Copper Trafficking: the Solution Structure of *Bacillus subtilis* CopZ†

Lucia Banci, Ivano Bertini,* Rebecca Del Conte, Jacob Markey, and Francisco Javier Ruiz-Dueñas‡

Centro di Risonanze Magnetiche and Department of Chemistry, University of Florence, Via Luigi Sacconi 6, 50019 Sesto Fiorentino, Italy

Received June 19, 2001

ABSTRACT: A sequence with a high homology (39% residue identity) with that of the copper-transport CopZ protein from *Enterococcus hirae* and with the same MXCXXC metal-binding motif has been identified in the genome of *Bacillus subtilis*, and the corresponding protein has been expressed. The protein, constituted by 73 amino acids, does bind copper(I) under reducing conditions and fully folded in both copper-bound and copper-free forms under the present experimental conditions. The solution structure of the copper-bound form was determined through NMR spectroscopy on an ^{15}N-labeled sample. A total of 1508 meaningful nuclear Overhauser effects, 38 dihedral ϕ angles, and 48 dihedral ψ angles were used in the structural calculations, which lead to a family of 30 conformers with an average rmsd to the mean structure of 0.32 ± 0.06 Å for the backbone and of 0.85 ± 0.07 Å for the heavy atoms. NMR data on the apoprotein also show that, also in this form, the protein is in a folded state and essentially maintains the complete secondary structure. Some disorder is observed in the loop devoted to copper binding. These results are compared with those reported for CopZ from *E. hirae* whose structure is well-defined only in the apo form. The different behaviors of copper-loaded *E. hirae* and *B. subtilis* are tentatively accounted for on the basis of the presence of dithiothreitol used in the latter case, which would stabilize the monomeric form. The comparison is extended to other similar proteins, with particular attention to the copper-binding loop. The nature and the location of conserved residues around the metal-binding site are discussed with respect to their relevance for the metal-binding process. Proposals for the role of CopZ are also presented.

The genome sequencing of organisms makes available a large number of proteins which require structural characterization. A reasonable approach to this requirement is that of studying the proteins involved in a given function. Among the many functions in living organisms, copper homeostasis is raising more and more attention (*1, 2*); copper is a trace element which is essential because of its catalytic properties in enzymes. This metal ion enters the cell and is ultimately incorporated into the appropriate active sites through processes which are being uncovered (*1, 3, 4*). On the other hand, free copper ions can cause serious cell damage through radical formation; therefore, the careful regulation of intracellular copper concentration is required (*5, 6*). Recently, metal-ion chaperones have been identified (*1, 7, 8*), which are part of the cellular metabolism in eukaryotes and guarantee that the metal ions are delivered to the various organelles and eventually to the target proteins, being, in all of the steps, always coordinated to a protein (*1*−*4*).

Less is known in prokaryotic organisms (*1, 4, 6, 9*). Furthermore, because of the different cellular organization of Gram-positive and Gram-negative organisms, different pathways and chains of proteins are expected. For example, in *Enterococcus hirae*,[1] a Gram-positive bacterium, one of the copper-transport processes is performed by a protein, CopZ (EhCopZ hereafter; *10*), that is structurally characterized in the metal-free reduced form (*11*). The EhCopZ was proposed to interact with a copper-responsive repressor, CopY (*12*), and two copper-transporting ATPases, CopA and CopB (*1, 13*).

Within the frame of a research aiming at the characterization of proteins involved in copper homeostasis, in order to understand the mechanism of copper trafficking, we have selected to browse the genome of *Bacillus subtilis*, which is again a Gram-positive bacterium. The search was carried out using as a template the sequence of *Saccharomyces cerevisiae* Atx1 (yAtx1 hereafter), a protein involved in copper transport in a yeast cell, as an example of eukaryotic organisms (*14*), of MerP, which is involved in mercury transport in Gram-negative bacteria (*8*), and of EhCopZ (*11*) itself. We found a gene encoding a protein, which we call BsCopZ hereafter, whose sequence is significantly homologous with that of the above proteins (with an identity of 28% to yAtx1, 35% to MerP, and 39% to EhCopZ). Then, it was decided to pursue its expression and purification in order to attempt to solve the solution structure of the copper-bound protein by taking advantage of our experience in handling copper-transport proteins (*15*). The findings are instructive in comparison with both the results available on EhCopZ and the structures of other copper-transport proteins.

EXPERIMENTAL PROCEDURES

Protein and DNA Sequence Analysis. The amino acid sequence encoded by the gene *yvgY* (*bscopz* in this article)

† This work was supported by the European Community (Contract No. HPRI-CT-1999-00009), by Italian CNR (Progetto Finalizzato Biotecnologie 99.00286.PF49), and by MURST-ex 40%.

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

‡ Present address: Department of Molecular Microbiology, Centro de Investigaciones Biológicas, CSIC, Velázquez 144, E-28006 Madrid, Spain.

[1] Abbreviations: *E. hirae*, *Enterococcus hirae*; *S. cerevisiae*, *Saccharomyces cerevisiae*; *E. coli*, *Escherichia coli*; *B. subtilis*, *Bacillus subtilis*; *H. pylori*, *Helicobacter pylori*; Cu−ATPase, copper-transporting P-type ATPase.

10.1021/bi0112715 CCC: $20.00
Published on Web 12/01/2001

in the genome of *B. subtilis* (*16*) was selected for its high homology with the amino acid sequence of EhCopZ, yeast Atx1 (yAtx1) and MerP and for the presence in the sequence of the MXCXXC metal-binding motif in the amino acid sequence. BLAST and EXPASY programs (in NCBI and SWISSPROT http servers) were used for the analysis, alignment, and comparison of nucleotide and amino acid sequences.

Primer Design and Preparation of the Expression Vector with the Gene bscopz. Oligonucleotides corresponding to the N-terminal sequence (Ntbscopz: 5′-CGggatccATCGAAGGTCGTATGGAACAAAAAACATTGCAAG-3′) and reverse and complementary to the C-terminal sequence (Ctbscopz: 5′-TATTctcgagACGACCTTCGATCTTGGCTACGTCATAGCCCTG-3′) of the gene *bscopz* were synthesized (Beckman oligo 1000 M synthesizer). Both of the oligonucleotides presented a factor Xa recognition site (underlined) and *Bam*HI and *Xho*I restriction enzyme sites, respectively, at the 5′ ends (lowercase). These oligonucleotides were used as primers and genomic DNA extracted from *B. subtilis*, following the protocol described by Marmur (*17*), as a template in polymerase chain reactions (PCR) for the amplification of gene *bscopz*. The reactions were carried out in 100 μL, using 1 μg of DNA, 20 pmol of each primer, and 2.5 U of Amplitaq (Perkin-Elmer Cetus) and 333 K as annealing temperature. The PCR products were separated on 1.5% agarose gels in a TAE buffer, purified using a Qiaex II gel extraction kit (Qiagen), and the *Bam*HI−*Xho*I fragment together with a sequence encoding an $(His)_6$-tag was cloned into pET21a (Novagen), yielding pET21−*bscopz*. Sequencing of the engineered *bscopz* gene was achieved using an automatic sequencer ABI 377.

