Moreover, highly mobile regions of the proteins are usually associated with molecular recognition, which is an essential step for polypeptide specific aggregation[19]. To evaluate such features, the measurement of $^{15}N$ relaxation rates of backbone amides in isotopically enriched proteins is usually employed. Different classes of dynamics are related to different kind of polypeptide motions, and can be investigated by mean of several NMR experiments, as it is reported in the figure below. Specifically, the detection of dynamics in the subnanoseconds range is crucial to individuate fast mobility of the backbone, and can be performed by mean of $T_1$, $T_2$ and NOE measurements. Conformational exchange phenomena can be revealed by $R_2$ values larger than the average and their time scale determined using CPMG and $T_{1\rho}$ experiments.

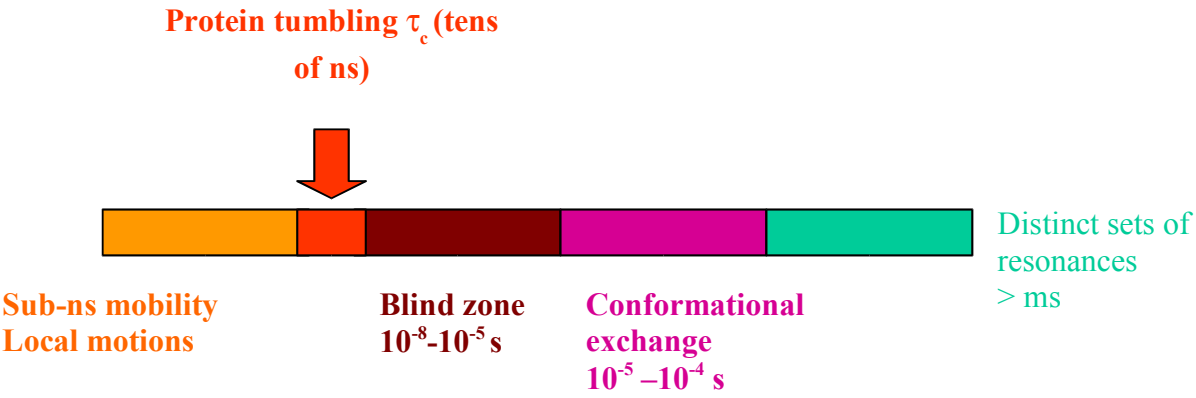



**Figure 11:** timescales of proteins motions

## Results

The mutants G37R and *as*S134N were thoroughly characterized by High Field NMR. On all of them, a preliminary comparison of the general structural features was performed by $^{15}$N-HSQC-NOESY at 900 MHz; mobility experiments ($T_1$, $T_2$ and NOE measurements) were also performed, at 600 MHz.

The acquired data evidenced a high structural homology, among the mutants and with *as*SOD, with only residues in the mutated regions undergoing relevant shifts. An heterogeneous state of metallation was found for G37R, whose spectrum revealed the presence of two differently metallated forms, characterized by unequal relative abundances. The doubling of signals was clustered in the region nearby the copper and zinc sites, with the shifts values of the less abundant species in better agreement with those found for the native protein, and involved all the binding histidines, with the exception of residue 80 (one of the zinc ligands) which could experience the same chemical shift in both forms. By comparison with the spectrum of *as*SOD, the less intense signals were tentatively assigned to the correct metallated state of the mutant.

Mobility experiments were performed to detect whether there were backbone regions characterized by a faster mobility with respect to the average, since *as*SOD is a quite rigid protein given of uniform mobility values all along the backbone and any deviation from such behaviour in the mutants is a sign of fast fast local motions. *as*S134N showed higher mobility in the ps-ns timescale for the backbone region around the mutation and this is in good agreement with its aggregating properties, while the non aggregating G37R was characterized by a high uniformity of $T_1$, $T_2$ and NOE, that are overally identical to native SOD ones.

For *as*S134N, several 3D experiments were performed, aiming at extensive spectra assignment; namely HNHA, HCCH-TOCSY, and $^{13}$C-HSQC-NOESY, on a Bruker Avance 900 MHz, while HNCOCA and HNCA spectra were registered at 500 MHz. The structural constraints were used to model a structure of the mutation surrounding region, using the already solved structure of *as*SOD as template. Such modelling evidenced that the only appreciable structural change is the distortion of the loop facing the mutation, so to give room for the bulkier residue

Concentration and time dependent experiments were performed at 800 MHz cryo-probe; H/D exchange kinetic was studied at 900 MHz cryo-probe.

For *as*S134N these studies were very useful in revealing the presence of oligomeric soluble aggregates; in fact, the dilution titration evidenced that residues around the mutation underwent relevant shift under dilution, a probable indication of their involvement in the formation of protein oligomers, while the H/D exchange confirmed the mobility data by showing an increase of the hydrogen exchange rate in the vicinity of the mutation, that resulted more solvent-exposed. Such data confirmed also the findings of the structural characterization, that evidenced the disappearing of the small helix in the region 132 – 137, which should exchange much slower if the structure is conserved.

Also for G37R mutant a dilution titration was performed, which showed no changes in the shifts of the polypeptide, even at very low concentrations. Since G37R didn't evidence modifications in the structure and mobility for the most part of the protein, H/D exchange was not necessary.

To better assess G37R metallation state, its apo form was titrated with both Cu and Zn and the metal uptake kinetics were followed in the aromatic region (10-16 ppm) of monodimensional $^{1}$H spectra, monitoring the evolution of imidazole NH protons signals. As a matter of comparison such experiments were performed also on apo *as*SOD, which showed an higher affinity towards the Cu ion with respect to it G37R mutant, although the kinetic of metal incorporation was very slow even in this case, spanning in the order of days. The second evidence, gathered by titration with two equivalent of Zn (i.e. to have $Zn_4$ G37R SOD) and one equivalent of Zn and one of Cu (i.e. to have $Cu_2Zn_2$ G37R SOD) on distinct apo G37R samples, confirmed the assumption made

on the basis of the $^{15}$N-HSQC that the sample is a combination of $Zn_4$ and $Cu_2Zn_2$G37R SOD species in a 3:1 ratio.

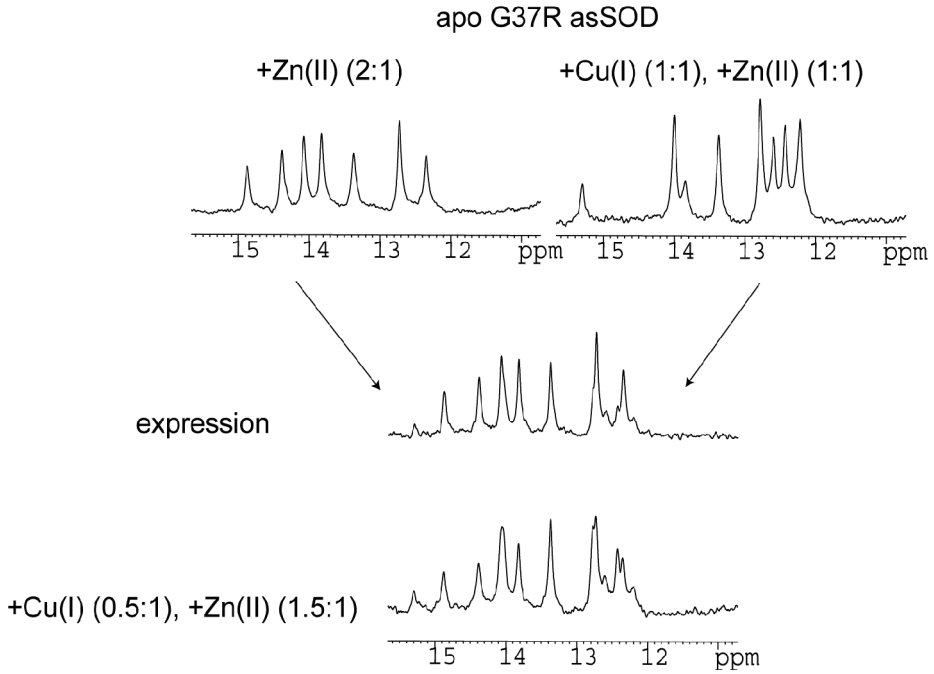



**Figure 12**: NMR spectra of imidazole amide region of G37R mutant The apo proteins was titrated with one equivalent of zinc(II) ion. Subsequently, another equivalent of either zinc(II) or Cu(I) was added. The 1D spectrum of G37R mutant resulting from expression in E.Coli (central) is clearly a combination of the spectra of the Zn-Zn and the Cu-Zn forms (B,top). The spectrum is also well reproduced by titrating the apo protein with 1.5 and 0.5 equivalents of zinc(II) and Cu(I), respectively (B,bottom).

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochemical and Biophysical Research Communications 328 (2005) 633–639
BBRC
www.elsevier.com/locate/ybbrc

# $^1$H nuclear magnetic relaxation dispersion of Cu,Zn superoxide dismutase in the native and guanidinium-induced unfolded forms☆

Elisa Libralesso[a], Kirill Nerinovski[a], Giacomo Parigi[b,*], Paola Turano[a]

[a] CERM and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy
[b] CERM and Department of Agricultural Biotechnology, University of Florence, P.le delle Cascine 28, 50144 Florence, Italy

Received 27 December 2004
Available online 19 January 2005

**Abstract**

Potentialities and limitations of the use of $^1$H NMRD technique for the characterization of the hydration properties of unfolded or partially folded states of proteins are discussed. The copper(I) form of monomeric Cu,Zn superoxide dismutase in its folded state and in the presence of 4 M guanidinium chloride is taken as case system. The dispersion profile, analyzed with an extended relaxation matrix analysis, indicates the presence of long-lived water molecules in the folded state. The observed increase in relaxation at high field upon addition of guanidinium chloride indicates an increase in the number of solvation protons interacting with the protein and exchanging with a time shorter than the protein reorientational time. The observed effect is consistent with an exposed protein surface of SOD in the presence of 4 M guanidinium chloride smaller than what could be expected for a random coil.


*Keywords:* Hydration; Unfolded proteins; SOD; Relaxometry; NMRD

The protein unfolding process is essentially a trade-off between the loss of intramolecular interactions and the gain of protein–solvent and protein–unfolding agent interactions (with the associated enthalpy and entropy variations). Understanding the details of the changes occurring upon unfolding of a specific protein implies drawing a satisfactory picture of all relevant interactions. High resolution NMR is able to shed light on the intramolecular interactions at the atomic level and also, to some extent, on the protein–solvent interactions as viewed from the protein side. A less widespread technique, nuclear magnetic relaxation dispersion (NMRD), can complement the information on the protein–solvent and protein–cosolvent interactions as viewed from the water side.

Since the pioneering studies of Koenig [1], the $^1$H NMRD profiles have been interpreted to obtain information on the reorientational time ($\tau_R$) of the protein in solution and on the number of hydration water molecules exchanging with a rate faster or slower than $\tau_R$. The observed proton relaxation is the sum of the contribution from all protons in the sample as a function of the magnetic field. In a protein solution, there are at least four pools of protons to be considered: (i) bulk water protons, (ii) protons of exchanging water molecules interacting with the protein, and (iii) exchangeable and (iv) non-exchangeable protein protons [2,3]. Due to the low sensitivity and low resolution of this technique, only bulk water protons actually affect the measured profiles in diluted protein samples. However, information can indirectly be obtained on the second and third pools of protons from the analysis of the profiles

☆ *Abbreviations:* $^1$H NMRD, nuclear magnetic resonance dispersion; SOD, superoxide dismutase; GdmCl, guanidinium chloride.

* Corresponding author. Fax: +39 055 4574253.
*E-mail address:* parigi@cerm.unifi.it (G. Parigi).

0006-291X/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.bbrc.2005.01.027

acquired in the presence and in the absence of the protein in solution [3,4].

In the last decade, hydration studies on proteins have been performed by using water $^{17}O$ and $^{2}H$ relaxation dispersions in isotopically enriched $H_2O$ to monitor directly the internal and external hydration of different proteins in native, molten globule, and unfolded states [5–7]. This approach has the advantage that the measured relaxation rates are not affected by the contribution of protein nuclei. The work by Halle and coworkers [8] has clarified that the main contribution to bulk water relaxation is provided by few water molecules trapped in protein cavities rather than by water molecules in the protein hydration layer as proposed in the past.

Here we show that, under certain pH conditions and in the assumption of rigid systems, the contribution of the exchanging water molecules interacting with the protein can be separated from that of the exchangeable protein protons also in $^{1}H$ NMRD profiles. Indeed, the contribution from these two pools can be calculated using a modified version of the program CORMA [9,10] taking into account the available protein structure and the exchange rates of protein protons. Implications deriving from water and protein dynamics are then discussed. The advantage of $^{1}H$ NMRD profiles is that they can be more easily obtained than $^{17}O$ or $^{2}H$ NMRD profiles for sensitivity reasons, and can now be collected with commercial instruments which permit the measurement of the relaxation rates over several decades of proton Larmor frequency, typically from <10 kHz to 50 MHz, called relaxometers [11]. Measurements at higher fields can be performed, if needed, using conventional NMR spectrometers. Such accessible range of frequencies permits the evaluation of the dispersion characterizing the spectral density functions, related to the reorientational rate of the protein. The latter is typically centered from a few tens of megaHertz, for proteins of thousand Dalton, to a few hundred kiloHertz, for proteins of hundred thousand Dalton.

We report here the $^{1}H$ NMRD profiles of Cu,Zn superoxide dismutase, monomeric form (SOD hereafter), in its natively folded state and in the unfolded state obtained in the presence of 4 M guanidinium chloride (GdmCl, hereafter). The unfolding of SOD is a process that is gaining increasing importance, given the recent view that identifies the formation of possibly misfolded SOD aggregates as responsible for the neurological disease familial amyotrophic lateral sclerosis (FALS) [12]. Folded Cu,ZnSOD is a dimeric protein constituted by two identical subunits characterized by a relatively rigid β-barrel structure [13–15], as exemplified by the $^{15}N$ NMR heteronuclear mobility data [16]. The essential structural and dynamic features are conserved also in the monomeric form [16,17]. The two metal ions are bound with high affinity and specificity to protein side chain donor atoms provided by His46, His48, and His120 for copper(I) and His63, His71, His80, and Asp83 for zinc(II). In the oxidized form of the enzyme His63 ligand bridges the two metal ions [18,19]. The protein form of monomeric SOD dominant at high concentrations of GdmCl (2.5 M or higher) is still relatively globular (with intact disulfide bridge between Cys57 and Cys146, and aspecifically bound metal ions) with a hydrodynamic radius 20% larger than in the native protein; it does not possess residual secondary structure; it is largely fluxional as indicated by a sensible drop in the $^{15}N$ $R_2$ and $\{^{1}H\}$–$^{15}N$ NOE values; and it also exhibits increased proton exchangeability (as monitored through CLEANEX-type experiments) [20].

