Characterization of the Recombinant Copper Chaperone (CCS) for the Copper/Zinc Superoxide Dismutase (CuZnSOD) from the plant *Glycine (G.) max*.

Sara Sagasti¹, Inmaculada Yruela¹, Maria Bernal^{1,&}, Maria A. Lujan¹, Susana Frago², Milagros Medina², Rafael Picorel^{1,\$}

¹Estación Experimental de Aula Dei (EEAD), Consejo Superior de Investigaciones Científicas (CSIC), Avda. Montañana 1005, E-50059 Zaragoza, Spain

(picorel@eead.csic.es;yruela@eead.csic.es;lujan@eead.csic.es;sagasti@eead.csic.es).

²Departamento de Bioquímica y Biología Molecular y Celular, Facultad de Ciencias and Instituto de Biocomputación y Física de Sistemas Complejos (BIFI), Universidad de Zaragoza, C/ Pedro Cerbuna 12, E-50009 Zaragoza, Spain (mmedina@unizar.es).

[&]Current address: LS Pflanzenphysiologie, Ruhr-Universität Bochum, ND Nord 3, Rm 3/71-73, Universitätsstraße 150, 44801 Bochum, Germany (maria.bernal@rub.de).

^{\$}**Corresponding author:** Rafael Picorel, Estación Experimental de Aula Dei (CSIC), Avda. Montañana 1005, E-50059 Zaragoza, Spain. Phone: 34-976-716053; FAX: 34(976)716145; e-mail:picorel@eead.csic.es.

SUMMARY

The goal of the present work was to characterize the copper chaperone from soybean. Very little is known about plant copper chaperones, which makes this study of current interest, and allows for a comparison with the better known homologues from yeast and human. To obtain sizeable amounts of pure protein suitable for spectroscopic characterization, we cloned and overexpressed the *G. max* CCS chaperone in *E. coli* in the presence of 0.5 mM copper and 0.5 mM zinc in the broth. A pure protein preparation was obtained by using two IMAC steps and pH gradient chromatography. The chaperone preparation thus obtained contained about 0.3 copper atoms per protein, indicating most of the chaperone was in its apo-form. The obtained protein preparation interacted with both copper and zinc ions in solution as assessed by absorption and fluorescence spectroscopy. The protein structure showed an unusual high content (i. e., over 40%) of loops, turns and random coil as determined both by circular dichroism and homology modelling, which indicated a very flexible structure. The homology 3-D model suggests the protein might fold in three structural protein domains. The domain I and III contain the two putative copper binding sites MXCXXC and CXC, respectively.

Key Words: Chaperone, cloning, copper, modelling, purification, zinc.

INTRODUCTION

Copper (Cu) is an essential transition metal used as a cofactor in many redox proteins in all living organisms. In plants, Cu-dependent metalloenzymes are found in multiple cellular locations including the cell wall, cytoplasmic membrane, chloroplast, cytosol, and mitochondrion among others. Eukaryotic cells use a family of Cu-carrying proteins named metallochaperones that deliver the metal ion to distinct intracellular targets requiring Cu as prosthetic group (O'Halloran and Culotta 2000; Arnesano et al. 2002). Cu-chaperones may also be directly involved in the binding of free Cu ions in cells to avoid their potential cellular damage. The first member of this protein family was identified and characterized in *Saccharomyces (S.) cerevisiae*, LYS7/yCcs1 (also named yCCS herein). The gene was originally identified as involved in lysine biosynthetic pathways (Horecka et al. 1995; Lin and Culotta 1995; Pufahl et al. 1997). Later on a gene for a similar protein was identified in humans (*hCCS*), and it was postulated to be required for Cu insertion into the active site of the cytosolic CuZn superoxide dismutase (CuZnSOD) enzyme (Culotta et al. 1997). Functional homolog proteins have been noted in several mammals, insects and plants. In plants, sequences encoding CCS have been identified in tomato (*Lycopersicon (L.) esculentum*; *LeCCS*) (Zhu et al. 2000), *Arabidopsis (A.) thaliana* (*AtCCS*) (Wintz and Vulpe 2002), potato (*Solanum (S.) tuberosum*; *StCCS*) (Trindade et al. 2003), and maize (*Zea (Z.) maize*; *ZmCCS*) (Ruzsa and Scandalios 2003).

Biochemical and crystallographic studies of the protein in yeast (yCCS) and human (*hCCS*) have identified some structural and functional aspects (Lamb et al. 1999; Schmidt et al. 2000; Strasser et al. 2007) of this chaperone family. CCS is a 26-30 kDa soluble protein with three main domains. Domain I at the N-terminal region is an ATX1-like polypeptide, which binds a single Cu⁺ ion at the MXCXXC metal binding-

motif. This domain is thought to sequester Cu under metal-starvation conditions. Domain II in the central region provides a high homology to its target protein, CuZnSOD. It is involved in recognition and binding to its CuZnSOD partner. Domain III at the C-terminal region is a highly disordered short tail. It contains a CXC Cu-binding site that can interact with domain I, and this interaction is thought to be crucial for Cu delivery to the apo-CuZnSOD *in vivo*. Studies of recombinant yCCS show that its apo-form is a monomer, while the Cu-yCCS form can exist in monomer/dimer equilibrium (Schmidt et al. 1999a,b). Several crystal structures of CCS (Lamb et al. 1999, 2000), CuZnSOD (Djinovic et al. 1992; Banci et al. 2003) and CCS-CuZnSOD heterodimer complex (Lamb et al. 2001) are already available. But details of the functional mechanisms of protein recognition and docking, and metal ion transfer are yet to be understood.

In plants, no CCS structural data are available, though a partial biochemical characterization of *Le*CCS is available (Zhu et al. 2000). In the current study, we use an *Escherichia (E.) coli*-based expression system for the overproduction of soybean CCS. The purified recombinant protein was characterized by immunochemical reactivity, mass spectrometry (MS), and several spectroscopic techniques (i.e., UV-Vis, circular dichroism (CD) and fluorescence emission). Moreover, a structural model for the *Gm*CCS is shown using yCCS as a template.