Expression and Purification of Recombinant BsCopZ. *Escherichia coli* DH5α (Life Technologies) was used as the host strain for cloning and for plasmid propagation. *E. coli* BL21 (DE3)pLysS (Novagen) was used as the host strain for the expression of BsCopZ. The BL21 (DE3)pLysS cells containing pET21−*bscopz* were grown at 310 K and 120 rpm in an induction medium containing 50 μg/mL of ampicillin to an optical density of 0.6 at 600 nm. Then, IPTG was added to a final concentration of 1 mM, and the cells were further incubated at 310 K for 5 h in an LB medium or for 24 h in an M9 medium (*18*). For the expression of the ^{15}N-labeled protein, the M9 medium contained 1 g/L of $^{15}NH_4(SO_4)_2$ as a nitrogen source. The cell pellet, obtained by the cultures, was resuspended in buffer A (20 mM Na_2HPO_4, 1 M NaCl, 5 mM imidazole, pH = 8.0) and stored at 253 K overnight. After being thawed, the suspension was sonicated and centrifuged. The soluble fraction containing the $(His)_6$-tagged recombinant BsCopZ was loaded onto a HiTrap 5 mL affinity column (Amersham Pharmacia Biotech) previously charged with Zn^{2+}. The column was consecutively washed with 50 mL of buffer B (20 mM Na_2HPO_4, 1 M NaCl, 20 mM imidazole, pH = 8.0) and 50 mL of buffer C (20 mM Na_2HPO_4, 1 M NH_4Cl, 20 mM imidazole, pH = 8.0). The proteins attached to the column, including $(His)_6$-tagged recombinant BsCopZ, were eluted with 10 mL of buffer D (20 mM Na_2HPO_4, 1 M NaCl, 20 mM imidazole, 50 mM EDTA, pH = 8.0). The restriction protease factor Xa (Roche diagnostic) was used to cut the $(His)_6$-tag, and the reaction product was dialyzed against buffer A and loaded again onto the HiTrap 5 mL affinity column. The minor contaminant proteins were retained in the column, and the BsCopZ without the $(His)_6$-tag was eluted in buffer A. The purity was checked by sodium dodecyl phosphate−polyacrylamide gel electrophoresis in 22% polyacrylamide gels after the staining of the protein bands with Coomasie Blue R-250. Up to 10 mg of the pure protein was obtained in the LB medium and up to 4 mg in the M9 medium per liter of bacterial culture. The molecular mass of the purified protein was 7826 or 7834 Da when it was produced in either the LB or M9 medium, respectively. This means that the protein obtained in the LB medium is 32 ± 2 Da higher than the theoretical molecular mass calculated from the amino acid sequence, which could correspond to the introduction of two oxygen atoms. A 32 Da excess is sometimes found in polypeptides containing methionine residues, as occurs in the present protein, because methionine can easily oxidize (*19*). When the protein is grown in an M9 medium, the molecular mass is 41 ± 2 Da higher than expected. This 41 Da increase is found for N-acetylation of either the amino terminus or the ϵ-NH_2 of one of the lysines in the protein (*20*). The molecular mass was not determined for the copper-loaded form as the copper ion is exposed to the solvent and therefore is easily lost during electrospray ionization mass spectroscopy (ESI-MS). The first five amino acids (MEQKT) were sequenced for the ^{15}N-labeled sample, which showed the same retention time as the standard amino acids, thus revealing that no modification exists in this segment composition. The value obtained for ϵ_{280} was 2200 M^{-1} cm^{-1}. The Cu/protein ratio was 0.77, similar to that found for the copper chaperone yAtx1 from *S. cerevisiae* (*21*). The copper-loaded form was the predominant species in the NMR samples, and its signals were well-identified with respect to the signals because of the apo form.

After the purification, the protein samples were kept under reduction conditions with dithiothreitol (DTT) in an inert atmosphere.

Protein Characterization. ϵ_{280} was estimated by the Bradford assay (Biorad) using immunoglobulin G as a standard, and the resulting value was used for the determination of the protein concentration (ϵ_{280} = 2200 M^{-1} cm^{-1}). The molecular mass was determined by ESI-MS using an API 365 ESI-MS instrument (Applied Biosystem) equipped with a Perkin-Elmer HPLC and an autosampler series 200. For the HPLC, a chromatography column Jupiter C5, 50 × 4.6 mm (Phenomenex), was used. The N-terminal sequence was obtained by automated Edman degradation in a Perkin-Elmer/Applied Biosystems Procise 494 sequencer.

Sample Preparations. The apoprotein sample was kept reduced by the addition of DTT in anaerobic conditions. The copper(I) derivative was prepared following the procedure described by Pufahl et al. (*14*) for the copper chaperone yAtx1 with some modifications. DTT (20-fold molar excess relative to protein) was used as a reductant. All of the manipulations were carried out under an inert atmosphere at 277 K. The copper content was checked through atomic absorption measurements with a GBC 903 instrument.

The NMR samples contained about 2 mM protein and were prepared in a 100 mM sodium phosphate buffer, pH = 7.0.

NMR Experiments. The NMR experiments for the signal assignment of BsCopZ were recorded on a Bruker Avance

700 spectrometer operating at 16.4 T. All of the experiments were performed with a triple-resonance 5 mm probe at 300 K on a ^{15}N-labeled sample. The NMR data were processed with the XWinNMR software, and the spectra analysis was done with XEASY (*22*).

For the Cu(I)-bond protein, the spectral window for both proton dimensions in homonuclear two-dimensional NOESY (2D NOESY; *23*; 100 ms mixing time) and 2D TOCSY (*24*; 80 ms spin lock time) was 8865 Hz for 1024 × 1024 data points. The 2D ^{15}N-HSQC experiment was collected with spectral windows of 8865 (^{1}H) × 3125 (^{15}N) Hz for 1024 × 256 data points, and phase sensitivity improvement using an Echo/Antiecho gradient selection was applied (*25*). The three-dimensional (3D) ^{15}N-NOESY-HSQC and ^{15}N-TOCSY-HMQC *26*) were collected with spectral windows of 8865 (^{1}H) × 3125 (^{15}N) × 8865 (^{1}H) Hz using 1024 (^{1}H) × 80 (^{15}N) × 256 (^{1}H) data points. To obtain $^3J_{HNH\alpha}$ coupling constants, a 3D HNHα experiment (*27*) was carried out using spectral windows of 9124 (^{1}H) × 3125 (^{15}N) × 9124 (^{1}H) Hz for 1024 (^{1}H) × 32 (^{15}N) × 256 (^{1}H) data points. In all of the experiments, water suppression was obtained with the Watergate sequence (*28*).

On the apo form, ^{15}N-HSQC, ^{15}N-TOCSY-HMQC, ^{15}N-NOESY-HSQC, and 2D NOESY spectra were recorded in the same conditions as those for the copper-loaded sample.

For all of the spectra, shifted sine bell functions were used to apodize the data. The data were zero-filled once in indirect dimensions before Fourier transformation.

