Identification and characterisation of a novel family of copper storage proteins from *Methylosinus trichosporium* OB3b

Semeli Platsaki

A Thesis Submitted for the Degree of Doctor of Philosophy

November 2014

Institute for Cell and Molecular Biosciences

Newcastle University

Declaration

I certify that this thesis contains my own work, except where acknowledged, and that no part of this work has been submitted in support of an application for other qualifications at this or any other institution.

Abstract

Methane oxidizing bacteria (MOB) use methane as their main source of carbon and energy. The main methane oxidizing enzyme in MOB is the copper-containing particulate methane monooxygenase (pMMO), a rare example of cytoplasmic copper enzyme. Some 'switchover' strains are capable of differentially expressing pMMO as well as a soluble iron-containing form (sMMO), and the switchover is regulated by copper. MOB secrete methanobactin (mb) which mediates copper uptake and is internalized in the cytoplasm. Despite this pathway for copper import, as well as copper regulating components such as CopA, CopZ and CueR being present in MOB, little is known on how these bacteria handle the large amounts of copper required for methane oxidation by pMMO.

Through metalloproteomic analysis of soluble extracts from the switchover MOB *M. trichosporium* OB3b a large number of soluble copper pools were visualized and a novel copper protein, Csp1, was identified. Two more homologues, Csp2 and Csp3, were identified in *M. trichosporium* OB3b through bioinformatics. In vitro characterization of Csp1 and the homologue Csp3 showed these proteins are tetramers of 4-helix bundles that bind 13 and 18 Cu(I) ions per monomer, respectively, all of which are stored inside the core of the 4-helix bundle and are coordinated mostly by Cys residues. Csp1 binds tightly at least 10 Cu(I) ions whereas Csp3 has an average Cu(I) affinity at the order of 10^{17} M^{-1}. Csp1 and Csp3 do not remove Cu(I) from Cu(I)-mb, however it is likely that apo-mb, which removes Cu(I) from these proteins, notably at very different rates, transports Cu(I) to pMMO. Csp1 is thought to be exported from the cytosol potentially to the intra-cytoplasmic membranes, where pMMO is localised, to store copper for the enzyme. Csp3 is thought to be cytosolic and either sequesters copper to prevent copper-induced toxicity or, more likely, supplies copper to unknown cytosolic copper enzymes, consistent with the large number of soluble copper pools visualized in the organism. Csp1 protein homologues are present in other bacteria, including MOB, while homologues of the cytosolic Csp3 are widely distributed in members of all major bacterial phyla. The presence of Csp3 in the bacterial cytosol implies either a function as a defense mechanism against copper-induced toxicity or more likely copper storage for supply to cytosolic copper enzymes, yet to be identified. The latter possibility challenges the present model according to which bacteria do not have a cytosolic requirement for copper.

Acknowledgements

I warmly thank my supervisor, Professor Christopher Dennison for giving me the opportunity to come to Newcastle for a Phd project that has been a learning experience on many levels and enabled me to become involved and contribute in a research area of great interest. Chris has scientifically guided and supported me throughout the degree with persistence and constructive feedback. The advice and support Chris has offered throughout the project and also during the writing of this thesis has been of critical importance for the completion of this work. I am also grateful to Dr Abdelnasser El Ghazouani who helped introduce me to the experimental methods necessary for the beginning of my project. I deeply thank Dr Kevin Waldron for his scientific support, his enthusiastic encouragement and a lot of helpful discussions on an informal level. I am thankful to Kevin for reading parts of this thesis and providing useful directions when writing reached certain bottlenecks. I must also thank Dr Arnaud Baslé for collecting crystallographic data and solving the structures of the Csp1 and Csp3 proteins. Arnaud's contribution to this thesis has been pivotal and I have appreciated the time and effort he took to introduce me to the intricacies of protein crystallography with great enthusiasm. Thanks also go to Owen Burbidge for performing the experiment shown in Figure 2.15, Dr Neil Patterson for performing the experiment shown in Figure 3.26, and Dr Joe Gray who has helped me with HPLC experiments and performed mass spectrometry for protein identification.

My lab mates deserve a very warm acknowledgement not only for their professional support but for also becoming great friends. Dr Stephen Allen has assisted me with numerous experiments and provided support with a great sense of humour and with being an excellent friend. Dr Nicolas Vita has been always willing to assist with experimental advice and a genuinely nice person to work and be friends with. Dr Dorota Kostrz has shown her interest in this work consistently throughout the years and has always maintained her belief in a positive outcome. Gianpiero Landolfi has been a wonderfully thoughtful lab mate and friend with constant words of encouragement and a lot of humour. Kerrie Brusby has been a source of extracurricular events and activities while Eliona Tsefou has selflessly provided her help on numerous occasions and the completion of this thesis would have been much harder without her assistance. Last but not least, Jack Stevenson has been a great friend during my time at Newcastle, never losing his positive take on things, making me laugh and helping me have a good time outside of work.

Abbreviations, acronyms and units

A	absorbance
Csp	Copper storage protein
DNA	deoxyribonucleic acid
ET	electron transfer
LMCT	ligand to metal charge transfer
mb	methanobactin
MOB	methane oxidising bacteria
MT	metallothionein
pMMO	particulate methane monooxygenase
RNA	ribonucleic acid
sMMO	soluble methane monooxygenase
ε	extinction coefficient ($mM^{-1} cm^{-1}$)

Measurements and techniques

AAS	atomic absorption spectroscopy
CD	circular dichroism
ICP-MS	inductively coupled plasma mass spectrometry
MALDI-TOF MS	matrix assisted laser desorption ionization time of flight mass spectrometry
LC/MS/MS	liquid chromatography tandem mass spectrometry
OD	optical density
PCR	polymerase chain reaction
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electrophoresis
UV-VIS	ultraviolet-visible

Units

Å	angrstom
°C	degrees Celsius
Da	Dalton
g	gram
h	hour

K	Kelvin
L	litre
M	mol$\times$dm^{-3}
min	minute
nm	nanometre
rpm	revolutions per minute
RT	room temperature
s	second
v/v	volume per volume
w/v	weight per volume

Selected chemicals and media

BCA	bicinchoninic acid
BCS	bathocuproine disulfonic acid disodium salt
DTT	dithiothreitol
EDTA	ethylenediaminetetraacetic acid
Hepes	4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid
IPTG	isopropyl β-D-1-thiogalactopyranoside
LB	Luria-Bertani medium
Mes	2-(N-morpholino)ethanesulphonic acid
NMS	nitrate minimal salts medium
Taps	N-tris(Hydroxymethyl)methyl-3-aminopropanesulfonic acid sodium-potassium salt
Tris	tris(hydroxymethyl)aminoethane

Amino acids and nucleic acids

Alanine	Ala	A	Methionine	Met	M
Cysteine	Cys	C	Asparagine	Asn	N
Aspartic acid	Asp	D	Proline	Pro	P
Glutamic acid	Glu	E	Glutamine	Gln	Q
Phenylalanine	Phe	F	Arginine	Arg	R
Glycine	Gly	G	Serine	Ser	S

Histidine	His	H	Threonine	Thr	T
Isoleucine	Ile	I	Valine	Val	V
Lysine	Lys	K	Tryptophan	Trp	W
Leucine	Leu	L	Tyrosine	Tyr	Y

Adenine	A	Guanine	G
Cytosine	C	Thymine	T

Table of Contents

Chapter 1:
Introduction

1.1 Biological functions of metals

The importance of metals for all forms of life is reflected by the large proportion of enzymes that utilize metals as cofactors (1, 2). The role these inorganic elements play in enzymes depends on whether the metal is redox active or not, in which case it acts to activate a substrate and/or stabilise an intermediate product of the catalytic reaction (2). Lewis acidity, i.e. how prone the metal is to accept a pair of electrons, is crucial for the functions the metal performs (2). Organisms have evolved to use certain metals based on their availability in different environmental niches and also their availability over time, which changed after the Great Oxidation Event (GOE) during which the presence of oxygen altered the solubility of certain metals over others (see section 1.2.1) (2, 3). Elements such as magnesium and calcium are abundant in bulk, whereas d block transition elements such as manganese, zinc, iron, copper, and also nickel and cobalt are used in smaller amounts and are referred to as trace elements (3).

Magnesium, calcium and zinc are metals that are redox inert, as they are abundant in a single oxidation state as divalents (2). Mg(II) is the most abundant metal in organisms, present at millimolar levels (4), and its large charge density, derived from the small radius of Mg(II) (0.72 Å), enables the electrostatic stabilisation of large negatively charged phosphate-containing macromolecules, such as DNA and RNA (2, 5). Moreover, Mg(II) can activate the P-O and C-O bonds of substrates, through its electrophilic nature (5). An example of this role, vital to all organisms, is the requirement of two Mg(II) ions for the catalytic activity of all DNA polymerases (4). In this case, a nucleotide-binding Mg(II) coordinates the three phosphate groups of the incoming dNTP while a catalytic Mg(II) coordinates the phosphate group of dNTP and the primer strand. Both Mg(II) ions are needed to induce the conformational changes necessary for the function of DNA polymerase and this is an example where magnesium is key in bringing together the catalytic groups, neutralise the negative charges of the reaction components and, therefore, stabilise the active site of the complex (6). Like Mg(II), Ca(II) also has the ability to polarise P-O and C-O bonds, through its electrophilic nature, however this element is used less often due to its smaller charge density, derived from its larger radius (1.00 Å) compared to Mg(II), and the consequent differences in coordination geometry (2, 7). While Mg(II) forms octahedral compounds, Ca(II) binds more ligands with irregular geometries, leading to use of this metal for its flexibility and ability to mediate signalling events. Phospholipase C is an example of an

enzyme that employs Ca(II) for its flexibility regarding coordination geometry that, in this example, varies between six and eight ligands during catalysis (8).

Zn(II) is less abundant in organisms than the above metals, at 0.1 mM levels (9, 10), and combines redox inactivity with stronger Lewis acid properties than Mg(II) and Ca(II) (2). Zn(II) catalyses hydrolysis directly or indirectly in hydrolytic enzymes such as peptidases, carbonic anhydrases, alcohol dehydrogenases and alkaline phosphatases and various models have been suggested for these reactions (11). On the other hand, zinc is also found in proteins that do not have enzyme activity, such as metallothioneins that are thought to be involved in transport and storage of Zn(II) (12). Zinc enables the correct folding of protein structures known as zinc fingers (13). In this case the role of Zn is structural and such examples include DNA and RNA polymerases as well as DNA repairing proteins (4). Manganese is a strong oxidising agent and is most stable as Mn(II), although the Mn(III) and Mn (IV) oxidation states are also observed. The oxidation chemistry of manganese is central in dioxygen production by photosystem II as well as in peroxide metabolism through enzymes like peroxidases, catalases and manganese-containing superoxide dismutases (4). The intermediate strength of Mn(II) as a Lewis acid, as well as a radius (0.83 Å) between those of Mg(II), Ca(II) and Zn(II), render Mn(II) interchangeable with these metals. In order to establish specificity for Mn(II) over the other metals, enzymes employ active sites that have a combination of hard and soft ligands, while cellular compartmentalisation ensures higher abundance of Mn(II) over other metals in certain locations to achieve enzyme selectivity (2, 14).

The biological functions of iron and copper are studied separately as they combine Lewis acid properties with redox activity (2-4). Iron, one of the most abundant transition metals with intracellular levels at 0.1 mM (9, 10), is mainly abundant as Fe(II) and Fe(III) and is found in a range of metal sites, such as haem groups, Fe-O-Fe and iron-sulfur clusters. The biological functions carried out by iron-containing enzymes include electron transfer, oxidation, dioxygen transport and nitrogen fixation (4). In iron-sulfur proteins, such as ferredoxins, iron is coordinated by S^- or S^{2-} ligands and the switch between Fe(II) and Fe(III) during electron transfer requires only minimal changes in the bond length (4, 15). Soluble methane monooxygenase is an example where Fe(II) is used in dioxygen activation and oxygen binding to the diiron center forms an intermediate Fe(III)-Fe(III)-peroxo complex (16, 17). A different category of iron proteins contains haem Fe and such proteins are involved in oxygen transfer, NO

binding, favoured by the porphyrin unit structure, and also ET and oxidation (4). Copper biochemistry takes advantage of the ability of both Cu(I) and Cu(II) to readily form complexes with organic molecules, while the Cu(I)/Cu(II) couple is used by enzymes for its adaptable redox potential (4). Copper is used by enzymes that mediate electron transfer as well as by oxidases and enzymes that bind dioxygen, such as haemocyanin and cytochrome oxidase. Copper is also found in enzymes that handle nitrogen oxides, like nitrous oxide reductase, and enzymes that protect cells from reactive oxygen species, like superoxide dismutase (4). Finally, metals like cobalt and nickel appear to be used mostly by early forms of life as their abundance is restricted. Cobalt is mainly used as a cofactor in vitamin B_{12} where it is used for its redox activity, and is also found in few bacterial porphyrins (2, 4, 18). Nickel is found in enzymes including urease, hydrogenase and some superoxide dismutases (4, 19).

1.2 Copper in biology

1.2.1 Copper biochemistry and bioavailability

Copper is a first row transition metal ($[Ar]3d^{10} 4s^1$), abundant in two stable oxidation states, Cu(II) ($[Ar]3d^9$) and Cu(I) ($[Ar]3d^{10}$) (4, 20). Copper is a Lewis acid and forms bonds with electron donor ligands, which act as Lewis bases, in order to form stable complexes. Cuprous Cu(I) is diamagnetic and favours the formation of tetrahedral complexes with donor ligands like thiols from Cys residues or thioether bonds from Met groups. Cupric Cu(II) is paramagnetic and forms mostly square planar complexes with ligands such as S from cysteine or methionine residues, N from His, and O donors from residues like Asp and Glu (4, 20). Although Cu(I) and Cu(II) form complexes with very different stereochemistry, the high electron affinity of the metal in both oxidation states renders the pair biologically relevant due to its adaptable redox potential that ranges between +0.2 mV and +0.8 mV (4).

The change in copper bioavailability, as a result of the Great Oxidation Event (GOE), had a dramatic impact on the utilisation of the metal by organisms. In the primitive form of the world, under anoxic and sulfidic conditions, copper was found as a Cu(I) sulfide, insoluble to water. The increase of atmospheric oxygen, during the GOE, 2.4-2.7 billion years ago, shifted the abundance of copper to the water-soluble Cu(II) form (21, 22). Concomitantly, organisms also faced the challenge of developing an oxygen metabolism as well as defence mechanisms against reactive oxygen species (21, 22). Due to its chemical properties as a redox-active metal, copper was incorporated in

enzymes involved in electron transfer, as part of the respiratory chain or in response to oxidative damage. Notably copper enzymes are known to have a broad range of reduction potentials, between +200 mV and +800 mV (20).

Electron transfer (ET) is an abundant role of copper in enzymes, in accordance with the flexible redox potential of the metal, and ET proteins are involved in a range of physiological functions, such as photosynthesis, denitrification and response to oxidative stress (23). There are two types of copper centers that can mediate ET reactions (mononuclear Type I sites and dinuclear Cu_A sites) and both are contained by a conserved protein structure known as the cupredoxin fold. The cupredoxin binding site favours neither the coordination geometry for Cu(II) (square planar) nor Cu(I) (tetrahedral). Instead, copper is usually coordinated in a distorted tetrahedral geometry, using S and N ligands provided by Cys, Met and His residues respectively, and the binding site shows minimal structural changes upon conversion between Cu(I) and Cu(II). These properties allow the protein fold to maintain a robust structure even without bound copper and act as a mobile electron scavenger. Examples of cupredoxins involved in these roles include plastocyanins and azurins (23).

Except for a role in enzymes that produce oxygen through photosynthesis, copper is also important in enzymes involved in oxygen management, examples of which are cytochrome oxidase and superoxide dismutase (4). Cytochrome oxidase is the final electron acceptor in the respiratory chain and contains two copper atoms, one of which is involved in electron transfer while the other promotes oxygen binding by the enzyme. Ultimately, cytochrome oxidase couples the reduction of oxygen to the proton transfer across the membrane, therefore acting as a proton pump. The translation of oxygen reduction to energy, which subsequently mediates proton flow, is the copper-dependent step of the process (4). Superoxide dismutase is part of the defence mechanism against reactive oxygen species (ROS) and evolved from the need of organisms to develop oxygen tolerance. This cytoplasmic enzyme, already existing in a Fe/Mn form before the GOE, developed to incorporate copper which is bound with very high affinity to the enzyme therefore achieving both minimal free copper concentrations in the cytoplasm and defence against oxygen species (4).

Our current knowledge of copper enzymes in bacteria is restricted to ten examples among which are the aforementioned plastocyanins, cytochrome c oxidase and

superoxide dismutase and also NADH dehydrogenase, nitrosocyanin, copper-containing nitrite reductases, tyrosinase, copper amine oxidases, copper-containing laccase and the particulate methane monooxygenase (24-33). Although this can be attributed to the lower complexity of bacteria compared to other organisms, the possibility of more copper enzymes yet to be discovered cannot be disregarded (22, 24). Interestingly, while some bacteria adapted to copper use as the bioavailability of the metal changed, others remained non-users. Regardless of this preference, all bacteria have developed systems for copper handling and transport while only one copper transport system is thought to be an importer. This indicates that copper transport systems are essential for copper efflux, in order to protect the organism against copper toxicity (22).

1.2.2 Copper toxicity

The same properties that make copper a bioelement employed for enzyme activity are also the reasons that cause copper toxicity, as the adaptable redox potential of the Cu(I)/Cu(II) couple can also perpetrate unwanted reactions. Although the exact mechanism for copper toxicity is not fully understood, the commonly suggested route involves the generation of reactive oxygen species (ROS) through the Fenton and Haber-Weiss reactions, shown in equations (i) and (ii) respectively (20, 21):

$$Cu(I) + H_2O_2 \rightarrow Cu(II) + OH^- + OH^\bullet \quad (i)$$

$$O_2^{\bullet-} + Cu(II) \rightarrow Cu(I) + O_2 \quad (ii)$$
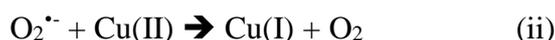
A second proposed mechanism for copper toxicity involves sulfhydryl depletion according to the reactions (iii) and (iv) (24):

$$2Cu(II) + 2RSH \rightarrow 2Cu(I) + RSSR + 2H^+ \quad (iii)$$

$$2Cu(I) + 2H^+ + O_2 \rightarrow 2Cu(II) + H_2O_2 \quad (iv)$$
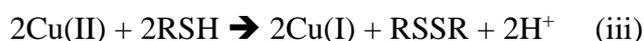
DNA has been regarded as the main target of oxidative damage (34) as well as other biomolecules like lipids and proteins. ROS production is thought to occur close to free copper ions, and the oxygen species cannot be scavenged by enzymes since they are extremely reactive and have a very short half-life (21). Extensive studies have suggested that copper-mediated oxidation of biomolecules results in the production of more ROS that, in turn, feed into the reaction cycle.

The model described above, however, has been questioned due to a number of observations on the nature of copper and its handling by organisms. To begin with copper is maintained at very low intracellular levels through tightly regulated homeostatic mechanisms. Moreover, copper has been shown to be less toxic under the aerobic conditions, required for all the above ROS generating reactions, while in recent cases *E. coli* was reported to be less prone to damage induced by H_2O_2 in the presence of copper (35) and DNA damage was not observed (35-37). This might be due to the cell compartmentalisation that keeps DNA protected in the cytosol where all copper is tightly controlled and/or due to the presence of ligands such as glutathione that bind copper and prevent it from binding inappropriately to biomolecules (36). Taking these points into consideration, new routes for copper toxicity have been suggested that do not involve oxidative stress.

Cu(I) is highly thiophilic and has been shown to disrupt iron metabolism, independently of oxygen, by disassembling solvent exposed [4Fe-4S] clusters in dehydratases, both in vivo and in vitro (36). In *E. coli* the iron-sulfur cluster in 6-phosphogluconate dehydratase is mismetallated by copper and, as a result, glucose catabolism is impaired. Another example in the same organism is isopropylmalate dehydratase, the inactivation of which impacts on biosynthesis of branched-chain amino acids (leucine, isoleucine and valine) resulting in auxotrophy (36). A separate study showed an additional target of Cu(I) toxicity is the heme-biosynthesis pathway (38). In this case, heme-containing enzymes such as catalases, peroxidases and nitric oxide reductases are inactivated thus rendering the bacterial cell susceptible to ROS and reactive nitrogen species (RNS) stress. The free labile Fe(II) that is released as a result of the [4Fe-4S] cluster can lead to further oxidative damage through iron Fenton reactions (39). Based on the above evidence, copper toxicity is not directly linked to oxidative damage but, instead, to the inactivation of iron-utilising enzymes, where the subsequent release of Fe(II) causes production of ROS as a side effect.

1.3 Bacterial copper homeostasis

In order to benefit from the use of copper and its redox active properties in enzymes but also protect themselves against copper toxicity, bacteria developed homeostatic mechanisms. These consist of a network of transporters, chaperones and regulators that ensure copper is delivered with precision to copper enzymes and any excess is transported outside the cell (40, 53, 64). The physiology and compartmentalisation of

the bacterial cell dictates the distribution of copper, which is more tolerated in the oxidising environment of the periplasm compared to the reducing cytoplasm where it is generally accepted that free copper is effectively zero (40, 53). It is interesting to note here that the different compartments of the bacterial cell also influence copper biochemistry, through the copper binding motifs favoured in each location (40). A common copper-binding motif, abundant in sensing and transporter proteins, is Cys-X-X-Cys and is found in the reducing cytoplasmic environment where disulfide bond formation is not favoured (40). This motif favours low coordination numbers and therefore helps the protein site discriminate between Cu(I) and Cu(II). Under the more oxidising conditions of the periplasm, the thioether moieties of Met residues prove to be more resistant to oxidation compared to the Cys thiols. Consequently, copper coordination via thioether Met bonds is often found in periplasmic proteins that are more associated with copper tolerance mechanisms (40).

In accordance to bacterial physiology, all known copper enzymes in bacteria are located in the periplasm, except for few known examples where there is a cytoplasmic need for copper. Such exceptions are found in cyanobacteria, where copper is imported to the thylakoid for the needs of plastocyanin and cytochrome c oxidase, two enzymes responsible for photosynthesis and aerobic respiration respectively (41, 42). Another example of a cytoplasmic copper enzyme is methane monooxygenase that is used by methane oxidising bacteria for the conversion of methane to methanol and is located in intra-cytoplasmic membranes (43).

1.3.1 Copper efflux mechanisms

The main machinery for copper transport across the plasma membrane consists of Cu(I)-ATPases, members of the P_{1B} subgroup of P-type ATPases, which are responsible for the transport of soft Lewis acids and are typically related to detoxification (44, 45). Consistent with this role, Cu(I)-ATPases are necessary for bacterial virulence, a feature that possibly developed from the need of bacteria for defence against phagocytosis (46), in order to survive. The overall structure of Cu(I)-ATPases consists of eight transmembrane helices, containing the metal binding sites, as well as an actuator and an ATP-binding domain, both of which are located in two intracellular loops (45). An additional cytoplasmic metal binding domain, containing the typical Cys-X-X-Cys motif, is found in the N-terminus of the transporter. The overall affinity of P_{1B} ATPases

is at the femtomolar range, consistent with the miniscule levels of available copper in the cytoplasm (44).

The mechanism by which Cu(I)-ATPases achieve copper transport across the membrane has some characteristic features common in all the members of the family (45). Transport of the metal is coupled to ATP hydrolysis and metal binding is essential for catalysis of ATP phosphorylation. The transporter can switch between two structural conformations (E1/E2), while the rate limiting step of the process is dephosphorylation and transport is regulated by metal binding to the cytoplasmic metal binding domain. Notably, metal acquisition by Cu(I)-ATPases depends on the interaction between the transporter and a metallochaperone, which is achieved by binding through the same ferredoxin fold (45).

A well characterised transporter of the family of P1-type ATPases is CopA, an internal membrane pump used for the transport of cytosolic Cu(I) to the periplasm (44). The overall crystal structure of CopA from *Legionella pneumophila* reveals the cytosolic region is composed of three domains, a nucleotide binding, a phosphorylation and an actuator domain, that have conserved cores as other homologues of the P-type ATPase family. The membrane part of the transporter is assembled of eight transmembrane helices as described earlier, although the exact position of the last two helices (MA and MB, specific to P1B ATPases) is controversial (44). In this structure the actuator domain is located between the membrane helices M2 and M3 and does not have the N terminus part (44) (Figure 1.1).

Two metal binding sites are contained in the membrane part of CopA and are accessible through a cytoplasmic opening formed at the N-terminus. Cu(I) ions are coordinated by six conserved residues, although the binding stoichiometry of the transporter remains uncertain (44). Comparison of these metal binding residues to the respective residues of class II ATPases indicates that in CopA Cu(I) transport might not be coupled to the transport of protons in the opposite direction. Additionally, the cytosolic heavy metal binding domain (HMBD) of the transporter protein is thought to pass on Cu(I) to the entry site of the transport channel of the pump. HMBD contains a characteristic ferredoxin fold $\beta\alpha\beta\beta\alpha\beta$ that is also present in copper chaperones and is responsible for the interaction of the two components. Although the exact position of HMBD has not been determined, due to insufficient electron density, one of the possible

positions of the domain indicates it could interact with the actuator domain and regulate the function of CopA. Nevertheless, HMBD is not considered to be necessary for ATPase activity of CopA (44).
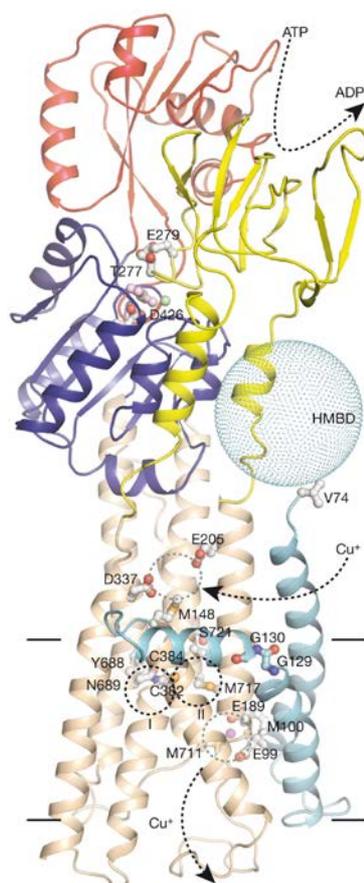
Figure 1.1 The overall structure of Cu(I)-ATPase from *L. pneumophila* is represented as a cartoon with the main domains coloured in yellow (A-), red (N-) and blue (P-).The transmembrane helices MA-MB and M1-M6 are shown in cyan and wheat, respectively. The HMBD region is marked in a sphere. Key residues are shown as ball-stick representations and the proposed route for Cu(I) transport and ATP turnover is shown with arrows, while putative Cu(I) binding sites are circled and show the entry (grey), membrane (black) and exit (grey) locations. Figure from reference 44.

The most interesting feature of the CopA structure is a platform that is formed by the interaction of the C terminus of the MB helix with the M1 helix. This platform has an amphipathic character, due to hydrophobic residues orientated towards the membrane and residues with positive charges towards the cytoplasm (44). Three residues of this platform, Met148, Glu205 and Asp337, are potential ligands for copper binding and this platform is possibly a docking site for HMBD or a copper chaperone. The delivery of

the metal to the platform might coincide with the transition of the transporter from E2 conformation to E1 (44), which results from the shift of helix M1 towards the extracellular space, while helix M4 moves towards the cytoplasm. Additional movement of residues is proposed to allow the dissociation of Cu(I) by increasing the distance of Cys 382, initially involved in Cu(I) coordination, from binding site II.

Although CopA is clearly a major component in copper homeostasis, deletion of the CopA gene does not lead to total loss of copper resistance, indicating another component is involved in copper detoxification. An identified metallochaperone from *Bacillus subtilis*, CopZ, was shown to interact in vivo with CopA and is, therefore, believed to be a cofactor in copper transport (47). Copper chaperones have structural features that aid interaction with their target metal binding domains. CopZ shares the same ferredoxin fold, βαββαβ, with the MBDs of CopA and contains the conserved motif MXCXXC responsible for Cu(I) binding (48, 49). The coordination geometry for Cu(I) has been suggested to be nearly linear S-Cu-S, with bonds between Cu(I) and the ligands deviating from linear by 26°, without the involvement of a third ligand (49). The distorted linear coordination geometry is thought to favour the observed dimerization of the chaperone upon Cu(I) binding (50).

The crystal structure of CopZ from *B. subtilis* showed the formation of a tetranuclear Cu(I) cluster inside the protein dimer (51) that consists of two types of Cu(I) sites with different coordination geometry. The outer couple of Cu(I) is coordinated in a trigonal geometry while the inner couple of Cu(I) is coordinated in distorted linear geometry. In both cases coordination is achieved via two Cys (Cys 13 and 16) and a His (His 15) residue, from each CopZ monomer. Each Cys residue contributes to the ligation of an inner and outer Cu(I) ion and His provides the third ligand in the case of trigonal geometry. Interactions with two H_2O molecules further stabilise the trigonal binding sites giving them a partial tetrahedral character. Notably, the His residues are conserved, indicating the importance of these in the novel coordination geometry described above. The interaction of metallochaperones with Cu(I) transporter proteins is essential for the efficient trafficking of the metal, as has been shown by experiments where copper deficiency is induced by the formation of tetrathiomolybdate (MoS_4^{2-}) (52). This compound binds to metallochaperones and forms a copper molybdenum cluster containing four coppers. In some cases tetrathiomolybdate can also form a heteromeric

complex with the chaperone and the transporter protein via the surface exposed MXCXXC motif, indicating the mechanism through which copper delivery is inhibited.

1.3.2   Copper sensing

Copper sensors are responsible for the regulation of copper transporter systems. In the case of CopA, the cytosolic copper sensor CueR regulates expression of the transporter (53). As shown by transcriptional assays, transcription regulation by CueR is independent of metal concentration and the efflux mechanism is by default 'on', as half maximal induction of CueR occurs at 10^{-20} M of free Cu(I). This indicates that the high content of ligands such as glutathione in the cytosol is clearly is not enough to ensure all cytosolic copper is tightly bound and it is the zeptomolar sensitivity of CueR to Cu(I) that ensures the levels of 'free' Cu(I) in the cytosol are minimal (53).

CueR is a member of the MerR family sensors and appears in the crystal structure as a dimer. Cu(I) coordinated by two Cys residues in the interface of the two monomers in a practically linear coordination geometry, with an S-Cu(I)-S angle of 176°. Each monomer consists of three domains involved in the dimerization, metal and DNA binding. The selectivity of CueR is determined by coordination, with low coordination geometry favouring monovalent ion binding rather than divalents. This is also confirmed by experiments where addition of Zn(II) or Hg(II) to CueR showed insignificant binding (53, 54). Hydrophobic and steric interactions help the restriction of the bound metal at low coordination geometry, while electrostatic interactions further stabilise the negative charge of the binding site, resulting in high affinity (54).

While CueR is the typical Cu(I)-responsive repressor in Gram-negative proteobacteria and has been studied in *E. coli*, two other families of regulators are known in Gram-positive bacteria, CopY and CsoR (24). CopY has been systematically studied in *Enterococcus hirae* where it acts in combination with the CopZ chaperone as part of the copper response mechanism that also includes CopA and CopB. This organism is a rare case where CopA is thought to be an importer and CopB is the efflux pump instead. In *E. hirae* CopY is found as a homodimer bound to the *cop* operon, encoding the respective copper efflux proteins, under stress free conditions. When copper concentration rises beyond the tolerated level, CopZ delivers copper to CopY inducing a structural change that results in decreased affinity of the regulator for the operon. CopY initially contains a Zn(II) ion that stabilises the structure of the regulator in its active

form. However, the higher affinity of CopY for Cu(I) drives the replacement of Zn(II) by two Cu(I) ions. Overall, this mechanism results in the release of the CopY homodimer from the DNA sequence and the subsequent expression of CopA and CopB (55, 56).

The discovery of CsoR, a Cu(I) sensing transcriptional regulator in *M. tuberculosis*, revealed a third distinct family of sensors that is much more common in prokaryotes than the previously described CueR and CopY (57). CsoR is widely distributed in Gram- positive bacteria and induces transcriptional derepression of genes responsible for copper resistance. Like all members of this family of sensors, CsoR has a characteristic residue motif x-C-H-C through which it coordinates Cu(I) (58). The crystal structure from *M. tuberculosis* showed CsoR is a dimer of dimers with a 4-helix bundle ($\alpha1$-$\alpha2$-$\alpha1$'-$\alpha2$') being the main structural unit of the protein. Cu(I) is coordinated in the protomer interface in a trigonal S_2N geometry by two Cys (Cys36 Cys65') and a His residue (His61') (57), while the structure is very similar in the characterised homologue from *B. subtilis* (59). The homologue from *B. subtilis* has been shown to bind one Cu(I) equivalent per monomer with an affinity at the range of 10^{21} M^{-1}. Although the sensor is capable of binding other metals than Cu(I), including Ni(II), Co(II) and Zn(II), the coordination geometry employed is different from the native one used for Cu(I) binding, and the affinity is much lower (59), indicating the strong preference of CsoR for Cu(I).

In the apo-form CsoR has a tight affinity for the DNA operator, which is decreased upon Cu(I) and, consequently, expression of copper resistance proteins is mediated (60). The exact mechanism of CsoR binding to DNA is not clear, however it has been suggested that the CsoR tetramer undergoes structural reorganisation when it binds copper, in order to detach from the DNA strand. In the Cu(I)-CsoR structure from *M. tuberculosis* $\alpha2$-helix of the sensor appears bent towards the Cu(I) ion, whereas this is not the case in apo-CsoR crystallised from *Thermus thermophilus* bound to DNA (61). The interruption of the $\alpha2$-helix or the insertion of a kink into it, in the Cu(I)-bound state, has been proposed in experiments comparing the NMR spectra of apo- and Cu(I)-bound forms of CsoR from *Geobacillus thermodenitrificans* (60) and is in agreement with the model where a conformational change drives binding to DNA and regulation of copper response (60).

1.3.3 Copper homeostasis in the periplasm

Once Cu (I) is exported in the periplasm, it is oxidised to Cu(II), which is less toxic, by CueO. This enzyme belongs to the multicopper oxidase family of proteins, along with ascorbate oxidase, laccase and cerruloplasmin, and can oxidise up to four Cu(I) ions to Cu(II), by coupling this reaction with the four- electron reduction of dioxygen to water (62, 63). The crystal structure of CueO clearly visualises a catalytic center containing four copper atoms. Two of the copper ions are coordinated in an almost linear geometry with an oxygen bridge and the third copper is at a distance of 3.2 Å from the bridging oxygen (63). The crystal structure of CueO revealed additional Cu(I) ions bind in a methionine-rich region near the substrate entry copper site (63). In *E. coli*, the group of copper homeostatic proteins described earlier (CopA, CopZ, CueR and CueO) is known as the 'Cu efflux' (cue) system, which is one of the two chromosomally regulated systems responsible for cell response to copper stress (62) (Figure 1.2). Under aerobic conditions, the cue system serves as the first response of the cell to copper stress, and half- maximal induction of the cue promoter occurs at 3μM copper (64). A second, 'Cu sensing' (cus) system (62), is expressed independently of cue and is thought to have a complementary role since it is expressed either when cue is overloaded by excess copper or under anaerobic conditions when CueO is inactive. In aerobic conditions, cus is only switched on at extreme copper stress levels, as the respective promoter reaches half maximal induction at approximately 200 μM copper (64). This indicates the cus system comes in use only when cue is overwhelmed by copper overload. However, under anaerobic conditions, cus has been shown to respond first. This seems reasonable as, in this case, CueO is inactive due to the absence of oxygen, which subsequently leads to accumulation of Cu(I) in the periplasm (62).

The cus system consists of the CusCFBA protein network, where CusCBA belong to the family of resistance- nodulation- division (RND) proteins (64) and CusF is a periplasmic copper chaperone or regulator (65). The cus system is regulated by a sensing (CusS) and a regulator component (CusR). CusS senses excess copper in the periplasm and activates CusR that subsequently induces CusCFBA transcription (66, 67). CusA is a transmembrane protein from the family of H$^+$ transporters located in the cytosolic membrane, while it contains a region that extends in the periplasm and interacts specifically with the membrane fusion protein CusB (68). CusA is connected with the outer membrane protein CusC, forming a transport channel across the

periplasm (66, 68), and uses methionine residues to bind Cu(I). Interestingly, the crystal structure of CusA suggests it can uptake metals both from the periplasm and the cytosol (69). The third component of the system is CusB is a membrane fusion protein and is responsible for binding the Cu(I) delivered by CusF (66).

The CusCBA complex spans the entire cell envelope and ejects copper to the extracellular space using proton motive force (64, 66) (Figure 1.2). It has been suggested that CusCBA mediates Cu(I) efflux either from the cytosol or the periplasm to the extracellular space (69). However, experiments using copA deletion strains have shown these mutants have increased sensitivity to copper, indicating that CopA can not be substituted by CusCBA. This finding, therefore, is consistent with the model where CusCBA exports copper from the periplasm to the extracellular space. Moreover, the regulation of CusCBA by CusS/R, that senses periplasmic copper, is also in agreement with cus being an independent system from cue, responsible for periplasmic copper efflux (63, 64).
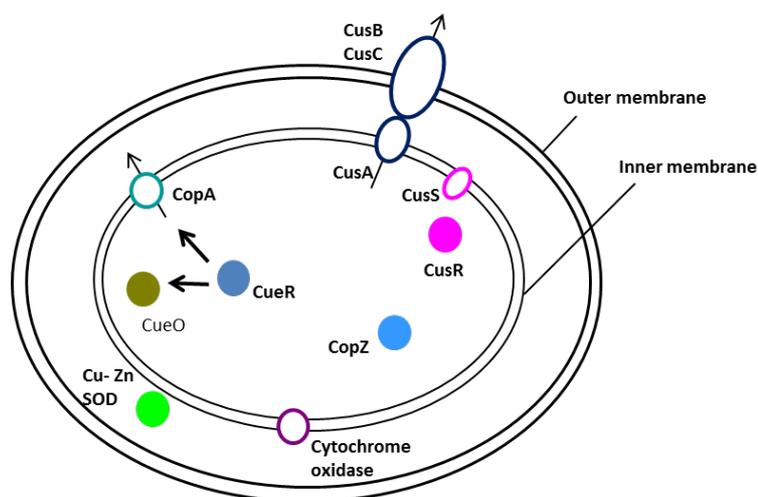
Figure 1.2 Copper proteins involved in the copper homeostasis of a Gram-negative bacterium including the components of the cue system (CueR, CopA, CopZ and CueO) and the cus system (CusABC and CusRS), copper/zinc superoxide dismutase and cytochrome oxidase (63, 64).

1.3.4 Copper import mechanisms

Although most copper proteins known thus far are periplasmic there are few cases of cytoplasmic copper proteins. Cyanobacteria, a characteristic example of organisms with cytoplasmic copper requirement, have developed a network of two transporters and a

chaperone that deliver copper to these enzymes. Plastocyanin, an electron shuttle for photosynthesis, and cytochrome oxidase (COX), the terminal electron acceptor in aerobic respiration, are localised in the thylakoid and the thylakoid membrane, respectively. CtaA, a P-type ATPase homologous of CopA, transports copper across the plasma membrane. Copper is then chaperoned across the cytoplasm by Atx1, a CopZ-type chaperone, and exported to the thylakoid by PacS, a P-type ATPase transporter situated in the thylakoid membrane (41, 42) (Figure 1.3).

In cyanobacteria the Atx1 chaperone binds two Cu(I) ions and forms a dimer upon Cu(I) binding, as observed in crystal structures and also by NMR in solution (41, 42, 70). This behaviour is consistent with that of the same type chaperone CopZ, as described earlier. Two different dimer formations, head-to-head for Cu(I)$_1$-Atx1 or side-to-side for Cu(I)$_2$-Atx1, are observed *in vitro*, depending on copper loading (41). The side to side dimer contains four Cu(I) ions arranged in a highly symmetrical [Cu$_4${μ_2-S$^\gamma$(Cys)}$_4$Cl$_2$]$^{2-}$ cluster. Copper exchange experiments show transfer of the metal is favoured between CtaA and Atx1 and the Cu(I)$_2$-Atx1 side-to-side dimer is more efficient in transferring copper to PacS, despite the process being reversible. The Cu(I)$_2$-Atx1 dimer therefore is considered to be the physiologically relevant form of the chaperone, while dimerization has been suggested to be a mechanism for regulating protein function and copper availability to PacS (41).
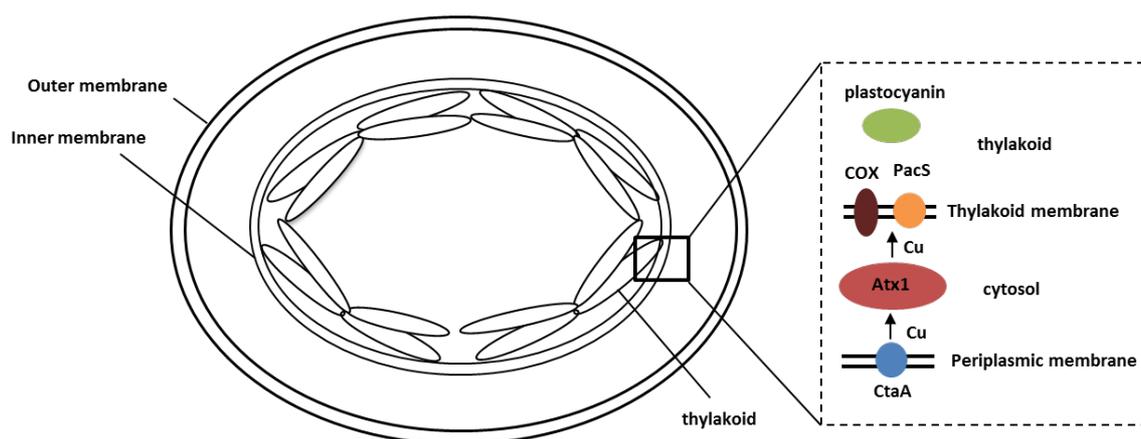
Figure 1.3 Schematic representation of the mechanism through which copper is thought to be imported in the cytosol and subsequently exported to the thylakoid for the needs of plastocyanin in *Synechocystis* PCC6803 (41, 42).

As with CopZ and CopA, the interaction of Atx1 with the P-type ATPase transporters is achieved through the MXCXXC motif present in the ferredoxin fold and intriguingly in *Synechocystis* PCC 6803 this interaction is preferably stabilised in the presence of Zn(II) rather than Cu(I) (71, 72). Atx1 from *Synechocystis* PCC 6803 binds Zn(II) and the crystal structure obtained is very similar to that of the Cu(I)$_2$-Atx1 side-to-side dimer (71). Moreover, the Atx1-Zn(II)-PacS complex is more stable compared to the respective Cu(I) bridged complex, a feature that can be attributed to the preference of Zn(II) for tetrathiolate coordination and the higher number of electrostatic interactions in the complex (71). The CtaA transporter has been shown to have an Zn(II) affinity that is 20-fold higher than that of Atx1 while the latter has been suggested to bind Zn(II) *in vivo* (71). The affinities of the above proteins for zinc combined with the stability of the Atx1-Zn(II)-PacS complex suggest that under different concentrations of zinc, Zn(II) transport might be favoured over Cu(I) transport. As a consequence, a role for zinc in the regulation of copper delivery to the two target enzymes, plastocyanin and cytochrome oxidase, has been suggested (71, 72).

Based on studies of a CtaA mutant, CtaA is the only Cu(I) transporter with a suggested role in import in bacteria so far (73). The homologous P-type ATPase in *Enterococcus hirae,* CopA, is also thought to function as an importer while CopB is responsible for Cu(I) efflux. This is based on experiments showing the growth of copA mutants is impaired under conditions of low copper availability, however conclusive evidence is not available. It is not clear what dictates the direction of Cu(I) transport by Cu(I)-ATPases, since the main structural features involved in Cu(I) efflux are conserved even in transporters like CopA from *E. hirae* for example . An alternative role has been proposed for transporters thought to mediate import and it involves assembly of copper proteins. This is based on the finding that a Cu(I)-ATPase- type transporter from *Pseudomonas aeruginosa* was only capable of Cu(I) efflux but at a very slow rate and, therefore, could not have a role in copper detoxification (74). According to this study, these Cu(I)-ATPases must serve a different function in the cell and, in the case of CtaA, the suggested role is copper protein assembly (45).

A different example of cytoplasmic copper requirement in bacteria is abundant in methane oxidising organisms where the main form of the methane oxidising enzyme is located in intracytoplasmic membranes. Methane monooxygenase (MMO) is abundant as membrane-bound, particulate, form (pMMO) in almost all methanotrophic bacteria

and requires copper for activity (32). The intra-cytoplasmic membranes of these bacteria have been studied by electron microscopy and two types of intracellular organisation have been observed (139). In the first type of membrane organisation, membrane bundles consisting of discrete vesicles were observed and, in some cases, the vesicles were attached to the cytoplasmic membrane, indicating they formed possibly by invaginations of this membrane. The second reported pattern was less ordered and consisted of membrane pairs arranged either around the periphery of the cytoplasmic membrane or throughout the cell. Based on the above, a definite conclusion cannot be made on whether the intracytoplasmic formations are continuous with the cytoplasmic membrane or form separate subcellular compartments. Despite the uncertainty about the exact nature of the membranes housing pMMO, the copper requirement of his enzyme suggests the existence of a copper import mechanism. Indeed, methanotrophs synthesize and excrete a peptide, known as methanobactin, which scavenges copper from the environment and is involved in copper uptake (35). Although methanobactin is a well-studied molecule in terms of its biochemical properties (section 1.3.7) and has been shown internalised in the cytoplasm (43) to deliver copper to pMMO, the mechanism of this process is unclear.

1.3.5   Copper storage

A family of proteins that has often been associated with detoxification of heavy metals as well as storage of essential metals, such as copper, is metallothioneins (75). This family of proteins was originally associated with cadmium resistance as the first metallothioneins in mammals and bacteria were identified from organisms growing in high levels of this metal (76). Metallothioneins have been well studied in mammals and are described as small proteins with high content of Cys and metal. These features provide high metal binding capacity and one proposed role of metallothioneins is storage (75). Only few metallothioneins are known in bacteria and, in some cases, have different features from the conventional description of the group.

The first characterised bacterial metallothionein came from *Synechococcus* PCC7942 were it is part of the smt operon. SmtA binds Zn(II) when bacterium is exposed to elevated concentrations of the metal and the structure of SmtA provided the first evidence that ligands other than Cys, such as His in this case, participate in metal binding. This feature provided new insights on the characteristics of metallothioneins (76). A more recent addition was the discovery of MymT from *M. tuberculosis*, a

metallothionein with a preference for Cu(I) binding, through Cys-X-Cys motifs, that is believed to participate in protecting the organism from copper toxicity (77). Although the low sequence similarity between metallothioneins, combined with their small size, makes identification of new members of the family difficult, more bacterial metallothioneins are likely to be discovered (76).

1.4 Methane oxidising bacteria

1.4.1 Methane oxidation

Methane oxidising bacteria (MOB) are Gram-negative organisms that use methane as a source of carbon and energy and are, therefore, an important component of the global carbon cycle. Methane is oxidised to CO_2 and the first step of this process, the conversion of methane to methanol, is catalysed by methane monooxygenase (MMO). Methanol is then converted to formaldehyde, and subsequently to formate which is finally oxidised to CO_2, with all reactions catalysed by the respective dehydrogenases (Figure 1.4) (78-80). Approximately 50% of the formaldehyde produced is used for carbon assimilation either through the serine pathway or via the ribulose monophosphate (RuMP) pathway (79). A third pathway for carbon assimilation by CO_2 fixation has emerged from the discovery of new methanotrophs within the verrucomicrobial phylum, and involves the Calvin-Benson-Bassham cycle (81, 82). The reactions of formaldehyde oxidation regenerate the reducing energy needed for the initial step of methane oxidation (80).
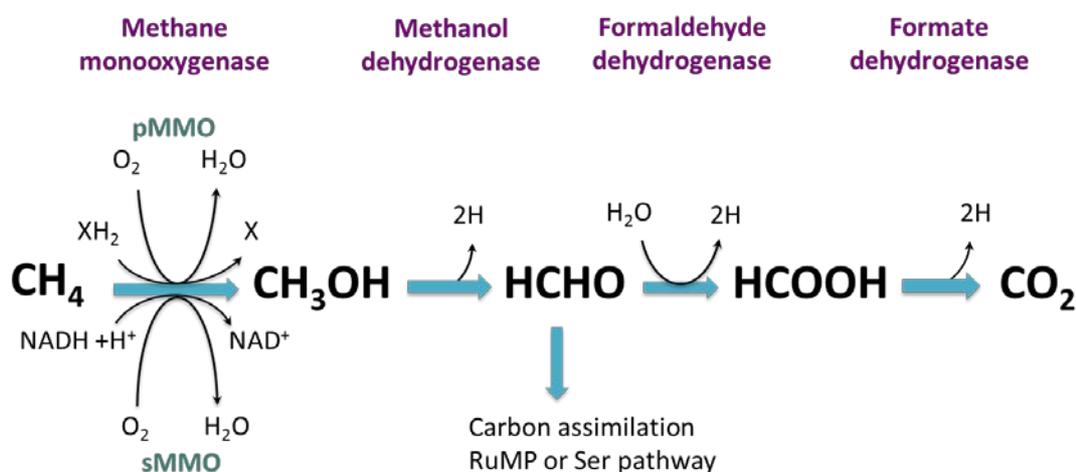
Figure 1.4 Metabolic pathway for methane oxidation in MOB (80).

Most methanotrophs are obligate organisms which means methane is their only source of carbon and energy (78). However, facultative methanotrophs have also been identified in the *Methylocella* species that can also grow on multi-carbon compounds, such as acetate (83). Methanotrophs possess enzymes that are unique to methane oxidation, such as MMO, methanol dehydrogenase and cytochromes, thought to mediate electron transfer for methanol oxidation, as well as enzymes specific to carbon assimilation. The genes for these enzymes can be recognised and used as probes for the determination of phylogenetic relationships between methanotrophs (79).

Two forms of MMO can be expressed, a soluble iron-containing form (sMMO) and a copper-containing particulate form (pMMO), which is located in intra-cytoplasmic membranes. MMO needs two reducing equivalents to break the bond of dioxygen: one resulting oxygen atom is then reduced to water and the other is incorporated into methane for oxidation to methanol (79). In the case of sMMO the reductant used is NADH, whereas the reductant used by pMMO under physiological conditions is unknown, although quinones are possibly involved (84). sMMO, expressed in copper limiting conditions, can oxidise a broad range of substrates including alkanes of up to eight carbons, cyclic alkanes, ethers and aromatic hydrocarbons (80). On the other hand, pMMO, the main methane oxidase form in MOB, represents 20% of the total cell protein (85) and is present in all known MOB except one (*Methylocella silvestris* BL2) (86). While pMMO has narrower substrate specificity compared to sMMO and will only oxidise alkanes with up to five carbons or alkenes of up to four carbons (32), it exhibits higher affinity for methane and is considered more efficient in methane oxidation (78). Notably, in organisms capable of expressing both MMO forms, the switch is copper dependent and sMMO has been proposed to act as an alternative route for methane oxidation aiding the survival of MOB under copper limiting conditions where the copper-requiring pMMO is not active (79).

1.4.2 Taxonomy and phylogeny

Previous studies have categorised MOB in three groups, based on physiological and functional characteristics. Type I MOB are gammaproteobacteria that have extended intra-cytoplasmic membranes structures throughout the cell, assimilate copper through the RuMP pathway and have characteristic phospholipid fatty acids of 14-16 carbons (78-79). Type II organisms are alphaproteobacteria that contain intra-cytoplasmic membranes localised along the periphery of the cytoplasmic membrane, assimilate

carbon via the serine pathway and have phospholipid fatty acids of 18 carbons. A third category, Type X, includes MOB strains with features from both other groups, although recently Type X organisms have been reclassified as Type I (78-79). A more relevant way of distinguishing MOB strains is based on the expression of MMO. According to this, most organisms express only pMMO whereas a few, 'switchover' organisms are capable of expressing sMMO as well under conditions of low copper availability in the environment (80).

Phylogenetically, MOB are studied based on the sequence similarities of 16S rRNA. According to this approach, MOB are widely distributed in alphaproteobacteria, gammaproteobacteria and, as more recently discovered, in Verrucomicrobia (82). Alphaproteobacteria include *Methylosinus, Methylocystis, Methylocapsa* and *Methylocella* strains, whereas much more diversity is found in gammaproteobacteria, with strains from *Methylomicrobium, Methylococcus* and *Methylothermus* strains, among others, but also methane oxidisers such as *Crenothrix polyspora* and *Clonothrix fusca* (Table 1.1). The branch of Verrucomicrobia contains *Methylacidiphilum infernorum*. Phylogenetic studies based on the sequence of *pmoA*, the gene responsible for the expression of the β-subunit of pMMO, provides additional information on how these organisms developed. For instance, although the *pmoA* sequence of *Crenothrix* is very different from the rest of the gammaproteobacteria strains, this did not occur as a result of horizontal gene transfer but, more likely, by differentiation of the species from the rest early on (78, 82).

In agreement with their phylogenetic diversity, MOB can be found in a range of environments including wetlands, freshwater, groundwater, marine sediments and sewage waste. These bacteria tolerate a wide range of temperatures, with some strains like *Methylosphaera hansonii* growing optimally at 10° and others, like *Methylothermus thermalis* at the opposite extreme of 57°. A wide range is also tolerated in pH and salt concentration, with alkalitolerant and halophilic strains from *Methylobacter* species optimally growing at pH 9 and at salt concentration >3%, and others like *Methylocella palustris* at pH 5-5.5. Nevertheless, a large proportion of MOB strains are mesophilic and neutrophilic, thriving at moderate temperatures and approximately neutral pH (78).

