Solution structure of the N-terminal domain of a potential copper-translocating P-type ATPase from Bacillus subtilis in the apo and Cu(I) loaded states

ARTICLE *in* JOURNAL OF MOLECULAR BIOLOGY · APRIL 2002

Impact Factor: 4.33 · DOI: 10.1006/jmbi.2002.5430 · Source: PubMed

7 AUTHORS, INCLUDING:

Simone Ciofi
University of Florence
71 PUBLICATIONS **2,704** CITATIONS
SEE PROFILE

Mariapina D'Onofrio
University of Verona
33 PUBLICATIONS **509** CITATIONS
SEE PROFILE

Frutos Carlos Marhuenda Egea
University of Alicante
50 PUBLICATIONS **750** CITATIONS
SEE PROFILE

Francisco J Ruiz-Dueñas
Centro de Investigaciones Biológicas
83 PUBLICATIONS **3,382** CITATIONS
SEE PROFILE

doi:10.1006/jmbi.2001.5430 available online at http://www.idealibrary.com on IDEAL® *J. Mol. Biol.* (2002) **317**, 415–429

JMB

Solution Structure of the N-terminal Domain of a Potential Copper-translocating P-type ATPase from *Bacillus subtilis* in the apo and Cu(I) Loaded States

Lucia Banci[1], Ivano Bertini[1]*, Simone Ciofi-Baffoni[1] Mariapina D'Onofrio[1], Leonardo Gonnelli[1], Frutos Carlos Marhuenda-Egea[2] and Francisco Javier Ruiz-Dueñas[3]

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

[2]*Department of Agrochemistry and Biochemistry, University of Alicante, 03080 Alicante Spain*

[3]*Department of Molecular Microbiology, Centro de Investigaciones Biológicas CSIC, Velázquez 144 E-28006 Madrid, Spain*

**Corresponding author*

A putative partner of the already characterized CopZ from *Bacillus subtilis* was found, both proteins being encoded by genes located in the same operon. This new protein is highly homologous to eukaryotic and prokaryotic P-type ATPases such as CopA, Ccc2 and Menkes proteins. The N-terminal region of this protein contains two soluble domains constituted by amino acid residues 1 to 72 and 73 to 147, respectively, which were expressed both separately and together. In both cases only the 73-147 domain is folded and is stable both in the copper(I)-free and in the copper(I)-bound forms. The folded and unfolded state is monitored through the chemical shift dispersion of ^{15}N-HSQC spectra. In the absence of any structural characterization of CopA-type proteins, we determined the structure of the 73-147 domain in the 1-151 construct in the apo state through ^{1}H, ^{15}N and ^{13}C NMR spectroscopies. The structure of the Cu(I)-loaded 73-147 domain has been also determined in the construct 73-151. About 1300 meaningful NOEs and 90 dihedral angles were used to obtain structures at high resolution both for the Cu(I)-bound and the Cu(I)-free states (backbone RMSD to the mean 0.35(±0.06) Å and 0.39(±0.07) Å, respectively). The structural assessment shows that the structures are accurate. The protein has the typical βαββαβ folding with a cysteine in the C-terminal part of helix α1 and the other cysteine in loop 1. The structures are similar to other proteins involved in copper homeostasis. Particularly, between BsCopA and BsCopZ, only the charges located around loop 1 are reversed for BsCopA and BsCopZ, thus suggesting that the two proteins could interact one with the other. The variability in conformation displayed by the N-terminal cysteine of the CXXC motif in a number of structures of copper transporting proteins suggests that this may be the cysteine which binds first to the copper(I) carried by the partner protein.

Keywords: CopA; NMR; folding; P-type ATPase; copper transporting protein

Introduction

Copper is an essential trace element required for a number of enzymes, including cytochrome oxidase, superoxide dismutase, and lysyl oxidase.[1] While trace amounts of copper are needed to sustain life, excess copper is extremely toxic, since the free metal ion can facilitate oxidative damage through generation of free radicals. It is likely that most, if not all, cells have mechanisms for copper homeostasis.[2–5] Furthermore, copper is present in

Abbreviations used: HSQC, heteronuclear single quantum coherence; RMSD, root-mean-square deviation; NOESY, nuclear Overhauser effect spectroscopy; TOCSY, total correlation spectroscopy; NOE, nuclear Overhauser effect; TPPI, time-proportional phase incrementation; CSI, chemical shift index; REM, restrain energy minimization; CR, cytoplasmic region; LB, Luria-Bertani; DTT, dithiothreitol; Tris, tris(hydroxymethyl)aminomethane; IPTG, isopropyl-β-D-thiogalactopyranoside.

E-mail address of the corresponding author: bertini@cerm.unifi.it

0022-2836/02/030415–15 $35.00/0

the cytoplasm at very low concentration[6] and therefore the copper-dependent proteins need systems which ''enhance'' the binding of copper to the protein. In eukaryotes, copper enters into the cell by specific uptake systems and is transferred to cytosolic chaperone proteins that eventually deliver copper to the copper-dependent proteins, leaving little free cytosolic copper ion.[6] In bacteria, copper uptake into the cell and its transport to the specific proteins is much less characterized.[7] The best studied copper transport systems in bacteria are the copper ATPase proteins CopA and CopB from *Enterococcus hirae*.[8] On the basis of gene-disrupted strain experiments, it has been suggested that CopB is a Cu(I)/Ag(I) efflux pump, while CopA is thought to be involved in copper uptake.[8] Close homologues to CopA have also been found in *Helicobacter pylori*,[9,10] *Escherichia coli*,[11] and in *Synechococcus* (CtaA and PacS).[12,13] It has been recently proved that CopA in *E. coli* catalyses copper extrusion.[11] Generally, it has been proposed that, in order to maintain low intracellular concentrations of copper, specific pumps remove excess copper ion from the cytosol, either by accumulation into organelles or extrusion out of the cell.[1]

In a wide range of organisms P-type ATPases are involved in metal homeostasis. They have been classified into five groups according to substrate specificity.[1] One of these groups contains the phylogenetically related P-type ATPases that catalyse transport of transition or heavy-metal ions, previously described as soft-metal transporting P-type ATPases or CPx-type ATPases.[14] These proteins have eight transmembrane domains[15] in contrast to the hard-metal P-type ATPases, which have ten transmembrane segments. Soft metal-transporting P-type ATPases can be further divided into subgroups that contain Cu(I)/Ag(I)-translocating ATPases and Zn(II)/Cd(II)/Pb(II)-translocating ATPases.[7] Members of the first subgroup most likely transport monovalent soft-metal cations, while members of the second subgroup are divalent soft-metal-cation pumps.

Within the frame of a research aiming at the characterization of P-type ATPases involved in copper homeostasis, in the goal of understanding the mechanism of copper trafficking in bacteria, we have selected to browse the genome of *Bacillus subtilis*. The search was carried out using as template the sequence of *Saccharomyces cerevisiae* Ccc2, a protein involved in copper transport in yeast cells as an example of eukaryotic organisms.[16] We found a gene, which we call *bscopa* hereafter, encoding a protein, whose sequence is significantly homologous to that of the Ccc2 protein and contains two N-terminal conserved MXCXXC motifs also identified in the two metal-binding domains of Ccc2. It was then decided to pursue the expression and purification of the soluble N-terminal region of this protein. Three different polypeptides have been produced and characterized by NMR in order to investigate the correlation between fold and structure with the length of the construct and the number of domains. The structures of the apo form of the BsCopA(73-147) domain in the 1-151 construct and of the Cu(I) form of the same domain but in the construct 73-151 have been solved by NMR. An instructive comparison with analogous proteins is also performed.

Results

The amino acid sequence of yeast Ccc2 was used in the search of sequences encoding equivalent proteins in the genome of *B. subtilis*, a Gram-positive bacteria. In particular, sequences were looked for with approximately the same length and with the presence of one or more MXCXXC metal-binding motifs. The sequence *yvgX* in the *B. subtilis* genome was selected. It encodes a protein which contains two MXCXXC metal-binding motifs in the N-terminal region and which we then called BsCopA. The corresponding *yvgX* gene was called *bscopa*. The hydropathy profiles predicted eight transmembrane segments. This model is in agreement with that obtained for yeast Ccc2 and also with the prediction for other bacterial P-type ATPases of *E. hirae*,[8] *H. pylori* and *H. felis*[9] and *E. coli*.[11] The sequence alignment of the N-terminal soluble segment of BsCopA with the corresponding segment in Ccc2 protein reveals 30 % identity, which increases to 53 % when conservative substitutions are considered.

The gene *bscopa* is immediately followed, in the genome of *B. subtilis*,[17] by a sequence encoding a protein (hereafter BsCopZ), whose NMR structure has been recently solved for the copper-bound form.[18] This protein is homologous to CopZ from *E. hirae* for which the solution structure of the apo form is available.[19] Since the genes which form the *cop* operon in other organisms where they have been described, such as in *H. pylori* and *H. felis*,[9] are organized in a similar way, BsCopA and BsCopZ might constitute the *cop* operon of *B. subtilis*. To get a better understanding of the pathways for copper transport in *B. subtilis*, we decided to express the soluble domains of BsCopA. Furthermore, in order to investigate the possible relations between the number of domains present and their structures, the following BsCopA constructs were prepared for the present study: BsCopA(1-72), BsCopA(73-147) and BsCopA(1-147). The same four-amino acid segment IEGR, corresponding to the restriction protease factor Xa recognition site used to remove the histidine tag, was engineered at the C terminus of all the constructs. Therefore, the lengths of the expressed constructs are 76, 80 and 151, respectively. For the shorter constructs, unlabelled and ^{15}N labelled samples were prepared, while for the two-domain protein also a $^{15}N/^{13}C$-labelled sample was produced.