## Materials and methods

*Sample preparation.* The monomeric form of human SOD, where the Phe50 and Gly51 residues at the subunit–subunit interface are replaced by two hydrophilic Glu residues and where further mutation at position 133 (Glu to Gln) was introduced, was expressed, isolated, and purified as reported [21,22]. Protein metallation and copper reduction were accomplished as described in detail in the literature [17,21]. NMRD measurements were performed on a set of 4 mM enzyme samples in 20 mM phosphate buffer (pH 5.0), with GdmCl concentrations varying from 0 to 4 M.

*Relaxation rate measurements.* Longitudinal water proton relaxation rates were measured with a prototype of a Stelar fast field cycling relaxometer (0.01–40 MHz proton Larmor frequency range) [23]. The instrument provides $R_1$ values with an error smaller than 1%. $^{1}H$ NMRD profiles were obtained by plotting proton relaxation rates as a function of applied magnetic field. Reported profiles are calculated after subtraction of relaxation rates measured for the solution in the absence of the protein from relaxation rate measured for the 4 mM protein solution.

*Theoretical background.* The typical water $^{1}H$ NMRD profile for a protein solution shows a low field plateau, a decrease—called dispersion—and a non-zero high field limit value. The qualitative interpretation of the profile is straightforward. Water proton relaxation is dominated by homonuclear dipole–dipole interactions. The longitudinal relaxation rate enhancement $R_1$ for a water proton, interacting for a time $\tau_M$ with the protein environment, is given by

$$R_1 = \frac{f_M}{T_{1M} + \tau_M}, \tag{1}$$

where $f_M$ is the molar fraction of bound protons with respect to the bulk water protons, and $T_{1M}^{-1} = R_{1M}$ is the relaxation rate of the proton in the bound state.

Under fast exchange of hydrogen nuclei between protein sites and bulk water, the water proton longitudinal relaxation rate can be written as in Eq. (2) [3,24,25]

$$R_1 = \beta(0.2J(\omega) + 0.8J(2\omega)), \tag{2}$$

where β is a constant related to the dipolar interaction between both protein and water exchangeable protons and their neighbors in the protein-bound state. These interactions include both the intermolecular $^{1}H$ water–$^{1}H$ protein interactions, and the intramolecular $^{1}H$ water–$^{1}H$ water and $^{1}H$ protein–$^{1}H$ protein interactions [26]. The spectral densities $J(\omega)$ are Lorentzian functions given by

$$J(\omega) = \frac{\tau}{1 + \omega^2\tau^2} \tag{3}$$

with $\omega$ equal to $2\pi$ times the proton Larmor frequency and $\tau$ is the correlation time which modulates the dipolar interaction. The correlation time $\tau$ is defined as

$$\tau^{-1} = \tau_R^{-1} + \tau_M^{-1}, \quad (4)$$

where $\tau_R$ is the reorientational time of protein molecules. Exchangeable protein protons exchange with a time usually much longer than the reorientational time, which is thus considered the correlation time for the dipolar interaction. Water molecules buried inside the protein can have lifetimes longer (long-lived) or shorter (short-lived) than the reorientational time, depending on their location and interaction with the surrounding protein matrix [4,25,27]. A sum of two terms of the type of Eq. (2) can thus account for the whole NMRD profile. A first term with $\tau = \tau_R$ reproduces the dispersion part of the profile, due to long-lived water molecules and exchangeable protein protons, whose dipolar interactions are modulated by the protein rotational correlation time. The second term, with $\tau = \tau_M \ll \tau_R$, can account for the high field non-dispersive plateau due to short-lived water molecules whose dipolar interactions are modulated by their short residence time and whose dispersion thus occurs beyond the accessible magnetic field range. This non-dispersive contribution is thus a constant often termed $\alpha$, and the whole NMRD profile is described by the following equation:

$$R_1 = \alpha + \beta(0.2J(\omega) + 0.8J(2\omega)). \quad (5)$$

As already mentioned, Eqs. (1) and (5) are valid in the assumption that proton residence times $\tau_M$ are longer than $\tau_R$ but at the same time shorter than their intrinsic relaxation time $T_{1M}$ (Eq. (1)). Therefore, contributions to $\beta$ are provided only by protons exchanging with a rate faster than ca. 100 $s^{-1}$. Differently, the increase in relaxation cannot be transmitted from the protons dipolarly coupled with protein protons to bulk water protons.

It is however well known that Eq. (5) may not be able to satisfactorily fit experimental $^1$H NMRD profiles, which often show a stretched dispersion. This is due to local mobility of protein groups (typically lysines, arginines, etc.) or bound water molecules [3], which may stretch the dispersion to the right, and to partial aggregation, that may increase the low field plateau and sensibly stretch the dispersion to the left. Therefore, a model-free approach was introduced by Halle [28], and the $J(\omega)$ functions in Eq. (1) are written as a sum of a finite number of Lorentzians, each of them characterized by a different correlation time and a normalized weighting factor [29].

In a different approach, relaxation rate profiles can be simulated by using a modified version [10] of the program CORMA [9], developed in our laboratory (see the Supplementary materials). This program calculates the complete relaxation matrix by including, in the modified version, long-lived water protons, and bulk water protons, besides the protein protons. Such an approach relies on the availability of a 3D protein structure and the position of long-lived water molecules. Estimates of the exchange rate constants between bulk water protons and long-lived water protons or protein protons must be provided, to be inserted in the relaxation matrix [10]. This is the approach we have chosen in the present work to simulate the dispersive term. In this way, besides treating without approximations all the dipolar interactions of long-lived exchangeable protons, the effect of a long $\tau_M$ on Eq. (1) can also be explicitly accounted for.

## Results and discussion

### *NMRD dispersion of the folded protein*

The $^1$H NMRD profile of monomeric SOD at pH 5 and 298 K in the absence of GdmCl is reported in Fig. 1 (filled symbols), after subtraction of the buffer relaxation rates acquired under the same conditions and normalization to 1 mM protein concentration. The profile has a good Lorentzian shape, with correlation time $\tau = 10.1 \pm 0.5$ ns. Such value is in accordance with previous findings based on $^{15}$N relaxation measurements, which provided an estimate for the reorientational correlation time of 9.3 ns [16]. The values of $\alpha$ and $\beta$ are $0.016 \pm 0.002$ $s^{-1}$ $mM^{-1}$ and $4.6 \pm 0.2 \times 10^6$ $s^{-2}$ $mM^{-1}$, respectively.

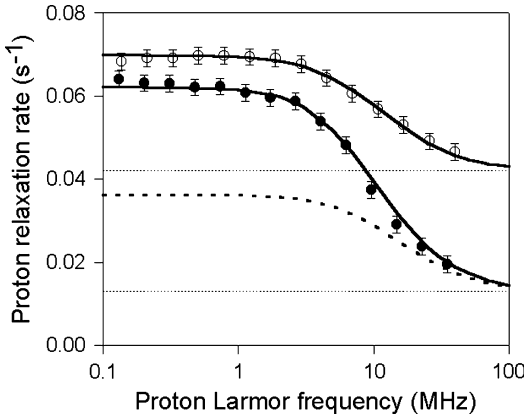



Fig. 1. Solvent $^1$H NMRD profiles for a solution containing the monomeric form of human Cu(I),ZnSOD at 298 K and pH 5 in the absence (filled symbols) and in the presence (empty symbols) of 4 M guanidinium chloride. The profiles have been normalized to 1 mM of protein concentration and have been subtracted of the relaxation rates measured in the absence of the protein. The curves indicate the profiles calculated with the modified CORMA program. The PDB structure 1MFM was considered, and side chain hydroxyl protons of threonine and serine amino acids were exchanged at $10^3$ $s^{-1}$. The dashed line indicates the contribution from side chain hydroxyl protons of threonine and serine amino acids of SOD in the fold form. The dotted horizontal lines represent the high field relaxation rates, due to short-lived or diffusing water protons, which contribution cannot be calculated by CORMA.

In order to investigate the origin of the contribution to relaxation rate from exchangeable protein and/or water protons, we simulated the experimental profiles with the modified version of the program CORMA [10]. At pH 5 only side chain hydroxyl protons of threonine and serine amino acids (the protein does not contain tyrosine residues) exchange fast enough to contribute to bulk water proton relaxation [30,31]. Considering an exchange rate $k_m$ for these protons of about $10^3$ $s^{-1}$, for externally oriented residues, assuming rigid side chains and a reorientational time of 11 ns, the NMRD profile has been calculated as reported in Fig. 1 by including the presence of one long-lived water molecule with and exchange rate larger than $10^3$ $s^{-1}$ (and smaller than $1/\tau_R$). Short-lived water molecules (exchanging much faster than the reorientational time) cannot contribute to the dispersive term, whereas long-lived water protons do. It was not possible to reproduce the data satisfactorily without at least one long-lived water molecule.

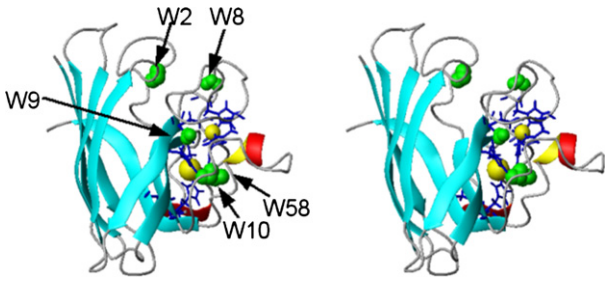



Fig. 2. Stereoview of the structure of monomeric Cu,ZnSOD as it results from the X-ray structure (PDB 1MFM) [32]. Spacefill representation is used for the metal ions and for the five crystallographic water molecules that meet the requirements for long-lived solvent molecules. Metal ligands are represented as sticks.

The presence of long-lived water molecules is consistent with the structural features of monomeric SOD, where up to five crystallographic water molecules (W2, W8, W9, W10, and W58 in the 1MFM pdb file (1.02 Å resolution) [32]) fit the requirements [27] to be considered long-lived (Fig. 2): these are three buried water molecules (W8, W9, and W10), one cleft water (W2), and one water molecule (W58) known to be semi-coordinated to the copper ion, and to exchange fast, as indicated by previous relaxation measurements performed on the oxidized protein [33]. It should be recalled that mobility of side chain protons and internal motions of buried waters may sizably decrease their contribution to the dispersive term. In fact, on moving consistently with the tumbling of the protein, protons may experience motions on the subnanosecond time scale, thus resulting in being only partially ordered. The contribution of a rigid water molecule or of a rigid side chain proton to the β term is thus scaled down by a factor ($S^2$) in the presence of local motions [34]. Therefore, even more than 1–2 water molecules may actually contribute to the observed dispersive term. Such indetermination in the number of long-lived water molecules due to their possible partial ordering has been reported also for other proteins investigated through $^2$H or $^{17}$O NMRD because the value of $S^2$ remains unknown [7]. In any case, it is clear that some contribution from long-lived water, besides that of exchangeable protons, is needed to account for the experimentally observed dispersion.

*NMRD dispersion of the unfolded protein*

The $^1$H NMRD profile of monomeric SOD at pH 5 and 298 K in the presence of GdmCl is also reported in Fig. 1 (empty symbols), after subtraction of the relaxation rates of an equimolar GdmCl solution in the same buffer, and normalization to 1 mM protein concentration. The profile of the unfolded protein shows similar features to those of the native one: the relaxation rate is constant at low fields, until a dispersion occurs. The low field plateau has similar relaxation values in the two profiles, and also the dispersion is positioned at a similar frequency. However, the jump in relaxation rates before and after the dispersion is clearly different, thus indicating different values for both the parameters $\alpha$ and $\beta$; furthermore, the dispersion is clearly more stretched than in the native protein. A preliminary fit with a single dispersion provides a $\tau$ value of about 9 ns, but the fit is less good with respect to the native protein; a fit with two dispersions yields a much better agreement, the resulting $\tau$ values being in the range 10–17 and 2–6 ns. The value of $\alpha$ is $0.044 \pm 0.002\ s^{-1}\ mM^{-1}$. A $\tau$ value of 17 ns would be in agreement with high resolution NMR data, which indicate an increase of about 20% in hydrodynamic radius upon unfolding. The height of the dispersion should be explained by the contribution from labile protein protons, the probable loss of long-lived waters, and the decrease in the order parameter. Again, we turned to CORMA simulations to better analyze the origin of the dispersion using as limit structures that of the native protein and that of a fully random polypeptide with the same sequence, and by imposing only one $\tau$ value of 17 ns. In Fig. 1, the profile obtained with the native protein structure, calculated without including any long-lived water molecule, is shown. The profile calculated using the completely unfolded structure shows a dispersion which is about 50% smaller than the experimental one, and thus the contribution of long-lived water molecules should be taken into account. This is obviously an unreasonable scenario. Inclusion of mobility effects would further reduce the contributions to the dispersion. The amplitude of the observed dispersion thus indicates that the protein cannot be completely unfolded.

The fact that only one dispersion with a $\tau$ value of 17 ns calculated using CORMA is capable to reproduce the experimental profile, whose dispersion can be fit to Eq. (5) with an effective $\tau$ value of 9 ns, is worth a further comment. As in the unfolded protein there is apparently no contribution from long-lived water molecules, the only contributors to the dispersive term are the exchangeable protons. Their exchange rates are typically [30,31] in a range that is fast enough to contribute to water proton relaxation, but not fast enough to be negligible with respect to $T_{1\,M}$ in Eq. (1), especially when $T_{1\,M}$ is shortest, i.e., at low field. As a consequence, the overall $R_1$ is more limited by exchange at low field than it is at high field, and the dispersion is distorted and seemingly shifted to higher field, yielding a smaller apparent $\tau$ value.

*The α term*

The α term, as explained earlier, is due to short-lived solvent molecules (water and guanidinium in this case)

that surround the protein. It also contains the contribution from protein protons and water protons experiencing fast (subnanoseconds) local motion. Water molecules in the hydration layer at the protein surface have typical lifetimes ranging from a few to hundreds of picoseconds, with a tail extending up to 1 ns [27,35,36], governed by hydrogen-bond dynamics. As an example, in the folded protein, assuming rigid side chains, we may think that out of the about 500 water molecules that can be estimated to be in contact with the protein surface, about 120 water molecules exchanging with a rate of $10^{10}\ s^{-1}$ (two orders of magnitude faster than the reorientational time) provide a contribution to $\alpha$ equal to $0.016\ s^{-1}\ mM^{-1}$ (as found for the protein under native conditions), each water molecule contributing to relaxation for about $1.3 \times 10^{-4}\ s^{-1}$.

The sizable increase in the $\alpha$ term in the unfolded protein (close to a factor 3) can be read as an indication that the contribution to relaxation ascribed to the pool of protons exchanging faster than the reorientational time of the protein increases. This increase should be analyzed in terms of short-lived water contribution, guanidinium contribution, and protein side chain mobility. In the unrealistic case that side chain mobility remains the same as in the native protein, and ignoring the guanidinium contribution, the solvent water molecules are the unique ones responsible for the increase of $\alpha$, which is equal to $0.030\ s^{-1}\ mM^{-1}$, and about 230 additional water molecules should be considered with a dipolar interaction modulated by a lifetime of $10^{-10}$ s. A factor 2.3–2.5 increase of the accessible surface is at the limit calculated for a family of randomly generated model structures where only the disulfide bond and the coordination bonds with the metal ions are maintained (the ligand side chains being free to rotate).