MATERIALS AND METHODS

Photosynthetic cell suspension growth

Photosynthetic cell suspensions from soybean (*Glycine max* var. Corsoy) SB-P line were grown as described by Rogers et al. (1987) with some modifications (Alfonso et al. 1996; Bernal et al. 2006). Cell suspensions were grown in photomixotrophic (KN¹)

medium under continuous light ($30 \pm 5 \mu\text{E m}^{-2} \text{ s}^{-1}$) and air with 5% CO_2 supplementation at 24 °C on a rotatory shaker (TEQ Model OSFT-LS-R, Massamá, Portugal) at 110 rpm. Cells used for experiments were collected from liquid cultures after 21 days of growth.

RNA preparation and cDNA synthesis

Cell suspensions were filtered through Miracloth paper and the resulting cells (0.4 g wet weight) were washed with fresh growth medium and pestled in liquid nitrogen. Total RNA was isolated and purified from soybean cells using the RNeasy Plant Mini Kit (Quiagen GmbH, Hiden, Germany). The RNA concentration was determined by measuring the O.D. at 260 nm. cDNA were synthetized from total RNA (5 μg) using 200 units of reverse transcriptase (M-MLV reverse transcriptase, Promega, Madison, WI, USA) and 1 μM oligo(dT)₁₂₋₁₈ from Invitrogen (Invitrogen Life Technologies, Carlsbad, CA, USA), according to the manufacturer's instructions.

Recombinant vector construction and *GmCCS* gene cloning

RT-PCR was used to isolate a DNA sequence corresponding to the entire soybean CCS (*GmCCS*) gene using cDNA. Specific gene primers were designed on the sequence of *GmCCS* (AF329816) to obtain an entire open reading frame with the putative signal peptide sequence (Forward: 5'-GGATCCATGGCATTCTGAGGTCA-3' with the *Bam*HI site underlined; Reverse: 5'-TCTAGACTATCAGACCTTGCTAGT-3' with the *Xba*I site underlined). The PCR reaction was done as follows: an initial denaturing step at 94 °C for 3 min, 35 cycles at 94 °C for 30 s, 54 °C for 30 s and 72 °C for 90 s. A final elongation step was performed at 72 °C for 10 min. The PCR product was cloned into the pGEM-T Easy vector (Promega, Madison, WI, USA) and subsequently sequenced

(CNIO Sequencing Service, Madrid, Spain). The plasmid was digested with *Bam*HI and *Xba*I before the isolated fragment was ligated into the modified pMALc2x expression vector (Pryor and Leiting 1992) with a His₆-Tag pretreated in the same way with both enzymes. The ligation mixture was transformed into *E. coli* DH5 α competent cells to generate ampicillin-resistant colonies. Bacterial colonies were cultured in LB medium supplemented with ampicillin. Plasmid DNA was isolated using High Pure Plasmid Isolation Kit (Roche, Mannheim, Germany), according to the manufacturer's instructions, and verified by sequencing.

Expression of the MBP-His₆-*Gm*CCS fusion protein

One colony of *E. coli* DH5 α competent cells bearing the modified plasmid pMALc2x-*Gm*CCS was grown overnight at 37 °C in 4-ml LB medium supplemented with 100 μ g ml⁻¹ ampicillin. A pre-culture was used to inoculate 500-ml fresh LB broth and the culture grown up to an O.D._{600nm} = 0.6 units at 37 °C. Then 0.3 mM IPTG, 0.5 mM CuSO₄, and 0.5 mM ZnSO₄ were added to the broth and the growth continued for an additional 3 h period. The supplementation of Cu and Zn ions to the broth was not toxic to the cells, and resulted in a significant increase in the expression of the fusion protein as we will describe later. Cells were harvested by centrifugation at 10000 x g for 5 min at 4 °C and kept frozen at -80 °C until use. Protein expression levels were estimated by SDS-PAGE and Coomassie Brilliant Blue R-250 protein staining (Sigma-Aldrich, Saint Louis, MO, USA).

Isolation and purification of the *Gm*CCS recombinant protein

Thawed cells were suspended in buffer containing 20 mM sodium phosphate, pH 7.4, and 500 mM NaCl with 1 mM pefabloc and 1 mM phenylmethylsulphonyl fluoride

(PMSF) as protease inhibitors, and then broken by sonication (6 x 1 min) (Sonicator Ultrasonic Processor XL Misonix, Farmingdale, NY, USA). One-min waiting intervals between sonication periods were used with the cell suspension container maintained cold on ice. Debris and big particles were separated by centrifugation at 14500 x g for 30 min at 4 °C. The supernatant was filtered through a 0.45 µm membrane and applied at a flow rate of 1 ml min⁻¹ onto a 5 ml Hitrap Ni²⁺-IMAC column (GE Healthcare), pre-equilibrated with 20 mM sodium phosphate, pH 7.4, and 500 mM NaCl. Then the column was washed with 5 column volumes of the same buffer and subsequently with 5 column volumes of 20 mM sodium phosphate, pH 7.4, 500 mM NaCl, and 50 mM imidazole. The fusion protein was eluted with a 50-450 mM imidazole linear gradient in the same buffer. Fractions of 0.75 ml each were collected at a flow rate of 0.5 ml min⁻¹ and analysed by SDS-PAGE (15%, w/v). Fractions enriched with MBP-His₆-*GmCCS* fusion protein were pooled and dialysed against the protein cleavage buffer containing 20 mM sodium phosphate, pH 7.0, and 140 mM NaCl. The fusion protein was cleaved by thrombin protease (10 units per mg protein) for 22 h at room temperature (around 22 °C). The resultant protein solution was loaded again onto a second Hitrap Ni²⁺-IMAC column and then eluted with a 0-450 mM imidazole linear gradient. Fractions containing *GmCCS* protein as assayed by SDS-PAGE were pooled and dialysed against 20 mM BIS-TRIS propane, pH 7.5, and 20 mM NaCl. The sample was then applied onto a weak anionic-exchange Toyopearl TSK DEAE 650s column (Tosoh Bioscience GmbH, Stuttgart, Germany) pre-equilibrated with the dialysis buffer, and washed with 5 column volumes of the same buffer. Sample elution was done with a 7.5-4.5 pH linear gradient in the same buffer at a flow rate of 1 ml min⁻¹ at 4 °C. The purity of *GmCCS* preparation was determined by SDS-PAGE analysis, and the protein concentration was

determined by using its theoretical molar extinction coefficient $\epsilon_{280\text{nm}}$ ($\text{M}^{-1} \text{cm}^{-1}$) = 15470 according to the deduced primary protein structure (Expasy Server).