The ^{1}H and the ^{15}N NMR spectra of the apo-BsCopA(1-76) polypeptide show a small dispersion of the signals, indicating that this polypeptide chain, predicted to form a soluble domain, is in an

unfolded state. Furthermore, the NMR lines are broader than expected for its molecular weight, thus suggesting that the protein might also be in an aggregate state. Changes in buffer and DTT concentrations did not produce any improvement in the spectra. Due to the unfolded state of this first domain, the BsCopA(1-151) construct in the apo form containing the two predicted domains, has been produced and characterized by NMR. The ^{15}N-HSQC spectrum of the apo form recorded on the freshly prepared sample shows a good dispersion of the signals indicating that the protein is in a folded state (Figure 1(a)). After a few days, the system evolved towards a state, characterized by some NMR peaks collapsed in the unfolded region of the ^{15}N-HSQC spectrum i.e. with ^{15}N and ^{1}H shifts grouped in the small 120-135 and 7.5-8.5 ppm range, respectively. Still, the majority of resonances are well dispersed and with linewidths consistent with a monomeric state for the protein (Figure 1(b)). This behaviour suggests the occurrence of unfolding for a segment of the protein.

Assignment of the resonances of apo-BsCopA(1-151) started from the analysis of the ^{15}N HSQC map which allowed the identification of the ^{15}N and ^{1}HN resonances. Then through the analysis of the 3D CBCA(CO)NH and CBCANH spectra, the sequence-specific assignment was performed. This was independently confirmed by sequential and medium-range NOEs.[20] Resonances for 115 of the 151 residues were assigned. In the apo-BsCopA(1-151) about 55 % of the proton resonances, 74 % of the ^{15}N resonances and 59 % of the ^{13}C resonances were assigned. These values increased respectively to 93 %, 98 % and 80 % when only residues 72-151 were considered. This result suggests that the first predicted domain collapses towards an unfolded state also when the complete N-terminal region of the ATPase is expressed and that the resonances of this first domain move to shift values typical of unfolded state after a few days. Direct evidence of this interpretation comes from the analysis of the NOESY spectra that shows the lack of both HN backbone resonances and medium and long range NOE cross-peaks. The analysis of the ^{15}N-HSQC map of the protein corresponding to only the second domain, apo-BsCopA(73-151), shows a good signal dispersion and linewidths consistent with a folded protein, present in a monomeric state (Figure 1(c)). The ^{15}N and ^{1}H resonances of the backbone HN are essentially identical to those observed for the second domain in the apo-BsCopA(1-151) construct, but resonances belonging to metal binding region (region 83-92) lose intensity or disappear. This latter difference led us to focus on the structure determination of the apo form of the second domain in the larger 1-151 construct. The second domain of this construct is hereafter called apo-BsCopAb. The differences observed in the ^{15}N-HSQC spectra of the two constructs (1-151, 73-151) indicate that the first domain in construct 1-151, even if it is in an unfolded state, has some

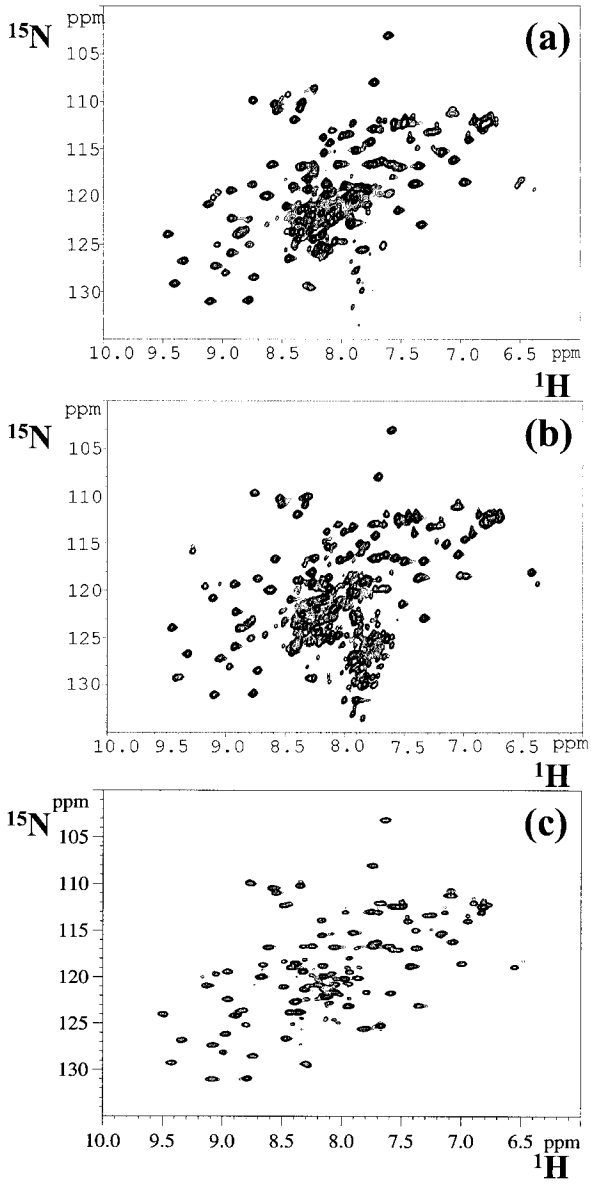

Figure 1. ^{15}N-HSQC NMR spectra at 700 MHz and 298 K of apo-BsCopA(1-151) immediately after preparation of the sample (a), after one week (b) and of apo-BsCopA(73-151) (c). For all the samples the protein concentration was about 1.5 mM, in 100 mM phosphate buffer.

effect on the structural arrangement of the second domain in the metal binding region.

For the two-domain construct, the binding of copper has also been characterized. The atomic absorption data of the Cu(I)-BsCopA(1-151) derivative shows that one copper ion is bound to the two-domain protein (metal/protein ratio 1.15). The ^{15}N-HSQC spectrum of the freshly prepared sample is similar to that of the apo form (Figure 2(a)). However, after some days, the protein evolves towards a completely unfolded state. Indeed, the resonances of the ^{15}N-HSQC spectrum

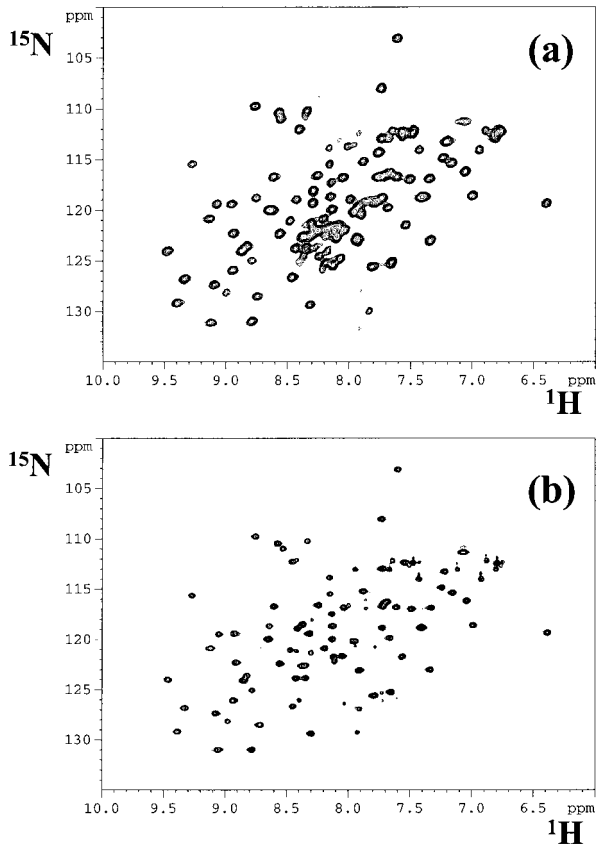

Figure 2. ^{15}N-HSQC NMR spectra at 700 MHz and 298 K of Cu(I)-BsCopA(1-151) immediately after preparation of the sample (a) and of the Cu(I)-BsCopA(73-151) (b). For all the samples the protein concentration was about 1.5 mM, in 100 mM phosphate buffer.

are essentially all grouped in the 7.5-8.5 ppm range, for ^{1}H, and 120-135 ppm, for ^{15}N, thus preventing us from characterizing the NMR structure of the copper-bound form for this construct. Changes in the experimental conditions (i.e. concentration of DTT, temperature and buffer) did not improve the stability of the copper-bound protein. The integrity of the polypeptide chain was also checked through electrophoresis, indicating that interaction with copper does not produce any break. However, when copper is bound to the second domain construct, i.e. apo-BsCopA(73-151), the protein is stable in a folded state for a longer period, as shown by the good dispersion of the signals of the ^{15}N-HSQC spectrum (Figure 2(b)). Therefore, the NMR structure of the Cu(I)-BsCopA(73-151) has also been solved. The atomic absorption data of the latter construct shows that one copper ion is bound to the protein.

NMR Structure of apo-BsCopAb

The pattern of assigned NOEs indicated the presence of a few secondary structure elements. Short and medium range backbone NOEs indicated the presence of two helices which are characterized by a high number of sequential and medium range connectivities such as $d_{NN}(i,i+1)$, $d_{NN}(i,i+2)$, $d_{\alpha N}(i,i+3)$, $d_{\alpha N}(i,i+4)$, $d_{\alpha\beta}$ $(i,i+3)$. The two helices involve residues 87-98 and 125-135. From the analysis of all backbone NOEs, four antiparallel β strands, involving residues 74-80, 102-105, 112-119, 138-144 had been individuated. The pattern of NOEs shows that the typical folding pattern of the copper chaperones, ''open-faced β-sandwich'' fold (β1-α1-β2-β3-α2-β4),[21–23] is present also in this protein.