However, we should also take into account the presence of 4 M guanidinium ions in solution. At variance with $^{17}O$ measurements, $^{1}H$ NMRD is sensitive to all kinds of solvent protons simultaneously. Short lived guanidinium ions on the protein surface can be present together, and actually compete, with water molecules [5]. Since the proton exchange between water and guanidinium at pH 5 in the bulk solution is of the order of $10–100\ s^{-1}$, it is expected that, from the NMRD point of view, a single proton pool is operative, comprising both water and guanidinium protons. Interestingly, each guanidinium ion bears three times the number of protons of a water molecule. In a limiting scenario, all water molecules responsible for the $\alpha$ term in the native protein could be substituted by guanidinium ions with the same lifetimes, and this would match with the factor 3 increase in the $\alpha$ term upon unfolding without any increase in protein solvation. Both the limit cases (no water molecules interacting with the protein, no guanidinium ions interacting with the protein) are unrealistic. Moreover, the presence of larger side chain mobility in the guanidinium-induced unfolded form will further decrease the contribution of protein solvation to the $\alpha$ value. Therefore, we are left with the picture that protein solvation is larger than that of the folded protein but sizably smaller than what could be expected for a randomly generated model structure.

The above considerations imply no or little change in the distribution of the lifetimes of water molecules on passing from a native to an unfolded state. As pointed out by Halle [5,6,8,37], these changes are hard to predict: while in the present case it is clear that the binding sites for long-lived water molecules are lost upon unfolding, it is possible that some sites remain, or even be created, where water and/or guanidinium molecules can be hosted for longer times than the average surface waters. In this case, the increase in the $\alpha$ term could in fact arise from a relatively few water/guanidinium molecules of intermediate lifetime, and this would imply a further more modest increase in solvent exposure.

## Conclusions

Information on hydration and dynamics of the investigated system is contained in all the main features of the $^{1}H$ NMRD profiles, which are (i) the position of the dispersion, determined by the average correlation time $\tau$, (ii) the high field (after the dispersion) value of relaxation rates, $\alpha$, and (iii) the increase in the low field value of relaxation rates with respect to the high field value.

The observed decrease in the value of $\beta$, and the increase in $\tau_R$ accompanying guanidinium-induced unfolding were reproduced with an extended relaxation matrix analysis (program CORMA), which permitted us to take into account the slow exchange regime for protein protons. The relatively small increase in the $\alpha$ value provides information on the increase in solvent exposed protein surface in the presence of denaturant.

The observed lengthening in the reorientational time of the protein upon addition of GdmCl compares well with that reported with other techniques for the transition from a folded to a molten globular structure. The 3-fold increase in the contribution from solvation protons in GdmCl is again consistent with an increase in the solvent exposed surface, which is much smaller than what could be expected for a randomly coiled peptide chain where only the disulfide bond and the coordination bonds with the metal ions are maintained.

In conclusion, the unfolding process monitored looking at the relaxation properties of water molecules indicates a relatively compact structure for the protein in the presence of high concentration of GdmCl, a picture which is consistent with the findings of high resolution NMR studies.

## Acknowledgments

We are grateful to Ivano Bertini and Claudio Luchinat for extensive discussion, corrections, and improvements of the manuscript. This work was supported by MIUR-FIRB, Italy, and the European Union (contract HPRI-CT-2001-50028).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bbrc.2005.01.027.

## References

[1] S.H. Koenig, R.D. Brown III, The dynamics of water–protein interactions. Results from measurements of nuclear magnetic relaxation dispersion, Progr. NMR Spectrosc. 22 (1991) 487–567.
[2] B.P. Hills, The proton exchange cross-relaxation model of water relaxation in biopolymer systems, Mol. Phys. 76 (1992) 489–508.
[3] K. Venu, V.P. Denisov, B. Halle, Water 1H magnetic relaxation dispersion in protein solutions. A quantitative assessment of internal hydration, proton exchange, and cross-relaxation, J. Am. Chem. Soc. 119 (1997) 3122–3134.
[4] B. Halle, V.P. Denisov, K. Venu, Multinuclear relaxation dispersion studies of protein hydration, Biol. Magn. Reson. 17 (1999) 419–483.
[5] V.P. Denisov, B.-H. Jonsson, B. Halle, Hydration of denatured and molten globule proteins, Nat. Struct. Biol. 6 (1999) 253–260.
[6] K. Modig, E. Kurian, F.G. Prendergast, B. Halle, Water and urea interaction with the native and unfolded forms of a beta-barrel protein, Protein Sci. 12 (2003) 2768–2781.
[7] V.P. Denisov, B. Halle, Thermal denaturation of Ribonuclease A characterized by water oxygen-17 and deuterium magnetic relaxation dispersion, Biochemistry 34 (1998) 9046–9051.
[8] K. Modig, E. Liepinsh, G. Otting, B. Halle, Dynamics of protein and peptide hydration, J. Am. Chem. Soc. 126 (2004) 102–114.
[9] B. Borgias, M. Gochin, D.J. Kerwood, T.L. James, Relaxation matrix analysis of 2D NMR data, Progr. NMR Spectrosc. 22 (1990) 83–100.
[10] L. Banci, S. Berners-Price, I. Bertini, V. Clementi, C. Luchinat, G.A. Spyroulias, P. Turano, Water–protein interaction in native and partially unfolded equine cytochrome *c* (Dedicated to Prof. R.R. Ernst), Mol. Phys. 95 (1998) 797–808.
[11] Ferrante, G. Sykora, S. Technical aspects of the fast field cycling, Adv. Inorg. Chem. 57 (2005) (in press).
[12] J.S. Valentine, P.J. Hart, Bioinorganic chemistry special feature: misfolded CuZnSOD and amyotrophic lateral sclerosis, Proc. Natl. Acad. Sci. USA 100 (2003) 3617–3622.
[13] E.D. Getzoff, J.A. Tainer, M.M. Stempien, G.I. Bell, R.A. Hallewell, Evolution of CuZn superoxide dismutase and the greek key β-barrel structural motif, Proteins Struct. Funct. Genet. 5 (1989) 322–336.
[14] J.A. Tainer, E.D. Getzoff, K.M. Beem, J.S. Richardson, D.C. Richardson, Determination and analysis of 2 A structure of copper zinc superoxide dismutase, J. Mol. Biol. 160 (1982) 181–217.
[15] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, M.S. Viezzoli, The solution structure of reduced dimeric copper zinc SOD: the structural effects of dimerization, Eur. J. Biochem. 269 (2002) 1905–1915.
[16] L. Banci, I. Bertini, F. Cramaro, R. Del Conte, A. Rosato, M.S. Viezzoli, Backbone dynamics of human Cu,Zn superoxide dismutase and of its monomeric F50/EG51E/E133Q mutant: the influence of dimerization on mobility function, Biochemistry 39 (2000) 9108–9118.
[17] L. Banci, I. Bertini, R. Del Conte, S. Mangani, M.S. Viezzoli, R. Fadin, The solution structure of a monomeric reduced form of human copper, zinc superoxide dismutase bearing the same charge as the native protein, J. Biol. Inorg. Chem. 4 (1999) 795–803.
[18] I. Bertini, C. Luchinat, R. Monnanni, Evidence of the breaking of the copper-imidazolate bridge in copper/cobalt-substituted superoxide dismutase upon reduction of the copper(II) centers, J. Am. Chem. Soc. 107 (1985) 2178–2179.
[19] I. Bertini, F. Capozzi, C. Luchinat, M. Piccioli, M.S. Viezzoli, Assignment of active site protons in the $^{1}$H NMR spectrum of reduced human Cu,Zn superoxide dismutase, Eur. J. Biochem. 197 (1991) 691–697.
[20] M. Assfalg, L. Banci, I. Bertini, P. Turano, P. Vasos, Superoxide dismutase folding/unfolding pathway: role of the metal ions in modulating structural and dynamical features, J. Mol. Biol. 330 (2003) 145–158.
[21] L. Banci, I. Bertini, C.Y. Chiu, G.T. Mullenbach, M.S. Viezzoli, Synthesis and characterization of a monomeric mutein of Cu/Zn superoxide dismutase with partially reconstituted enzymatic activity, Eur. J. Biochem. 234 (1995) 855–860.
[22] E.D. Getzoff, D.E. Cabelli, C.L. Fisher, H.E. Parge, M.S. Viezzoli, L. Banci, R.A. Hallewell, Faster superoxide dismutase mutants designed by enhancing electrostatic guidance, Nature 358 (1992) 347–351.
[23] I. Bertini, C. Luchinat, G. Parigi, Solution NMR of Paramagnetic Molecules, Elsevier, Amsterdam, 2001.
[24] A. Abragam, The Principles of Nuclear Magnetism, Oxford University Press, Oxford, 1961.
[25] V.P. Denisov, B. Halle, Direct observation of calcium-coordinated water in Calbindin D9k by nuclear magnetic relaxation dispersion, J. Am. Chem. Soc. 117 (1995) 8456–8465.
[26] S. Kiihne, R.G. Bryant, Protein-bound water molecule counting by resolution of 1H spin–lattice relaxation mechanisms, Biophys. J. 78 (2000) 2163–2169.
[27] V.P. Denisov, B. Halle, Protein hydration dynamics in aqueous solution, Faraday Discuss. 103 (1996) 227–244.
[28] B. Halle, H. Jóhannesson, K. Venu, Model-free analysis of stretched relaxation dispersions, J. Magn. Reson. 135 (1998) 1–13.
[29] I. Bertini, M. Fragai, C. Luchinat, G. Parigi, $^{1}$H-NMRD profiles of diamagnetic proteins: a model-free analysis, Magn. Reson. Chem. 38 (2000) 543–550.
[30] E. Liepinsh, G. Otting, Proton exchange rates from amino acid side chains—implication for image contrast, Magn. Reson. Med. 35 (1996) 30–42.
[31] K. Wüthrich, NMR of Proteins and Nucleic Acids, Wiley, New York, 1986.
[32] M. Ferraroni, W. Rypniewski, K.S. Wilson, M.S. Viezzoli, L. Banci, I. Bertini, S. Mangani, The crystal structure of the monomeric human SOD mutant F50/G51E/E133Q at atomic resolution. The enzyme mechanism revisited, J. Mol. Biol. 288 (1999) 413–426.
[33] L. Banci, I. Bertini, C. Luchinat, R. Monnanni, A. Scozzafava, Water $^{1}$H nuclear magnetic relaxation dispersion (NMRD) of $Cu_2Zn_2$SOD with some anions and $^{1}$H NMR spectra of $Cu_2Co_2$SOD in the presence of $CN^-$, Inorg. Chem. 27 (1988) 107–109.

[34] G. Lipari, A. Szabo, Model-free approach to the interpretation of nuclear magnetic resonance relaxation in macromolecules. 2. Analysis of experimental results, J. Am. Chem. Soc. 104 (1982) 4559–4570.

[35] A. Van-Quynh, S. Willson, R.G. Bryant, Protein reorientation and bound water molecules measured by (1)H magnetic spin–lattice relaxation, Biophys. J. 84 (2003) 558–563.

[36] B. Halle, M. Davidovic, Biomolecular hydration: from water dynamics to hydrodynamics, Proc. Natl. Acad. Sci. USA 100 (2003) 12135–12140.

[37] S. Kumar, K. Modig, B. Halle, Trifluoroethanol-induced $\beta \rightarrow \alpha$ transition in β-lactoglobulin: hydration and cosolvent binding studies by 2H, 17O, and 19F magnetic relaxation dispersion, Biochemistry 42 (2003) 13708–13716.

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 280, NO. 43, pp. 35815–35821, October 28, 2005


# Fully Metallated S134N Cu,Zn-Superoxide Dismutase Displays Abnormal Mobility and Intermolecular Contacts in Solution*[S]

Received for publication, June 17, 2005 Published, JBC Papers in Press, August 16, 2005, DOI 10.1074/jbc.M506637200

**Lucia Banci[‡§¶], Ivano Bertini[‡§¶1], Nicola D'Amelio[‡2], Elena Gaggelli[‡‖], Elisa Libralesso[‡], Irena Matecko[‡3], Paola Turano[‡§], and Joan Selverstone Valentine[**]**

*From the [‡]Magnetic Resonance Center, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy, [§]Department of Chemistry, University of Florence, Via Della Lastruccia 5, 50019 Sesto Fiorentino, Italy, [¶]FiorGen Foundation, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Florence, Italy, [‖]Department of Chemistry, University of Siena, Via Aldo Moro, 53100 Siena, Italy, and [**]Department of Chemistry and Biochemistry, University of California Los Angeles, Los Angeles, California 90095-1569*

**S134N copper-zinc superoxide dismutase (SOD1) is one of the many mutant SOD1 proteins known to cause familial amyotrophic lateral sclerosis. Earlier studies demonstrated that partially metal-deficient S134N SOD1 crystallized in filament-like arrays with abnormal contacts between the individual protein molecules. Because protein aggregation is implicated in SOD1-linked familial amyotrophic lateral sclerosis, abnormal intermolecular interactions between mutant SOD1 proteins could be relevant to the mechanism of pathogenesis in the disease. We have therefore applied NMR methods to ascertain whether abnormal contacts also form between S134N SOD1 molecules in solution and whether Cys-6 or Cys-111 plays any role in the aggregation. Our studies demonstrate that the behavior of fully metallated S134N SOD1 is dramatically different from that of fully metallated wild type SOD1 with a region of subnanosecond mobility located close to the site of the mutation. Such a high degree of mobility is usually seen only in the apo form of wild type SOD1, because binding of zinc to the zinc site normally immobilizes that region. In addition, concentration-dependent chemical shift differences were observed for S134N SOD1 that were not observed for wild type SOD1, indicative of abnormal intermolecular contacts in solution. We have here also established that the two free cysteines (6 and 111) do not play a role in this behavior.**

Over 100 different mutations in the superoxide dismutase 1(SOD1)[4] gene have been linked to the inherited form of amyotrophic lateral sclerosis (ALS), a fatal neurodegenerative disease characterized by progressive death of motor neurons and consequent paralysis. The individual mutations have been shown to exert their pathological effects by a gain of function mechanism, implying that the copper-zinc superoxide dismutase variant (Cu,Zn-SOD) expressed from the mutated gene has in some way become toxic. Recent evidence suggests that SOD1-linked ALS, like many other neurodegenerative diseases, is a protein misfolding disorder characterized by abnormal deposits of aggregated proteins in neural tissues (1–3). Implicated also in the toxicity of ALS-mutant SOD1 proteins are accelerated oxidative damage and inhibition of proteasome, chaperone, or mitochondrial function (2, 4, 5).

Relatively large, insoluble proteinaceous inclusions are observed in neural tissue in ALS and other neurodegenerative diseases, but it has been proposed that, rather than causing the disease, these deposits of aggregated protein may be formed as a part of a defensive mechanism to sequester misfolded and oligomerized protein when excessive amounts accumulate. More recently it has been suggested that the toxic species are smaller, high molecular weight oligomerized proteins formed from abnormal contacts of misfolded proteins resulting in protofibrils, pores, or other toxic species (6–10). High molecular weight forms of oligomerized mutant SOD1 have been observed in several studies using ALS-SOD1 transgenic mice (11–13).