Domain: Bacteria/ Kingdom:Eubacteria				
Phylum: *Proteobacteria*				*Verrucomicrobia*
Class: *Gammaproteobacteria*			*Alphaproteobacteria*	*Verrucomicrobiae*
Family: *Methylococcaceae*			*Methylocystaceae*	*Methylacidiphilaceae*
Methylobacter	*Methylothermus*	*Methylocaldum*	*Methylocystis*	*Methylacidiphilum*
Methylomicrobium	*Methylomarinum*	*Methylogaea*	*Methylosinus*	
Methylosarcina	*Methylococcus*	*Methylohalobius*	*Methylocella*	
Methylosphaera	*Crenothrix*	*Methylovulum*	*Methylocapsa*	
Methylosoma	*Clonothrix*	*Methylomonas*	*Methyloferula*	

Table 1.1 Phylogenetic classification of MOB based on 16S rRNA (82).

1.4.3   Applications of MOB

Due their broad environmental distribution and the range of substrates they can oxidise, MOB represent an excellent candidate for various biotechnological applications. Firstly, considering MMO is one of only two known enzymes for methane oxidation alongside ammonia monooxygenase, it is of great interest for the development of synthetic catalysts. The C-H bond of methane is the most stable hydrocarbon bond (104 kcal/ mol) (87, 88) and its catalytical oxidation is the main challenge for fuel production. As methane is the main component of natural gas, its oxidation to liquid methanol, which is easily stored and transported, has been the subject of ongoing research with emphasis on catalysts able to oxidise C-H at low temperature and pressure conditions (89). Liquid methanol contains 12.6% hydrogen that can be made available for fuel through, low temperature methanol dehydrogenation technology (90, 91). Therefore, MMO in combination with the available technologies can be used to facilitate the use of hydrogen as fuel, providing a clean energy source as an alternative to petroleum.

Methane accounts for 20-30% of global warming, as it is a more effective greenhouse gas than CO_2 due to its higher capacity for trapping radiation (81). Considering methane concentrations have been annually increasing by 1% over the last two centuries, MOB offer a direct solution to mitigating methane emissions (92). Methanotrophs are known to oxidise between 50-90% of the methane produced underground before it reaches the atmosphere, whereas they also account for 5% of the atmospheric methane sink (81). Taking these into account, technologies are being developed in order to take full advantage of MOB for cost-effective mitigation of methane emissions from landfills that are responsible for approximately 24-30% of human methane production. Such solutions are based on the use of MOB populated biofilters or biocovers through which landfill gas could be actively of passively vented to minimize the release of methane in the atmosphere (93).

Another aspect of MOB applications is based on the range of substrates oxidized by MMO, which can provide a solution to bioremediation. sMMO has been shown to oxidise trichloroethylene and chloroform, common water pollutants generated by the use of chlorinated aliphatic compounds in industry and agriculture (94). Cells expressing sMMO are capable of degrading such pollutants at high rates, compared to those expressing mainly pMMO, although sMMO activity becomes limited by product toxicity. On the other hand, at higher pollutant concentrations, pMMO expressing cells

grow faster and have the competitive advantage in degradation rate (95). Polluted groundwater usually also contains high concentrations of copper indicating that in these cases pMMO, being a copper enzyme, might be more efficient in bioremediation than sMMO, which is repressed in these conditions (96). These findings suggest conditions of MOB growth can be optimized in order to control the expression and activity of MMO for bioremediation applications.

1.4.4   Structure and regulation of sMMO

sMMO has been isolated from *Methylococcus capsulatus* Bath and *Methylosinus trichosporium* OB3b, and is a non- heme iron-containing enzyme consisting of three components: a hydroxylase MMOH, a reductase MMOR and a regulating protein MMOB (97) (Figure 1.5). Methane oxidation takes place in MMOH, which consists of 3 subunits arranged in a $\alpha_2\beta_2\gamma_2$ heterodimer. Subunit α contains a dinuclear iron center bridged by glutamate- carboxylate and hydroxide. The electrons needed for methane oxidation are generated by MMOR, an iron-sulfur flavoprotein, from the oxidation of NADH. MMOB controls the activity of MMOH by forming complexes with that component that affect the structure of the di-iron site (97, 98). MMOB has been shown to bind to the hydroxylase in the interface of subunits α and β with the N-terminus of MMOB forming a ring-shape on MMOH that is further stabilised through hydrophobic interactions and hydrogen bonds (98). The conformational changes in MMOH, resulting from the formation of the MMOH-MMOB complex, control access of oxygen, methane and protons to the catalytic site.

A cluster of genes encodes sMMO known as *mmoX*, *mmoY* and *mmoZ*, responsible for the expression of the α, β, γ subunits of MMOH, and also *mmoB* and *mmoC* for MMOB and MMOR respectively. An open reading frame (OrfY) coding for a protein of unknown function is located between the *mmoZ* and *mmoC* gene. A recent study reported the isolation and characterisation of MMOD, an additional component of sMMO, which is encoded by OrfY (Figure 1.5). MMOD has been suggested to inhibit sMMO activity as it binds to MMOH with affinity comparable to that of MMOB and MMOR (99). In *M. Capsulatus* (Bath) all the aforementioned genes are controled by a single promoter that is copper regulated. However, the exact mechanism by which the expression of these genes is controlled at different copper concentrations is not understood. It is possible that this occurs through a sensing/ regulating copper binding

protein that, subsequently, binds to the promoter and controls RNA polymerase binding (80, 97).

Interestingly, expression of the mmoXYBZDC operon is activated only under low copper: biomass ratio and a gene encoding what is thought to be a transcriptional activator (MmoR), *mmoR*, has been identified upstream of the sMMO operon in *M. trichosporium* OB3b, with homologues also present in other MOB (100). Binding of MmoR to DNA does not lead to transcription and it seems unlikely that the activator is regulated directly by copper ions, as copper binding motifs have not been identified. A second identified gene, *mmoG* is believed to encode a chaperone that interacts in vivo with MmoR. This theory is further supported by studies showing that mutant strains of *M. trichosporium* OB3b, lacking *mmoR* and *mmoG*, were incapable of expressing sMMO and MmoG may be necessary for MmoR binding to DNA (100).
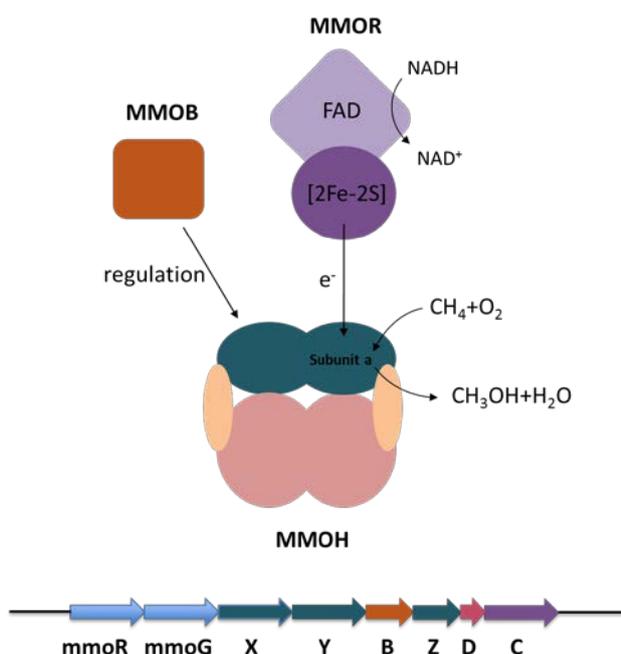
Figure 1.5 Schematic representation of the MMOR, MMOH andMMOB components of sMMO and the genes involved in the expression and regulation of the enzyme (97, 100).

1.4.5 pMMO structure and active site

The particulate form of MMO has been crystalised from *M. capsulatus* (Bath), *M. trichosporium* OB3b and *Methylocystis* strain M (88, 101, 102). The overall structure of pMMO from all organisms is the same and consists of three subunits arranged in a

$\alpha_3\beta_3\gamma_3$ trimer. The pmoB subunit has two cupredoxin folds situated in the N- terminal and the C- terminal respectively, which are the only soluble domains of the enzyme, while subunits pmoA and pmoC are composed mainly of transmembrane helices. Structures from *M. trichosporium* OB3b and *Methylocystis* strain M show an additional transmemrane region close to the pmoC subunit, which is not present in pMMO from Bath (32). An electron microscopy structure also visualised pores, formed in the periplasmic soluble regions of pmoB, which are not obvious in the crystal structure and are suggested to be the locations for substrate access to the active site of the enzyme (103).

Three metal centers have been observed in the crystal structures of pMMO, although with slight differences between the three strains. In pMMO from *M. capsulatus* (Bath), two metal centers occupied by copper where modelled in the soluble domains (spmoBd1 and spmoBd2 in Figure 1. 6) of the pmoB subunit. One of these centers is thought to be dinuclear, although the resolution of the structure is not sufficient to distinguish the two coppers at such proximity, while the second metal center is mononuclear (32, 104). In the dinuclear center, which is located near the membrane interface, copper is coordinated by His residues (His 33, 137 and 139) that are conserved in all methanotrophic pmoB sequences, except for Verrucomicrobia (82, 104). The mononuclear center is located in the interface of the two cupredoxin domains and copper is coordinated by His 48 and His 72 residues (32, 104). A third metal center is situated in the transmembrane part of pMMO and is occupied by zinc that is derived from the crystalisation buffer (88) (Figure 1.6). Notably, the model described is in agreement with the copper binding stoichiometry reported for pMMO from *M. capsulatus* (Bath), where 2-3 copper ions bind per the $\alpha\beta\gamma$ protomer (32).

The crystal structures from *M. trichosporium* OB3b and *Methylocystis* strain M present some differences in the metal centers described above. In pMMO from *Methylocystis* strain M the majority of the metal sites where modelled as mononuclear. Moreover, from the copper coordinating residues of the mononuclear site in pMMO (Bath) His 48 is replaced by an Asn in pMMO from both *M. trichosporium* OB3b and *Methylocystis* strain M, while this site doesn't contain copper in neither of these structures. The intramembranous metal center is occupied by zinc in pMMO from *Methylocystis* strain M while in *M. trichosporium* OB3b this site is ocuppied by copper. However, this is not necessarily of physiological relevance as the protein was purified in the presence of

copper. In terms of copper binding, the stoichiometry for pMMO from these two organisms is slightly lower than pMMO (Bath) at approximately 2 coppers per protomer (101, 102).
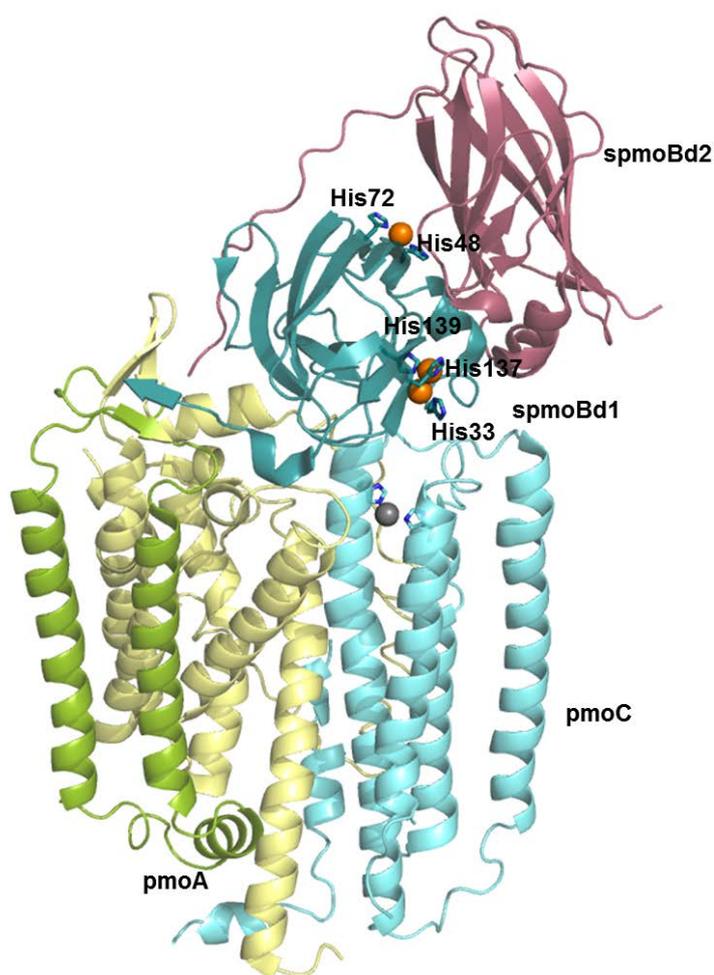
Figure 1.6 Crystal structure of the pMMO protomer (αβγ) from *M. capsulatus* (Bath) (pdb code: 1YEW). The soluble cupredoxin domains of pmoB at the N-terminus (spmoBd1) and the C-terminus (spmoBd2) are shown in dark cyan and purple, respectively, and the transmembrane helices are shown in green. Copper ions are shown as orange spheres and zinc is shown as a grey sphere. The pmoA and pmoC subunits are shown in pale yellow and pale blue, respectively (104).

Various models have been proposed regarding the active site of pMMO and its metal content including, in some cases, different copper binding stoichiometries to those already mentioned. According to one model, pMMO binds a total of 15 copper ions that are arranged in catalytic and electron transfer clusters (105, 106). Based on EPR data,

the catalytic cluster is thought to contain three copper ions that form a trinuclear center, suggested as the active site. The activation mechanism of such a trinuclear copper cluster by dioxygen has been investigated and a model for oxygen binding, which is necessary for methane oxidation, has been proposed for this catalytic cluster (107). The remaining 10-12 copper ions form an electron transfer cluster and act as a reservoir that mediates the reduction of the active site coppers after methane oxidation (105). However, the trinuclear copper cluster and the additional copper ions have not been observed crystallographically. Notably, older reports also suggest pMMO is binds 10-15 copper equivalents, but most of these copper ions were proven to be associated with methanobactin, a copper sequestering peptide secreted by MOB to mediate copper uptake (see section 1.3.7), which has been shown to co-purify with the enzyme in some cases (108).

In other studies pMMO is reportedly isolated with 2 copper ions (109), in good agreement with the data from the Rosenzweig group, discussed earlier. Despite the possible discrepancy between copper stoichiometry observed in the crystal structures and that determined for purified proteins in solution, the consistent observation of copper in the dinuclear center of pMMO suggests it is likely to be the active site of the enzyme. While previously an iron-containing site, similar to the one present in sMMO and located where the zinc site has been observed, was suggested as the active site of pMMO (85), detailed studies of the soluble domains of the pmoB subunit (spmoBd1 and d2) provide evidence this is not the case. Experiments performed on metal depleted pMMO from *M. capsulatus* (Bath) showed that addition of 2-3 coper equivalents per pMMO protomer restored enzyme activity while iron had no impact (104). Moreover, the copper binding properties and the methane oxidation activity of the soluble domains of pmoB were investigated. These experiments showed not only that copper is the essential metal for catalytic activity of pMMO but also that within spmoB, only spmoBd1 combined copper binding and methane oxidation activity (104). These data, combined with experiments showing the dinuclear copper site is capable of oxygen binding, provide strong indications it is the active site of pMMO (110).

1.4.6 Copper metabolism in MOB

MOB have an unusually high requirement in copper, as a result of the main methane oxidizing form, pMMO, being a copper enzyme. Copper has a critical role in the differential expression between sMMO and pMMO in 'switchover' MOB, although the

mechanism is not understood (78). At low copper: biomass ratio MOB use sMMO for mehane oxidation, while at high copper: biomass levels the formation of extended intra-cytoplasmic membranes that house pMMO is induced. A number of copper-responsive genes have been identified in MOB and are believed to regulate the expression of the two MMO forms. These include *mmoR* and *mmoG*, identified in *M. trichosporium* OB3b but also present in *M. capsulatus* (Bath) (100), as well as *mmoQ* and *mmoS*, all of which are possibly involved in sMMO expression. Both MmoG and MmoR are thought to be expressed in the absence of copper, where MmoR facilitates transcription of the sMMO genes and MmoG is proposed to be a chaperone mediating copper binding and, therefore, regulation of MmoR. MmoS and MmoQ are homologous to two- component signaling systems, where MmoS is proposed to be the sensor that phosphorylates MmoQ, the regulator, as a response to changes in copper levels (111). However, as there is no evidence for direct copper binding to MmoS, it remains unknown how signal transduction is achieved (78, 100, 111). Less is known about the regulation of the *pmoCAB* operon, responsible for the expression of pMMO.

Another aspect of the copper metabolism of MOB is portrayed by a number of copper responsive proteins that have been identified and characterized from *M. capsulatus* Bath and *Methylomicrobium album* BG8. In *M. capsulatus* Bath surface-associated proteins are known to respond to copper levels in the environment and therefore have a role in the ability of the organism to adapt to the environment (112). The first protein to be identified from this organism was MopE, a protein that is associated with the outer membrane in a non-covalent manner and is surface exposed. Most interestingly, a truncated form of the protein lacking the N-terminus, MopE*, is also secreted by the organism and has been isolated from the culture media (113). MopE* is expressed under copper limiting conditions and binds Cu(I) with high affinity using a unique ligand, kynurenine, which is derived from tryptophan (114). The oxidation of tryptophan is a feature specific of the wild type protein and is thought to be physiologically relevant, since it is not present in recombinant MopE* (115). MopE* is also capable of binding Cu(II), although with lower affinity (114), and due to this flexibility the protein is suggested to act as copper chaperone. Alternatively, MopE* might be one of multiple copper uptake mechanisms of *M. capsulatus* (Bath) at different copper concentrations.

The only protein with significant sequence homology to MopE identified so far is CorA in *Methylomicrobium album* BG8. *M. trichosporium* OB3b, a member of the α-proteobacteria and a Type II organism, does not encode homologues of either of these proteins, indicating they are not involved in the copper-mediated switch between the two MMO forms. CorA is repressed by copper and is essential for the survival of *Methylomicrobium album* BG8 (116), an organism depending entirely on pMMO for methane oxidation. CorA binds Cu(I) with the same ligands as MopE*, all fully conserved including kynurenine, while a bound Ca(II) ion is thought to contribute to the stability of the protein structure. Although CorA is also bound to the outer membrane and surface-exposed like MopE*, it is not present in the spent medium (117).

In both *M. capsulatus* (Bath) and *Methylomicrobium album* BG8 the genes coding for MopE and CorA are in transcriptional units with genes encoding copper repressible c-type haem cytochrome peroxidases. In *M. capsulatus* (Bath) the second protein (MCA2590) is surface associated, while CorB, the protein encoded downstream of CorA in *Methylomicrobium album* BG8 is periplasmic (118, 119). Although the two proteins share some sequence similarity, CorB is thought to represent a new family of di-haem cytochrome peroxidases. The periplasmic localization of CorB means that any potential interaction with CorA is limited by the export of CorA to the surface of the cell (118). CorB and MCA2590 are believed to have a physiological role similar to MauG proteins that mediate electron transfer reactions, and therefore could be involved in the modification of tryptophan to kynurenin, present in both CorA and MopE (118).

The only identified component for copper uptake in MOB so far is methanobactin (mb), a small peptide (1.217 kDa) that is produced and secreted by the cells in order to scavenge copper. Mb is thought to solubilise copper minerals from the environment, thus making the metal available to the cells (120-122). The primary target for copper delivery by mb in MOB is pMMO, although the mechanism through which this is achieved, or other protein components participating in this process, are unknown. In some cases mb has been reportedly associated with pMMO in preparations of the protein, indicating it might interact directly with the enzyme (108, 123). However, although there is evidence of mb being internalized by the cells, direct in vivo interaction with pMMO has not been verified (43). In other studies, mb has been shown to mediate the switchover in the differential expression of pMMO and sMMO in a

concerted mechanism with MMOD, a protein with a suggested role in DNA binding and regulation of the *mmo* operon (124, 126).

1.4.7 Methanobactin

A copper uptake system for MOB was first suggested in the case of *M. trichosporium* OB3b mutants, not expressing pMMO, where unusually high concentrations of copper accumulated in the culture medium (96). Copper binding ligands were subsequently identified from *M. trichosporium* OB3b (122) and *Methylococcus capsulatus* (Bath) (108) while the further study and characterisation of these compounds led to renaming this compound methanobactin (mb). Mb has also been characterised from other MOB strains including *Methylocystis* strain SB2, strain M and *hirsuta* CSC1, as well as the 'non switchover' strain *Methylocystis rosea* (125, 126).

1.4.7.1 Primary sequence and synthesis

Mb from *M. trichosporium* OB3b has been thoroughly characterised and is a small peptide consisting of seven amino acid residues, two of which are modified, as seen in the primary sequence: 1-(N-[mercapto-{5-oxo-2-(3-methylbutanoyl)oxazol-(Z)-4-ylidene}methyl]-Gly^1-L-Ser^2-L-Cys^3-L-Tyr^4)-pyrrolidin-2-yl-(mercapto-[5-oxo-oxazol-(Z)-4-ylidene]methyl)-L-Ser^5-L-Cys^6-L-Met^7. The modified residues contain alkilydene oxazolone rings (127), previously falsely identified as hydroxymidazolate rings (128), and suggest that mb is either synthesised through a non ribosomal peptide synthetase (NRPS) or after modification of a ribosomaly synthesised peptide (129). Interestingly, two putative NRPS genes are encoded in the genome of *M. capsulatus* (Bath). However, when mb from *M. trichosporium* OB3b and from *Methylocystis* strain SB2 is subjected to acid-hydrolysis followed by decarboxylation, amino acids are produced suggesting the peptide is synthesised ribosomally. This senario is supported further by the identification of a potential mb precursor gene from the genome of *M. trichosporium* OB3b, LCGSCYPCSCM (124, 125). Deletion of the respective gene from the organism confirmed the hypothesis as mb production was not observed in the culture media of these mutants (124).

1.4.7.2 Cu(I) binding and structure

Biochemical characterisation of mb from *M. trichosporium* OB3b showed the peptide binds Cu(I), regardless of whether it is loaded with Cu(I) or Cu(II), with a 1:1 stoichiometry (130) and with affinity at the range of $6\text{-}7 * 10^{20}$ M^{-1} (131). Determination

of the affinity for Cu(II) from the reduction potential of methanobactin (640 mV) demonstrates the clear preference of the peptide for Cu(I) (131). The crystal structure of Cu(I)-methanobactin from *M. trichosporium* OB3b (Figure 1.7 A) shows the molecule is arranged in a pyramid-like shape (132) and the copper ion is located at the base. Cu(I) is coordinated in a distorted tetrahedral geometry by the nitrogens of the two oxazolone rings and the sulfurs of the enethiolate groups, and the binding site is shielded from the solvent (131). Additional interactions stabilise the copper binding site including hydrogen bonds between the coordinating sulfurs and the backbone amide of Cys[3] and Met[7], as well as a π interaction between the first oxazolone ring and the phenol group of Tyr[4]. The thiols of the two Cys residues form a disulfide bond, however this is not coupled to Cu(I) binding. Reduction of the disulfide bond results in binding of a second Cu(I) although with lower affinity (131).
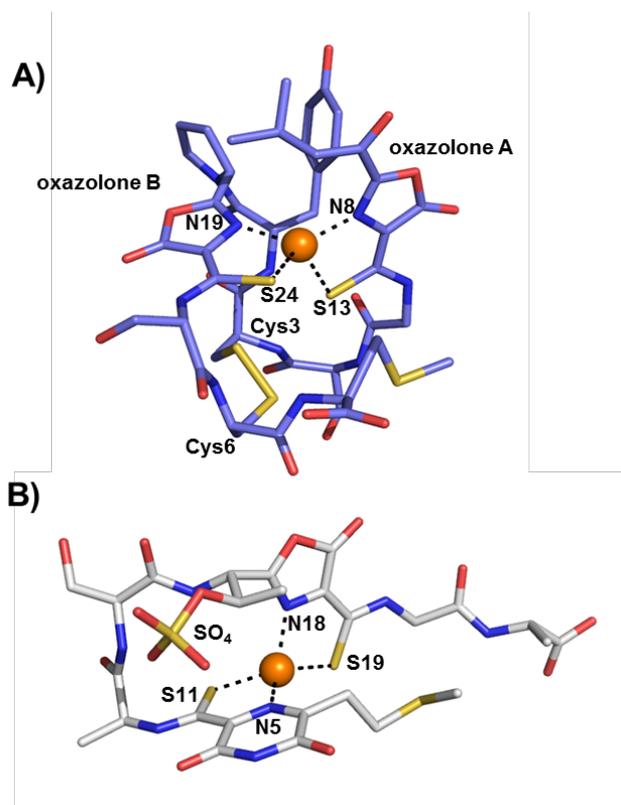
Figure 1.7 Crystal structures of Cu(I)-mb from *M. trichosporium* OB3b (A) and Cu(I)-mb-Thr from *M. hirsuta* CSC1 (B) with the Cu(I) ions represented as orange spheres (pdb codes 2xjh and 2ygi, respectively)) (126, 131).

The crystal structure of a Cu(I)-mb form produced by *Methylocystis hirsuta* CSC1(Figure 1.7 B) showed a different arrangement of the peptide. This molecule has a backbone consisting of four conventional amino acid residues and is rich in alanine,

while two additional modified residues include oxazolone and pyrazinedione rings (126). The overall arrangement of the molecule forms a hairpin-like structure, less compact than that from *M. trichosporium* OB3b. Cu(I) is coordinated in a distorted tetrahedral geometry as before, by a set of N_2S_2 ligands. The presence of a sulfate group attached to the threonine side chain in this mb creates additional stabilising interactions, as it forms a hydrogen bond with the backbone amide of Ser^2 and also a π-anion interaction with the pyrazinedione ring, which result in Cu(I) affinity one order of magnitude higher, compared to *M. trichosporium* OB3b, at 10^{21} M^{-1} (126).

1.4.7.3 MOB produce different mb forms

 Except for the full length form, described so far, *M. trichosporium* OB3b produces a second shorter form of mb, which is lacking the C-terminal Met residue. This shorter form is not a degradation product of full length mb but instead is produced by the organism, although the growth conditions that induce the production of each form have not been determined. Nevertheless, the affinity of both forms for Cu(I) is very high at the range of 6-7 * 10^{20} M^{-1} which is consistent with the role of mb in copper uptake (131). Production of more than one methanobactin forms with high affinity for Cu(I) has also been observed in *Methylocystis* strains, with strains *hirsuta* CSC1 and M producing a full length mb form and an mb that is missing the C-terminal Thr residue. The broad range in reduction potential of the *Methylocystis* strains has been proposed to correlate with the mechanism by which copper is released from mb (126).

1.4.7.4 Role of methanobactin

 Mb has been suggested to mediate copper uptake in MOB as secretion of the peptide has been observed under copper limiting growth conditions (130). C(I)-mb is also thought to mediate the switch between sMMO and pMMO expression in experiments where pmoA and mmoX transcript levels were monitored along with sMMO activity (131, 133). One suggested role of mb is the solubilisation of mineral copper from the environment and studies have shown that the effect of mb on the methane oxidation rate of cells exposed to mineral copper sourses (such as borosilicate glass) was comparable to those of cells grown on soluble copper (133). Despite these observations, how exactly mb delivers copper to methanotrophs is not understood. pMMO is the main methane oxidizing form and requires copper for activity, which renders this enzyme the primary target for copper delivery inside the cytoplasm of MOB, while some reports mention that mb has been found associated to pMMO in purifications of the enzyme (134, 135).

Based on these data mb has been thought to directly interact with pMMO in vivo and studies have reported the peptide is necessary for pMMO activity as it increases electron flow to the enzyme (135). An additional role for reducing copper toxicity has been suggested for mb on the basis of the peptide having superoxide dismutase activity (134) as well as preventing the accumulation of unchelated copper. In experiments where transcription of 16S-rRNA was monitored in correlation to copper levels in the culture media, addition of mb to the media at a 1:1 ratio to copper delayed the decrease of transcription, indicating the role of mb in regulating copper uptake (133).

Direct evidence for the internalization of methanobactin is provided by a study where uptake of copper and the peptide was monitored by isotopic and fluorescent labelling (43). In this study, the copper content of *M. trichosporium* OB3b cells increased when the cells were exposed to ^{65}Cu-mb, indicating that the intact peptide is internalized. Furthermore, labeling with a fluorescent probe enabled the visualization of mBBr-Cu-methanobactin in the cytoplasm of *M. trichosporium* OB3b. Treatment of the cells with inhibitors specific to porins and TonB-dependent transporters indicated that the copper peptide is actively transported across the outer membrane, whereas copper ions passively diffuse though porins (43) (see also section 2.1.1). The only known requirement for cytoplasmic copper in MOB is derived by pMMO that is localized in intra-cytoplasmic membranes. Nevertheless, it is unclear whether the intra-cytoplasmic membranes that house pMMO are invaginations of the cytoplasmic membrane or form separate subcellular compartments (139). A further implication for copper trafficking arises from the high Cu(I) affinity of mb forms. Known protein components for copper homeostasis (CopA, CueR CopZ), are present in methanotrophs, as indicated by bioinformatic analysis (32, 137, 138), however little is known on how MOB handle the large amounts of copper they require for pMMO activity (Figure 1.8).
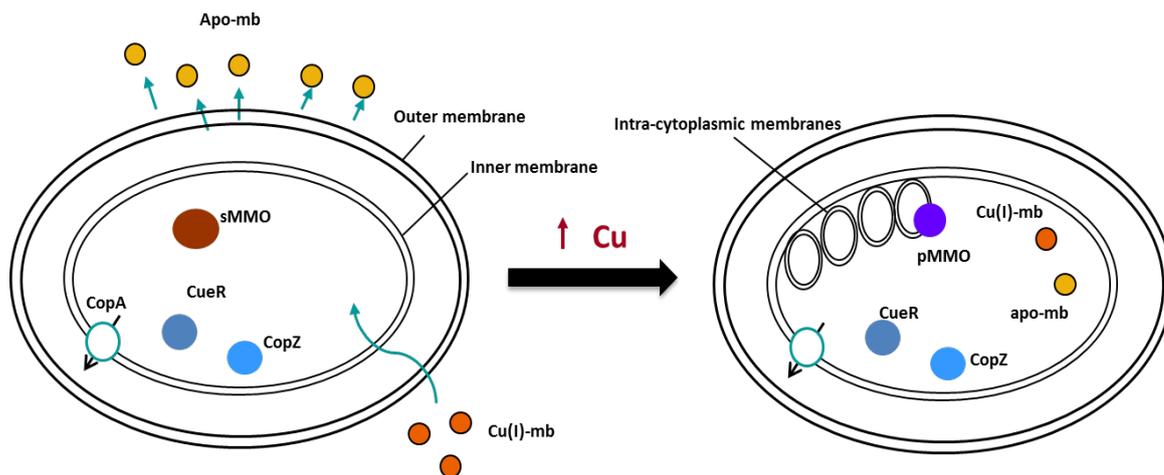
Figure 1.8 Schematic representation of copper metabolism in MOB and copper-dependent 'switchover' between the iron-containing sMMO and the copper-containing pMMO, which is located in intra-cytoplasmic membranes.

1.5 Aims and Objectives

The biotechnological applications of MOB, including the potential of hydrogen fuel production from liquid methanol, renders these organisms environmentally important. The high requirement of MOB for copper is due to the copper-dependent enzyme pMMO that is primarily responsible for methane oxidation, and this copper demand is satisfied by the production and secretion of mb, a well-characterised peptide with high affinity for Cu(I). How copper is transported from internalised Cu(I)-mb to pMMO is not understood and it is possible that unknown copper proteins participate in this process. In this light, the current work aimed to:

- Investigate whether mb provides copper to any soluble copper proteins, which may facilitate copper transport from mb to pMMO, and identify these.
- Characterise *in vitro* the identified proteins with emphasis on their structure and copper binding properties.
- Determine the crystal structures of the identified proteins and study the coordination of copper in order to gain insight in the mechanism of copper binding and release.
- Investigate whether homologues of the proteins studied from *M. trichosporium* OB3b are present in other organisms, including MOB.

1.6 References

1. Bertini I., Cavallaro G., Metals in the "omics" world: copper homeostasis and cytochrome c oxidase assembly in a new light. J. Biol. Inorg. Chem., **2008**, 13, 3-14

2. Andreini C., Bertini I., Cavallaro G., Holliday G. L., Thornton J. M., Metal ions in biological catalysis: from enzyme databases to general principles, J. Biol. Inorg. Chem, **2008**, 13, 1205-1218

3. Williams R. J. P., Fraústo do Silva J. R. R., The chemistry of evolution: the development of our ecosystem, Elsevier, **2006**

4. Fraústo do Silva J. R. R., Williams R. J. P., The biological chemistry of the elements: the inorganic chemistry of life, Oxford university press, Second edition **2001**

5. Lüthi D., Günzel D., McGuigan J. A., Mg-ATP binding: its modification by spermine, the relevance to cytosolic Mg2+ buffering, changes in the intracellular ionized Mg2+ concentration and the estimation of Mg2+ by 31P-NMR, Exp. Physiol., **1999**, 84, 231-252

6. Yang L., Arora K., Beard W. A., Wilson S. H., Schlick T., Critical role of magnesium ions in DNA polymerase β's closing and active site assembly, J. Am. Chem. Soc., **2004**, 126, 8441-8453

7. Carafoli E., Calcium signalling: a tale for all seasons, Proc. Natl. Acad. Sci., 2002, 99, 1115-1122

8. Essen L-O, Perisic, O., Katan M., Wu Y., Roberts M. F., Williams R. L., Structural mapping of the catalytic mechanism for a mammalian phosphoinositide-specific Phospholipace C, Biochemistry, **1997**, 36, 1704-1718

9. Giedroc D. P., Arunkumar A. I., Metal sensor proteins: nature's metalloregulated allosteric switches, Dalton Trans., 2007, 29, 3107-3120

10. Finney L. A., O'Halloran T. V., Transition metal speciation in the cell: insights from the chemistry of metal ion receptors, Science, **2003**, 300, 931-936

11. Kimura E., Roles of Zn(II) ion in zinc enzymes, Pure and Appl. Chem., **1993**, 65, 355-359

12. Romero-Isart N., Vasak M., Advances in the structure and chemistry of metallothioneins, J. Inorg. Biochem., **2002**, 88, 388-396

13. Sénèque O., Latour J-M., Coordination properties of zinc finger peptides revisited: ligand competition studies reveal higher affinities for zinc and cobalt, J. Am. Chem. Soc., **2010**, 132, 17760-17774

14. Pittman J. K., Managing the manganese: molecular mechanisms of manganese transport and homeostasis, New Phytol., **2005**, 167, 733-742

15. Banci L., Bertini I., Gori Savellini G., Luchinat C., Individual reduction potentials of the iron ions in Fe_2S_2 and high-potential Fe_4S_4 ferredoxins, Inorg. Chem., **1996**, 35, 4248-4253

16. Kopp D. A., Lippard S. J., Soluble methane monooxygenase: activation of dioxygen and methane, Curr. Opin. Chem. Biol., **2002**, 6, 568-576

17. Kovaleva E. G., Neibergall M. B., Chakrabarty S., Lipscomb J. D., Finding intermediates in the O_2 activation pathways in non-heme iron oxygenases, Acc. Chem., Res., **2007**, 40, 475-483

18. Banerjee R., Ragsdale S. W., The many faces of vitamin B_{12}: catalysis by cobalamine-dependent enzymes, Annu. Rev. Biochem., **2003**, 72, 209-247

19. Boer J. L., Mulrooney S. B., Hausinger R. P., Nickel-dependent metalloenzymes, Arch. Biochem. Biophys., **2014**, 0, 142-152

20. Chaturvedi K. S., Henderson J. P., Pathogenic adaptations to host-derived antibacterial copper, Front. Cell. Infect. Microbiol., **2014**, 3

21. Dupont C. L., Grass G., Rensing C., Copper toxicity and the origin of bacterial resistance- new insights and applications, Metallomics, **2011**, 3 1109-1118

22. Ridge P. G., Zhang Y., Gladyshev V. N., Comparative genomic analyses of copper transporters and cuproproteomes reveal evolutionary dynamics of copper utilisation and its link to oxygen, PLoS ONE, **2008**, 1

23. Dennison C., Investigating the structure and functions of cupredoxins, Coord. Chem. Rev., **2005**, 249, 3025-3054

24. Solioz M., Abicht H. K., Mermod M., Mancini S., Response of gram-positive bacteria to copper stress, J. Biol. Inorg. Chem., **2010**, 15, 3-14

25. Cavet J. S., Borrelly G. P., Robinson N. J., Zn, Cu and Co in cyanobacteria: selective control of metal availability, FEMS Microbiol Rev, **2003**, 27, 165–181

26. Rapisarda V. A., Chehin R. N., De Las R. J., Rodriguez-Montelongo L., Farias R. N., Massa E. M., Evidence for Cu(I)-thiolate ligation and prediction of a putative copper-binding site in the *Escherichia coli* NADH dehydrogenase-2, Arch. Biochem. Biophys., **2002**, 405, 87–94

27. Rodriguez-Montelongo L., Volentini S. I., Farias R. N., Massa E. M., Rapisarda V. A., The Cu(II)-reductase NADH dehydrogenase-2 of *Escherichia coli* improves the bacterial growth in extreme copper concentrations and increases the resistance to the damage caused by copper and hydroperoxide Arch Biochem Biophys, **2006**, 451, 1–7

28. Arciero D. M., Pierce B. S., Hendrich M. P., Hooper A. B., Nitrosocyanin, a red cupredoxin-like protein from Nitrosomonas europaea, Biochemistry, **2002**, 41, 1703-1709

29. Ellis M. J., Grossmann J. G., Eady R. R., Hasnain S. S., Genomic analysis reveals widespread occurrence of new classes of copper nitrite reductases, J. Biol. Inorg. Chem., **2007**, 12, 1119-1127

30. Tsai T. Y., Lee Y. H., Roles of copper ligands in the activation and secretion of Streptomyces tyrosinase, J. Biol. Chem., **1998**, 273, 19243–19250

31. Brazeau B. J., Johnson B. J., Wilmot C. M., Copper-containing amine oxidases. Biogenesis and catalysis; a structural perspective, Arch. Biochem. Biophys., **2004**, 428, 22–31

32. Culpepper M. A., Rosenzweig A. C., Architecture and active site of particulate methane monooxygenase, Crit. Rev. Biochem. Mol. Biol., **2012**, 47, 483-492

33. Hullo M-F., Moszer i., Danchin A., Martin-Verstraete I., CotA from Bacillus subtilis is a copper-dependent laccase, J. Bacteriol., **2001**, 183, 5426-5430

34. Imlay J. A, Chin S. M., Stuart Linn, Toxic DNA damage by hydrogen peroxide through the fenton reaction in vivo and in vitro, Science, **1988**, 240, 640-642

35. Macomber L., Rensing C., Imlay J. A., Intracellular copper does not catalyse the formation of oxidative DNA damage in E. coli, J. Bacteriol., **2007**, 189, 1616-1626

36. Macomber L., Imlay J. A., The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity, Proc. Natl. Acad. Sci., **2009**, 106, 8344-8349

37. Bird L. J., Coleman M. L., Newman D. K., Iron and copper act synergistically to delay anaerobic growth of bacteria, Appl. Environ. Microbiol., **2013**, 79, 3619-3627

38. Djoko K. Y., McEwan A. G., Antimicrobial action of copper is amplified via inhibition of heme biosynthesis, ACS Chem. Biol., **2013**, 8, 2217-2223

39. Dixon S. J., Stockwell B. R., The role of iron and reactive oxygen species in cell death, Nat. Chem. Biol., **2014**, 10, 9-17

40. Davis A. V., O'Halloran T. V., A place for thioether chemistry in cellular copper

ion recognition and trafficking, Nat. Chem. Biol., **2008**, 4, 148-151

41. Badarau A., Firbank S. J., McCarthy A. A., Banfield M. J., Dennison C., Visualising the metal-binding versatility of copper trafficking sites, Biochemistry, **2010**, 49, 7798-7810

42. Borrelly G. P., Blindauer C. A., Schmid R., Butler C. S, Cooper C. E., Harvey I., Sadler P. J., Robinson N. J., A novel copper site in a cyanobacterial chaperone, Biochem. J., **2004**, 378, 293-297

43. Balasubramanian R., Kenney G. E., Rosenzweig A. C., Dual pathways for copper uptake by methanotrophic bacteria, J. Biol. Chem., **2011**, 286, 37313-37319

44. Gourdon P., Liu X. Y., SkjørRinge T., Morth J. P., Møller L. B., Pedersen B. P., Nissen P., Crystal structure of a copper- transporting PIB-type ATPase, Nature, **2011**, 475, 59-64

45. Raimunda D., González-Guerrero M., Leeber B. W. 3rd, Argüello J. M., The transport mechanism of bacterial Cu+-ATPases: distinct efflux rates adapted to different function, Biometals, **2011**, 24, 467-475

46. White C., Lee J., Kambe T., Fritsche K., Petris M. J., A role for the ATP7A copper-transporting ATPase in macrophage bactericidal activity, J. Biol. Chem., **2009**, 284, 33949-33956

47. Radford D. S., Kihlken M. A., Borrelly G. P., Harwood C. R., LeBrun N. E., Cavet J. S., CopZ from Bacillus subtilis interacts in vivo with a copper exporting CPx-type ATPase CopA, FEMS Microbiol. Lett., 2003, 220, 105-112

48. Rodriguez-Granillo A., Wittung-Stafshede P., Tuning of copper loop flexibility in B. subt CopZ: role of conserved residues, J. Phys. Chem. B, **2009**, 113, 1919-1932

49. Banci L., Bertini I., Del Conte R., Solution structure of apo-CopZ from Bacillus subtilis: further analysis of the changes associated with the presence of copper, Biochemistry, **2003**, 42, 13422-13428

50. Kihlken M. A., Leech A. P., Le Brun N. E., Copper-mediated dimerization of CopZ, a predicted copper chaperone from Bacillus subtilis, Biochem. J., **2002**, 368, 729-739

51. Hearnshaw S., West C., Singleton C., Zhou L., Kihlken M. A., Strange R. W., Le Brun N. E., Hemmings A. M., A tetranuclear Cu(I) cluster in the metallochaperone protein CopZ, Biochemistry, **2009**, 48, 9324-9326

52. Alvarez H. M., Xue Y., Robinson C. D., Canalizo-Hermández M. A., Marvin R. G., Kelly R. A., Mondragón A., Penner-Hahn J. E., O'Halloran T. V.,

Tetrathiomolybdate inhibits copper trafficking proteins through metal cluster formation, Science, **2010**, 327, 331-334

53. Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V., Mondragón A., Molecular basis of the metal- ion selectivity and zeptomolar sensitivity by CueR, Science, **2003**, 301, 1383-1387

54. Rao L. Cui Q., Xu X., Electronic properties and desolvation penalties of metal ions plus protein electrostatics dictate the metal binding affinity and selectivity in the copper efflux regulator, J. Am. Chem. Soc., **2010**, 132, 18092-18102

55. Cobine P. A., George G. N., Jones C. E, Wickramasinghe W. A. Solioz M., Dameron C. T., Copper transfer from the Cu(I) chaperone, CopZ, to the repressor, Zn(II)CopY: metal coordination environments and protein interactions, Biochemistry, **2002**, 41, 5822-5829

56. Cobine P. A., Jones C. E, Dameron C. T., Role for zinc(II) and copper(I) regulated protein CopY, J. Inorg. Biochem., **2002**, 88, 192-196

57. Liu T., Ramesh A., Ma Z., Ward S. K., Zhang L., George G. N., Talaat A. M., Sacchettini J. C., Giedroc DP, CsoR is a novel Mycobacterium tuberculosis copper sensing transcriptional regulator, Nat. Chem. Biol., **2007**, 3. 60-68

58. Higgins K. A., Giedroc DP, Insights into protein allostery in the CsoR/RcnR family of transcriptional repressors, Chem. Lett., **2014**, 43, 20-25

59. Ma Z., Cowart D. M., Scott R. A., Giedroc DP, Molecular insights into the metal selectivity of the Cu(I)-sensing repressor CsoR from Bacillus subtilis, Biochemistry, **2009**, 48, 3325-3334

60. Coyne H. J. 3rd, Giedroc DP, Backbone resonance assignments of the homotetrameric (48 kD) copper sensor CsoR from Geobacillus thermodenitrificans in the apo- and Cu(I)-bound states: insights into copper-mediated allostery, Biomol. NMR Assign., **2013**, 7, 279-283

61. Sakamoto K., Agari Y., Agari K., Kuramitsu S., Shinkai A., Structural and functional characterisation of the transcriptional repressor CsoR from Thermus thermophilus HB8, Microbiology, **2010**, 156, 1993-2005

62. Singh S. K., Roberts S. A., McDevitt S. F., Weichsel A., Wildner G. F., Grass G. B., Rensing C., Montfort W. R., Crystal structures of multicopper oxidase CueO bound to copper (I) and silver (I): functional role of a methionine-rich sequence, J. Biol. Chem., **2011**, 286, 37849-37857

63. Roberts S. A., Weichsel A., Grass G., Thakali K., Hazzard J. T., Tolin G., Rensing C., Montfort W. R., Crystal structure and electron transfer kinetics of CueO, a multicopper oxidase required for copper homeostasis in Escherichia coli, Proc. Natl. Acad. Sci., **2002**, 99, 2766-2771

64. Outten F. W., Huffman D. L., Hale J. A., O'Halloran T. V., The independent *cue* and *cus* systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli, J. Biol. Chem., **2001**, 276, 30670-30677

65. Xue Y., Davis A. V., Balakrishnan G., Stasser J. P., Staehlin B. M., Focia P., Spiro T. G., Penner-Hahn J. E., O'Halloran T. V., Cu(I) recognition via cation π and methionine interactions in CusF, Nat. Chem. Biol., **2008**, 4, 107-109

66. Kim E. H., Nies D. H., McEvoy M. M., Rensing C., Switch or funnel: how RND-type transport systems control periplasmic metal homeostasis, J. Bacteriol., **2011**, 193, 2381-2387

67. Munson G. P., Lam D. L., Outten F. W., O'Halloran T. V., Identification of a copper-responsive two component system on the chromosome of Escherichia coli K-12, J. Bacteriol., **2000**, 182, 5864-5871

68. Su C. C., Long F., Zimmermann M. T., Rajashankar K. R., Jernigan R. L., Yu E. W., Crystal structure of the CusBA heavy-metal efflux complex of Escherichia coli, Nature, **2011**, 470, 558-562

69. Long F., Su C. C., Zimmermann M. T., Boyken S. E., Rajashankar K. R., Jernigan R. L., Yu E. W., Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport, Nature, **2010**, 467, 484-488

70. Banci L., Bertini I., Ciofi-Baffoni S., Su X. C., Borrelly G. P., Robinson N. J., Solution structures of a cyanobacterial metallochaperone: Insight into an atypical copper-binding motif, J. Biol. Chem., **2004**, 279, 27502-27510

71. Badarau A., Baslé A., Firbank S. J., Dennison C., Crosstalk between Cu(I) and Zn(II) homeostasis via Atx1 and cognate domains, Chem. Commun., **2013**, 49, 8000-8002

72. Badarau A., Dennison C., Thermodynamics of copper and zinc distribution in the cyanobacterium Synechocystis PCC 6803, Proc. Natl. Acad. Sci., **2011**, 108, 13007-13012

73. Tottey S., Rich P. R., Rondet S. A., Robinson N. J., Two Menkes-type ATPases supply copper for photosynthesis in Synechocystis PCC 6803, J. Biol. Chem., **2001**, 276, 19999-20004

74. González-Guerrero M., Raimunda D., Cheng X., Argüello J. M., Distinct functional roles of homologous Cu^+ efflux ATPases in Pseudomonas aeruginosa, Mol. Microbiol., **2010**, 78, 1246-1258

75. Romero-Isart N., Vasak M., Advances in the structure and chemistry of metallothioneins, J. Inorg. Biochem., **2002**, 88, 388-396

76. Blindauer C. A., Bacterial metallothioneins: past, present, and questions for the future, J. Biol. Inorg. Chem., **2011**, 16, 1011-1024

77. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

78. Semrau J. D., DiSpirito A. A., Yoon S., Methanotrophs and copper, FEMS Microbiol. Rev., **2010**, 34, 496-531

79. Hanson R. S., Hanson T. E., Methanotrophic bacteria, Microbiol. Rev., **1996**, 60, 439-471

80. Murrell J. C., McDonald I. R., Gilbert B., Regulation of expression of methane monooxygenases by copper ions, Trends in Microbiol., **2000**, 8, 221-225

81. Nazaries L., Murrell J. C., Millard P, Baggs L., Singh B. K., Methane, microbes and models: fundamental understanding of the soil methane cycle for future predictions, Environ. Microbiol., **2013**, 15, 2395-2417

82. Op den Camp H. J. M., Islam T, Stott M. B., Harhangi H. R., Hynes A., Schouten S., Jetten M. S. M., Birkeland N-K., Pol1 A., Dunfield P. F., Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia, Environ. Microbiol. Rep., **2009**, 1, 293-306

83. Debysh S. N., Knief C., Dunfield P. F., Methylocella species are facultatively methanotrophic, J. Bact., **2005**, 187, 4665-4670

84. Hakemian A. S., Rosenzweig A. C., The biochemistry of methane oxidation, Annu. Rev. Biochem., **2007**, 76, 223-241

85. Martinho M., Choi D. W., DiSpirito A. A., Antholine W. E, Semrau J. D., Munck E., Mössbauer studies of a membrane-associated methane monooxygenase from Methylococcus capsulatus Bath: evidence for a diiron center, J. Am. Chem. Soc., **2007**, 129, 15783-15785

86. Theisen A. R., Ali M. H., Radayewski S., Dumont M. G., Dunfield P. F., McDonald I. R., Debysh S. N., Miguez C. B., Murrell J. C., Regulation of methane

oxidation in the facultative methanotroph Methylocella silvestris BL2, Mol. Microbiol., **2005**, 58, 682-692

87. Lieberman R. L., Rosenzweig A. C., Biological methane oxidation: regulation, biochemistry and active site structure of pMMO, Crit. Rev. Biochem. Mol. Biol., **2004**, 39, 147-164

88. Lieberman R. L., Rosenzweig A. C., Crystal structure of a membrane bound metalloenzyme that catalyses biological oxidation of methane, Nature, **2005**, 434, 177-182

89. Cargnello M., Deldado Jaén J. J., Hernández Garrido J. C., Bakhmutsky K., Mortini T., Calvino Gámez J. J., Gorte R. J., Fornasiero P., Exceptional activity for methane combustion over modular $PdCeO_2$ subunits on functionalised Al_2O_3, Science, **2012**, 337, 713-717

90. Nielsen M., Alberico E., Baumann W., Drexler H. J., Junge H., Gladiali S., Beller M., Low- temperature aqueous- phase methanol dehydrogenation to hydrogen and carbon dioxide, Nature, **2013**, 495, 85-89

91. Stephan D. W., A step closer to a methanol economy, Nature, **2013**, 495, 54-55

92. Baani M., Liesack W., Two isoenzymes of particulate methane monooxygenase with different methane oxidation kinetics are found in Methylocystis sp. Strain SC2, P. Natl. Ac. Sci., **2008**, 105, 10203-10208

93. Huber-Humer M., Gebert J., Hilger H., Biotic systems to mitigate landfill methane emissions, Waste Manag. Res., **2008**, 26, 33-46

94. Alvarez-Cohen L., McCarty P. L., Boulygina E.., Hanson R. S., Brusseau G. A., Tsien H. C., Characterisation of a methane utilising bacterium from a bacterial consortium that rapidly degrades trichloroethylene and chloroform, Applied and environmental microbiology, **1992**, 58, 1886-1893

95. Lee S.W., Keeney D. R., Lim D. H., Dispirito A. A., Semrau J. D., Mixed pollutant degradation by M. trichosporium OB3b expressing either soluble or particulate methane monooxygenase: can the tortoise beat the hare, Applied and Environ. Microbiol., **2006**, 72, 7503-7509

96. Phelps P. A., Agarwal S. K., Speitel G. E., Georgiou G., Methylosinus trichosporium OB3b mutants having constitutive expression of sMMO in the presence of high levels of copper, App. Environ. Microbiol, **1992**, 58, 3701-3708

97. Merkx M., Kopp D. A., Sazinsky M. H., Blazyk J. L., Müller J., Lippard S. J., Dioxygen activation and methane hydroxylation by sMMO: a tale of two irons and three proteins, Angew. Chem. Int. Ed., **2001**, 40, 2782-2807

98. Cho U., Control of substrate access to the active site in methane monooxygenase, Nature, 2013,

99. Merkx M., Lippard S. J., Why OrfY? Characterisation of MMOD, a long overlooked component of the soluble methane monooxygenase from Methylococcus capsulatus (Bath), J. Biol. Chem., **2002**, 277, 58-58-5865

100. Scanlan J., Dumont M. G., Murrell J.C., Involvement of MmoR and MmoG in the transcriptional activation of soluble methane monooxygenase genes in Methylosinus trichosporium OB3b, FEMS Microbiol. Lett., **2009**, 301, 181-187

101. Smith S. M., Rawat S., Telser J., Hoffman B. M., Stemmler T. L., Rosenzweig A. C., Crystal structure and characterisation of particulate methane monooxygenase from Methylocystis species strain M, Biochemistry, **2011**, 50, 10231-10240

102. Hakemian A. S., Kondapalli K. S., Telser J., Hoffman B. M., Stemmler T. L., Rosenzweig A. C., The metal centers of particulate methane monooxygenase from *Methylosinus trichosporium* OB3b, Biochemistry, **2008**, 47, 6793-67901

103. Kitmitto A. Myronova N., Basu P., Dalton H., Characterisation and structural analysis of an active particulate methane monooxygenase trimer from Methylococcus capsulatus (Bath), Biochemistry, **2005**, 44, 10954-10965

104. Balasubramanian R., Smith S. M., Rawat S., Yatsunyk L. A., Stemmler T. L., Rosenzweig A. C., Oxidation of methane by a biological dicopper center, Nature, **2010**, 465, 115-119

105. Yu S. S., Ji C. Z., Wu Y. P., Lee T. L., Lai C. H., Lin S. C., Yang Z. L., Wang V. C., Chen K. H., Chan S. I., The C-terminal Aqueous exposed domain of the 45 kDa subunit of the particulate methane monooxygenase in Methylococcus capsulatus (Bath) is a Cu(I) sponge, Biochemistry, **2007**, 46, 13762-13774