A total of 1582 NOESY cross-peaks were assigned, mostly integrated in the 3D spectra, and transformed in upper distance limits with the program CALIBA.[24] They corresponded to 1413 unique upper distance limits, of which 1278 were found to be meaningful. The number of NOEs per residue is shown in Figure 3(a). The average number of NOEs per residue is 18 for apo-BsCopAb, of which 16 are meaningful. 51 dihedral ϕ angle constraints and 43 ψ angle constraints were measured and used in the structural calculations. In the course of the structure calculations, a total of 22 proton pairs were stereospecifically assigned through the program GLOMSA.[24] Chemical shift

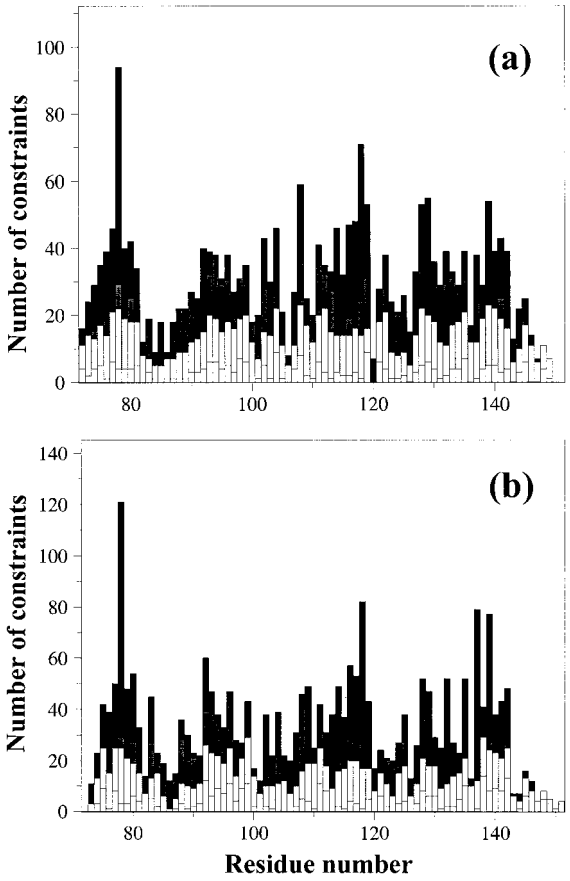

Figure 3. Number of meaningful NOEs per residue for apo-BsCopAb (a) and Cu(I)-BsCopA(73-151) (b). White, gray and black bars indicate intraresidue, sequential and medium/long range connectivities, respectively.

indices (CSI),[25] which is strongly related to the presence of secondary structure elements, were also used in the structure calculations. In the apo-BsCopAb domain, CSI predicts the secondary structure elements in the following regions: 73-82 (β1), 87-97 (α1), 113-119 (β3), 125-133 (α2), 137-140 (β4). The strand β2 is not predicted on the basis of CSI. The CSI results are then converted into constraints on the backbone dihedral angles, using the standard parameters for α helix and β sheet elements.

The 30 conformers constituting the final DYANA family had an average target function of 0.58(±0.09) Å^2 and average RMSD values over residues 73-145 with respect to the mean structure of 0.39(±0.07) Å for the backbone and of 0.92(±0.09) Å for the heavy atoms. The family of conformers was then subjected to further refinement through energy minimization.[26] The REM family has contributions to the average target function of 0.20(±0.03) Å^2 and of 0.04(±0.01) rad^2 from NOEs and the dihedral angle constraints, respectively. The average RMSD values for the family with respect to the mean structure are 0.39(±0.07) Å for the backbone and 0.95(±0.08) Å for the heavy atoms over residues 73-145. The RMSD values per residue of the final REM family to the mean structure are shown in Figure 4(a).

The final family of conformers was analysed with PROCHECK-NMR[27] and results are shown in Table 1. According to this program, the secondary structure elements in the energy minimized mean structure involve residues 74-81 (β1), 86-98 (α1), 101-104 (β2), 112-119 (β3), 124-135 (α2), and 138-142 (β4). Analysis of the NOE patterns has led to somewhat similar conclusions (see above). In the energy minimized average structure, 83.1 % of the residues are in the most favoured regions of the Ramachandran plot, 14.1 % of the residues are in the allowed regions and 2.8 % of residues are in the generously allowed region. No residues are in the disallowed regions (Table 1).

The polypeptide fold of apo-BsCopAb consists of a twisted four-stranded antiparallel β sheet and two α helices that are both located on the same side of the β sheet (Figure 5(a)). Helix α2 and all the antiparallel β strands of apo-BsCopAb are well defined, with an average backbone RMSD lower than 0.35 Å. Helix α1 is less defined than the other secondary structure elements. This increase of RMSD value (Figure 4(a)) is due to the relatively low number of NOEs for the initial part of the helix α1 of the second domain (region 86-89) (Figure 3(a)). The high RMSD values of residues in loop 1, containing MXCXXC metal-binding motif, are also due to the paucity of NOEs (Figure 3(a)). On the contrary, the other loop regions are quite well defined (Figure 5(a)). The number of residues

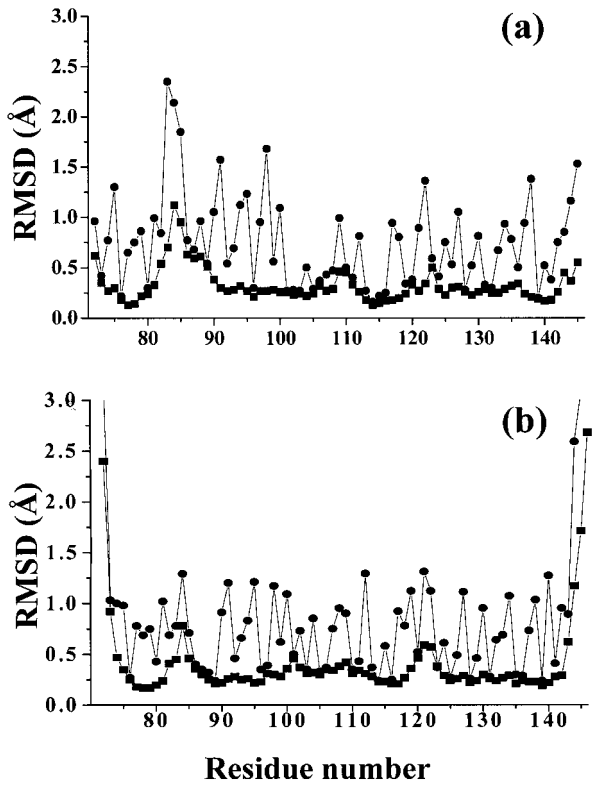

Figure 4. RMSD per residue to the mean structure of apo-BsCopAb (a) and of Cu(I)-BsCopA(73-151) (b) for the backbone (filled squares) and all heavy atoms (filled circles) of the REM structure family of 30 conformers.

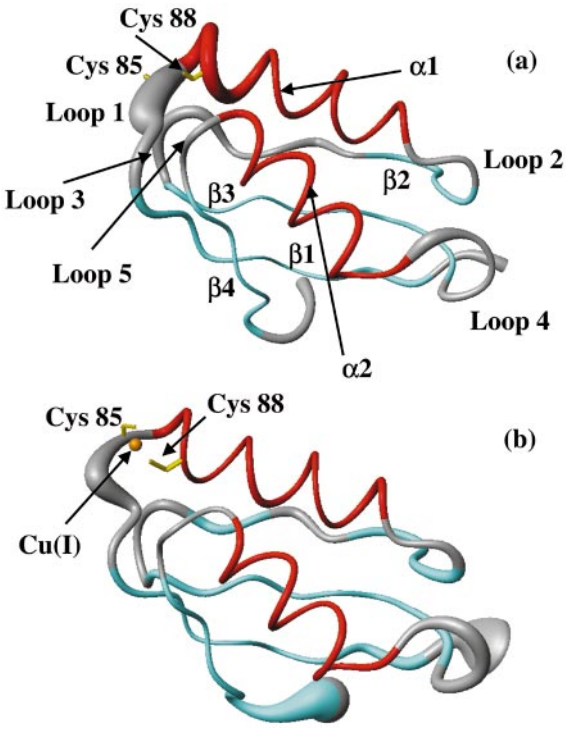

Figure 5. Backbone atoms for residues 72-145 of the solution structures apo-BsCopAb (a) and Cu(I)-BsCopA(73-151) (b) as a tube with variable radius, proportional to the backbone RMSD value of each residue. The side-chains of Cys85, Cys88 and the Cu(I) ion are also shown. The Figure was generated with the program MOLMOL.[51]

Table 1. Statistical analysis of the final REM family and the mean structure of apo-BsCopAb from *B. subtilis*

	REM (30 Structures)	⟨REM⟩
RSM violations per experimental distance constraint (Å)[a]		
Intraresidue (225)	0.0151 ± 0.0026	0.0136
Sequential (412)	0.0082 ± 0.0015	0.0097
Medium range[b] (293)	0.0086 ± 0.0022	0.0074
Long range (348)	0.0114 ± 0.0017	0.0109
Total (1278)	0.0108 ± 0.0010	0.0104
Average number of violations per structure		
Intraresidue	9.4 ± 2.3	10
Sequential	6.6 ± 1.6	6
Medium range[b]	5.0 ± 1.7	5
Long range	9.5 ± 1.9	9
Total	30.5 ± 3.9	30
Average no. of NOE violations larger than 0.3 Å	0.00 ± 0.00	0
Average no. of ϕ, ψ violations larger than 5°	4.1 ± 1.7	5
Largest residual NOE violation (Å)	0.21	0.13
Average NOE deviations (Å^2)	0.20 ± 0.03	0.22
Structural analysis[c]		
% of residues in most favourable regions	79.4	83.1
% of residues in allowed regions	16.5	14.1
% of residues in generously allowed regions	3.2	2.8
% of residues in disallowed regions	0.8	0.0

REM indicates the energy minimized family of 30 structures, ⟨REM⟩ is the energy minimized average structure obtained from the coordinates of the individual REM structures.

[a] The number of experimental constraints for each class is shown in parenthesis.

[b] Medium range distance constraints are those between residues $(i,i+2)$, $(i,i+3)$, $(i,i+4)$ and $(i,i+5)$.

[c] As it results from the Ramachandran plot analysis.

forming the strand β2 is lower than that formed in other metallochaperone structures.[22,28] This might be due to the presence of Pro106, which gives rise to a conformation such as to prevent the following residues to form the β strand. The C-terminal region is disordered as a consequence of the lack of medium and long range NOEs involving residues 147-151, which correspond to the factor Xa recognition site (Figure 3(a)).

The backbone and side-chain of Cys85, which is located in loop 1 and which constitutes a potential copper ligand is disordered and spans different conformations, due to the small number of NOEs. The other possible copper ligand, Cys88, which belongs to helix α1, is better defined (RMSD BB 0.61 Å, RMSD HA 0.96 Å). Still, RMSD values higher than the average indicates some disorder in this region.