Constraints Used in the Structure Calculations. The peaks used for the structure calculations were integrated in the 2D NOESY and 3D ^{15}N-NOESY-HSQC experiments. The intensities of dipolar connectivities were converted into upper distance limits to be used as input for structure calculations by using the approach provided by the program CALIBA (*29*). All of the NOE cross peaks were divided into different classes according to the nature of the protons involved, and each class was calibrated independently. The calibration curves were adjusted iteratively as the structure calculations proceeded. Stereospecific assignments were obtained using the program GLOMSA (*29*).

$^3J_{HNH\alpha}$ coupling constants were obtained from the ratio between the intensity of the diagonal peak and that of the cross peak of the HNHα map. The scalar coupling constants were correlated to the backbone torsion angles ϕ by means of the appropriate Karplus curve (*27*). For $^3J_{HNH\alpha}$ values of >7.5 and <5 Hz, the ϕ angle was assumed to be between $-155°$ and $-80°$ and between $-70°$ and $-30°$, respectively. ψ torsion angle constraints, for residue $i-1$, were determined by measuring the ratio of the intensity of the $d_{\alpha N}(i-1, i)$ and $d_{N\alpha}(i, i)$ NOEs found on the ^{15}N plane of residue i in the ^{15}N-NOESY-HSQC spectrum (*30*). For ratios smaller than 1 and larger than 1, the ψ angles were assumed to be between 60° and 175° and between $-85°$ and $-10°$, respectively. ϕ and ψ torsion angles were used as constraints in the structure calculations.

Structure Calculations. Structure calculations were performed using the program DYANA (*31*). The copper ion was included in the calculations using special linkers (pseudoresidues) called LLN, as previously described (*32*). Then, the sulfur atoms of Cys13 and Cys16 were linked to the metal ion through upper-distance limits of 2.35 Å. This approach does not impose any fixed orientation of the ligands with respect to the copper. To prove the conformation of the coordination site of copper(I), calculations were first performed without the copper ion.

A total of 500 random conformers were annealed in 10 000 steps using the above constraints. The 30 conformers with the lowest target function constituted the final family.

Refinement was performed as restrained energy minimization (REM) with the Sander module of the program AMBER 5.0 to each member of the family (*33*). The force-field parameters for the copper(I) ion are taken as those in similar systems (*32*). The electrostatic potentials mapped onto the surface of the proteins were calculated with the program MOLMOL (*34*).

The program CORMA (*35*), which is based on relaxation matrix calculations, was used to back calculate the NOESY cross peaks from the structure at various stages of its calculation, to assign a few more NOE cross peaks between already assigned resonances, and to qualitatively assess the whole NOESY map assignment on the basis of the agreement between the NOEs calculated on the final structures and those from the experiment.

The quality of the structure was checked through the Ramachandran analysis using the program PROCHECK-NMR (*36*).

RESULTS AND DISCUSSION

The amino acid sequences of EhCopZ, yAtx1, and MerP were used in the search of nucleotide sequences codifying equivalent proteins in the genome of *B. subtilis*. Such sequences correspond to the only soluble copper chaperones for which the structure is known, either for the copper-depleted or the copper-loaded form or for both (*11, 37, 38*). The sequence *yvgY* in the *B. subtilis* genome, which was initially described as a probable mercuric ion-binding protein (*16*), was individuated. It encodes a protein which contains the MXCXXC metal-binding motif and which we then called BsCopZ. We called the corresponding *yvgY* gene the *bscopz* gene. The identity degree between protein BsCopZ and EhCopZ is 39%, which increases to 52% when conservative substitutions are considered. These values are higher than those found with the yeast copper chaperone yAtx1 (28% and 38%) and the mercury resistance periplasmic protein MerP (35% and 51%), as is reasonably expected (Figure 1).

The gene *bscopz* is immediately preceded, in the genome of *B. subtilis*, by a sequence encoding a putative metal-transporting ATPase (*16*). Because the genes which form the *cop* operon are organized in a similar way in other organisms where they have been determined, such as in *H. pylori* (*39*), the BsCopZ protein and the ATPase preceding in the operon, hereafter called BsCopA, could constitute the *cop* operon of *B. subtilis*.

Production, Purification, and Characterization of BsCopZ. Nucleotide sequences corresponding to the N-terminal part and those that are reverse and complementary to the C-terminal part of the protein BsCopZ were designed and used as primers in PCR reactions. A fragment of 249 bp was amplified from the genomic DNA of *B. subtilis*, cloned into pET21a (an expression vector), and used to transform *E.coli* BL21 (DE3)pLysS. The protein encoded by this DNA fragment was overexpressed in two different culture media

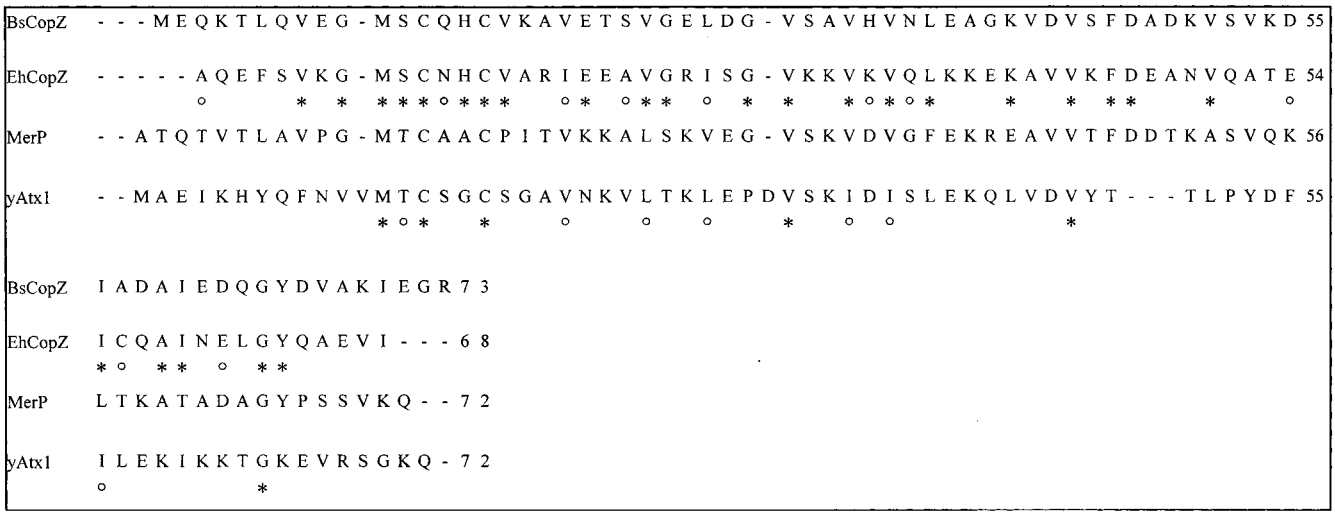

FIGURE 1: Sequence alignment of metallochaperones containing the MXCXXC motif and similar length in the amino acid sequence. Proteins (accession numbers): BsCopZ of *B. subtilis* (Z99121; this article); EhCopZ, copper-chaperone of *E. hirae* (Z46807); MerP, mercury chaperone of *Pseudomonas aeruginosa* (Z00027); yAtx1, copper chaperone of *S. cerevisiae* (L35270). The asterisks and circles indicate the conserved and conservatively substituted residues in EhCopZ (first row) and in all the three of the proteins with respect to the BsCopZ protein, respectively.