106. Nguyen H-H. T., Nakagawa K. H., Hedman B., Elliott S. J., Lidstrom M. E., Hodgson K. O., Chan S. I., X-ray absorption and EPR studies on the copper ions associated with particulate methane monooxygenase from M. capsulatus (Bath). Cu(I) ions and their implications, J. Am. Chem. Soc., **1996**, 118, 12766-12776

107. Chan S. I., Yu S. F., Controlled oxidation of hydrocarbons by the membrane-bound methane monooxygenase: the case for a tricopper cluster, Acc. Chem. Res., **2008**, 41, 969-979

108. Zahn J. A., DiSpirito A. A., Membrane associated methane monooxygenase from Methylococcus capsulatus (Bath), J. Bacteriol., **1996**, 178, 1018-1029

109. Basu P., Katterle B., Andersson K. K., Dalton H., The membrane associated form of methane mono-oxygenase from Methylococcus capsulatus (Bath) is a copper/iron protein, Biochem. J., **2003**, 369, 417-427

110. Culpepper M. A., Cutsail G. E., Hoffman B. M., Rosenzweig A.C., Evidence for oxygen binding at the active site of pMMO, J. Am. Chem. Soc., 2012, 134, 7640-7643

111. Ukaegbu U. E., Henery S., Rosenzweig A. C., Biochemical characterisation of MmoS, a sensor protein involved in copper-dependent regulation of soluble methane monooxygenase, Biochemistry, **2006**, 45, 10191-10198

112. Karlsen O. A., Larsen Ø., Jensen H. B., The copper responding surfaceome of Methylococcus capsulatus Bath, FEMS Microbiol. Lett., **2011**, 323, 97-104

113. Karlsen O. A., Berven F. S., Stafford G. P., Larsen Ø., Murrell J. C., Jensen H. B., Fjellbirkeland A., The surface-associated and secreted MopE protein of Methylococcus capsulatus (Bath) responds to changes in the concentration of copper in the growth medium, Appl. Environ. Microbiol., **2003**, 69, 2386-2388

114. Ve T., Mathisen K., Helland R., Karlsen O. A., Fjellbirkeland A., Røhr Å. K., Andersson K. K., Pendersen R. B., Lillehaug J. R., Jensen H. B., The Methylococcus capsulatus (Bath) secreted protein, MopE*, binds both reduced and oxidised copper, Plos One, **2012**, 7

115. Helland R., Fjellbirkeland A., Karlsen O. A., Ve T., Jensen H. B., An oxidised tryptophan facilitates copper binding in Methylococcus capsulatus-secreted protein MopE, J. Biol. Chem., **2008**, 283, 13897-13904

116. Berson O., Lindstrom M. E., Cloning and characterisation of corA, a gene encoding a copper-repressible polypeptide in the type I methanotroph, Methylomicrobium albus BG8, FEMS Microbiol. Lett., **1997**, 148, 169-174

117. Johnson K. A., Ve T., Larsen Ø., Pedersen R. B., Lillehaug J. R., Jensen H. B., Helland R., Karlsen O. A., CorA is a copper repressible surface-associated copper (I)-binding protein produced in Methylomicrobium album BG8, Plos One, **2014**, 9

118. Karlsen O. A., Larsen Ø., Jensen H. B., Identification of a bacterial di-haem cytochrome c peroxidase from Methylomicrobium album BG8, Microbiology, 2010, 156, 2682-2690

119. Karlsen O. A., Kindingstad L., Angelskår S. M., Bruseth L. J., Straume D., Puntervoll P., Fjellbirkeland A., Lillehaug J. R., Jensen H. B., Identification of a copper-repressible c-type heme protein of Methylococcus capsulatus (Bath), FEBS J., **2005**, 272, 6324-6335

120. Téllez C. M., Gauz K. P., Graham D. W., Arnold R. G., Guzman R. Z., Isolation of copper biochelates from M. trichosporium OB3b and sMMO mutants, Appl. Environ. Microbiol., **1998**, 64, 1115-1122

121. Kim H. J., Galeva N., Larive C. K., Alterman M., Graham D. W., Purification and physical- chemical properties of methanobactin: a chalkophore from M. trichosporium OB3b, Biochemistry, **2005**, 44, 5140-5148

122. DiSpirito A. A., Zahn J. A., Graham D. W., Kim H. J., Larive C. K., Derrick T. S., Cox C. D., Taylor A., Copper binding compounds from M. trichosporium OB3b, J. Bacteriol., **1998**, 180, 3606-3613

123. Choi D. W., Kunz R. C., Boyd D. S., Antholine W. E., Han J. I., Zahn J. A., Boyd J. M., de la Mora A. M., Semrau J. D., DiSpirito A. A., The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH: quinone oxidoreductase complex from Methylococcus capsulatus Bath, J. Bacteriol., **2003**, 185, 5755-5764

124. Semrau J. D., Jagadevan S., DiSpirito A. A., Khalifa A., Scanlan J., Bergman B. H., Freemeler B. C., Baral B. S., Bandow N. L., Vorobev A., Haft D. H., Vuilleumier S., Murrell J. C., Methanobactin and MmoD work in concert to act as the 'copper-switch' in methanotrophs, Environ. Microbiol., **2013**, 15, 3077-3086

125. Krentz B. D., Mulheron H. J., Semrau J. D., DiSpirito A. A., Bandow N. L., Haft D. H., Vuilleumier S., Murrell J. C., McEllistrem M. T., Hartsel S. C., Gallagher W. H., A comparison of methanobactins from Methylosinus trichosporium OB3b and Methylocystis strain SB2 predicts methanobactins are synthesized from diverse peptide precursors modified to create a common core for binding and reducing copper ions, Biochemistry, **2010**, 49, 10117-10130

126. El Ghazouani A., Baslé A., Gray J., Graham D. W., Firbank S. J., Dennison C., Variations in methanobactin structure influence copper utilization by methane-oxidising bacteria, Proc. Natl. Acad. Sci., **2012**, 109, 8400-8404

127. Behling L. A., Hartsel S. C., Lewis D. E., DiSpirito A. A., Choi D. W., Masterson L. R., Veglia G., Gallagher W. H., NMR, mass spectrometry and chemical evidence reveal a different chemical structure for methanobactin that contains oxazolone rings, J. Am. Chem. Soc., **2008**, 130, 12604-12605

128. Kim H. J., Graham D. W., DiSpirito A. A., Alterman M. A., Galeva N., Larive C. K., Asunskis D., Sherwood P. M. A., Methanobactin, a copper acquisition compound from methane oxidising bacteria, Science, **2004**, 305, 1612-1615

129. Kenney G. E., Rosenzweig A. C., Chemistry and biology of the copper chelator methanobactin, ACS Chem. Biol, **2012**, 7, 260-268

130. Hakemian A. S., Tinberg C. E., Kondapalli K. C., Telser J., Hoffman B. M., Stemmler T. L., Rosenzweig A. C., The copper chelator methanobactin from Methylosinus trichosporium OB3b binds copper (I), J. Am. Chem. Soc., **2005**, 127, 17142-17143

131. El Ghazouani A., Baslé A., Firbank S. J., Knapp C. W., Gray J., Graham D. W., Dennison C., Copper-binding properties and structures of methanobactins from Methylosinus trichosporium OB3b, Inorg. Chem., **2011**, 50, 1378-1391

132. Balasubramanian R., Rosenzweig A. C., Copper methanobactin: a molecule whose time has come, Curr. Opin. Chem. Biol., **2008**, 12, 245-249

133. Knapp C. W., Fowle D. A., Kulczycki E., Roberts J. A., Graham D. W., Methane monooxygenase gene expression mediated by methanobactin in the presence of mineral copper sources, Proc. Natl. Acad. Sci., **2007**, 104, 12040-12045

134. Choi D. W., Semrau J. D., Antholine W. E., Hartsel S. C., Anderson R. C., Carey J. N., Dreis A. M., Kenseth E. M., Renstrom J. M., Scardino L. L., Van Gorden G. S., Volkert A. A., Wingad A. D., Yanzer P. J., McEllistrem M. T., de la Mora A. M., DiSpirito A. A., Oxidase, superoxide dismutase and hydrogen peroxide reductase activities of methanobactin from types I and II methanotrophs, J. Inorg. Biochem., **2008**, 102, 1571-1580

135. Choi D. W., Antholine W. E., Do Y. S., Semrau J. D., Kisting C. J., Kunz R. C., Campbell D., Rao V., Hartsel S. C., DiSpirito A. A., Effect of methanobactin on the activity and electron paramagnetic resonance spectra of the membrane associated methane monooxygenase in M. capsulatus Bath, Microbiology, **2005**, 151, 3417-3426

136. Ward N., Larsen Ø., Sakwa J., Bruseth L., Khouri H., Durkin A. S., Dimitrov G., Jiang L., Scanlan D., Kang K. H., Lewis M., Nelson K. E., Methé B., Wu M., Heidelberg J. F., Paulsen I. T., Fouts D., Ravel J., Tettelin H., Ren Q., Read T., DeBoy R. T., Seshadri R., Salzberg S. L., Jensen H. B., Birkeland N. K., Nelson W. C., Dodson R. J., Grindhaug S. H., Holt I., Eidhammer I., Jonasen I., Vanaken S., Utterback T., Feldblyum T. V., Fraser C. M., Lillehaug J. R., Eisen J. A., Genomic

insights into methanotrophy: the complete genome sequence of Methylococcus capsulatus (Bath), PLoS Biol., **2004**, 2, 1616-1628

137. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., **2010**, 192, 6497-6498

138. Stein L. Y., Bringel F., Dispirito A. A., Han S., Jetten M. S., Kalyuzhnaya M. G., Kits K. D., Klotz M. G., Op den Camp H. J., Semrau J. D., Vuilleumier S., Bruce D., Cheng J. F., Davenport K. W., Goodwin L., Han S., Hauser L., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Genome sequence of the methanotrophic alphaproteobacterium Methylocystis sp. Strain Rockwell (ATCC49242), J. Bacteriol., **2011**, 193, 2668-2669

139. Davies S. L., Whittenbury R., Fine structure of methane and other hydrocarbon-utilizing bacteria, J. Gen. Microbiol., **1970**, 61, 227-232

CHAPTER 2:

Identification of a novel copper protein from *M. trichosporium* OB3b

2.1 Introduction

2.1.1 Methanobactin mediates copper uptake in MOB

Methanobactin (mb) has a key role in copper uptake by methane oxidising bacteria (MOB) (1-3) as it has been shown to solubilise copper from mineral sources, thus making the metal available to MOB, while it is thought to mediate the switch between the expression of the copper-containing pMMO and the iron-containing sMMO (3, 4). The primary target for copper delivery in MOB is pMMO, a rare example of cytoplasmic copper enzyme that is housed in intra-cytoplasmic membranes (1, 5). However, the mechanism through which mb mediates copper uptake, which other protein components are involved in this process, and how copper is delivered to pMMO remain unclear. Different approaches have been used to study the uptake of Cu(I)-mb by a MOB in order to investigate the internalisation of the peptide (1, 2).

In one study the copper content of *M. trichosporium* OB3b cells exposed to either unchelated copper, provided as Cu(II), or Cu(I)-mb was shown increase by twofold, compared to the copper content of untreated cells, indicating that both copper forms are taken up equally by the organism. Subsequently ^{65}Cu was used in order to track the uptake of ^{65}Cu(I)-mb by *M. trichosporium* OB3b. When exposed to a mixture of either ^{65}Cu and Cu(I)-mb or Cu and ^{65}Cu(I)-mb, the copper content of *M. trichosporium* OB3b cells was again found to increase, while the ratio of ^{63}Cu/^{65}Cu decreased compared to that of naturally occurring copper (1). This experiment provided additional evidence of the simultaneous uptake of both unchelated ^{65}Cu and ^{65}Cu(I)-mb. Regardless of the high Cu(I) affinity of mb, which makes dissociation of Cu(I) from the peptide unlikely (2, 6), these experiments are not solid evidence for the internalisation of Cu(I)-mb by *M. trichosporium* OB3b. This was addressed by labelling the peptide with monobromobimane (mBBr), a probe that is fluorescent when bound to thiols. Confocal microscopy images visualised mBBr-Cu(I)-mb inside the cytoplasm of *M. trichosporium* OB3b where the most likely protein target for copper delivery is pMMO. Through this method mBBr-apo-mb was also visualised in the cytoplasm of the organism (1). While the experiment described showed apo- and Cu(I)-mb internalised in the cytoplasm of *M. trichosporium* OB3b, the nature of the fluorescent labelling probe (mBBr), which cleaves the disulfide bond of intact mb, should not be overlooked as it inevitably alters the peptide.

In a different study, the internalisation of intact Cu(I)-mb was investigated by monitoring the presence of mb in the culture medium and the ability of the peptide to induce switchover from sMMO to pMMO in *M. trichosporium* OB3b and *M. hirsuta* CSC1 (2). Although both strains were capable of internalising Cu(I)-mb forms produced by the other strain, switchover between sMMO and pMMO and copper uptake was most efficient when each MOB strain was treated with its native Cu(I)-mb (Cu(I)-mb produced by the strain itself). Moreover, a slower rate of copper uptake was observed in the case of shorter Cu(I)-mb forms, also produced by these bacterial strains (see section 1.4.7.3), and could be attributed to specific amino acid sequences that help in the recognition of the native mb by each organism (2). This study showed that mb-mediated copper uptake is responsible for the switchover between sMMO and pMMO but the mechanism is still not clear. The structural similarities of mbs from the nonswitchover organisms *Methylocystis* strains M and *rosea*, which do not express sMMO, with the ones used in the copper uptake experiment may imply that Cu(I)-mb does not participate directly in the switchover mechanism (2).

The possible routes for the internalisation of mb were investigated by using transport inhibitors to target two different import mechanisms of the outer membrane. The inhibition of porin transporters by spermine had an effect on the uptake of unchelated copper, indicating that copper ions passively diffuse through porins, whereas no difference was observed in the uptake of Cu(I)-mb (1). On the other hand, upon treatment of *M. trichosporium* OB3b with methylamine, a molecule that disrupts active transport by inhibiting inner membrane proton motive force, Cu(I)-mb uptake was inhibited indicating the import of Cu(I)-mb is an energy coupled process (1). Transporters that use the energy generated from the inner membrane proton motive force, such as TonB-dependent transporters, are present in the genome of *M. trichosporium* OB3b (7) and other MOB (8, 9) and are therefore a possible route for Cu(I)-mb import (1).

While the above experiments showed that mb is internalised by MOB, currently the mechanism of copper release from the peptide is unclear and may involve oxidation of Cu(I) to Cu(II) or degradation of the peptide (2). The Cu(I) affinity of mbs from various MOB strains have been determined and are in the range of 10^{20}-10^{21} M^{-1}, implying Cu(I) release from the peptide is not favoured (2, 6). Copper homeostasis proteins identified in other organisms are present in MOB (7-9), including a regulator

homologous to the copper sensor CueR from *E. coli*, the CopA copper transporting P-type ATPase and the metallochaperone CopZ. CueR regulates the expression of CopA and the multicopper oxidase CueO, both of which are part of the cue system (Cu efflux) that is responsible for copper stress response (10, 11), while CopZ acts with CopA to mediate copper export (12) (section 1.3.1-2). The Cu(I) affinities of CopZ (~10^{18} M^{-1}) (13-17) and CueR (~10^{20} M^{-1}) (10) are lower compared to that of mb indicating Cu(I) exchange between these proteins is unlikely to occur.

2.1.2 Ag(I) used as a probe for Cu(I)

Mb is part of the copper uptake system of MOB, which is necessary in order for these bacteria to satisfy their high copper requirement for the main methane oxidising enzyme, pMMO (1-5). It is not clear whether mb delivers Cu(I) directly to pMMO or whether there are other protein components involved in this process. In an attempt to visualise internalised mb inside *M. trichosporium* OB3b and investigate possible protein targets to which mb delivers Cu(I), Ag(I) was employed in the present work, based on its ability to act as a Cu(I) mimic.

Ag(I) acts as a Cu(I) mimic due to the analogous electronic configuration of the two elements ([Kr]$4d^{10}5s^1$ and [Ar] $3d^{10}4s^1$, respectively) as they belong to the same group of d-block elements. However, in contrast to copper that can be found in the Cu(I) or Cu(II) oxidation state, silver is redox-inert and is found predominantly as Ag(I) (18). Silver is one of the non-essential metals (metals that are not used by organisms for physiological functions and only induce toxicity) and is known to be toxic to bacteria, which use P-type ATPase transporters for metal detoxification and export Ag(I) (19, 20). On the other hand, ABC–type transporters, located in the inner membrane of Gram-negative bacteria, are known to import only essential metals in the bacterial cytoplasm (20). Ag(I) owes its toxicity in bacteria to the interaction with sulfhydryl groups resulting in the formation of Ag(I)-S bonds (21). As thiols are essential for the function of numerous enzymes, Ag(I) binding disrupts respiration and electron transfer mechanisms and, consequently, disables the proton motive force of the cell membrane (21). Additionaly, Ag(I) binds to [4Fe-4S] clusters leading to inactivation of the cluster-containing enzyme and release of iron from the disrupted cluster that can, subsequently, damage DNA and RNA (21, 22).

An example of how Ag(I)-binding to Cu(I)-proteins can induce toxicity is shown in the case of *E. coli*, which expresses to independent systems, cue (Cu efflux) and cus (Cu sensing) that are responsible for Cu(I) efflux (section 1.3.1-3) (23). The cue system consists of the the Cu(I)-binding proteins CueR, CopA and CueO (11, 23). The mismetallation of CueO by Ag(I) results in inhibition of the oxidase activity of the enzyme, as Ag(I) binds to the substrate-binding site of CueO that is necessary for full enzyme activity, and also to two more sites located in the methionine-rich region, thought to enable access to the substrate-binding site (11). Ag(I) also binds to CueR, where it is coordinated at the same site as Cu(I) with very similar geometry. As in the case of Cu(I), Ag(I) binding to CueR results in the transcriptional activation of CopA expression that exports Ag(I) in order to overcome toxicity (10). Ag(I) binding at the Cu(I) sites of the proteins of the cue system is thought to compromise the ability of the cue system for Cu(I) efflux, resulting in accumulation of Cu(I) in the cytoplasm. Cu(I) toxicity is overcome through activation of the cus system, which consists of the CusCBA and CusF proteins, and is switched on in order to expel Ag(I), or Cu(I), from the bacterial cell (23). This role is supported by studies showing that CusF as well as CusB and CusA bind Ag(I) (24, 25). In the case of CusA binding of Ag(I) induces the same conformational changes to the protein as Cu(I), which involves movement of the horizontal helix of the protein that allows the binding of the metal ion and subsequent transport across the channel for export (25). In the reports mentioned Ag(I) was used as a probe for studying the function and structure of Cu(I) proteins. A similar approach was used in the case *M. trichosporium* OB3b in this study, where Ag(I) was used in metalloproteomics experiments with *M. trichosporium* OB3b in order to target Cu(I) proteins to which mb may deliver copper.

2.1.3   Metalloproteomics

Metalloproteomics is a term used to describe the array of analytical techniques used to detect and identify metalloproteins in an organism. While a large number of metal-requiring proteins are predicted through bioinformatics (26), there is growing interest of this field in the detection and identification of novel metalloproteins. The interaction of proteins with metals depends on a number of factors that should be taken into account when analysing and interpreting metalloproteomic data. Such factors include exposure to oxygen during cell lysis, especially in the case of redox-active metalloproteins that contain reduced thiols from Cys residues (17, 27), the pH-dependent metal affinity of

proteins due to the pKa of certain residues, the chemical environment of proteins which inevitably changes due to lysis and metal contamination (27). A number of techniques are employed in metalloproteomics in order to resolve the soluble extract of the organism, quantify the metal content of these proteins and finally, identify the metal-bound protein. These techniques include native or denaturing gel electrophoresis, liquid chromatography, inductively coupled plasma mass spectrometry (ICP-MS), immobilised metal affinity chromatography, radiolabelling and mass spectrometry methods, the main advantages and drawbacks of which are discussed in the following paragraphs.

In one approach, ^{64}Cu radioisotopes were used to label metalloproteins in *E. coli* that were subsequently separated by native isoelectric focusing combined with blue native electrophoresis (28). Proteins are resolved in the first dimension of electrophoresis based on their isoelectric point, under native conditions, while the second electrophoresis provides further resolution based on molecular mass. Autoradiography was used to quantify ^{64}Cu on the gels and the concentration of the copper peaks measured was compared to the intensity of the protein spots on the 2D gel. Potential copper proteins were excised from the gel, digested with trypsin and analysed by liquid chromatography electrospray ionisation (LC-ESI) quadrupole time-of-flight (Q-TOF) tandem mass spectrometry (MS/MS) (28). The use of radiolabelling enables high sensitivity in the detection of copper proteins which are analysed under non-denaturing conditions in order to remain folded and with the native metal bound. However, the resolution of blue native electrophoresis is limited, compared to sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) (29), and a second electrophoresis step is needed. Another example where 2D native electrophoresis was used is in the case of *Ferroplasma acidiphilum* where a large number of iron-associated proteins were resolved. The metalloproteomic approach involved visualising of proteins on the 2D gel with the use of a chemiluminescent stain and identification of the excised proteins by matrix assisted laser desorption ionisation (MALDI)-TOF MS/MS, while the metal content was quanitified by ICP-MS (30).

An alternative approach for resolving the soluble proteins of an organism involves one or more steps of liquid chromatography. Such a method was used in *Pyrococcus furiosus*, several metalloproteins of which were purified by multiple chromatographic steps (31). This study focused on the identification of metals that were incorporated in

the organism, some of which were expected, such as cobalt, nickel and iron, while others are not known to be used by this organism such as lead, uranium and vanadium. The initial resolution of the metal pools of the organism was achieved by two steps of size-exclusion chromatography followed by ICP-MS. Subsequently, specific metal peaks were chosen for further purification by multiple chromatographic steps in order to obtain homogenous protein samples with stoichiometric amounts of metal bound (31). Although proteins were identified by this method, it must be noted that multiple chromatographic steps improve resolution but also compromise the yield of the obtained metalloprotein or lead to the dissociation of the metal from the protein of interest.

Finally, immobilised metal affinity chromatography (IMAC) is often used to detect proteins with affinity for specific metals, as in the case of *Synechococcus sp*. WH8102, where columns charged with Fe(III), Co(II) or Ni(II) were used (32). This technique resulted in the identification of rubrerythrin, a protein that contains an iron site. Although IMAC provides a useful tool for capturing of proteins with affinity for a certain metal, it also comes with some inherent limitations. The retention of nickel or cobalt affinity columns, for example, is dependent mainly on the presence of His residues in a protein (33), while neighbouring residues also influence binding (34). This can lead to false positive results and contamination of the protein sample as His residues are also involved in copper and zinc binding and, therefore, unoccupied His-containing sites of a copper or zinc metalloprotein could bind on an IMAC column (33, 35). In contrast, genuine nickel or cobalt metalloproteins could escape the IMAC column if they are isolated saturated with metal from the organism being analysed, or because the immobilisation of the metal on the column restricts the binding sites of the metalloprotein, leading to false negative results.

For analysing the soluble fraction of *M. trichosporium* OB3b a modified version of a published method (36) was used. The original method employed two dimensions of native chromatography, anion-exchange and size-exclusion, for resolving periplasmic extracts from S*ynechocystis* PCC 6803. The minimal chromatographic steps aimed at minimising contamination and false positive/negative results, while achieving efficient protein resolution. The elution fractions were analysed for various metals by ICP-MS, providing simultaneous quantification of multiple metals with high sensitivity, and the protein content of metal-containing fractions was visualised by SDS-PAGE. Principle component analysis was used to estimate the correlation between the intensity of protein

bands on the SDS-PAGE gels with metal concentration profiles. In the case of *M. trichosporium* OB3b the method aimed to investigate the copper binding proteins of the organism and potentially identify internalised mb. The use of Ag(I) as a probe for Cu(I)-containing soluble pools was tested and Ag(I) was also employed for the detection of internalised mb in *M. trichosporium* OB3b. The complexity of the proteins profiles visualised by SDS-PAGE gels did not allow the use of principle component analysis and the correlation of metal concentration to protein intensity was performed manually. Finally, in order to achieve better resolution of copper pools of interest, a linear NaCl gradient was introduced in the anion-exchange chromatography step.

2.2   Materials and Methods

2.2.1   *M. trichosporium* OB3b cultures

 M. trichosporium OB3b was grown from glycerol stocks as described in section 6.3.1 (6).   In order to prepare the cultures that were used for the small scale (2 L) metalloproteomics experiments (cell profiles) of *M. trichosporium* OB3b exposed to copper and silver, a culture of *M. trichosporium* OB3b was started in a 5 L fermentor (27 ℃, stirred at 250 rpm), in nitrate minimal salts (NMS) medium supplemented with 2 μM Cu(II) and 10 μM Fe(II) (culture A). When an OD_{600} of ⸴0.8 was reached, 1.6 L of the culture A was transferred to each of two smaller fermentors (3 L) and supplemented with 10 μM Fe(II) plus 1 μM Cu(II) (culture A1) or 1 μM Ag(I) (culture A2), respectively. Cultures A1 and A2 were incubated for 24 h and 1 L of each was harvested at $\sim OD_{600}$ 1.3. Cells were harvested by centrifugation and washed as described in section 6.3.2. For the cell profile of *M. trichosporium* OB3b exposed to Ag(I)-mb a culture of *M. trichosporium* OB3b was grown in a 3 L fermentor and 1 μM Ag(I)-mb was added when the culture reached an OD_{600} 0.8 (culture B). 1 L of culture B was harvested at OD_{600} 1.3, and the cell pellet was washed. In order to grow sMMO-active *M. trichosporium* OB3b, copper was omitted from the growth medium and the cells were assayed for sMMO activity (see below). Two cultures of *M. trichosporium* OB3b were grown in 3 L fermentors supplemented with 10 μM Fe(II) and either 5 μM Cu(II) (culture C) or no copper (culture D), for the sMMO-active cell profile, and were incubated for 24 h. 2 L of each of the cultures C and D were harvested at a $\sim OD_{600}$ 2 and were tested for sMMO activity. The cell profile of sMMO-active *M. trichosporium*

OB3b cells was compared to the cell profile of the organism grown under 5 μM Cu(II), which was found to be sMMO-inactive.

For the large scale (≥10 L) cell profiles of *M. trichosporium* OB3b, which aimed to maximise the concentrations of copper and protein detected, the organism was grown in NMS media supplemented with 10 μM Fe(II) and high Cu(II) concentration. In the case of the 10 L cell profile of *M. trichosporium* OB3b, pellets were combined from 6 L of culture, grown at 5 μM Cu(II), and 4 L grown at 1 μM Cu(II). The 10 L cell profile resulted in the identification of a novel copper protein from *M. trichosporium* OB3b and subsequently a 16 L cell profile, where all pellets were from cultures grown at 5 μM Cu(II), was performed aiming to optimise the purification of this protein. In all cases cells were collected by centrifugation and pellets were washed.

2.2.2 sMMO activity assay

M. trichosporium OB3b cultures were tested for sMMO activity using the o-dianisidine/naphthalene spectrophotometric assay as described in section 6.4. Pelleted cells (2 ml culture) were resuspended in 10 mM phosphate buffer pH 7 plus 10 mM sodium formate and diluted appropriately to a final volume of 2 ml and OD_{600} 0.3-0.7. The reaction was initiated by adding crushed naphthalene crystals and vigorously shaking. 50 μl of o-dianisidine dye (5 mg/ml stock) were added to the solution and the absorbance at 528 nm, corresponding to the formation of napthol, was monitored for 30-40 mins. The change in absorbance at 528 nm was monitored over time (ΔA_{528nm}/min) and converted to ng of napthol min^{-1} mg^{-1} cells (2, 37).

2.2.3 Purification and quantification of methanobactin

Mb was purified from cell-free medium isolated from *M. trichosporium* OB3b cultures (section 6.5) by reverse phase HPLC eluted using 10 mM ammonium acetate pH 7 (buffer A) and 10 mM ammonium acetate in 80% acetonitrile (buffer B) in a linear gradient from 5% to 40% (buffer B) over 40 minutes (6). Purified mb forms were lyophilized overnight and dried mb was resuspended in 20 mM Hepes pH 7.5. Apo- and Cu(I)-mb were quantified by UV-VIS spectroscopy using ε_{340}= 21.8 mM^{-1} cm^{-1} and ε_{290}= 16.4 mM^{-1} cm^{-1}, respectively (6).

2.2.4 Ag(I) binding by methanobactin

The Ag(I)-binding properties of mb were investigated by titrating Ag(I) into apo-mb (13.2 μM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl and the experiment was

monitored by UV-VIS spectroscopy (section 6.6). The ability of Ag(I) to replace Cu(I) in mb was tested by anaerobically incubating Cu(I)-mb (8 μM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl with a stoichiometric amount of Ag(I) in the presence of 40 μM of the Cu(I) chelator bathocuproine disulfonate (BCS) that has high affinity for Cu(I) ($\log\beta_2=20.8$ (38)) (section 6.6). The formation of $[Cu(BCS)_2]^{3-}$ was monitored at 483 nm.

2.2.5 Cell profiles of *M. trichosporium* OB3b

Cell profiles of *M. trichosporium* OB3b were performed by analysing the soluble extract of the organism according to a modified version of a previous protocol (36) (section 6.7). Soluble extracts from *M. trichosporium* OB3b were resolved by two dimensions of chromatography, first by anion-exchange chromatography (HiTrap Q HP column) eluted with 20 mM Hepes pH 8.8 containing increasing NaCl concentrations (100, 200, 300, 400 and 1000 mM), in which proteins are resolved according to pI value (section 6.7.2)., followed by size-exclusion chromatography on either a SW3000, a G100 or a S75 Superdex column, where proteins are resolved based on their size (section 6.7.3). Notably, for large scale (≥10 L) cell profiles of *M. trichosporium* OB3b a NaCl gradient was used in the anion-exchange chromatography step. The metal content (copper, silver, iron, zinc, manganese) of eluted fractions was quantified by ICP-MS (section 6.7.4). and the proteins of metal-containing fractions were visualised by SDS-PAGE (section 6.11.2). Proteins with intensity on SDS-PAGE gels that correlated with metal concentration, as quantified by ICP-MS, were excised from gels, digested with trypsin and subjected to nano liquid chromatography tandem mass spectrometry (nano LC/MS/MS) (39). The parent proteins from which detected peptides originated were identified with the Mascot MS/MS ion search tool against the NCBI non-redundant protein sequence database of alphaproteobacteria (section 6.7.5). Searches using the Mascot MS/MS ion search tool are based on probability and a Mascot score ($-10 \times \log(P)$) is quoted for the results, where P is the probability of the observed match being a random result. For all protein identifications Mascot scores greater than 0.70 were quoted as indicative of identity or sequence homology. The expectation value (E) for the search performed is also quoted and represents the number of matches that can occur for the same search by chance. For data representation SigmaPlot was used to create three dimensional plots (surface plots) representing the

two chromatographic dimensions on axes x and y and the metal concentration quantified by ICP-MS on the z axis.

In order to assess where mb would elute from the columns used for anion-exchange and size-exclusion chromatography in cell profiles of *M. trichosporium* OB3b, control experiments with apo-, Cu(I)- and Ag(I)-mb were performed. The three forms of mb (20-40 µM) were loaded onto an anion-exchange (HiTrap Q HP, 1 ml) or a size-exclusion (SW3000) column. Elution from the anion-exchange column was performed with 20 mM Hepes pH 8.8 containing 100, 200, 300, 400 and 1000 mM NaCl. Size-exclusion chromatography was performed in 5 mM Hepes pH 7.5 plus 50 mM NaCl. UV-VIS spectra of the eluted fractions from both columns were recorded in order to test whether they contained mb.

2.2.6 Bioinformatics

In order to search for protein homologues of a novel copper protein identified from one of the cell profiles of *M. trichosporium* OB3b, the protein sequence was BLAST searched against the *M. trichosporium* OB3b genome (7), using the NCBI protein blast tool (40). Only hits with an expect value (E) lower or equal to 10^{-3} were considered as reliable. SignalP was used to check for signal peptides (41) and Clustal Omega (42) was used for aligning protein sequences. Protein homologues whose metal-free structures have been solved were identified by blast searching the full sequence of the novel copper protein from *M. trichosporium* OB3b on the pdb database (Section 6.8).

2.3 Results

2.3.1 Identification of bacterioferritin from a cell profile of *M. trichosporium* OB3b

The aim of the cell profiles of *M. trichosporium* OB3b performed initially was to use Ag(I) to probe the soluble Cu(I) pools where mb possibly delivers copper and potentially identify mb inside *M. trichosporium* OB3b. In order to apply this idea, *M. trichosporium* OB3b was first exposed to 1 µM Ag(I) to assess whether Ag(I) is internalised by the organism and whether it co-migrated with soluble copper pools. The elution fractions that resulted from this experiment were analysed for copper, silver and also zinc, iron and manganese. A strikingly high iron concentration (16 µM) eluting in

the size-exclusion purification of the 200 mM NaCl eluate from the anion-exchange column was observed in this metal analysis. A complex mixture of proteins in the iron-containing fractions was resolved by SDS-PAGE (Figure 2.1 A) and a distinct protein band running just below the 21.5 kDa marker on the gel had intensity that clearly peaked at the 7.5 ml fraction, the same as the iron concentration (Figure 2.1 B). This protein was excised from the gel and the digested peptides underwent nano LC/MS/MS. Nine peptides were detected and matched and the parent protein was identified as bacterioferritin (18.569 kDa) from *M. trichosporium* O3Bb, with a Mascot score of 139 and an E value of 4.4×10^{-8}.

2.3.2 Internalisation of Ag(I) by *M. trichosporium* OB3b

The surface plot in Figure 2.2 A shows the soluble silver pools detected in the soluble extract of *M. trichosporium* OB3b exposed to 1 μM Ag(I). This indicates that silver, a non-native metal, is internalised by *M. trichosporium* OB3b and a maximal silver concentration of 0.36 μM was detected. A comparison of the soluble silver and copper pools detected in this cell profile, shown in Figure 2.2 A and B, respectively, indicates the two metals co-elute in the 100 mM, 200 mM and 400 mM NaCl anion-exchange fractions. However the elution of the soluble silver and copper pools is not identical. The extent to which the two metals co-elute can be understood better by looking at the metal analysis of the elution fractions from the size-exclusion purification of the 200 mM and 400 mM NaCl anion-exchange fractions in Figures 2.2 C and D, respectively. In Figure 2.2 C (200 mM), two copper peaks eluting at 11 and 13 ml were resolved, containing 0.43 μM and 0.65 μM copper, respectively. Silver (0.21 μM) co-migrates only with the first copper peak. Figure 2.2 D (400 mM) shows a sharp copper peak (0.19 μM) eluting at 6.5 ml and a second broader copper peak (0.17 μM), likely to be composed of more than one component, eluting at 9 ml. Poorly resolved copper species also elute at 13 and 14.5 ml. Silver peaks (0.10 μM) co-elute with copper at 6.5 and 8.5 ml. The proteins of the copper- and silver-containing fractions shown in Figure 2.2 C were resolved by SDS-PAGE (Figure 2.3). A protein band running just above the 21.5 kDa marker on the SDS-PAGE gel (Figure 2.3 A) appeared to correlate in intensity to the copper and possibly the silver concentration of the fractions, peaking (Figure 2.3 B), but identification was not possible in this case.

2.3.3 Ag(I) binding by methanobactin

The ability of mb to bind Ag(I) was tested in order to subsequently perform a cell profile of *M. trichosporium* OB3b exposed to Ag(I)-mb. The UV-VIS spectra recorded for the titration of Ag(I) into apo-mb (Figure 2.4 A) indicated that mb binds Ag(I). The absorbance at 394 nm decreases during Ag(I) addition (Figure 2.4 B) until a plateau is reached at ~1 molar equivalent of Ag(I), indicating that mb binds one Ag(I) ion as is also the case for Cu(I) (6). Incubation of Cu(I)-mb (8 μM) with a five-fold excess of BCS (40 μM) did not result in the removal of Cu(I) from mb (Figure 2.4 C). However, when 1 molar equivalent of Ag(I) (8 μM) was added to the sample, the formation of the $[Cu(BCS)_2]^{3-}$ was observed (Figure 2.4 C). This ability of Ag(I) to displace Cu(I) from mb indicates that mb also has a high affinity for Ag(I).

2.3.4 Ag(I) internalisation into *M. trichosporium* OB3b is mediated by methanobactin

Once the Ag(I)-binding ability of mb was confirmed, a cell profile of *M. trichosporium* OB3b exposed to 1 μM Ag(I)-mb was performed, aiming to investigate whether mb delivers copper to any soluble copper pools and possibly identify internalised Ag(I)-mb. Soluble silver pools were detected in the soluble extract of *M. trichosporium* OB3b, although at very low concentrations (0.006 μM) as shown in the surface plot in Figure 2.5 A. The silver concentrations in the size-exclusion purification of the 200 mM NaCl (Figure 2.5 B) and 400 mM NaCl (Figure 2.5 C) anion-exchange fractions were compared for the cell profiles of *M. trichosporium* OB3b exposed to silver either as Ag(I)-mb or as unchelated Ag(I) (from Figure 2.2 C and D). In Figure 2.5 B (200 mM) a silver peak elutes from 9 to 14 ml peaking at 10 ml for the Ag(I)-mb cell profile and at 11 ml for the Ag(I) cell profile. In Figure 2.5 C (400 mM) silver peaks elute from 5.5 to 12.5 ml and peak at 6.5 and 10.5 ml for the Ag(I)-mb cell profile and at 6.5 and 8.5 ml for the Ag(I) cell profile. The elution of silver peaks in the same fractions for the two cell profiles of *M. trichosporium* OB3b (exposed to Ag(I)-mb and to Ag(I)) indicates the silver concentrations detected in the Ag(I)-mb cell profile, although extremely low, are not background noise.

2.3.5 Control analysis of mb by anion-exchange and size-exclusion chromatography

In order to assess which fractions mb would be expected to elute in from the anion-exchange and size-exclusion columns used for the cell profiles described, control

experiments were performed with different mb forms (apo-mb, Cu(I)-mb and Ag(I)-mb). Apo-mb elutes in the 400 mM NaCl fraction from the anion-exchange column, whereas the metal-bound forms elute at 200 mM NaCl. When subjected to size-exclusion chromatography on a SW3000 column all mb forms eluted from 14 to 14.5 ml. In light of these data the soluble silver and copper pools in the size-exclusion purification of the 200 mM NaCl eluate from the anion-exchange column were re-examined for both cell profiles of *M. trichosporium* OB3b exposed to Ag(I) (Figure 2.2 B) and Ag(I)-mb (Figure 2.5 B). Nevertheless, silver or copper peaks were not observed at ~14 ml, not enabling the identification of internalised Ag(I)-mb or Cu(I)-mb. Although mb was not visualised in the cell profiles of *M. trichosporium* OB3b exposed to Ag(I)-mb, the detection of soluble silver pools (Figure 2.5) implies mb is involved in Ag(I) uptake by *M. trichosporium* OB3b.

2.3.6 Visualisation of soluble copper pools in *M. trichosporium* OB3b

While the cell profiles described so far did not lead to the identification of mb in the soluble extract of *M. trichosporium* OB3b they provided a method for the identification of metalloproteins, as confirmed by the case of bacterioferritin (see section 2.3.1). Moreover, an unexpected number of soluble copper pools were visualised in the cell profile of *M. trichosporium* OB3b exposed to Ag(I) (Figure 2.2 B), which contained appreciable copper (0.65 μM). Therefore, in order to investigate the soluble copper pools of *M. trichosporium* OB3b a soluble extract of *M. trichosporium* OB3b exposed to 1 μM Cu(II) was analysed next. The cell profile revealed an abundance of soluble copper pools in *M. trichosporium* OB3b that elute in different salt concentrations, as shown in the surface plot in Figure 2.6 A. The highest copper concentration (1.5 μM) eluted at 10.5 mL from the size-exclusion purification of the 200 mM NaCl eluate from the anion-exchange column (Figure 2.6 B). Notably, the most abundant soluble copper pool from the cell profile of *M. trichosporium* OB3b exposed to Ag(I) (Figure 2.2 C in black) eluted also at 200 mM NaCl from the anion-exchange column. Figure 2.6 C, corresponding to size-exclusion purification of the 1 M NaCl anion-exchange fraction, contains at least three copper peaks, at copper concentrations lower than 0.6 μM, eluting at 6.5, 10 and 14 ml. The proteins of the copper-containing fractions of the most abundant soluble copper pool (shown in Figure 2.6 B) were resolved by SDS-PAGE (Figure 2.7 A) revealing a high-complexity protein mixture. The intensity of several proteins bands, running below the 14.4 kDa marker on the gel, appears to follow the

copper concentration in the fractions eluting from 9.5 to 13 ml (Figure 2.7 B), which peaks at 10.5 ml. The protein concentration on the gel, however, was not sufficient to allow identification.

2.3.7 Identification of metalloproteins from the main soluble copper pool of *M. trichosporium* OB3b

Aiming to improve the resolution of the most abundant soluble copper pool of *M. trichosporium* OB3b eluting at 200 mM NaCl from the anion-exchange column, as shown by both cell profiles in sections 2.3.2 (Figure 2.2 B) and 2.3.6 (Figure 2.6 B and 2.7), a NaCl gradient was introduced in the anion-exchange chromatography, focusing on the 0-500 mM NaCl range. An additional objective was to maximise copper and protein levels obtained and, for this reason, the soluble extract from a larger pellet of *M. trichosporium* OB3b, corresponding to 10 L of culture grown under high copper concentration was analysed. The copper analysis of the anion-exchange fractions shows a well resolved peak eluting at 28 ml, which corresponds to ~175 mM NaCl (equivalent to the 200 mM anion-exchange eluate from the cell profile described in section 2.5), that contained the highest copper concentration (30 μM) (Figure 2.8). A second copper peak eluting at 48 ml was also present but contained significantly less copper and eluted over more fractions, while other copper peaks of lower resolution eluted between 500 mM and 1 M NaCl. The protein content of the copper-containing anion-exchange fractions eluting from 20 to 33 ml was analysed by SDS-PAGE and revealed a protein mixture of high complexity (Figure 2.8). A protein running just below the 14.4 kDa marker on the SDS-PAGE gel, as was also observed in the cell profile described in section 2.3.6 (Figure 2.7), had a relative intensity that matched copper concentration and peaked in the 28 ml fraction. It is interesting to note that, while the anion-exchange fractions eluting from 20 to 33 ml were also analysed for manganese and zinc, these metals do not peak in the same fractions as the intensity of the protein band of interest, which is the case only for copper (Figure 2.9). Due to the high protein content of the analysed fractions, further purification was required to verify this observation.

The 28 ml fraction was therefore applied to a G100 size-exclusion column. ICP-MS analysis revealed a single copper peak l (Figure 2.10) that contained fewer proteins. The intensity of at least two proteins (running below the 14.4 kDa marker, and between the 21.5 and the 31 kDa markers on the gel) correlates with the copper concentration profile, and they were cut from the gel and identified through nano LC/MS/MS. Seven

peptides were detected and matched for the protein running between the 21.5 and the 31 kDa marker. The closest match corresponds to an iron-manganese superoxide dismutase (22.857 kDa), however the reliability of this hit is low with a Mascot score of 68 and an E value of 0.71. The protein running below the 14.4 kDa marker was identified as a 'conserved hypothetical protein of unknown function' (14.966 kDa) with a Mascot score of 82 and an E value of 0.0024. Notably, in this case, a single tryptic fragment was detected from the 13 peptides expected by digestion with tryspin. The analysis of the full protein sequence by SignalP (41) suggested the first 24 residues (from Met1 to Ala24) are a predicted twin arginine translocation (Tat) leader peptide, which secretes folded proteins from the cytosol (Figure 2.11) (43). The theoretically predicted mass of the mature protein (without the signal peptide) is 12.591 kDa, consistent with its migration on SDS-PAGE gels.

2.3.8 Semi-purification of an identified hypothetical protein of unknown function from *M. trichosporium* OB3b

The identification of the 'conserved hypothetical protein of unknown function' from a copper-containing size-exclusion fraction of the resolved soluble extract of *M. trichosporium* OB3b, described in section 2.3.7, led to an attempt to further purify this protein in larger amounts. Taking into account that the identification of the novel protein resulted from a cell profile of *M. trichosporium* OB3b grown under elevated copper concentrations, the organism this time was grown under 5 μM Cu(II) and tested to confirm sMMO-inactivity. The soluble extract analysed came from a cell pellet corresponding to 16 L of *M. trichosporium* OB3b culture and was resolved by a NaCl gradient that focused on the 0-250 mM NaCl range, in order to further resolve the identified hypothetical protein of interest from the copper pool eluting at ~175 ml NaCl in section 2.3.7 (Figure 2.8). Copper analysis of the elution fractions from the anion-exchange column shows a well resolved peak eluting at 36 ml, corresponding to ~113 mM NaCl, that contains 85 μM copper (Figure 2.12). A poorly resolved peak containing all other soluble copper proteins elutes between 300 mM and 1 M NaCl. Selected fractions across the copper peak at 36 ml were run on SDS-PAGE gels in order to evaluate the amount of the protein of interest compared to the level of contaminants they contain (Figure 2.12). The same copper-containing fractions were also analysed for zinc and manganese and, notably, the fractions where the intensity of the identified hypothetical protein peaks (around 36 ml) do not contain significant amounts of these

metals (Figure 2.13). The fraction at 32 ml was chosen for further purification by size-exclusion chromatography on a Superdex S75 column, a column of higher resolving capacity (3 to 70 kDa) than the G100 (4 to 150 kDa) used previously (Figure 2.14 A).

The copper analysis of the eluted fractions from the Superdex S75 column showed two copper peaks eluting at 12 and 17 ml while the concentration of zinc in the same fractions is effectively zero (Figure 2.14 B). The absorbance at 280 nm in Figure 2.14 A showed protein eluting at 11.3 ml, which corresponds to an apparent molecular weight of 39.3 kDa. The resolution of the elution fractions (from 8 to 20 ml) by SDS-PAGE showed the only protein with intensity matching the copper concentration in these fractions, is the one running just below the 14.4 kDa marker on the gel (Figure 2.14 C). The identity of this protein was verified by nano LC/MS/MS confirming it is the same 'conserved hypothetical protein of unknown function' identified in section 2.3.7. On average, the hypothetical protein identified elutes from the Superdex S75 column at 11.1 ml (average of 20 measurements ranging from 10.8 ml to 11.4 ml), which corresponds to an average apparent molecular weight of 43.7 kDa. However, under denaturing conditions the protein runs at below the 14.4 kDa marker on the gel, at ~12 kDa (Figure 2.14 C), indicative of the protein forming a tetramer in solution. In this experiment the identity of the previously identified hypothetical protein (section 2.3.7) was confirmed. The protein was semi-purified from sMMO-inactive *M. trichosporium* OB3b grown under high copper concentration and shown to be isolated from the organism loaded with appreciable amounts of copper. By estimating the protein concentration (~0.23 μM) from the intensity of the protein band on the gel (by comparison to the intensity of the markers used, the concentration of which is known) and based on the copper concentration by ICP-MS (2.3 μM), an estimation of the copper to protein stoichiometry at 10:1 was made. As a consequence of the results presented here a role of the identified hypothetical protein in copper storage (copper storage protein 1, Csp1 herein) is suggested.

2.3.9   Copper-dependent expression of Csp1

In order to test the hypothesis that the expression of Csp1 is copper-dependent, the soluble extracts from sMMO-inactive, grown under 5 μM Cu(II), and sMMO-active (714 ng napthol min^{-1} mg^{-1} cells) *M. trichosporium* OB3b were analysed (data by Owen Burbidge, undergraduate project student 2012). The copper analysis of the size-exclusion fractions from the resolution of the 200 mM NaCl anion-exchange fraction on

a Superdex S75 column showed the copper peak eluting at 11 ml, which corresponds to Csp1, was present in the case of sMMO-inactive cells and contained a remarkably high amount of copper (17 μM) whereas the copper concentration in the same fractions of the sMMO-active cells was negligible (Figure 2.15 A). When the proteins of the size-exclusion fractions from these cell profiles were resolved by SDS-PAGE, the protein band running below the 14.4 kDa marker on the gel, with intensity that correlates to the copper concentration of these fractions, was present in the sMMO-inactive cell profile (Figure 2.15 B). Semi-purified Csp1, obtained from the experiment described in section 2.3.8, was used as an additional marker on the gel and confirms the protein band present in the sMMO-inactive cell profile is Csp1. In contrast, the protein band was absent on the SDS-PAGE gel for the respective fractions, in the case of the sMMO-active cell profile (Figure 2.15 C), consistent with the effectively zero copper concentration detected (Figure 2.15 A in black). This finding suggests that Csp1 expression is regulated by copper and is induced under high copper conditions.