NMR structure of Cu(I)-BsCopA(73-151)

Assignments of the resonances of Cu(I)-BsCopA(73-151) protein started from the analysis of the ^{15}N HSQC map. Then through the analysis of the 3D NOESY-HSQC and of 2D NOESY and TOCSY, the sequence-specific assignment was performed. In the copper-bound protein about 97% of the proton resonances could be located in the maps and all the ^{15}N backbone resonances have been assigned, with the exception of the first two residues and Ala86. The pattern of assigned NOEs indicated the presence of the same secondary structure elements observed in the apo form. A total of 2721 NOESY cross-peaks were assigned, mostly integrated in the 2D spectra, and transformed in upper distance limits with the program CALIBA.[24] They corresponded to 1904 unique upper distance limits, of which 1415 are meaningful. The average number of NOEs per residue is 24, of which 18 are meaningful. The number of NOEs per residue is shown in Figure 3(b). 44 dihedral ϕ angle constraints and 43 ψ angle constraints were measured and used in the structural calculations. In the course of the structure calculations, a total of 34 proton pairs were stereospecifically assigned through the program GLOMSA.[24] The 30 conformers constituting the final DYANA family had an average target function of 0.48(±0.05) Å^2 and an average RMSD value over residues 74-143 with respect to the mean structure of 0.37(±0.07) Å for the backbone and of 0.82(±0.07) Å for the heavy atoms. The family of conformers was then subjected to further refinement through energy minimization.[26] The REM family has contributions to the average target function of 0.23(±0.03) Å^2 and of 0.03(±0.01) rad^2, from the NOEs and the dihedral angle constraints, respectively. The average RMSD values for the family with respect to the mean structure are 0.35(±0.06) Å, for the backbone, and 0.84(±0.06) Å for the heavy atoms, for residues 73-143. The RMSD values per residue of the final REM family to the mean structure are shown in Figure 4(b).

The final family of conformers was analyzed with PROCHECK-NMR[27] and the results are shown in Table 2. The secondary structure elements in the energy minimized mean structure involve residues 74-81 (β1), 86-98 (α1), 101-104 and

Table 2. Statistical analysis of the final REM family and the mean structure of Cu(I)-BsCopA(73-151) from *B. subtilis*

	REM (30 structures)	⟨REM⟩
RSM violations per experimental distance constraint (Å)[a]		
Intraresidue (302)	0.0169 ± 0.0025	0.0184
Sequential (399)	0.0117 ± 0.0016	0.0121
Medium range[b] (295)	0.0058 ± 0.0016	0.0054
Long range (419)	0.0090 ± 0.0012	0.0126
Total (1415)	0.0115 ± 0.0009	0.0129
Average number of violations per structure		
Intraresidue	7.8 ± 1.9	7
Sequential	10.6 ± 1.9	11
Medium range[b]	2.2 ± 0.9	1
Long range	6.7 ± 1.5	9
Total	27.3 ± 3.4	28
Average no. of NOE violations larger than 0.3 Å	0.00 ± 0.00	0
Average no. of ϕ, ψ violations larger than 5°	3.3 ± 1.4	5
Largest residual NOE violation (Å)	0.23	0.22
Average NOE deviations (Å^2)	0.23 ± 0.03	0.27
Structural analysis[c]		
% of residues in most favourable regions	78.4	78.9
% of residues in allowed regions	16.9	16.9
% of residues in generously allowed regions	3.9	4.2
% of residues in disallowed regions	0.8	0.0

REM indicates the energy minimized family of 30 structures, ⟨REM⟩ is the energy minimized average structure obtained from the coordinates of the individual REM structures.
[a] The number of experimental constraints for each class is shown in parenthesis.
[b] Medium range distance constraints are those between residues $(i,i+2)$, $(i,i+3)$, $(i,i+4)$ and $(i,i+5)$.
[c] As it results from the Ramachandran plot analysis.

107-108 (β2), 112-119 (β3), 124-134 (α2), and 138-144 (β4). At variance with the structure of the copper-free form, strands β2 and β4 are longer by two residues, even if strand β2 is not present in all the conformers of the family. In the energy minimized average structure, 78.9 % of the residues are in the most favoured regions of the Ramachandran plot, 16.9 % of the residues are in the allowed regions and 4.2 % of residues are in the generously allowed region. No residues are in the disallowed regions (Table 2).

The structure of Cu(I)-BsCopA(73-151), shown in Figure 5(b), is well defined all over its sequence, with the exception of the C-terminal region, which is disordered as in the apo form. The most striking difference between the apo and Cu(I) structures is observed in loop 1, which includes the copper binding ligands. Indeed, the backbone RMSD values with respect to the mean structure for residues 82-89 belonging to loop 1 and the first turn of helix α1 decrease from 0.71 Å, in copper-free form, to 0.40 Å, in the copper-bound form. A dramatic decrease in RMSD value is also observed for the side-chain of Met83, which is well ordered in the copper-bound form (RMSD on SC 0.88 Å) but highly disordered in the copper-free form (RMSD on SC 3.32 Å). A similar behaviour is observed also for Cys85, characterized by higher RMSD in the copper-free form (RMSD on SC 2.91 Å) with respect to the copper-bound form (RMSD on SC 0.87 Å).

Discussion

Comparison between the structures of apo- and Cu(I)-BsCopA

The overall folding of the copper-bound and the copper-free forms are essentially identical, the global backbone RMSD value between the two minimized mean structures being 0.96 Å (Figure 6(a)). All the secondary structure elements are well superimposed, while the largest RMSD values are found in loop 1 and loop 3. Loop 5, which contains Tyr137, does not show any conformational change. Loop 1 shows the largest difference in the local RMSD, with the most significant conformational change involving Cys85, which moves close to the copper ion in the copper-bound form, while Cys88, located in helix α1, does not change position.

In the Cu(I)-BsCopA(73-151) structure the copper ion is coordinated by two cysteine residues, Cys85 and Cys88 (Figure 6(b)). In this structure calculation, no S-Cu-S angle constraints were applied. In the final family, the S-Cu-S angle is 132(±20)°, suggesting that Cu(I) in this environment may be coordinated by a third ligand. The value of this angle is similar to those obtained for the copper(I)-bound forms of Atx1,[28] Ccc2a[22] and BsCopZ.[18] In addition to one of the copper(I)-binding cysteine residues, loop 1 also contains a threonine residue, Thr84, whose oxygen is pointing toward the metal ion (O^{γ}-Cu(I) distance 3.45 Å in the energy minimized average structure) (Figure 6(b)). Therefore, this residue may be considered a long distance

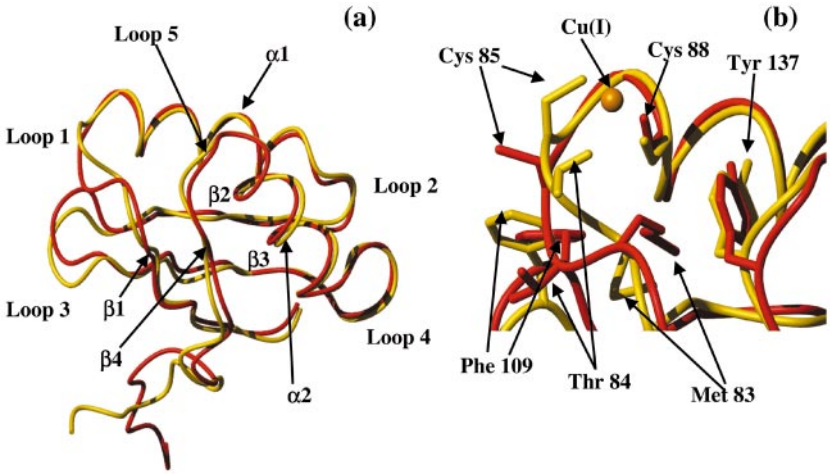

Figure 6. (a) Comparison between the backbone of apo-BsCopAb (red) and Cu(I)-BsCopA(73-151) (gold). (b) Comparison of the metal binding region. The cysteine residues involved in the copper binding, the Cu(I) ion and other relevant residues are shown.

metal ligand. This residue is, indeed, highly conserved as either Thr or Ser in all the bacterial and eukaryotic copper transporter sequences.[29] In the energy minimized average structure of the ATPase domain, Cu(I)-Ccc2a, the oxygen of Thr12 is at 6.5 Å from the metal.[22] A third coordination donor may also come from an exogenous ligand, like DTT, which has been used to reduce the protein. This is indeed what was proposed for Cu(I)-Atx1[30] and then found from EXAFS and NMR[18] studies on BsCopZ.

In loop 1, Met83 has a slightly different conformation in the two protein states, being more ordered in the copper-bound form. The larger number of NMR constraints defining Met83 in the Cu(I) form might be the result of a reduced conformational disorder of this residue when copper is bound. Furthermore, long-range NOE constraints between the side-chain of Met83 and the ring protons of Phe109 are only present in the copper-bound form. Phe109 has a completely different conformation in the two structures, supported by differences in NOE pattern. In the apo form the side-chain moves closer to the surface of the protein, while in the copper-bound form it is oriented towards the binding loop (Figure 6(b)). This residue is conserved as either Phe or Leu in all the bacterial CopA domain sequences located up to now[29] and is also highly conserved as a Phe in the MerP sequences, a protein displaying the same fold and involved in the transport of mercury ions.[31] The comparison between the structures of the mercury-free and mercury-bound forms of MerP showed that this residue (Phe38 in the MerP sequence) displays a similar behaviour with a sizeable change in conformation of the Phe side-chain. This rearrangement has been proposed to be important for protein-protein interactions with its partner MerT.[32] Therefore, the movement of Phe could be relevant to drive a specific interaction with the physiological partner.