Many of the ALS-linked Cu,Zn-SOD variant proteins have been expressed, purified, and characterized in an effort to identify properties that they hold in common that differ from the same properties of wild type Cu,Zn-SOD (5). These studies have revealed a wide variability in their biophysical and biochemical properties. Those of the Cu,Zn-SOD variant proteins with modifications far from the metal binding region have normal SOD activities and spectroscopic characteristics. Their three-dimensional crystal structures are overall quite similar to those of wild type Cu,Zn-SOD and thus give no clues as to possible mechanisms of misfolding and aggregation (14, 15). For that reason, it has been hypothesized that those mutant proteins may adopt a toxic form either prior to folding and metallation or subsequent to covalent modification or partial degradation (5, 16–19).

By contrast, the Cu,Zn-SOD variant proteins with modifications near the metal-binding sites, *i.e.* in either the zinc or the electrostatic loops, show altered SOD activities and spectroscopic properties. Some of them, specifically metal-deficient S134N and H46R SOD1, have been shown to crystallize in filament-like arrays with abnormal contacts between the individual protein dimers (20, 21). The presence of abnormal protein-protein contacts in the crystal structures of the ALS-linked Cu,Zn-SOD variant proteins suggests the possibility that similar intermolecular contacts may be made in solution in the earliest steps of the protein oligomerization that ultimately leads to aggregation.

In the current study, several spectroscopic techniques have been

* This work was supported in part by Ministero Italiano dell'Universita' e della Ricerca PRIN-2003, by European Union Contract LSHG-CT-2004-512052 "UPMAN-Understanding Protein Misfolding and Aggregation by NMR", and grants from the National Institute of General Medical Sciences (GM28222) and from the Amyotrophic Lateral Sclerosis Association (to J. S. V.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Fig. S1 and supplemental Tables S1 and S2.

[1] To whom correspondence should be addressed. Tel.: 39-055-4574272; Fax: 39-055-4574271; E-mail: ivanobertini@cerm.unifi.it.

[2] Recipient of a fellowship provided by Consorzio Interuniversitario per le Risonanze Magnetiche di Metalloproteine Paramagnetiche.

[3] Recipient of a Marie Curie host fellowship for early stage research training (contract MEST-CT-2004-504391 "NMR in Inorganic Structural Biology").

[4] The abbreviations used are: SOD, superoxide dismutase; *as*SOD, SOD in which Cys-6 and Cys-111 have been mutated to Ala and Ser, respectively; ALS, amyotrophic lateral sclerosis; HSQC, heteronuclear single quantum coherence; NMRD, nuclear magnetic relaxation dispersion; NOESY, nuclear Overhauser effect spectroscopy; $R_1$, longitudinal relaxation rate; $R_2$, transversal relaxation rate; TOCSY, total correlation spectroscopy; EPR, electron paramagnetic resonance.

applied to probe the structure and dynamics of the fully metallated S134N SOD1 protein in solution. Studies were carried out using both S134N *as*SOD1 and S134N *wt*SOD1. ("S134N *as*SOD1" or "*as*S134N" denotes a protein in which two additional mutations have been introduced, *i.e.* the 2 cysteine residues not involved in the intramolecular disulfide bond, Cys-6 and Cys-111, have been mutated to Ala and Ser, respectively. Those cysteines have not been removed in *wt*S134N or S134N *wt*SOD1). The former was used to ensure that any intermolecular contacts observed in solution were not due to formation of intermolecular disulfide bonds by air oxidation of the sample.

## MATERIALS AND METHODS

*Molecular Biology and Protein Characterization*—The mutation S134N was introduced into *as*SOD1 by PCR mutagenesis, using the QuikChange® mutagenesis kit (Stratagene) and pPCuZnSODII$^q$ as template. A plasmid encoding *wt*SOD1 and containing the same mutation was obtained by back mutation on pPCuZnSODII$^q$ to get the *wt*SOD1 gene again and then inserting the mutation S134N as previously described. All the resulting constructs were sequenced to confirm the mutations. *Escherichia coli* TOPP1 strains were transformed with the obtained plasmids and with pPCuZnSODII$^q$.

To get large amounts of enzymes, the protein was expressed in 2xYT complex medium (22); to obtain uniformly $^{15}$N- and, when needed, $^{13}$C-labeled proteins for the sake of NMR studies, $^{15}$N- and $^{13}$C-enriched M9 chemically defined medium was used as described (23). An exception was the case of *wt*S134N, which was expressed without addition of copper sulfate to the culture because some expression tests revealed a sharp decrease in the protein yield upon addition of such salt. In the case of S134N *as*SOD1, the protein was isolated and purified as reported (24). Stock solutions of the proteins were stored at 4 °C, for short-term use, or at −20 °C. Further *in vitro* addition of Cu and Zn to the enzymes was not performed. Extraction and purification of S134N *wt*SOD1 followed the procedure of Hayward *et al.* (25) for insect cell cultures, adapted for use in *E. coli*.

Because the $^{15}$N uniformly labeled *wt*S134N mutant showed lower solubility in ammonium sulfate than the unlabeled one, this specific enzyme was extracted from the cells using the same lysis buffer as the unlabeled one, but without NaCl, and was then purified by anionic exchange chromatography using the same protocol as for the *as*SOD1 mutants. In this case, 2 mM dithiothreitol was added to the elution buffers to keep the cysteines reduced during purification. Further *in vitro* addition of Cu and Zn was performed on the $^{15}$N-labeled enzyme following a well known procedure (26) because it appeared to be largely in the apo form when checked by bidimensional NMR.

Metal content analysis of *wt*S134N and *as*S134N, obtained both from complex and chemically defined media, were performed using a Spectro Ciros charge-coupled device inductively coupled plasma optical emission spectrometer (Spectro Analytical Instruments) in combination with a Lichte nebulizer and a peristaltic pump for sample introduction. Concentrated stocks of proteins were diluted to a concentration of 3–5 $\mu$M in milliQ water. The inductively coupled plasma was programmed to detect three wavelengths for the Zn (202, 206, 213 nm) and three for the Cu (219, 219b, 324 nm), each measurement being repeated three times. The standardization curve was made using standard solutions in the range 0–10 $\mu$M Zn and Cu in milliQ water; a blank was also registered, using the protein buffer diluted in the same way as the samples themselves. Protein concentrations after dilution were determined using the Coomassie assay (Pierce Perbio) with bovine serum albumin as standard.

*Spectroscopic Characterization*—Spectra of S134N *as*SOD1 and S134N *wt*SOD1 were recorded both in the visible and in the UV range of wavelengths, using samples of concentration $10^{-4}$ M in the first case and diluting them 10-fold in the latter. The samples, in 20 mM phosphate buffer, pH 5, were centrifuged before performing the measurements.

Circular dichroism was used to characterize the secondary structure content of the mutants as well as the copper coordination environment. For the near-UV, spectra were collected on samples with concentrations ranging from 10 to 25 $\mu$M, at 298 K in the wavelength range 250–190 nm, on a Jasco J-810 spectrophotometer equipped with a Peltier unit. In the visible region spectra have been acquired on a 500-$\mu$M sample at 298 K in the wavelength 250–1000 nm, in 20 mM phosphate buffer at pH 5.

EPR spectra were recorded at 180 K on an Elexsys E500 spectrometer (Bruker), equipped with X-band microwave bridge (microwave frequency 9.45 GHz) and an ER 4131 VT unit for temperature control. 1-mM protein samples in 20 mM phosphate buffer at pH 5.0 were used.

NMRD measurements were performed on a set of 3 to 4 mM protein samples in 20 mM phosphate buffer, pH 5.0. Two distinct sets of measurements were performed, on the freshly prepared samples and on the same samples incubated at room temperature for a week, to promote oligomerization of the mutants. All measurements were made at a constant temperature of 298 K. To avoid the paramagnetic effect, reduction of Cu(II) to Cu(I) was achieved by addition of 1.2–1.5 equivalents of *p*-isoascorbic acid dissolved in degassed 20 mM phosphate buffer, pH 5, under reducing conditions. Longitudinal water proton relaxation rates were measured with a Stelar prototype of fast field cycling relaxometer (proton Larmor frequency 0.01–40 MHz). $^1$H NMRD profiles were obtained by plotting water proton relaxation rates, after subtracting the relaxation rates of the buffer and normalizing the rates to 1 mM protein concentration, as a function of applied magnetic field.

*High Field NMR*—All the protein samples were in 20 mM phosphate buffer, pH 5.0, and the Cu(II) ion was reduced according to the procedure described above. The protein concentration (calculated according to the monomer) in the $^{15}$N samples was in the range 1.0–1.5 mM, whereas in $^{13}$C, $^{15}$N samples were in the range 2.0–2.5 mM.

For mutant *as*S134N, several experiments were performed with the aim of achieving extensive spectra assignment: HNHA, HCCH-TOCSY, $^{13}$C-HSQC-NOESY, and $^{15}$N-HSQC-NOESY on a Bruker Avance 900 MHz and HNCOCA and HNCA at 500 MHz.

Mobility experiments were performed on both mutants using a Bruker 600-MHz spectrometer. For concentration- and time-dependent experiments, an 800 MHz spectrometer equipped with cryoprobe was used. For the purpose of hydrogen/deuterium exchange, a sample of *as*S134N was reduced as previously described; the *p*-isoascorbic acid was then discarded by using a Pd-10 column under a reducing atmosphere, and the enzyme was lyophilized at −50 °C and ~0.5 bar. The sample was resuspended in $D_2O$, and *p*-isoascorbic acid was again added. $^1$H-$^{15}$N HSQC spectra were acquired every hour for a total of 72 h using a 900-MHz spectrometer equipped with a cryoprobe.

*Structure Modeling*—NOE cross-peaks in three-dimensional $^{15}$N and $^{13}$C NOESY-HSQC spectra were integrated for the region flanking the mutation (residues 120–140) and for residues spatially close in the solution structure of the non-mutated protein (residues 70–76, 43–46, 61–63, 85, 86). Volumes were converted into 691 upper distance limits for inter-proton distances with the program CALIBA (27). For the remaining part of the protein upper distances were evaluated by the solution structure and used as additional constraints. Some NOE cross-peaks regarding this part of the protein were integrated in order to calibrate the calculated distances in the region surrounding the muta-

tion. Structure calculations were performed using the program DYANA (28). Two hundred random conformers of one subunit were annealed using the above constraints in 10,000 steps. The final family was made up of 20 structures with the lowest values for the target function.

## RESULTS

The S134N *as*SOD1 protein was isolated with two copper and two zinc ions (supplemental Table S1) as it has been seen previously for the *as*SOD1 in our expression system (24). It is in the dimeric form, as indicated by gel filtration measurements (data not shown) and as further confirmed by $^{15}$N NMR relaxation data and field-dependent water relaxation rates on freshly prepared samples (see below).

The mutant was characterized at 298 K and pH 5 using freshly prepared protein samples. The copper(II) form was used for the electronic absorption, circular dichroism, and electron paramagnetic resonance spectra to take advantage of the spectroscopic features of this metal ion for the characterization of the metal-binding site, whereas the copper(I) form of the protein was used for NMR and NMRD characterization to avoid paramagnetic effects that would complicate the spectral analysis. These last two sets of data were used to monitor the structural and dynamic features of the protein. The experimental data on the two oxidation states of the protein were then analyzed to obtain a complete picture of the mutant.

*The Copper-binding Site*—The copper coordination environment is only slightly altered with respect to that of the *as*SOD1 protein. Indeed, the CD spectrum in the visible region is essentially superimposable on that of native SOD1, with a negative band at 13,200 $cm^{-1}$ and a positive band at 16,600 $cm^{-1}$ (29). The EPR spectrum (supplemental Fig. S1) is characterized by $g_{//} = 2.26$, $g_{\perp} = 2.06$, and $A_{//} = 154 \times 10^{-4}$ $cm^{-1}$, *i.e.* parameters very close to those of native SOD1 (29). However, at variance with the native protein, which possesses a rhombic spectrum, the spectrum of *as*S134N is essentially axial. Axial EPR spectra have been observed for some other SOD1 mutant proteins, not related to ALS (29), indicating a more tetragonal geometry for copper(II) bound in these mutants.

*Structural and Dynamic Properties*—The secondary β-structure is essentially conserved, as suggested by CD measurements in the far UV region. The $^{15}$N-$^{1}$H HSQC spectra of the *as*S134N and *as*SOD1 proteins are substantially superimposable (Fig. 1*A*), with the exception of a few resonances. Meaningful shift differences between *as*SOD1 and the *as*S134N variant are clustered in the surroundings of the mutation as summarized in Fig. 1*B*, from which is it clear that residues 127–140 sense a different environment in the mutant. The chemical shift for residues 130–140 approaches those typical for isolated amino acids as for random coil polypeptides, and the $^{3}J_{HNHa}$ values are consistently in the range 4.6–7.2 Hz. The $^{1}$H, $^{15}$N, and $^{13}$C resonance assignments for *as*S134N are reported in supplemental Table S2.

The differences in chemical shifts for the backbone amides in this protein area are accompanied by differences in the dynamic properties of residues in the same area, in particular for residues 129–137 (Fig. 2). $^{15}$N relaxation rates of proteins and {$^{1}$H}-$^{15}$N heteronuclear NOEs are sensitive to the overall protein tumbling and to local motions either faster than the protein tumbling or much slower, *i.e.* in the range of tens of microseconds to milliseconds that is typical of conformational equilibria. In the present case, the average values for $R_1$, $R_2$, and NOE, if the 129–137 fragment is not considered, were 0.80 ± 0.10 $s^{-1}$, 25.8 ± 4.5 $s^{-1}$, and 0.78 ± 0.14, respectively, which are very close to those of the *as*SOD1 protein (30). $R_1$ rates for the 129–137 fragment are sensibly larger, with average values of 1.10 ± 0.10 $s^{-1}$. $R_2$ rates and heteronuclear NOEs decrease dramatically to values in the range 11–23 $s^{-1}$ for $R_2$ and to values smaller than 0.55 for the NOE. The $^{15}$N relaxation parameters were instead uniform along the entire sequence in the protein with the native amino acid in position 134. The present data therefore point at extensive mobility on the nanosecond time scale in the protein segment 129–137, where an Asn replaces the native Ser.