2.3.10 Bioinformatics

An initial bioinformatics search for Csp1 homologues revealed the existence of two more homologous proteins in *M. trichosporium* OB3b; Csp2, which also contains a predicted Tat-leader peptide, and Csp3 that does not have a Tat-leader. Sequence alignments of the homologues show that the mature forms of Csp1 and Csp2 share 57.38% sequence identity, while Csp3 is significantly different as it has only 18.3% sequence identity to Csp1 (Figure 2.16) (42). Csp1 and Csp2 contain 13 Cys residues, while Csp3 contains 18 Cys residues, and these are arranged in Cys-X-X-Cys and Cys-X-X-X-Cys motifs in all three proteins. Protein homologues are also present in *Pseudomonas aeruginosa (*PA2107 protein) and *Nitrosospira multiformis* (Nmul_A1745 protein) and the structures of the metal-free proteins have been determined by the Northeast Structural Genomics consortium (Forouhar, F., Lew, S., Seetharaman, J., Sahdev, S., Xiao, R., Ciccosanti, C., Lee, D., Everett, J.K., Nair, R., Acton, T.B., Rost, B., Montelione, G.T., Tong, L., Hunt, J.F.) (pdb codes 3KAW and 3LMF for *Pseudomonas aeruginosa* and *Nitrosospira multiformis*, respectively), while functional data for these proteins are not available. The structure of the homologous protein from *Nitrosospira multiformis*, which shares 25% sequence identity with Csp1, is a 4-helix bundle in which all Cys residues face towards the core of the bundle and do not form disulfide bonds (Figure 2.17).
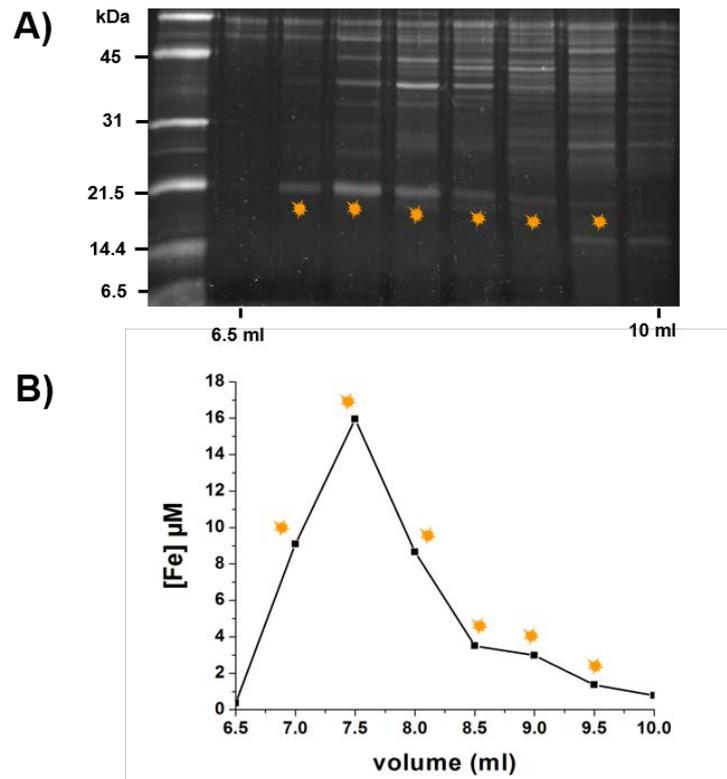
Figure 2.1 (A) SDS-PAGE showing the resolved proteins of the iron-containing fractions (B) eluting from 6.5 to 10.0 ml from the size-exclusion purification (SW3000 column) of the 200 mM NaCl eluate from the anion-exchange column. The intensity of the protein band running just below the 21.5 kDa marker correlates to the iron concentration peaking at 7.5 ml and was identified as bacterioferritin.
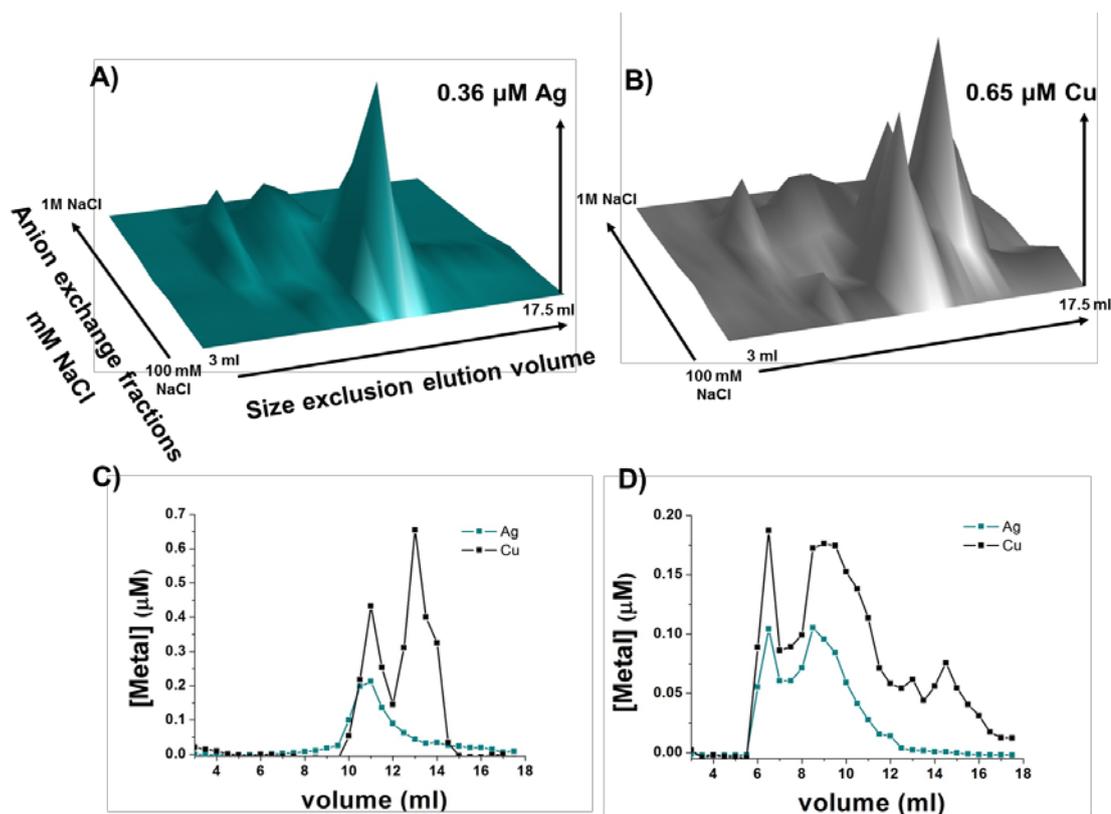
Figure 2.2 The soluble extract of *M. trichosporium* OB3b (1 L culture grown in the presence of 1 μM Ag(I) for 24 h) eluted from an anion-exchange column with 100, 200, 300, 400 and 1000 mM NaCl. These were subsequently resolved on a SW3000 size-exclusion column. The silver (A) and copper (B) concentrations detected in the elution fractions from anion-exchange and size-exclusion chromatography are represented in surface plots. Chromatograms in C) and D) represent sections of the surface plots and show the copper (black) and silver (cyan) analysis of elution fractions from the SW3000 size-exclusion purification of the 200 mM (C) and 400 mM NaCl (D) anion-exchange fractions.
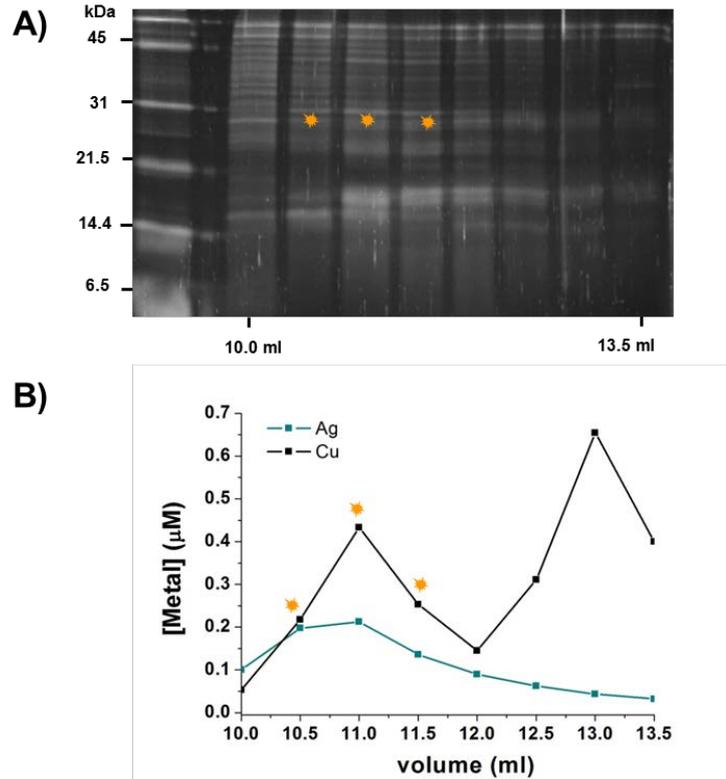
Figure 2.3 (A) SDS-PAGE showing the resolved proteins of the copper- (black) and silver- (cyan) containing fractions (B) eluting from 10.0 to 13.5 ml from the size-exclusion purification (SW3000 column) of the 200 mM NaCl eluate from the anion-exchange column (metal analysis taken from Figure 2.2 C). The intensity of the protein band running just above the 21.5 kDa marker appears to correlate with the copper and possibly the silver concentration peaking at 11 ml, however identification was not possible.
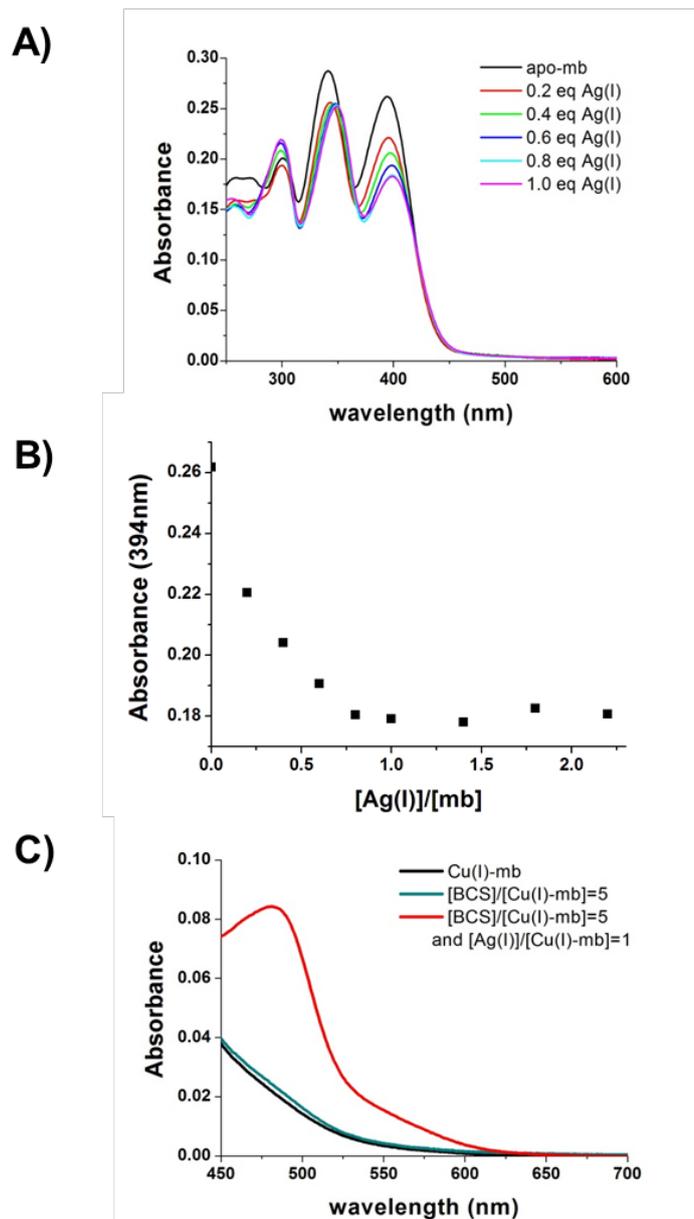
Figure 2.4 A) UV-VIS spectrum of apo-mb (13.2 µM) in 20 mM Hepes pH 7 (black line) and the influence of the addition of 0.2 (red), 0.4 (green), 0.6 (blue), 0.8 (magenta) and 1 (light blue) molar equivalents of Ag(I). B) Plot of the absorbance at 394 nm, taken from the spectra shown in A), against [Ag(I)]/[mb] ratio. C) UV-VIS spectrum of Cu(I)-mb (8 µM) (black line) and of the sample after the addition of 40 µM BCS (corresponding to 5 molar equivalents to Cu(I)-mb, cyan line) and the subsequent addition of 8 µM Ag(I) (corresponding to 1 molar equivalent to Cu(I)-mb, red line).
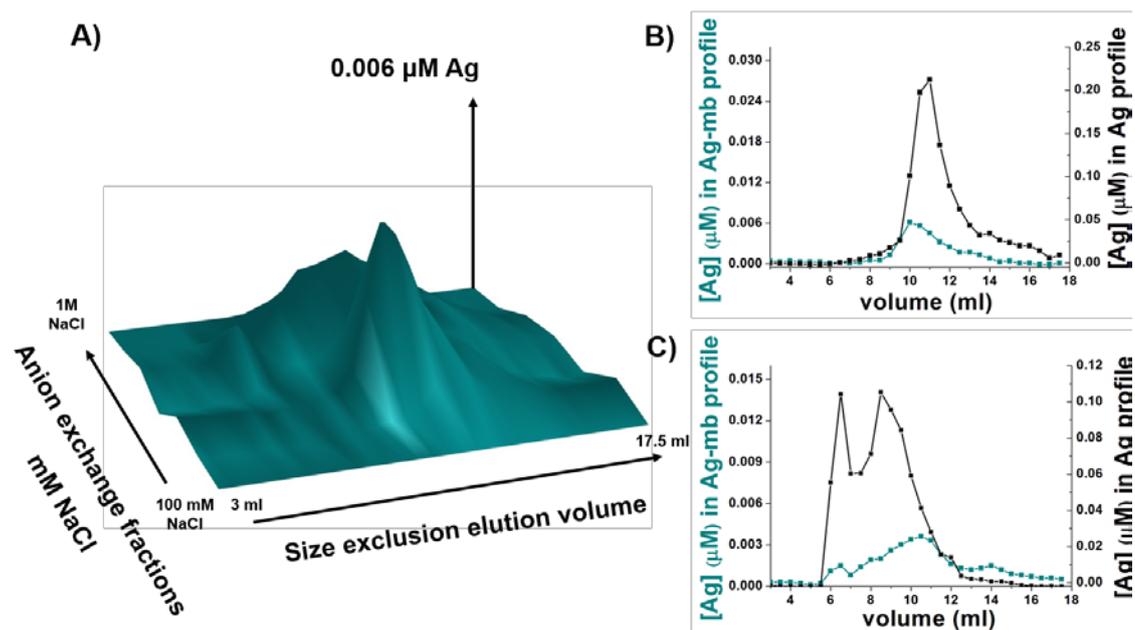
Figure 2.5 The soluble extract from *M. trichosporium* OB3b (1 L culture grown in the presence of 1 µM Ag(I)-mb for 24 h) eluted from an anion-exchange column with 100, 200, 300, 400 and 1000 mM NaCl. These were subsequently loaded onto a SW3000 size-exclusion column. The silver concentration detected in the elution fractions from anion-exchange and size-exclusion chromatography is represented in the surface plot shown in (A). Chromatograms in B) and C) represent sections of the surface plot and show the size-exclusion purification of the 200 mM (B) and 400 mM (C) anion-exchange fractions on a SW3000 column. The silver concentration is compared between soluble extracts of *M. trichosporium* OB3b from cells exposed to 1 µM silver either as Ag(I)-mb (cyan) or unchelated Ag(I) (black) (from Figure 2.2). Silver peaks from the two cell profiles elute in the same fractions (B and C).
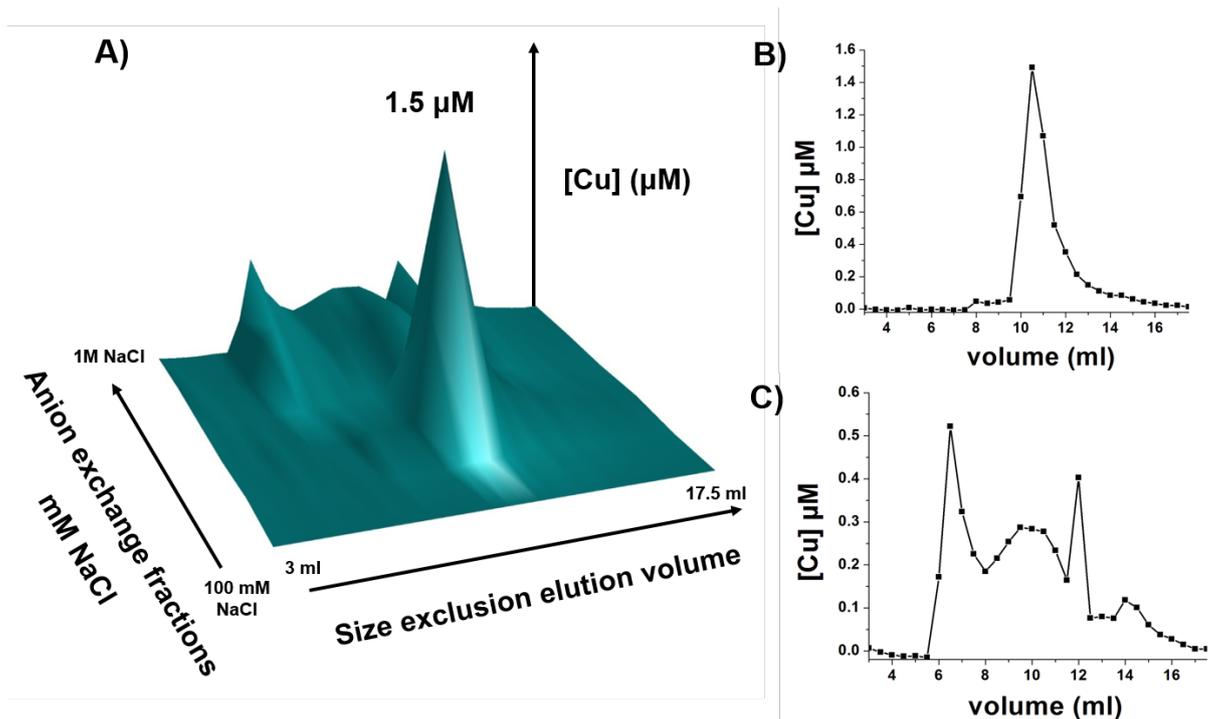
Figure 2.6 The soluble extract from *M. trichosporium* OB3b (1 L culture grown in the presence of 1 μM Cu(II) for 24 h) was eluted from an anion-exchange column with 100, 200, 300, 400 and 1000 mM NaCl. These were subsequently loaded onto a SW3000 size-exclusion column. The copper concentration detected in the elution fractions from anion-exchange and size-exclusion chromatography is represented in the surface plot shown in (A). Chromatograms in B) and C) represent sections of the surface plot and show the copper analysis of elution fractions from the SW3000 size-exclusion purification of the 200 mM (B) and 1 M NaCl (C) anion-exchange fractions.
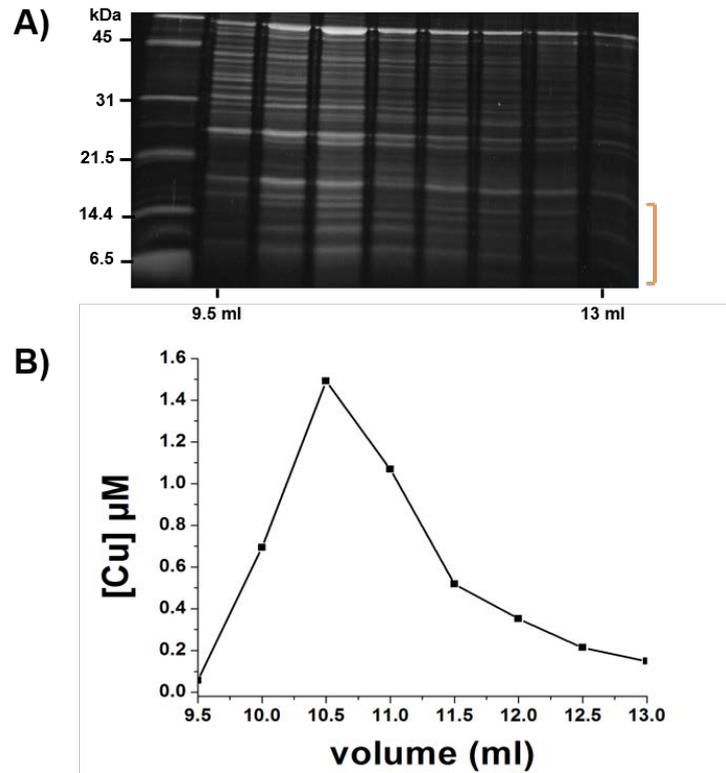
Figure 2.7 (A) SDS-PAGE showing the resolved proteins of the copper-containing fractions (B) eluting from 9.5 to 13.5 ml from the size-exclusion purification (SW3000 column) of the 200 mM NaCl eluate from the anion-exchange column (copper analysis taken from Figure 2.6 B). A number of protein bands with intensity matching the copper concentration peaking at 10.5 ml are present on the gel, running below the 14.4 kDa marker, but identification was not possible.
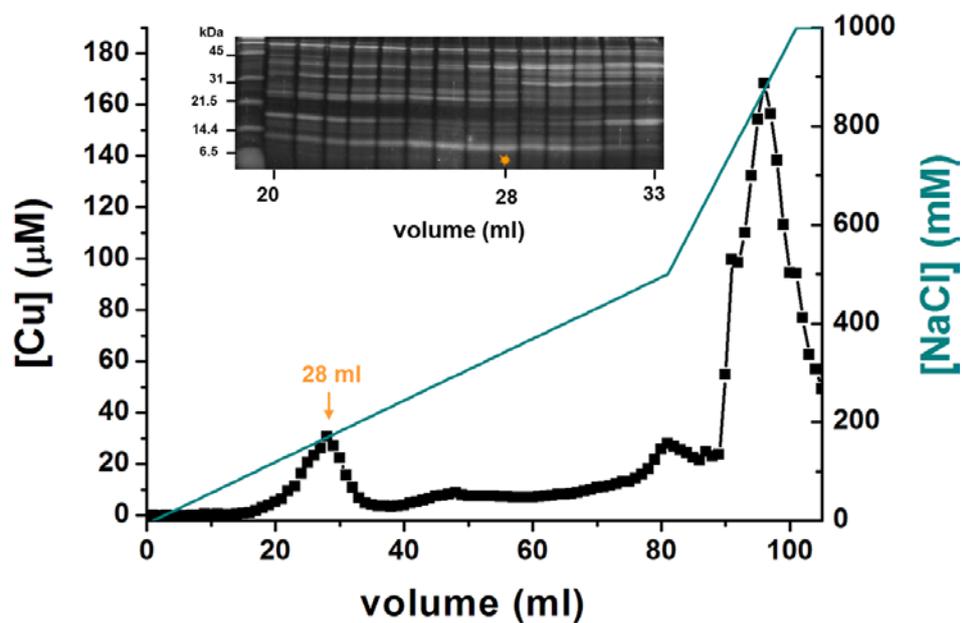
Figure 2.8 The soluble extract from *M. trichosporium* OB3b (10 L culture grown under elevated Cu(II)) was eluted from an anion-exchange column with a linear NaCl gradient (cyan line) and the copper content of the eluted fractions was analysed. The proteins of the copper-containing fractions eluting from 20 to 33 ml were visualised by SDS-PAGE. The intensity of the band just below the 14.4 kDa marker correlates with the copper concentration profile peaking at 28 ml.
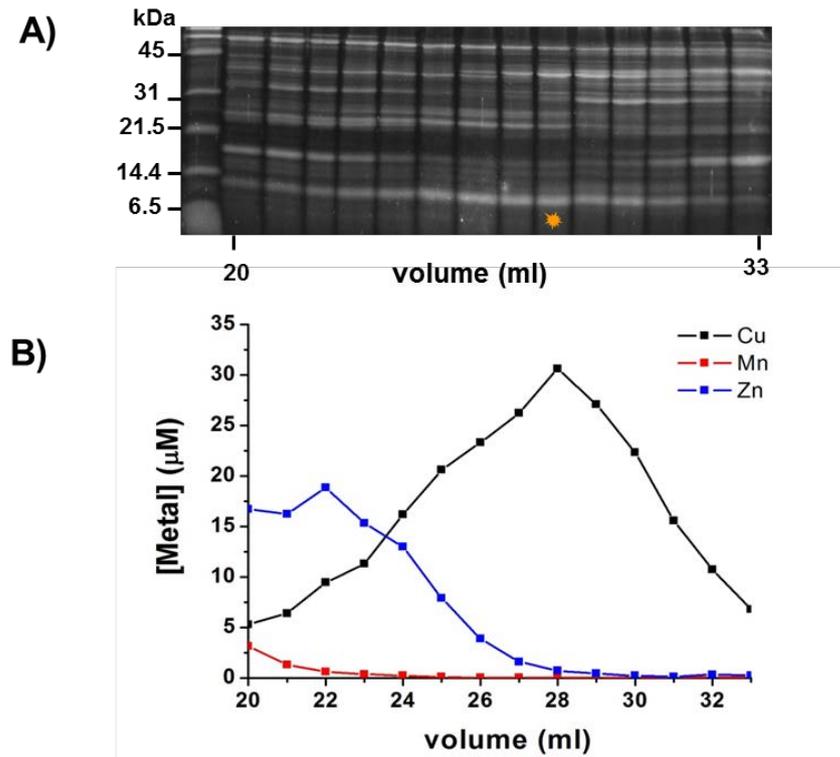
Figure 2.9 The protein content of the anion-exchange fractions eluting from 20 to 33 ml, containing the most abundant soluble copper pool eluting at 28 ml (corresponding to 175 mM NaCl, shown in Figure 2.8), was resolved by SDS-PAGE (A). Except for copper (black), fractions were also analysed for zinc (blue) and manganese (red) (B). The highlighted protein band running below the 14.4 kDa marker on the gel (orange) has intensity that follows the copper concentration profile and peaks at 28 ml, while zinc and manganese peak in earlier fractions.
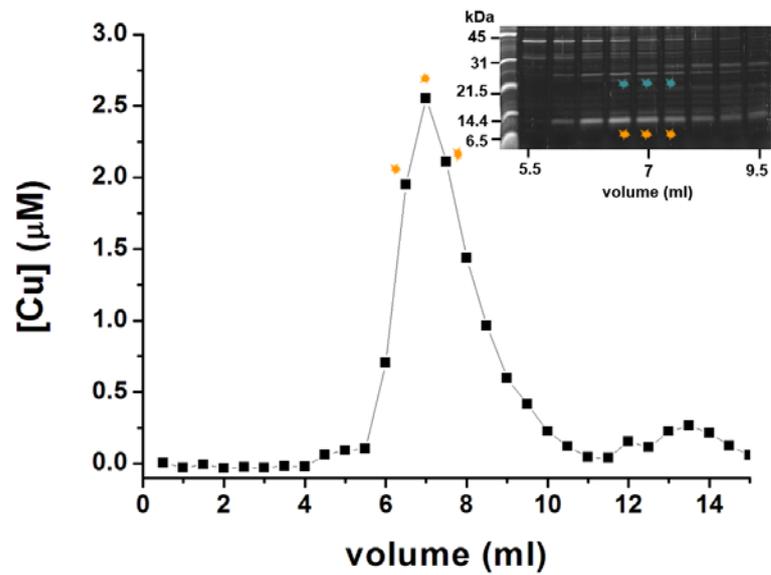
Figure 2.10 The copper concentration of the size-exclusion purification of the anion-exchange fraction 28 (shown in Figure 2.8) on a G100 column. Also shown are the resolved proteins (SDS-PAGE) of the copper-containing fractions eluting from 5.5 to 9.5 ml. The intensity of the protein band running below the 14.4 kDa marker on the gel (orange) correlates to the copper concentration profile peaking at 7 ml and was identified as a 'conserved hypothetical protein of unknown function'. The protein band running between the 21.5 and the 31 kDa markers on the gel (cyan) was also identified as iron-manganese superoxide dismutase.

```
          10        20        30        40        50        60

MERRDFVTAFGALAAAAAASSAFAGEDPHAGHKMSHGAKYKALLDSSSHCVAVGEDCLRHCFEMLAMND

70        80        90        100       110       120       130

ASMGACTKATYDLVAACGALAKLAGTNSAFTPAFAKVVADVCAACKKECDKFPSIAECKACGEACQACAE

140

ECHKVAA
```

Figure 2.11 The protein sequence of the 'conserved hypothetical protein of unknown function' identified from *M. trichosporium* OB3b with the predicted Tat-leader peptide shown in italics. The 13 Cys residues are highlighted in yellow and the Cys-X-X-Cys and Cys-X-X-X-Cys motifs are underlined. The single tryptic fragment identified by MS/MS analysis is shown in bold and represents 11% sequence coverage of the mature protein (without the Tat-leader peptide). This is the only tryptic fragments of the mature protein anticipated to be detected by MS, as the other predicted trypric fragments of the protein were of small mass or contained Cys residues.
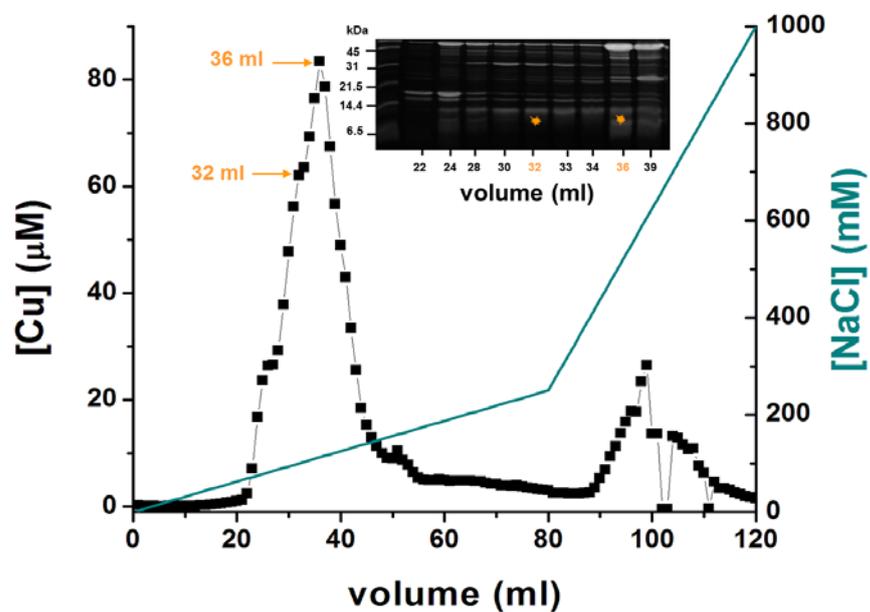
Figure 2.12 The copper content of anion-exchange fractions (linear NaCl gradient shown as a cyan line) from the soluble extract of sMMO-inactive *M. trichosporium* OB3b (16 L culture grown under 5 μM Cu(II)). Selected copper-containing fractions from 22 to 39 ml were analysed by SDS-PAGE and contain the protein band running below the 14.4 kDa marker on the gel. Fraction 32 contains fewer protein contaminants compared to fraction 36 and was further purified by size-exclusion chromatography (see Figure 2.14).
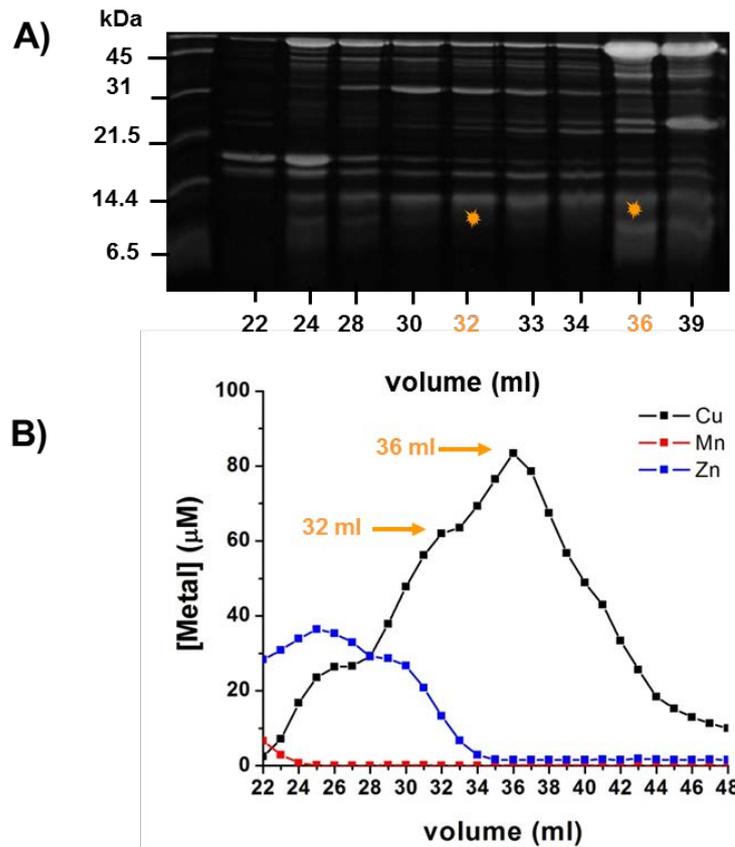
Figure 2.13 The protein content of selected anion-exchange fractions eluting from 22 to 48 ml, containing the most abundant soluble copper pool eluting at 36 ml (corresponding to 113 mM NaCl, shown in Figure 2.12), was resolved by SDS-PAGE (A). The fractions were analysed for copper (black), zinc (blue) and manganese (red) (B). The copper-containing fractions across the peak at 32, 33, 34, 36 and 39 ml contain the protein band running below the 14.4 kDa marker on the gel and do not contain significant concentrations of zinc or manganese.
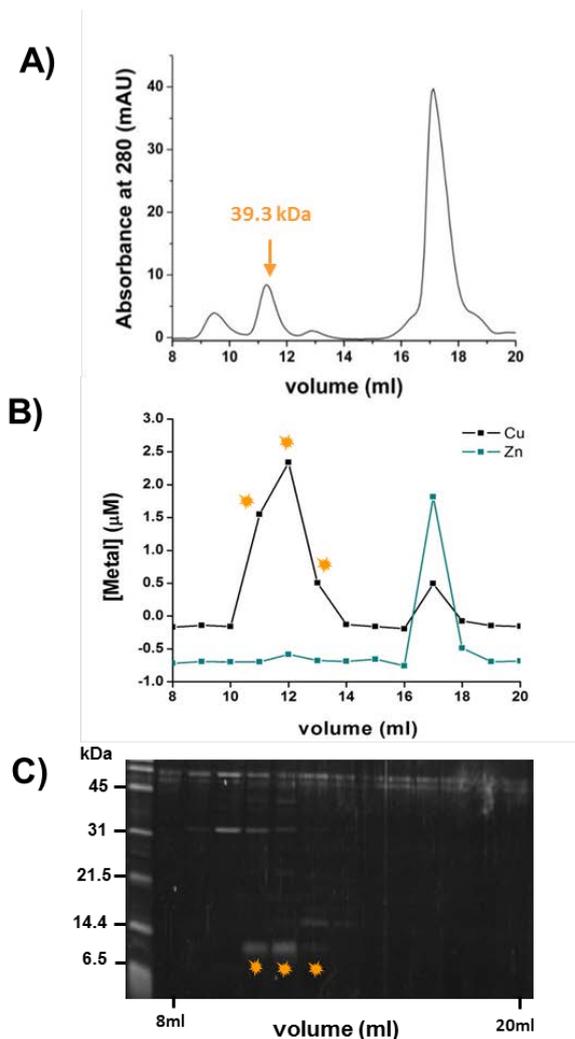
Figure 2.14 (A) Size-exclusion purification of the anion-exchange fraction 32 (shown in Figure 2.12) on a Superdex S75 column. The elution fractions from 8 to 20 ml were analysed for copper (black) and zinc (cyan), shown in (B), and for protein (C). Only two copper peaks are present in these fractions, elute at 12 and 17 ml, while the zinc concentration is negligible (B). The only protein band co-eluting with the copper concentration profile peaking at 12 ml is the one running just below the 14.4 kDa marker on the gel (C). The identity of this protein was confirmed as the same 'conserved hypothetical protein of unknown function' identified in Figure 3.10. The identified protein corresponds to the peak eluting at 11.3 ml (39.3 kDa apparent molecular weight) in (A).
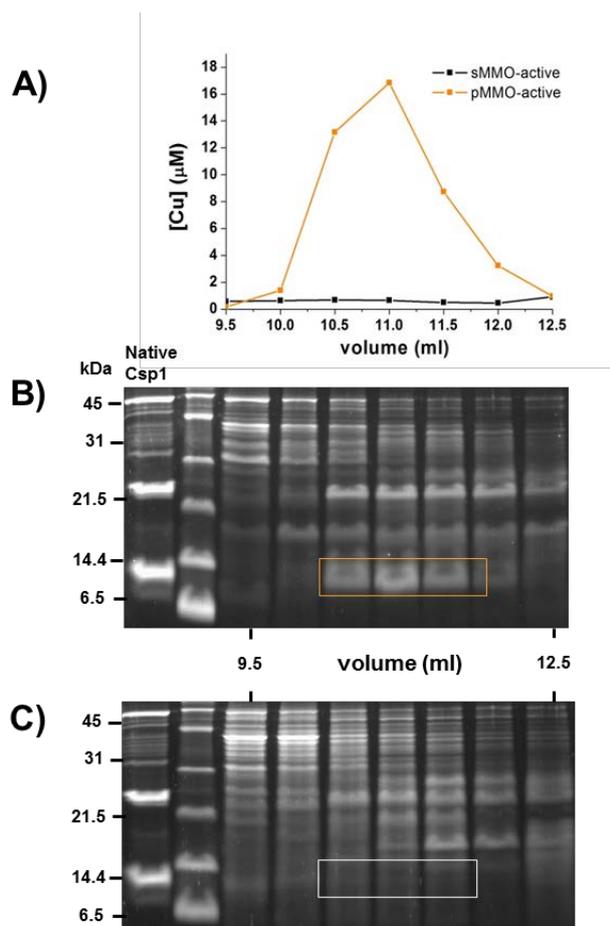
Figure 2.15 Soluble extracts from *M. trichosporium* OB3b (2 L cultures grown in the presence of 5 μM Cu(II) (sMMO-inactive) and in the absence of copper (sMMO-active) for 24 h) were eluted from an anion-exchange column with 100, 200, 300, 400 and 1000 mM NaCl. These were subsequently loaded onto a Superdex S75 size-exclusion column. Copper analysis of elution fractions (from 9.5 to 12.5 ml) from the size-exclusion purification of the 200 mM anion-exchange fraction of the sMMO-inactive (orange) and sMMO-active (black) is shown in (A). The proteins of the size-exclusion fractions from 9.5 to 12.5 ml from the sMMO-inactive (B) and the sMMO-active (C) cell profile were resolved by SDS-PAGE. The protein band corresponding to Csp1, running just below the 14.4 kDa marker on the gel, is present on the gel with intensity that correlates with the copper concentration peaking at 11 ml (B). This protein band is absent in the gel of the sMMO-active profile (C), consistent with the absence of copper in these fractions (shown in A in black). Semi-purified Csp1 (native Csp1), isolated from the 16 L cell profile (see section 2.3.8 and Figure 2.12) was used in addition to the standard protein markers both gels. Data from Owen Burbidge (undergraduate student, 2012).

```
Csp1  ---------GEDPHAGHKMSHGAKYK--ALLDSSSHCVAVGEDCLRHCFEMLAMNDASMGAC--TKATYDLVA----ACGALAKLA----

Csp2  -----QTTQGLAPG-APVHHHPAKYH--ALMETSAKCVSTGNECLRHCFGMLSMNDTSMADC--TKASYDLVA----ACAALETLS----

Csp3  ----------MHVEAMISKHPQARG--QTDRSLVQCVEMCFDCAQTCAACADA-CLGEDKVADLRHCIRLNLDCAEICVAAGSIASRAA

3KAW  MTRAINDPGNEDPGSLLETDADALLGGAAAQAPEERCRLAAQACIRACERYLAL-CTES-----SREQRQHAGDCADLCRLAALLL----

3LMF  --------------------MFLYT--ETDQNLQACIDACNHCYRTCLRMAMNHCLEAGGKHVEADHLRLMMNCAEICQTSLNFM----


Csp1 GTNSAFTPAFAKVVADVCAACKKECDKFPSI-AECKACGEACQACAEECHKVAA---

Csp2 AVNSSATPALAKTVYDVCMACKKECDRFPQY-SECKNCGDACKACADECQRVSS---

Csp3 GTEESILRTMLQTCAEMCRMCEEECRRHAGNHEHCRICADVCKECETACRSATGLTH

3KAW ERRSPWAPAACELAARYALACAERCDGDEPLERE---CAGACRRFVEACRPLLPA--

3LMF LSGSRFSPKVCGVCAEICDACAKSCEQLDGM-EEC---VQTCRQCAEHCRKMAA---
```

Figure 2.16 Alignment of the mature protein sequences of the Csp1 protein homologues from *M. trichosporium* OB3b and the protein homologues from *Nitrosospira multiformis* and *Pseudomonas aeruginosa* whose metal-free structures have been determined (pdb codes 3KAW and 3LMF, respectively). Conserved Cys residues are highlighted in yellow while additional Cys residues of each protein are highlighted in green. The Cys-X-X-Cys and Cys-X-X-X-Cys motifs are underlined. Csp1 and Csp2 share 57.38% sequence identity and both contain a predicted Tat-leader peptide (not included here) and 13 Cys residues. Csp3 has only 18.3% sequence identity to Csp1, does not contain a Tat-leader, and has 18 Cys residues. The homologous proteins from *Nitrosospira multiformis* and *Pseudomonas aeruginosa* share 25% and 23% sequence identity with Csp1, respectively, and do not contain Tat-leader peptides.
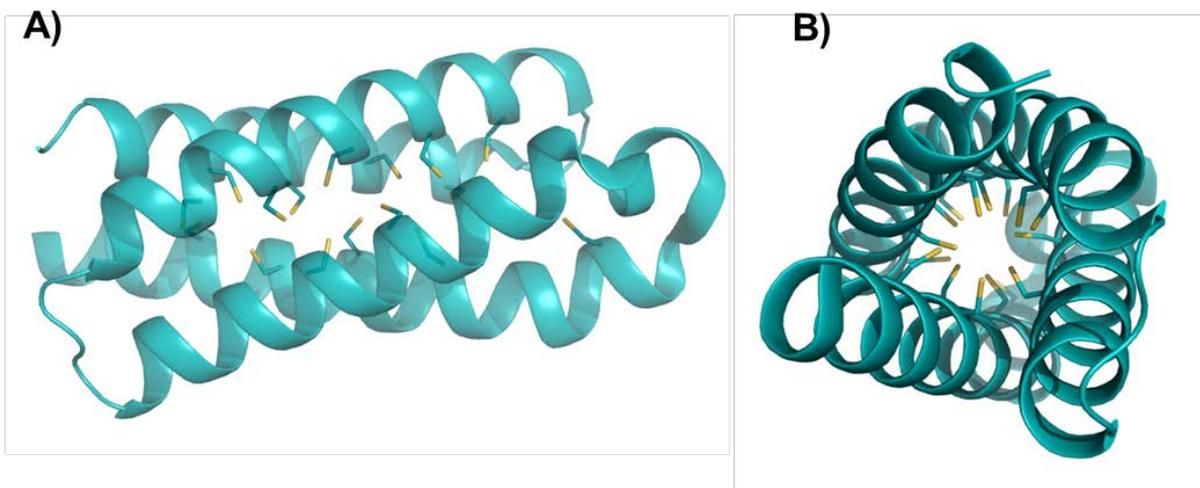
Figure 2.17 Structure of the metal-free form of the protein homologue from *Nitrosospira multiformis* (pdb code: 3LMF), determined by the Northeast Structural genomics consortium, showing the protein is a 4- helix bundle (A) which contains 16 Cys residues, represented as sticks with the side chain shown in yellow. All the Cys residues face towards the core of the 4-helix bundle (B) and do not form disulfide bonds.

2.4 Discussion

2.4.1 Metalloproteomic studies of *M. trichosporium* OB3b

The only characterised component for copper uptake in *M. trichosporium* OB3b is mb, a peptide that sequesters copper from the environment as part of the unique prokaryotic copper uptake system of MOB (1-4). Homologues of well characterised copper homeostasis proteins, such as CopA, CopZ and CueR, are present in *M. trichosporium* OB3b (7). MOB have a high copper requirement as they mainly use pMMO, a copper-dependent enzyme, for methane oxidation. However, the mechanism through which pMMO acquires copper is not understood especially since they have a copper efflux system, also present in other bacteria. In *M. trichosporium* OB3b, capable of expressing both the copper-containing pMMO and the iron-containing sMMO, the switchover between sMMO and pMMO expression is regulated by copper (1-3) and, most importantly, mb has been shown to mediate this switchover (2, 3). In this context, the metalloproteomic studies of *M. trichosporium* OB3b aimed to investigate whether mb is internalised by *M. trichosporium* OB3b, and to identify soluble copper proteins that are potentially involved in copper uptake.

While investigating the internalisation of mb in cell profiles of *M. trichosporium* OB3b, a number of proteins were visualised in metal-containing fractions (Figures 2.1, 2.3 and 2.7) and more than one protein candidates with intensity on the gels that correlated to the metal concentrations in the same fractions were observed. Although proteins with intensity that correlated to the copper and silver concentration of certain fractions were observed in the cell profiles with 1 μM Ag(I) and 1 μM Cu(II) (Figures 2.3 and 2.7), the low resolution of the protein mixture visualised did not allow the proteins of interest to be identified. However, in one case (cell profile with 1 μM Ag(I)) the levels of iron detected were significantly higher than any other metal and the correlation of the iron concentration to the intensity of the protein bands visualised in the analysed fractions enabled the identification of bacterioferritin (Figure 2.1). This was the first protein identified by this method in *M. trichosporium* OB3b and is proof of principle for the metalloproteomic approach. Bacterioferritin is responsible for iron storage as Fe(III) in a FeO(OH) mineral core (44) The identification of bacterioferritin from a profile in which cells grew under low copper levels is interesting. At low copper to biomass ratio *M. trichosporium* OB3b uses the iron-containing sMMO for methane oxidation (45) and bacterioferritin could be responsible for supplying sMMO with iron

that is necessary for catalytic activity. On the other hand, as *M. trichosporium* OB3b was exposed to 1 μM Ag(I) for this profile, it is possible that elevated iron may be a result of the toxic effect of Ag(I) disrupting [4Fe-4S] clusters (22) and bacterioferritin is produced to sequester the excess iron.

The suitability of Ag(I) as a probe for Cu(I) in *M. trichosporium* OB3b was first tested by a cell profile aimed at assessing whether Ag(I) was internalised by the organism. The detection of soluble silver pools (Figure 2.2 A) confirmed this non-native metal is internalised by *M. trichosporium* OB3b possibly by passive diffusion through porins, as is the case for unchelated copper (1). The co-migration of copper and silver, in most cases, confirmed Ag(I) can be used as a probe for the identification of soluble copper proteins (Figure 2.2 C and D). Notably, however, silver co-migrated with only certain copper pools (Figure 2.2 C) which may be a result of copper proteins acquiring both metals, as is the case with components of the cue system (23), which disrupts their function, while other proteins still function to overcome the Ag(I)-induced toxicity and therefore are visualised with only Ag(I) bound.

The localisation of silver in the same soluble pools as copper, in most cases (Figure 2.2), combined with the Ag(I)-binding properties of mb (Figure 2.4), were used in an attempt to visualise the peptide inside the cells but, most importantly, investigate the potential soluble copper proteins to which mb may deliver copper. Indeed, soluble silver pools were detected in the soluble fraction of *M. trichosporium* OB3b exposed to 1 μM Ag(I)-mb but they contained very low silver concentrations (Figure 2.5 A). Although remarkably low, the silver peaks elute in the same fractions as in the silver profile (Figure 2.5 B and C), which indicates the silver levels detected in this profile are not background noise of the ICP-MS. Given the affinity of mb for Ag(I) (Figure 2.4), the fact that some silver was internalised by the organism suggests that mb is involved in the import of the metal, as it is unlikely that Ag(I) dissociated from the peptide. A possible reason for detecting a maximum of only 6 nM of silver in *M. trichosporium* OB3b may be that Ag(I)-mb is not recognised by the active import mechanism used for Cu(I)-mb (1) resulting in poor internalisation of Ag(I)-mb. It may also be the case that mb does not release Ag(I) as efficiently as Cu(I), which is thought to be released either by oxidation to Cu(II) or after the peptide has been degraded (2). Regardless, these results indicate mb is involved in the internalisation of Ag(I) by *M. trichosporium* OB3b and possibly delivers the metal to soluble copper proteins. Although internalised mb

was not identified from the cell profile of *M. trichosporium* OB3b exposed to Cu(II) either, consistently with the Ag(I) cell profile, the most abundant soluble copper pool was present in the 200 mM eluate from the anion-exchange column and contained a large number of proteins (Figure 2.6 and 2.7). This observation led to the cell profiles that followed in order to resolve the proteins of this soluble copper pool.

2.4.2   Identification of a novel copper protein from *M. trichosporium* OB3b

The presence of the most abundant soluble copper pool consistently in the 200 mM NaCl eluate from the anion-exchange column and the large number of proteins this copper pool contained (Figure 2.2 B and C and Figures 2.3, 2.6 and 2.7) resulted in focusing the subsequent experiments on improving the resolution of this soluble copper pool. The linear NaCl gradient for eluting the anion-exchange column achieved better resolution of the copper pool of interest (Figure 2.8) and a protein band running just below the 14.4 kDa marker on the gel, that was also visualised previously in the cell profile with 1 μM Cu(II) (Figure 2.7), was observed with intensity that followed the copper concentration. Notably, the concentration of zinc and manganese was not peaking in the fractions where the intensity of this protein band peaked, indicating it is potentially a copper protein (Figure 2.9). The peaking copper-containing fraction was further purified by size-exclusion chromatography (G100 column) (Figure 2.10), which resulted in the initial identification of a potentially novel copper protein. The full protein sequence contains 13 Cys residues (Figure 2.11), which explains the detection of a single tryptic peptide fragment (from a total of 13) as Cys ionisation is poor and the Cys thiols can form disulfide bridges (46, 47). The complexity of the SDS-PAGE gel shown in Figure 2.10 does not allow the definite correlation of this single metalloprotein to copper, since other metalloproteins are clearly present in the same fractions with intensity on the gel that correlates to the copper concentration of these fractions. The second protein identified from this experiment is such an example, however this protein did not contain copper and was identified as an iron-manganese superoxide dismutase.

By focusing the NaCl gradient used for anion-exchange chromatography to a narrower concentration range (Figure 2.12) and by using a size-exclusion column of higher resolving ability (Superdex S75), a much cleaner fraction of the novel protein was isolated and the identity was confirmed. The identified protein is the only protein on the gel with staining intensity that peaks in the same fractions as copper (Figure 2.14 B and

C) as the concentration of zinc in these fractions is negligible and the soluble manganese pool peaks earlier (Figure 2.13). The level of purity of this fraction allowed an estimate of the copper to protein ratio at 10:1, based on the copper content of this fraction quantified by ICP-MS (Figure 2.14 B) and the protein concentration estimated from the intensity of the band visualised by SDS-PAGE (Figure 2.14 C). The isolation and identification of this novel protein from pMMO-active *M. trichosporium* OB3b grown under high copper levels, together with the fact that the protein was isolated loaded with significant amounts of copper, point towards a role in copper storage was suggested for this protein (Copper storage protein, Csp1).

The comparison of cell profiles from sMMO-active and sMMO-inactive *M. trichosporium* OB3b (Figure 2.15, data shown from Owen Burbidge, undergraduate student 2012) indicated the Csp1 levels present in *M. trichosporium* OB3b depend on the copper concentration in the growth media. The intensity of the protein band in the sMMO-inactive, grown under high copper concentrations, profile indicated the expression of the protein is up-regulated under high copper concentrations, consistent with the absence of an obvious protein band in the sMMO-active profile (Figure 2.15 C). The isolation and identification of Csp1 from pMMO-active *M. trichosporium* OB3b grown under high copper levels, together with the fact that Csp1 was isolated loaded with significant amounts of copper, point towards a role in copper homeostasis. The copper-responsive expression of Csp1 is also in agreement with this hypothesis.

Csp1 contains 13 Cys residues some of which are arranged in Cys-X-X-Cys and Cys-X-X-X-Cys motifs (Figure 2.11), that are also found Cu(I)-binding metallothioneins (48, 49). Bioinformatics revealed two Csp1 homologues in the *M. trichosporium* OB3b genome (7) herein named Csp2 and Csp3 (Figure 2.16). Csp2 contains a predicted Tat-leader peptide and is expected, therefore, to function outside the cytosol (43), while Csp3 is expected to be cytosolic, as it does not contain a Tat-leader peptide. Notably, neither of Csp2 or Csp3 were identified in the metalloproteomic studies of *M. trichosporium* OB3b. Two more protein homologues were identified through bioinformatics from *Nitrosospira multiformis* and *Pseudomonas aeruginosa* with structures available for the metal-free forms. The protein homologue from *Nitrosospira multiformis* (pdb code 3LMF) is a 4-helix bundle with the Cys residues pointing towards the core of the bundle (Figure 2.17). Assuming the structure of Csp1 is similar

to 3LMF, the arrangement of the Cys residues inside the core of a 4-helix bundle provides a shielded environment that appears to have a high copper binding capacity.

The target for copper delivery in *M. trichosporium* OB3b is pMMO, the main enzyme for methane oxidation that is copper-dependent is housed in intra-cytoplasmic membranes that develop at high copper to biomass ratios (1, 5). However, it is not clear whether the pMMO-containing intracytoplasmic membranes are continuous with the cytoplasmic membrane (50) and whether they form separate subcellular compartments. The predicted Tat-leader peptide present in Csp1 and Csp2 indicates these proteins fold in the cytosol before being exported (43). It is possible that Csp1 and Csp2 are exported from the cytosol to where pMMO is localised, in order to serve as copper storage for pMMO. The possibility that Csp1 and Csp2 acquire copper in the cytosol, where they fold before being exported, must also be noted as these two proteins may serve as a mechanism of exporting copper from the cytosol to avoid copper-induced toxicity. Csp3, on the other hand, is expected to be cytosolic and, as indicated by the 18 Cys residues of the protein sequence, is expected to have a higher copper-binding capacity than Csp1 or Csp2. The presence of a copper storage protein in the cytosol is intriguing as bacteria are known to export copper from this compartment (10-17, 23). Csp3 may store copper in the cytosol of *M. trichosporium* OB3b to prevent toxicity or to supply the metal to other unknown protein targets. This possibility challenges the present model for bacterial copper homeostasis, according to which copper is exported from the cytosol.

2.5   References

1. Balasubramanian R., Kenney G. E., Rosenzweig A. C., Dual pathways for copper uptake by methanotrophic bacteria, J. Biol. Chem., **2011**, 286, 37313-37319

2. El Ghazouani A., Baslé A., Gray J., Graham D. W., Firbank S. J., Dennison C., Variations in methanobactin structure influence copper utilization by methane-oxidising bacteria, Proc. Natl. Acad. Sci., **2012**, 109, 8400-8404

3. Semrau J. D., Jagadevan S., DiSpirito A. A., Khalifa A., Scanlan J., Bergman B. H., Freemeler B. C., Baral B. S., Bandow N. L., Vorobev A., Haft D. H., Vuilleumier S., Murrell J. C., Methanobactin and MmoD work in concert to act as the 'copper-switch' in methanotrophs, Environ. Microbiol., **2013**, 15, 3077-3086

4. Knapp C. W., Fowle D. A., Kulczycki E., Roberts J. A., Graham D. W., Methane monooxygenase gene expression mediated by methanobactin in the presence of mineral copper sources, Proc. Natl. Acad. Sci., **2007**, 104, 12040-12045

5. Hakemian A. S., Rosenzweig A. C., The biochemistry of methane oxidation, Annu. Rev. Biochem., **2007**. 76, 223–241

6. El Ghazouani A., Baslé A., Firbank S. J., Knapp C. W., Gray J., Graham D. W., Dennison C., Copper-binding properties and structures of methanobactins from Methylosinus trichosporium OB3b, Inorg. Chem., **2011**, 50, 1378-1391

7. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., **2010**, 192, 6497-6498

8. Stein L. Y., Bringel F., Dispirito A. A., Han S., Jetten M. S., Kalyuzhnaya M. G., Kits K. D., Klotz M. G., Op den Camp H. J., Semrau J. D., Vuilleumier S., Bruce D., Cheng J. F., Davenport K. W., Goodwin L., Han S., Hauser L., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Genome sequence of the methanotrophic alphaproteobacterium Methylocystis sp. Strain Rockwell (ATCC49242), J. Bacteriol., **2011**, 193, 2668-2669

9. Ward N., Larsen Ø., Sakwa J., Bruseth L., Khouri H., Durkin A. S., Dimitrov G., Jiang L., Scanlan D., Kang K. H., Lewis M., Nelson K. E., Methé B., Wu M., Heidelberg J. F., Paulsen I. T., Fouts D., Ravel J., Tettelin H., Ren Q., Read T., DeBoy R. T., Seshadri R., Salzberg S. L., Jensen H. B., Birkeland N. K., Nelson W.