Comparison between the solution structures of BsCopA and Ccc2a for both apo and Cu(I) states

The apo-BsCopAb and apo-Ccc2a structures, which have 24 % of residue identity (Figure 7(a)), experience an overall backbone RMSD of 1.40 Å. When copper is bound, BsCopA(73-151) and Ccc2a have an overall backbone RMSD value of 0.99 Å. α helices and β strands are well superimposed with the exception of strand β4, which experiences a minor translation from one structure to the other (Figure 7(b) and (c)). The largest differences in the apo forms are observed for loops 1 and 3, where large conformational changes are present, while loops 2 and 5 are well superimposed (Figure 7(b)). In loop 5, the conserved aromatic residue Tyr137 has an orientation similar to that of Phe64 in Ccc2a (Figure 7(b) and (c)). Loop 4 shows similar conformation but the presence of Pro120 in the BsCopAb sequence (see Figure 7(a)) determines a more extended loop in the apo- and Cu(I)-BsCopA structure. In the two apo forms, the N-terminal copper binding cysteine has a different conformation (Figure 7(b)). In the apo-Ccc2a, this cysteine is the first residue of helix α1, at variance with all the other metallochaperone-like proteins. The conformation of the second copper binding Cys (88 and 16 in BsCopA and Ccc2, respectively) located in helix α1 is similar in the apo structures. In the copper-bound state the biggest difference is observed in loop 1, in particular in the vicinity of the conserved Thr84, which is, in the Cu(I)-Ccc2a structure, farther from the copper ion (Figure 7(c)). At variance with the copper-free forms, loop 3 is now well superimposed and the conformations of both cysteine residues are very similar between the two structures (Figure 7(c)).

The methionine in the MXCXXC metal binding region (Met83) is highly conserved in the N-terminal end of the soluble domains of membrane-bound ATPases as well as in soluble copper transporter proteins.[29] It has been suggested that this methionine in the Ccc2a protein (Met11) acts as a

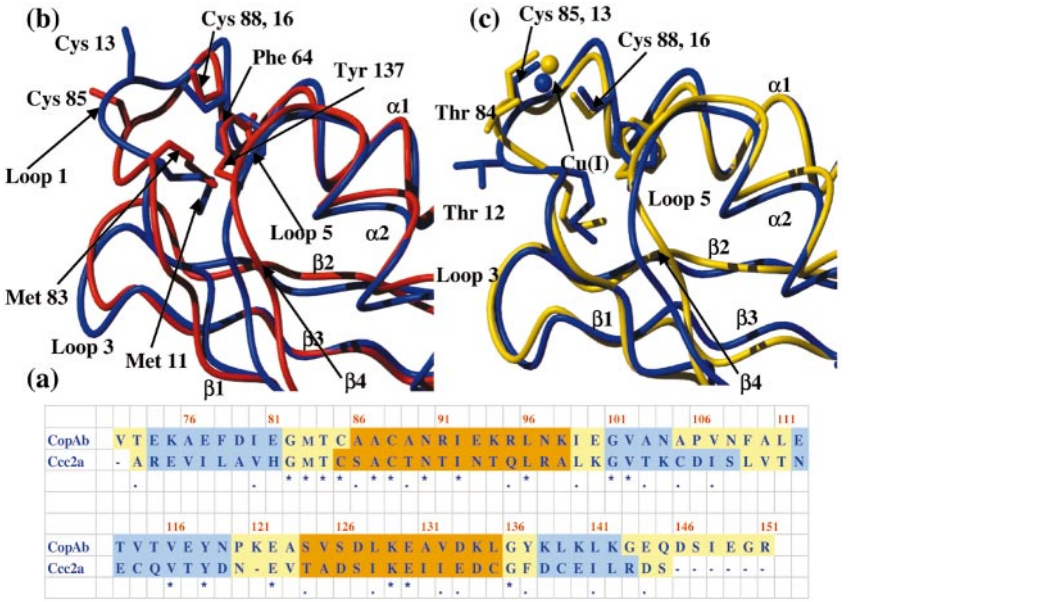

Figure 7. (a) Sequence alignment of the BsCopAb from *B. subtilis* with the second metal-binding domain of yeast Ccc2. The residues are colour shaded according to the secondary structure element (β strands are in blue, α helices are in orange, and loop regions are in yellow), as found in their apo structures. Identical or conservatively substituted residues are indicated by the symbols * and ●, respectively, below the sequences. (b) Comparison of the backbone of apo-BsCopAb (red) and apo-Ccc2a (blu).[22] The cysteine residues involved in copper binding and other relevant residues in the metal binding loop are also shown. (c) Comparison of the backbone of Cu(I)-BsCopA(73-151) (gold) and Cu(I)-Ccc2a (blu).[22] The cysteine residues involved in copper binding, the copper ion and other relevant residues in the metal binding region are also shown.

hydrophobic tether that anchors the metal-binding loop *via* hydrophobic interactions.[22,29] The side-chains of Met83 in the present two structures and of Met11 in Ccc2a show similar conformations, both pointing towards the hydrophobic core of the protein (Figure 7(b) and (c)). However, the side-chain of Met83 is very disordered in the apo-BsCopAb structure (RMSD SC 3.32 Å) with respect to the side-chain of Met11 in apo-Ccc2a (RMSD SC 0.99 Å). Indeed, Met83 does not experience medium- and long-range NOE constraints with the conserved residue Ile92, which were observed in the apo-Ccc2a structure between the corresponding residues. Significantly, also loop 1 is more disordered in apo-BsCopAb than in apo-Ccc2a. This behaviour suggests an influence of the methionine conformation on the extent of order of the entire loop 1, when copper is absent.

Comparison with the solution structures of a potential partner, CopZ

As mentioned before, a potential partner of CopA is CopZ. The latter protein from *B. subtilis* has been recently expressed and characterized in terms of solution structure in the copper-bound form.[18] Also the apo form had been characterized,[18] while the solution structure of the apo state is available for an homologous CopZ from *E. hirae* (hereafter EhCopZ).[19] The sequence alignment of BsCopAb with EhCopZ and BsCopZ reveals 28 % and 27 % residue identity, respectively (Figures 8(a) and 9(a)). As for all the other copper transporting proteins, also the structures of these proteins show the same ferrodoxin-like type of folding (Figures 8(b) and 9(b)). The overall backbone RMSD value of apo-BsCopAb is 1.32 Å with apo-EhCopZ and 1.96 Å with Cu(I)-BsCopZ. All the secondary structure elements and loops 2 and 4 are well superimposed. The largest differences are observed in the metal binding loops 1 and 3 for both proteins (Figures 8(b) and 9(b)). Between BsCopA, in both apo- and copper-bound forms, and Cu(I)-BsCopZ also loop 5 and helix α1, which has a skewed orientation, show changes in conformation (Figures 8(b) and (c) and 9(b)). Therefore, the largest differences are observed in the region around the copper binding site. In particular, in loop 1, the copper-binding Cys85 in apo-BsCopAb has a different orientation when compared to Cys11 in apo-EhCopZ, while the conformation of the other cysteine (Cys88 and 14 in apo-BsCopAb and apo-EhCopZ, respectively) is very similar (Figure 9(b)). The same behaviour is present comparing the apo-BsCopZ with the apo-BsCopA (Figure 8(b)). On the other hand, in the copper-bound forms of BsCopA and BsCopZ, both metal-coordinating cysteine residues are pointing toward the copper ion in a similar way (Figure 8(c)). The latter results indicate that the conformation of the second cysteine is prearranged to bind the metal, while the other metal binding ligand adapts its conformation to the presence or absence of the metal.

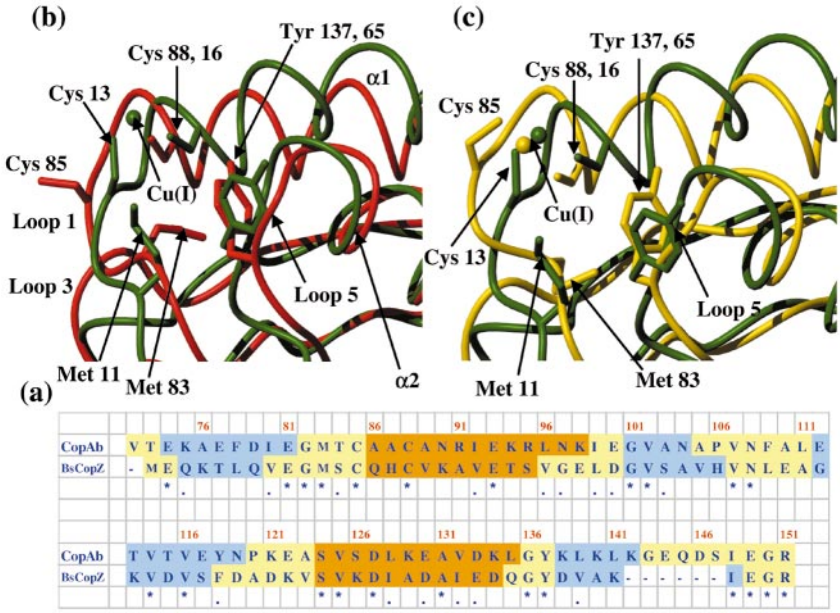

Figure 8. (a) Sequence alignment of the BsCopAb with the CopZ, both from *B. subtilis*. The colour codes are as in Figure 7. (b) Comparison of the backbone of apo-BsCopAb (red) and Cu(I)-CopZ (green).[18] The cysteine residues involved in the copper binding, the Cu(I) ion and other relevant residues are also shown. (c) Comparison of the backbone of Cu(I)-BsCopA(73-151) (gold) and Cu(I)-CopZ (green).[18]

Significantly, Met in loop 1 moves much closer to the copper ion in Cu(I)-BsCopZ than in Cu(I)-BsCopA (Figure 8(c)). Its possible role of protection of the copper ion through hydrophobic interactions, as suggested for BsCopZ,[18] can be ruled out in the present structure. It might be possible that ATPase domains do not need this stabilizing interaction with respect to their chaperone partners.