Metal content determination by inductively coupled plasma-mass spectrometry (ICP-MS)

After extensive dialysis to remove adventitiously protein bound metals, metal concentrations were determined by ICP-MS (Scientific-Technical Services, University of Barcelona, Barcelona, Spain). For a typical determination, 50 μl of a 6 μM protein solution were digested with 100 μl HNO_3 (Baker Instra) for 7 h at 60 $^\circ\text{C}$. After protein digestion and before measurements, the final volume was raised up to 2 ml by adding 1.850 ml of DI water (Elga Purelab Ultra). The determination was done with an Agilent 7500 CE Instrument using Rh as internal standard.

Matrix-assisted laser desorption-time of flight (MALDI-TOF) analysis

GmCCs protein band was cut off from the SDS-PAGE gel and analyzed with a MALDI-TOF/TOF 4700 Proteomics Analyzer, Applied Biosystems (Proteomics Platform Service, Parc Científic of Barcelona, Barcelona, Spain).

Determination of the N-terminal amino acid sequence

The purified protein was separated in SDS-PAGE, electroblotted to a polyvinylidene difluoride (PVDF) membrane and stained with Coomassie Blue. The band containing the *GmCCS* was cut off and stored at 4 $^\circ\text{C}$ until use. N-terminal automated sequence analysis was performed with a Perkin Elmer Procise 494 (CIB-CSIC Service, Madrid, Spain) and with an Applied Biosystems Procise 492 protein sequencer (Barcelona Autonomous University, Barcelona, Spain).

***Gm*CCS antibody production**

A rabbit anti-*Gm*CCS polyclonal antibody (Genosphere Biotechnologies, Paris, France) was produced using a 15 amino acid synthetic peptide. The selected peptide corresponded to the sequence –CVPEDFLISAAVSEF- in structural domain I of the mature protein. The anti-*Gm*CCS antibody was subsequently purified by affinity chromatography using the same produced synthetic peptide (Genosphere Biotechnologies, Paris, France).

Western blotting

Protein denaturing procedure was carried out by boiling the samples for 5 min using a sample buffer containing 2.3% (w/v) SDS, 10% (v/v) glycerol, 0.02% (w/v) bromophenol blue, 5% (v/v) β -mercaptoethanol, and 62 mM TRIS-HCl, pH 6.8. Protein concentration was determined according to Bradford (1976) using BSA as a standard. Sample proteins were separated by SDS-PAGE and transferred onto nitrocellulose membranes with a transfer system (Biorad, Hercules, CA, USA). After blocking the membrane with 5% (v/v) skimmed milk, it was incubated with the specific anti-*Gm*CCS antibody described above or with a commercial monoclonal antibody anti-His₆. Bands were developed with Amplified Alkaline Phosphatase Goat Anti-rabbit Immune-Blot Assay kit (Biorad) for the anti-*Gm*CCS or Horseradish Peroxide for the anti-His₆, according to the manufacturer's instructions.

Characterization of metal binding to *Gm*CCS

Purified protein was transferred into adequate buffer depending on the experiment (see herein) by three cycles in 10 times more dialysis solution volume, and subsequent concentration with a Centricon-10 centrifugal filter (Millipore, Bedford, MA, USA).

UV titration of a *GmCCS* solution (3 μM final protein concentration) with Cu^+ ions added as 100 μM or 1 mM CuCl (Sigma-Aldrich) in saturated NH_4Cl was performed in 20 mM phosphate, pH 7.0, and 150 mM NaCl at 25 $^\circ\text{C}$. The absorbance change in the 220-320 nm spectral range was monitored. Difference absorption spectra were generated after subtracting the protein solution alone and the Cu^+ solution alone at the corresponding metal concentrations from the protein/metal mixture solutions. The double reciprocal plot representation of the absorption at 240 nm ($\text{Abs}_{240\text{nm}}$) versus Cu^+ concentrations was used to estimate the dissociation constant (K_d) of the Cu^+ -*GmCCS* complex. UV-Vis titration of a *GmCCS* solution (5 μM final protein concentration) with Zn^{2+} was performed in the presence of 10 μM mag-fura-2 probe (Molecular Probes, Eugene, OR, USA) in buffer 20 mM HEPES, pH 7.0, 150 mM NaCl , and 10 mM ascorbic acid at 25 $^\circ\text{C}$. The absorbance in the 300-400 nm spectral range was monitored. The double reciprocal plot representation of the absorption at 322 nm *minus* 400 nm ($\Delta\text{Abs}_{322\text{nm}-400\text{nm}}$) between different Zn^{2+} concentrations and the control (without metal) was used to estimate the K_d of the Zn^{2+} -*GmCCS* complex. In both titrations, all solutions were bubbled with nitrogen for several min before titration measurements. The spectra were recorded with a Beckman DU-640 photodiode array spectrophotometer (Beckman Instruments, Fullerton, CA, USA).

Cu^+ binding to *GmCCS* (2 μM final protein concentration) was also studied by monitoring changes in the intrinsic protein fluorescence during metal titration in 20 mM phosphate, pH 7.0, and 150 mM NaCl at 25 $^\circ\text{C}$. Prior to metal titration, the solutions were bubbled with argon for several min. Intrinsic tryptophan (Trp) fluorescence

emission was measured after excitation at 290 nm. At this specific wavelength, Trp is preferentially excited, whereas the extinction coefficient of the other protein fluorescent residues, such as tyrosine (Tyr) and phenylalanine (Phe), are considerably lower. Resonance fluorescence energy transfer may also occur between these two residues and Trp. These considerations allow us to ascribe the observed fluorescence emission to Trp. Fluorescence spectra were recorded with an Aminco-Bowman Series 2 spectrofluorimeter in the 300-400 nm spectral range.