(LB and M9), yielding approximately 90% of recombinant protein in solution and 10% in inclusion bodies into the cytoplasm.

NMR Spectra and Resonance Assignment of BsCopZ−Cu-(I). The ^{1}H and ^{15}N NMR spectra of the copper-bound species show a good dispersion of the signals, indicating that essentially the entire protein is in a folded state. Furthermore, the spectra indicate the presence of a single well-defined species with the exception of few residues (see later).

The assignment of the protein signals was performed on the ^{15}N-labeled sample from the analysis of the ^{15}N-HSQC, 3D ^{15}N-TOCSY-HMQC, 3D ^{15}N-NOESY-HSQC, and 2D NOESY and TOCSY maps. The 2D TOCSY and ^{15}N-TOCSY-HMQC experiments were used to identify the spin patterns of the intraresidue connectivities of the amino acids. Then, the sequential assignment was performed using the interresidue connectivities available from sequential characteristic NOESY peaks in the 2D and 3D maps. The 3D experiments allowed us to solve problems with cross-peak overlap that, for some regions of the 2D spectra, are quite extensive. About 93% of the total proton resonances were assigned. All of the expected 72 peptidic NH cross peaks were identified in the ^{15}N-HSQC spectra. The resonance assignments are reported as Supporting Information. A strong NH cross peak is present in the spectrum, which could correspond to an acetylated Lys ϵ-NH and is tentatively assigned to Lys18 (see Experimental Procedures).

Two sets of signals with an intensity ratio of 2:1 can be detected for both the ^{15}N and ^{1}H resonances of the NH amide groups for seven residues, from 12 to 18, in the ^{15}N-HSQC map of the labeled sample. The two forms are called A and B hereafter; both sets of signals were analyzed. On the contrary, an unlabeled protein shows only one set of resonances for all of the residues, which corresponds to the B form. The presence of two conformations, therefore, is related to the culture medium in which bacteria are grown to produce the protein, which suggests the presence of some amino acid modification during the cell growth in stress conditions. This behavior has been reported in a variety of cases for protein produced by bacteria grown in different conditions, for example, in minimal media (*20*, *40*). In the following analysis, the structure of the A form is determined in the presence of some apo BsCopZ (because the protein could not be saturated with copper) and some B form. However, the two A and B forms have the same conformation, within the resolution of the present structure. Indeed, all of the residues, and in particular those around Lys18 and Lys18 itself, experience the same NOEs in the two forms. The species distribution was taken into account when evaluating the NOE intensities to be converted into distance constraints.

Structure Calculations and Refinement. A total of 1713 upper-distance limits (of which 1508 were meaningful), 38 dihedral ϕ angle constraints, and 48 ψ angle constraints were measured and used in the structural calculations with the program DYANA (*31*). The average number of meaningful structural constraints per residue is 20.7. The number of experimental NOEs per residue, subdivided according to their class, is reported in Figure 2A. A total of 15 stereospecific assignments were obtained through the program GLOMSA (*29*).

A family of 30 conformers was obtained with an average target function of 0.78 ± 0.07 Å^2 (the best structure has a target function of 0.63 Å^2) and an rmsd value, to the mean structure over residues 4−70, of 0.33 ± 0.06 Å for the backbone and 0.83 ± 0.06 Å for the heavy atoms. This family of conformers was then refined by performing restrained energy minimization on each member. The final family is characterized by an average rmsd value to the mean structure (for residues 4−70) of 0.32 ± 0.06 Å for the backbone and 0.85 ± 0.07 Å for the heavy atoms and a penalty function for the distance constraints, averaged over the family, of 0.31 Å^2. In Figure 2B, the rmsd values per residue of the final REM family are reported. The NMR solution structure is shown in Figure 3. A statistical analysis of the conformers family and of the average structure is reported in Table 1.

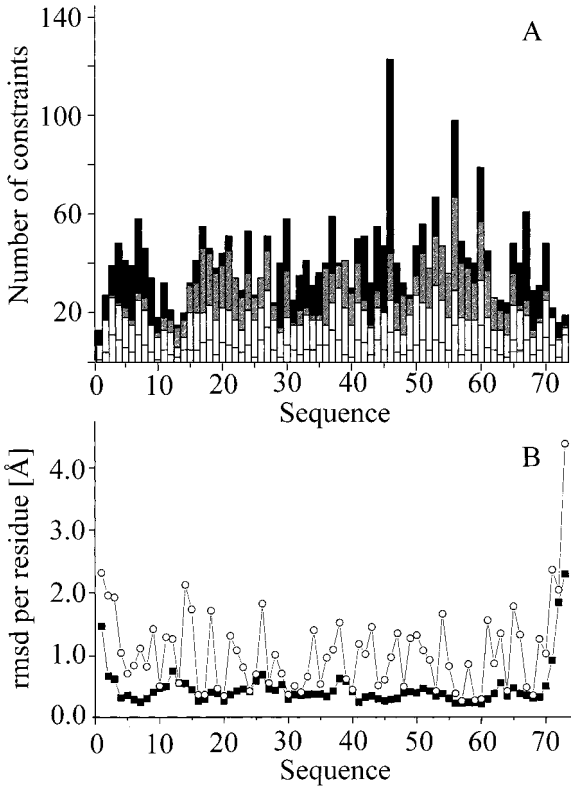

FIGURE 2: (A) Number of intraresidue (white), sequential (light gray), medium-range (gray), and long-range (black) NOEs per residue in *B. subtilis* CopZ−Cu(I). (B) Backbone (■) and heavy atom (○) rmsd values per residue for the family of 30 conformers with respect to the average structure of BsCopZ−Cu(I).

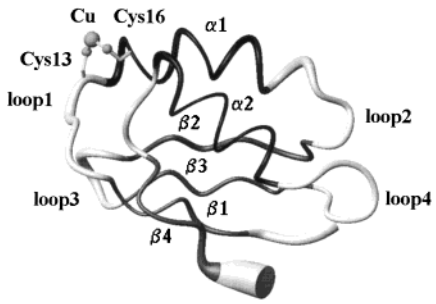

FIGURE 3: Tube representation of the family of 30 lowest target function conformers of *B. subtilis* CopZ−Cu(I) obtained with DYANA calculations and refined with restrained energy minimization calculations. Structures were aligned by a best-fit superimposition of C, Cα, and N atoms of residues 4−8 and 15−70. The elements of the secondary structure are indicated with a black color for the α helices and a gray color for the β sheets.

Description of the Structure. The elements of the secondary structure were identified through the analysis of the structure with PROCHECK-NMR (*36*) and of the pattern of backbone NOEs (*41*). The BsCopZ−Cu(I) fold consists of a twisted four-stranded antiparallel β sheet (4−8, 30−36, 41−46, and 67−71) and two α helices (14−24 and 54−63) that are both located on the same side of the β sheet (Figure 3); the secondary structure reported is related to the energy-minimized mean structure.