Tyr137/63 (apo-BsCopAb/apo-EhCopZ), which is presumably important for maintaining an optimal metal-binding loop conformation and which is conserved between the two proteins, does not show significant conformational changes in the

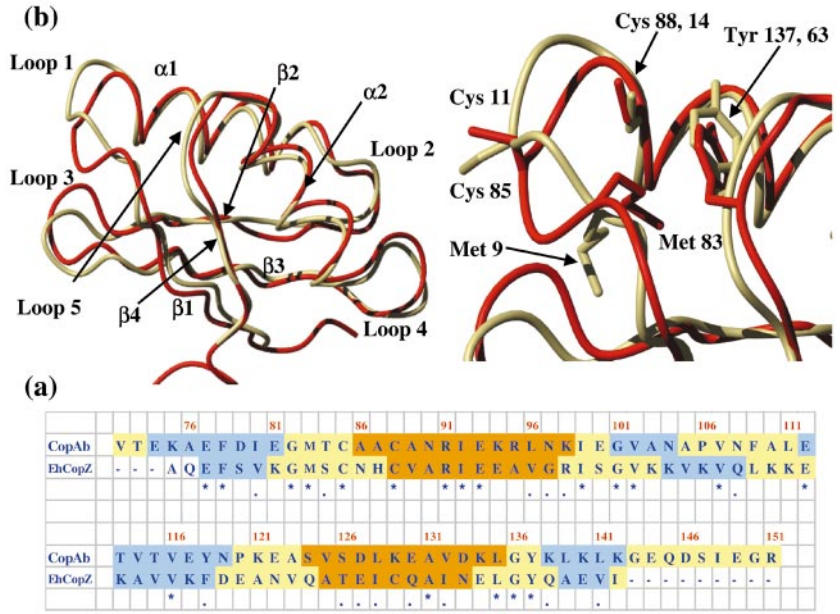

Figure 9. (a) Sequence alignment of BsCopAb from *B. subtilis* with the copper chaperone CopZ from *E. hirae*. The colour codes are as in Figure 7. (b) Comparison of the backbone of apo-BsCopAb (red) and apo-CopZ of *E. hirae* (khaki).[19] The cysteine residues involved in the copper binding and other relevant residues are also shown.

two structures (Figure 9(b)), while a small movement towards the surface of the protein is observed in the Cu(I)-BsCopZ (Figure 8(b) and (c)).

Structural features and biological process

The comparison of the present structures with other related eukaryotic and prokaryotic metallochaperone-like proteins shows that the same fold is chosen by nature in a wide range of species for different types of proteins involved in the same biological function. Significantly, even a metallochaperone and its partner share the same fold. The structural differences are limited to conformational changes in the regions close to the metal binding site, which might be related to the specificity and efficiency of the metal transport process. Long distance electrostatic interactions play a fundamental role in the molecular recognition process. It has been proposed that the interaction between the copper chaperone Atx1 and one domain of its target protein Ccc2 could be determined by the complementary attractions between the cluster of positive residues in Atx1 and the negative region in Ccc2a.[21] A structural model of the interaction between Atx1 and Ccc2a domain has been recently characterized.[33] The surface electrostatic potential distributions on Cu(I)-BsCopZ and on apo-BsCopAb and on a model for the adduct are shown in Figure 10. The distribution of charged residues on the surface of apo-BsCopAb produces a positively charged patch close to the metal binding site, opposite in electrostatic potential to the same side on BsCopZ. Therefore a protein-protein adduct, similar to that already observed for Ccc2 and Atx1, can be proposed for the BsCopA/BsCopZ interaction, since the fold for all these proteins is the same. Furthermore, the fact that always one and the same cysteine in the sequence, i.e. the first in the sequence, experiences changes in the conformation from one species to the other and from one protein to its partner might suggest that this cysteine is that initially involved in the copper transfer process. This free cysteine in the apo form interacts with the copper ion bound to the partner protein, occupying its third coordination position. In this way the bond with the second cysteine in the sequence of copper-loaded protein is weakened at first and eventually broken and the copper transfer process completed. In support of this interpretation we can also note that this cysteine is also the one experiencing the largest changes

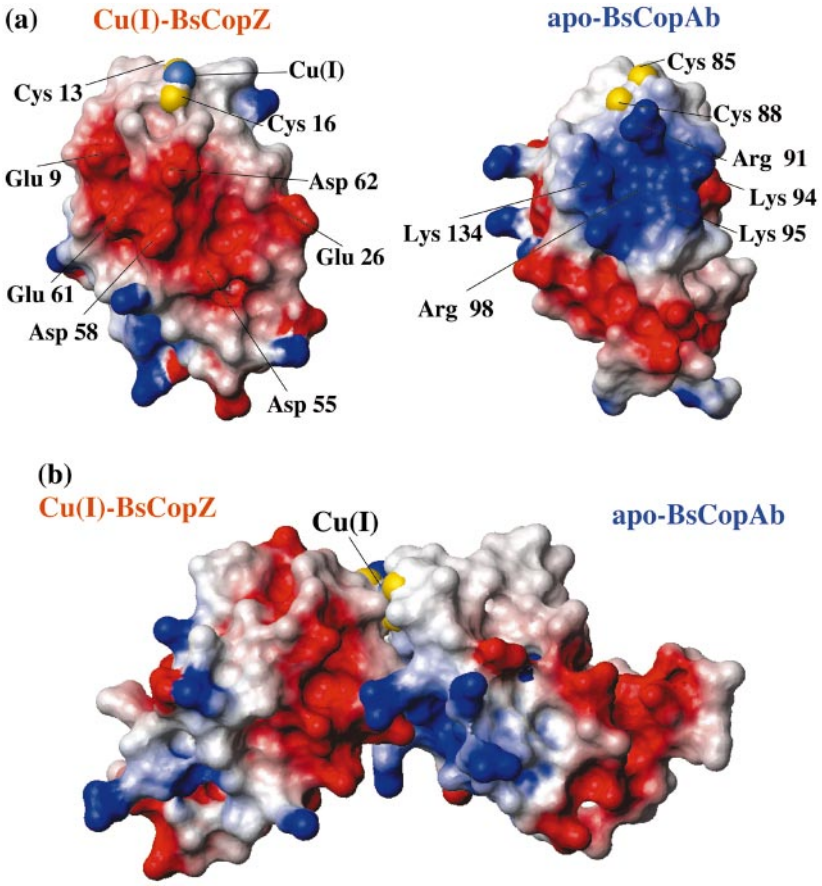

Figure 10. (a) Electrostatic potential surface of apo-BsCopAb and Cu(I)-BsCopZ. The positively and negatively charged and neutral amino acid residues are represented in blue, red and white, respectively. (b) Schematic drawing of the CopAb:CopZ adduct model based on the Ccc2a:Atx1 adduct model.

between the copper-free and the copper-bound forms.

Concluding Remarks

Here the solution structures of a soluble domain of BsCopA from *B. subtilis*, both in the copper-bound and in the copper-free forms, are resolved at a satisfactory degree of resolution. The copper binding site is similar to all other NMR available structures, within the limits of the technique which does not give information on the copper ion itself. Besides an angular arrangement of the S-Cu-S moiety, the oxygen of Thr84 is found close to copper, although possibly beyond bond distance. In the other copper transporting proteins, the oxygen atoms of the corresponding Thr or Ser are located at a much longer distance from copper. The protein belongs to the same operon where BsCopZ is encoded and possibly the two proteins are partners in the copper transport mechanism of the bacteria, as also proposed for the corresponding proteins in *E. hirae*.[34] The soluble domain of BsCopA, which is an ATPase, has the same βαββαβ structure as all proteins involved in the copper transport characterized so far. It appears that the nature of the residues located around the metal binding site, largely determines protein-protein recognition. It is proposed that when the proteins get close, one Cys binds copper(I) of the partner protein by completing adduct formation. Among the two Cys residues of the CXXC motif, following earlier suggestions,[33] the N-terminal Cys85, which belongs to loop 1, may be the first to be involved in the metal transfer. As a matter of fact, this is the Cys more disordered in all copper-depleted structures and, more important, which experiences the largest variability in the side-chain orientation among the copper transport protein structures and between the copper-bound and the copper-free forms. Mobility studies on the millisecond time scale showed that in the yeast Ccc Za:Atx1 adduct N-terminal Cys is not involved in conformational exchange process, while it displays conformational exchange in the apo form of Atx1.[33] On the other hand, the other Cys, belonging to helix α1, has less variable orientations and in the yeast Ccc2a:Atx1 adduct is never involved in conformational exchange process.[33]

Materials and Methods

Protein and DNA sequence analysis

The amino acid sequence encoded by the gene *yvgX* (*bscopa* in this article), initially described in the genome of *B. subtilis* as a putative metal-transporting ATPase,[16] was selected for high homology with the amino acid sequence of yeast Ccc2 and for the presence in the sequence of two MXCXXC metal-binding motifs in the amino acid sequence. BLAST and EXPASY programs (in NCBI and SWISSPROT http servers) were used for analysis, alignment and comparison of nucleotide and amino acid sequences.

Plasmids, bacterial strains and media

E. coli DH5α (Life Technologies) was used as the host strain for cloning and plasmid propagation and *E. coli* BL21(DE3)pLysS (Novagen) as the host strain for the expression vector pET21a (Novagen). Both strains were routinely grown in Luria-Bertani (LB hereafter) broth[35] or on plates of LB agar, supplemented with ampicillin (100 μg/ml) for transformed DH5α strain or with ampicillin (50 μg/ml) and chloramphenicol (17 μg/ml) for transformed BL21(DE3)pLysS strain. Induction was carried out in M9 medium[35] supplemented with 2 ml/l of solution Q (40 mM HCl, 5 g/l $FeCl_2 \cdot 4H_2O$, 184 mg/l $CaCl_2 \cdot 2H_2O$, 64 mg/l H_3BO_3, 18 mg/l $CoCl_2 \cdot 6H_2O$, 340 mg/l $ZnCl_2$, 605 mg/l $Na_2Mo_4 \cdot 2H_2O$, 40 mg/l $MnCl_2 \cdot 4H_2O$) and 10 ml/l of vitamin mix (500 mg/l thiamine, 100 mg/l D-biotin, 100 mg/l choline chloride, 100 mg/l folic acid, 100 mg/l niacinamide, 100 mg/l D-pantothenic acid, 100 mg/l pyridoxal, 10 mg/l riboflavin), including 2 g/l $^{15}NH_4(SO_4)_2$ or/and 2 g/l ^{13}C-labelled glucose as nitrogen and carbon sources, respectively, for labelled protein expression.