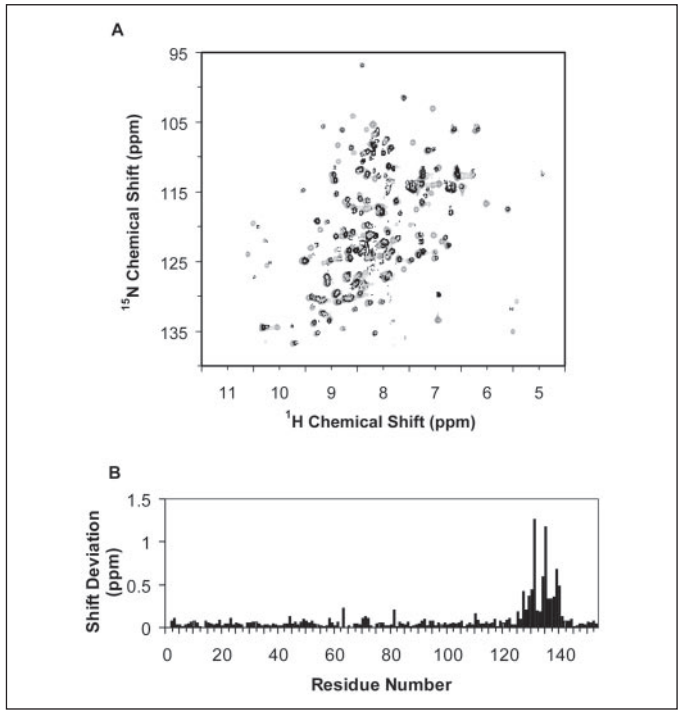



FIGURE 1. **$^{15}$N-HSQC superimposition of *as*SOD1 (*light gray*) and *as*S134N mutant (*black*) (*A*) and chemical shift deviations along the primary structure (*B*) computed as $((\Delta\delta_N^2/5+\Delta\delta_H^2)/2)^{1/2}$ where $\Delta\delta$ is the observed chemical shift change.**

The average rotational correlation time for a protein can in principle be estimated from the $R_2/R_1$ ratio. However, the ratio $R_2/R_1$ is also affected by the presence of conformational exchange and by motional anisotropy. On the other hand, the product $R_1R_2$ is independent of the overall correlation time and motional anisotropy and therefore can be used to highlight residues undergoing chemical exchange (31). A plot of $R_1R_2$ values for *as*S134N is shown in Fig. 2*D*, where an increase in the $R_1R_2$ values due to the presence of conformational exchange on the microsecond-millisecond time range is observed for a number of residues, consistent with the *as*SOD1 data (30). Additionally, reduced $R_1R_2$ values for residues 129–137 are observed, which further confirms the presence of fast motions for this protein fragment. From the ratio $R_2/R_1$, with the exclusion of residues involved in local fast motions or conformational exchange, an average correlation time for the protein tumbling of 18.1 ± 1.9 ns was estimated. Residues 129–137 experience an average correlation time of 12.3 ± 2.0 ns, which represents an estimate of the effective correlation time for the dipolar interactions in the 129–137 fragment.

The $^{1}$H NMRD profiles report the water relaxation rates as a function of the magnetic field of millimolar solutions of protein and monitor the overall protein tumbling time. The profiles for the freshly prepared *as*S134N and *as*SOD1 are shown in Fig. 3, *A* and *B*. Their best fit was performed using the Model-free approach (32) (*solid line*), which provides an estimate of correlation time for the protein tumbling. For the *as*S134N variant this value is 20.0 ns and compares well with that of the *as*SOD1, for which a correlation time of 20.5 ns is calculated, thus clearly confirming that the mutant is also in the dimeric state. More-

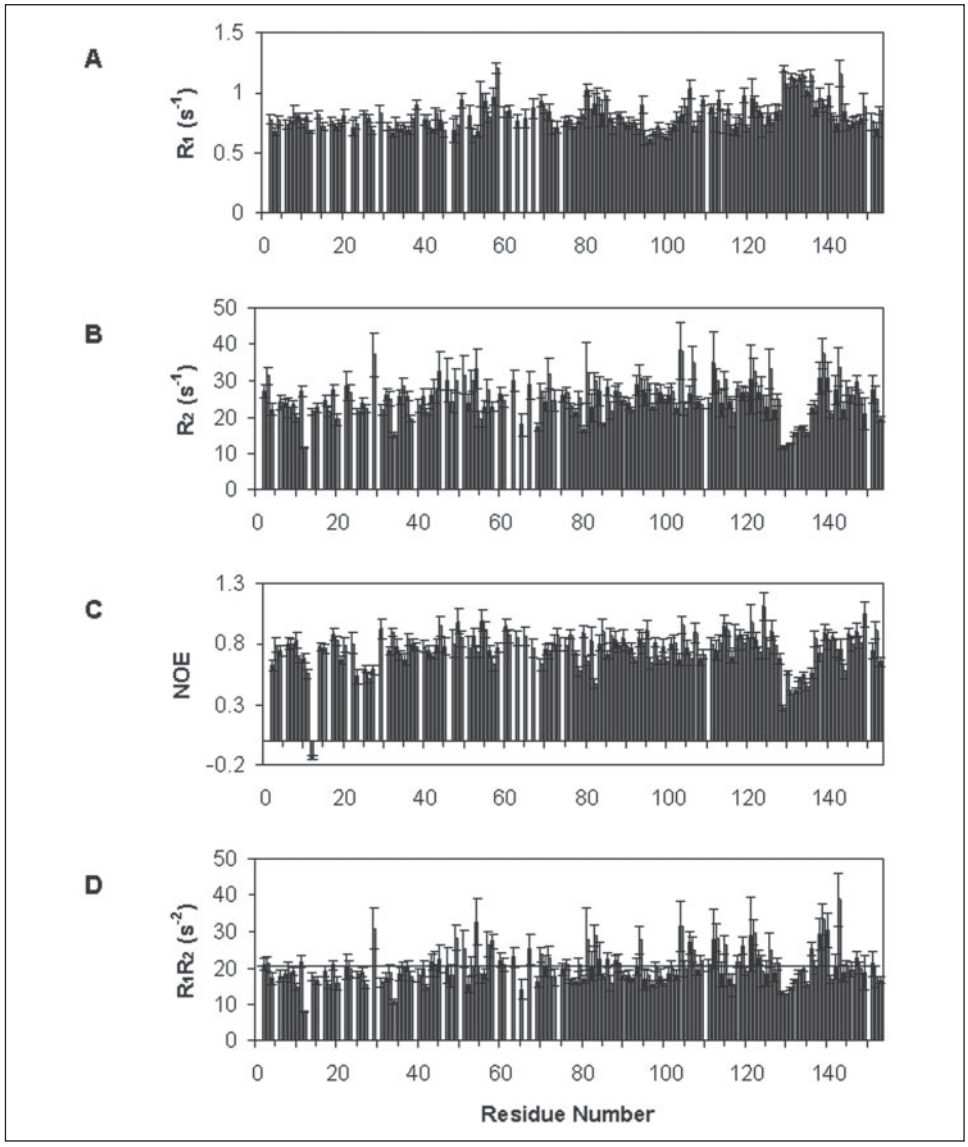



FIGURE 2. $^{15}N$ relaxation data (*A–C*) measured for S134N *as*SOD1. *D*, the values of the $R_1R_2$ product are reported for each residue; the *horizontal line* indicates the average value of $R_1R_2$ calculated excluding the mobile part of the protein around the mutation.

over, these correlation times are completely consistent with the values found from $^{15}N$ relaxation rates.

The time dependence of the hydrogen/deuterium exchange of backbone NH resonances provides information on their solvent accessibility. The per-residue exchange behavior of the *as*S134N mutant, as followed through $^{1}H$-$^{15}N$ HSQC spectra recorded every hour for 72 h after dissolving in $D_2O$ a lyophilized sample of the protein, is similar to that of the *as*SOD1 studied under the same conditions (33) with the exception of the area surrounding the mutation. Indeed, in the *as*S134N variant, the backbone NHs of residues 127–143 are fully exchanged within the first hour, with residue 131 being completely exchanged within 2 h. In the case of the *as*SOD1, the residues in these areas exchange more slowly; complete exchange was reported to occur after ~6 h for Gly-129, Asn-131, and Thr-133 and after more than 12 h for Ser-134 and Lys-136 (33). Faster exchange in the variant in the area of the mutation is consistent with the presence of fast motions and lack of secondary structure, which make the backbone NHs more solvent accessible.

The number of NOEs/residue as measured from the $^{1}H$-$^{15}N$ HSQC-NOESY spectrum of *as*S134N mutant for the protein area around the mutation site is comparable with that observed for the native protein (33), although most of the NOEs are sensibly reduced in intensity. This effect is consistent with the presence of fast motions in this protein region as discussed above. Furthermore, the backbone sequential NOEs, as well as the $^{3}J_{HNH\alpha}$ and CSI values, do not indicate the presence of the short helix present in the native structure. Differences in long-range NOE patterns were not detected, except for a generic reduction in intensity that in some cases makes undetectable the NOEs that were already weak in *as*SOD1. In particular, the long-range NOEs between the 129–137 region characterized by fast internal motions and the protein fragment facing it in the structure of *as*SOD1 (residues 125–127) are weaker. However, these contacts are not lost, with the exception of those between the NH of Asp-125 and the side chain of residue 134 (the mutated residue). The introduction of a new residue slightly changed the NOE patterns for residue 134.

The structural model, calculated as described under "Materials and Methods," shows the region flanking the mutation quite disordered with backbone root mean square deviation values for the fragment 127–140 of 1.58 Å.

*Tendency to Oligomerize*—A macroscopic difference between the native protein and the *as*S134N mutant is the tendency of the latter to give rise to high molecular weight species in concentrated protein solutions. Such high molecular weight species are clearly observable as

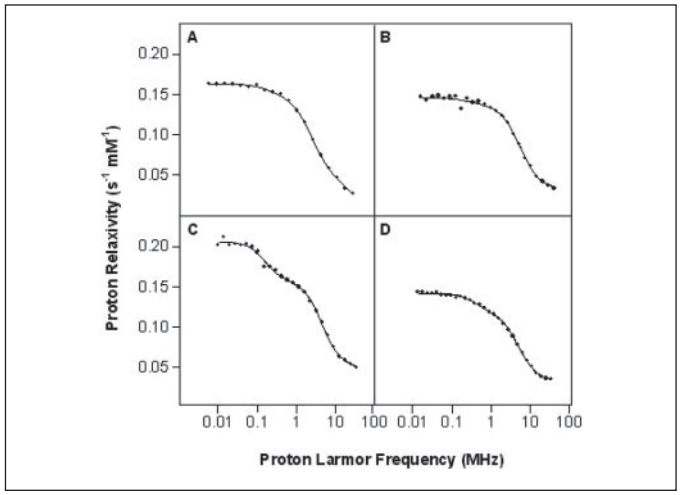



FIGURE 3. NMRD profiles of S134N *as*SOD1 freshly prepared (*A*) and after 3 weeks (*C*). Experiments were repeated for *as*SOD1 (*B* and *D*, respectively). The concentration of the protein samples was 3 mM at pH 5. Data were fitted using the Model-free approach (*solid line*). Fitting parameters were: *A*, $\alpha = 0.028\ s^{-1}$, $\beta = 6.7 \times 10^6\ s^{-2}$, $\tau = 20.0$ ns; *B*, $\alpha = 0.035\ s^{-1}$, $\beta = 5.4 \times 10^6\ s^{-2}$, $\tau = 20.5$ ns; *C*, $\alpha = 0.047\ s^{-1}$, $\beta = 5.8 \times 10^6\ s^{-2}$, $\tau_1 = 548$ ns, $\tau_2 = 19.0$ ns, $c_1 = 1.6$, $c_2 = 98.4$ (with $c_i$ expressing the relative population of the two species with different molecular weight); *D*, $\alpha = 0.032\ s^{-1}$, $\beta = 5.0 \times 10^6\ s^{-2}$, $\tau = 21.6$ ns.

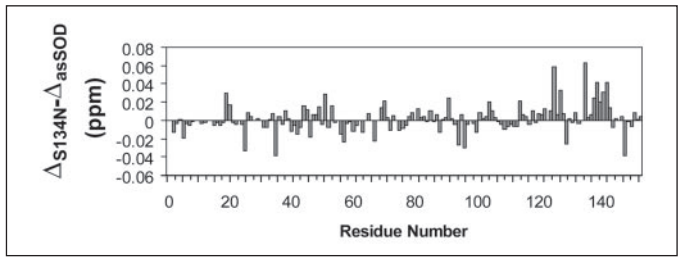



FIGURE 4. **Differences in chemical shift deviations between S134N *as*SOD1 and *as*SOD1 along the protein sequence, upon diluting the protein samples from 1.5 to 0.125 mM.** The chemical shift deviations are computed as $((\Delta\delta_N^2/5+\Delta\delta_H^2)/2)^{1/2}$ where $\Delta\delta$ is the observed chemical shift change.

insoluble entities in the NMR tube after a few days. The presence of large molecular weight protein forms was monitored through NMRD profiles. Although the freshly prepared protein has a profile very similar to that of the *as* form (Fig. 3, *A* and *B*), with time a new dispersion at lower field appears (Fig. 3*C*), indicating the presence of species with slow tumbling rates, *i.e.* of larger molecular weight. The relaxation dispersion profiles provide two independent correlation times, one as an average value for the large molecular weight species and one for the dimeric protein. For the latter a correlation time corresponding to that of the freshly prepared samples was obtained (19.0 ns). The extent of the oligomerization depends on protein concentration and time, but the dispersion at lower fields can always be interpreted as due to the presence of aggregates with molecular weight ~20–30 times larger than that of the dimeric protein. This effect does not have a counterpart in the *as* protein (Fig. 3*D*).

Another noticeable feature of the *as*S134N mutant is that the chemical shifts of residues around the mutation change depending on the sample concentration. By analyzing the chemical shift variations upon protein dilution for the *as*SOD1 and the S134N *as*SOD1 variants (stepwise dilution from 1.5 to 0.06 mM, followed at 800 MHz with a cryoprobe), some changes in chemical shift are observed for a number of residues. Nevertheless, a selective effect was detected in the S134N mutant for residues in the mutation area that does not correspond with the non-mutated protein under the same experimental conditions. The changes in chemical shift are larger at concentrations lower than 0.25 mM. The differences in chemical shift variation for the S134N *as*SOD1 and *as*SOD1 observed upon diluting the sample from 1.5 to 0.125 mM are reported in Fig. 4. The 0.125 mM was considered as final concentration because in the 0.06-mM sample some signal intensity was below the detection limit.

*Removing the as Mutation*—The results reported above describe the effect of the mutation in position 134 through the characterization of the mutated and non-mutated thermostable *as*SOD1 forms. It is clearly evident that the *as*S134N variant possesses a higher tendency toward aggregation although sharing with the non-mutated *as*SOD1 most of the structure and dynamics features; the main differences are limited to a very small protein fragment around the mutation site that becomes less structured and more flexible, with mobility faster than the overall tumbling.

The *wt*S134N is very similar to the *as*S134N, with essentially superimposable EPR spectra for the copper(II) and an electronic absorption spectrum with a maximum for the copper band at 655 nm. The differences in chemical shift observed through $^{15}N$-$^{1}H$ HSQC spectra in the protein area containing the mutation with respect to the *wt*SOD1 are reproduced also in the *wt*S134N derivative (data not shown). More importantly, the tendency toward aggregation is again displayed only by the *wt*S134N variant, thus proving that the effect of the S134N mutation is the same on the two *as* and *wt* protein scaffolds and does not depend on the presence of Cys-6 or Cys-111.