The secondary structure is very well-defined. The average rmsd values to the mean structure for the protein segments of the four-stranded antiparallel β sheet are 0.26 ± 0.06 and 0.79 ± 0.09 Å for the backbone and all of the heavy atoms, respectively, and the rmsd values for the two α helices are 0.22 ± 0.06 and 0.84 ± 0.08 Å for the backbone and the heavy atoms, respectively. This indicates that the β sheet and the α helices are characterized by a lower disorder than that of the loops connecting them. Other than the N- and C-terminal parts, the most disordered region is loop 1 (residues 9−14), which contains one copper ligand (Cys13), and is close to the second one (Cys16), which is located in helix α_1. The global average rmsd values of the family for the backbone and the heavy atoms decrease to 0.31 ± 0.06 Å and 0.83 ± 0.07 Å, respectively, when they are calculated without residues 9−14.

Table 1: Statistical Analysis of the Restrained Energy Minimized (REM) Family and of the Mean Structure of CopZ−Cu(I) from *B. subtilis*[a]

	REM (30 conformers)	⟨REM⟩ (av)
RMS Violations per Experimental Distance (Å) and Angle (deg) Constraints[b]		
intraresidue (194)	0.0238 ± 0.0025	0.0247
sequential (441)	0.0101 ± 0.0012	0.0097
medium-range[c] (401)	0.0087 ± 0.0021	0.0116
long-range (472)	0.0128 ± 0.0011	0.0124
total (1508)	0.0133 ± 0.0009	0.0138
backbone torsion angle φ (38)	0.67 ± 0.47	0.6289
backbone torsion angle ψ (48)	0.22 ± 0.17	0.0000
Average Number of NOE Violations per Structure		
intraresidue	13.6 ± 2.0	14.0
sequential	10.1 ± 2.3	9.0
medium-range[c]	5.1 ± 1.6	6.0
long-range	11.6 ± 2.0	12.0
total	40.5 ± 4.4	41.0
backbone torsion angle φ (38)	1.1 ± 0.8	1.0
backbone torsion angle ψ (48)	0.3 ± 0.2	0.0
violations of NOE larger than 0.3 Å	0	0
violations of NOE between 0.1 and 0.3 Å	7.7 ± 2.3	8.0
average NOE penalty function (Å²)	0.31 ± 0.04	0.32
average BB torsion angle penalty function (deg²)	1.49 ± 0.54	1.02
Structural Analyses		
% of residues in disallowed regions	0.1	0.0
% of residues in generously allowed regions	3.7	3.1
% of residues in allowed regions	24.5	21.5
% of residues in most favorable regions	71.6	75.4

[a] REM indicates the energy-minimized family of 30 structures, and ⟨REM⟩ is the energy-minimized average structure obtained from the coordinates of the individual REM structures. [b] Number of experimental constraints for each class is reported in parentheses. [c] Medium-range distance constraints are those between residues (i, $i+2$), (i, $i+3$), (i, $i+4$), and (i, $i+5$).

The conformation of the copper ligands (two cysteines) was also determined using only experimental NOEs and without the bonds to the copper ion. The two cysteines have the same conformation, and both of the sulfur atoms are oriented toward the same point (i.e., toward the location occupied by the copper ion) in the calculations with or without bonds to copper ion, thus indicating that the cysteine conformation is well-defined and correctly defined by the experimental structural constraints.

The Cys−Cu(I)−Cys bonds make an angle of 115 ± 26°, consistent with the occurrence of three coordination for the copper. The latter condition has been recently observed for BsCopZ−Cu(I) on the basis of EXAFS measurements (*42*) and has also been found for the yAtx1−Cu(I) system (*14*). The third donor atom is likely to be a sulfur atom from an exogenous moiety, such as DTT, which is added to the solution to maintain anaerobic conditions.

The metal-binding loop (loop 1) connects helix α_1 and sheet β_1. In addition to one copper(I)-binding cysteine, loop 1 also contains a histidine residue, His15, whose side-chain ring is pointing toward the solvent. In the present system containing Cu(I), the side chain of His15 is defined by 10 NOEs, which definitely rule out a possible interaction with the metal ion. This residue, however, might be a potential metal ligand if it reorients toward the protein interior. It might also have a role in assisting the protein in its copper-transferring function because it can interact with the incoming or leaving metal ion in the metal-transfer process. Besides the two cysteines, potential donor atoms, if some structural rearrangement occurs, are OH of the Tyr65 side chain (Cu−OH = 8.85 Å), Nϵ of the side chain of Gln63 (Cu−Nϵ = 9.63 Å) that belong to loop 5, and Sγ of Met11 (Cu−S = 5.41 Å). Particularly, the latter residue is quite close to copper and also is within hydrophobic contact with Tyr65. Within 10 Å from the copper ion, a few other relevant residues are present. One conserved residue is Ser12 whose side-chain OH group is detectable by NMR. This indicates that it has a slow exchange with the bulk water as a consequence of the establishment of a hydrogen bond with the carboxylate group of Glu38.

Comparison with the Reduced Apo Forms. The ^{15}N-labeled apoprotein was also characterized through ^{15}N-HSQC, and the backbone resonances were assigned through 3D ^{15}N-TOCSY-HMQC and 3D ^{15}N-NOESY-HSQC spectra. The ^{15}N-labeled apoprotein displays A and B forms in the same ratio as the copper-loaded protein. The apoprotein shows well-dispersed spectra similar to those of the copper-loaded form. In the ^{15}N-HSQC spectrum, the NH backbone signals of all of the residues, with the exception of those of Cys13 and His15, were assigned. The observed ^{15}N- and ^{1}H-backbone chemical shifts are very close to those in the copper-loaded protein. The backbone NOE patterns obtained through the analysis of ^{15}N-NOESY-HSQC confirm the presence and the extent of all of the elements of the secondary structure of BsCopZ−Cu(I), with the exception of helix α_1. The NOEs are also essentially the same for all of the loops except for loop 1. Indeed, the observed NOEs of loop 1 and of the N-terminal part of helix α_1 are few. The comparison with the copper-loaded protein indicates that most of the expected NOEs are below the threshold of detectability. A decrease of NOE intensities could be due to mobility, which leads to both a signal broadening of side-chain nuclei resonances and to subnanosecond movements. The further investigation of this aspect is not pursued here. The situation is similar to that of yAtx1, whose apo form shows few or no NOEs for the residues constituting loop 1 and the N-terminal part of helix α_1 (*37*), and that of the fourth metal-binding domain of Menkes copper-transporting ATPase, in which the metal-binding loop is better defined in the Ag(I)-loaded protein with respect to the reduced apo form. Indeed, in the latter state, the first cysteine experiences a disordered conformation, and the NH resonances of two other residues in the loop are not detected (*43*).

EhCopZ shows a relatively ordered conformation only for the apo form; therefore, the structure was solved only in the absence of copper(I) (*11*). The residue identity between the BsCopZ and EhCopZ proteins is 39%. The global-fold and the secondary-structure elements are maintained, as shown in Figure 4. The two structures have an rmsd for the backbone of 2.06 Å, excluding the residues in the copper-binding loop, and a backbone rmsd of 1.13 and 1.60 Å for the β sheet and the α helices, respectively.