Far-UV CD, in the 200-250 nm spectral range, was measured in a 0.1-cm path-length cuvette (Chirascan Spectropolarimeter, Applied Photophysics Ltd., Letherhead, Surrey, UK). *GmCCS* (14 μ M) in 5 mM phosphate, pH 7.5, at 25 °C was bubbled with nitrogen for several min. Background spectra obtained with buffer solutions without protein and supplemented with the indicated concentrations of CuCl₂, were subtracted as blanks. The data were analyzed in the Dichroweb site (<http://www.cryst.bbk.ac.uk/cdweb/html/home.html>) using the K2d analysis algorithm (Whitmore and Wallace 2004, 2008). The spectra in the 200-250 nm range were used for the secondary structure simulation.

Multiple alignment and homology modelling of *GmCCS*

GmCCS sequence comparison was performed using CLUSTAL W (Thompson et al. 1994). The template for homology modelling of *GmCCS* was yCCS [1jk9_B.pdb] (Lamb et al. 2001). The best aligned template provided by PDB-RPS server, pdb_1jk9, was used to drive homology modelling with Modeller (Eswar et al. 2006). Protein structure was visualized with PyMol Molecular Viewer (Schrödinger, LLC).

RESULTS AND DISCUSSION

Cloning, overexpression and purification of *GmCCS* protein

The full-length cDNA sequence encoding the *GmCCS* was generated by RT-PCR and inserted into a modified expression vector, pMALc2x, with a His₆-Tag and a Thrombin protease site. The strategy to clone the *GmCCS* gene and overexpress the pMALc2x-*GmCCS* fusion protein in *E. coli* is described in Fig. 1. Transformed *E. coli* DH5 α competent cells without induction (control cells) showed no fusion protein (Fig. 2a, lane 1). After induction with 0.3 mM IPTG, the extract showed certain degree of expression of the fusion protein (Fig. 2a, lane 2). But maximum levels of expression were obtained after induction with 0.3 mM IPTG in the presence of 0.5 mM CuSO₄ and 0.5 mM ZnSO₄ in the broth (Fig. 2a, lane 3). The apparent molecular weight of the MBP-His₆-*GmCCS* fusion protein was about 80 kDa. A MBP-His₆ truncated protein was also induced with an apparent molecular weight of about 47 kDa (Fig. 2b). The molecular weight based on their derived amino acid sequences for MBP-His₆-*GmCCS* and MBP-His₆ were 75.6 and 52.6 kDa, respectively. Note the high immunoreactivity of the truncated polypeptide compared to the fusion protein (Fig. 2b), which may indicate a higher exposure of the His₆-Tag to the antibody (Ab) in the truncated polypeptide. This is a clear example of potential misleading results when using an Ab for protein quantitation without proper controls.

Cells were broken by ultrasonication, and the clarified crude extract was fractionated on the first Hitrap Ni²⁺-IMAC column. Two main bands were observed in a SDS-PAGE and Coomassie Blue staining, which corresponded to the fusion and truncated polypeptides, respectively. Another less intense band is clearly visible at around 33 kDa (Fig. 3a, lane 2). Most of other extracted proteins did not bind to the IMAC column. The resultant protein solution was then treated with Thrombin protease for 22 h at room temperature. After thrombin digestion, the fusion protein band almost disappeared, but

the truncated polypeptide band increased with a new band appearing at around 37 kDa (Fig. 3a, lane 3). This later band and that at around 33 kDa described earlier correspond to *GmCCS* since both immunoreacted with the obtained Ab against a synthetic peptide of the domain I of the chaperone (Fig. 3b, lane 3). This specific synthetic peptide was chosen within the domain I of the chaperone to avoid cross-reactivity with the CuZnSOD. Note that the domain II of the chaperone exhibits a high homology with the CuZnSOD as described in the Introduction. The digested protein solution was loaded onto a second Hitrap Ni²⁺-IMAC column aiming to eliminate the truncated polypeptide due to the presence of the His₆-Tag (Fig. 3a, lane 4). The fractions containing the *GmCCS* were pooled, dialyzed and applied onto a weak anionic-exchange Toyopearl TSK DEAE 650s column, and the proteins were eluted with a 7.5-4.5 pH linear gradient. This last chromatography yielded a highly pure *GmCCS* protein forming a doublet band in the SDS-PAGE (Fig. 3a, lane 5). The overall yield of the purification procedure was about 0.5 ml of pure protein solution of 60 µM from 0.5 L of bacterial culture.

ICP-MS analyses proved that the isolated protein was mostly metal-free. For a typical metal to protein ratio determination, measured as described in Materials and Methods, we obtained around 2.9 ppb Cu from a protein solution of 0.15 µM, which corresponded to a final metal to protein ratio of less than 0.3 Cu atoms per *GmCCS* protein. MALDI-TOF analysis identified both bands observed as a doublet in the SDS-PAGE as *GmCCS* chaperone, confirming the western blot results. MALDI-TOFF analysis of the upper SDS-PAGE band resulted in a protein with a molecular weight of 32687 Da, which corresponded very well with that from the deduced protein amino acid sequence, and a pI of 5.55. The identification of that band as a *GmCCS* was very clear with 47% sequence coverage, score 325, and p<0.05. Moreover, N-terminal sequence

analysis revealed that the observed molecular weight difference between both SDS-PAGE bands was due to the loss of the first 40-50 amino acid residues of the N-terminal sequence (Fig. 4). This smaller band was not due to a non-specific cleavage of the Thrombin protease since that band already appeared in the material from the first IMAC column and before the protease treatment (Fig. 3, lane 2). Moreover, immunoblots using the Ab against the *GmCCS* detected the presence of that band in the bacterial crude extract (not shown). We suggest endogenous *E. coli* cell proteases cleaved the recombinant protein in three different close-by positions located at the end of the signal peptide of *GmCCS*.

Characterization of protein metal binding

All experiments were performed under anaerobic conditions and the spectra were immediately monitored after mixing the reactants. The metal-protein K_d s were estimated using the double reciprocal plot representation. The difference electronic absorption spectra for a Cu^+ titration of *GmCCS* are shown in Fig. 5a. Addition of Cu^+ ions to the protein solution (3 μM final protein concentration) lead to an absorption increase in the 220-320 nm spectral region, with a maximum at around 240 nm. This is characteristic of Cu^+ -thiolate charge-transfer transitions (see Materials and Methods for more details). The plot of $1/\text{Abs}_{240}$ versus $1/[\text{Cu}]$ could be fit by two straight lines (Fig. 5b), whose slopes were $K_{d1} \sim 0.18 \mu\text{M}$ and $K_{d2} \sim 6.08 \mu\text{M}$. These results suggest the presence of two Cu-binding sites with different metal affinities.