## DISCUSSION

S134N *wt*SOD1 is partially metal deficient when isolated from its yeast or baculovirus expression systems and has been observed to crystallize in linear, amyloid-like filaments in which disorder in the loop regions has resulted in a deprotection of exposed $\beta$-strands and formation of abnormal intermolecular protein-protein contacts (20). These contacts are facilitated by the loss of zinc in the zinc site, which induces conformational heterogeneity in the zinc and electrostatic loops. Indeed, these loops are found to be significantly disordered in the x-ray structure, and residues 125–131 play an important role in filament formation as they participate to H-bonds and apolar interactions with the $\beta$-barrel of a neighboring protein molecule. This structure suggests that similar intermolecular contacts might be forming in solution as the first step in the protein oligomerization that leads to the toxicity associated with the disease.

We carried out the current study to determine the effect of the S134N mutation on the structure and dynamics of this region of the SOD1 protein and to look for evidence of abnormal intermolecular protein-protein contacts. We performed most of our studies on the S134N *as*SOD1 mutant in order to eliminate the possibility of aggregation due to formation of intermolecular disulfide bonds via Cys-6 or Cys-111. Comparative studies of S134N *as*SOD1 and S134N *wt*SOD1 confirmed that the *as* mutations had little, if any, effect on the structure and mobility of the proteins in solution. These results are in accord with numerous previous studies of *as*SOD1 mutants. The C6A and C111S mutations increase the thermostability of the sample, but they do not have discernible effects on the activity, structure, or mobility of SOD1. In particular, extensive data are available for the dimeric, fully metallated *as*SOD1 protein (33), for the monomeric, fully metallated *as*F50EG51EE133Q SOD1 derivative (30), for the apo (34) and copper-free states (35) of this monomeric form, and for the fully metallated familial amyotrophic lateral sclerosis-related mutant *as*G93A SOD1 (36).

Fully metallated *as*SOD1, $Cu(I)_2Zn_2$SOD, is a relatively rigid protein with rather uniform relaxation parameters (30). About 15 residues, 3 of which (residues 122, 126, and 129) are in the region of interest for the

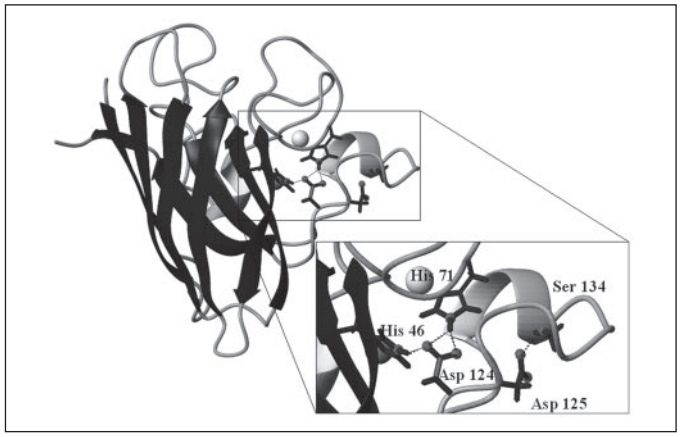



FIGURE 5. **H-bond network involving residues His-46, His-71, Asp-124, Asp-125, and Ser-134 in the wild type protein.** The figure has been generated from the x-ray crystal structure 1SOS (37). Zn and Cu ions are represented as spheres (*light* and *dark gray*, respectively).

present work, experience conformational exchange. The engineered monomeric protein in the fully metallated form behaves very similarly, with the exception of the area of loop IV where point mutations have been introduced to obtain the monomerization and which displays subnanosecond mobility (30).

The current study was carried out on fully metallated S134N proteins, and therefore the observed differences from the wild type protein can be attributed entirely to the mutation in position 134 and not to vacancies in metal-binding sites. We found here the behavior of fully metallated S134N *as*$Cu(I)_2Zn_2$SOD to be dramatically different from that of fully metallated *as*$Cu(I)_2Zn_2$SOD, either in the dimeric or the monomeric form. Indeed, we see high mobility in loop VII, comparable with that observed for apo-*as*SOD1, *despite the presence of zinc in the zinc site* in the *as*S134N variant. The absence of copper in monomeric zinc-only SOD derivative has previously been shown to have only limited effects on the structure and mobility of the *as*SOD1 protein, with slightly reduced order parameter values for the residues involved in copper binding (35). By contrast, in the monomeric apoprotein, where copper and zinc are both absent, loop IV (residues 45–85) and loop VII (residues 121–142) are largely unstructured and characterized by subnanosecond mobility, with $\{^1H\}$-$^{15}N$ heteronuclear NOE and $R_2$ values sensibly smaller than the average and $R_1$ values larger than the average (34). Lack of zinc has, therefore, a destabilizing effect on loops IV and VII. We conclude that zinc binding to the zinc site of SOD1 normally causes a sequence of structural changes that results in a pinning of loop VII to the rest of the protein structure. The non-native structural plasticity of the fully metallated, familial amyotrophic lateral sclerosis-linked S134N mutant is a relevant difference with respect to the non-mutated *as* and *wt* proteins and might represent the key to understanding the peculiar behavior of this variant observed *in vitro*.

In the crystal structure of *as*SOD1, the native Ser-134 OH forms an H-bond with the carboxylate of Asp-125 (the distance between the donor oxygen and the acceptor oxygen being of the order of 2.7 Å), as shown in Fig. 5 (37). This H-bond belongs to an extensive H-bond network that also involves Asp-124 carboxylate H-bonded to the NH of two histidines, His-46 that is a copper ligand, and His-71 that is a zinc ligand. Based on our results, the substitution of the native Ser-134 with an Asn destabilizes this H-bond network in such a way that the presence of zinc is no longer sufficient to pin the electrostatic loop down, leaving it rather flexible. Consistent with this view, the major difference in the NOE patterns observed for the electrostatic loop (as discussed under "Results") concerns the loss of the dipolar connectivity between the NH of Asp-125 and the side chain of the amino acid in position 134. Moreover, the distances between the oxygen atoms of the carboxylate of Asp-125 and Nδ of Asn-134 are on average both longer than 4.5 Å in our model structure. Another hydrogen bond that is important in maintaining the conformation of the electrostatic loop in the non-mutated protein involves Asp-125 and Asn-139; this appears to be maintained in S134N protein.

In the crystal structure of metal-deficient S134N *wt*SOD1 (20), Leu-126 forms a small hydrophobic contact with Leu-42 of a neighboring SOD1 molecule. Leu-42 is located on a β structure and become deprotected upon mobilization of the zinc loop in the metal-deficient protein. A similar contact does not appear to be formed by fully metallated S134N in solution. However, in our solution model of the fully metallated *as*S134N, Leu-126 becomes more solvent exposed as a result of the S134N mutation. (The solvent accessible area goes from an average of 3.2% in the family of solution structures of *as*SOD1 to an average value of 9.7% in the family of the model structure of *as*S134N.) Meaningful chemical shift differences upon dilution are observed in the area around the mutation, and none of our data point to a possible involvement of Leu-42 in interprotein contacts in solution. We conclude, therefore, that the intermolecular protein-protein interaction does not involve that latter region in solution but is confined to the electrostatic loop.

Interestingly, abnormal contacts between different SOD1 molecules, similar to those described for the S134N derivative, are also observed in the crystal structure of the H46R familial amyotrophic lateral sclerosis-related variant (21) where the mutation site again involves a residue belonging to the above described H-bonding network.

It is quite clear from our results that the S134N variant, even in the fully metallated form, tends to generate protein oligomers. This finding implies that this derivative is prone to formation of intermolecular interactions. Protein surface areas characterized by fast mobility have been reported in the literature to be important for molecular recognition in macromolecular complexes (38), and a flexible area could also be important for protein-protein recognition between molecules of the same protein.

The observed concentration-dependent changes in chemical shift of S134N further support this picture. The change in resonance frequency is indicative of an altered structural environment consistent with an involvement of this protein area in weak intermolecular interactions that cannot be revealed by other means due to their transient nature. At micromolar protein concentrations, the chemical shifts probably reflect the situation in the absence of interactions, whereas at higher concentrations the interactions with other molecules affect the chemical shift. This could be a local effect, or it could be a selection of some of the possible protein conformations in these highly flexible regions upon interaction with other molecules.

In general, the observed chemical shift differences due to intermolecular interaction detectable by NMR are usually not large even when the protein-protein complex becomes the dominant species in solution (39). The fact that in the present SOD1 variant the observed variations in average chemical shift are smaller than 0.1 ppm and no indications of complex formation could be detected from the relaxation rates analysis implies that interactions among different SOD1 molecules involve very small amounts of the protein molecules in solution. However, such protein-protein interactions could represent a nucleation event that, once started, leads to quite large oligomers, no longer detectable by standard solution NMR but revealed in our NMRD studies. This latter process may provide insights at the molecular level into the early stages of the formation of the soluble oligomers or aggregates that have been proposed to be responsible for the protein toxicity.

*Acknowledgment—We thank Dr. P. John Hart for helpful discussions.*

**REFERENCES**


1. Bruijn, L. I., Houseweart, M. K., Kato, S., Anderson, K. L., Anderson, S. D., Ohama, E., Reaume, A. G., Scott, R. W., and Cleveland, D. W. (1998) *Science* **281,** 1851–1854
2. Bruijn, L. I., Miller, T. M., and Cleveland, D. W. (2004) *Annu. Rev. Neurosci.* **27,** 723–749
3. Rowland, L. P., and Shneider, N. A. (2001) *N. Eng. J. Med.* **344,** 1688–1700
4. Cleveland, D. W., and Rothstein, J. D. (2001) *Nat. Rev. Neurosci.* **2,** 806–819
5. Valentine, J. S., Doucette, P. A., and Potter, S. Z. (2005) *Annu. Rev. Biochem.* **76,** 563–569
6. Kirkitadze, M. D., Bitan, G., and Teplow, D. B. (2002) *J. Neurosci. Res.* **69,** 567–577
7. Sherman, M. Y., and Goldberg, A. L. (2001) *Neuron* **29,** 15–32
8. Lashuel, H. A., Hartley, D., Petre, B. M., Walz, T., and Lansbury, P. T. (2002) *Nature* **418,** 291
9. Morrison, B. M., Morrison, J. H., and Gordon, J. W. (1998) *J. Exp. Zool.* **282,** 32–47
10. Johnston, J. A., and Madura, K. (2004) *Prog. Neurobiol.* **73,** 227–257
11. Wang, J., Xu, G. L., Gonzales, V., Coonfield, M., Fromholt, D., Copeland, N. G., Jenkins, N. A., and Borchelt, D. R. (2002) *Neurobiol. Dis.* **10,** 128–138
12. Johnston, J. A., Dalton, M. J., Gurney, M. E., and Kopito, R. R. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 12571–12576
13. Wang, J., Xu, G., and Borchelt, D. R. (2002) *Neurobiol. Dis.* **9,** 139–148
14. Hough, M. A., Grossmann, J. G., Antonyuk, S. V., Strange, R. W., Doucette, P. A., Rodriguez, J. A., Whitson, L. J., Hart, P. J., Hayward, L. J., Valentine, J. S., and Hasnain, S. S. (2004) *Proc. Natl. Acad. Sci. U. S. A.* **101,** 5976–5981
15. Hart, P. J., Liu, H., Pellegrini, M., Nersissian, A. M., Gralla, E. B., Valentine, J. S., and Eisenberg, D. (1998) *Protein Sci.* **7,** 545–555
16. Valentine, J. S., and Hart, P. J. (2003) *Proc. Natl. Acad. Sci. U. S. A.* 3617–3622
17. Lindberg, M. J., Tibell, L., and Oliveberg, M. (2002) *Proc. Natl. Acad. Sci. U. S. A.* **99,** 16607–16612
18. Arnesano, F., Banci, L., Bertini, I., Martinelli, M., Furukawa, Y., and O'Halloran, T. V. (2004) *J. Biol. Chem.* **279,** 47998–48003
19. Ray, S. S., Novak, R. J., Brown, R. H., Jr., and Lansbury, P. T., Jr. (2005) *Proc. Natl. Acad. Sci. U. S. A.* **102,** 3639–3644
20. Elam, J. S., Taylor, A. B., Strange, R., Antonyuk, S., Doucette, P. A., Rodriguez, J. A., Hasnain, S. S., Hayward, L. J., Valentine, J. S., Yeates, T. O., and Hart, P. J. (2003) *Nature Struct. Biol.* **10,** 461–467
21. Antonyuk, S., Elam, J. S., Hough, M. A., Strange, R. W., Doucette, P. A., Rodriguez, J. A., Hayward, L. J., Valentine, J. S., Hart, P. J., and Hasnain, S. S. (2005) *Protein Sci.* **14,** 1201–1213
22. Banci, L., Bertini, I., Chiu, C. Y., Mullenbach, G. T., and Viezzoli, M. S. (1995) *Eur. J. Biochem.* **234,** 855–860
23. Banci, L., Benedetto, M., Bertini, I., Del Conte, R., Piccioli, M., Richert, T., and Viezzoli, M. S. (1997) *Magn. Reson. Chem.* **35,** 845–853
24. Getzoff, E. D., Cabelli, D. E., Fisher, C. L., Parge, H. E., Viezzoli, M. S., Banci, L., and Hallewell, R. A. (1992) *Nature* **358,** 347–351
25. Hayward, L. J., Rodriguez, J. A., Kim, J. W., Tiwari, A., Goto, J. J., Cabelli, D. E., Valentine, J. S., and Brown, R. H., Jr. (2002) *J. Biol. Chem.* **277,** 15923–15931
26. Beem, K. M., Rich, W. E., and Rajagopalan, K. V. (1974) *J. Biol. Chem.* **249,** 7298–7305
27. Güntert, P., Braun, W., and Wüthrich, K. (1991) *J. Mol. Biol.* **217,** 517–530
28. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol.* **273,** 283–298
29. Banci, L., Bertini, I., Cabelli, D. E., Hallewell, R. A., Tung, J. W., and Viezzoli, M. S. (1991) *Eur. J. Biochem.* **196,** 123–128
30. Banci, L., Bertini, I., Cramaro, F., Del Conte, R., Rosato, A., and Viezzoli, M. S. (2000) *Biochemistry* **39,** 9108–9118
31. Kneller, J. M., Lu, M., and Bracken, C. (2002) *J. Amer. Chem. Soc.* **124,** 1852–1853
32. Bertini, I., Fragai, M., Luchinat, C., and Parigi, G. (2000) *Magn. Reson. Chem.* **38,** 543–550
33. Banci, L., Bertini, I., Cramaro, F., Del Conte, R., and Viezzoli, M. S. (2002) *Eur. J. Biochem.* **269,** 1905–1915
34. Banci, L., Bertini, I., Cramaro, F., Del Conte, R., and Viezzoli, M. S. (2003) *Biochemistry* **42,** 9543–9553
35. Banci, L., Bertini, I., Cantini, F., D'Onofrio, M., and Viezzoli, M. S. (2002) *Protein Sci.* **11,** 2479–2492
36. Shipp, E., Cantini, F., Bertini, I., Valentine, J. S., and Banci, L. (2003) *Biochemistry* **42,** 1890–1899
37. Parge, H. E., Hallewell, R. A., and Tainer, J. A. (1992) *Proc. Natl. Acad. Sci. U. S. A.* **89,** 6109–6114
38. Dyson, H. J., and Wright, P. E. (2005) *Nat. Rev. Mol. Cell. Biol.* **6,** 197–208
39. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem.* **276,** 41365–41376

# 5. Characterization of G37R Cu,Zn-superoxide dismutase, a WTL mutant with an aberrant metallation state

In preparation

# 6. List of publications

## $^{1}$H nuclear magnetic relaxation dispersion of Cu,Zn superoxide dismutase in the native and guanidinium-induced unfolded forms

Libralesso E., Nerinovski K, Parigi G, Turano P

Biochem Biophys Res Commun. 2005 Mar 11;328(2):633-9.