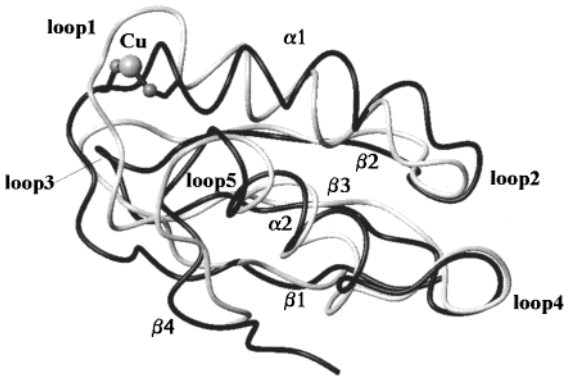

FIGURE 4: Comparison of the BsCopZ−Cu(I) structure (black), with the apo form of the EhCopZ (gray; *11*).

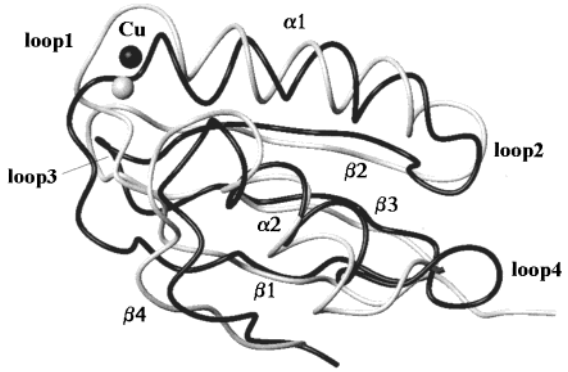

FIGURE 5: Comparison of the BsCopZ−Cu(I) structure (black) with the yAtx1−Cu(I) (gray; *37*).

Differences are observed in the metal-binding loop (loop 1), which may be related to the structural changes occurring upon copper binding, and on the loops 3 and 5 that are both located on the same side as loop 1. Loop 1 assumes a more external conformation in the copper-free protein, being extensively projected toward the solvent (Figure 4). A similar movement in the metal-binding loop was suggested also for yAtx1 on the basis of changes in a few of the NOEs involving the side chains (*37*). A slightly different orientation of helix α_1 is found for BsCopZ−Cu(I) with respect to the EhCopZ apoprotein. In the former present structure, the conformation of helix α_1 is well-defined by 91 NOEs with protons not belonging to the sequential helix.

Comparison with yAtx1−Cu(I), Ccc2−Cu(I), and MerP−Hg(II). It may be appropriate to compare the present structure also with that of the other homologous proteins mentioned previously and with the fourth domain of the Menkes copper-transporting ATPase coordinated with the Ag(I) ion (*43*). The global foldings of the yAtx1−Cu(I) protein and BsCopZ−Cu(I) are very similar, with the residue identity being 28% (Figure 5). The rmsd value for backbone over the aligned region of the two proteins, with the exclusion of loop 1, is 2.79 Å, dropping to 1.54 Å for the α-helix regions. The relatively large rmsd, despite the same fold, could be due to four insertions in regions close to β-strand elements.

Also, the N-terminal soluble domain of the Cu−ATPase Ccc2 (*15*) and MerP (*38*) exhibit the same fold as that of BsCopZ (Figure 6). The rmsd values for the backbone with

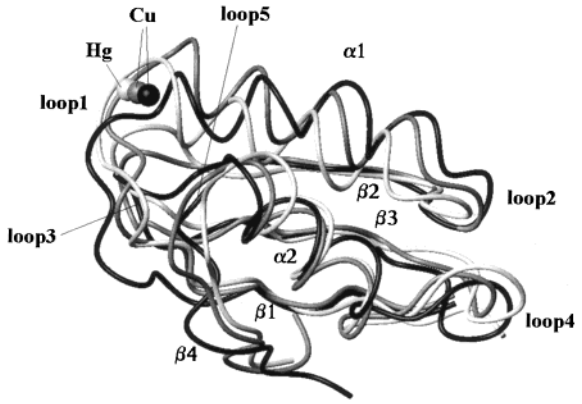

FIGURE 6: Comparison of the structure of BsCopZ−Cu(I) (black) with that of the soluble domain of the Cu−ATPase Ccc2−Cu(I) (gray; *15*) and with MerP−Hg(II) (light gray; *38*).

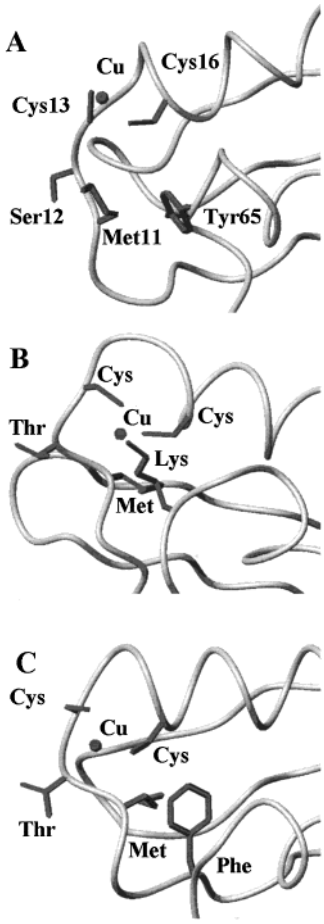

FIGURE 7: Comparison of (A) the active site of BsCopZ−Cu(I) with that of (B) yAtx1−Cu(I) (*37*) and with that of (C) Ccc2−Cu(I) (*15*).

the exclusion of loop 1 between BsCopZ−Cu(I) and Ccc2−Cu(I) and MerP−Hg(II) are 1.84 and 2.67 Å, respectively.

A comparison of the metal-binding region of BsCopZ−Cu(I), yAtx1−Cu(I), and Ccc2−Cu(I) is shown in Figure 7. It can be seen that loop 1 has a different conformation in the three proteins but that the two metal-coordinating cysteines are pointing toward the copper ion in a similar way. The coordination angle is also the same for the three systems (115 ± 26° for BsCopZ, 120 ± 40° for yAtx1, and 120 ± 40° for Ccc2), indicating a similar coordination geometry which may be three-coordinated by the participation of an exogenous ligand, as suggested by EXAFS measurements for BsCopZ (*42*) and yAtx1 (*21*). Significantly, Met11, which is very highly conserved in several classes of these metal-chaperone proteins, moves much closer to the copper ion in BsCopZ−Cu(I) than in yAtx1−Cu(I) and in Ccc2−Cu(I). A similar behavior occurs in several metal-chaperone proteins, for example, in the Menkes copper-transporting ATPase−Ag(I) domain where none of the sulfur atoms of two methionine residues, one present in loop 1 and the other in loop 5, are in a position suitable for coordination to the metal ion already bound to the two cysteine sulfur atoms (*43*). This residue is close and makes hydrophobic contacts with a Phe/Tyr in loop 5 (Tyr65 in the present system) in bacterial copper chaperones and in the soluble domains of Cu−ATPase, while in yAtx1 and eukaryotic copper chaperones, a lysine is invariably found in position 65. When an aromatic residue is present in position 65, it is hydrophobically interacting with Met11. It has also been proposed that this hydrophobic contact contributes to maintenance of the correct conformation of the metal-binding loop (*44*).