Addition of Zn^{2+} ions to a *GmCCS* protein solution (5 μM final protein concentration) in the presence of a mag-fura-2 probe, lead to an increase of the absorbance at around 322 nm and a decrease over 340 nm with a well defined isosbestic point at around 340 nm due to the formation of the Zn^{2+} -*GmCCS* complex (Fig. 6a).

The double reciprocal plot representation of the $\Delta\text{Abs}_{322-400\text{nm}}$ (see Materials and Methods for more details) as a function of the Zn^{2+} concentration resulted in a single straight line, whose slope corresponded to a $K_d \sim 9.15 \mu\text{M}$ (Fig. 6b). This suggests the presence of only one Zn^{2+} -binding site in the *GmCCS* chaperone. Furthermore, the presence of a well defined isosbestic point is also a clear indication of a single type of Zn^{2+} -*GmCCS* complex formation in agreement with the double reciprocal plot calculations.

This set of experiments might suggest *GmCCS* binds two Cu^+ (high and low metal affinity) and one Zn^{2+} per protein. As a working hypothesis, we may think that the low affinity copper corresponds to the one that is transferred to the target, CuZnSOD, located at the $\text{C}_{284}\text{XC}_{286}$ metal-binding motif of the chaperone domain III. The high affinity should correspond to the $\text{M}_{80}\text{XC}_{82}\text{XXC}_{85}$, metal-binding site of domain I. This makes sense since domain I of the chaperone is an ATX1-like polypeptide and has a strong affinity for Cu ions. On the other hand, the data also suggest the presence of only one Zn^{2+} -binding site, which is localized most probably within the domain II of *GmCCS* as is the case for yeast and human. The role of the Zn^{2+} ion should be structural. Recent site-directed mutagenesis studies within domain I and III have suggested that one Cu^+ ion is bound to domain I in a three-coordinate geometry composed of two cysteines and another ligand, while a second Cu^+ ion forms an intermolecular cluster between domain III of two different *hCCS* monomers. Apo-*hCCS* is a dimer, with interactions occurring through domain II. But in the presence of Cu^+ , the protein rearranges into a new dimeric species *via* cluster formation at the domain III interface forming a holodimer. This would leave the domain II dimer interface accessible for complex formation with the incoming CuZnSOD dimer target protein (Stasser et al. 2005, 2007).

Protein structure and modelling

Fluorescence emission spectra

After excitation at 290 nm, control *GmCCS* protein sample showed a fluorescence emission spectrum with maxima at 322 and 338 nm due to the intrinsic Trp (Fig. 7, 0 μM Cu^+). A great deal of information has been published relating the Stokes shift of the protein intrinsic Trp fluorescence to the environment surrounding the fluorochrome (Lakowicz 2006). In general, it can be said that a more polar environment and a more exposition of the fluorochrome to the solvent result in a stronger Stokes shift. Smaller Stokes shifts are normally observed when Trp is located in a more hydrophobic environment as normally occurs when it is buried inside of the protein matrix. Considering there are only two Trps in the *GmCCS* sequence (see Fig. 4) and their location in the homology model depicted herein, we ascribe the peak at 322 nm to Trp₁₈₈ in the domain II and the one at 338 nm to Trp₂₉₂ in the domain III. Low copper concentration addition causes a slight decrease of the 322 nm peak and a slight increase of the main peak at 338 nm (Fig. 7, 2.1 μM Cu^+). But further additions induced a strong quenching of the fluorescence emission (Fig. 7, 10.5 and 14.7 μM Cu^+). Although it is difficult to interpret the slight fluorescence increase at lower Cu concentrations, the decrease of the 322 nm peak and the strong fluorescence quenching at higher Cu^+ concentrations is most probably due to an increase of the medium dielectric constant affecting the dipole moment of Trp microenvironments. Moreover, the addition of Cu^+ also induces a red shift of the 338 nm peak, which is in agreement with an increase of polarity surrounding the fluorochrome.

Circular dichroism (CD) spectra

The secondary structure of *GmCCS* was monitored by far-UV CD in the 200-250 nm spectral range (Fig. 8). The CD spectra exhibited negative signals at 208-220 nm

spectral region characteristic of ordered β -sheet and α -helix protein secondary structure conformations. The fitting of the control CD spectrum (see Materials and Methods) indicated 7% α -helical (a.h.), 51% β -sheet (b.s.), whereas the rest of the structure (i.e., 42%) appears to be composed of turns, loops and random coil (t.l.rc.). Although these calculations are not very accurate, the results let us to conclude that the structure of the chaperone is not very compact. This flexibility would, in principle, favour the adaptation of the chaperone to its target, CuZnSOD, to deliver the Cu^+ ion to specific site. The addition of Cu^+ ions to the protein solution had no major effects on the *GmCCS* secondary structure (Fig. 8). This is consistent with the idea that structural modifications may occur after addition of Cu^+ at specific sites without affecting the overall secondary structure of the protein as indicated by CD spectroscopy.

Alignment and modelling

An alignment of the deduced primary structure of several CCSs and the obtained predicted structure of the *GmCCS* using the *yCCS* as a template are shown in Fig. 9a and 9b, respectively. The three main structural domains (I, II, III) might be identified in the resultant model. The calculated percentage of the model secondary structure confirmed the presence of high content (over 40%) of turns, loops and random coil, and very small content (less than 10%) of α -helix structure. The similarity of protein secondary structures obtained by CD and modelling indicates that the purified recombinant protein described in this work maintains its overall native structure. Note that our protein solution is a mixture of protein with and without signal peptide and that the close similarity of the secondary structure obtained from CD and modelling data indicates that the presence of the signal peptide does not significantly affect the overall protein structure.

CONCLUSIONS

Cloning, overexpression and purification of native, complete CCS chaperone from soybean was accomplished successfully. Sufficient amounts of pure protein were obtained, which allowed basic spectroscopic characterization and metal-chaperone complex *K_d* estimations. This allowed us to suggest putative assignments of the protein metal-binding sites and their potential functionality. Since available information on plant CCS is still very scarce, the isolation and characterization of native as in this work and future mutated recombinant CCS protein studies from a plant species will improve and further extend our knowledge of this interesting metal chaperone family.