C., Dodson R. J., Grindhaug S. H., Holt I., Eidhammer I., Jonasen I., Vanaken S., Utterback T., Feldblyum T. V., Fraser C. M., Lillehaug J. R., Eisen J. A., Genomic insights into methanotrophy: the complete genome sequence of Methylococcus capsulatus (Bath), PLoS Biol., **2004**, 2

10. Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V., Mondragón A., Molecular basis of the metal- ion selectivity and zeptomolar sensitivity by CueR, Science, **2003**, 301, 1383-1387

11. Singh S. K., Roberts S. A., McDevitt S. F., Weichsel A., Wildner G. F., Grass G. B., Rensing C., Montfort W. R., Crystal structures of multicopper oxidase CueO bound to copper (I) and silver (I): functional role of a methionine-rich sequence, J. Biol. Chem., **2011**, 286, 37849-37857

12. Radford D. S., Kihlken M. A., Borrelly G. P. M., Harwood C. R., Le Brun N. E., Cavet J. S., CopZ from *Bacillus subtilis* interacts in vivo with a copper exporting CPx-type ATPase CopA, FEMS Microbiol. Lett., **2003**, 220, 105-112

13. Badarau A., Dennison C., Thermodynamics of copper and zinc distribution in the cyanobacterium Synechocystis PCC 6803, Proc. Natl. Acad. Sci., **2011**, 108, 13007-13012

14. Zhou L., Singleton C., LeBrun N. E., High Cu(I) and low proton affinities of the CXXC motif of Bacillus subtilis CopZ, Biochemistry, **2008**, 413, 459-465

15. Xiao Z., Brose J., Schimo S., Ackland S. M., La Fontaine S., Wedd A. G., Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards, J. Biol. Chem., **2011**, 286, 11047-11055

16. Xiao Z., Loughlin F., George G. N., Howlett G. J., Wedd A. G., C-Terminal Domain of the Membrane Copper Transporter Ctr1 from Saccharomyces cerevisiae Binds Four Cu(I) Ions as a Cuprous-Thiolate Polynuclear Cluster: Sub-femtomolar Cu(I) Affinity of Three Proteins Involved in Copper Trafficking, J. Am. Chem. Soc., **2004**, 126, 3081-3090

17. Badarau A., Dennison C., Copper trafficking mechanism of CXXC-containing domains: Insight from the pH dependence of their Cu(I) affinities, J. Am. Chem. Soc., **2011**, 133, 2983-2988

18. Bagai I. Liu W., Rensing C., Blackburn N. J., McEvoy M. M., Substrate- linked conformational change in the periplasmic component of a Cu(I)/Ag(I) efflux system, J. Biol. Chem., **2007**, 282, 35695-35702

19. Rosenzweig A. C., Argüello J. M., Toward a molecular understanding of metal transport by P_{1B}-type ATPases, Curr. Top. Membr., **2012**, 69, 113-136

20. Klein J. S., Lewinson O., Bacterial ATP-driven transporters of transition metals: physiological roles, mechanisms of action, and roles in bacterial virulence, Metallomics, **2011**, 3, 1098-1108

21. Mijnendonckx K., Leys N., Mahillon J., Silver S., Van Houdt R., Antimicrobial silver: uses, toxicity and potential for resistance, Biometals, **2013**, 2, 609-621

22. Xu F. F., Imlay J. A., Silver(I), Mercury(II), Cadmium(II), and Zinc(II) target exposed enzymic iron-sulfur clusters when they toxify Escherichia coli, App. Environ. Microbiol., **2012**, 78, 3614-3621

23. Outten F. W., Huffman D. L., Hale J. A., O'Halloran T. V., The independent *cue* and *cus* systems confer copper tolerance during aerobic and anaerobic growth in *Escherichia coli*, J. Biol. Chem., **2001**, 276, 30670-30677

24. Xue Y., Davis A. V., Balakrishnan G., Stasser J. P., Staehlin B. M., Focia P., Spiro T. G., Penner-Hahn J. E., O'Halloran T. V., Cu(I) recognition via cation π and methionine interactions in CusF, Nat. Chem. Biol., **2008**, 4, 107-109

25. Long F., Su C-C., Zimmermann M. T., Boyken S. E., Rajashankar K. R., Jernigan R. L., Yu E. W., Crystal structures of the CusA efflux pump suggest methionine-mediated metal transport, Nature, **2010**, 477, 484-488

26. Andreini C., Bertini I., Cavallaro G., Holliday G. L., Thornton J. M., Metal ions in biological catalysis: from enzyme databases to general principles, J. Biol. Inorg. Chem, **2008**, 13, 1205-1218

27. Barnett J. P., Scanlan D. J., Blindauer C. A., Protein fractionation and detection for metalloproteomics: challenges and approaches, Anal. Bioanal. Chem., **2012**, 402, 3311-3322

28. Sevcenco A-M., Krijger G. C., Pinske M. W. H., Verhaert P. D. E. M., Hagen W. R., Hagedoorn P-L., Development of a generic approach to native metalloproteomics: application to the quantitative identification of soluble copper proteins in Escherichia coli, J. Biol. Inorg. Chem., **2009**, 14, 631-640

29. Schägger H., Von Jagow G., Blue native electrophoresis for isolation of membrane protein complexes in enzymatically active form, Anal. Biochem., **1991**, 199, 223-231

30. Ferrer M., Golyshina O. V., Beloqui A., Golyshin P. N., Timmis K. N., The cellular machinery of Ferroplasma acidiphilum is iron-protein-dominated, Nature, **2007**, 445, 91-94

31. Cvetkovic A., Menon A. L., Thorgersen M. P., Scott J. W., Poole II F. L., Jenney Jr F. E., Lancaster W. A., Praissman J. L., Shanmukh S., Vaccaro B. J., Trauger S. A., Kalisiak E., Apon J. V., Siuzdak G., Yannone S. M., Tainer J. A., Adams M. W. W., Microbial metalloproteomes are largely uncharacterised, Nature, **2010**, 466, 779-784

32. Barnett J. P., Scanlan D. J., Blindauer C. A., Fractionation and identification of metalloproteins from a marine cyanobacterium, Anal. Bioanal. Chem., **2012**, 402, 3371-337

33. Sun X., Yu G., Xu Q., Li N., Xiao C., Yin X., Cao K., Han J., He Q-Y., Putative cobalt- and nickel-binding proteins and motifs in Streptococcus pneumoniae, Metallomics, **2013**, 5, 928-935

34. Kozlov I. A., Kermani B. G., Melnyk P. C., Barker D. L., Zhao C., Hachmann J. P., Lebl M., Retention of histidine-containing peptides on a nickel-affinity column, J. Chrom. Sci., **2007**, 45, 207-211

35. Robichon C., Luo J., Causey T. B., Benner J. S., Samuelson J. C., Engineering Escherichia coli BL21 (DE3) derivative strains to minimize E. coli protein contamination after purification by immobilized metal affinity chromatography, App. Env. Microbiol., **2011**, 77, 4634-4646

36. Tottey, S., Waldron, K. J., Firbank, S. J., Reale, B., Bessant, C., Katsuko, S., Cheek, T. R., Gray, J., Banfield, M. J., Dennison, C., Robinson, N. J., Protein-folding location can regulate manganese-binding versus copper- or zinc-binding, Nature, **2008**, 455, 1138-1142

37. DiSpirito A. A., Zahn J. A., Graham D. W., Kim H. J., Larive C. K., Derrick T. S., Cox C. D., Taylor A., Copper binding compounds from Methylosinus trichosporium OB3b, J. Bacteriol., **1998**, 180, 3606-3613

38. Bagchi P., Morgan M. T., Bacsa J., Fahrni C. J., Robust affinity standards for Cu(I) biochemistry, J. Am. Chem. Soc., **2013**, 135, 18549–18559

39. Hellman U., Wernstedt C., Gonez J., Heldin C. H., Improvement of an 'in-gel' digestion procedure for the micropreparation of internal protein fragments for amino-acid sequencing, Anal. Biochem., **1995**, 224, 451-455

40. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T. L., Software BLAST+: architecture and applications, BMC Bioinformatics, **2009**, 10, 421

41. Bendtsen J. D., Nielsen H., Widdick D., Palmer T., Brunak S., Prediction of twin-arginine signal peptides, BMC Bioinformatics, **2005**, 6

42. Sievers F., Wilm A., Dineen D. G., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D., Higgins D., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol., **2011**, 7

43. Palmer T., Berks B. C., Moving folded proteins across the bacterial cell membrane, Microbiology, **2003**, 149, 547-556

44. Yung X., Le Brun N. E., Thomson A. J., Moore G. R., Chasteen N. D., The iron oxidation and hydrolysis chemistry of Escherichia coli Bacterioferritin, Biochemistry, **2000**, 39, 4915-492

45. Murrell J. C., McDonald I. R., Gilbert B., Regulation of expression of methane monooxygenases by copper ions, Trends Microb., **2000**, 8, 221-225

46. Paulech J., Solis N., Cordwell S. J., Characterization of reaction conditions providing rapid and specific cysteine alkylation for peptide-based mass spectrometry, Biochim. Biophys. Acta, **2013**, 1834, 372-379

47. Liu F., van Breukelen B., Heck A. J. R., Facilitating protein disulfide mapping by a combination of pepsin digestion, Electron Transfer higher energy Dissociation (EThcD) and a dedicated search algorithm SlinkS, Mol. Cell Proteomics, **2014**, 13, 2776-2786

48. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

49. Sutherland D. E. K., Stillman M.J., The 'magic' numbers of metallothionein, Metallomics, **2011**, 3, 444-463

50. Davies S. L., Whittenbury R., Fine structure of methane and other hydrocarbon-utilizing bacteria, J. Gen. Microbiol., **1970**, 61, 227-232

CHAPTER 3:

In vitro characterisation of recombinant Csp1 and Csp3 from *M. trichosporium* OB3b

3.1 Introduction

3.1.1 The identified Csp1 from *M. trichosporium* OB3b is hypothesised to serve copper storage

The metalloproteomic studies of *M. trichosporium* OB3b, presented in Chapter 2, led to the identification of Csp1. Csp1 was identified from *M. trichosporium* OB3b grown under high copper levels and was subsequently semi-purified from a soluble extract of pMMO-active *M. trichosporium* OB3b, grown under 5 µM Cu(II). This final experiment resulted in the isolation of Csp1 with appreciable amounts of copper bound. The primary sequence of Csp1 revealed the protein has a predicted twin arginine translocation (Tat)-leader peptide responsible for the export of folded proteins from the cytosol (1), and contains 13 Cys residues some of which are arranged in Cys-X-X-Cys and Cys-X-X-X-Cys motifs. Bioinformatics identified two more homologous proteins to Csp1 in *M. trichosporium* OB3b, Csp2 (57% sequence identity) which also has 13 Cys and a predicted Tat-leader peptide, and Csp3 (18% sequence identity), which is expected to be cytosolic, as it does not contain a Tat-leader, and has 18 Cys residues. Bioinformatics also identified two Csp1 protein homologues from *Nitrosospira multiformis* and *Pseudomonas aeruginosa* whose crystal structures have been solved for the metal-free proteins (pdb codes 3LMF and 3KAW, respectively). The structure of 3LMF, which has the highest sequence identity (25%) to Csp1, is a 4-helix bundle with all the Cys residues facing into the core of the bundle and none of which are involved in disulfide bonds. The identification of Csp1 with copper bound, together with the high number of Cys residues in Csp1 and the structure of the homologous 3LMF are intriguing and point towards a novel family of proteins designed for copper storage. Known protein families which are related to metal storage include metallothioneins (MTs) and ferritins. MTs are also small proteins that are rich in Cys residues and bind large amounts of metal (2), while ferritins are known to store iron in the form of a FeO(OH) mineral core, that is deposited inside a hollow shell formed by 24 4-helix bundle units (3-5).

3.1.2 Metallothioneins

Metallothioneins (MTs) are usually described as low molecular weight proteins which contain a large number of Cys residues (~20%,) and can bind significant amounts of metal, through thiol coordination (2). Originally identified in mammals, MTs are known to strongly bind metals with d^{10} electronic configurations, such as Cu(I), Ag(I), Au(I) or

Zn(II), Cd(II) and Hg(II). Mammalian MTs isolated from the liver often contain Zn(II), while forms isolated from the kidney may contain a mixture of Cu(I), Zn(II) or Cd(II) (2, 6). Because of their ability to bind a large number of metal ions and their metal-dependent expression (7), suggested roles for MTs involve metal storage, metal detoxification and protection against metal-induced oxidative stress (6, 8). The rapid degradation of apo-MTs by proteases, compared to metal-loaded forms, indicates that MTs are most likely a short-term response against stress caused by elevated metal concentrations (9-11).

The mechanism of metal binding by MTs was initially studied using ^{113}Cd nuclear magnetic resonance (12) in an attempt to elucidate the structure of the metal complexes inside MTs, while in other studies MTs were reconstituted in the presence of certain metals (11). Studies of rabbit MT originally revealed the formation of two metal clusters (A and B) (12). In MT from rabbit liver Zn(II) and Cd(II) bind to the two clusters in an ordered way, with cluster A occupied first by four Cd(II) ions, while cluster B is subsequently occupied by two Zn(II) and one Cd(II) ion (11). In the case of MT from rat liver, addition of Cu(I) to apo-MT revealed the metal shows binding preference for cluster B possibly through a cooperative mechanism, where binding of one metal ion favoured the binding of the next metal ion (11). On the other hand, when Cu(I) was added to Zn-Cd-loaded MT from rabbit liver metal replacement occurred in a distributed way where the binding of each metal ion is independent of other metal-binding events (6, 13). Notably, in this study emission spectroscopy was used to monitor the formation of Cu(I) clusters in rabbit liver MT (13). Cu(I) binding to MT resulted in emission at ~600 nm at room temperature, when the sample was excited as 300 nm, and the emission was found to be dependent on Cu(I) loading of the protein (13).

The crystal structure of a truncated Cu(I)-MT form from yeast (*S. cerevisiae*) (protein sequence shown in Figure 3.1 A) revealed the nature of the octanuclear Cu(I) cluster, in which six Cu(I) ions are coordinated in trigonal and two Cu(I) ions in digonal geometry, by ten Cys residues in total. The two digonally coordinated Cu(I) ions appear to be labile and, therefore, could be readily exchanged with other copper binding proteins (14). In further research, X-ray absorption spectroscopy and density functional theory were used to investigate the full form of the protein that was crystallised from yeast (15). X-ray absorption spectroscopy confirmed the presence of a Cu(I)-thiolate cluster

and density functional theory was used to model the structure of the full protein on the crystal structure previously solved for the truncated protein form (14). This study showed the additional residues present in the full form do not make a difference in the formation and coordination of the Cu(I) cluster (15). Moreover, the full protein loaded with only four Cu(I) ions was modelled and the Cu(I) ions were found to be arranged in a stable $[Cu_4S_6]^{2-}$ cluster (15), indicating that Cu(I) binding may occur in a non-cooperative manner where the binding of each Cu(I) ion is independent of the binding of the next Cu(I) ion (6, 16).

Unlike eukaryotic MTs, few bacterial MTs are known (8, 17). In *Synechoccocus elongatus* PCC7942 the smt operon has been identified that encodes an MT, named SmtA, and SmtB, a repressor of SmtA transcription (18). SmtA is expressed under high Zn(II) and is thought to be involved in zinc homeostasis. Nine Cys residues are responsible for binding four Zn(II) ions that form a cluster (19). In contrast to eukaryotic MTs, SmtA also employs two His residues in Zn(II) binding that play a role in stabilising the overall structure of the protein, through their ability to form hydrogen bonds and bind to Zn(II) with high affinity while maintaining an overall low negative charge for the cluster (20). Notably, one of the four Zn(II) ions in SmtA participates in the formation of a zinc finger and is inert to metal exchange (21). Homologues of SmtA are present in a number of other bacteria including *Pseudomonas aeruginosa*, *Pseudomonas putida* and *Anabaena* PCC7120, with almost all the zinc-binding residues conserved (10, 22). The second type of bacterial MT to be identified is MymT from mycobacteria, which is involved in the sequestration of excess copper (8). MymT does not have obvious sequence similarity to SmtA (8) and contains seven Cys residues. Cys-X-Cys and Cys-X-His motifs are involved in Cu(I) binding (Figure 3.1 B). Between four and six Cu(I) ions bind to MymT and form a cluster which is shielded from the solvent and is responsible for the luminescence the complex exhibits (8). The expression of MymT is strongly induced by copper and cadmium, but also by exposure to cobalt, nickel and zinc (8).

The formation of metal clusters is a common theme in MTs and they are often critical for the stability of the protein as apo-MTs degrade rapidly (9) and lack secondary structure (8, 14). Cys residues arranged in motifs such as Cys-X-X-Cys, Cys X-Y-Cys or Cys-X-Cys and in the case of SmtA and MymT also Cys-X-His (8, 22, 23). Coordination geometries may vary from tetrahedral, in the case of divalent ions such as

Zn(II) and Cd(II), to digonal or trigonal for monovalents, such as Cu(I) or Au(I) (23). Circular dichroism (CD) spectroscopy and UV-VIS spectroscopy provide useful tools for studying the secondary structure and the metal binding of proteins, respectively. The formation of a bond between Cu(I) and a reduced thiol group results in ligand to metal charge transfer bands in the 220-400 nm region (24-27), while the formation of metal clusters that are shielded from the solvent results in luminescence of the protein, a feature that can be studied by emission spectroscopy (23, 28).

3.1.3 Iron storage in ferritins

Ferritins are a family of proteins that store iron through a mechanism of Fe(II) oxidation that is followed by hydrolysis to form FeO(OH). Ferritins consist of 24 units, each one of which is a 4-helix bundle, and are packed to form a shell the core of which is hollow (5). In bacterioferritin from *E. coli*, each 4-helix bundle contains a ferroxidase center that facilitates the oxidation of Fe(II) to Fe(III) and a nucleation site mediating the subsequent hydrolysis reaction (3, 4). Notably, the formed FeO(OH) in ferritins is deposited inside the shell formed by the 24 4-helix bundle units and not inside the core of each 4-helix bundle (3-5). Iron storage by ferritin serves reserving iron for use by iron-utilising proteins but also confining the metal to protect the cell from iron-induced toxicity (5).

3.1.4 Metal binding by 4-helix bundles

Metal binding inside a well-defined protein structure, such as a coiled coil, is of great interest due to the simplicity and the functional diversity these structures offer (29, 30). A coiled coil consists of a bundle of α-helices (usually two, three or four, in which case a 4-helix bundle is formed) that are arranged in a parallel or antiparallel manner and form superhelical structures (30). In the bundle the α-helix is formed by repeats of a heptad of amino acid residues (abcdefg)$_n$ (30). A number of synthetic coiled coils have been designed to bind single metal ions or metal clusters containing copper, zinc or iron (29, 31), and motifs such as Cys-X-X-Cys have been incorporated in order to create metal binding sites. His residues have also been shown to contribute to metal binding in a coiled coil as their position in the α-helix can alter the coordination geometry of a certain metal and the metal binding specificity of the α-helical bundle (29).

4-helix bundles have been synthesised using residue sequences present in naturally occurring heme proteins, and have been shown to bind between one and four hemes by

ligation through His residues (32). Such synthetic 4-helix bundles have been described as 'heme-maquettes' (32, 33) and have been used for the further design of peptides that bind heme and also [4Fe-4S] clusters found in ferredoxins (33). In this case, Cys residues have been incorporated in the peptide, arranged in Cys-X-X-Cys motifs, for the binding of [4Fe-4S] clusters (33). In the absence of the [4Fe-4S] cluster the peptide showed no secondary structure, as monitored by CD spectroscopy (33). Moreover, the presence of the Cys residues proved to be indispensable for the formation of the [4Fe-4S] cluster, as replacement of three Cys residues by Ala resulted in the loss of the cluster formation. The intermediate residues (X) of the motif were also shown to have significant contribution in the formation of the [4Fe-4S] cluster (33). The incorporation of Cys residues in a 4-helix bundle for binding of a [4Fe-4S] cluster is interesting as naturally occurring Cys ligands in ferredoxins, usually come from β-sheets or, in the case they come from helices, they are located at the end of the helix, rather than in an intermediate motif (31). The role of Cys residues in the formation of α-helical secondary structure has also been shown in a synthetic hydrogenase where replacement of the Cys residues of a Cys-X-X-Cys motif, involved in the coordination of a diiron site, by Ala hinders the assembly of the α-helix (34).

Studies of copper binding to synthetic 4-helix bundle proteins show different types of copper sites can be engineered in the peptide chain. A stable apo-protein consisting of an antiparallel 4-helix bundle was synthesised with a His_2Cys ligand motif that, together with a weaker fourth ligand, such as Met residue, participates in the coordination of either a square planar or a distorted tetrahedral Cu(II) site (35). The secondary coordination sphere of the copper binding sites can be fine-tuned to improve the protein stability (36). Notably, naturally occurring proteins where a thiolate-coordinated mononuclear copper site is located in a 4-helix bundle are not known. For the synthesis of such a 4-helix bundle only a motif abundant in natural β-sheet proteins, (His_2Cys), was used (36). In another study, a synthetic peptide found as a random coil in the metal-free form was shown to assemble into a 4-helix bundle upon binding of one Cu(I) ion per synthetic peptide monomer (37). As a result, the peptide formed a tetramer in solution that contained a Cu_4S_4 cluster. The primary sequence of this peptide was designed to contain a Cys-X-X-Cys motif, however, only one Cys residue per monomer participated in Cu(I) binding (37). Consistent with the formation of a Cu(I) cluster,

Cu(I) binding by the peptide was followed by luminescence at 600 nm (37), as also shown for MTs (8, 13).

In the case of Csp1, a role in copper storage has been hypothesized on the basis of the identification of Csp1 as a copper protein from the soluble extract of *M. trichosporium* OB3b. Interestingly, the structure of the Csp1 protein homologue from *Nitrosospira multiformis* (3LMF) is a 4-helix bundle with all the Cys residues buried in the core of the bundle. The possibility of copper storage inside the core of a 4-helix bundle may be unique to this new family of proteins. In order to investigate the above hypothesis, Csp1 and the homologous Csp3, which is significantly different (18% sequence identity with Csp1, does not have a predicted Tat-leader and contains 18 Cys) were cloned, overexpressed and purified. The copper binding behaviour of Csp1 and Csp3 was studied while copper exchange experiments between Csp1 or Csp3 and the high affinity Cu(I) chelator methanobactin (mb) (38, 39) from *M. trichosporium* OB3b, which is secreted by MOB and mediates Cu(I) uptake (39, 40), were performed. Finally, Csp1 and Csp3 were crystallised and structures were determined, in order to provide insight in the relationship between structure and function for the two proteins.

A) S. cerevisiae MFSELINFQNEG**HECQCQCGSCKNNEQCQKSCSCPTGCNSDDKCPCGN**KSEETKKSCCSGK

B) SmtA   MTSTTLVKCACEPCLCNVDPSKAIDRNGLYYCSEACADGHTGGSKGCGHTGCNCHG----

   M. marinum   -------------------------MTWMREIPMATHEAGTMLTCGHEGCGCRVRIEV

   M. avium   -----------------------------MSHMATYESGTLLTCGHEGCGCRVRIEV

   M. tuberculosis   ----------------------------MRVIRMTNYEAGTLLTCSHEGCGCRVRIEV

   M. bovis BCG   ----------------------------------MTNYEAGTLLTCSHEGCGCRVRIEV

SmtA   ----------------------

   M. marinum   PCHCSGSGEPYRCTCGDALVPVQ

   M. avium   PCHCSGAGEEYRCTCGDALTPVK

   M. tuberculosis   PCHCAGAGDAYRCTCGDELAPVK

   M. bovis BCG   PCHCAGAGDAYRCTCGDELAPVK

Figure 3.1 A) The full protein sequence of the MT from *Saccharomyces cerevisiae*, where the truncated form that was crystallised is shown in bold (14). Cys residues are highlighted in yellow. B) Alignment of bacterial MT protein sequences including SmtA from *Synechococcus elongatus* PCC 7942 and some of the MymTs identified in different mycobacteria (8, 10). Conserved Cys and His residues are highlighted in yellow and green, respectively. Conserved Cys-X-His and Cys-X-Cys motifs are underlined.

3.2 Materials and methods

3.2.1 Cloning of Csp1 and Csp3

Csp1 without its Tat-leader peptide (Gly25 to Ala146) was amplified from *M. trichosporium* OB3b genomic DNA, using GCGCATATGGGAGAGGATCCTCATGC (forward) and GCGCCATGGTCAGGCGGCGACCTTATGGC (reverse) primers. For the polymerase chain reaction (PCR) Phusion polymerase was used, according to the manufacturer's protocol, and the PCR mix was prepared as described in section 6.10.2. The PCR samples were initially incubated for 5 min at 95 ℃, followed by 30 cycles consisting of 1 min at 95 ℃, 1 min at 60 ℃ and 1.2 min at 72 ℃ incubations. Finally the samples were incubated for 10 min at 72 ℃. The PCR product was A-tailed and cloned into pGEMT. Both strands of the gene were verified by sequencing. The Csp1 gene was, subsequently, cloned into the NdeI and NcoI sites of pET29a that introduced a Met residue at the N-terminus (pET29a_Csp1). The gene for Csp3 gene (Met1 to His133) was amplified with the same method as for Csp1, using GCGCATATGCATGTGGAAGCC (forward) and GCGCCATGGCTAATGCGTGAGCCCCGTCGC (reverse) primers. The PCR product was A-tailed, cloned into pGEMT and, subsequently, into the NdeI and NcoI sites of pET29a (pET29a_Csp3).

3.2.2 Expression and purification of Csp1 and Csp3

Csp1 and Csp3 were expressed and purified as described in section 6.13. The molecular weight of the intact proteins was verified by matrix assisted laser desorption ionisation (MALDI) time of flight (TOF) mass spectrometry as described in section 6.14 and the copper and zinc content of the purified proteins was determined by atomic absorption spectroscopy (AAS) as described in section 6.15.

3.2.3 Protein quantification

The concentration of apo-proteins was determined by denaturing the proteins in 20 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (Hepes) pH 7.5 plus 200 mM NaCl containing 8 M urea (final concentration at 7.4 M) plus 1 mM ethylenediaminetetraacetic acid (EDTA) and performing thiol quantification using dithiobis(2-nitrobenzoic acid) (DTNB, Ellman's reagent) and an $\varepsilon_{412nm}=14.15$ mM^{-1} cm^{-1} (41), as described in section 6.16. Csp1 and Csp3 concentrations were quantified assuming 13 and 18 thiols, respectively. The concentration of Cu(I)-Csp1 was

determined by Bradford assay (section 6.16). For Cu(I)-Csp3 the concentration was determined by DTNB assay after Cu(I) was removed by bathocuproine disulfonic acid (BCS) in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing 6.8 M guanidine hydrochloride (section 6.16).

3.2.4   Far-UV Circular Dichroism spectroscopy

Far UV (180-250 nm) Circular Dichroism (CD) spectra for folded proteins (0.3-0.7 mg/ml) were recorded in 100 mM phosphate pH 8 at 20 °C, as described in section 6.17 (42, 43). Unfolding of apo- and Cu(I)$_{18}$-Csp3, loaded with 18 molar equivalents of Cu(I), (0.15-0.4 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing either 8 M urea (at a final concentration $\geq$6.4 M) or 6.8 M guanidine hydrochloride (at a final concentration $\geq$5.4 M) were monitored by CD spectroscopy. The α-helical content of the proteins was determined from the experimental CD data using the mean residue ellipticity (mre) at 222 nm (44, 45), according to the following equation:

$$[([\theta]_{222}\text{-}3,000)/(\text{-}36,000\text{-}3,000)] \times 100$$

The α-helical content calculated from the experimental CD data were compared to that of the crystal structures of the proteins when processed on the STRIDE interface (46).

3.2.5   Analytical gel-filtration chromatography

Analytical gel-filtration chromatography of apo- and Cu(I)$_{12}$-Csp1 and Cu(I)$_{18}$-Csp3, loaded with 12 and 18 molar equivalents of Cu(I), respectively, (5-100 μM) was performed on a Superdex 75 GL 10/300 column (GE healthcare) equilibrated in 20 mM Hepes pH 7.5 plus 200 mM NaCl, as described in section 6.18 (25, 47). The column was calibrated using blue dextran (2000 kDa), albumin (67 kDa), ovalbumin (43 kDa), chymoptrypsinogen (25 kDa) and ribonuclease (13 kDa) in the same buffer.

3.2.6   Copper (I) binding stoichiometry of Csp1 and Csp3 and average Cu(I) affinity estimation for Csp3

The Cu(I) binding stoichiometries were estimated by titrating Cu(I) into apo-Csp1 (15 μM) or Csp3 (5 μM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, as described in section 6.19. For Csp3 the experiment was also repeated in the same buffer in the absence of NaCl (25). In order to verify the stoichiometry of tightly bound Cu(I) to Csp1 or Csp3, and also estimate the Cu(I) affinity in the case of Csp3, Cu(I) additions into Csp1 (11 μM) or Csp3 (2.5 μM) were also performed in the presence (100, 250,

500, 1000 μM) of bathocuproine disulfonic acid (BCS) ($\log\beta_2$=20.8) or bicinchoninic acid (BCA) ($\log\beta_2$=17.7) (48) (100, 250, 500, 1000 μM), using ε_{483nm}=12.5 mM^{-1} cm^{-1} and ε_{562nm}=7.7 mM^{-1} cm^{-1} for the quantification of $[Cu(BCS)_2]^{3-}$ and $[Cu(BCA)_2]^{3-}$, respectively (42). For Csp3 the experiment was also repeated for 100 μM BCA in 20 mM Mes pH 6.5 or 20 mM Taps pH 8.5, both containing 200 mM NaCl.

In the case of Csp3, protein solutions containing the appropriate concentration BCA or BCS were prepared and aliquoted in anaerobic quartz cuvettes into which Cu(I) additions corresponding to increasing Cu(I), by 3 molar equivalents to protein per cuvette (for BCA) and by 6 molar equivalents to protein per cuvette (for BCS), were made separately. According to an earlier experiment, in which the equilibration time was tested over time for Cu(I) additions (made by 4 molar equivalents to protein per cuvette) into apo-Csp3 (2.5 μM) in the presence of 100 μM BCA, samples were found to reach equilibration after 24 h. Based on this, therefore, the samples containing Csp3, with various BCS or BCA concentrations, in which Cu(I) additions were made were incubated in anaerobic cuvettes inside the anaerobic chamber, and equilibration was monitored by UV-VIS from 24 h onward.

In order to estimate the average Cu(I) affinity ($K_{association}$) of Csp3 the Hill equation (49) was used to fit the data for Cu(I) additions into Csp3 in the presence of 1000 μM BCA or 100 μM BCS:

$$Y= [Cu(I)]_{free}^{n} / (K_{Cu}^{n} + [Cu(I)]_{free}^{n}) \qquad (i)$$

Where Y represents the fractional occupancy of the protein, $[Cu(I)]_{free}$ is the concentration of free copper, K_{Cu} is the dissociation constant ($K_{dissociation}$=1/$K_{association}$), n is the Hill coefficient, which is a measure of the cooperativity for binding multiple Cu(I) ions. These parameters result from the following calculations. The concentration of Csp3-bound Cu(I) ($[Cu(I)]_{bound}$) is calculated from the total copper concentration $[Cu(I)]_{total}$ added into each sample after subtraction of the Cu(I) bound to BCA ($[Cu(BCA)_2]^{3-}$)

$$[Cu(I)]_{bound} =[Cu(I)]_{total} -[Cu(BCA)_2]^{3-} \qquad (ii)$$

Subsequently the value calculated from (ii) is used to calculate the occupancy of Csp3 corresponding to each Cu(I) addition, according to equation (iii):

$$Occupancy= [Cu(I)]_{bound}/ [Csp3] \qquad (iii)$$

The calculation of the fractional occupancy is obtained by dividing the occupancy of Csp3 for each Cu(I) addition by the maximum occupancy observed at the final point of the experiment:

$$\text{Fractional occupancy} = \text{Occupancy}/ \text{Occupancy}_{max} \quad (iv)$$

The total concentration of the ligand (BCA or BCS) ($[L]_{total}$) present in the samples is known and from this, and the $[Cu(BCA)_2]^{3-}$ the concentration of unbound ligand $[L]_{unbound}$ is calculated as follows:

$$[L]_{unbound} = [L]_{total} - 2 \text{ x } [Cu(L)_2]^{3-} \quad (v)$$

Finally the concentration of free Cu(I) ($[Cu(I)]_{free}$) corresponding to each Cu(I) addition is calculated from the equation:

$$[Cu(I)]_{free} = [Cu(L)_2]^{3-} (M)/ ([L]_{unbound} (M))^{\wedge}2 \text{ x } \beta_2 (M^{-2}) \quad (vi)$$

The relative occupancy is then plotted against $[Cu(I)]_{free}$ and the data is fitted in a non-linearly to the Hill equation (i) with n and K_{Cu}^n as variables.

3.2.7   Fluorescence

Emission spectra during the titration of Cu(I) into apo-Csp3 (5 μM) were recorded by exciting at 300 nm and monitoring the emission at 400-700 nm, as described in section 6.20 (25).

3.2.8   Cu(I) removal by BCS

Cu(I) removal from Cu(I)$_{13}$-Csp1, loaded with 13 molar equivalents of Cu(I), and Cu(I)$_{18}$-Csp3, loaded with 18 molar equivalents of Cu(I), (1-1.2 μM) by BCS (2.5 mM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl was performed as described in section 6.21, in the presence or absence of 8 M urea (at a final concentration of 7.4 M) or 6.8 M guanidine hydrochloride (at a final concentration of 6.3 M).

3.2.9   Isolation of methanobactin from *M. trichosporium* OB3b

Methanobactin (mb) was isolated from the cell-free medium of *M. trichosporium* OB3b cultures as described in section 6.5 (38).

3.2.10   Copper transfer experiments

For copper transfer between either Cu(I)-Csp1 or Cu(I)-Csp3 and full length apo-methanobactin (apo-mb) from *M. trichosporium* OB3b, samples were prepared as

described in section 6.22.1. 1.2 μM of $Cu(I)_{13}$-Csp1, loaded with 13 molar equivalents of Cu(I), was mixed with 15.4 μM (corresponding to 13 molar equivalents to $Cu(I)_{13}$-Csp1), 32.3 μM (corresponding to 27 molar equivalents to $Cu(I)_{13}$-Csp1) or 46.1 μM apo-mb (corresponding to 38.4 molar equivalents to $Cu(I)_{13}$-Csp1) in 20 mM Hepes pH 7.5 plus 200 mM NaCl and incubated for a total of 16 h. The samples were monitored over the first 3.5 h of the reaction at a set wavelength (394 nm). 0.95 μM $Cu(I)_{18}$-Csp3, loaded with 18 molar equivalents of Cu(I), was incubated with either 17 μM (corresponding to 18 molar equivalents to $Cu(I)_{18}$-Csp3) or 34 μM apo-mb (corresponding to 36 molar equivalents to $Cu(I)_{18}$-Csp3), in 20 mM Hepes pH 7.5 plus 200 mM NaCl, for 15 days. The amount of Cu(I)-mb formed in each case was quantified using the difference between the extinction coefficients at 394 nm for apo-mb and Cu(I)-mb, $\Delta\varepsilon_{394\ nm}= 7.2\ mM^{-1}\ cm^{-1}$. For copper exchange between either apo-Csp1 or apo-Csp3 and the full length of Cu(I)-mb from *M. trichosporium* OB3b, samples were prepared as described in section 6.22.2. Apo-Csp1 (72.5 μM or 234 μM) and apo-Csp3 (98 μM or 243 μM) were incubated with 2.6 μM Cu-mb, in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and incubated for up to 20 h.

3.2.11 Crystal trials

Crystal trials of apo- Csp1 and Csp3 and Cu(I)-loaded Csp1 and Csp3 were set up as described in section 6.23. Typically, apo-Csp1 and apo-Csp3 samples in 20 mM Hepes pH 7.5 were concentrated using a spinning centrifugal device. A 100-fold dilution of the final sample was checked by Bradford assay for protein concentration and the sample was diluted appropriately at a concentration of 20 mg/ml. For preparation of Cu(I)-loaded protein samples, apo-Csp1 (70-75 μM) was typically incubated with 12-14 molar equivalents of Cu(I) in 20 mM Hepes pH 7.5 plus 200 mM NaCl to give Cu(I)-Csp1 ($Cu(I)_{13}$-Csp1 for crystal structure). The sample was concentrated and a dilution of the final sample was made in order to assay the protein and copper concentration. The concentration of the Cu(I)-loaded protein was determined by Bradford assay. Cu(I)-Csp3 samples were prepared as above by loading apo-Csp3 (53-70 μM) with 2, 8, 16 or 18-19 molar equivalents of Cu(I) (to give $Cu(I)_2$-Csp3, $Cu(I)_8$-Csp3, $Cu(I)_{16}$-Csp3 and $Cu(I)_{19}$-Csp3, respectively). Protein concentration was determined by Bradford assay and Cu(I) concentration was assayed by AAS and by Cu(I) removal by BCS (section 6.15 and 6.21). Crystallisation screens on 96-well (using the sitting drop method of vapour diffusion) or 24-well plates (using the hanging drop method of vapour diffusion)

were set with a crystallisation robot or manually, respectively, and the plates were stored at 20 ˚C either aerobically for apo-Csp1 and apo-Csp3 or in the anaerobic chamber (where they were sealed) for Cu(I)-Csp1 and Cu(I)-Csp3. Diffraction data were collected at 100 K at the Diamond Light Source, U.K.. Data collection, model building and refinement were performed by Dr. Arnaud Baslé and the phase was determined using single-wavelength anomalous dispersion for copper. In the case of Cu(I)-Csp1 the oxidation state of copper was verified by X-ray absorption near edge spectroscopy (XANES), performed by Dr Neil Patterson, DLS). The analysis of the structures of the two proteins (Csp1 and Csp3) was performed by Semeli Platsaki with the help of Prof. Christopher Dennison who first analysed the structural data for Csp1. Secondary structure matching (ssm) superimposition was used to overlay and compare structures, referring to the root mean square deviation of the Cα atomic coordinates.

Protein	Concentration (Bradford assay) (mg/ml)	Protein: well solution ratio	Total drop volume (µl)	Condition
Apo-Csp1	20	1:1	2	0.1 M bis-Tris pH 6.5, 25% PEG 3350
Apo-Csp3	20	1:1	0.2	40% pentaerythritol propoxylate 5/4 PO/OH, 0.1M Hepes-NaOH pH 7, 200mM NaSCN
$Cu(I)_{13}$-Csp1	9.55	2:1	0.3	0.03 M $MgCl_2$, 0.03 M $CaCl_2$, 0.1 M Tris-Bicine pH 8.5, 37.5% 2- methyl- 2,4 pentanediol (MPD) (racemic) plus PEG 1000 plus PEG 3350
$Cu(I)_{XANES}$-Csp1	10.8	2:1	0.9[a]	0.025 M $MgCl_2$, 0.025 M $CaCl_2$, 0.1 M Tris-Bicine pH 8.5, 40.5% 2- methyl- 2,4 pentanediol (MPD) (racemic) plus PEG 1000 plus PEG 3350
$Cu(I)_{16}$-Csp3	13.5	2:1	0.3	0.2 M $MgCl_2$, 0.1 M Na Hepes pH 7.5, 30% PEG 400
$Cu(I)_{19}$-Csp3	16.6	1:1	2[b]	0.2 M $MgCl_2$, 0.1 M Na Hepes pH 7.5, 30% PEG 400
$Cu(I)_2$-Csp3	9.76	2:1	0.9[c]	0.2 M $MgCl_2$, 0.1 M Na Hepes pH 7.5, 30% PEG 400
$Cu(I)_8$-Csp3	12	2:1	0.9[c]	0.2 M $MgCl_2$, 0.1 M Na Hepes pH 7.5, 30% PEG 400

a: 96-well plate where the condition in which the $Cu(I)_{13}$-Csp1 crystal grew was optimised by varying the concentration of $MgCl_2$, $CaCl_2$ and MPD and by pipetting 300 nl drops. The $Cu(I)_{XANES}$-Csp1 crystal was used for the X-ray absorption near edge structure experiment. b: 24-well plate where the condition in which the $Cu(I)_{16}$-Csp3 crystal grew was optimised by varying the concentration of $MgCl_2$ and PEG 400 and by pipetting 1 µl drops. c: 96-well plate with the condition in which the $Cu(I)_{16}$-Csp3 crystal grew, where 300 nl drops were pipetted.

Table 3.1 Crystallisation conditions for Csp1 and Csp3.

3.3 Results

3.3.1 Protein purification and initial characterisation

The purity of Csp1 and Csp3 was determined by SDS-PAGE as ~90% (Figure 3.2) and the molecular weights determined by MALDI-TOF were within 0.5 Da of expected values (Table 3.2). Csp1 and Csp3 were isolated with negligible amounts of copper and zinc, as determined by AAS.

3.3.2 CD Spectroscopy

As shown in Figure 3.3, the CD spectra of apo-Csp1 and $Cu(I)_{13}$-Csp1 have two minima at 208 and 222 nm, which indicate mainly of α-helical secondary structure (44). The α-helical contents were calculated from the mean residue ellipticity at 222 nm (45) and are 78% and 76% for apo-Csp1 and $Cu(I)_{13}$-Csp1, respectively (Table 3.3). These values are in agreement with the α-helical content calculated from the crystal structures (described in sections 3.3.7.1 and 3.3.7.2) of the proteins (75% and 76% for apo-Csp1 and $Cu(I)_{13}$-Csp1, respectively) using the STRIDE interface (46) (Table 3.3). The CD spectra recorded for apo- and $Cu(I)_{18}$-Csp3 show minima at 208 nm and 222 nm (Figure 3.4) indicative of an α-helical secondary structure (44, 45). The α-helical content is 74% and 76% for apo-Csp3 and $Cu(I)_{18}$-Csp3 respectively, which is almost identical to the values (76% for both apo-Csp3 and $Cu(I)_{18}$-Csp3) calculated by STRIDE (46) for the crystal structures of apo-Csp3 and $Cu(I)_{18}$-Csp3 (described in sections 3.3.7.1 and 3.3.7.3) (Table 3.3).

For apo-Csp3 (0.4 mg/ml) unfolding in 6.7 M urea is complete after 15 mins (Figure 3.5 A) whereas Cu(I)-Csp3 in 6.4 M urea is not fully unfolded after 48 h (Figure 3.5 B), as denaturation is dependent on the excess of denaturant over protein. In the presence of 5.7 M guanidine hydrochloride apo-Csp3 (0.37 mg/ml) also unfolds rapidly (Figure 3.6 A) whereas the unfolding rate for $Cu(I)_{18}$-Csp3 depends on the ratio of guanidine hydrochloride over protein. At 0.37 mg/ml $Cu(I)_{18}$-Csp3, in the presence of 5.4 M guanidine hydrochloride, is not fully unfolded after 24 h (Figure 3.6 B) whereas at 0.15 mg/ml $Cu(I)_{18}$-Csp3, in the presence of 6.1 M guanidine hydrochloride, almost all secondary structure of the protein is lost after 24 h (Figure 3.6 C).

3.3.3 Analytical gel-filtration chromatography

The analysis of Csp1 on a semi-preparative Superdex S75 column shows apo-Csp1 elutes at 10.7 ml (Figure 3.7 A). $Cu(I)_{12}$-Csp1 elutes as a single peak in 10.8-10.9 ml

(Figure 3.7 B). The elution volumes for apo-Csp1 and Cu(I)$_{12}$-Csp1 are summarised in Table 3.4 and correspond to apparent masses of 51.8 kDa and 47.6-49.6 kDa, respectively, based on the calibration of the column. These are indicative of the formation of a tetramer in solution, given the mass of Csp1 is 12.591 kDa (Table 3.4). Apo- and Cu(I)$_{18}$-Csp3 were loaded onto the Superdex S75 column at concentrations that ranged from 100 μM to 5 μM (Figure 3.8) and the elution volumes are summarised in Table 3.5. At 100 μM apo-Csp3 elutes mainly at 11 ml, which corresponds to an apparent mass of 44.6 kDa, indicative of a trimer, whereas a small peak at 12 ml, corresponding to an apparent mass of 28.5 kDa (dimer) was also present. At lower concentrations the ratio of the two peaks changes and at 5 μM apo-Csp3, the peak corresponding to the dimer is the main peak (Figure 3.8 A). This behaviour indicates that apo-Csp3 may dissociate in solution in a concentration-dependent manner. In the case of Cu(I)$_{18}$-Csp3, however, the protein elutes as a single peak at all concentrations (Figure 3.8 B) and the elution volumes (11.2-11.6 ml) with the corresponding apparent molecular weight are shown in Table 3.5.

3.3.4   Cu(I) binding stoichiometry

Figure 3.9 A shows the UV-VIS difference spectra for apo-Csp1 upon addition of Cu(I) and the spectral changes observed in the range of 250-400 nm are due to the formation of ligand (S(Cys)) to metal charge transfer (LMCT) bands (24-27). The absorbance at 260 nm, 275 nm and 300 nm increases during Cu(I) addition (Figure 3.9 B) and starts to level at a plateau at ~11-12 Cu(I) ions to protein. These data indicate Csp1 binds approximately 11-12 Cu(I) ions. The Cu(I) titration into apo-Csp1 (11 μM) was repeated in the presence of 330 μM BCA in order to evaluate the stoichiometry of tightly bound Cu(I) to the protein. [Cu(BCA)$_2$]$^{3-}$ starts to form only after 10 Cu(I) ions have been bound to Csp1 (Figure 3.9 C), indicating tight binding of at least 10 Cu(I) ions to Csp1.

Cu(I) binding to Csp3 was monitored by UV-VIS and the difference spectra in Figure 3.10 A show the formation of S(Cys)→Cu(I) LMCT bands in the 250-400 nm region. The absorbance at 250, 275 and 315 nm was plotted as a function of the Cu(I) to Csp3 ratio (Figure 3.10 B), and indicates Csp3 binds approximately 17-18 Cu(I) ions at pH 7.5 in the presence of NaCl. The shoulder at 315 nm seen in the difference spectra (Figure 3.10 A) is a spectral feature that has previously been observed in the case of Atx1 and has been attributed to binding of chloride to a Cu$^I_2$-Atx1 dimer (25). For this

reason, the Cu(I) binding behaviour of Csp3 was also investigated in the absence of NaCl (Figure 3.11 A and B) the behaviour was almost identical to that in the presence of NaCl. Cu(I) binding to Csp3, both in the presence and absence of NaCl, resulted in luminescence at approximately 550 nm. Figure 3.10 C shows the intensity of the emission at 550 nm plotted against the ratio of Cu(I) to Csp3 in buffer containing NaCl. The emission intensity is maximum when 8 equivalents of Cu(I) have been added to Csp3 and subsequently decreases to effectively zero by the time Csp3 is loaded with ~18 Cu(I) ions. When the experiment was performed in the absence of NaCl (Figure 3.11 C) the intensity of the emission was lower than before and peaked at the addition of 10 molar equivalents of Cu(I) to Csp3.

In order to estimate the stoichiometry of tightly bound Cu(I) to Csp3, Cu(I) additions were made into apo-Csp3 (2.5 µM) in the presence of 100 µM BCA showing that $[Cu(BCA)_2]^{3-}$ starts to form only after Csp3 has bound approximately 17-18 Cu(I) ions confirming the stoichiometry of the protein, while significant differences in the Cu(I) binding stoichiometry of Csp3 were not observed at pH 6.5 and 8.5 (Figure 3.12). Cu(I) additions into apo-Csp3 (2.6 µM) at pH 7.5 were also made in the presence of various BCA concentrations (100, 250, 500 and 1000 µM) aiming to calculate the average Cu(I) binding affinity of Csp3. Figure 3.13 shows the $[Cu(BCA)_2]^{3-}$ formed as a function of the [Cu(I)]/[Csp3] ratio at different time points and at different BCA concentrations. Clearly longer equilibration times are required as BCA concentration increases. Plotting of the data after the samples were fully equilibrated (Figure 3.14 A) shows that at higher BCA concentrations the formation of $[Cu(BCA)_2]^{3-}$ starts at lower [Cu(I)]/[Csp3] ratios, indicating stronger competition between BCA and Csp3 for binding of the metal. The dataset corresponding to 1000 µM BCA from this experiment (Figure 3.14 A) was used to calculate the concentration of free Cu(I) ($[Cu(I)]_{free}$) and the occupancy of Csp3 at each addition point. The maximum occupancy of Csp3 was found to be 18.1 and by plotting the calculated fractional occupancy against $[Cu(I)]_{free}$ the data was fitted non-linearly to the Hill equation (49), which results in an average Cu(I) affinity ($1/K_{Cu}$) at the order of 10^{17} M^{-1} with a Hill coefficient of n=1.67 (Figure 3.14 B).

Subsequently, Cu(I) additions into apo-Csp3 (2.5 µM) were also made in the presence of BCS (100, 250, 500 and 1000 µM) which has higher affinity for Cu(I) ($log\beta_2$=20.8) compared to BCA ($log\beta_2$=17.7) (48). Figure 3.15 shows the concentration of $[Cu(BCS)_2]^{3-}$ formed against the [Cu(I)]/[Csp3] ratio, just after the samples were

prepared and after 72 h. Figure 3.16 A, were the points at 72 h from Figure 3.15 have been plotted against [Cu(I)]/[Csp3] for all BCS concentrations, indicates that BCS competes strongly with Csp3 for Cu(I) binding. From the data shown for 100 μM BCS a maximal occupancy of 17.6 was calculated for Csp3. When the fractional occupancy is plotted against the calculated $[Cu(I)]_{free}$ and the data were fitted to the Hill equation (49) (Figure 3.16 B) resulting in an average Cu(I) affinity $(1/K_{Cu})$ at the order of 10^{17} M^{-1} with a Hill coefficient of n=1.55.

3.3.5 Cu(I) removal by BCS

Cu(I) can be removed from $Cu(I)_{13}$-Csp1 by a large excess of BCS. The reaction is completed in 40 min (Figure 3.17). The amount of $[Cu(BCS)_2]^{3-}$ quantified corresponds to complete Cu(I) removal from Csp1 and is consistent between the two conditions (i.e. in the presence and absence of 7.4 M urea). In the case of Csp3, a large excess of BCS can efficiently remove Cu(I) from $Cu(I)_{18}$-Csp3 only when protein is unfolded in the presence of 6.3 M guanidine hydrochloride, and the reaction is completed in 20 min (Figure 3.18 A and B). For folded $Cu(I)_{18}$-Csp3 the reaction is very slow (Figure 3.18 C and D) and is not complete after 23 days.

3.3.6 Copper transfer experiments

The ability of Csp1 and Csp3 to exchange copper with mb was investigated with copper transfer experiments in both directions, i.e. between $Cu(I)_{13}$-Csp1 or $Cu(I)_{18}$-Csp3 and apo-mb, and between apo-Csp1 or apo-Csp3 and Cu(I)-mb. $Cu(I)_{13}$-Csp1 was incubated with 13, 27 or 38 molar equivalents of apo-mb and spectra of the samples were monitored over time (Figure 3.19 A-C). The absorbance at 394 nm rapidly decreases and reaches a plateau after 50 min for all three samples (inserts in Figure 3.19). In the case of the sample initially containing $Cu(I)_{13}$-Csp1 and 13 molar equivalents of apo-mb (Figure 3.19 A), effectively all Cu(I) is removed from $Cu(I)_{13}$-Csp1 by apo-mb and 12.2 μM Cu(I)-mb is formed corresponding to 80% of the Cu(I) removed from $Cu(I)_{13}$-Csp1. For the samples containing $Cu(I)_{13}$-Csp1 and 27 or 38 molar equivalents of apo-mb (Figure 3.19 B and C, respectively), the final spectra correspond to a mixture of Cu(I)-mb, apo-Csp1 and apo-mb. In these cases, more Cu(I)-mb than Cu(I) available from $Cu(I)_{13}$-Csp1 appeared to be formed. This observation, however, was not dependent on the amount of apo-mb present, as similar Cu(I)-mb concentrations were obtained in both cases (22.1 and 21.8 μM Cu(I)-mb for the samples containing 27 and 38 molar equivalents of of apo-mb, respectively). To check for

copper transfer in the opposite direction, Cu(I)-mb was incubated with a large excess of apo-Csp1 (Figure 3.20). The samples were incubated for 20 h and the spectrum of Cu(I)-mb remains constant indicating no copper was transferred to Csp1.

Cu(I)$_{18}$-Csp3 was incubated anaerobically with 18 or 36 molar equivalents of apo-mb (Figure 3.21 A and C). Both reactions were monitored over 15 days (inserts in Figure 3.21 A and C), at which point 19.8 μM and 22 μM Cu(I)-mb was quantified, respectively (corresponding to 115% and 130% of Cu(I) initially present in Cu(I)$_{18}$-Csp3, respectively). Control experiments with samples that contained the same concentration of apo-mb and the equivalent amount of Cu(I) to Cu(I)$_{18}$-Csp3 (Figure 3.21 B and D) resulted in the formation of 17.4 μM and 25 μM Cu-mb, respectively. Copper transfer in the opposite direction was studied by incubating a large excess of apo-Csp3 (Figure 3.22) with Cu(I)-mb for up to 20 h and no copper transfer was observed.

3.3.7 Structural characterisation of Csp1 and Csp3

3.3.7.1 *Crystal structure of apo-Csp1 and apo-Csp3*

The crystal structures of apo-Csp1 and apo-Csp3 were solved (by Dr. Arnaud Baslé) at 1.50 Å and 1.19 Å, respectively. Both proteins are tetramers in the asymmetric unit (Figure 3.23 A and 3.24 A). Figure 3.23 A shows the tetramer for apo-Csp1 which consists of a 4-helix bundle. In the tetramer, two sets of monomers are aligned in an antiparallel manner and the two pairs of monomers are rotated at an angle. All Cys residues point towards the centre of the bundle and, most importantly, none of the Cys residues are involved in disulfide bonds (Figure 3.23 B). Apo-Csp1 has three Met and one His residue at one end of the molecule, which might play a role in acquiring metal, while the opposite end of the bundle, where the N- and C- termini are located, is hydrophobic. The arrangement of the monomers, in the case of apo-Csp3, to form the tetramer are similar to apo-Csp1. In the apo-Csp3 monomer an additional small α-helix is present at the N-terminus and is connected to the 4-helix bundle by a loop which has not been modelled (Figure 3.24 B). The position of the additional α-helix of Csp3 can also be seen in Figure 3.25, where the crystal structure of apo-Csp3 has been superimposed on that of apo-Csp1 (Cα rmsd 1.29 Å), while the two 4-helix bundle cores of the proteins overlay well. The end of apo-Csp3 where the N- and C-termini located, is rich in Leu residues and, therefore, hydrophobic, whereas the opposite end of the

helical bundle is expected to be the opening of the molecule (Figure 3.24 B). As for apo-Csp1, all Cys residues point towards the core of the bundle (Figure 3.24 C) and do not form disulfide bonds. As expected, the superimposed structures of the two proteins (Figure 3.25) reveal the additional Cys residues of apo-Csp3 are distributed throughout the core of the 4-helix bundle to accommodate more copper.