In yAtx1, Lys65 has been proposed to have a role in the metal-transfer process (*44, 45*) because its proximity to the copper ion could protect it from possible oxidants present in solution, in addition to the stabilization of the overall negative charge of the Cu(I)−bisthiolate center. In the apo form, Lys65 has moved away from the metal center, thus producing a translation of the entire helix α_2 (*37*).

In the present BsCopZ−Cu(I) protein, Met11 is much closer to the copper ion, still maintaining its hydrophobic interaction with Tyr65, which follows the translation of Met11. Met11 together with Tyr65 might play, in BsCopZ, a role similar to that proposed for Lys65 in yAtx1 (*44*) (i.e., the protection of the Cu(I) from possible oxidants present in solution and of the cysteine sulfur from the solvent. It is tempting to propose that Met11, highly conserved and close to copper but whose coordination to copper is definitely ruled out, plays a role of gate for access to the copper site.

Comparison with Copper-Loaded EhCopZ. The copper-loaded EhCopZ reduced with $Na_2S_2O_4$ tends to aggregate and to precipitate (*11*). Our procedure is that of using DTT for reduction. Such a reductant possibly binds copper, as indicated by XAS studies on these samples showing a copper coordination with three sulfur atoms (*46, 47*). Indeed, the S−Cu(I)−S angle is 115 ± 26°. EXAFS studies have also shown that both the present protein and other copper chaperones (*46*−*48*) in the absence of DTT still maintain three coordination with a fourth heavy atom at a larger distance. This situation is reached through dimerization (*46*−*48*). A dimerization process slow on the NMR time scale was suggested to account for the data published on copper-loaded EhCopZ (*11*).

Structural Features and Biological Process. The comparison of the present structure with others of related proteins shows that all of these metallochaperone proteins experience the same fold and that meaningful differences, in any case minor, essentially involve only the metal-binding loop, the relative position and orientation of some helices which point to the metal-binding site, and the loop which is located close to the metal-coordination site.

The orientation of helix α_1 as well as the conformation of loop 1 and possibly loop 5 might be related to the efficiency

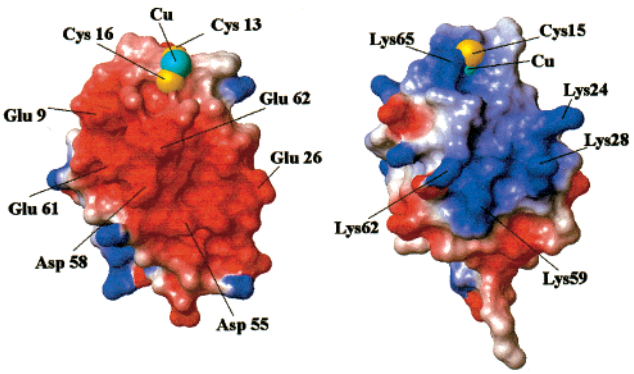

FIGURE 8: Electrostatic potential surface on the copper-containing protein face in BscopZ−Cu(I) (left panel) and in yAtx1−Cu(I) (right panel).

of the metal-transfer process and to the specificity in molecular recognition with the partner.

The copper site is well-ordered and well-defined with the copper ion possibly in three coordination. It is also experiencing a hydrophobic protection by a Met/Tyr moiety, which sizably reduces its solvent accessibility. The three coordination and this partial protection might be relevant for determining the correct conformation at the copper site for the incoming copper-receiving protein. Indeed, it has been noted in EhCopZ that the copper form tends to aggregate with possible interactions with the copper-binding loops (*11*). In that protein, the potential third copper ligand, DTT, was not present in solution; therefore, the copper ion was more accessible.

The conformation of the external loops around the copper site and the nature of the residues present there determine the electrostatic potential on the protein sample. The distribution of charged residues on the surface of BsCopZ−Cu(I) produces two negative-charged faces on the molecule. The first region contains a cluster of four glutamic acid residues (9, 26, and 61−62) and two aspartic acid residues (55 and 58), most of which are located in helix α_2. The second region is formed by three glutamic acid residues (21, 26, and 38) and one aspartic acid residue (43) and is located opposite to helix α_2. Figure 8 shows the electrostatic potential surface of BsCopZ−Cu(I), on the protein side containing the copper ion, together with the electrostatic potential surface of the same side on the protein of yAtx1. It is clear that the two proteins have an opposite electrostatic potential, with that of yAtx1 being positive. The partner of yAtx1, Ccc2, has a negative surface potential, which makes the electrostatic energy the major driving factor which brings the two proteins to interact with each other. For EhCopZ, which has a negative surface, a partner with a positive electrostatic potential surface would be expected.

BsCopZ is encoded in the putative *cop* operon of *B. subtilis* together with a putative copper-transporting P-type ATPase CopA, BsCopA (*16*). Given the high homology between Ccc2 and BsCopA (53% of homology) and the structural similarity between Atx1 and BsCopZ, a transfer process, and therefore a protein−protein interaction similar to that already suggested for Ccc2 and yAtx1, might also be expected between BsCopA and BsCopZ. To confirm this hypothesis, the determination of the structure of BsCopA is in progress in our laboratory.

CONCLUDING REMARKS

In this paper, the gene codifying a possibly copper chaperone in the genome of *B. subtilis* has been identified on the basis of the homology with the sequences of similar proteins in other organisms and of the presence of the "classical" MXCXXC consensus sequence. The corresponding protein (BsCopZ) has been expressed and produced, and its structure, in the copper-bound form, has been determined. Presumably, it has a function similar to that of analogous proteins in eukaryotes, of which yAtx1 is an example. However, the overall charge is negative and opposite to that of the analogous eukaryotic proteins. The BsCopZ apoprotein has a behavior similar to that of apo yAtx1, because both proteins (yAtx1 and BsCopZ) maintain a tertiary structure and show disorder in the copper-binding loop. The analogous protein from *E. hirae* is reported to be stable only in the apo form, while it has some tendency to aggregate in solutions lacking possible copper-coordinating molecules, such as DTT. In the present protein, the copper-coordination site has been characterized in detail, providing information on the geometrical features and on the residue interactions relevant for the metal-transfer process.

The overall folding of these chaperones as well as that of the partner Cu−ATPases is remarkably similar. A gene of a protein, which we call BsCopA, has been found in the *cop* operon of *B. subtilis* and is under investigation as a possible partner.

ACKNOWLEDGMENT

We thank Prof. G. Moneti for recording mass spectra and Prof. Piccardi for the recording of the absorbance spectra, Dr. Elena Molteni for help in the analysis of some genome organisms, and Dr. Fabio Arnesano for the helpful discussion on the comparison with other structures.