## Fully Metallated S134N Cu,Zn-Superoxide Dismutase Displays Abnormal Mobility and Intermolecular Contacts in Solution.

Banci L, Bertini I, D'Amelio N, Gaggelli E, Libralesso E, Matecko I, Turano P, Valentine JS.

J Biol Chem. 2005 Oct 28;280(43):35815-35821. Epub 2005 Aug 16.

# 7. General discussion and perspectives

The approach chosen during this PhD for the characterization of the fALS mutants was the detailed analysis, often at atomic level, of the structure and dynamics of these polypeptides, with the goal of finding out any discrepancy with respect to native SOD and of investigating the relationship between the enzymes aberrant properties and their toxic gain of function . Aiming for this result, two fALS mutants belonging respectively to "metal binding site" and "wild type-like" class of mutants were thoroughly investigated, mainly by mean of NMR techniques and circular dichroism; while for a third one, featuring a mutation in one of the Cu ligand histidines, the preliminary characterization is beginning.

As enlightened in the previous sections, the *in vitro* studies performed by our group evidenced that the mutants manifest highly varying behaviours in the aggregating tendency as well as in the metallation degree - although they were all expressed and purified following the same protocol - being the general structural architecture the only essentially conserved property.

The MBR mutant S134N, in fact, shows fast (ns-ps) mobility in the backbone region surrounding the mutation, in a similar fashion to the apo form of *as*SOD, although the mutant is isolated in its correct metallation state from the expression system. Associated to this, a net aggregating tendency differentiates it from its non mutated analogue. These macroscopic variation in its behaviour were expected, since the metal-binding regions mutants appear to be primed for aggregation and specifically S134N, as reported, crystallizes in filament-like arrays[13], but their linking with conspicuous dynamical properties is an important finding for the deciphering of said aggregating process itself.

On the contrary, WTL mutant G37R shows an aberrant metallation state, while its dynamic and aggregating properties are identical to *as*SOD ones. The first property is quite unusual for a mutant belonging to the WTL class, but it could have been expected, given the subunit asymmetry found in the crystal structure[15], that accounting for a certain distortion with respect to the native geometry could determine a diminished affinity for the Cu ion itself. We would have expected also to find an altered mobility in the region surrounding the mutation, since it takes places in an area of the molecule where the three-dimensional architecture is considered critical for maintaining the protein stability. The residue 37 is in fact located at the end of strand 3, in a portion of the three dimensional fold where the hydrogen bonding interactions facilitate the "plug" end of the β-barrel. The clustering of fALS mutations in this region of the β-barrel further suggests that the wild-type packing arrangement of apolar residues is crucial for correct enzymatic function and protein stability [102] and for G93A, a mutation taking place at the end of

the adjacent strand 6, aberrant mobility of all this structural element was found[11], although they took place in a slower timescale with respect to the one studied during our mobility experiments.
Also denaturation with TFE showed large differences, both among the mutants and between the apo and holo form of each variant, in relation to their relative stability and their unfolding intermediates. Stability of G37R and S134N, in the metallated form, well fits in the overall picture delineated by High Field NMR and FFCR. Fully metallated S134N, in facts, shows a decreased stability with respect to native SOD, and at variance with respect to the latter and to the mutants G37R and H48Q, a net α propensity that can be correlated with its aggregating tendency. TFE induced denaturation, in fact, provides partly folded states, which may resemble in some features early intermediates of folding and might be prone to aggregation or misfolding, especially when their secondary structure content is different from the folded protein one[103]. G37R, at variance, behaves very similarly to native SOD and this also correlates well with its lack of oligomerization. When examining the apo forms of mutants, on the contrary, S134N has a very similar unfolding pathway with respect to apo *as*SOD, while G37R shows an even larger α-propensity. Since this could be a symptom of aggregating tendency for the apo form, FFCR studies are in progress on the demetallated protein in order to assess whether it oligomerizes in the same fashion as S134N. H48Q, finally, shows elevate stability and low α-propensity also in its apo form, although oligomerization of the demetallated enzyme upon TFE addition is appreciable even by eyes. This is not unexpected, since mutant H46R, also mutated in a Cu binding residue, shows fibrillar aggregates in its crystal structure[13]. Even in this case, the characterization of molecular weight distributions of species in solution by FFCR can help to shed light on its oligomerization tendency.
To speak about the Cu binding site, the differences in CD and EPR spectra indicate deviations from the native coordination, with H48Q the more evident case since in this case the mutation completely disrupts the tetragonal coordination, although G37R and S134N present a certain grade of distortion, too. Similar results have been yielded from the previous characterization of some SOD mutants, not related to fALS, that showed a change in the Cu centre coordination, upon mutations in residues belonging to the charged Cu cavity[85]. The case of G37R is probably the most unsexpected, since mutation S134N takes place in a crucial region near the electrostatic loop, while position 37 is quite far from the metal binding region. To assess the variations from the native coordination, $J^2$ experiments are in progress, on both $Zn_4$ and $Cu_2Zn_2$G37R SOD.
To conclude, all the evidences gathered are in agreement with the rest of the quickly increasing body of evidences about fALS mutants, that predict different enzymes behaviour when regarding to MBR or WTL mutants with a certain uniformity within each of these classes, although the disease mechanisms are far from being deciphered. The most odd feature of fALS disorder is that

single mutations that in vitro show a so large variety of behaviours, cause so uniform disease symptoms. For this reason, more studies are needed in order to find common traits for the different enzyme variants. Also the record of cases should be widened, in order to ascertain whether the differences observed among WTL and MBR mutants, and apo and holo forms of the enzymes, are recurrent and whether they can be related to the processes occurring in vivo.

The investigation of the fundamentals of fibril formation by mean of *in vitro* studies of recombinant proteins, like the ones performed during this PhD, is a powerful approach to understand some of the molecular mechanisms involved in the development of pathological conditions of this type. Basically, the use of a model system allows to extrapolate general principles in order to interpret the behaviour of peptides and proteins that form fibrils *in vivo.* However, when exploiting this kind of approach, a researcher can become doubtful about the validity of his work. This is basically due to the inherent difficulty of isolating a protein in the same biochemical state as it is when it's expressed in the organism of interest, and to the great simplification imposed to the huge variety of interactions between molecules that can occur in biological cells and extracellular spaces, when dealing with a system like a recombinant protein/s sample. For these reasons, the *in vitro* study of cell cultures and the creation of mouse models expressing human mutated SOD1 are the naturally following steps once a deep understanding of the enzymes properties, needed to ascertain the intrinsic mechanism of their action, is achieved.

If referring to animal models of fALS, it can be said that up to now 12 mutant SOD1-fALS mice expressing different fALS mutant proteins as well as 2 fALS rats have been reported to develop the disease[74,104-106] WTL mutant SOD1 transgenic mice (e.g. G93A, G37R) present, at cellular level, prominent disease symptoms different from the ones developed by MBR (e.g. G85R, H46R, H46R/H48Q, H46R/H48Q/H63G/H120G) mutant SOD1 ones. Thus it is firmly established that the expression of ALS-mutant SOD1 proteins is the ultimate cause of motor neuron death; however the mechanism by which these SOD1 mutant proteins confer toxicity is still unknown. Again the number of transgenic mice bearing different mutations is quite limited and still it must be established whether these differences reflect a real change in the disorder pathogenesis as it is due to WTL or MBR mutants.

A lot of work has been spent on the characterization of fALS mutants, since the disease was mapped in the SOD1 chromosome in 1993, but it still seems very little, in comparison with the increasing number of disorder-related mutations and the high non-homogeneity of the resulting mutant proteins. Hopefully, the results emerged from the researches performed during this PhD course will help to shed some light on the intricate pathogenesis of this neurodegenerative disorder.

# 8. Bibliography


1. Uversky,V.N. & Fink,A.L. Conformational constraints for amyloid fibrillation: the importance of being unfolded. *Biochimica et Biophysica Acta-Proteins and Proteomics* 1698, 131-153 (2004).

2. Bertram,L. & Tanzi,R.E. The genetic epidemidogy of neurodegenerative disease. *Journal of Clinical Investigation* 115, 1449-1457 (2005).

3. Lander,E.S. The new genomics: Global views of biology. *Science* 274, 536-539 (1996).

4. Kelly,J.W. The alternative conformations of amyloidogenic proteins and their multi-step assembly pathways. *Curr. Opin. Struct. Biol.* 8, 101-106 (1998).

5. Alonso,A.D., Grundke-Iqbal,I., Barra,H.S. & Iqbal,K. Abnormal phosphorylation of tau and the mechanism of Alzheimer neurofibrillary degeneration: sequestration of microtubule-associated proteins 1 and 2 and the disassembly of microtubules by the abnormal tau. *Proc. Natl. Acad. Sci. USA* 94, 298-303 (1997).

6. Dobson,C.M. Protein misfolding, evolution and disease. *Trends in Biochemical Sciences* 24, 329-332 (1999).

7. Bellotti,V., Mangione,P. & Stoppini,M. Biological activity and pathological implications of misfolded proteins. *Cellular and Molecular Life Sciences* 55, 977-991 (1999).

8. Rochet,J.C. & Lansbury,P.T. Amyloid fibrillogenesis: themes and variations. *Curr. Opin. Struct. Biol.* 10, 60-68 (2000).

9. Bhatia,R., Lin,H. & Lal,R. Fresh and globular amyloid beta protein (1-42) induces rapid cellular degeneration: evidence for A beta P channel-mediated cellular toxicity. *FASEB J.* 14, 1233-1243 (2000).

10. Nilsberth,C. *et al.* The 'Arctic' APP mutation (E693G) causes Alzheimer's disease by enhanced A beta protofibril formation. *Nature Neuroscience* 4, 887-893 (2001).

11. Shipp,E., Cantini,F., Bertini,I., Valentine,J.S. & Banci,L. Dynamic properties of the G93A mutant of copper-zinc superoxide dismutase as detected by NMR spectroscopy: implications for the pathology of familial amyotrophic lateral sclerosis. *Biochemistry* 42, 1890-1899 (2003).

12. DiDonato,M. *et al.* ALS mutants of human superoxide dismutase form fibrous aggregates via framework destabilization. *J. Mol. Biol.* 332, 601-615 (2003).

13. Elam,J.S. *et al.* Amyloid-like filaments and water-filled nanotubes formed by SOD1 mutant proteins linked to familial ALS. *Nature Struct. Biol.* 10, 461-467 (2003).

14. Ellerby,M.L., Cabelli,D.E., Graden,J.A. & Valentine,J.S. Copper-Zinc Superoxide Dismutase: Why Not ph-Dependent? *J. Am. Chem. Soc.* 118, 6556-6561 (1996).

15. Hart,P.J. *et al.* Subunit asymmetry in the three-dimensional structure of a human CuZnSOD mutant found in familial amyotrophic lateral sclerosis. *Protein Sci* 7, 545-555 (1998).

16. Hayward,L.J. *et al.* Decreased metallation and activity in subsets of mutant superoxide dismutases associated with familial ALS. *J. Biol. Chem.* (2002).

17. Rodriguez,J.A. *et al.* Familial ALS-associated mutations decrease the thermal stability of distinctly metallated species of human copper-zinc superoxide dismutase. *J. Biol. Chem.* (2002).

18. Elam,J.S. *et al.* An alternative mechanism of bicarbonate-mediated peroxidation by copper-zinc superoxide dismutase - Rates enhanced via proposed enzyme-associated peroxycarbonate intermediate. *Journal of Biological Chemistry* 278, 21032-21039 (2003).

19. Dyson,H.J. & Wright,P.E. Intrinsically unstructured proteins and their functions. *Nat. Rev. Mol. Cell Biol.* 6, 197-208 (2005).

20. Assfalg,M., Banci,L., Bertini,I., Turano,P. & Vasos,P.R. Superoxide dismutase folding/unfolding pathway: role of the metal ions in modulating structural and dynamical features. *J. Mol. Biol.* 330, 145-158 (2003).

21. Banci,L., Bertini,I., Chiu,C.Y., Mullenbach,G.T. & Viezzoli,M.S. Synthesis and characterization of a monomeric mutein of Cu/Zn superoxide dismutase with partially reconstituted enzymatic activity. *Eur. J. Biochem.* 234, 855-860 (1995).

22. Hallewell,R.A. *et al.* Thermostabilization of recombinant CuZn superoxide dismutases by replacement of free cysteines. *Biochem. Biophys. Res. Commun.* 181, 474-480 (1991).

23. De Beus,M.D., Chung,J.Y. & Colon,W. Modification of cysteine 111 in Cu/Zn superoxide dismutase results in altered spectroscopic and biophysical properties. *Protein Sci.* 13, 1347-1355 (2004).

24. Kim,Y.J., Nakatomi,R., Akagi,T., Hashikawa,T. & Takahashi,R. Unsaturated fatty acids induce cytotoxic aggregate formation of amyotrophic lateral sclerosis-linked superoxide dismutase 1 mutants. *Journal of Biological Chemistry* 280, 21515-21521 (2005).

25. Chiti,F. *et al.* Designing conditions for in vitro formation of amyloid protofilaments and fibrils. *Proceedings of the National Academy of Sciences of the United States of America* 96, 3590-3594 (1999).

26. Chiti,F. *et al.* Mutational analysis of the propensity for amyloid formation by a globular protein. *Embo Journal* 19, 1441-1449 (2000).

27. McCord,J.M., Keele,B.B. & Fridovich,I. An enzyme-based theory of obligate anaerobiosis: the physiological function of superoxide dismutase. *Proc. Natl. Acad. Sci. USA* 68, 1024-1027 (1971).

28. Klug,D., Rabani,J. & Fridovich,I. A Direct Demonstration of the Catalytic Action of Superoxide Dismutase through the Use of Pulse Radiolysis. *J. Biol. Chem.* 247, 4839-4842 (1972).

29. Halliwell,B. Free radicals and antioxidants: a personal view. *Nutr. Rev.* 52, 253-265 (1994).

30. Richter,C. *et al.* Oxidants in mitochondria: from physiology to diseases. *Biochim. Biophys. Acta* 1271, 67-74 (1995).

31. Davies,K.J. Oxidative stress: the paradox of aerobic life. *Biochem. Soc. Symp.* 61, 1-31 (1995).

32. Estévez,A.G. *et al.* Induction of Nitric Oxide-Dependent Apoptosis in Motor Neurons by Zinc-Deficient Superoxide Dismutase. *Science* 286, 5449 (1999).