Primer design and preparation of the expression vectors

The DNA sequence encoding the BsCopA cytoplasmic region (CR hereafter) of 147 amino acid residues, containing two MXCXXC metal-binding motifs, was amplified by PCR using oligonucleotides corresponding to the CR N-terminal sequence (CRNt: 5′-GGGAATTCcatatg TTGAGTGAACAAAAGGAAATCGCG-3′), and reverse and complementary to the CR C-terminal sequence (CRCt: 5′-TATTctcgagACGACCTTCGATACTGTCCTG TTCACCCTTCAACTTTAGC-3′). The DNA sequence encoding the two metal-binding domains (D1 and D2) corresponding to the first 72 and the last 75 amino acid residues of CR, respectively, harbouring one metal-binding motif each, were amplified by PCR following the same strategy employed for CR amplification. Oligonucleotides CRNt (above) and D1Ct (5′-TATTctcgag ACGACCTTCGATAACGACGTGATAGCCCAGCTTCT C-3′) were used for D1 domain amplification; and oligonucleotides D2Nt (5′-GGGAATTCcatatgACAGAAAA AGCTGAATTCGATATTGAGGGC-3′) and CRCt (above) for D2 domain amplification. The four oligonucleotides presented *Nde*I (CRNt and D2Nt oligonucleotides) or *Xho*I (CRCt and D1Ct oligonucleotides) restriction enzyme sites at the 5′ ends (lowercase), and CRCt and D1Ct oligonucleotides presented a sequence encoding a factor Xa recognition site (underlined). The PCR reactions were carried out in 100 μl volume, using 0.1 μg genomic DNA extracted from *B. subtillis*[36] as template, 20 pmol of each oligonucleotide and 2.5 units of Amplitaq (Perkin-Elmer Cetus). A ''hot start'' of five minutes at 95 °C was followed by five cycles of denaturation (35 seconds at 95 °C), annealling (one minute at 50 °C) and extension (two minutes at 72 °C); and 20 more cycles of denaturation (35 sec at 95 °C), annealling (one minute at 60 °C) and extension (two minutes at 72 °C) using a programmable heating block (Perkin-Elmer Gene Amp System 9600). The PCR products were separated on 1.5 % (w/v) agarose gels in TAE buffer, purified using Qiaex II Gel extraction kit (Qiagen) and the three *Nde*I-*Xho*I fragments were cloned into pET21a (Novagen) beside a sequence encoding a $(His)_6$-tag at the C-terminal end yielding pET21-CR, pET21-D1 and pET21-D2. Sequencing of the engineered DNA fragments was achieved using an automatic sequencer ABI 377.

Expression and purification of the recombinant BsCopA cytoplasmic region and the two metal-binding domains

BL21(DE3)pLysS cells containing the expression vector (pET21-CR or pET-D1 or pET21-D2) were grown at 37 °C and 130 r.p.m. in induction medium containing 50 μg/ml ampicillin to an absorption of 0.6 at 600 nm. Then IPTG was added to a final concentration of 1 mM and the cells were further incubated at 37 °C for 4 hours. Cultures were then centrifuged at 4 °C and 5000 *g* for ten minutes and the cell pellet resuspended in buffer A (20 mM Na_2HPO_4, 1 M NaCl, 5 mM imidazole, pH 8.0) and stored at −20 °C overnight. After thawing, the suspension was sonicated and centrifugated at 4 °C and 40,000 r.p.m. in a Beckman 60 Ti rotor for ten minutes. The soluble fraction containing the $(His)_6$-tagged recombinant protein was loaded onto a HiTrap 5 ml affinity column (Amersham pharmacia biotech) previously charged with Zn^{2+}. The column was consecutively washed with 100 ml of buffer B (20 mM Na_2HPO_4, 1 M NaCl, 20 mM imidazole, pH 8.0) and 100 ml of buffer C (20 mM Na_2HPO_4, 1 M NH_4Cl, 20 mM imidazole, pH 8.0). Proteins attached to the column, including $(His)_6$-tagged recombinant protein, were eluted with 30 ml of buffer D (20 mM Na_2HPO_4, 1 M NaCl, 20 mM imidazole, 50 mM EDTA, pH 8.0) and then dialysed against 100 mM NaCl, 50 mM Tris-HCl, 1 mM $CaCl_2$ (pH 8.0). Restriction protease factor Xa (Roche diagnostic) was added to the solution (30 μg of enzyme for each 20 mg of recombinant protein) to cut the $(His)_6$-tag. After incubation for 24 hours at 4 °C, the solution was dialysed against buffer A and loaded again onto the HiTrap 5 ml affinity column charged with Zn^{2+}. The minor contaminant proteins were retained into the column and the recombinant protein without the $(His)_6$-tag was eluted in buffer A. The purity was checked by SDS-PAGE in 22 % polyacrylamide gels after staining of protein bands with Coomasie Blue R-250.

The hydropathy profiles were calculated by the Dense Alignment Surface method.[37]

NMR sample

The apo protein sample was reduced by addition of DTT (20-fold molar excess relative to protein) in anaerobic conditions and then washed with 100 mM sodium phosphate buffer (pH 7.0).

The copper (I) derivative was prepared following the procedure described by Pufahl R.A. *et al*.[38] for the copper chaperone Atx1 with some modifications. Copper content was checked through atomic absorption measurements with a Perkin Elmer 2380 instrument.

The NMR samples contained about 1.5 mM protein. All manipulations were carried out under inert atmosphere at 4 °C.

NMR spectroscopy

NMR spectra were acquired on Avance 800, 700 and 600 Bruker spectrometers operating at a proton nominal frequency of 800.13 MHz, 700.13 MHz and 600.13 MHz, respectively, and equipped with ($^1H/^{13}C/^{15}N$) triple resonance probes with pulsed field gradients. The NMR data were processed with XWINNMR software and the spectral analysis was done with XEASY.[39] On the $^{15}N,^{13}C$ apo-BsCopAb sample, three-dimensional CBCA (CO)NH and CBCANH[40] NMR spectra were used to obtain the backbone sequential assignments. Side-chain resonances were assigned using 3D ^{13}C H(C)CH-TOCSY and (H)CCH-TOCSY experiments.[41] A ^{15}N NOESY-HSQC and a ^{13}C NOESY-HSQC[42] (100 ms mixing times) spectra have been collected at 700 MHz to obtain dipolar connectivities and a HNHA[43] experiment has been performed to determine the $^3J_{HNH\alpha}$ coupling constants. Watergate 2D NOESY (100 ms mixing time) and TOCSY (80 ms spin lock time) experiments have been registered at 700 MHz to identify connectivities involving aromatic residues of the protein. On the ^{15}N Cu(I)-BsCopA(73-151) sample, 3D and 2D NOESY, TOCSY experiments and a HNHA experiment have been collected at the 700 MHz.

For all the 3D experiments, quadrature detections in the indirect dimensions were performed in echo-antiecho and States-TPPI modes. For all the 2D experiments, quadrature detection in the indirect dimensions was performed in the TPPI mode.[44] Water suppression was achieved through WATERGATE sequence.[45] All 3D and 2D spectra were collected at 298 K.

NMR structure calculations

The cross-peak volumes of NOESY spectra were converted into upper distance limits, to be used as input for structure calculations, by using the approach provided by the program CALIBA.[24] The calibration curves were adjusted iteratively as the structure calculations proceeded. Stereospecific assignments of diastereotopic protons were obtained using the program GLOMSA.[24] $^3J_{HNH\alpha}$ coupling constants were correlated to the backbone torsion angle ϕ by means of the appropriate Karplus curve.[43] For $^3J_{HNH\alpha}$ values larger than 7 Hz the ϕ angle ranges between −155° and −85° while for values lower than 4.5 Hz it ranges between −70° and −30°. Backbone dihedral angle ψ for residue $(i-1)$ have been determined from the ratio of the intensity of the $d_{\alpha N}(i-1,i)$ and $d_{N\alpha}(i,i)$ NOE found on the ^{15}N plane of residue (i) in the ^{15}N NOESY-HSQC.[46] Ratio values of the residue $(i-1)$ larger than 1 are characteristic of β sheet, with ψ values ranging between 60° and 180°, while values smaller than 1 indicate a right-handed α helix, with ψ values between −60° and −20°.[46] The program CSI[25] was also used to determine backbone dihedral angles from the chemical shift index (CSI). The copper ion was included in the calculations by adding a new residue in the amino acid sequence, formed by a chain of dummy atoms which have their van der Waals radii set to zero so that it can freely penetrate into the protein and one atom with a radius of 1.4 Å, which mimics the copper ion. The sulphur atoms of Cys85 and Cys88 were linked to the metal ion through upper distance limits of 2.5 Å. This approach does not impose any fixed orientation of the ligands with respect to the copper. In particular, no constraint on the S(Cys85)-Cu-S(Cys88) angle was used. Structure calculations were performed using DYANA.[47] 300 random conformers were annealed in 10,000 steps using NOE and both ^3J value (when available) constraints. The 30 conformers with the lowest target function constitute the final family.

Restrained energy minimization (REM) was then applied using the SANDER module of AMBER 5.0 program package.[26] The force field parameters for the copper(I) ion were adapted from similar systems.[48] The NOE and torsion angle constraints were applied with

force constants of 50 kcal mol^{-1} $Å^{-2}$ and 32 kcal mol^{-1} rad^{-2}, respectively.

The program CORMA[49] was used to check the agreement between the experimental and the back-calculated NOESY cross-peaks, obtained from the calculated structure. The quality of the structure was evaluated through Ramachandran analysis, performed using the programs PROCHECK[50] and PROCHECK-NMR.[27]

Acknowledgements

We thank Professor G. Piccardi for the recording of the atomic absorption spectra, Dr Elena Molteni for help in the analysis of the genomes of some organisms. This work was supported by the European Community (contract HPRI-CT-1999-00009), by Italian CNR (contract number 01.00238.PF49) and by MURST-ex 40 %.