KNOWLEDGEMENTS

This work was supported by Grants from the PN I+D+i of Spain (BFU2005-04722-C02-01 and AGL2008-00377) and the Aragón Government (DGA-GC E33). We would like to thank to Dr. Michael Seibert for editing the manuscript. S. S. was a predoctoral fellowship of the Ramón Areces Foundation (Madrid, Spain); M. B. was a predoctoral fellowship of the CSIC (I3P Program, financed by the European Social Fund); and S. F. was a predoctoral fellowship of the FPI Program (PN I+D+i, Spain).

REFERENCES

Alfonso M, Pueyo JJ, Gaddour K, Etienne A-L, Kirilovsky D, Picorel R (1996) Induced new mutation of D1 serine-268 in soybean photosynthetic cell cultures produced atrazine resistance, increased stability of $S_2Q_B^-$ and $S_3Q_B^-$ states, and increased sensitivity to light stress. *Plant Physiol* 112: 1499-1508

Arnesano F, Banci L, Bertini I, Ciofi-Baffoni S, Molteni E, Huffman DL, O'Halloran TV (2002) Metallochaperones and metal-transporting ATPases: A comparative analysis of sequences and structures. *Genome Res* 12: 255-271

Banci L, Bertini I, Cramaro F, del Conte R, Viezzoli MS (2003) Solution structure of Apo Cu,Zn superoxide dismutase: role of metal ions in prn folding. *Biochemistry* 42: 9543-9553

Bernal M, Ramiro MV, Cases R, Picorel R, Yruela I (2006) Excess copper effect on growth, chloroplast ultrastructure, oxygen-evolution activity and chlorophyll fluorescence in *Glycine max* cell suspensions. *Physiol Plantarum* 127: 312-325

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of dye binding. *Anal Biochem* 72: 248-259

Casareno RLB, Waggoner D, Gitlin JD (1998) The copper chaperone CCS directly interacts with copper/zinc superoxide dismutase. *J Biol Chem* 273: 23265-23628

Culotta VC, Klomp LW, Strain J, Casareno RL, Krems B, Gitlin JD (1997) The copper chaperone for superoxide dismutase. *J Biol Chem* 272: 23469-23472

Djinovic K, Gatti G, Coda A, Antolini L, Pelosi G, Desideri A, Falconi M, Marmocchi F, Rotilio G, Bolognesi M (1992) Crystal structure of yeast Cu,Zn superoxide dismutase. Crystallographic refinement at 2.5 Å resolution. *J Mol Biol* 225: 791-809

Eisses JF, Stasser JP, Ralle M, Kaplan JH, Blackburn NJ (2000) Domains I and III of the human copper chaperone for superoxide dismutase interact *via* a cysteine-bridged dicopper(I) cluster. *Biochemistry* 39: 7337-7342

Eswar N, Marti-Renom MA, Webb B, Madhusudham MS, Eramian D, Shen M, Pieper U, Sali A (2006) Comparative Protein Structure Modelling with MODELLER, in:

Current Protocols in Bioinformatics, Jonh Wiley & Sons, Inc., Supplement 15: 5.6.1-5.6.30

Hall LT, Sánchez RJ, Holloway SP, Zhu H, Stine JE, Lyons TJ, Demeler B, Schirf V, Hansen JC, Nersissian AM, Valentine JS, Hart PJ (2000) X-ray crystallographic and ultracentrifugation analyses of truncated and full-length copper chaperones for SOD (LYS7). A dimer-dimer model of LYS7-SOD association and copper delivery. *Biochemistry* 39: 3611-3623

Horecka J, Kinsey PT, Sprague GF (1995) Cloning and characterization of the *Saccharomyces cerevisiae* LYS7 gen evidence for function outside of lysine biosynthesis. *Gene* 162: 87-92

Lakowicz J R (2006) Principles of Fluorescence Spectroscopy, third ed., Plenum Press, New York, NY USA

Lamb AL, Wernimont AK, Pufahl RA, Culotta VC, O'Halloran TV, Rosenzweig AC (1999) Crystal structure of the copper chaperone for superoxide dismutase. *Nat Struct Biol* 6: 724-729

Lamb AL, Torres AS, O'Halloran TV, Rosenzweig AC (2000) Heterodimer formation between superoxide dismutase and its copper chaperone. *Biochemistry* 39: 14720-14727

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

Lin SJ, Culotta VC (1995) The ATX1 gene of *Saccharomyces cerevisiae* encodes a small metal homeostasis factor that protects cells against reactive oxygen toxicity. *Proc Natl Acad Sci USA* 92: 3784-3788

O'Halloran TV, Culotta VC (2000) Metallochaperones, an intracellular shuttle service for metal ions. *J Biol Chem* 275: 25057-25060

Pryor KD, Leiting B (1997) High-level expression of soluble protein in *Escherichia coli* using a his₆-tag and maltose-binding-protein double-affinity fusion system. *Protein Express and Purif* 10: 309-319

Pufahl RA, Singer CP, Peariso KL, Lin SJ, Schmidt PJ, Culotta VC, Penner-Hahn JE, O'Halloran TV (1997) Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science* 278: 853-856

Rae TD, Torres AS, Pufahl RA, O'Halloran TV (2001) Mechanism of Cu,Zn-superoxide dismutase activation by the human metallochaperone hCCS. *J Biol Chem* 276: 5166-5176

Rogers SMD, Ogren WL, Widholm JM (1987) Photosynthetic characteristics of a photoautotrophic cell suspension culture of soybean. *Plant Physiol* 84: 1451-1456

Ruzsa SM, Scandalios JG (2003) Altered Cu metabolism and differential transcription of Cu/ZnSod genes in a Cu/ZnSOD-deficient mutant of maize: Evidence for a Cu-responsive transcription factor. *Biochemistry* 42: 1508-1516

Schmidt PJ, Rae TD, Pufahl RA, Hamma T, Strain J, O'Halloran TV, Culotta VC (1999a) Multiple protein domains contribute to the action of the copper chaperone for superoxide dismutase. *J Biol Chem* 274: 23719-23725