3.3.7.2 *Crystal structure of Cu(I)-Csp1*

In order to verify the oxidation state of copper in the Csp1 crystal (Cu(I)$_{XANES}$-Csp1), a fresh crystal was used to perform X-ray absorption near edge spectroscopy (XANES) (by Dr Neil Patterson, DLS). The data acquired showed a peak at 8984 eV, characteristic of a copper 1s→4p transition, indicating two/three coordinate Cu(I) (Figure 3.26) (50, 51). The crystal structure of Cu(I)$_{13}$-Csp1 was solved (by Dr. Arnaud Baslé) at a 1.90 Å resolution by using single-wavelength anomalous dispersion for copper to determine the phase (51). The anomalous difference density data below the copper-edge show the structure contains 13 Cu(I) ions, at fully occupied sites, spanning the core of the 4-helix bundle. In their majority the copper ions are coordinated by thiolate sulfurs from Cys residues (Figure 3.27 A and Table 3.6), while Cu(I) binding does not significantly alter the structure of the protein as seen by the superimposition of the Cu(I)$_{13}$-Csp1 structure on that of apo-Csp1 (Cα rmsd 0.42 Å) (Figure 3.28). Eleven Cu(I) ions are within 2.7 Å from another Cu(I) site and some are close to more than one Cu(I) sites (Cu7 and Cu11) (Table 3.7). The Cu5 and Cu9 sites have larger distances from other Cu(I) sites at 2.8 and 2.9 Å, respectively (Table 3.7).

Most of the Cu(I) ions are coordinated in a digonal geometry where the Cu(I)-S(Cys) bonds and the S(Cys)-Cu(I)-S(Cys) angles range from 2.0 to 2.3 Å and from 158° to 176°, respectively. At Cu4, Cu11 and Cu13 the metal is coordinated in a trigonal geometry. At the Cu4 site Cu(I) is coordinated by three S(Cys) ligands with Cu(I)-S(Cys) bonds and the S(Cys)-Cu(I)-S(Cys) ranging from 2.2 to 2.4 Å and from 88° to 145° respectively (Figure 3.27 B). The Cu11 and Cu13 sites have atypical Cu(I) coordination (Figure 3.27 C). In Cu11, Cu(I) is coordinated by Cys51(2.3 Å), Cys103 (2.2 Å) and also Met48 (2.4 Å bond length), with bond angles ranging from 100° to 145° (Table 3.6). Cu13 is coordinated by Cys37 (2.1 Å) while His36 (N$^\delta$, 2.1 Å) and Met48 (2.6 Å) also provide ligands. These sites are located towards the opening of the 4-helix bundle and are potentially responsible for recruiting Cu(I) together with the

neighbouring Met40 and Met43 residues (Figure 3.27 A). Interestingly, in four Cu(I) sites (Cu1, Cu6, Cu8 and Cu12) S(Cys) ligands are involved in Cys-X-X-X-Cys motifs (Figure 3.27 B and C and Table 3.6) and the first Cys residue also interacts with the Cu(I) ion through the backbone carbonyl (2.0-2.3 Å distances). In other Cu(I) sites (Cu2, Cu3 and Cu4 as well as Cu10 and Cu11) the ligating Cys residues are located in adjacent helices (Figure 3.27 B and C).

3.3.7.3 *Crystal structures for Cu(I)-Csp3*

Crystal structures were obtained (by Dr. Arnaud Baslé) for Cu(I)-Csp3 loaded with different molar equivalents of Cu(I), in order to provide insight into the mechanism of Cu(I) binding by Csp3, and single-wavelength anomalous dispersion for copper was used to determine the phase as before. The crystal structure for Cu(I)$_2$-Csp3 was solved at 1.54 Å resolution and is that of a 5-helix bundle (Figure 3.29). The anomalous difference density data below the copper-edge show 4 partially occupied Cu(I) sites, inside the core of the 5-helix bundle, with a sum of occupancy factors at 1.4 (Figure 3.29 A). The Cu(I) sites with the highest occupancy factors are considered to be the most likely Cu(I) binding sites (51) and are assigned as Cu1a (0.4 occupancy factor) Cu2a (0.5 occupancy factor) (shown in orange in Figure 3.29 A and B), while the Cu(I) sites with the lowest occupancy factors (Cu1b and Cu2b both with 0.25 occupancy factor) are shown in grey. Cu(I) is coordinated either in trigonal or digonal geometry and the distance between Cu(I) sites is within 2.7 Å (Figure 3.29 B). Cu1a, Cu2a and Cu2b are coordinated digonally by S(Cys) which in the case of Cu1a and Cu2b come from Cys-X-X-X-Cys motifs. Cu(I)-S(Cys) bonds range from 1.9 Å to 2.3 Å and S(Cys)-Cu(I)-S(Cys) angles range from 152° to 157° (Figure 3.29 B and Table 3.8). In the case of Cu1a and Cu2b, Cys97 and Cys114, respectively, interact with the Cu(I) ions through the backbone carbonyl (2.4 Å) (Table 3.8). Cu1b is trigonally coordinated by Cys101 and Cys114, with Cu(I)-S bond lengths that range from 2.0 to 2.2 Å, and also by (O)Asn58 (2.4 Å bond length) with (O)Asn-Cu(I)-S(Cys) angles ranging from 97° to 116° (Figure 3.29 B and Table 3.8).

In the case of Cu(I)$_8$-Csp3, the crystal structure was solved at 1.55 Å resolution (by Dr. Arnaud Baslé) and the anomalous difference density data below the copper-edge revealed 18 partially occupied Cu(I) sites inside the core of the 5-helix bundle, with a sum of occupancy factors at 7.6, indicative of 8 Cu(I) ions (Figure 3.30 A and Table 3.9). Based on the occupancies observed for these sites (Table 3.9) and the intensity of

the anomalous difference density for copper, sites Cu1-Cu8 are more likely to be the actual Cu(I) binding sites. The most likely binding sites for the eight Cu(I) ions are shown as orange spheres in Figure 3.30 A, while the other sites are shown as grey spheres. Figure 3.30 B shows the Cu8 site, as an example of a Cu(I) site with higher occupancy (occupancy factor 0.55) and stronger signal from the anomalous difference density for copper. Cu(I) is trigonally coordinated by (S)Cys from Cys101 and Cys114, with Cu(I)-S(Cys) ranging from 2.1 to 2.2 Å, and by (O)Asn (2.4 Å) with O(Asn)-Cu(I)-S(Cys) ranging from 98° to 110° (Table 3.9). Sites Cu9 and Cu11 are shown as examples of sites with lower occupancy (0.4 and 0.3 occupancy factors, respectively) where the anomalous difference density for copper is weak (Figure 3.30 B). The Cu(I)$_{16}$-Csp3 crystal structure, solved at 1.15 Å resolution (by Dr. Arnaud Baslé) revealed 22 Cu(I) sites inside the core of the 5-helix bundle, based on the anomalous difference density data below the copper-edge, with a sum of total occupancies at 16, indicative of 16 Cu(I) ions (Figure 3.31 A). 10 Cu(I) sites are fully occupied (occupancy factor 1.00) while in 6 sites Cu(I) is distributed among two partially occupied sites (assigned as A and B) with occupancy factors that are shown in Table 3.10. Figure 3.31 B shows the coordination in two partially occupied sites for each of Cu14 and Cu16 Cu(I) ions.

The structure of the fully loaded Cu(I)$_{19}$-Csp3 was solved (by Dr. Arnaud Baslé) at lower resolution (2.30 Å) and the anomalous difference density data below the copper-edge revealed shows 19 fully occupied Cu(I) sites inside the core of the 5-helix bundle (Figure 3.32 A and Table 3.11). The superimposed structures of Cu(I)$_{19}$-Csp3 and apo-Csp3 (Cα rmsd 0.31 Å) (Figure 3.33) show Cu(I) binding does not significantly change the structure of the protein. In all the Cu(I) sites, with the exception of Cu3 and Cu12, Cu(I) is within 2.7 Å from another metal, while sites Cu5, Cu7, Cu9, Cu10 and Cu13 are in proximity to more than one other Cu(I) sites (Table 3.12). Cu(I) ions in the majority of the Cu(I) sites are coordinated digonally by S(Cys) ligands with Cu(I)-S(Cys) bonds ranging from 1.9 Å to 2.3 Å and S-Cu(I)-S angles ranging from 123° to 169° (examples are Cu12, Cu14-17). In Cu1, Cu3, Cu6, Cu8, Cu10, Cu12, Cu14 and Cu16 the coordinating Cys residues are involved in Cys-X-X-X-Cys motifs (Figure 3.32 B and C) and backbone carbonyl interactions (2.0-2.4 Å distances) from the first Cys residue also interact with the Cu(I) ion (Table 3.11). Site Cu13 is trigonally coordinated by S(Cys) ligands, with Cu(I)-S(Cys) bonds at 2.2 Å, and by O(Asn58) (2.4 Å) (Figure

3.32 B and Table 3.11). Site Cu18 is also trigonally coordinated by S(Cys) ligands, with Cu(I)-S(Cys) bonds ranging from 2.2 to 2.2 Å, and by N(His104) (2.3 Å) (Figure 3.32 C and Table 3.11). Cu19 is coordinated at the opening of the 5-helix bundle by His110 (2.1 Å) and Cys111 (2.1 Å) and with an N(His)-Cu(I)-S(Cys) angle of 158° (Figure 3.32 C and Table 3.11).

Protein	Experimental mass (kDa)	Theoretical mass (kDa)
Csp1	12.590,9	12591.4 (-Met)
Csp3	14.525,0	14.524.6 (-Met)

Table 3.2 Theoretical and experimental mass (MALDI-TOF) for recombinant Csp1 and Csp3.

	Percentage of α-helix			
	Apo-Csp1	Cu(I)-Csp1	Apo-Csp3	Cu(I)-Csp3
Far-UV CD spectrum	78%	76%	74%	76%
Crystal structure*	75%	76%	76%	76%

*The crystal structures were processed on the Stride interface to assess the percentage of residues participating in the formation of α-helices in the crystal structure.

Table 3.3 α-helical content of Csp1 and Csp3.

Protein	Concentration (μM)	Elution volume (ml)	Apparent molecular weight (kDa)
apo-Csp1	100	10.7	51.8
	20	10.7	51.8
	5	10.7	51.8
Cu(I)-Csp1	20	10.8	49.6
	5	10.9	47.6

Table 3.4 Gel-filtration analysis of Csp1.

Protein	Concentration (μM)	Elution volume (ml)		Apparent molecular weight (kDa)
apo-Csp3	100	Peak 1	11.0	44.6
		Peak 2	12.0	28.5
	60	Peak 1	10.9	45.8
		Peak 2	12.0	28.5
	40	Peak 1	11.0	43.7
		Peak 2	12.1	27.5
	20	Peak 1	10.9	46.2
		Peak 2	12.0	28.1
	5	Peak 1	10.8	47.1
		Peak 2	12.0	28.1
Cu(I)-Csp3	100	11.1		42.4
	20	11.1		41.8
	5	11.6		34.6

Table 3.5 Gel filtration analysis of Csp3.

Cu(I) site	Ligands and Cu(I)-ligand bond length (Å)			Bond angles (°)	Backbone carbonyl interactions from Cys, Cu(I)-O bond lengths (Å)
1	Cys113 2.1	Cys117 2.0		158	Cys113 2.3
2	Cys87 2.2	Cys117 2.1		170	
3	Cys62 2.1	Cys113 2.3		171	
4	Cys26 2.2	Cys62 2.4	Cys87 2.3	89, 125, 145	
5	Cys26 2.3	Cys110 2.0		165	
6	Cys106 2.1	Cys110 2.0		158	Cys106 2.3
7	Cys90 2.2	Cys106 2.0		176	
8	Cys90 2.2	Cys94 2.2		157	Cys90 2.0
9	Cys51 2.2	Cys94 2.1		166	
10	Cys33 2.2	Cys103 2.3		170	
11	Met48 2.4	Cys51 2.3	Cys103 2.2	100, 115, 145	
12	Cys33 2.1	Cys37 2.2		154	Cys33 2.1
13	His36 2.1	Cys37 2.1	Met48 2.6	92, 103, 162	

Table 3.6 The coordination ligands and other interactions for the Cu(I) sites in Cu(I)$_{13}$-Csp1 crystal structure, including bond lengths and bond angles

Cu(I) site	1	2	3	4	5	6	7	8	9	10	11	12	13
1		2.4	2.8										
2	2.4		3.1										
3	2.8	3.1		2.5									
4			2.5		2.8								
5						2.8							
6					2.8		2.5						
7						2.5		2.7					
8									2.9				
9								2.9			2.9		
10											2.6	2.6	
11									2.9	2.6			
12													2.8
13												2.8	

Table 3.7 The distances (Å) between the Cu(I) sites in the Cu(I)$_{13}$-Csp1 crystal structure.

Cu(I) site	Occupancy factor	Ligands and Cu(I)-ligand bond length (Å)			Bond angles (°)	Backbone carbonyl interactions from Cys, Cu(I)-O bond lengths (Å)
Cu1a	0.4	Cys97 2.0	Cys101 2.1		156	Cys97 2.4
Cu1b	0.25	Asn58 2.4	Cys101 2.2	Cys114 2.0	97, 116, 147	
Cu2a	0.5	Cys97 2.0	Cys118 2.1		152	
Cu2b	0.25	Cys114 1.9	Cys118 2.3		157	Cys114 2.4

The file used for this structure is CSP3-1974-01_molrep1_refmac10.pdb.

Table 3.8 The coordination ligands and other interactions for the Cu(I) sites in Cu(I)$_2$-Csp3 crystal structure, including bond lengths and bond angles.

Cu(I) site	Occupancy factor	Ligands and Cu(I)-ligand bond length (Å)			Bond angles (°)	Backbone carbonyl interactions from Cys, Cu(I)-O bond lengths (Å)
1	0.60	Cys121 2.1	Cys125 2.2		152	Cys121 2.3
2	0.70	Cys90 2.1	Cys125 2.1		151	
3	0.60	Cys90 2.1	Cys94 2.3		157	Cys90 2.4
4	0.45	Cys61 2.2	Cys65 2.0		161	Cys61 2.5
5	0.50	Cys31 2.1	Cys35 2.1		154	Cys31 2.3
6	0.55	Cys97 2.1	Cys101 2.2		152	Cys197 2.4
7	0.55	Cys114 2.2	Cys118 2.2		157	Cys114 2.4
8	0.55	Asn58 2.4	Cys101 2.2	Cys114 2.1	98, 110, 149	
9	0.40	Cys38 2.1	Cys101 2.3		156	
10	0.25	Cys38 2.1	Cys42 2.1		163	Cys38 2.5
11	0.30	Cys38 2.1	His110 1.93		94	
12	0.35	Cys35 2.2	Asn58 2.5	Cys97 2.1	94, 95, 165	
13	0.30	Cys97 2.2	Cys118 1.9		90	
14	0.45	Cys61 2.2	Cys118 2.1		93	
15	0.30	Cys31 2.1	Cys94 2.1		102	

122

16	0.25	Cys94 1.9	Cys121 2.0	163
17	0.50	Cys31 2.0	Cys65 2.2	153
18	0.45	Cys65 2.2	Cys121 2.0	170

The file used for this structure is CSP3-1974-09_molrep1_refmac19.pdb.

Table 3.9 The coordination ligands and other interactions for the Cu(I) sites in Cu(I)$_8$-Csp3 crystal structure, including bond lengths and bond angles.

Cu(I) site	Occupancy factor	Ligands and Cu(I)-ligand bond length (Å)			Bond angles (°)	Backbone carbonyl interactions from Cys, Cu(I)-O bond lengths (Å)
1	1.00	Cys121 2.1	Cys125 2.2		150	Cys121 2.2
2	1.00	Cys90 2.2	Cys125 2.2		157	
3	A 0.66 B 0.33	Cys65 2.2 Cys94 1.9	Cys121 2.1(A) 2.2(B)		168 158	
4	1.00	Cys90 2.2	Cys94 2.2		157	Cys90 2.3
5	A 0.66 B 0.33	Cys31 2.2(A) 2.1(B)	Cys94 2.1 Cys65 2.2		159 156	
6	1.00	Cys31 2.1	Cys35 2.1		152	Cys31 2.3
7	A 0.80 B 0.20	Cys61 2.2 Cys35 2.2	Cys118 2.2 Cys61 2.1		154 147	
8	1.00	Cys61 2.2	Cys65 2.1		157	Cys61 2.4
9	A 0.80 B 0.20	Cys35 2.2 Cys97 2.3	Cys97 2.1 Cys118 1.9		166 152	
10	1.00	Asn58 2.3	Cys101 2.2	Cys114 2.2	99, 110, 147	
11	1.00	Cys97 2.1	Cys101 2.2		151	Cys97 2.3

12	1.00	Cys114 2.2	Cys118 2.2	156	Cys114 2.3
13	1.00	Cys38 2.2	Cys101 2.3	139	
14	A 0.80 B 0.20	Cys42 2.2 Cys54 2.2	Cys54 2.1 Cys111 1.9	166 152	
15	1.00	Cys38 2.1	Cys42 2.2	155	
16	A 0.50 B 0.50	Cys38 2.0 Cys111 1.9	His110 2.1(A) 2.0(B)	172 168	

The file used for this structure is CSP3-CuI_1973-12_refmac14.pdb.

Table 3.10 The coordination ligands and other interactions for the Cu(I) sites in $Cu(I)_{16}$-Csp3 crystal structure, including bond lengths and bond angles.

Cu(I) site	Ligands and Cu(I)-ligand bond length (Å)			Bond angles (°)	Backbone carbonyl interactions from Cys, Cu(I)-O bond lengths (Å)
1	Cys24 2.2	Cys28 2.2		123	Cys24 2.0
2	Cys24 2.1	Cys125 2.1		169	
3	Cys121 2.1	Cys125 2.1		162	Cys121 2.3
4	Cys28 2.0	Cys90 2.2			
5	Cys65 2.2	Cys121 2.2		161	
6	Cys90 2.2	Cys94 2.1		135	Cys90 2.2
7	Cys31 2.1	Cys94 2.1		145	
8	Cys61 2.0	Cys65 2.1		160	Cys61 2.3
9	Cys61 2.2	Cys118 2.3		147	
10	Cys31 2.1	Cys35 2.2		145	Cys31 2.0
11	Cys35 2.1	Asn58 2.6	Cys97 2.1	90, 91, 167	
12	Cys114 1.9	Cys118 2.0		163	Cys114 2.4
13	Asn58 2.4	Cys101 2.2	Cys114 2.2	101, 110, 149	
14	Cys97 2.2	Cys101 2.3		151	Cys97 2.3
15	Cys38 2.3	Cys101 2.0		156	
16	Cys38 2.0	Cys42 2.2		148	Cys38 2.2

17	Cys42 2.0	Cys54 2.2		174
18	Cys54 2.3	Cys111 2.2	His104 2.3	121, 134, 104
19	His110 2.1	Cys111 2.1		158

All Cu(I) sites have been renumbered, with Cu1 located nearest to the C-terminus of the protein (CSP3-18eq_1969_13_3dii_molrep_refmac17.pdb was used for this structure).

Table 3.11 The coordination ligands and other interactions for the Cu(I) sites in Cu(I)$_{19}$-Csp3 crystal structure, including bond lengths and bond angles.

Cu(I) site	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1		2.5																	
2	2.5		3.0																
3		3.0		3.1															
4			3.1		2.6														
5				2.6		2.7	2.8												
6					2.7														
7					2.8			2.7	2.6	3.2									
8							2.7												
9							2.6			2.5									
10							3.2		2.5		2.7								
11										2.7		3.0	3.2	3.0					
12											3.0		2.8						
13											3.2	2.8		2.5	2.7				
14											3.0		2.5						
15													2.7			3.2			
16															3.2		2.9		2.7
17																2.9		2.5	
18																	2.5		
19																2.7			