33. Rosselli,M., Keller,P.J. & Dubey,R.K. Role of nitric oxide in the biology, physiology and pathophysiology of reproduction. *Human Reproduction Update* 4, 3-24 (1998).

34. Sies,H. Oxidative Stress - from Basic Research to Clinical-Application. *American Journal of Medicine* 91, S31-S38 (1991).

35. Halliwell,B. & Whiteman,M. Measuring reactive species and oxidative damage in vivo and in cell culture: how should you do it and what do the results mean? *British Journal of Pharmacology* 142, 231-255 (2004).

36. Bertini,I., Mangani,S. & Viezzoli,M.S. Advanced Inorganic Chemistry. Sykes,A.G. (ed.), pp. 127-250 (Academic Press, San Diego, CA, USA,1998).

37. Banci,L. *et al.* Solution structure of reduced monomeric Q133M2 Copper, Zinc Superoxide Dismutase. Why is SOD a dimeric enzyme? *Biochemistry* 37, 11780-11791 (1998).

38. Getzoff,E.D. *et al.* Elecrostatic Recognition between Superoxide and Copper, Zinc Superoxide Dismutase. *Nature* 306, 287-290 (1983).

39. Tainer,J.A., Getzoff,E.D., Richardson,J.S. & Richardson,D.C. Structure and Mechanism of Copper, Zinc Superoxide Dismutase. *Nature* 306, 284-287 (1983).

40. Valentine,J.S. & Pantoliano,M.W. Copper Proteins. Spiro,T.G. (ed.), pp. 291 (Wiley, New York,1981).

41. Tainer,J.A., Getzoff,E.D., Beem,K.M., Richardson,J.S. & Richardson,D.C. Determination and Analysis of 2 A Structure of Copper Zinc Superoxide Dismutase. *J. Mol. Biol.* 160, 181-217 (1982).

42. Banci,L., Bencini,A., Bertini,I., Luchinat,C. & Piccioli,M. $^{1}$H NOE and ligand field studies of copper-cobalt superoxide dismutase with anions. *Inorg. Chem.* 29, 4867-4873 (1990).

43. Lepock,J.R., Frey,H.E. & Hallewell,R.A. Contribution of conformational stability and reversibility of unfolding to the increased thermostability of human and bovine superoxide dismutase mutated at free cysteines. *J. Biol. Chem.* 265, 21612-21618 (1990).

44. Parge,H.E., Hallewell,R.A. & Tainer,J.A. Atomic structures of wild-type and thermostable mutant recombinant human Cu,Zn superoxide dismutase. *Proc. Natl. Acad. Sci. USA* 89, 6109-6114 (1992).

45. Hongbin,L. *et al.* Copper(2+) binding to the surface residue cysteine 111 of His46Arg human copper-zinc superoxide dismutase, a Familial Amyotrophic Lateral Sclerosis mutant. *Biochemistry* 39, 8125-8132 (2000).

46. Petrovic,N., Comi,A. & Ettinger,M.J. Identification of an apo-superoxide dismutase (Cu,Zn) pool in human lymphoblasts. *Journal of Biological Chemistry* 271, 28331-28334 (1996).

47. Petrovic,N., Comi,A. & Ettinger,M.J. Copper incorporation into superoxide dismutase in menkes lymphoblasts. *Journal of Biological Chemistry* 271, 28335-28340 (1996).

48. Waggoner,D.J. *et al.* Brain copper content and cuproenzyme activity do not vary with prion protein expression level. *J. Biol. Chem.* 275, 7455-7458 (2000).

49. Rossi,L., Marchese,E., DeMartino,A., Rotilio,G. & Ciriolo,M.R. Purification of a fully metal-depleted Cu, Zn superoxide dismutase from copper-deficient rat liver. *Biometals* 10, 257-262 (1997).

50. Furukawa,Y. & O'Halloran,T.V. 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 (2005).

51. Forman,H.J. & Fridovich,I. On the stability of bovine superoxide dismutase. *J. Biol. Chem.* 248, 2645-2649 (1973).

52. Bertini,I., Luchinat,C., Monnanni,R., Scozzafava,A. & Borghi,E. Investigation of zinc-deprived bovine superoxide dismutase. *Inorg. Chim. Acta* 91, 109-111 (1984).

53. Banci,L. *et al.* Backbone Dynamics of Human Cu, Zn Superoxide Dismustase and of its Monomeric F50/EG51E/E133Q Mutant: The influence of Dimerization on Mobility and Function. *Biochemistry* 39, 9108-9118 (2000).

54. Banci,L., Bertini,I., Cramaro,F., Del Conte,R. & Viezzoli,M.S. The solution structure of reduced dimeric copper zinc SOD: the structural effects of dimerization. *Eur J. Biochem.* 269, 1905-1915 (2002).

55. Strange,R.W. *et al.* The Structure of Holo and Metal-deficient Wild-Type Human Cu,Zn Superoxide Dismutase and its Relevance to Familial Amyothrophic Lateral Sclerosis. *J. Mol. Biol.* 28, 877-882 (2003).

56. Bertini,I., Piccioli,M., Viezzoli,M.S., Chiu,C.Y. & Mullenbach,G.T. A spectroscopic characterization of a monomeric analog of copper-zinc superoxide dismutase. *Eur. J. Biophys.* 23, 167-176 (1994).

57. Getzoff,E.D. *et al.* Faster Superoxide Dismutase Mutants designed by Enhancing Electrostatic Guidance. *Nature* 358, 347-351 (1992).

58. Banci,L., Bertini,I., Cantini,F., D'Onofrio,M. & Viezzoli,M.S. Structure and dynamics of copper-free SOD: The protein before binding copper. *Protein Sci.* 11, 2479-2492 (2002).

59. Banci,L., Bertini,I., Cramaro,F., Del Conte,R. & Viezzoli,M.S. Solution structure of apo Cu,Zn superoxide dismutase: the role of metal ions in protein folding. *Biochemistry* 42, 9543-9553 (2003).

60. Rosen,D.R. *et al.* Mutation in Cu,Zn superoxide dismutase gene are associated with familial amyotrophyc lateral sclerosis. *Nature* 362, 59-62 (1993).

61. Cleveland,D.W. & Rothstein,J.D. From Charcot to Lou Gehrig: deciphering selective motor neuron death in ALS. *Nat. Rev. Neurosci.* 2, 806-819 (2001).

62. Eisen,A. Recent considerations in the etiopathogenesis of ALS. *Suppl Clin Neurophysiol.* 57, 187-190 (2004).

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

64. Andersen,P.M. *et al.* Sixteen novel mutations in the Cu/Zn superoxide dismutase gene in amyotrophic lateral sclerosis: a decade of discoveries, defects and disputes. *Amyotrophic Lateral Sclerosis and Other Motor Neuron Disorders* 4, 62-73 (2003).

65. Andersen,P.M. *et al.* Amyotrophic-Lateral-Sclerosis Associated with Homozygosity for An Asp90Ala Mutation in Cuzn-Superoxide Dismutase. *Nature Genetics* 10, 61-66 (1995).

66. Jonsson,P.A. *et al.* CuZn-superoxide dismutase in D90A heterozygotes and dominant from recessive ALS pedigrees. *Neurobiology of Disease* 10, 327-333 (2002).

67. Bruneteau,G., Demeret,S. & Meininger,V. Physiopathology of ALS: therapeutic approach. *Revue Neurologique* 160, 235-241 (2004).

68. Smith,R.G. & Appel,S.H. Molecular Approaches to Amyotrophic-Lateral-Sclerosis. *Annual Review of Medicine* 46, 133-145 (1995).

69. Horner,R.D. *et al.* Occurrence of amyotrophic lateral sclerosis among Gulf War veterans. *Neurology* 61, 742-749 (2003).

70. Piazza,O., Siren,A.L. & Ehrenreich,H. Soccer, neurotrauma and amyotrophic lateral sclerosis: is there a connection? *Current Medical Research and Opinion* 20, 505-508 (2004).

71. Inge,P.G. & Codd,G.A. Return of the cycad hypothesis - does the amyotrophic lateral sclerosis/parkinsonism dementia complex (ALS/PDC) of Guam have new implications for global health? *Neuropathol Appl Neurobiol.* 31, 345-353 (2005).

72. Bensimon,G. *et al.* A study of riluzole in the treatment of advanced stage or elderly patients with amyotrophic lateral sclerosis. *Journal of Neurology* 249, 609-615 (2002).

73. Gurney,M.E. *et al.* Motor Neuron Degeneration in Mice That Express a Human Cu,Zn Superoxide Dismutase Mutation. *Science* 264, 1772-1775 (1994).

74. Reaume,A.G. *et al.* Motor neurons in Cu/Zn superoxide dismutase-deficient mice develop normally but exhibit enhanced cell death after axonal injury. *Nat Genetic* 13, 43-47 (1996).

75. Wong,P.C. *et al.* An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria. *Neuron* 14, 1105-1116 (1995).

76. Bruijn,L.I. *et al.* ALS-linked SOD1 mutant G85R mediates damage to astrocytes and promotes rapidly progressive disease with SOD1-containing inclusions. *Social History of Medicine* 327-338 (1997).

77. Stathopulos,P.B. *et al.* Cu/Zn superoxide dismutase mutants associated with amyotrophic lateral sclerosis show enhanced formation of aggregates in vitro. *Proc. Natl. Acad. Sci. U. S. A* 100, 7021-7026 (2003).

78. Wood,J.D., Beaujeux,T.P. & Shaw,P.J. Protein aggregation in motor neurone disorders. *Neuropathology and Applied Neurobiology* 29, 529-545 (2003).

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

80. Juneja,T., Pericak-Vance,M.A., Laing,N.G., Dave,S. & Siddique,T. Prognosis in familial amyotrophic lateral sclerosis: progression and surivial in patients with Glu100Gly and Ala4Val mutations in Cu,Zn superoxide dismutase. *Neurology* 48, 55-57 (1997).

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

82. Rabizadeh,S. *et al.* Mutations associated with amyotrophic lateral sclerosis convert superoxide dismutase from an antiapoptotic gene to a proapoptotic gene: studies in yeast and neural cells. *Proc. Natl. Acad. Sci. USA* 92, 3024-3028 (1995).

83. Liochev,S.I., Chen,L.L., Hallewell,R.A. & Fridovich,I. Superoxide-dependent peroxidase activity of H48Q: A superoxide dismutase variant associated with familiar amyotrophic lateral sclerosis. *Archives of Biochemistry and Biophysics* 346, 263-268 (1997).

84. Macura,S., Wüthrich,K. & Ernst,R.R. The relevance of *J* cross-peaks in two-dimensional NOE experiments of macromolecules. *J. Magn. Reson.* 47, 351-357 (1982).

85. Banci,L., Bencini,A., Bertini,I., Luchinat,C. & Viezzoli,M.S. The angular overlap analysis of the spectroscopic parameters of copper zinc SOD and its mutants. Dedicated to Prof. Angelo Mangini. *Gazz. Chim. Ital.* 120, 179-185 (1990).

86. Cole,K.S. & Cole,R.H. Dispersion and absorption in dielectrics. I. Alternating-current characteristics. *J. Chem. Phys.* 9, 341-351 (1941).

87. Hallenga,K. & Koenig,S.H. Protein Rotational Relaxation as Studied by Solvent $^{1}H$ and $^{2}H$ Magnetic Relaxation. *Biochemistry* 15, 4255-4264 (1976).

88. Lopiano,L. *et al.* Nuclear magnetic relaxation dispersion profiles of substantia nigra pars compacta in Parkinson's disease patients are consistent with protein aggregation. *Neurochemistry International* 37, 331-336 (2000).

89. Halle,B., Jóhannesson,H. & Venu,K. Model-Free Analysis of Stretched Relaxation Dispersions. *J. Magn. Reson.* 135, 1-13 (1998).

90. Borgias,B., Thomas,P.D. & James,T.L. COmplete Relaxation Matrix Analysis (CORMA). (5.0). 1989. San Francisco, CA, University of California.
Ref Type: Computer Program

91. Borgias,B., Gochin,M., Kerwood,D.J. & James,T.L. Relaxation matrix analysis of 2D NMR data. *Progr. NMR Spectrosc.* 22, 83-100 (1990).

92. Perczel,A., Park,K. & Fasman,G.D. Analysis of the circular dichroism spectrum of proteins using the convex constraint algorithm: a practical guide. *Anal. Biochem.* 203, 83-93 (1992).

93. Provencher,S.W. & Glockner,J. Estimation of globular protein secondary structure from circular dichroism. *Biochemistry* 20, 33-37 (1981).

94. Hennessey,J.P., Jr. & ,a.J.W.C.Jr. Information content in the circular dichroism of proteins. *Biochemistry* 20, 1085-1094 (1981).

95. Sreerama,N. & Woody,R.W. A self-consistent method for the analysis of protein secondary structure from circular dichroism. *Anal Biochem* 32-44 (1993).

96. Manavalan,P. & Johnson,W.C., Jr. Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. *Anal Biochem* 167, 76-85 (1987).

97. Andrade,M.A., Chacon,P., Merelo,J.J. & and Moran,F. Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network. *Protein Eng.* 6, 389-390 (1993).

98. Bohm,G., Jaenicke,R. & Muhr,R. Quantitative analysis of protein far UV circular dichroism spectra by neural networks. *Protein Eng.* 5, 191-195 (1992).

99. Sreerama,N., Venyaminov,S.J. & Woody,R.W. Analysis of Protein Circular Dichroism Spectra Based on the Tertiary Structure Classification. *Anal. Biochem.* 299, 271 (2001).

100. Mei,G. *et al.* Denaturation of Human Cu/Zn Superoxide Dismutase by Guanidine Hydrochloride: A Dynamic Fluorescence Study. *Biochemistry* 31, 7224-7230 (1992).

101. Banci,L., Bertini,I. & Del Conte,R. The solution structure of apo CopZ from *Bacillus subtilis*: a further analysis of the changes associated with the presence of copper. *Biochemistry* 42, 13422-13428 (2003).

102. Deng,H.X. *et al.* Two novel SOD1 mutations in patients with familial amyotrophic lateral sclerosis. *Hum. Molec. Genet.* 4, 1113-1116 (1995).

103. Gast,K., Zirwer,D., Muller-Frohne,M. & Damaschun,G. Trifluoroethanol-induced conformational transitions of proteins: Insights gained from the differences between alpha-lactalbumin and ribonuclease A . *Protein Sci.* 8, 625-634 (1999).

104. Turner,B.J., Lopes,E.C. & Cheema,S.S. Inducible superoxide dismutase 1 aggregation in transgenic amyotrophic lateral sclerosis mouse fibroblasts. *Journal of Cellular Biochemistry* 91, 1074-1084 (2004).

105. Nagai,M. *et al.* Rats expressing human cytosolic copper-zinc superoxide dismutase transgenes with amyotrophic lateral sclerosis: Associated mutations develop motor neuron disease. *Journal of Neuroscience* 21, 9246-9254 (2001).

106. Shibata,N. Transgenic mouse model for familial amyotrophic lateral sclerosis with superoxide dismutase-1 mutation. *Neuropathology* 21, 82-92 (2001).