Schmidt PJ, Ramos-Gómez M, Culotta VC (1999b) A gain of superoxide dismutase (SOD) activity obtained with CCS, the copper metallochaperone for SOD1. *J Biol Chem* 274: 36952-36956

Stasser JP, Eisses JF, Barry AN, Kaplan JH, Blackburn NJ (2005) Cysteine to serine mutants of the human copper chaperone for superoxide dismutase reveal a domain III dimmer interface. *Biochemistry* 44: 3143-3152

Stasser JP, Siluvai GS, Barry AN, Blackburn NJ (2007) A multinuclear copper (I) cluster forms the dimerization interface in copper-loaded human copper chaperone for superoxide dismutase. *Biochemistry* 46: 11845-11856

Thompson JD, Higgins DJ, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22: 4673-4680

Trindade LM, Horvath BM, Bergervoet MJE, Visser RGF (2003) Isolation of a gene encoding a copper chaperone for the copper/zinc superoxide dismutase and characterization of its promoter in potato. *Plant Physiol* 133: 618-629

Whitmore L, Wallace BA (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. *Nucleic Acids Res* 32: W668-W673

Whitmore L, Wallace BS (2008) Protein secondary structure analyses from circular dichroism spectroscopy: Methods and reference databases. *Biopolymers* 89: 392-400

Wintz H, Vulpe C (2002) Plant copper chaperones. *Biochem Soc Trans* 30: 732-735

Zhu H, Shipp E, Sanchez RJ, Liba A, Stine JE, Hart PJ, Gralla EB, Nersissian AM, Valentine JS (2000) Cobalt(2+) binding to human and tomato copper chaperone for superoxide dismutase: Implications for the metal ion transfer mechanism. *Biochemistry* 39: 5413-5421

FIGURE LEGENDS

Figure 1. *GmCCS* gene cloning strategy. (a) Cloning and Thrombin sites in a modified pMALc2X vector. (b) Construction of pMALc2Xmodif-*GmCCS*.

Figure 2: (a) SDS-PAGE (12%, w/v) analysis of the *GmCCS* expression in *E. coli*. The gel was stained with Coomassie Blue. M, molecular weight markers; 1, control cell extract (without induction); 2, cell extract with induction at 37 °C for 3 h with 0.3 mM IPTG; 3, cell extract with induction at 37 °C for 3 h with 0.3 mM IPTG, 0.5 mM Cu²⁺, and 0.5 mM Zn²⁺. (b) Western blot with anti-His₆ monoclonal antibody of cell extract of lane 3.

Figure 3. (a) SDS-PAGE (15%, w/v) protein analysis of the *GmCCS* purification process. Gels were stained with Coomassie Blue. M, molecular weight standards; lane 1, cell extract after induction for 3 h with 0.3 mM IPTG, 0.5 mM CuSO₄, and 0.5 mM ZnSO₄; lane 2, protein profile after the first Hitrap Ni²⁺-IMAC column; lane 3, protein profile after treatment with the protease Thrombin; lane 4, protein profile after a second Hitrap Ni²⁺-IMAC column; lane 5, purified protein *GmCCS* after pH gradient chromatography on a Toyopearl TSK DEAE 650s column. (b) Western blot of lane 3 of 3a with the anti-*GmCCS* antibody.

Figure 4. Heterogeneous processing of *GmCCS* at the N-terminal. The chloroplast targeting sequence is underlined. The three different N-terminal sequences (6 amino acid residues) obtained in the analysis are shown at the top. Arrows indicate their localization in the amino acid sequence.

Figure 5. Cu⁺ binding to *GmCCS*. (a) Difference absorption spectra in the 220-320 nm spectral range of a *GmCCS* solution (3 μM final protein concentration) in the presence of increasing amounts of Cu⁺ (see Materials and Methods for more details). The arrow indicates the wavelength, 240 nm, at which the increase in absorbance was monitored as

a function of the addition of Cu^+ ions. The Cu^+ final concentrations corresponding to each spectrum were: 0.4, 0.7, 5.6, 8.4, 14.1, 19.9, and 34.1 μM . (b) Double reciprocal plot representation of $\text{Abs}_{240\text{nm}}$ versus Cu^+ concentration to estimate the dissociation constants, K_{d1} and K_{d2} , of Cu^+ -*GmCCS* complex. Data shown are representative of at least two independent replicates.

Figure 6. Zn^{2+} binding to the recombinant *GmCCS*. (a) Electronic absorption spectra in the 300-400 nm spectral range of a *GmCCS* solution (5 μM final protein concentration) with a mag-fura-2 probe mixture containing no addition (control) and with addition of increasing concentrations of Zn^{2+} . The final Zn^{2+} concentrations corresponding to each spectrum were: 0, 0.6, 1.1, 1.7, 2.3, 4.6, 5.6, and 11.3 μM . (b) Double reciprocal plot representation of $\Delta\text{Abs}_{322-400\text{nm}}$ (i.e., absorption at 322nm–400nm at each Zn^{2+} concentration minus absorption at 322nm–400nm of the control) versus Zn^{2+} concentration to estimate the K_d of the Zn^{2+} -*GmCCS* complex. Data shown are representative of at least two independent replicates.

Figure 7. Intrinsic Trp fluorescence emission spectra in the 300-400 nm spectral range of purified recombinant *GmCCS* chaperone (2 μM final protein concentration) after excitation at 290 nm. The final Cu^+ concentration in solution for each spectrum is indicated in the figure.

Figure 8. Far-UV CD spectra in the 200-250 nm spectral range of our purified recombinant *GmCCS* protein with increasing concentrations of Cu^+ in the samples. Inset: calculated percentage of the protein secondary structures (see Materials and Methods).

Figure 9. (a) Alignment of several known CCS chaperone sequences. The putative signal peptide of *GmCCS* is boxed, and the Cys, Trp, and Tyr residues are in red. The two metal-binding motifs are indicated in domain I and III. (b) Modelling of *GmCCS*

chaperone using the yCCS (pdb 1jk9) as a template. Putative Cu⁺-binding cysteines (Cys₈₂ and Cys₈₅) and (Cys₂₈₄ and Cys₂₈₆) are indicated as well as tryptophans (Trp₁₈₈ and Trp₂₉₂) and tyrosines (Tyr₂₃₅, Tyr₂₅₅, and Tyr₂₈₀). The possible three protein.