References

1. Pena, M. M. O., Lee, J. & Thiele, D. J. (1999). A delicate balance: homeostatic control of copper uptake and distribution. *J. Nutr.* **129**, 1251-1260.
2. Huffman, D. L. & O'Halloran, T. V. (2001). Function, structure, and mechanism of intracellular copper trafficking proteins. *Annu. Rev. Biochem.* **70**, 677-701.
3. Harrison, M. D., Jones, C. E., Solioz, M. & Dameron, C. T. (2000). Intracellular copper routing: the role of copper chaperones. *Trends Biochem. Sci.* **25**, 29-32.
4. Rosenzweig, A. C. (2001). Copper delivery by metallochaperone proteins. *Acc. Chem. Res.* **34**, 119-128.
5. O'Halloran, T. V. & Culotta, V. C. (2000). Metallochaperones: an intracellular shuttle service for metal ions. *J. Biol. Chem.* **275**, 25057-25060.
6. Rae, T., Schmidt, P. J., Pufahl, R. A., Culotta, V. C. & O'Halloran, T. V. (1999). Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. *Science,* **284**, 805-808.
7. Rensing, C., Ghosh, M. & Rosen, B. P. (1999). Families of soft-metal-ion-transporting ATPases. *J. Bacteriol.* **181**, 5891-5897.
8. Odermatt, A., Suter, H., Krapf, R. & Solioz, M. (1993). Primary structure of two P-type ATPase involved in copper homeostasis in *Enterococcus hirae*. *J. Biol. Chem.* **268**, 12775-12779.
9. Bayle, D., Wängler, S., Weitzenegger, T., Steinhilber, W., Volz, J., Przybylski, M. *et al*. (1998). Properties of the P-type ATPase encoded by the CopAP operons of *Helicobacter pylori* and *Helicobacter felis*. *J. Bacteriol.* **180**, 317-329.
10. Melchers, K., Herrmann, L., Mauch, F., Bayle, D., Heuermann, D., Weitzenegger, T. *et al*. (1998). Properties and function of the P type ion pumps cloned from *Helicobacter pylori*. *Acta Physiol. Scand. Suppl.* **643**, 123-135.
11. Rensing, C., Fan, B., Sharma, R., Mitra, B. & Rosen, B. P. (2000). CopA: an *Escherichia coli* Cu(I)-translocating P-type ATPase. *Proc. Natl Acad. Sci. USA,* **97**, 652-656.
12. Kanamaru, K., Kashiwagi, S. & Mizuno, T. (1994). A copper-transporting P-type ATPase found in the thylakoid membrane of cyanobacterium *Synechococcus* species PCC7942. *Mol. Microbiol.* **13**, 369-377.
13. Phung, L. T., Ajlani, G. & Haselkorn, R. (1994). P-type ATPase from the cyanobacterium Synechococcus 7942 related to the human Menkes and Wilson disease gene products. *Proc. Natl Acad. Sci. USA,* **91**, 9651-9654.
14. Solioz, M. & Vulpe, C. D. (1996). CPx-type ATPases: a class of P-type ATPases that pump heavy metals. *Trends Biochem. Sci.* **21**, 237-241.
15. Melchers, K., Weitzenegger, T., Buhmann, A., Steinhilber, W., Sachs, G. & Schafer, K. P. (1996). Cloning and membrane topology of a P-type ATPase from *Helicobacter pylori*. *J. Biol. Chem.* **271**, 446-457.
16. Vulpe, C. D., Levinson, B., Whitney, S., Packman, S. & Gitschier, J. (1993). Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. *Nature Genet.* **3**, 7-13.
17. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V. *et al*. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. *Nature,* **390**, 249-256.
18. Banci, L., Bertini, I., Del, Conte R., Markey, J. & Ruiz-Dueñas, F. J. (2001). Copper trafficking: the solution structure of *Bacillus subtilis* CopZ. *Biochemistry,* **40**, 15660-15668.
19. Wimmer, R., Herrmann, T., Solioz, M. & Wüthrich, K. (1999). NMR structure and metal interactions of the CopZ copper chaperone. *J. Biol. Chem.* **274**, 22597-22603.
20. Wüthrich, K. (1986). *NMR of Proteins and Nucleic Acids*, Wiley, New York.
21. Rosenzweig, A. C., Huffman, D. L., Hou, M. Y., Wernimont, A. K., Pufahl, R. A. & O'Halloran, T. V. (1999). Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. *Struct. Fold. Des.* **7**, 605-617.
22. Banci, L., Bertini, I., Ciofi-Baffoni, S., Huffman, D. L. & O'Halloran, T. V. (2001). Solution structure of the yeast copper transporter domain Ccc2a in the apo and Cu(I)-loaded states. *J. Biol. Chem.* **276**, 8415-8426.
23. Gitschier, J., Moffat, B., Reilly, D., Wood, W. I. & Fairbrother, W. J. (1998). Solution structure of the fourth metal-binding domain from the Menkes copper-transporting ATPase. *Nature Struct. Biol.* **5**, 47-54.
24. Güntert, P., Braun, W. & Wüthrich, K. (1991). Efficient computation of three-dimensional protein structures in solution from Nuclear Magnetic Resonance data using the program DIANA and the supporting programs CALIBA, HABAS and GLOMSA. *J. Mol. Biol.* **217**, 517-530.
25. Wishart, D. S. & Sykes, B. D. (1994). The ^{13}C chemical shift index: a simple method for the identification of protein secondary structure using ^{13}C chemical shift data. *J. Biomol. NMR,* **4**, 171-180.
26. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M. *et al*. (1997). *AMBER 5.0*, University of California, San Francisco.
27. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R. & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR,* **8**, 477-486.
28. Arnesano, F., Banci, L., Bertini, I., Huffman, D. L. & O'Halloran, T. V. (2001). Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. *Biochemistry,* **40**, 1528-1539.
29. Arnesano, F., Banci, L., Bertini, I., Ciofi-Baffoni, S., Molteni, E., Huffman, D. L. & O'Halloran, T. V. (2002). Metallochaperones and metal-transporting

ATPases: a comparative analysis of sequences and structures. *Genome Res.* **12**, 255-271.
30. Pufahl, R. A., Singer, C. P., Peariso, K. L., Lin, S.-J., Schmidt, P. J., Fahrni, C. J. *et al*. (1997). Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science,* **278**, 853-856.
31. Foster, T. J. (1987). The genetics and biochemistry of mercury resistance. *CRC Crit. Rev. Microbiol.* **15**, 117-140.
32. Steele, R. A. & Opella, S. J. (1997). Structures of the reduced and mercury-bound forms of MerP, the periplasmic protein from the bacterial mercury detoxification system. *Biochemistry,* **36**, 6885-6895.
33. Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L. & O'Halloran, T. V. (2001). Characterization of the binding interface between the copper chaperone Atx1 and the first cytosolic domain of Ccc2 ATPase. *J. Biol. Chem.* **276**, 41365-41376.
34. Multhaup, G., Strausak, D., Bissig, K. D. & Solioz, M. (2001). Interaction of the CopZ copper chaperone with the CopA copper ATPase of *Enterococcus hirae* assessed by surface plasmon resonance. *Biochem. Biophys. Res. Commun.* **288**, 172-177.
35. Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
36. Marmur, M. J. (1961). A procedure for the isolation of deoxyribonucleic acid from micro-organisms. *J. Mol. Biol.* **3**, 208-218.
37. Cserzo, M., Wallin, E., Simon, I., vio Heijne, G. & Elofsson, A. (1997). Prediction of transmembrane alpha helices in prokaryotic membrane proteins: the dense alignment surfase method. *Protein Eng.* **10**, 673-676.
38. Pufahl, R., Singer, C. P., Peariso, K. L., Lin, S.-J., Schmidt, P. J. & Fahrni, C. J., *et al*. (1997). Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science,* **278**, 853-856.
39. Bartels, C., Xia, T. H., Billeter, M., Güntert, P. & Wüthrich, K. (1995). The program XEASY for computer-supported NMR. *J. Biomol. NMR*, **5**, 1-10.
40. Kay, L. E., Ikura, M., Tschudin, R. & Bax, A. (1990). Three-dimensional triple-resonance NMR spectroscopy of isotopically enriched proteins. *J. Magn. Reson.* **89**, 496-514.
41. Kay, L. E., Xu, G. Y., Singer, A. U., Muhandiram, D. R. & Forman-Kay, J. D. (1993). A gradient-enhanced HCCH-TOCSY experiment for recording side-chains ^{1}H and ^{13}C correlations in H_2O samples of proteins. *J. Magn. Reson. Ser. B*, **101**, 333-337.
42. Wider, G., Neri, D., Otting, G. & Wüthrich, K. (1989). A heteronuclear three-dimensional NMR experiment for measurements of small heteronuclear coupling constants in biological macromolecules. *J. Magn. Reson.* **85**, 426-431.
43. Vuister, G. W. & Bax, A. (1993). Quantitative *J* correlation: a new approach for measuring homonuclear three-bond $J(H^{N}H^{\alpha})$ coupling constants in ^{15}N enriched proteins. *J. Am. Chem. Soc.* **115**, 7772-7777.
44. Marion, D. & Wüthrich, K. (1983). Application of phase sensitive correlated spectroscopy (COSY) for measurements of proton-proton spin-spin coupling constants in proteins. *Biochem. Biophys. Res. Commun.* **113**, 967-974.
45. Piotto, M., Saudek, V. & Sklenar, V. (1992). Gradient-tailored excitation for single quantum NMR spectroscopy of aqueous solutions. *J. Biomol. NMR,* **2**, 661-666.
46. Gagne', R. R., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B. & Sykes, B. D. (1994). Quantification of the calcium-induced secondary structural changes in the regulatory domain of troponin-C. *Protein Sci.* **3**, 1961-1974.
47. Güntert, P., Mumenthaler, C. & Wüthrich, K. (1997). Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283-298.
48. Banci, L., Benedetto, M., Bertini, I., Del, Conte R., Piccioli, M. & Viezzoli, M. S. (1998). Solution structure of reduced monomeric Q133M2 copper, zinc superoxide dismutase. Why is SOD a dimeric enzyme? *Biochemistry,* **37**, 11780-11791.
49. Borgias, B., Thomas, P. D. & James, T. L. (1989). *Complete Relaxation Matrix Analysis (CORMA)*, University of California, San Francisco.
50. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993). PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallog.* **26**, 283-291.
51. Koradi, R., Billeter, M. & Wüthrich, K. (1996). MOLMOL: a program for display and analysis of macromolecular structure. *J. Mol. Graph.* **14**, 51-55.

Edited by M. F. Summers

(Received 1 August 2001; received in revised form 16 January 2002; accepted 16 January 2002)

http://www.academicpress.com/jmb

Supplementary Material comprising Tables of ^{1}H, ^{15}N and ^{13}C chemical shifts, experimental NOEs and angle constraints, and stereospecific assignment is available on IDEAL