Table 3.12 The distances (Å) between the Cu(I) sites in the Cu(I)$_{19}$-Csp3 crystal structure.
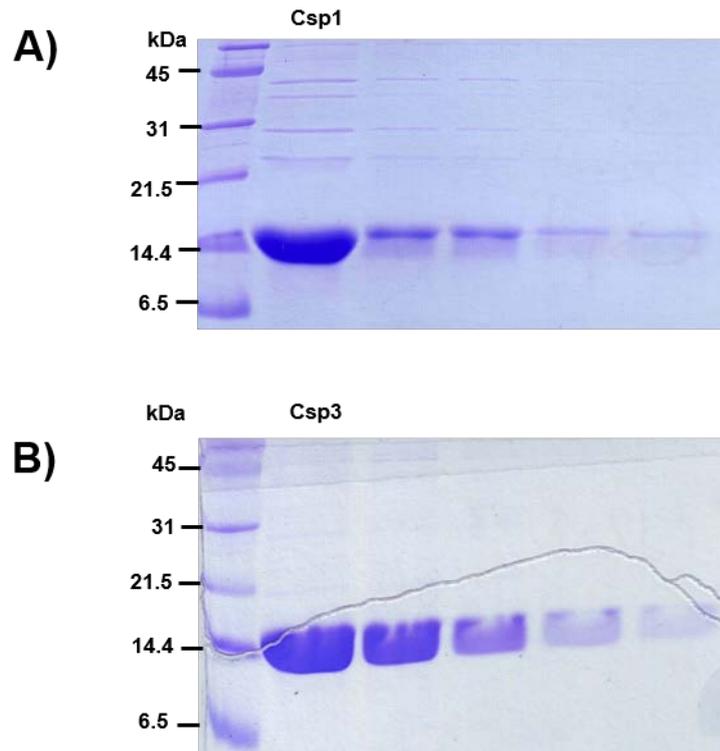
Figure 3.2 SDS-PAGE gel showing the purity of recombinant Csp1 and Csp3.
Concentrated fractions of purified Csp1 (A) and Csp3 (B), as well as serial dilutions of
these (twofold, fourfold, eightfold and sixteenfold), have been run on SDS-PAGE to
estimate the purity of the obtained recombinant proteins.
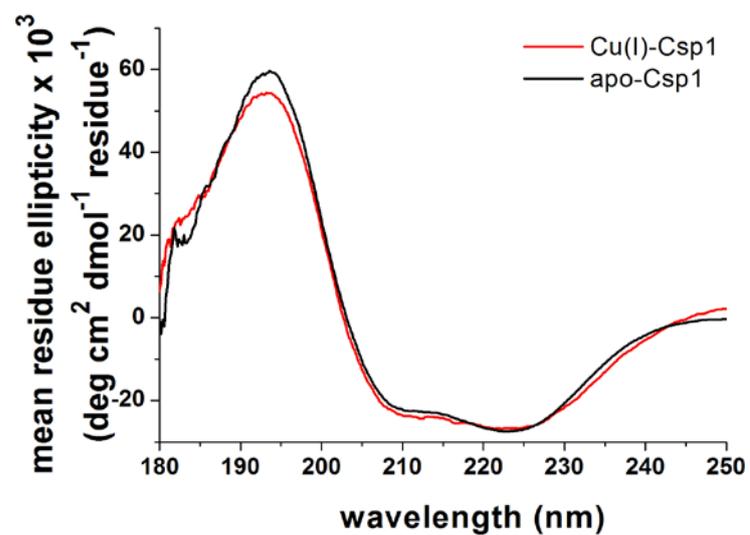
Figure 3.3 CD spectra of Csp1. Far-UV CD spectra for apo-Csp1 (black line) and Cu(I)$_{13}$-Csp1 (red line) in 100 mM phosphate pH 8 at 0.7 and 0.3 mg/ml respectively.
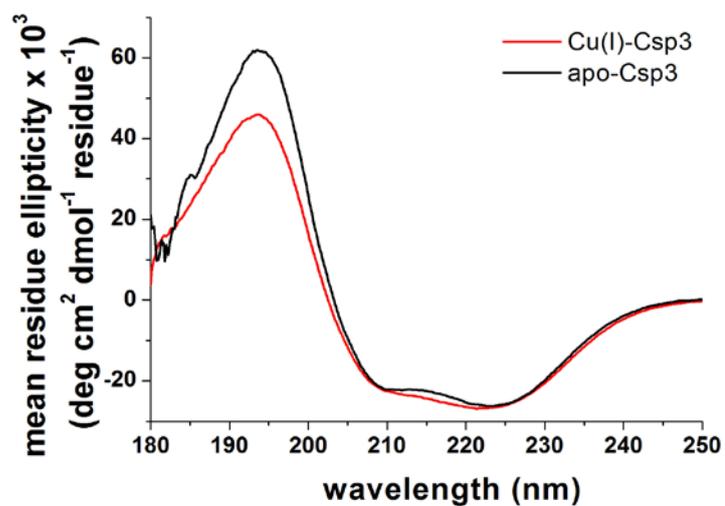
Figure 3.4 CD spectra of Csp3. Far-UV CD spectra for apo-Csp3 (black line) and Cu(I)$_{18}$-Csp3 (red line) in 100 mM phosphate buffer pH 8, at 0.55 mg/ml and 0.68 mg/ml respectively.
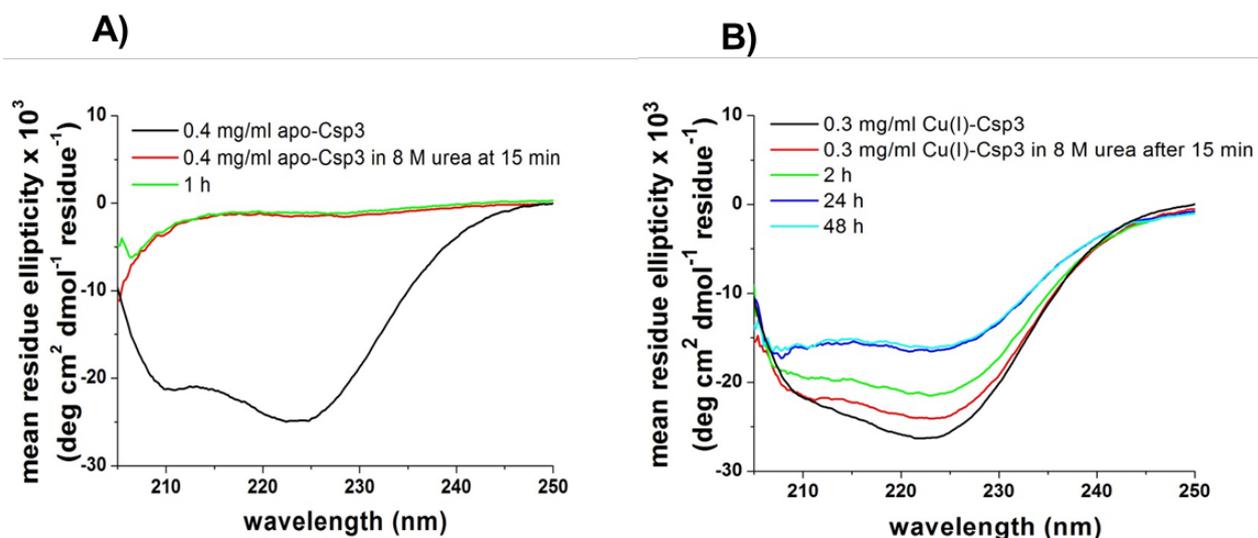
Figure 3.5 Unfolding of Csp3 in urea. A) Far-UV CD spectra of apo-Csp3 (0.4 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl (black line) and in the same buffer containing 8 M urea (6.7 M final concentration) after 15 min (red line) and 1 h (green line). B) Far-UV CD spectra of Cu(I)$_{18}$-Csp3 (0.3 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl (black line) and in the same buffer containing 8 M urea (6.4 M final concentration) after incubating for 15 min (red line), 2 (green line), 24 (blue line) and 48 (cyan line) h.
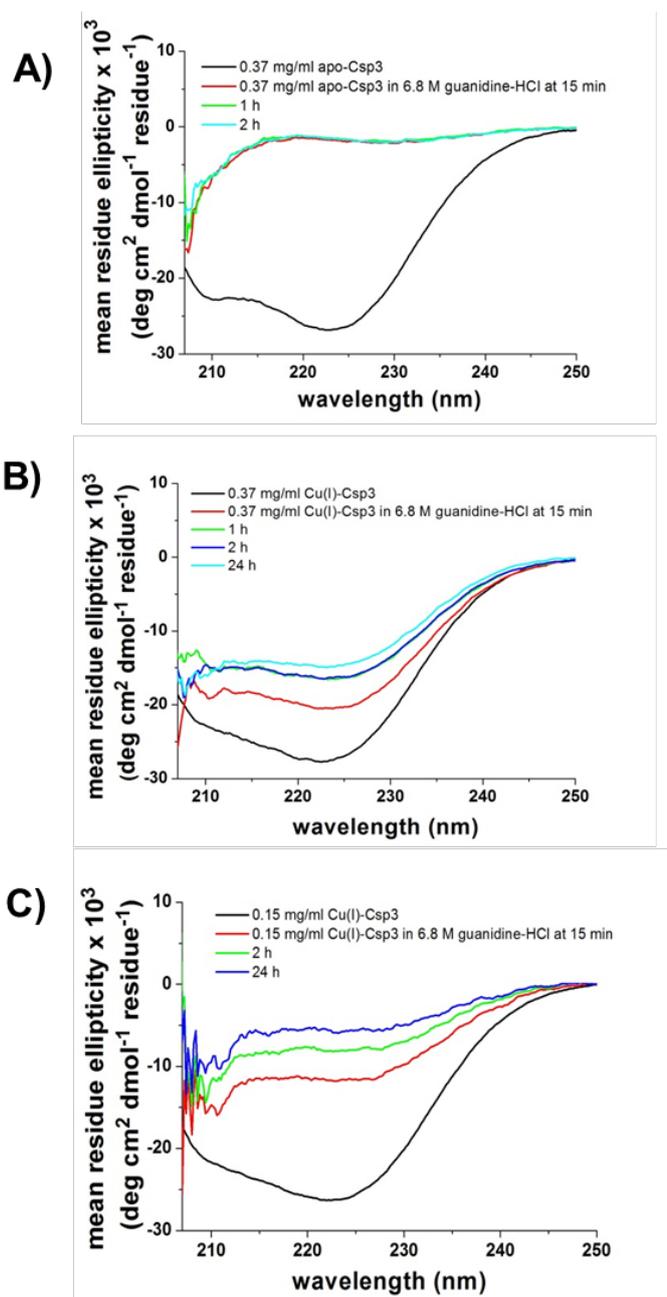
Figure 3.6 Unfolding of Csp3 in guanidine hydrochloride. A) Far-UV CD spectra of apo-Csp3 (0.37 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl (black line) and in the same buffer containing 6.8 M guanidine hydrochloride (5.7 M final concentration) after 15 min (red line), 1 h (green line), 2 h (cyan line). B) Far-UV CD spectra of $Cu(I)_{18}$-Csp3 (0.37 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl (black line) and in the same buffer containing 6.8 M guanidine hydrochloride (5.4 M final concentration) after 15 min (red line), 1 (green line), 2 (blue line) and 24 (cyan line) h. C) Far-UV CD spectra of $Cu(I)_{18}$-Csp3 (0.15 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl (black line) and in the same buffer containing 6.8 M guanidine hydrochloride (6.1 M final concentration) after 15 min (red line), 2 (green line) and 24 (blue line) h.
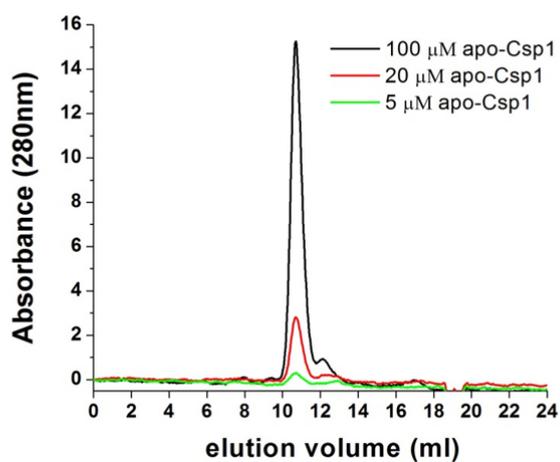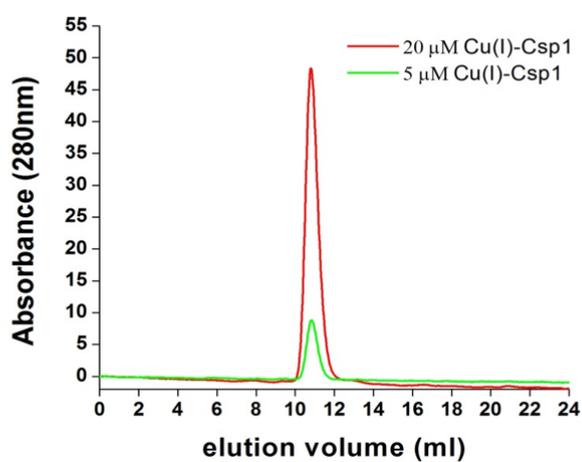
Figure 3.7 Analytical gel-filtration for Csp1. Analytical gel-filtration chromatograms of apo-Csp1 (A) at 100 (black line), 20 (red line) and 5 (green line) μM and Cu(I)$_{12}$-Csp1 (B) at 20 (red line) and 5 μM (green line) in 20 mM Hepes pH 7.5 plus 200 mM NaCl. The absorbance was monitored at 280 nm.
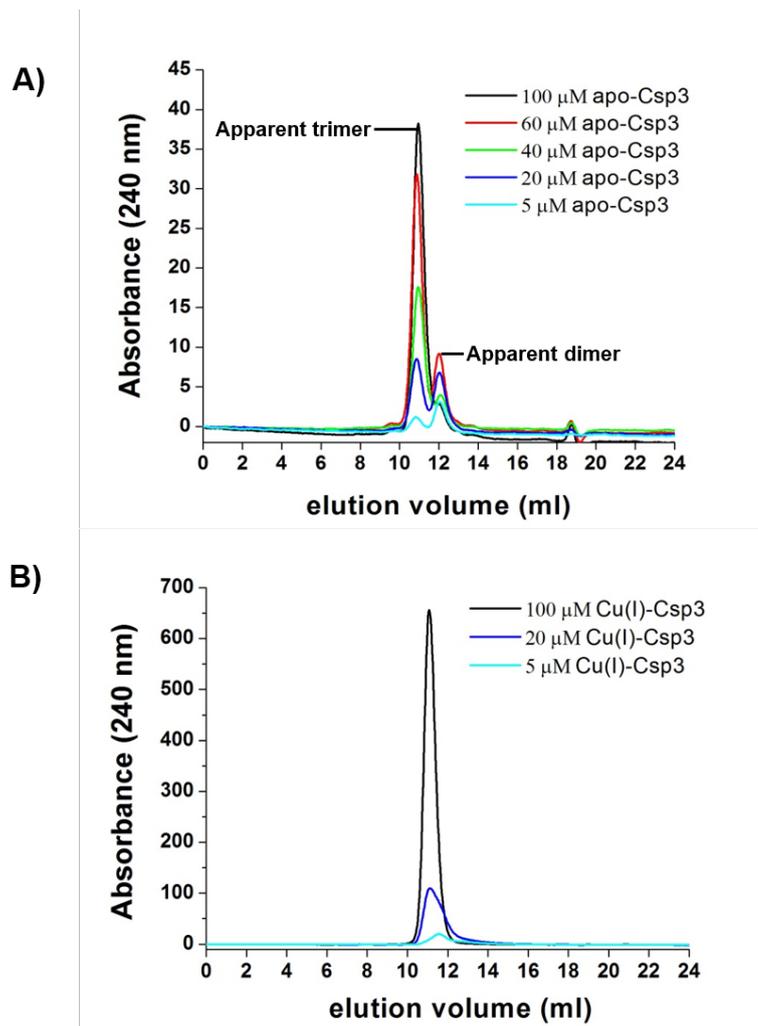
Figure 3.8 Analytical gel-filtration for Csp3. Analytical gel-filtration chromatograms in 20 mM Hepes pH 7.5 plus 200 mM NaCl for apo-Csp3 (A) at 100 (black line), 60 μM (red line), 40 μM (green line), 20 μM (blue line) and 5 μM (cyan line) and Cu(I)$_{18}$-Csp3 (B) at 100 μM (black line), 20 μM (blue line) and 5 μM (cyan line). The absorbance was monitored at 240 nm.
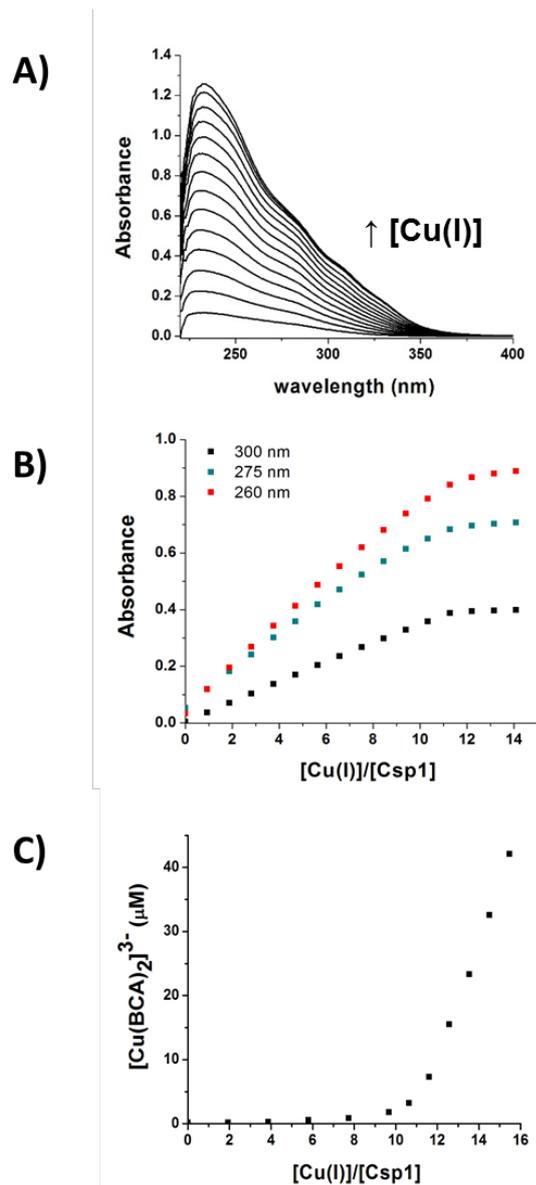
Figure 3.9 Cu(I) titration into apo-Csp1. A) UV-VIS difference spectra upon addition of Cu(I) to apo-Csp1 (15 µM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl. B) Plots of absorbance at 260 (red squares), 275 (cyan squares), 300 (black squares) nm against the [Cu(I)]/[Csp1] ratio. C) Cu(I) addition to apo-Csp1 (11 µM) in the presence of 330 µM BCA in 20 mM Hepes pH 7.5 plus 200 mM NaCl.
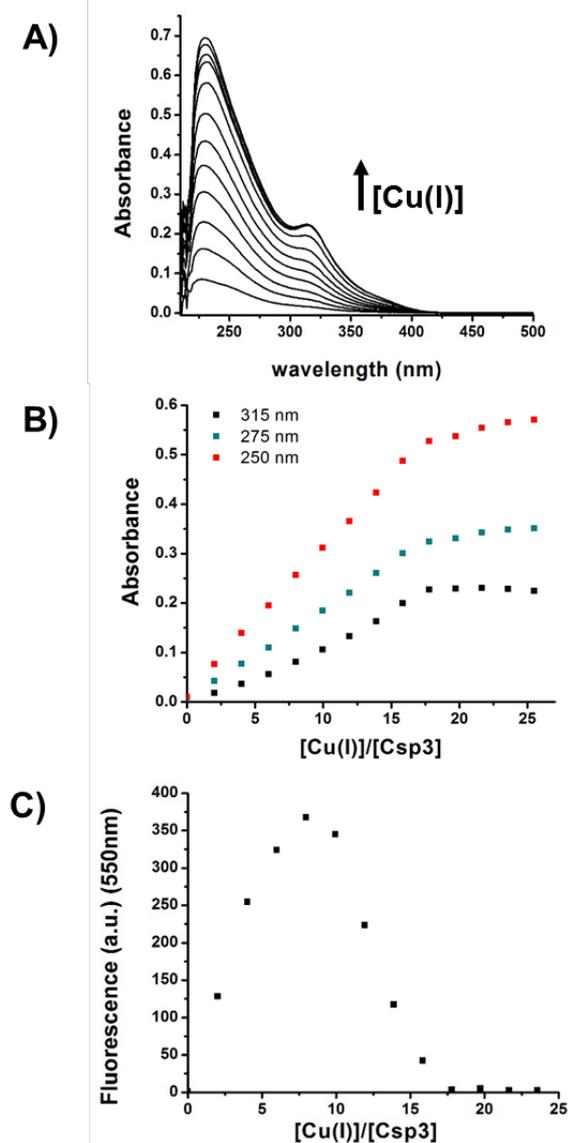
Figure 3.10 Cu(I) titration into apo-Csp3. A) UV-VIS difference spectra upon addition of Cu(I) to apo-Csp3 (5 μM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl. B) Plots of absorbance at 250 nm (red squares), 275 nm (cyan squares), 315 nm (black squares) against [Cu(I)]/[Csp3] ratio. C) Fluorescence emission monitored at 550 nm upon addition of Cu(I) to apo-Csp3 and excitation of the sample at 300 nm.
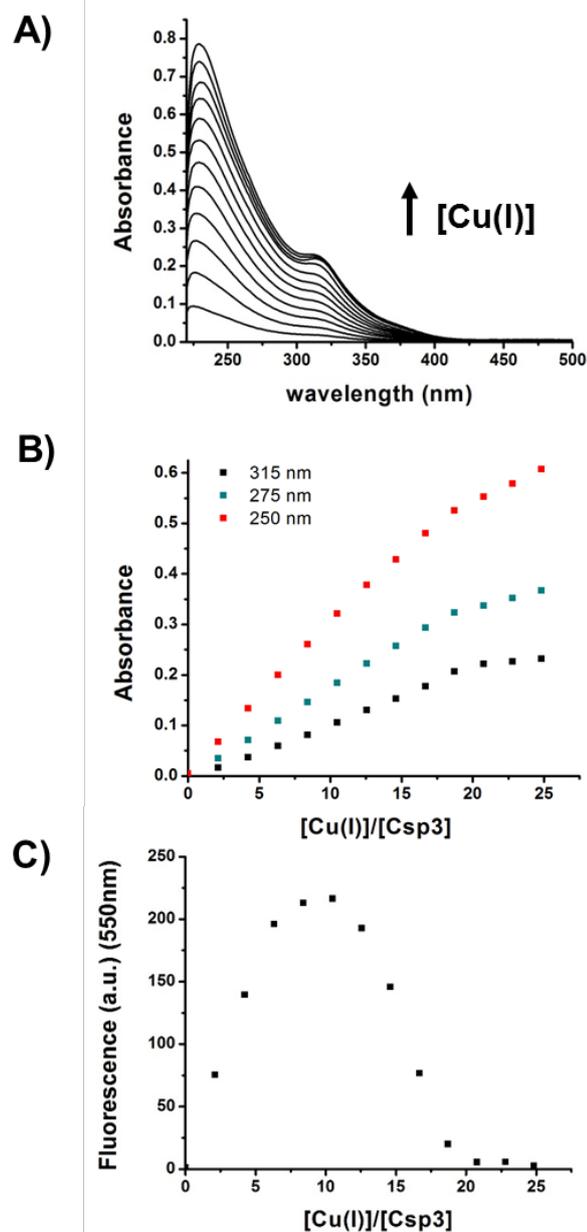
Figure 3.11 Cu(I) titration into apo-Csp3 in the absence of NaCl. A) UV-VIS difference spectra upon addition of Cu(I) to apo-Csp3 (5 μM) in 20 mM Hepes pH 7.5. B) Plots of absorbance at 250 nm (red squares), 275 nm (cyan squares), 315 nm (black squares) against [Cu(I)]/[Csp3] ratio. C) Fluorescence emission monitored at 550 nm upon addition of Cu(I) to apo-Csp3 and excitation of the sample at 300 nm.
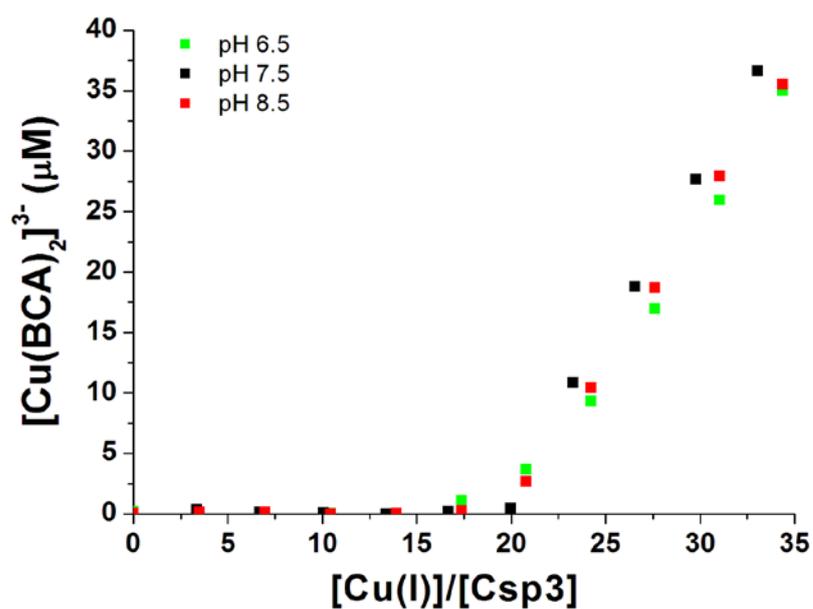
Figure 3.12 Cu(I) additions into apo-Csp3 in the presence of 100 μM BCA at varying pH. Cu(I) addition to apo-Csp3 (2.5 μM) in the presence of 100 μM BCA in 20 mM Mes pH 6.5 (green squares) or Hepes pH 7.5 (black squares) or Taps pH 8.5 (red squares), all containing 200 mM NaCl.
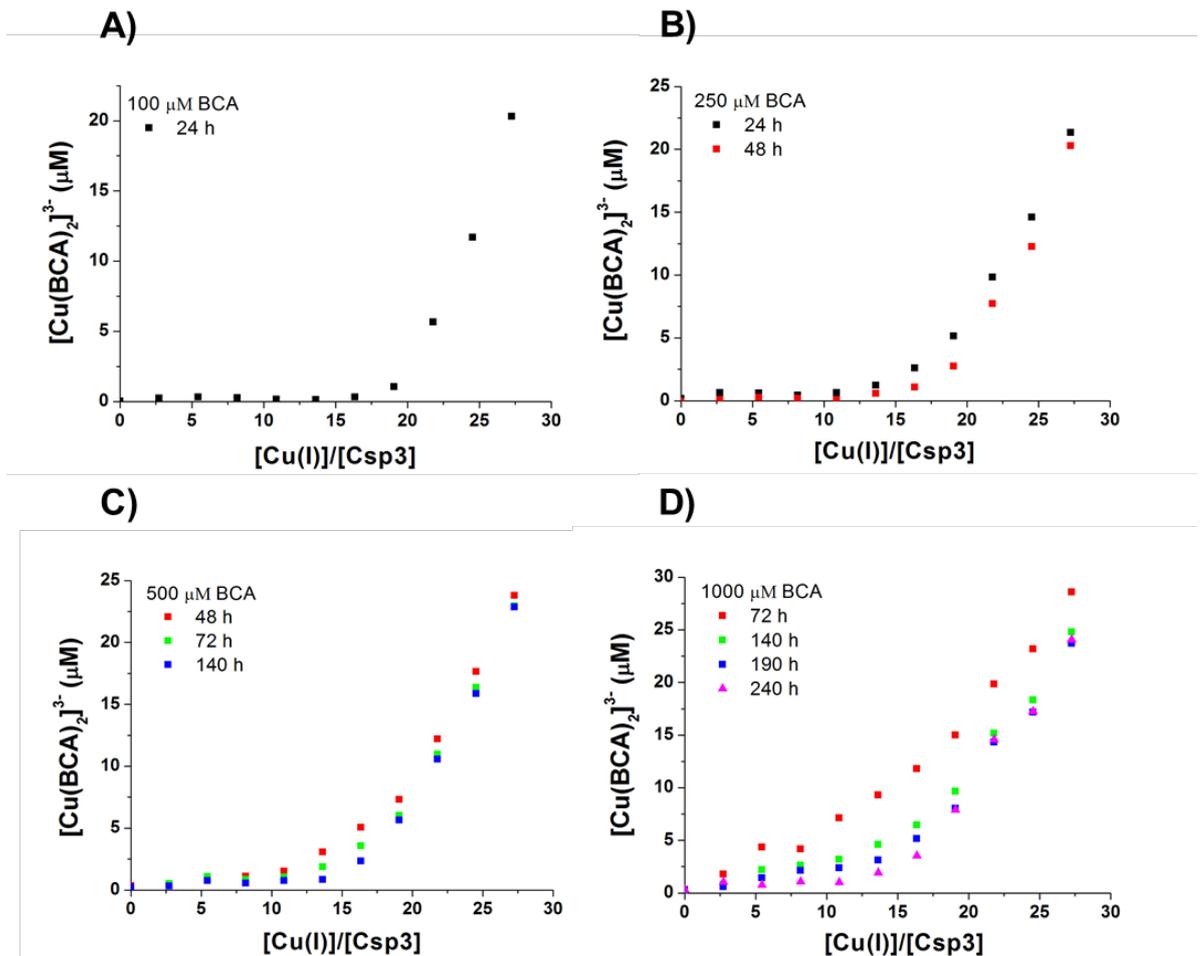
Figure 3.13 Equillibration time for Cu(I) additions into apo-Csp3 in the presence of varying BCA concentrations. Cu(I) was added, by 3 molar equivalents per addition, into apo-Csp3 (2.6 μM) in the presence of 100 (A), 250 (B), 500 (C) and 1000 (D) μM BCA in 20 mM Hepes pH 7.5 plus 200 mM NaCl and samples were scanned at 24 h (black squares), 48 h (red squares), 72 h (green squares), 140 h in C) or 190 h in D) (blue squares), and 240 h (magenta triangles) to monitor the formation of $[Cu(BCA)_2]^{3-}$.
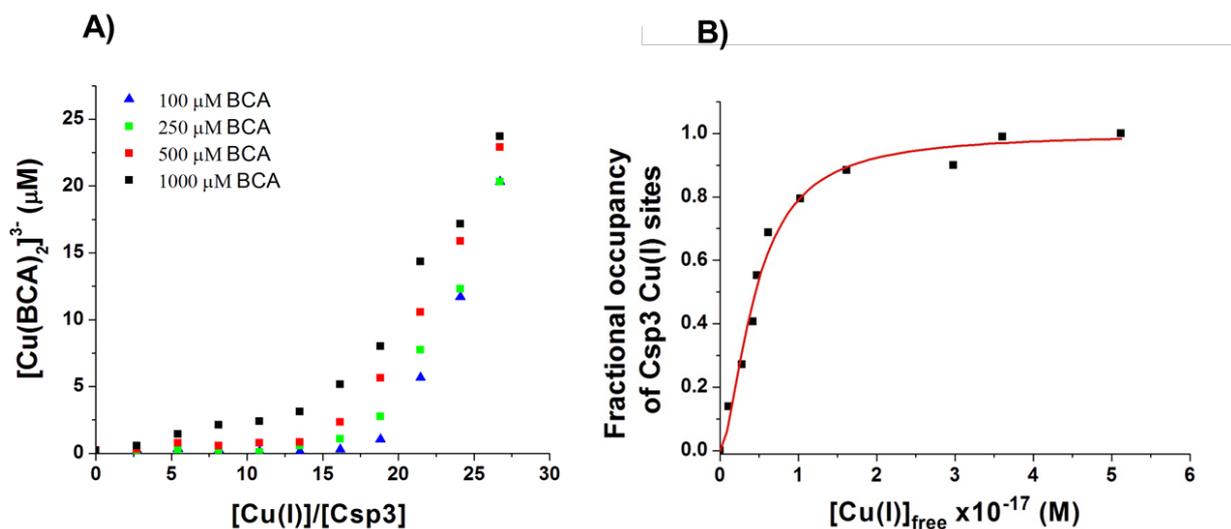
Figure 3.14 Estimation of Cu(I) affinity of Csp3 from BCA data. A) Cu(I) additions, by 3 molar equivalents per addition, into apo-Csp3 (2.6 μM) in the presence of 100 (blue triangles), 250 (green squares), 500 (red squares) and 1000 (black squares) μM BCA in 20 mM Hepes pH 7.5 plus 200 mM NaCl. The samples were allowed to equilibrate (final points taken from Figure 3.13) and the formed $[Cu(BCA)_2]^{3-}$ has been plotted against [Cu(I)]/[Csp3]. B) The dataset corresponding to 1000 μM BCA from A) (black) was used to calculate the fractional occupancy of the Cu(I) sites of Csp3 and the [Cu(I)]$_{free}$ for each addition point. The data was fitted non-linearly to the Hill equation from which an average Cu(I) affinity at the order of 10^{17} M^{-1} was calculated.
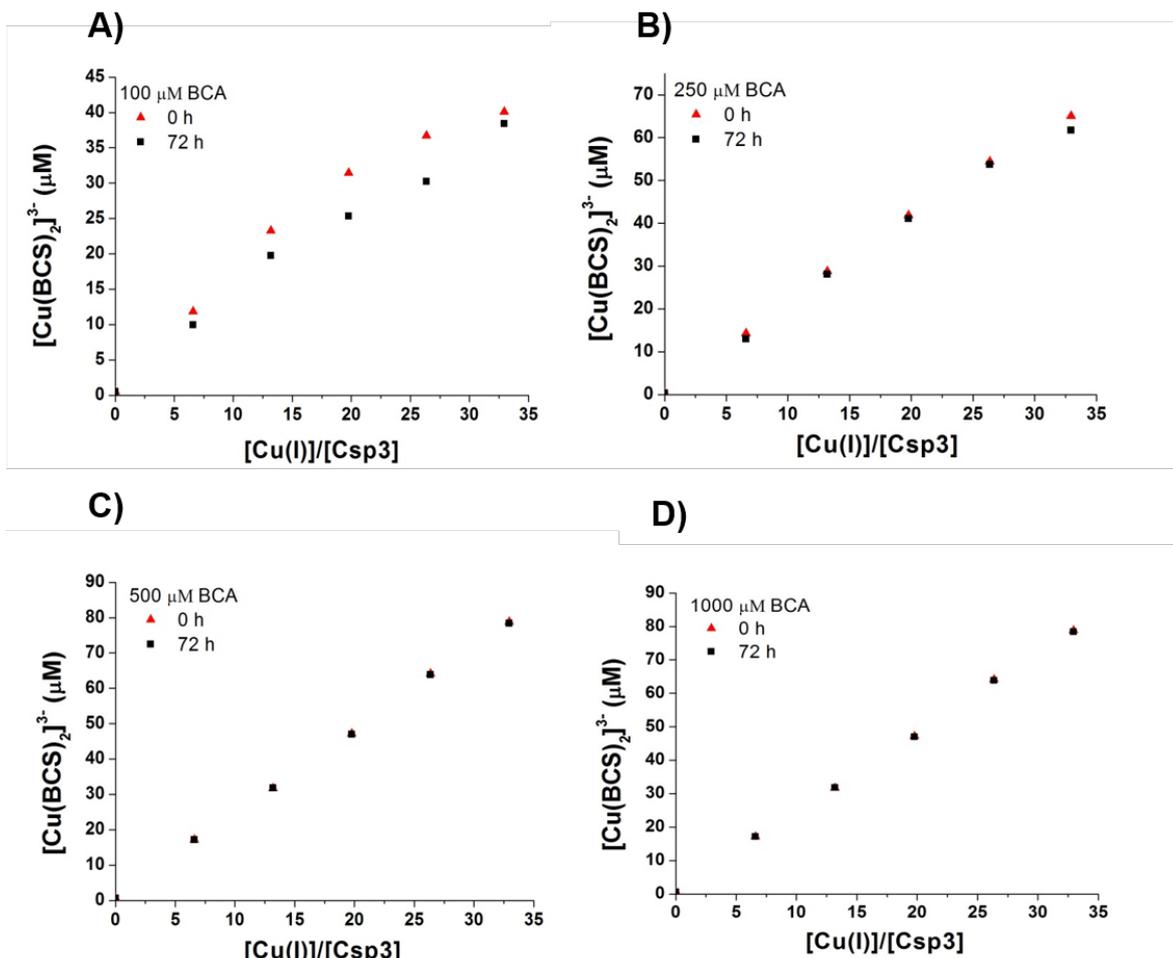
Figure 3.15 Equillibration time for Cu(I) additions into apo-Csp3 in the presence of varying BCS concentrations. Cu(I) was added, by 6 molar equivalents per addition, into apo-Csp3 (2.5 μM) in the presence of 100 (A), 250 (B), 500 (C) and 1000 (D) μM BCS in 20 mM Hepes pH 7.5 plus 200 mM NaCl and the samples were scanned immediately after the Cu(I) addition (red triangles) and after 72 h (black squares) to monitor the formation of $[Cu(BCS)_2]^{3-}$.
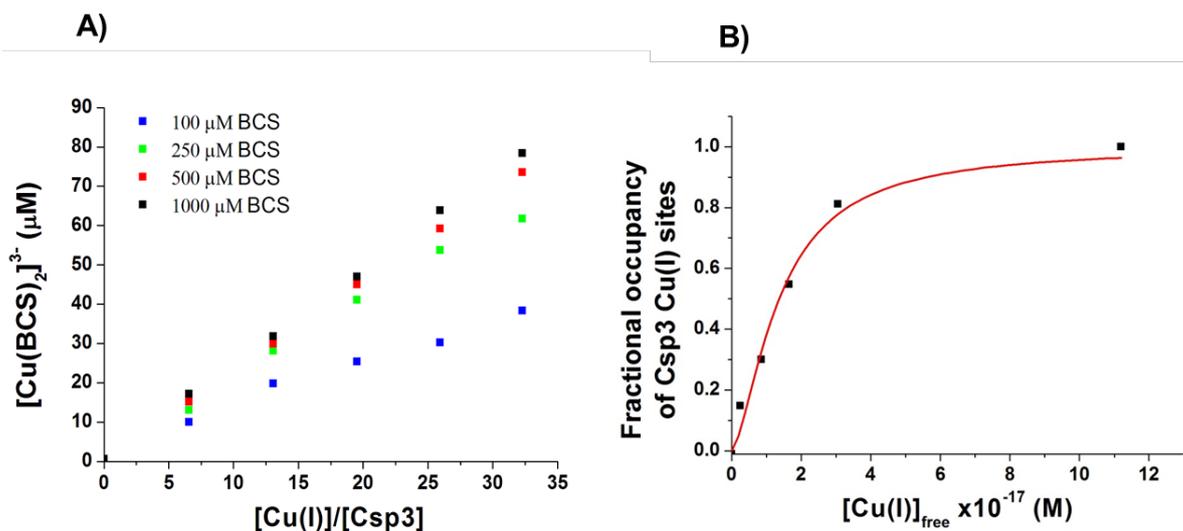
Figure 3.16 Estimation of Cu(I) affinity of Csp3 from BCS data. A) Cu(I) additions, by 6 molar equivalents per addition, into apo-Csp3 (2.5 μM) in the presence of 100 (blue squares), 250 (green squares), 500 (red squares) and 1000 (black squares) μM BCS in 20 mM Hepes pH 7.5 plus 200 mM NaCl. The formed $[Cu(BCS)_2]^{3-}$ has been plotted against [Cu(I)]/[Csp3] at 72 h (points taken from Figure 3.10). B) The dataset corresponding to 100 μM BCS from A) (blue) was used to calculate the fractional occupancy of the Cu(I) sites of Csp3 and the $[Cu(I)]_{free}$ for each addition point. The data was fitted non-linearly to the Hill equation from which an average Cu(I) affinity at the order of 10^{17} M^{-1} was calculated.
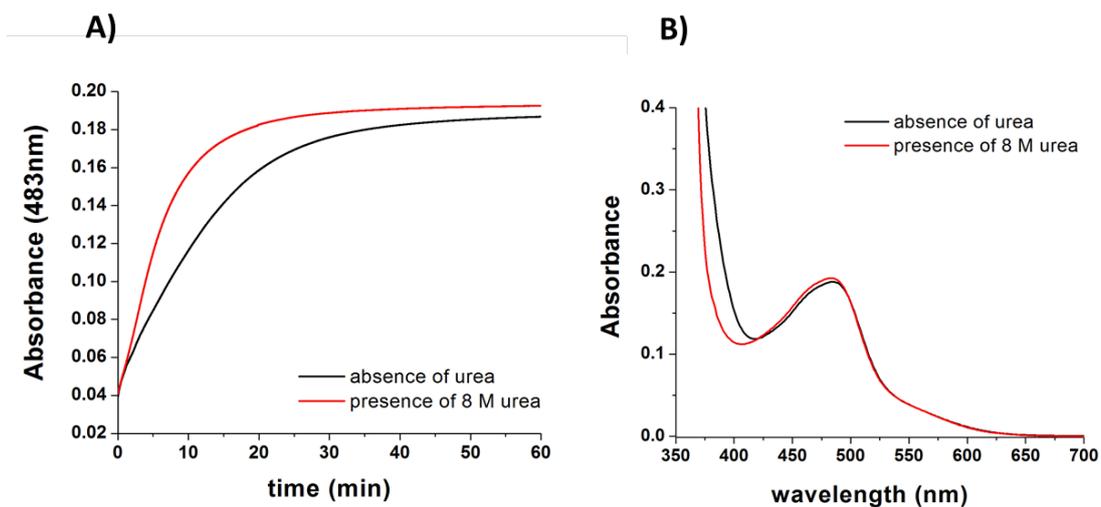
Figure 3.17 Cu(I) removal from Cu(I)-Csp1 by BCS. A) Time-course showing the removal of Cu(I) from Cu(I)$_{13}$-Csp1 (1.2 μM) by an excess of BCS (2.5 mM) and the formation of the [Cu(BCS)$_2$]$^{3-}$ complex in 20 mM Hepes pH 7.5 plus 200 mM NaCl in the presence (red line) and absence (black line) of 8 M urea (at a 7.4 M final concentration). B) UV-VIS spectra of the samples from A at the end of the time-course (red and black lines).
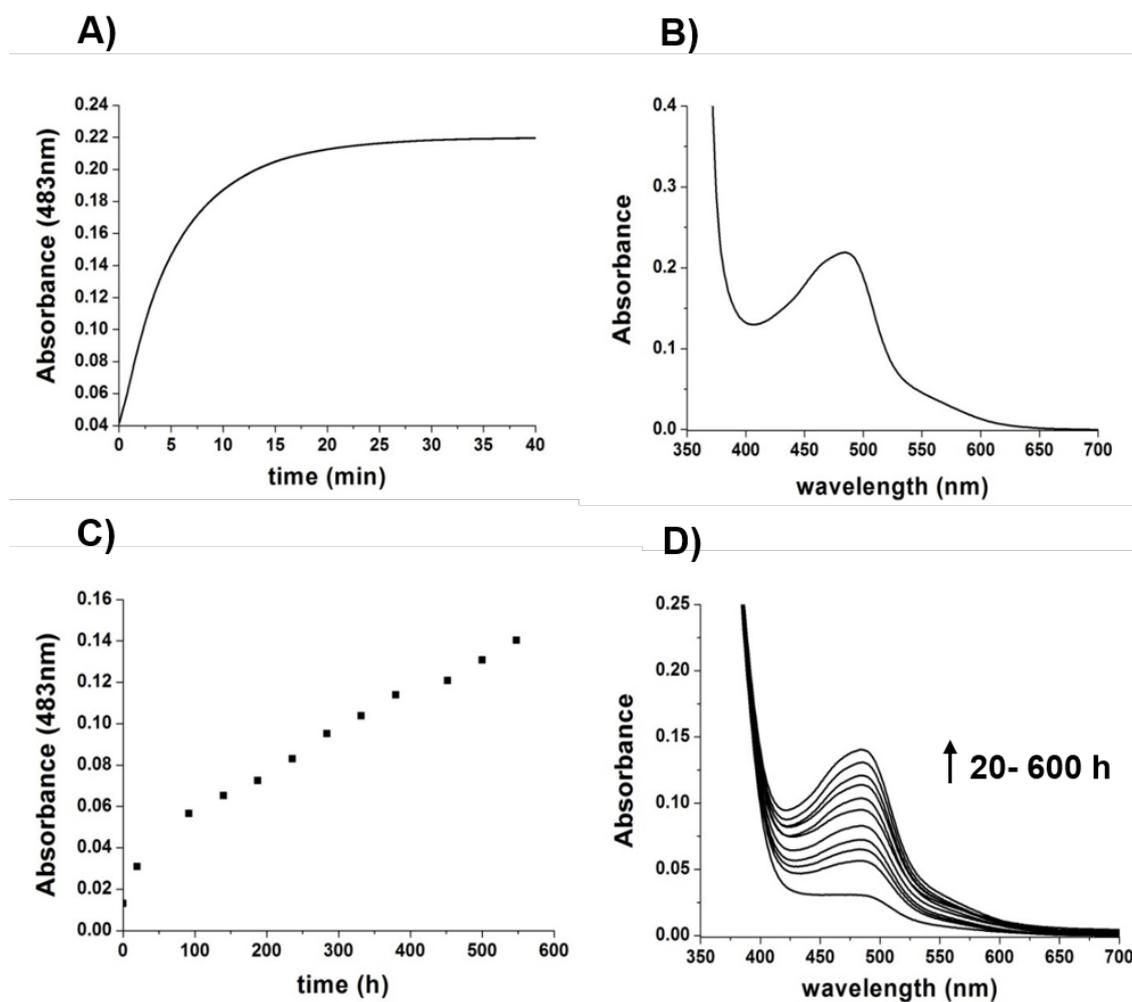
Figure 3.18 Cu(I) removal from Cu(I)-Csp3 by BCS. A) Time-course showing the removal of Cu(I) from $Cu(I)_{18}$-Csp3 (1 μM) by an excess of BCS (2.5 mM) and the formation of the $[Cu(BCS)_2]^{3-}$ complex in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing 6.8 M guanidine (6.3 M final concentration). B) UV-VIS spectrum of the sample from A) at the end of the time-course. C) Time-course showing the removal of Cu(I) from $Cu(I)_{18}$-Csp3 (1 μM) by an excess of BCS (2.5 mM) and the formation of the $[Cu(BCS)_2]^{3-}$ complex in 20 mM Hepes pH 7.5 plus 200 mM NaCl. D) UV-VIS spectra of the sample from C) at the time points shown in graph C).
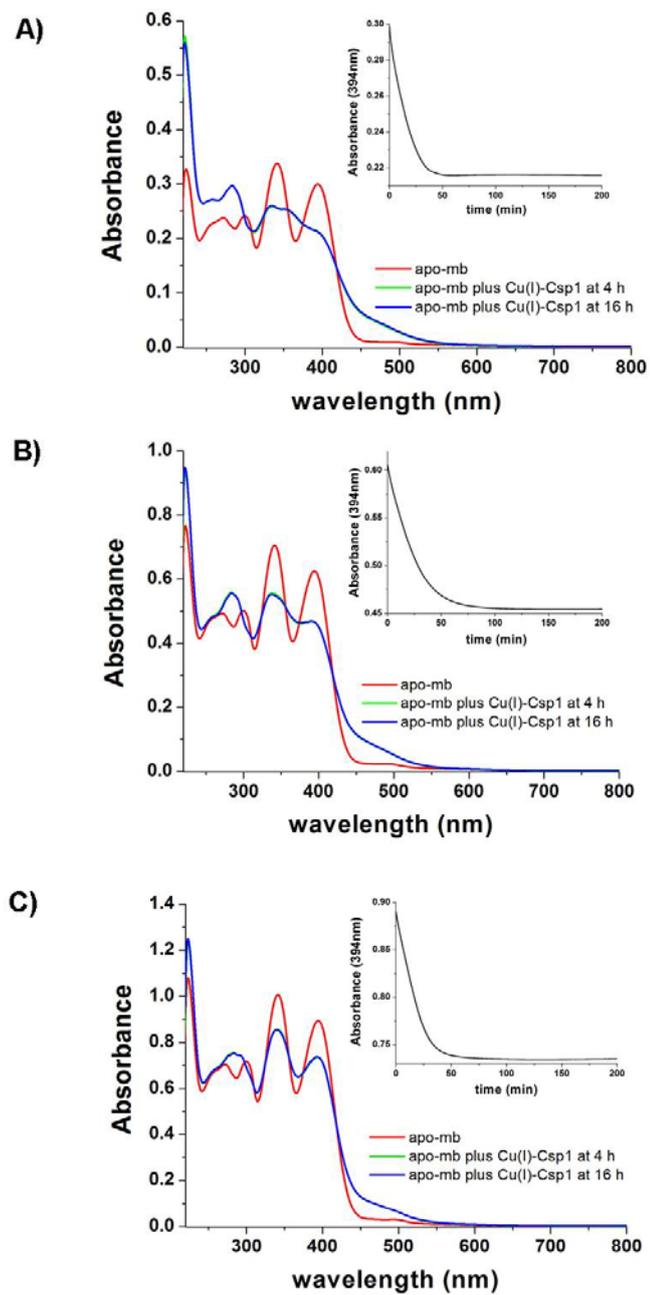
Figure 3.19 Cu(I) transfer between Cu(I)-Csp1 and apo-mb. UV-VIS spectra of apo-mb (red lines) at 15.4 μM (A), 32.3 μM (B) and 46.1 μM (C) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and after the addition of 1.2 μM Cu(I)$_{13}$-Csp1 and incubation of the mixture for 4 (green lines) and 16 (blue lines) h. The inserts show the absorbance at 394 nm as a function of time.
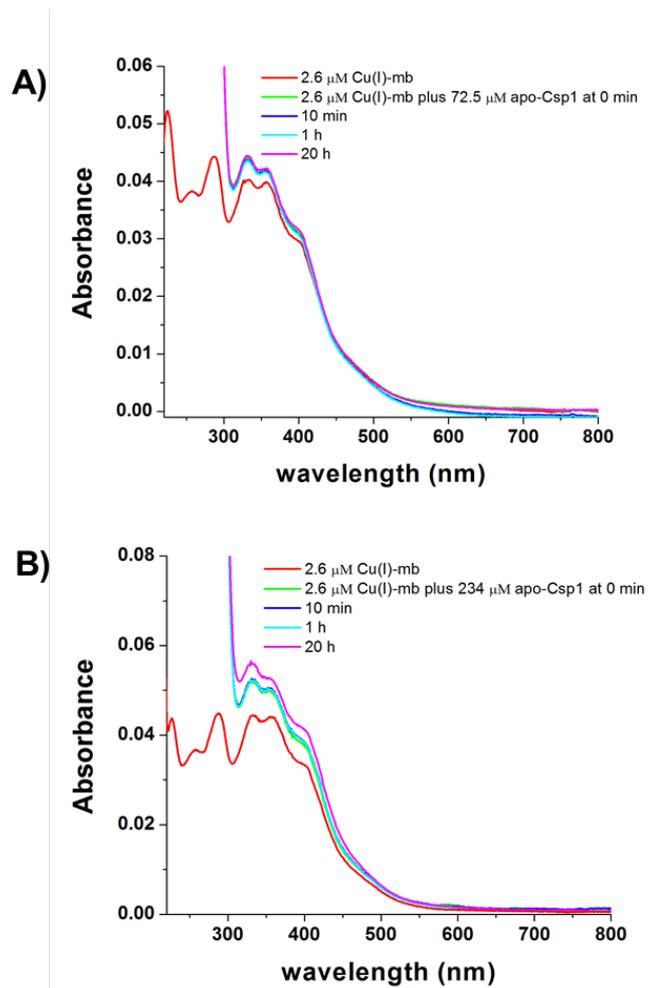
Figure 3.20 Cu(I) transfer between Cu(I)-mb and apo-Csp1. UV-VIS spectra of Cu(I)-mb at 2.6 μM (red line) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and spectra after the addition of 72.5 μM and 234 μM apo-Csp1 respectively. Spectra have been monitored directly after the addition of apo-Csp1 (green line) and after incubation for 10 min (blue line), 1 h (cyan line) and 20 (magenta line) h.
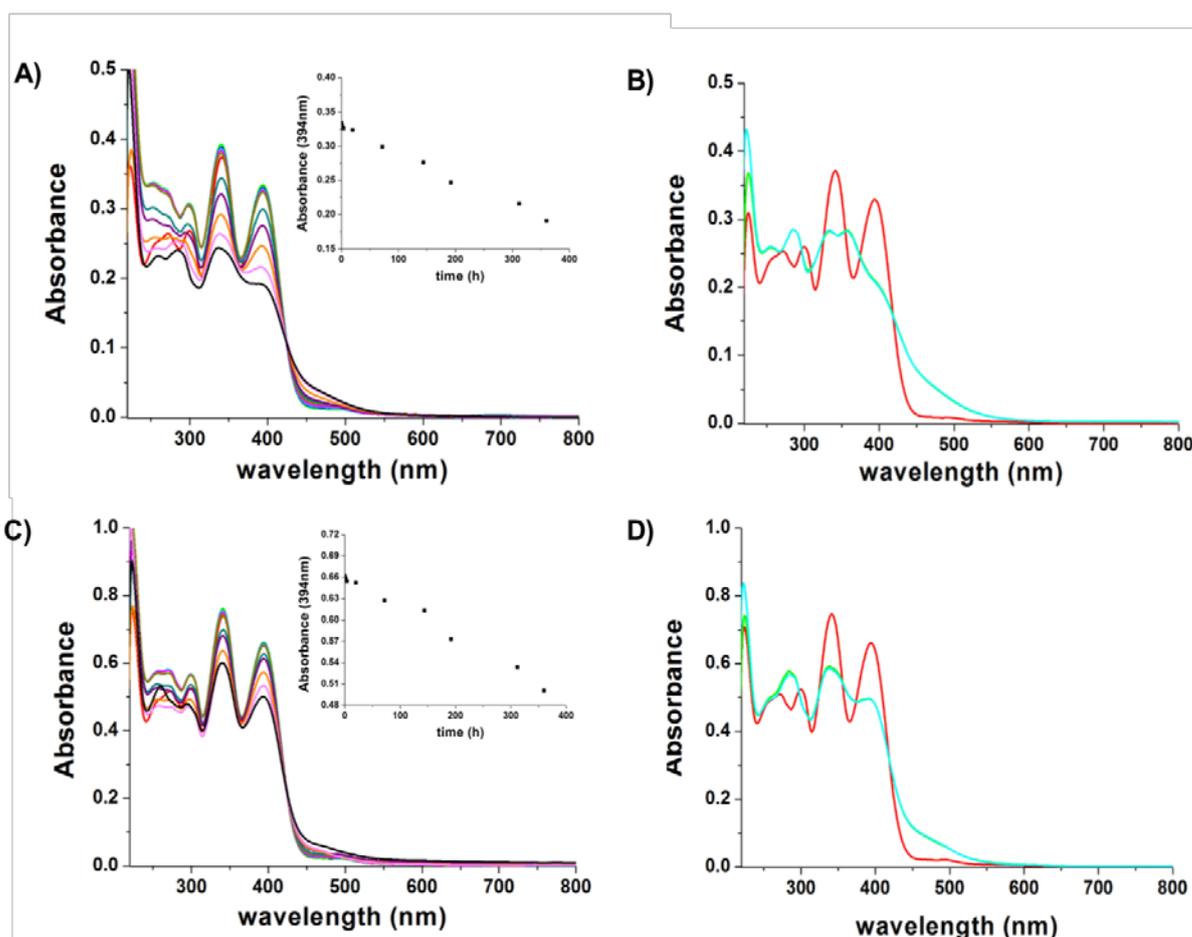
Figure 3.21 Cu(I) transfer between Cu(I)-Csp3 and apo-mb. A) and C) show UV-VIS spectra of apo-mb at 17 μM and 34 μM respectively (red lines) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and spectra directly after the addition of 0.95 μM Cu(I)-Csp3 (green lines) and after incubation of the mixture for 1h (blue lines), 2h (cyan lines), 4h (magenta lines), 20h (dark yellow lines), 3 days (dark cyan lines), 6 days (purple lines), 8 days (orange lines), 13 days (pink lines) and 15 days (black lines). The inserts in A) and C) show the absorbance at 394 nm at various time-points. B) and D) show UV-VIS spectra of apo-mb at 17 μM and 34 μM respectively (red lines) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and spectra directly after the addition of 17 μM Cu(I) (green lines) and after incubation of the mixture for 16h (cyan lines).
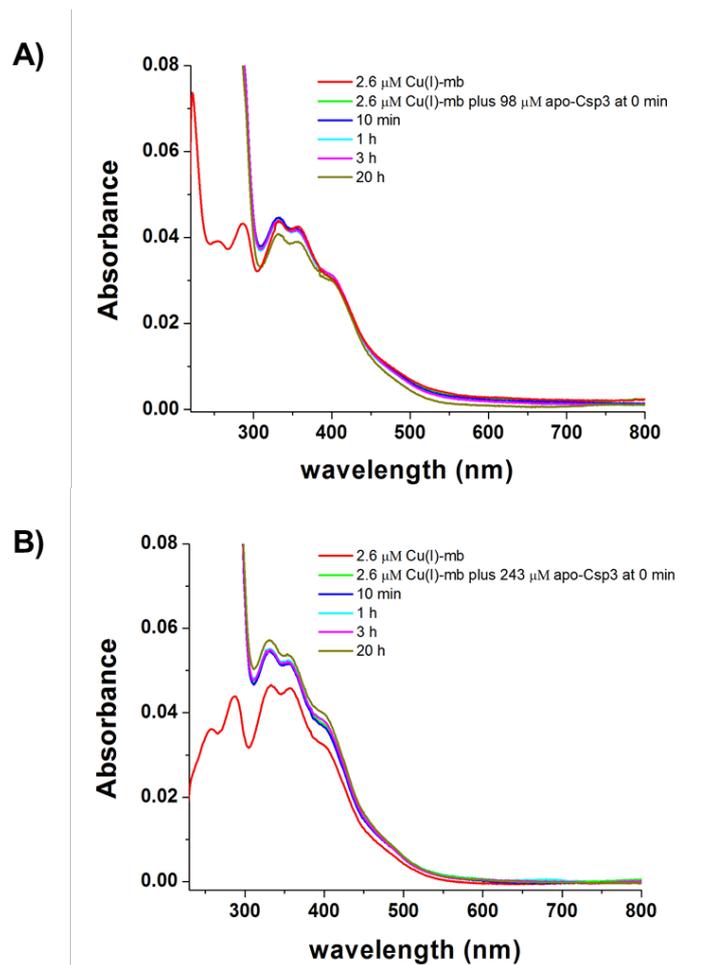
Figure 3.22 Cu(I) transfer between Cu(I)-mb and apo-Csp3. A) and B) show UV-VIS spectra of Cu(I)-mb at 2.6 μM (red lines) in 20 mM Hepes pH 7.5 plus 200 mM NaCl, and spectra after the addition of 98 μM and 243 μM apo-Csp3 respectively. Spectra have been monitored directly after the addition of apo-Csp3 (green lines) and after incubation for 10 min (blue lines), 1h (cyan lines), 3h (magenta lines) and 20h (dark yellow lines).
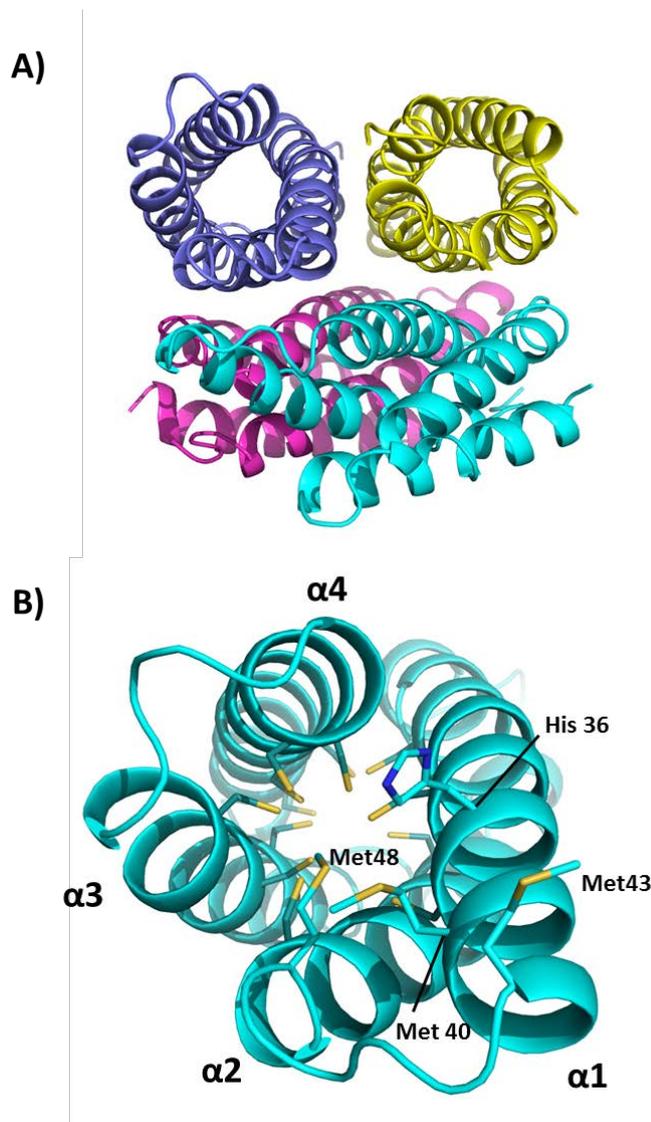
Figure 3.23 Crytal structure of apo-Csp1. A) Csp1 in the asymmetric unit of the crystal structure is a tetramer of 4-helix bundles consisting of two sets of anti-parallel monomers that are rotated at an angle. B) The opening into the core of the apo-Csp1 monomer with the side chains of the Cys, and other significant residues, shown as sticks. All Cys residues face the core of the 4-helix bundle, while His36, Met40, Met43 and Met48 are located at the opening of the 4-helix bundle.
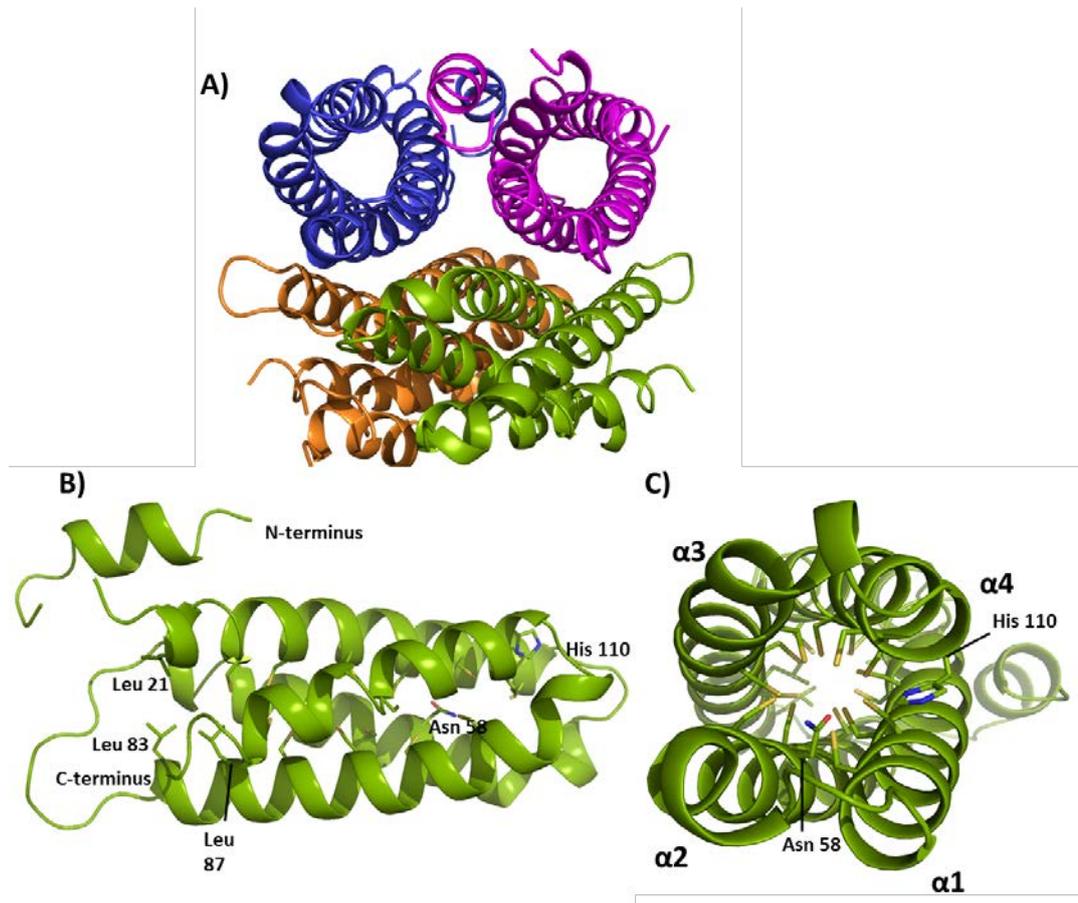
Figure 3.24 Crystal structure of apo-Csp3. A) Csp3 in the asymmetric unit of the crystal structure is a tetramer of 5-helix bundles consisting of two sets of anti-parallel monomers that are rotated at an angle. B) The apo-Csp3 monomer with the side chains of the Cys and other significant residues shown as sticks. The end of the molecule where the N- and C- termini are located is hydrophobic due to Leu residues. C) The apo-Csp3 monomer with all the Cys residues facing towards the core of the 4-helix bundle and His110 facing outwards.
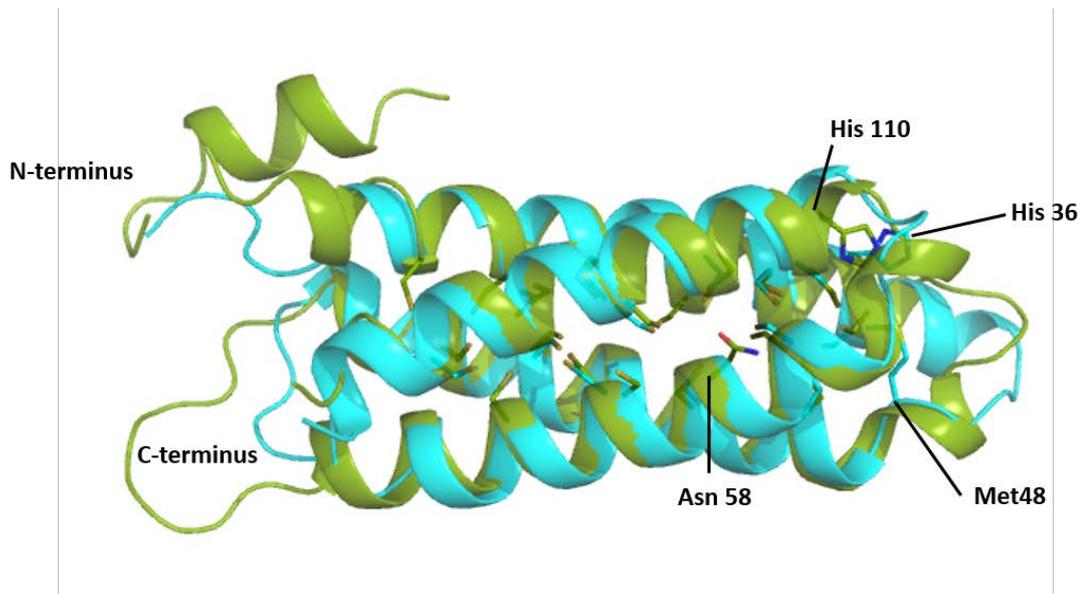
Figure 3.25 Superimposition of apo-Csp3 and apo-Csp1. The crystal structure of apo-Csp3 (green) has been overlayed on that of apo-Csp1 (cyan). The two proteins share approximately 18% sequence identity and the structures have an rmsd of 1.29 Å, referring to the Cα atomic coordinates.
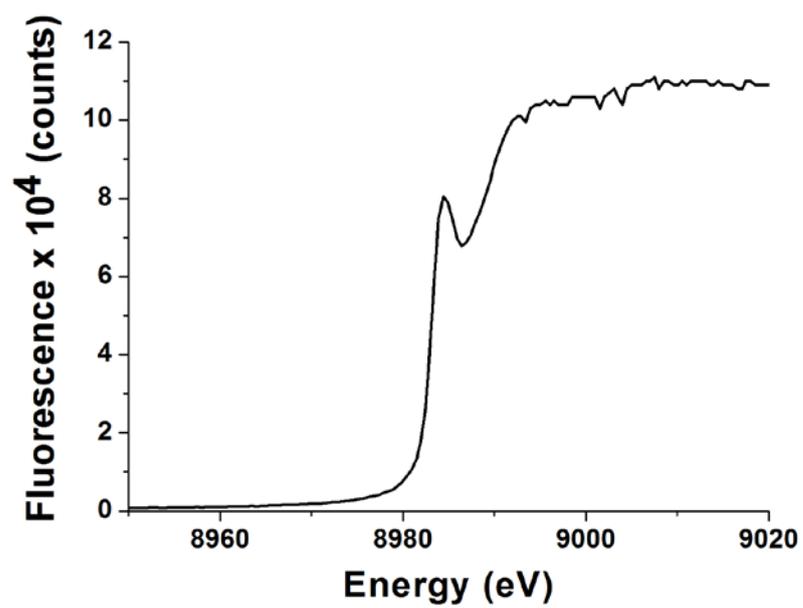
Figure 3.26 XANES for Cu(I)-Csp1. The X-ray absorption near edge spectrum of a fresh Cu(I)-Csp1 crystal was acquired to test the oxidation state of copper in the crystal.
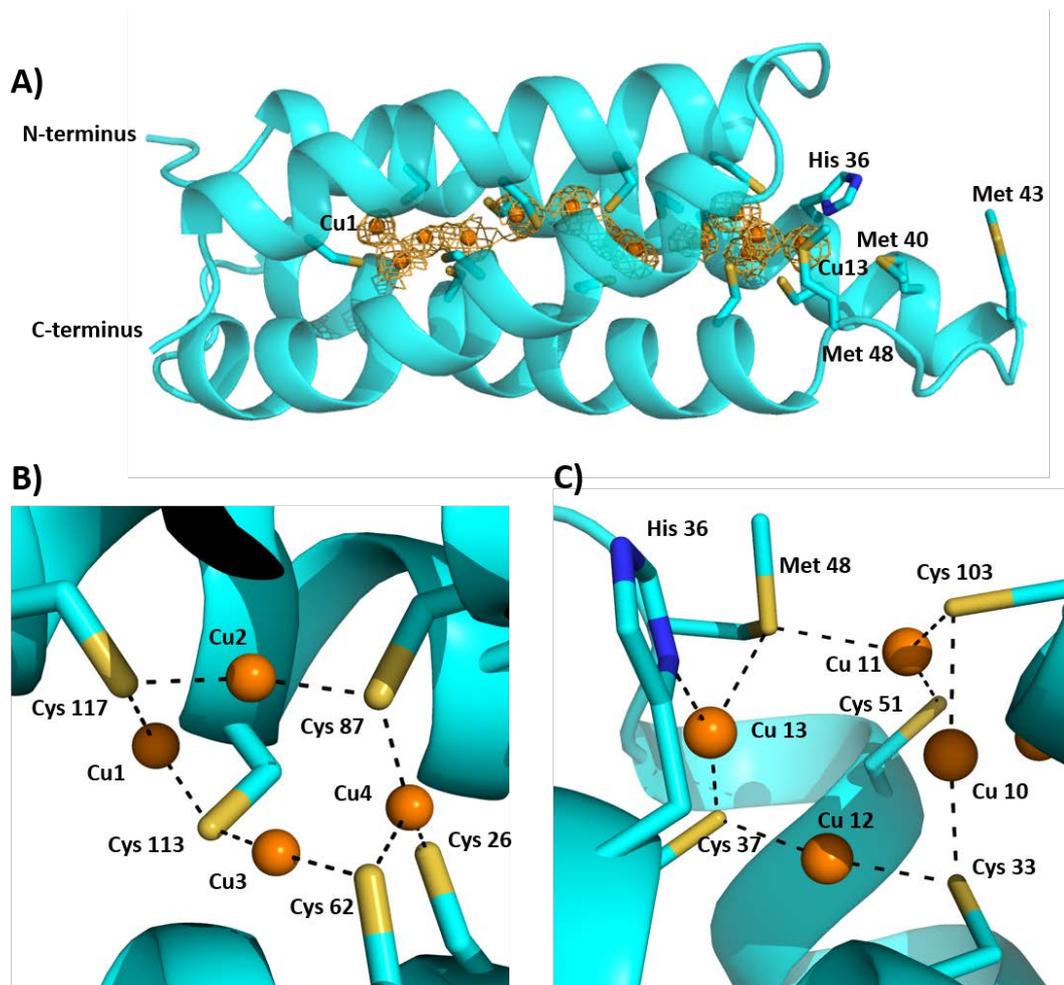
Figure 3.27 Crytal structure of Cu(I)₁₃-Csp1. A) The structure of Cu(I)$_{13}$-Csp1 with Cu(I) ions shown as orange spheres and the side chains of the Cys and other key residues shown as sticks. The anomalous difference density for copper is shown as an orange mesh at 3.0 σ. B) and C) show the coordination of the Cu(I) sites at the two ends of the molecule.
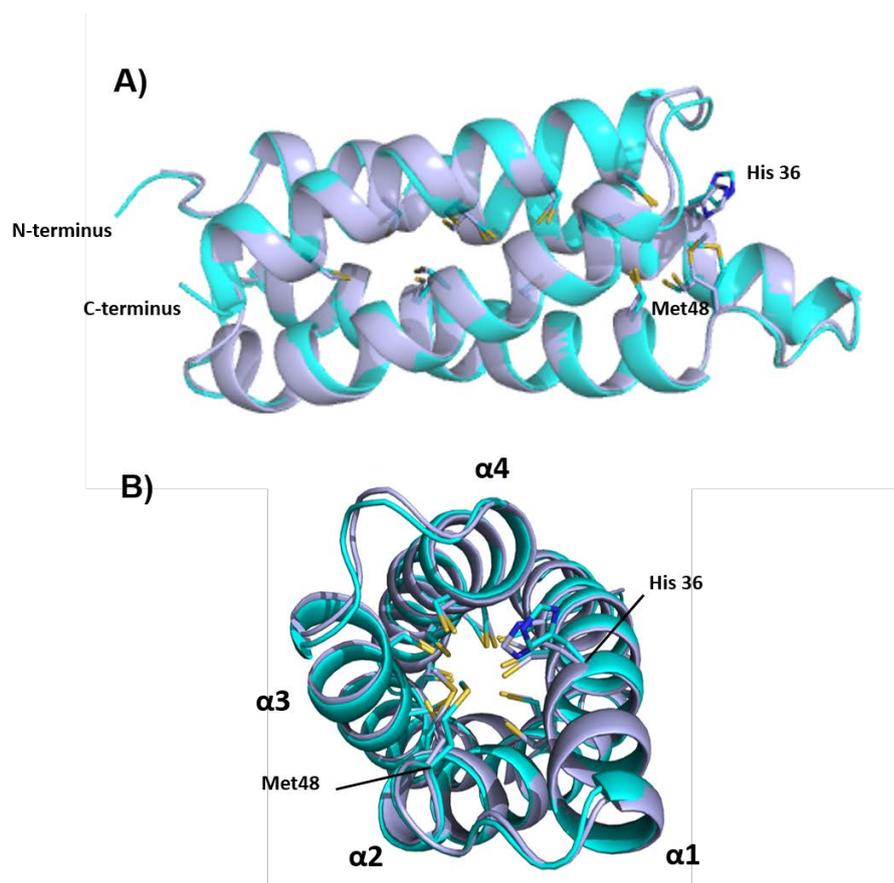
Figure 3.28 Superimposition of Cu(I)$_{13}$-Csp1 and apo-Csp1. The crystal structure of Cu(I)$_{13}$-Csp1 (ice grey, Cu(I) ions not shown) has been overlayed on that of apo-Csp1 (cyan). The rmsd of the two proteins is 0.42 Å, referring to the Cα atomic coordinates.
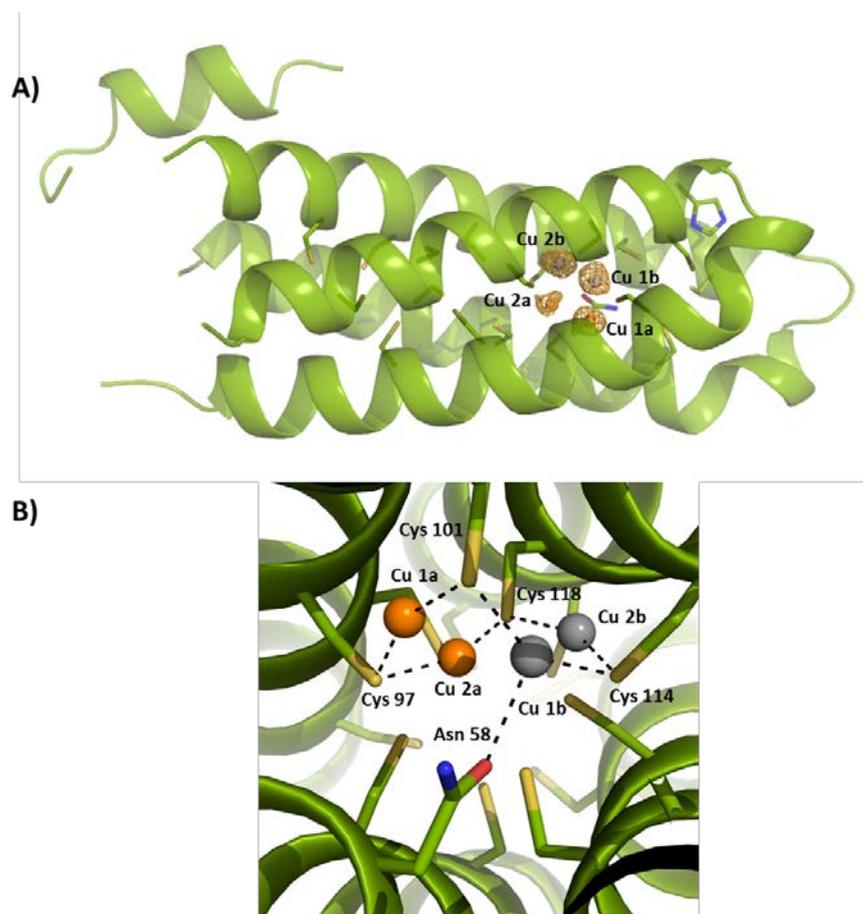
Figure 3.29 Crytal structure of Cu(I)$_2$-Csp3. A) The structure of Cu(I)$_2$-Csp3 contains two Cu(I) ions distributed among four partially occupied sites that are shown in orange (highest occupancy factors) and grey spheres (lowest occupancy factors). The side chains of the Cys and other key residues are shown as sticks and the anomalous difference density for copper is shown as an orange mesh at 3.0 σ. B) The coordination of the four Cu(I) sites in Cu(I)$_2$-Csp3.
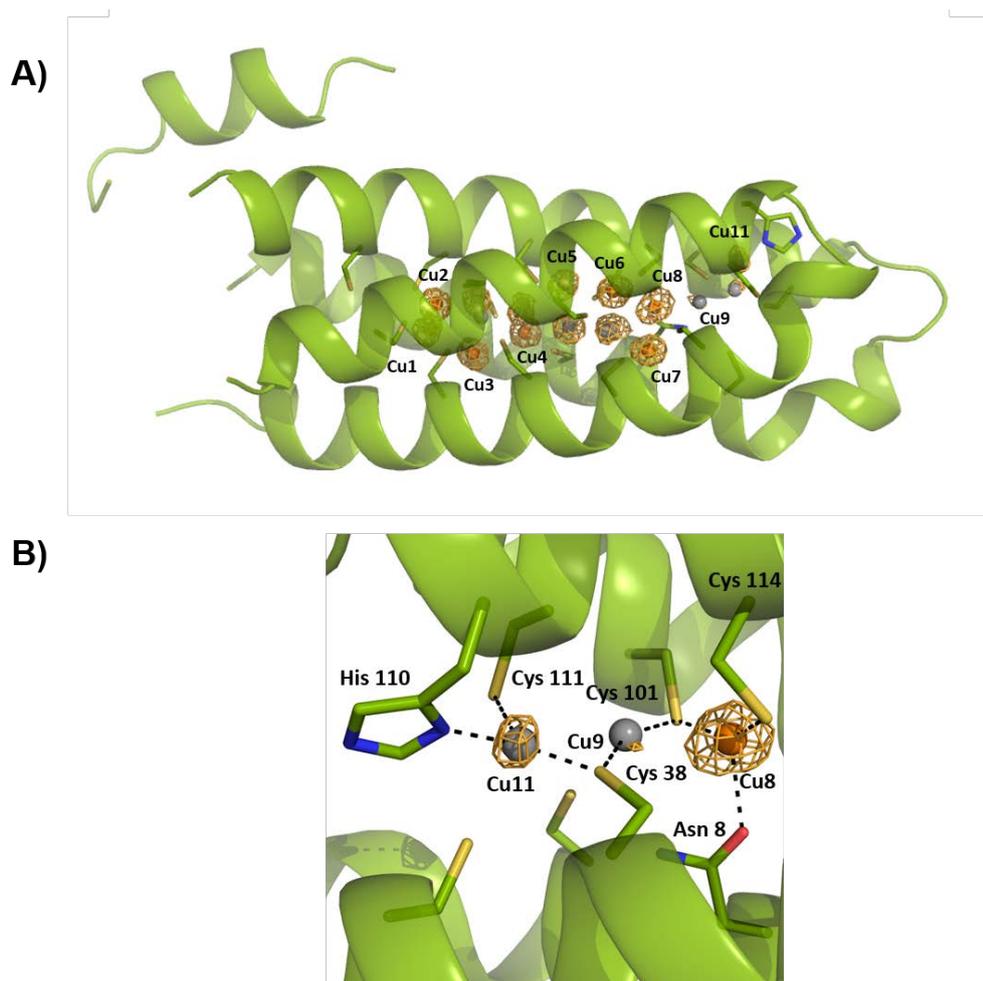
Figure 3.30 Crytal structure of Cu(I)$_8$-Csp3. A) The structure of Cu(I)$_8$-Csp3 contains 8 Cu(I) ions distributed among 18 partially occupied sites that are shown in orange (highest occupancy factors) and grey spheres (lowest occupancy factors) (also see Table 3.9). The side chains of the Cys and other key residues are shown as sticks and the anomalous difference density for copper is shown as an orange mesh at 3.5 σ. B) The coordination of Cu8, as an example of a Cu(I) site with higher occupancy (0.55) and stronger signal, and of Cu9 and Cu11, as examples of lower occupancy sites (0.4 and 0.3 occupancy factors, respectively), in Cu(I)$_8$-Csp3.
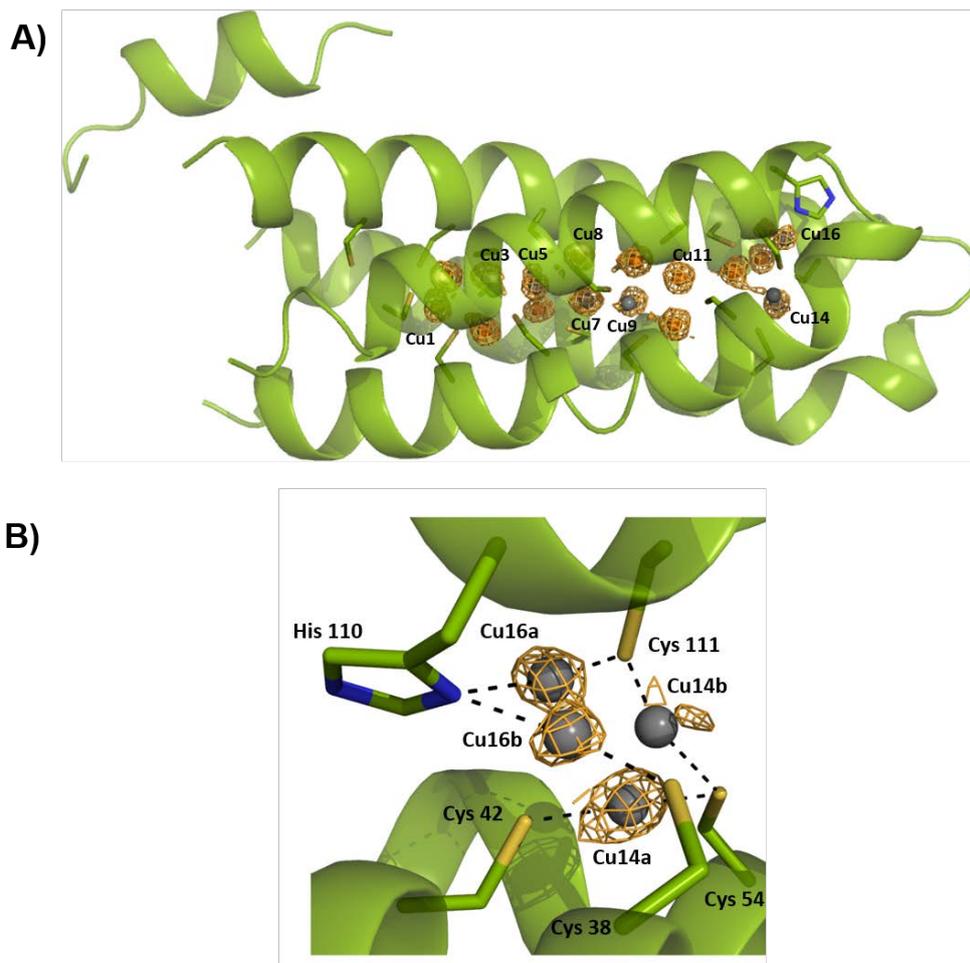
Figure 3.31 Crytal structure of Cu(I)$_{16}$-Csp3. The structure of Cu(I)$_{16}$-Csp3 contains 16 Cu(I) ions, 10 of which are located in fully occupies sites (shown as orange spheres) while 6 are distributed among two partially occupied sites (shown as grey spheres) with occupancy factors shown in Table 3.10. The side chains of the Cys and other key residues are shown as sticks and the anomalous difference density for copper is shown as an orange mesh at 3.5 σ. B) The coordination of Cu16 and Cu14, where Cu(I) ions are distributed between two sites (a and b).
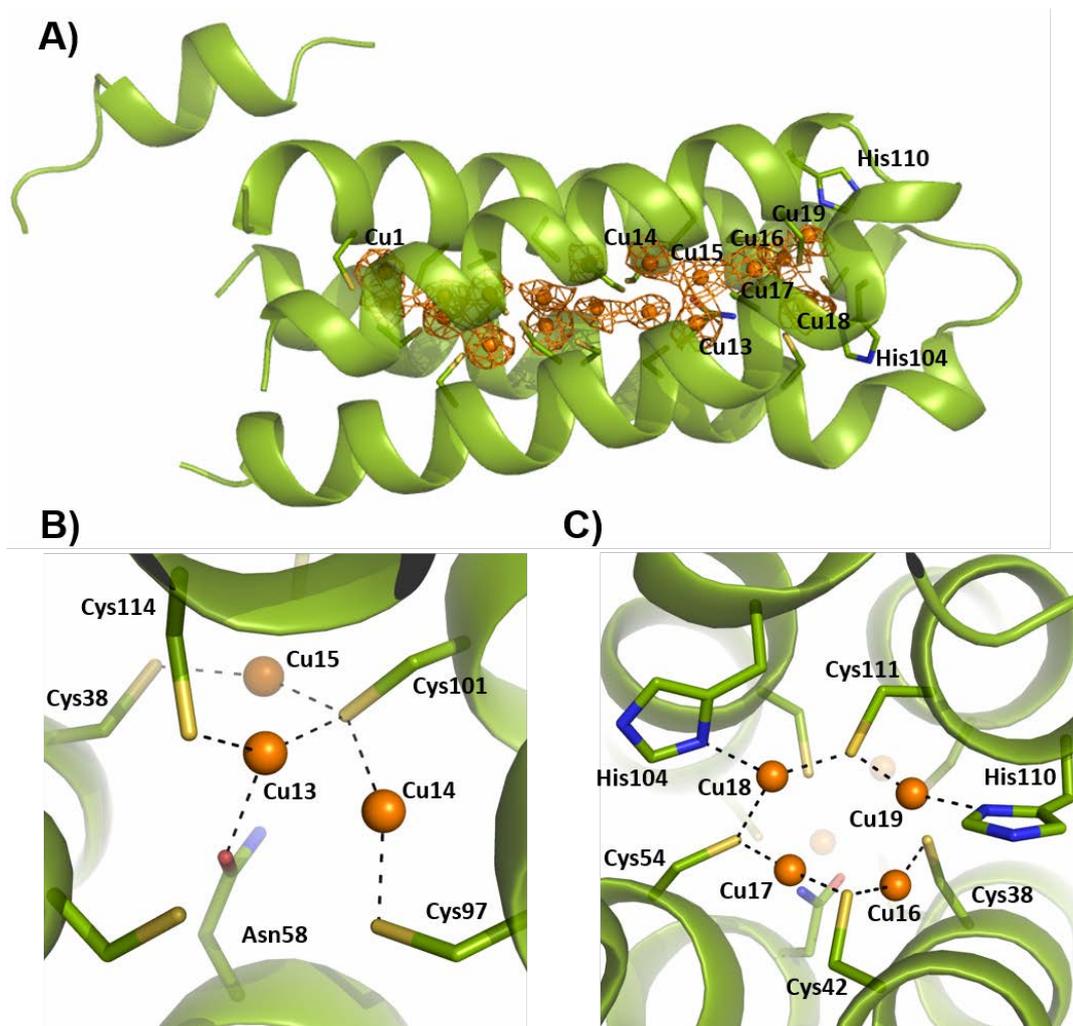
Figure 3.32 Crytal structure of Cu(I)₁₉-Csp3. A) The structure of Cu(I)$_{19}$-Csp3 with Cu(I) ions shown as orange spheres and the side chains of the Cys and other key residues shown as sticks. The anomalous difference density for copper is shown as an orange mesh at 3.0 σ. B) and C) show the coordination of the Cu(I) sites at the opening of the molecule.
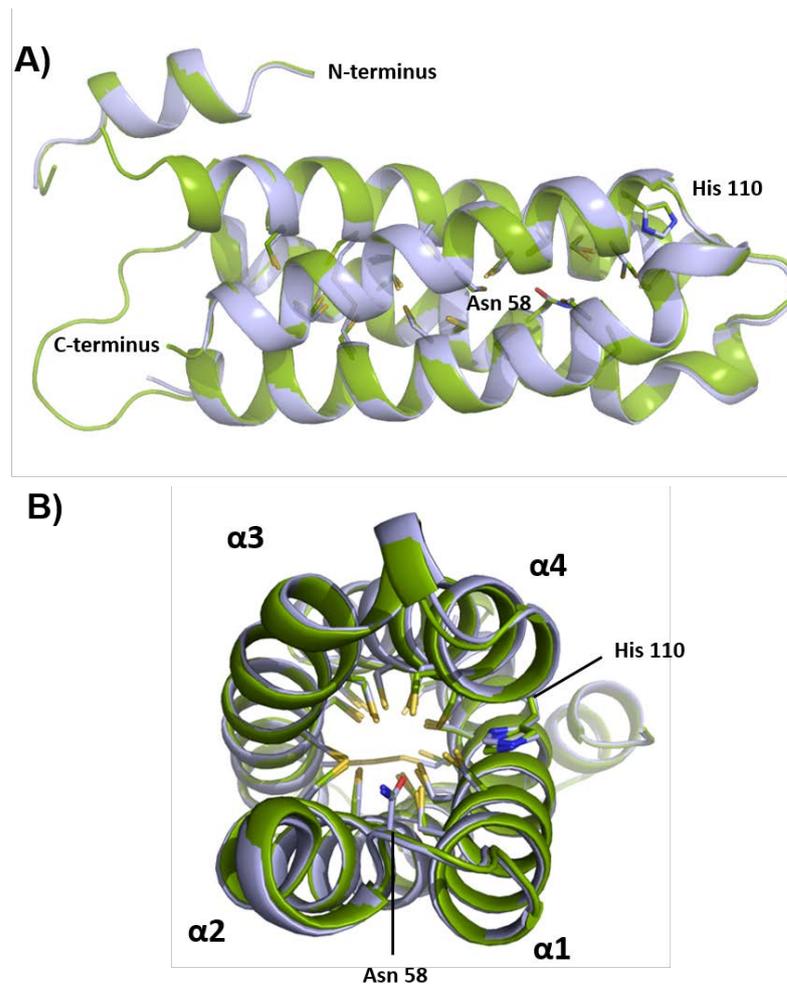
Figure 3.33 Superimposition of Cu(I)$_{19}$-Csp3 and apo-Csp3. The crystal structure of Cu(I)$_{19}$-Csp3 (ice grey, Cu(I) ions not shown) has been overlayed on that of apo-Csp3 (green).The rmsd of the two structures is 0.31 Å, referring to the Cα atomic coordinates.