SUPPORTING INFORMATION AVAILABLE

^{15}N and ^{1}H resonance assignments for the BsCopZ (conformation A) in 100 mM phosphate buffer (pH = 7.0) at 300 K and the experimental NOE intensity constraints with a list of stereospecific assignments, used for the structure calculations. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Rosenzweig, A. C., and O'Halloran, T. V. (2000) *Curr. Opin. Chem. Biol. 4*, 140−147.
2. Harrison, M. D., Jones, C. E., and Dameron, C. T. (1999) *J. Biol. Inorg. Chem. 4*, 145−153.
3. Askwith, C., and Kaplan, J. (1998) *Trends Biochem. Sci. 23*, 135−138.
4. Pena, M. M. O., Lee, J., and Thiele, D. J. (1999) *J. Nutr. 129*, 1251−1260.
5. Nishihara, E., Furuyama, T., Yamashita, S., and Mori, N. (1998) *Mol. Neurosci. 9*, 3259−3263.
6. Koch, K. A., Pena, M. M. O., and Thiele, D. J. (1997) *Chem. Biol. 4*, 549−560.
7. Valentine, J. S., and Gralla, E. B. (1997) *Science 278*, 817−818.
8. Morby, A. P., Hobman, J. L., and Brown, N. L. (1995) *Mol. Microbiol. 17*, 25−35.
9. Cooksey, D. A. (1994) *FEMS Microbiol. Rev. 14*, 381−386.
10. Odermatt, A., Suter, H., Krapf, R., and Solioz, M. (1992) *Ann. N. Y. Acad. Sci. 671*, 484−486.

11. Wimmer, R., Herrmann, T., Solioz, M., and Wüthrich, K. (1999) *J. Biol. Chem. 274*, 22597−22603.
12. Cobine, P., Wickramasinghe, W. A., Harrison, M. D., Weber, T., Solioz, M., and Dameron, C. T. (1999) *FEBS Lett. 445*, 27−30.
13. Odermatt, A., Suter, H., Krapf, R., and Solioz, M. (1993) *J. Biol. Chem. 268*, 12775−12779.
14. Pufahl, R., Singer, C. P., Peariso, K. L., Lin, S.-J., Schmidt, P. J., Fahrni, C. J., Cizewski Culotta, V., Penner-Hahn, J. E., and O'Halloran, T. V. (1997) *Science 278*, 853−856.
15. Banci, L., Bertini, I., Ciofi-Baffoni, S., Huffman, D. L., and O'Halloran, T. V. (2001) *J. Biol. Chem. 276*, 8415−8426.
16. Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursier, L., Brans, A., Braun, M., Brignell, S. C., Bron, S., Brouillet, S., Bruschi, C. V., Caldwell, B., Capuano, V., Carter, N. M., Choi, S. K., Codani, J. J., Connerton, I. F., Danchin, A., and et al. (1997) *Nature 390*, 249−256.
17. Marmur, M. J. (1961) *J. Mol. Biol. 3*, 208−218.
18. Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989) *Molecular cloning: a laboratory manual*, Cold Spring Harbor Laboratory Press, Plainview, NY.
19. Barker, P. D., Nerou, E. P., Freund, S. M. V., and Fearnley, I. M. (1995) *Biochemistry 34*, 15191−15203.
20. Smith, V. F., Schwartz, B. L., Randall, L. L., and Smith, R. D. (1996) *Protein Sci. 5*, 488−494.
21. Pufahl, R. A., Singer, C. P., Peariso, K. L., Lin, S.-J., Schmidt, P. J., Fahrni, C. J., Cizewski Culotta, V., Penner-Hahn, J. E., and O'Halloran, T. V. (1997) *Science 278*, 853−856.
22. Eccles, C., Güntert, P., Billeter, M., and Wüthrich, K. (1991) *J. Biomol. NMR 1*, 111−130.
23. Macura, S., Wüthrich, K., and Ernst, R. R. (1982) *J. Magn. Reson. 47*, 351−357.
24. Bax, A., and Davis, D. G. (1985) *J. Magn. Reson. 65*, 355−360.
25. Kay, L. E., Keifer, P., and Saarinen, T. (1992) *J. Am. Chem. Soc. 114*, 10663−10665.
26. Wider, G., Neri, D., Otting, G., and Wüthrich, K. (1989) *J. Magn. Reson. 85*, 426−431.
27. Vuister, G. W., and Bax, A. (1993) *J. Am. Chem. Soc. 115*, 7772−7777.
28. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−666.
29. Güntert, P., Braun, W., and Wüthrich, K. (1991) *J. Mol. Biol. 217*, 517−530.
30. Gagné, R. R., Tsuda, S., Li, M. X., Chandra, M., Smillie, L. B., and Sykes, B. D. (1994) *Protein Sci. 3*, 1961−1974.
31. Güntert, P., Mumenthaler, C., and Wüthrich, K. (1997) *J. Mol. Biol. 273*, 283−298.
32. Banci, L., Benedetto, M., Bertini, I., Del Conte, R., Piccioli, M., and Viezzoli, M. S. (1998) *Biochemistry 37*, 11780−11791.
33. Pearlman, D. A., Case, D. A., Caldwell, J. W., Ross, W. S., Cheatham, T. E., Ferguson, D. M., Seibel, G. L., Singh, U. C., Weiner, P. K., and Kollman, P. A. (1997) *AMBER 5.0*, University of California, San Francisco, CA.
34. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
35. Borgias, B., Thomas, P. D., and James, T. L. (1989) *Complete Relaxation Matrix Analysis (CORMA)*, University of California, San Francisco, CA.
36. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR 8*, 477−486.
37. Arnesano, F., Banci, L., Bertini, I., Huffman, D. L., and O'Halloran, T. V. (2001) *Biochemistry 40*, 1528−1539.
38. Steele, R. A., and Opella, S. J. (1997) *Biochemistry 36*, 6885−6895.
39. Bayle, D., Wängler, S., Weitzenegger, T., Steinhilber, W., Volz, J., Przybylski, M., Schäfer, K. P., Sachs, G., and Melchers, K. (1998) *J. Bacteriol. 180*, 317−329.
40. Wilkins, M. R., Gasteiger, E., Gooley, A. A., Herbert, B. R., Molloy, M. P., Binz, P. A., Ou, K. S., Bairoch, A., Williams, K. L., and Hochstrasser, D. F. (1999) *J. Mol. Biol. 289*, 645−657.
41. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.
42. Banci, L., Bertini, I., Del Conte, R., Mangani, S., and Meyer-Klaucke, W. (2002) *X-ray Absorption Spectroscopy Study of CopZ, a Copper Chaperone in Bacillus subtilis*, submitted for publication.
43. Gitschier, J., Moffat, B., Reilly, D., Wood, W. I., and Fairbrother, W. J. (1998) *Nat. Struct. Biol. 5*, 47−54.
44. Arnesano, F., Banci, L., Bertini, I., Ciofi-Baffoni, S., Molteni, E., Huffman, D. L., and O'Halloran, T. V. (2002) *Metallochaperones and metal transporting ATPases: a comparative analysis of sequences and structures*, submitted for publication.
45. Portnoy, M. E., Rosenzweig, A. C., Rae, T., Huffman, D. L., O'Halloran, T. V., and Cizewski Culotta, V. (1999) *J. Biol. Chem. 274*, 15041−15045.
46. Srinivasan, C., Posewitz, M. C., George, G. N., and Winge, D. R. (1998) *Biochemistry 37*, 7572−7577.
47. Eisses, J. F., Stasser, J. P., Ralle, M., Kaplan, J. H., and Blackburn, N. J. (2000) *Biochemistry 39*, 7337−7342.
48. Graden, J. A., Posewitz, M. C., Simon, J. R., George, G. N., Pickering, I. J., and Winge, D. R. (1996) *Biochemistry 35*, 14583−14589.

BI0112715