Fig. 1

Fig. 2

LATPPS...
 TPPSAL...
 LVKTLA...

↓ ↓ ↓

MAFLRSIATTAIATIPAALAFSSSSSSSFPRSSQSPNPQNRLGLVKTLATPPSALHMDHKLSSQPDA
 VLPELLTEFMVDMKCEGCVNAVKNKLNEINGVKNVVVDLSNQVVRLGSTPVKTMTEALEQTG
 RKARLIGQGVPEDFLISA AVSEFKGPDIFGVVRLAQVNMELARIEANFGGLSPGKHGWSINEFGD
 LTRGAASTGKMFPVNEENSKEPLGDLGTLEANEKGEAFYSGVKEKLRVADLIGRSVVVYATE
 DKSEHGITA AVIARSAGVGENYKKLCTCDGTTIWEATDTDFVTSKV

Fig. 3

Fig. 4

Fig. 5

(a)	GmCCS	1	-MAFLRSIAT-----TAIATIPAAALAFSSS--SSSS--FPRSSQ-----SP-NPQNRLGLVKTLA--T	50
	AtCCS	1	MASILRSVATTSAVVAAASAIPIAIAFSSSS--SSSS--TNPKSQSLNFSFLSRSSPR-LLGLSRSFVSSP	66
	StCCS	1	-MAFFRSIVTA--KTTAIAAAIPAAAFSAASSISSSSSQFERPSKNIKFGSISSSNPIFQLSFAKNLQKTS	67
	LeCCS	1	--AFLRSIVTA--KTTAIAAAIPAAAFVSSISSSS--QFERPLKNLKFSGSISSSNSILQLSFAKNLQKKS	65
	OsCCS	1	MVGFLRALTAAS--AVPAAAVAAVALSTNSSSSSRLRLPSPASLPSLSAYAAAPASGSARKPNVPPM	68
	yCCS	1	-----	1
			MxCxxC	
	GmCCS	51	PPSALHMD-----HKLSSQPDVPIPELLTEFMVDMKCEGCVNAVKNKLNEINGVKNVEVDLSNQVVRILG	115
	AtCCS	67	MATALTS-----RNLHQEDRAMPELLTEFMVDMTCCEGCVNAVKNKLETIEGIEKVEVDLSNQVVRILG	130
	StCCS	68	PPSALHMETPSSNHQTSSDNEVVLPELLTEFMVDMSCQGCVCNAVKSCLKQTVEGVKNVDVDLDNQVVRILG	137
	LeCCS	66	PPSALHMETHSSNHQTSSDNGVVLPELLTEFMVDMSCQGCVCNAVKSCLKQTVEGVKNVDVDLDNQVVRILG	135
	OsCCS	69	AAAAATAD-----LSAAADKGAALPELMTEFMVDMKCDGCVTAVKNKFQTLLEGIKNIEVDLNNQVVRVLG	133
	yCCS	1	---MTTN-----DTYEATYAIPMHCENCVNDIKACLKNVPGINSLNFDIEQQIMSVES	50
	GmCCS	116	STPVKTMTEALEQTGRKARLIGQGVPE-----DFLISAADVSEFKGPDIFGVVRLAQVNMELARIEANF	178
	AtCCS	131	SSPVKAMTQALEQTGRKARLIGQGVPE-----DFLVSAAVAEFKGPDI FG VVRF AQVSMELARIEANF	193
	StCCS	138	SSPVKTMTEALEQTGRKARLIGQGVPE-----DFLISAADVAEFKGPDI FG VVRL AQVNMELTRIEANF	200
	LeCCS	136	SSPVKTMTEALEQTGRKARLIGQGVPE-----DFLISAADVAEFKGPDI FG VVRL AQVNMELTRIEANF	198
	OsCCS	134	SLPVNTMLDTLHQTRDARLIGQGNPN-----DFLVSAAVAEFKGPVIFGVVRLAQVNMELAIVEATF	196
	yCCS	51	SVAPSTIINTLRNCKDAIIRGAGKPNSSAVAILETFQKYITIDQKKDTAVRGLARIVQGENKTLFDITV	120
	GmCCS	179	SGLSPG-KHGWTSINEFGDLTRGAASTGKMFNPVNEENSKEPLGDLGTLEANKEGEAFYSGVKEKLRVADL	247
	AtCCS	194	TGLSPG-THSWCINEYGDLTNGAASTGSLYNPFQDQTGTPLGDLGTLEADKNGEAFYSGKKEKLVADL	262
	StCCS	201	SGLSPG-KHAWTSINEFGDLTRGAASTGKLYS-----PPLGDLVTLEVDEKGEAFYTGPKKEKVRVADL	261
	LeCCS	199	SGLSPG-KHAWTSINEFGDLTRGAASTGKLYS-----LPLGDLGTLVDVEKGEAFYSGPKEKLRVADL	259
	OsCCS	197	SGLSPG-KHGWTSINEFGDLTRGAESTGKVYNP-SDYRSNKPLGDLGTLEAGEKGEAQFSASKEKLVVDL	264
	yCCS	121	NGVPEAGNTHASIEHKGDVSKGVESTGKVTHHKFD-----EPIECFNESDLGKNLYSGKTFLSAPLPTQL	185
	GmCCS	248	IGRSVVVYATED--KSEHG-----ITAAVIARSAGVGGENYKKICTCDGTTIWEATD TDFVTSKV	304
	AtCCS	263	IGRAVVVYKTDDN-KSGPG-----LTAAVIARSAGVGGENYKKICTSCDGTVIWEATNSDFVASKV	320
	StCCS	262	IGRAIAVYATED--KTDPG-----LTAAVIARSAGVGGENYKKICTACDGTTIWEATN-----KL	312
	LeCCS	260	IGRAIAVYATED--KSDPG-----LTAAVIARSAGVGGENYKKICTCDGTTIWEATS-----KI	310
	OsCCS	265	IGRSIALYATED--RSDPG-----IAAAVIARSAGVGGENYKKICTCDGVTTIWESS-----	312
	yCCS	186	IGRSFVISKSLNHPENEPSSVKDYVFLGVIAARSAGVWENNKQVCACTGKTVWEERKDALANNIK	249
			CxC

Fig. 6