3.4   Discussion

3.4.1   Overall structure and oligomeric state of Csp1 and Csp3

The secondary structures of Csp1 and Csp3 are mostly composed by α-helices (Table 3.3), forming a 4-helix bundle and a 5-helix bundle in the case of Csp1 and Csp3, respectively, and the secondary structure for both proteins is not altered by Cu(I) binding (Figures 3.3 and 3.4, and also 3.28 and 3.33). This behaviour distinguishes Csp1 and Csp3 from MTs, where metal binding induces the folding of a disordered protein (6). The α-helical content determined by CD is in good agreement with that of the crystal structures of the two proteins. The shielded nature of the Cys residues inside the core of the helical bundle becomes clear from the crystal structures of apo-Csp1 and apo-Csp3 (Figure 3.23 and 3.24) and is consistent with the observation that free thiols can only be observed when the proteins are unfolded, which is why quantification of the two proteins was only possible by DTNB assay in urea.

The analysis of Csp1 by gel-filtration chromatography shows the protein forms a tetramer in solution, in both apo- and Cu(I)-forms (Table 3.4 and Figure 3.7), also consistent with the crystal structure where Csp1 forms a tetramer in the asymmetric unit (Figure 3.23). In the case of Csp3, the elution volume from the gel-filtration column corresponds to an apparent trimer (Table 3.5). The oligomeric state of apo-Csp3 is dependent on protein concentration and at 5 μM the protein appears to be mostly in a dimeric form. The two oligomeric forms (apparent trimer/dimer) elute as separate peaks in the case of apo-Csp3, whereas for Cu(I)-Csp3 the exchange equilibrium between the two forms appears to be faster and a single peak is observed (Table 3.5 and Figure 3.8). In the crystal structure however, Csp3 is a tetramer in the asymmetric unit (Figure 3.24) and the larger discrepancy, compared to Csp1, of the apparent molecular weight from a tetramer could possibly be attributed to the non-globular shape of the Csp3, which is longer than the Csp1 molecule, compared to the protein standards used for calibration of the column (52).

3.4.2   Cu(I) binding by Csp1 and Csp3

Cu(I) titrations showed Csp1 and Csp3 bind significant amounts of Cu(I), at ~13 and ~18 molar equivalents per monomer, respectively (Figures 3.9, 3.10 and 3.10). The Cu(I) binding stoichiometry of the two proteins was also tested in the presence of the Cu(I) binding chelator BCA and the results were in good agreement with the Cu(I)

titration experiments (Figures 3.9 C, 3.14 and 3.16). Moreover, the crystal structures of the fully loaded forms of the two proteins are also in accordance with the stoichiometry (Figures 3.27 and 3.32 A). Based on the quaternary structure of Csp1 and Csp3 being a tetramer, the Cu(I) binding stoichiometry of 13 and 18 molar equivalents per monomer, respectively, translates to binding of 52 and 72 Cu(I) ions per tetramer for Csp1 and Csp3, respectively. As shown by the crystal structures, the Cu(I) ions are stored inside the core of the 4-helix bundle, for Csp1 and Csp3, and are coordinated mostly by Cys residues (Figures 3.27 and 3.32 A). A known family of proteins, which are made up of 4-helix bundles and store a metal ion are ferritins (3). However, in this case Fe(III) is stored in a FeO(OH) mineral core inside a shell formed by 24 identical units, each one of which is a 4-helix bundle, rather than in the core of the 4-helix bundle itself (4). Cu(I) binding inside a 4-helix bundle has not been reported to date for a protein and renders the Csp1 and Csp3 proteins unique in that respect, while the high Cu(I) binding capacity of the proteins suggests a role in Cu(I) storage.

In contrast to naturally occurring proteins, binding of four Cu(I) ions in a cluster inside the core of an 4-helix bundle has been previously reported for synthetic peptides (37, 53). Interestingly, Cu(I) addition to a synthetic peptide containing Cys residues, arranged in a Cys-X-X-Cys motif, induces folding of four of the originally random coil peptides into a 4-helix bundle. Cu(I) is bound as a tetranuclear Cu_4S_4 cluster that is coordinated by one Cys residue from each of the four synthetic peptides that constitute the 4-helix bundle and the complex is luminescent at room temperature like MTs (37, 53). Cu(I) ions in Csp1, on the other hand, are coordinated in many cases by both Cys residues of Cys-X-X-X-Cys motifs, which is likely a result of the longer length of this motif, compared to a Cys-X-X-Cys motif, that allows both Cys from the same motif to participate in Cu(I) coordination in the same binding site. Cu(I) binding to Csp1, however, does not give rise to emission, in contrast to the Cu_4S_4 cluster of the synthetic peptide mentioned. This is not the case with Csp3, in which Cu(I) binding is followed by luminescence at 550 nm (Figures 3.10 C and 3.11 C). Notably, Cu(I)-Csp3 luminescence reaches a maximum when the protein is half-loaded (8-10 molar equivalents of Cu(I)) and subsequently the emission decreases until the protein is fully loaded at which point the emission is totally lost. Quenching of emission has been associated with solvent exposure of the Cu(I) clusters leading to energy degradation of the excited state (28). Loss of emission due to solvent exposure has also been previously

reported when excess metal, beyond saturation, was added to the protein causing local structural reorganisation that results in exposing the metal centres to the solvent (53-55) and examples in the literature include a metal-responsive transcription factor (55) and a copper chaperone (54) that bind Cu(I).

The crystal structures of Cu(I)-Csp3 may provide some insight into the decrease in emission upon Cu(I) addition between semi- and fully-loaded protein is obtained (Figure 3.10 C and 3.11 C). The Cu(I)$_2$-Csp3 structure revealed that the first two Cu(I) ions added to the protein are bound in the centre of the molecule (Figure 3.29). Subsequent Cu(I) addition to Csp3 resulted in the Cu(I)$_8$-Csp3 structure where 18 partially occupied Cu(I) sites are present, which may indicate high mobility of the Cu(I) ions. The location of the sites with the highest occupancy factors indicates the eight Cu(I) ions are distributed along the core of the molecule, without occupying all the available Cu(I) sites at the hydrophobic end of Csp3 (where the N- and C- termini are located) (Figure 3.30 A). It could be that, up to the addition of < 8 Cu(I) ions into Csp3, Cu(I) binding is achieved in an ordered manner, possibly forming solvent-shielded Cu(I) centres at the first half of the molecule, where Cu5-Cu8 are in Figure 3.30. However, accumulation of more Cu(I) ions may induce reorganisation of the Cu(I) ions inside the protein core that now need to occupy the Cu(I) sites at the hydrophobic end of the molecule in order for the protein to accommodate more Cu(I). This structural reorganisation at the opening of the 5-helix bundle could be responsible for rendering the Cu(I) sites of Cu(I)$_8$-Csp3 more accessible by solvent and, consequently, for the observed decrease in emission until full loading. It is unclear why emission was not observed at all upon Cu(I)-binding by Csp1.

3.4.3   Cu(I) release from Csp1 and Csp3

Csp1 binds tightly at least 10 Cu(I) ions and Csp3 has an average Cu(I) binding affinity at the order of 10^{17} M^{-1} (Figure 3.14 and 3.16) with a Hill coefficient greater than 1 that possibly indicates a cooperative mechanism in Cu(I) binding by Csp3. The Cu(I) affinity of Csp3 is lower than that of the metallochaperone CopZ (10^{18} M^{-1}), which delivers copper to the copper transporter CopA, (42, 56) and the Cu(I) sensor CueR (10^{20} M^{-1}) (57), which are encoded in the genome of *M. trichosporium* OB3b (58). Moreover, Csp1 and Csp3 show drastically different rates of Cu(I) removal by BCS. Csp1 readily releases Cu(I) in the presence of BCS and denaturing conditions do not have a drastic effect on the rate of Cu(I) removal (Figure 3.17). In contrast, effective

Cu(I) removal from Cu(I)-Csp3 by BCS takes place only when the protein is unfolded in guanidine hydrochloride (Figure 3.18). This difference in the kinetics of Cu(I) release is also observed in the Cu(I) transfer experiments with mb, the naturally produced Cu(I) binding chelator with high Cu(I) affinity (38, 39) in *M. trichosporium* OB3b. Mb removes Cu(I) almost stoichiometrically from $Cu(I)_{13}$-Csp1 within 1 h (Figure 3.19), whereas the same reaction in the case of the cytosolic $Cu(I)_{18}$-Csp3 does not appear to be physiologically relevant as the reaction takes weeks to reach completion (Figure 3.21). This observation is also consistent with the effective Cu(I) removal from $Cu(I)_{18}$-Csp3 only when the protein is unfolded in the presence of guanidine hydrochloride.

It is not understood how copper is released from Cu(I)-mb and suggested mechanisms include oxidation of Cu(I) to Cu(II) or degradation of the peptide by proteases (39). Csp1 and Csp3 do not appear to remove Cu(I) from Cu(I)-mb (Figure 3.20 and 3.22), however it is possible that apo-mb, which removes stoichiometrically Cu(I) from Csp1 and Csp3, yet with strikingly different rates, may extract Cu(I) from these proteins for the primary copper target of *M. trichosporium* OB3b, pMMO (39). Mb has been shown to be localised in the cytoplasm of *M. trichosporium* OB3b (40) and is thought to mediate Cu(I) transport to pMMO. Csp1 has a predicted Tat-signal peptide that indicates the protein folds in the cytosol before being exported (1). Folding of Csp1 in the reducing environment of the cytosol is also consistent with preventing the formation of disulfide bonds between the Cys residues (59). It is also possible that Csp1 acquires Cu(I) in the cytosol before being exported and, therefore, transports the metal to the destination where it is exported, possibly the intra-cytoplasmic membranes that form at high copper: biomass ratios and house pMMO. The Csp2 protein homologue that was identified through bioinformatics also contains a predicted Tat-leader peptide for export from the cytosol (1). It is unclear whether the intra-cytoplasmic membranes where pMMO is located, form separate subcellular compartments or are invaginations of the cytoplasmic membrane (60), and it is possible that Csp1 and Csp2 are exported from the cytosol to different destinations (the intra-cytoplasmic membranes and the periplasm for instance). The export of Csp1 may serve copper storage for the needs of pMMO, while Csp2 may be exported to the periplasmic space, or vice versa. Csp3 on the other hand is cytosolic, as it does not contain a Tat-leader peptide, and acts as a store of copper there. The presence of a copper storage protein in the cytosol indicates it acts as either a defence mechanism against copper-induced toxicity, or more likely stores copper for the

needs of an unknown cytoplasmic proteins. The visualisation of a surprisingly large number of soluble copper pools in cell profiles of *M. trichosporium* OB3b may also imply that some copper proteins are present in the cytosol. This possibility directly challenges the current perception of copper being exported from the cytosol and the absence of cytosolic need for copper.

3.5 References

1. Palmer T., Berks B. C., Moving folded proteins across the bacterial cell membrane, Microbiology, **2003**, 149, 547-556

2. Romero-Isart N., Vasak M., Advances in the structure and chemistry of metallothioneins, J. Inorg. Biochem., **2002**, 88, 388-396

3. Yang X., Le Brun N. E., Thomson A. J., Moore G. R., Chasteen N. D., The iron oxidation and hydrolysis chemistry of Escherichia coli bacterioferritin, Biochemistry, **2000**, 39, 4915-4923

4. Baaghil S., Lewin A., Moore G. R., Le Brun N. E., Core formation in Escherichia coli bacterioferritin requires a functional ferroxidase center, Biochemistry, **2003**, 42, 14047-14056

5. Harrison P. M., The structure and function of ferritin, **1986**, Biochem. Educ., 14, 154-162

6. Sutherland D. E. K., Stillman M.J., The 'magic' numbers of metallothionein, Metallomics, **2011**, 3, 444-463

7. Davis S. R., Cousins R. J., Metallothionein expression in animals: a physiological perspective on function, J. Nutr., **2000**, 130, 1085-1088

8. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

9. Klaassen C. D., Choudhuri S., McKim J. M. Jr, Lehman-McKeeman L. D., Kershaw W. C., In vitro and in vivo studies on the degradation of metallothionein, Environ. Health Persp., **1994,** 102, 141-146

10. Blindauer C. A., Sadler P.J., How to hide zinc in a small protein, Acc. Chem. Res., **2005**, 38, 62-69

11. Nielson K. B., Winge D. R., Preferential binding of copper to the β domain of metallothionein, J. Biol. Chem., **1984,** 259, 4941-4946

12. Otvos J. D., Armitage I. M., Structure of the metal clusters in rabbit liver metallothionein, Proc. Natl. Acad. Sci., **1980**, 77, 7094-7098

13. Green A. R., Presta A., Gasyna Z., Stillman M. J., Luminescent probe of copper-thiolate cluster formation with mammalian metallothionein, Inorg. Chem., **1994**, 33, 4159-4168

14. Calderone V., Dolderer B., Hartmann H. J., Echner H., Luchinat C., Del Bianco C., Mangani S., Weser U., The crystal structure of yeast copper thionein: the solution

of a long lasting enigma, Proc. Natl. Acad. Sci., **2004**, 102, 51-56

15. Zhang L., Pickering I. J., Winge D. R., George G. N., X-ray absorption spectroscopy of cuprous-thiolate clusters in Saccharomyces cerevisiae metallothionein, Chem. Biodivers., **2008**, 5: 2042-2049

16. Byrd J., Berger R. M., McMillin D. R., Wright C. F., Hamer D., Winge D. R., Characterisation of the copper-thiolate cluster in yeast metallothionein and two truncated mutants, J. Biol. Chem., **1988**, 263, 6688-6694

17. Blindauer C. A., Bacterial metallothioneins: past, present, and questions for the future, J. Biol. Inorg. Chem., **2011**, 16, 1011-1024

18. Morby A. P., Turner J. S., Huckle W., Robinson N.J., SmtB is a metal-dependent repressor of the cyanobacterial metallothionein gene smtA: identification of a Zn inhibited DNA-protein complex, Nuc. Acids Res., **1993**, 21, 921-925

19. Blindauer C. A., Harrison M. D., Parkinson J. A., Robinson A. K., Cavet J. S., Robinson N. J., Sadler P. J., A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity, Proc. Natl. Acad. Sci., **2001**, 98, 9593-9598

20. Blindauer C. A., Razi M. T., Campopiano D. J., Sadler P.J., Histidine ligands in bacterial metallothionein enhance cluster stability, J. Biol. Inorg. Chem., **2007**, 12, 393-405

21. Blindauer C. A., Polfer N. C., Keiper S. E., Harrison M. D., Robinson N. J., Langridge-Smith P. R. R., Sadler P.J., Inert site in a protein zinc cluster: isotope exchange by high resolution mass spectrometry, J. Am. Chem. Soc., **2003**, 125, 3226-3227

22. Blindauer C. A., Harrison M. D., Robinson A. K., Parkinson J. A., Bowness P. W., Sadler P. J., Robinson N.J., Multiple bacteria encode metallothioneins and SmtA-like zinc fingers, Mol. Microbiol., **2002**, 45, 1421-1432

23. Chan J., Huang Z., Merrifield M. E., Salgado M. T., Stillman M. J., Studies of metal binding reactions in metallothioneins by spectroscopic, molecular biology, and molecular modelling techniques, Coord. Chem. Rev., **2002**, 233-234, 319-339

24. Stillman M. J., Presta A., Gui Z., Jiang D. T., Spectroscopic studies of copper, silver and gold-metallothioneins, Met. Based Drugs, **1994**, 1, 375-394

25. Badarau A., Firbank S. J., McCarthy A. A., Banfield M. J., Dennison C., Visualising the metal-binding versatility of copper trafficking sites, Biochemistry, **2010**, 49, 7798-7810

26. Poutney D. L., Schauwecker I., Zarn J., Vašák M., Formation of mammalian Cu_8-metallothionein in vitro: Evidence for the existence of two Cu(I)$_4$-thiolate clusters, Biochemistry, **1994**, 33, 9699-9705

27. Cobine P. A., George G. N., Jones C. E., Wickramasinghe W. A., Solioz M., Dameron C. T., Copper transfer from the Cu(I) chaperone, CopZ, to the repressor, Zn(II)CopY: metal coordination environments and protein interactions, Biochemistry, **2002**, 41, 5822–5829

28. Lytle F. E., Solution luminescence of metal complexes, Appl. Spectrosc., **1970**, 24, 319-326

29. Kiyokawa T., Kanaori K., Tajima K., Koike M., Mizuno T., Oku J-I., Tanaka T., Binding of Cu(II) or Zn(II) in a de novo designed triple-stranded α-helical coiled-coil toward a prototype for a metalloenzyme, J. Peptide Res., **2004**, 63, 347-353

30. Lupas A. N., Gruber M., The structure of α-helical coiled coils, Adv. Protein Chem., **2005**, 70, 37-78

31. Grzyb J., Xu F., Weiner L., Reijerse E., Lubitz W., Nanda V., Noy D., De novo design of a non-natural fold for an iron-sulfur protein: alpha-helical coiled-coil with a four-iron four-sulfur cluster binding site in its central core, Biochim. Biophys. Acta, **2010**, 1797, 406-413

32. Robertson D. E., Farid R. S., Moser C. C., Urbauer J. L., Mulholland S. E., Pidikiti R., Lear J. D., Wand J., DeGrado W., Dutton P. L., Design and synthesis of multi-haem proteins, Nature, **1994**, 368, 425-432

33. Gibney B. R., Mulholland S. E., Rabanal F., Dutton P. L., Ferredoxin and ferredoxin- heme maquettes, Proc. Natl. Acad. Sci., **1996**, 93, 15041-15046

34. Jones A. K., Lichtenstein B. R., Dutta A., Gordon G., Dutton P. L., Synthetic hydrogenases: incorporation of an iron carbonyl thiolate into a designed peptide, J. Am. Chem. Soc., **2007**, 129, 14844-14845

35. Schnepf R., Hörth P., Bill E., Wieghardt K., Hildebrandt P., Haehnel W., De novo design and characterization of copper centers in synthetic four-helix-bundle proteins, J. Am. Chem. Soc., **2001**, 123, 2186-2195

36. Schnepf R., Haehnel W., Wieghardt K., Hildebrandt P., Spectroscopic identification of different types of copper centers generated in synthetic four-helix bundle proteins, J. Am. Chem. Soc., **2004**, 126, 14389-14399

37. Kharenko O. A., Kennedy D. C., Demeler B., Maroney M. J., Ogawa M. Y., Cu(I) luminescence from the tetranuclear Cu_4S_4 cofactor of a synthetic 4-helix bundle, J.

Am. Chem. Soc., **2005**, 127, 7678-7679

38. El Ghazouani A., Baslé A., Firbank S. J., Knapp C. W., Gray J., Graham D. W., Dennison C., Copper-binding properties and structures of methanobactins from Methylosinus trichosporium OB3b, Inorg. Chem., **2011**, 50, 1378-1391

39. El Ghazouani A., Baslé A., Gray J., Graham D. W., Firbank S. J., Dennison C., Variations in methanobactin structure influence copper utilization by methane-oxidising bacteria, Proc. Natl. Acad. Sci., **2012**, 109, 8400-8404

40. Balasubramanian R., Kenney G. E., Rosenzweig A. C., Dual pathways for copper uptake by methanotrophic bacteria, J. Biol. Chem., **2011**, 286, 37313-37319

41. Riddles P. W., Blakeley R. L., Zerner B., Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid) - a re-examination, Anal. Biochem., **1979**, 94, 75–81

42. Badarau A., Dennison C., Thermodynamics of copper and zinc distribution in the cyanobacterium Synechocystis PCC 6803, Proc. Natl. Acad. Sci., **2011**, 108, 13007-13012

43. Allen S., Badarau A., Dennison C., The influence of protein folding on the copper affinities of trafficking and target sites, Dalton Trans., **2013**, 42, 3233–3239

44. Kelly S. M., Jess T. J., Price N. C., How to study proteins by circular dichroism, Biochim. Biophys. Acta, **2005**, 1751, 119-139

45. Barrow C. J., Yasuda A., Kenny P. T. M., Zagorski M. G., Solution conformations and aggregational properties of synthetic amyloid β-peptides of Alzheimer's disease- Analysis of Circular Dichroism Spectra, J. Mol. Biol., **1992**, 225, 1075-1093

46. Frishman D., Argos P., Knowledge-based protein secondary structure assignment, Proteins, **1995**, 23, 566-579

47. Allen S., Badarau A., Dennison C., Cu(I) affinities of the domain 1 and 3 sites in the human metallochaperone for Cu,Zn-superoxide dismutase, Biochemistry, **2012**, 51, 1439–1448

48. Bagchi P., Morgan M. T., Basca J., Fahrni C. J., Robust affinity standards for Cu(I) biochemistry, J. Am. Chem. Soc., **2013**, 135, 18549-18559

49. Banci L., Bertini I., Ciofi-Baffoni S., Kozyreva T., Zovo K., Palumaa P., Affinity gradients drive copper to cellular destinations, Nature, **2010**, 465, 645-648

50. Kau L. S., Spira-Solomon D. J., Penner-Han J. E., Hodgson K. O., Solomon E. I., X-ray absorption edge determination of the oxidation state and coordination number of copper: applications to the type 3 site in Rhus vernicifera laccase and its

reaction with oxygen, J. Am. Chem. Soc., **1987**, 109, 6433–6442

51. Rupp B., Biomolecular crystallography: principles, practice, and application to structural biology, New York: Garland science, **2010**

52. Feng Y. Q., Sligar S. G., Effect of heme binding on the structure and stability of Escherichia coli apo-cytochrome b562, Biochemistry, **1991**, 30, 10150-10155

53. Xie F., Sutherland D. E. K., Stillman M. J., Ogawa M. Y., Cu(I) binding properties of a designed metalloprotein, J. Inorg. Biochem., **2010**, 104, 261-267

54. Srinivasan C., Posewitz M. C., George G. N., Winge D. R., Characterization of the copper chaperone Cox17 of Saccharomyces cerevisiae, Biochemistry, **1998**, 37, 7572-7577

55. Chen X., Hua H., Balamurugan K., Kong X., Zhang L., George G. N., Georgiev O., Schaffner W., Giedroc D. P., Copper sensing function of Drosophila metal-responsive transcription factor-1 is mediated by a tetranuclear Cu(I) cluster, Nuc. Acids Res., **2008**, 36, 3128-3138

56. Zhou L., Singleton C., LeBrun N. E., High Cu(I) and low proton affinities of the CXXC motif of Bacillus subtilis CopZ, Biochemistry, **2008**, 413, 459-465

57. Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V., Mondragón A., Molecular basis of the metal- ion selectivity and zeptomolar sensitivity by CueR, Science, **2003**, 301, 1383-1387

58. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., 2010, 192, 6497-6498

59. Davis A. V., O'Halloran T. V., A place for thioether chemistry in cellular copper ion recognition and trafficking, Nat. Chem. Biol., **2008**, 4, 148-151

60. Davies S. L., Whittenbury R., Fine structure of methane and other hydrocarbon-utilizing bacteria, J. Gen. Microbiol., **1970**, 61, 227-232

CHAPTER 4:

A bioinformatics study of Csp1 and Csp3 protein homologues

4.1 Introduction

4.1.1 Csp1 and Csp3: a system of copper storage proteins in *M. trichosporium* OB3b

Csp1 has been identified from *M. trichosporium* OB3b grown under elevated copper concentration and the protein was isolated loaded with appreciable amounts of copper. While the Csp2 homologue, identified through bioinformatics has high sequence identity (57%) with Csp1 and is similar in that it contains 13 Cys and a predicted twin arginine translocation (Tat)-leader peptide, Csp3 shares only 18% sequence identity with Csp1. Consistent with this, the in vitro characterisation of recombinant Csp1 and Csp3 revealed significant differences between the two proteins. Csp1 binds ~13 Cu(I) ions per monomer and is expected to be exported from the cytosol, as it contains a Tat-leader peptide (1). Csp3 on the other hand, has significantly higher Cu(I) binding capacity, at ~18 molar equivalents of Cu(I) per monomer, and does not contain a predicted Tat-leader and, therefore is expected to be cytosolic (1). Except for differences in their Cu(I) binding capacity, Csp1 and Csp3 differ in the kinetics of Cu(I) release, with Csp1 readily releasing effectively all bound Cu(I) in contrast with the slow kinetics of Cu(I) release from Csp3. Neither of the two proteins removed Cu(I) from methanobactin (mb), however apo-mb readily removes all Cu(I) from Csp1 within an hour, which may imply apo-mb extracts Cu(I) from Csp1 to transport it to pMMO, the main copper-dependent methane oxidising enzyme used by methane oxidising bacteria (MOB) that is localised in the intra-cytoplasmic membranes (2, 6). Notably, the rate of Cu(I) exchange between apo-mb and Csp3 is strikingly slower.

The predicted localisation of Csp1 outside of the cytosol, combined to its Cu(I) binding capacity, notably Csp1 can store 52 Cu(I) ions as a tetramer, renders this protein a candidate for copper storage, possibly for pMMO. The protein target that Csp3 may supply copper to, as a cytosolic copper storage protein (the Csp3 tetramer can bind 72 Cu(I) ions), is unknown, however a role of this protein as defence against copper toxicity, similar to that of metallothioneins (MTs), cannot be ruled out (3, 4). Csp1 and Csp3 are novel in the fact that they store large amounts Cu(I) in the core of a 4-helix bundle. While it is interesting to speculate on the physiological role of Csp1 and Csp3 in *M. trichosporium* OB3b, the possible presence of Csp1 and Csp3 protein homologues in other MOB, or other bacteria, may provide insight into the role of these proteins.

4.1.2 Genome-encoded information for MOB

Methylotrophs are bacteria that are capable of oxidising only one-carbon compounds that they use as a source of carbon and energy and are spread between Proteobacteria and Firmicutes (5). MOB are methylotrophs since they oxidise methane and are characterised as obligate, capable of oxidising solely methane, or facultative, which oxidise other single-carbon compounds as well. MOB are the only biological sink for the mitigation of atmospheric methane levels and are found among Proteobacteria (6) and Verrucomicrobia (7). Complete genomes are available for a number of MOB (8-17) in the alphaproteobacteria and gammaproteobacteria subdivisions, synonymous to Type II and Type I respectively (7), while the *Methyloversatilis* species has been suggested to belong taxonomically to betaproteobacteria (5, 16) (section 1.4.2).

The information derived from the available genomes reveals MOB share, as expected, genes responsible for methane oxidation, but also variations between organisms. The *pmoCAB* operon which encodes pMMO is present in all MOB genomes, except that of the facultative methanotroph *Merthylocella silvestris* BL2 (13, 18, 19). *Methylocella silvestris* BL2 lacks the intra-cytoplasmic membranes usually housing pMMO and employs the iron-containing sMMO for methane oxidation. Moreover, *Methylocella silvestris* BL2 can also oxidise two- three- or four- carbon compounds and is the only methanotroph that is phylogenetically closer to a non-methanotrophic organism, *Beijerinckia indica*, than to other methanotrophs (19). In *Methylocystis* strain SC2, *Methylocystis parvus* OBBP and *Methylococcus capsulatus* Bath *pmoCAB* can be present in multiple copies (9, 11, 17) and both *pmoCAB* operons need to be expressed for maximal efficiency in methane oxidation in *Methylocystis strain* SC2. Except for two copies of *pmoCAB*, *Methylocystis strain* SC2 encodes a second type of pMMO (*pmoCAB2* operon). The two types of pMMO have different kinetics for methane oxidation and therefore function under different methane concentrations (20). *pmoCAB2* is present in other MOB genomes within the alphaproteobacteria but not the gammaproteobacteria (20, 21). A different operon (*pxmABC*) is present in the genomes of *Methylobacter tundripaludum* SV96 and *Methylomicrobium album* BG8 and is thought to encode a copper membrane MMO of unknown substrate specificity. The different gene organisation in this operon (*pxmABC* in contrast to *pmoCAB*) and sequence alignments indicate the MMO is distinct from the pMMOs identified so far (10, 15, 22).

Other genes related to methane oxidation are present in MOB genomes including genes encoding methanol-, formaldehyde-, and formate dehydrogenases, which are essential for the further oxidation of methane after conversion to methanol (6, 8-17, 23). The *mmoXYBZDC* operon responsible for sMMO expression is present in switchover organisms (24). In *Methylocystis* strain Rockwell (ATCC49242) genes related to copper homeostasis (*copCD*) are located downstream of pMMO (8), an arrangement that has also been observed in the ammonia-oxidising *Nitrosomonas eutropha* C91 (25). Two copies of the copCD operon have also been identified in *Methylocystis* strain SC2 along with several copies of genes encoding P-type ATPases responsible for copper transport (11). Most MOB genomes also encode an open reading frame encoding a peptide which is similar to the mb backbone (26), indicating that mb results from post-translational modification of a ribosomally synthesised precursor peptide. Notably the MbnA precursor peptide is also present in a number of organisms that do not oxidise methane (27).

Except for genes serving methane oxidation, a number of genes related to nitrogen metabolism are present in MOB genomes. The ability to contribute to nitrification and denitrification has been recorded for MOB (6, 28) and appears to be widely distributed in these organisms. Genes responsible for the expression of ammonium (*amtB*), nitrate (*nark*) and nitrite (*nasFED*) transporters are encoded by MOB. The main enzymes involved in nitrogen metabolism are also predicted in MOB (8-17, 29). The *nifH* gene, encoding for the highly conserved iron-containing subunit of nitrogenase, involved in nitrogen fixation, is present in MOB genomes, with the exception of the *Methylomicrobium* species (30). pMMO, structurally and functionally homologous to the ammonia monooxygenase (AMO) of ammonia oxidising bacteria (AOB), can oxidise ammonia (31) and genes (*haoAB*) encoding enzymes homologous to hydroxylamine oxidoreductase for oxidation of hydroxylamine (32) are found in MOB genomes (8-17, 29). Moreover, nitrite reductase and nitric oxide reductase are also encoded by many MOB by the *nirS*/ *nirK* and *norCB* genes, respectively (10, 12, 28). The electron transfer required for all the above reactions is thought to be mediated by a number of proteins of the cytochrome family. *Methylococcus capsulatus* Bath is thought to employ cytochrome P460, encoded by *cytL*, for electron transfer during oxidation/reduction of nitrogenous oxides while a cytochrome *c'* is thought to be the electron acceptor from cytochrome P460 (9, 33). *Methylomonas methanica* MC09 also

encodes cytochrome P460 (12), while *Methylosinus trichosporium* OB3b encodes a cytochrome *c'*-alpha (29).

MOB and AOB are believed to share common descent, as portrayed by the homology between pMMO and AMO, which are thought to have evolved from a common ancestor enzyme the specificity of which adapted to the environmental needs of the bacteria (31). Other genes related to nitrogen metabolism are broadly distributed in bacteria, as a result of horizontal gene transfer events that facilitate the adaptation of bacteria to their environment (22, 34-36). The *nirK* gene responsible for the expression of the copper-containing nitrite reductase NirK is widely distributed among nitrifying bacteria (35) and other copper-dependent enzymes, are AMO and the nitrous oxide reductase (37, 31). The diversity of the organisms that contribute to nitrogen metabolism through nitrification and denitrification (22, 34, 35, 36) renders the understanding of their evolutionary relationship challenging. Nevertheless, the evolution of copper-dependent enzymes involved in the above processes coincided with the changes in the levels of atmospheric oxygen and the subsequent availability of copper (38). In order to handle the copper utilising proteins that evolved after the great oxidation event (see introduction), bacteria developed mechanisms to tightly handle copper and ensure the metal is tightly bound to proteins and not free inside the bacterial cell.

Based on the currently available research evidence, bacteria use homeostatic systems, consisting of sensors (such as CueR) (39), copper chaperones (such as CopZ) (40), copper transporting systems (such as the cue system, composed by the CueR sensor, the Cu(I) transporter CopA and the multicopper oxidase CueO (39, 41, 42), or the cus system composed of the CusCFBA components (41)) in order to avoid copper-induced toxicity and ensure that the levels of free Cu(I) in the cytoplasm are effectively zero. Consistent with this model, most known copper enzymes in bacteria are periplasmic, with few exceptions. These include plastocyanin and cytochrome c oxidase in cyanobacteria, both of which are localised in the thylakoid compartment, and pMMO in MOB that is localised in intra-cytoplasmic membranes (38). The cyanobacterial thylakoid is a rare example of cytoplasmic copper requirement (import is potentially mediated by CtaA) and subsequently is chaperoned by Atx1 to PacS that exports it to the thylakoid (43-46). Although it is not clear how thylakoid membranes form and whether they are continuous to the cytoplasmic membrane in all cases (47), the import of Cu(I) in the cytoplasm and the subsequent export to the thylakoid is well documented

(43-46). In MOB the nature of the intra-cytoplasmic membranes in MOB is not well understood (48) and neither is the mechanism through which pMMO acquires copper. Intra-cytoplasmic membranes have also been observed in members of purple phototrophic bacteria and are believed to function to maximise the surface area housing membrane enzymes involved in photosynthesis (49). In the case of *Rhodobacter sphaeroides* intra-cytoplasmic membranes that were completely detached from the cytoplasmic membrane, therefore functioning as a separate organelle, were identified (49). Despite the ambiguity on whether intra-cytoplasmic membranes form discrete cytoplasmic compartments, the presence of certain enzymes (plastocyanin, cytochrome c oxidase, pMMO, AMO and potentially other monooxygenases) in these membranes opens the possibility that copper requirement in the cytoplasm is not a feature of cyanobacteria or MOB only. In *M. trichosporium* OB3b, Csp1 is expected to be exported from the cytosol, due to the predicted Tat-leader the protein contains, where the protein possibly acquires copper before folding and export. Csp3 is expected to be cytoplasmic, as it does not contain a Tat-leader peptide, indicating the need of *M. trichosporium* OB3b for cytoplasmic Cu(I) storage. In order to gain some perspective on these possibilities mentioned, bioinformatics searches for protein homologues of Csp1 and Csp3 in other bacteria were performed.

4.2 Materials and methods

4.2.1 Identification of Csp1, Csp2 and Csp3 homologues in MOB and other bacteria

Protein homologues of Csp1 and Csp3 were identified by BLAST searching the protein sequences of Csp1, Csp2 or Csp3 against MOB for which complete genomes are available (8-17) using the NCBI protein-protein blast tool (49) (see Table 4.1). In order to identify protein homologues of Csp1 and Csp3 in other bacteria, the BLAST search was repeated against the non-redundant database for bacteria using the NCBI protein-protein blast tool (50). In all cases the size of the protein sequence (number or residues), the number of conserved Cys residues and an expect value (E) lower or equal to 10^{-3} were used as criteria for the identification of true homologues within the list of all the generated hits. SignalP (51) and TatP (52) were used to identify Tat-leader peptides. The protein sequences of the Tat-leader peptide containing Csp1 homologues, not

including the Tat-leader peptides, were aligned using Clustal Omega (53). The same programme was used for the alignment of the Csp3 homologues which do not contain Tat-leader peptides.

4.3 Results

4.3.1 Identification of Csp1 and Csp3 protein homologues in MOB

The blast search of the Csp1, Csp2 and Csp3 protein sequences against published MOB genomes (8-17) and against all bacteria in the non-redundant database of NCBI resulted in the identification of the proteins shown in Figure 4.1. Homologues of Csp1 are found in representatives of alphaproteobacteria (*Methylosinus* and *Methylocystis* genera) and also in the alphaproteobacterial methylotrophs *Methylobacterium populi* BJ001 and *Methylobacterium extroquens* DM4 (23, 54). A Csp1 protein homologue is also encoded in the methylotrophic *Methyloversatilis* sp. NVD, which has been suggested to belong in the betaproteobacteria group (5). Except for *M. trichosporium* OB3b, three other MOB strains, *Methylocystis* strain Rockwell (ATCC49242), *Methylocystis* SC2 and *Methylosinus* LW4, have both Csp1 and Csp3 homologues, while the gammaproteobacteria *Methylobacter tundripaludum* SV96 and *Methylomicrobium album* BG8 (10, 15) encode only Csp3 protein homologues. The Cys residues of Csp1 and Csp3 are almost fully conserved in Csp1 and Csp3 protein homologues, respectively (Figure 4.2), and the same is true for some of the Cys-X-X-X-Cys motifs that participate in Cu(I) binding in the two proteins (Figure 4.1).

4.3.2 Identification of Csp1 and Csp3 homologues in other bacteria

The blast search for Csp1 and Csp3 protein homologues against all bacteria in the NCBI non-redundant database resulted in the identification of protein sequences in a large number of organisms. Csp1 protein homologues were found among the alpha-, beta- and gammaproteobacteria (Figure 4.2) with some organisms being rhizobia (*Nitrobacter* sp. Nb-311A, *Hyphomicrobium zavarzinii*, *Rhodopseudomonas palustris* TIE-1, *Burkhoderiales* bacterium JOSHI_001, *Dechlorosoma suillum* PS) and pathogens (*Cardiobacterium hominis*, *Kingella kingae*, *Neisseria gonorrhoea*) (55, 56). Csp1 protein homologues were also identified in the metabolically versatile soil bacterium *Pseudomonas putida* (57, 58), the iron-oxidising *Gallionella*

capsiferriformans (59), the unclassified betaproteobacterium of the *Rhodocyclaceae* family bacterium R294 and the neuston gammaproteobacterium *Nevskia ramosa* (Figure 4.2) (60). The 13 Cys residues of Csp1 are almost fully conserved in all the protein homologues (Figure 4.2). The Cys-X-X-X-Cys motifs that participate in Cu(I) binding in Csp1 (involving Cys33/Cys37, Cys90/Cys94, Cys106/Cys110 and Cys113/Cys117) are also nearly fully conserved and so is His36 and Met48 (Figure 4.2). Csp3 protein homologues were identified in bacteria within the phyla of proteobacteria (alpha-, beta-, gamma-, delta-), cyanobacteria, firmicutes, actinobacteria, spirochaetes, bacteroidetes, planctomycetes, deinococcus/thermus and verrucomicrobia (Figure 4.3) (55, 56). The 18 Cys residues of Csp3 are highly conserved in the Csp3 protein homologues, as are the Cys-X-X-X-Cys motifs (involving Cys24/Cys28, Cys31/Cys35, Cys61/Cys65, Cys90/Cys94, Cys97/Cys101 and Cys114/Cys118) that participate in Cu(I) binding in Csp3 (Figure 4.3).

MOB species	Strain	Complete genome	Publication reference	Csp1/ Csp3 homologues	Tat leader
Methylomonas	sp. 11b	✔	-	-	
	sp. MK1	✔	-	-	
	MC09	✔	(12)	-	
Methylocystis	sp. ATCC 49242	✔	(8)	Csp1 Csp3	✔
	Rosea	✔	-	Csp1	✔
	sp. LW5	✔	-	-	
	sp. SC2	✔	(11)	Csp1 Csp3	✔
	parvus OBBP	✔	(17)	-	
Methylocella	silvestris BL2	✔	(13)	-	
Methylomicrobium	buryatense 5G	✔	-	-	
	album BG8	✔	(10)	Csp3	
	alcaliphilum 20Z	✔	(14)	-	
Methylococcus	capsulatus ATCC 19069	✔	-	-	
	capsulatus Bath	✔	(9)	-	
Methylosinus	sp. PW1	✔	-	-	
	sp. LW3	✔	-	-	
	sp. LW4	✔	-	Csp1 Csp3	✔
	trichosporium OB3b	✔	(29)	Csp1 Csp2 Csp3	✔ ✔

Methylobacterium	populi BJ001	✔	(23)	Csp1	✔
	extorquens DM4	-	(54)	Csp1	✔
Methyloversatilis	sp. NVD	✔	(5, 16)	Csp1	✔
Methylobacter	tundripaludum SV96	✔	(15)	Csp3	

Table 4.1 MOB strains with complete genomes are shown together with the Csp1 or Csp3 protein homologues they encode. Reference numbers are provided, wherever available.

```
Csp1      MERRDFVTA-FGALAAAAAAASSAFAGEDPHAG----HKMSHGA--------KYKALLDSSSHCVAVGEDCLRHCFEMLAMND--ASMGACTKATYDLVAACGAL----AKLAGTNSAFTP
Csp2      MERRQFVAA-IGAAAAAASASRAFAQTTQGLA-PGAPVHHHPA--------KYHALMETSAKCVSTGNECLRHCFGMLSMND--TSMADCTKASYDLVAACAAL----ETLSAVNSSATP
M Rock-1  MERRDFIAT-LGAVAAAATVSSARAEEGK------KVAHHHPP--------KYKALSDAAAKCVLEGNNCLRHCFGMLSMDD--SSMAECTKATFETIAACAAL----ESLSSTNSNFTP
M SC2-1   MERREFIA--VGTAAAVVSATQAFAQATGQQDGAAKMEDMHPP--------LYKNLQNASLECVATGNDCLRHCFGMFAMKD--TSMAECAEAAYQLVAACNAL----ASLAAANSSHVG
M LW4-1   MERREFVAA-AVGAAAALSATRALAEDRT-AD-PHAGHMMHGP--------KFQALMKTSAECVSTGNECLRHCFGMLSMND--ASMAGCTKAAYDLVSACAAL----ESLAAVNSTFTP
M rosea   MERREFIA--VGTAAAVASATQAFAQATGQQDSAAKMEDMHPP--------LYKNLEHASIECVSTGNDCLRHCFGMFAMKD--TSMAECADAAYQLVAACNAL----ASLAAVNSSHVG
M extroq  MRRAEMQDRRWVIGSGLALATAGLSVRLAEAQQPGAMPGMGGG---------AMTIQECVDSCLKSHAMCLDTGR-YCTEQGGRHVAAAHLALLLDCAEMCQMT----ANSLLRRSPQHA
M populi  MERRDFITA-LGAAAGLASMAAVRAQAAEHDHGHDHGHDHGAAGGHPEPATPFAALKEASAHCVSTGNDCLNHCFAMLAMKD--TSMTACMRATYDLVHACAAL----ETLAAVNSPHTR
M NVD     MDRRNAIGAGIAASLGGLLLAGNAAAADGHEHHHHAQAKPGGP-------HKHAALLEGTTNCLRAGEICHAECLRVLATGD--KDMAACAVTVSDMLATCDAL----LKLAAADSRHLP
Csp3      ------------------------MHVEAMISKHPQARGQT---------DRSLVQCVEMCFDCAQTCAACAD-ACLGEDKVADLRHCIRLNLDCAEICVAAGSIASRAAGTEESILR
M Rock-3  -------------------------------------MHKM---------SKEMQSCVDECLRCYQMCFGMAMTHCLETGGDHVKPKHFRAMISCAEMCRNA----AHMMLMKSPQAR
M SC2-3   -------------------------------------MHMI---------PKEMQACIDACLNCYQMCFGLAMTHCLEKGGEHVKPKHFRAMIACADMCRNS----TQMMLMNSPLAK
M LW4-3   ---------MGALAAATAATGTLAAEDPHA-----HHHGGGA--------KYKALFESTTKCVAAGEECLRHCFEMLAAND--ASMGPCTKSTFDVVAACKAL----ASLSGTASALTP
M tundr   ----------------------------------MKQPTHT---------EHAMQACIEACSHCHQVCLQSAMNHCLKTGGKHVEAEHFRLMISCAEICQTS----ANLQLSSSKFSH
M album   --------------------------MYTPQTEPQKTSTAA---------FQSMQPCIDNCNRCAQTCLQTAMNQCLEMGGRHVEPEHFRLMICCAEICRLS----ANFMLSSSPFHT
```

```
Csp1                                  AFAKVVADVCAACKKECDKFPSIA-ECKACGEACQACAEECHKVAA-----
Csp2                                  ALAKTVYDVCMACKKECDRFPQYS-ECKNCGDACKACADECQRVSS-----
Methylocystis Rockwell ATCC 49242-1  AMAKVVAGVCEACKKECDKFPEVA-ECNAMGAACKACADECKKIAA-----
Methylocystis SC2-1                  PLAKTVAMICEDCKKQCDKFPKVK-ECVECGKSCQKCADECRKV-------
Methylosinus sp. LW4-1               TLAKAVGDICMACKKECDRFPQYS-ECKACGDSCKACADECRKVSV-----
Methylocystis rosea                  PLAKTVAMICDDCKKQCDKFPKVK-ECVECGKSCQKCADECRKV-------
Methylobacterium extorquens DM4      VVCDACARLCEACARDCEGFAP-DPQMQRCARTCQDCARSCRDMA------
Methylobacterium populi BJ001        AMARAVAEVCAACETECAKFPDIA-ACRACRDSCRTCAAECRKTAA-----
Methyloversatilis sp. NVD            KMAALAAAVCDDCEKECRKHEKMHSECRQCAEACVECARACRKVAA-----
Csp3                                 TMLQTCAEMCRMCEEECRRHAGNHEHCRICADVCKECETACRSATGLTH--
Methylocystis Rockwell ATCC 49242-3  HICEDCAEACEACAKECDALP----DMKDCAAQCRRCAEACRKMAGQKMAA
Methylocystis SC2-3                  QMCALCAEACETCAKECDPIP----DMKECADECRRCAEECRRMSGQKMAA
Methylosinus sp. LW4-3               SFAKAVAQACLACKKECDKFPNIA-ECKACGDACKACADECEKVAA-----
Methylobacter tundripaludum SV96     RLCEVCAEVCEACAADCKKIG----DMECVEACEKCAKSCREMAS-AQH-
Methylomicrobium album BG8           RTCEVCAEICEACAKDCASIG----DMDECVSICRECAESCKQMASMATH-
```

Figure 4.1 Alignment of the full protein sequences (including Tat-leader peptides, shown in italics) of Csp1 and Csp3 protein homologues found in MOB. Cys residues known to participate in Cu(I) binding in Csp1 and Csp3 are highlighted in yellow, with conserved Cys-X-X-X-Cys motifs underlined. In strains that encode both Csp1 and Csp3 protein homologues the Csp1 homologue is numbered as (1) and the Csp3 homologue is numbered as (3). The abbreviated names of the organisms in the upper panel are shown in full in the lower panel of the alignment.

181

```
Csp1     -----------------------GED-PHAGHKMSHGAKYKALLDSSSHCVAVGEDCLRHCFEMLAMNDASMGA-CTKATYDLVAACGALAKLAGTNSAFTPAFAKVVADVCAACKKECD
Csp2     -------------------QTTQGLAPGAPV-HHHPAKYHALMETSAKCVSTGNECLRHCFGMLSMNDTSMAD-CTKASYDLVAACAALETLSAVNSSATPALAKTVYDVCMACKKECD
N Nb311A -----------------------NKAPAPAPAESMHPAKYKALEDATSHCVSKGDACMRHCLGMLSMKDASMAG-CTNAAYQMMAACGALRTLAAVNSPHVPALAKVIAQVCVDCQKECE
H zav.   ------------QEDPHAGHEH-H-AEMHGAEGHASSQHRALIEAAQKCVASAEACVPHCLGLIAKGDTSLAE-CLESVLAMSPVCTAVVRVASLQATRLKELVKVCGDICADCEKACR
R palust ------------------AEDH-S-AHMHAAAAGAAPANAKLIETASDCVKTGQACIAHCLQSFAAGDTSLAA-CAKSVDQMLSVCATLQKLASLGSPNLPAMAKVALAVCEDCEKECR
B bact   -------------------------HAHHHPAGGGKYDAVMQSAAHCVMTGETCLAHCHIVLAEGEKDMAE-CAKSVNELIAVCTALRSLAAQDAPRLPALAKVAIDTCLACEKECR
D suill  ------------------AAQDH-Q-NHHHDHGAMGGKSYSALIDSSSDCQKTGEACLAHCLVLLGQGDKEMAA-CAQSVSELLATCNALMKLALQGSKFTPAMAKVAADVCAACEKECR
R bacter ----------------DDHAHH-H-HDHGAGKPAGPHKRAALIASTSECLQAGQICLAECHRVLATGDKDMAA-CAITVTEMLATCDALLKLAAADSRHLPKLASLALAVCEDCEKECR
G capsi  -----------------AEHDH-M-AHEHHHGAAGN---AALAKSAADCVQTGEVCLAHCLILLGDGEKEMAA-CAKSVSQMLALCGALQQLANQNSKLLPKLAALAQDACNQCEEACR
K kingae ------------------HNHA--G-HAHHGAAAASTNPYESIRLAAGHCYAAGNVCLAHCIRLLSQGDTSMKD-CATNVNQMLNLCGALSNLAAQQSSLVPALAKVALQACRACAEACR
N gonor  ---------------HGHAD-Y-HHHHDMQPAAASAYTAVRQTAAHCLDAGQVCLTHCLSLLTQGDTSMSD-CAVAVRQMLALCGAVHDLAAQNSPLTRDAAKVCLEACKQCAKACK
C hominis----------------------E-E-HHHHAHGAAAANPNDALIKATAACLSAGRACLAHCIRLLSEGDKSMAD-CAKAVNQMLALCDATNSLAAQNASLLAAVAKQCAEACKQCADACK
P putida ------------------ATHE-H-AHQHVHGTAGARPFAGLIDSSSDCQKTGEACLAHCLVLLGQGDKQMAA-CAQSVSELLATCSALMKLALQGSRFTPAMAKVAADVCASCETQCR
N ramosa DDPRNARNDVAAAKAKPAAAVDHSA-HQGHAAAAAAPGRYAALATVYADCAARATECITHCQSVLATGDKSLGD-CLKTALACQTTCTAVAQLARLNSDFAPALARDTVPVMKACLDTCR


Csp1                                  KFPS-IAECKACGEACQACAEECHKVAA-
Csp2                                  RFPQ-YSECKNCGDACKACADECQRVSS-
Nitrobacter sp. Nb-311A              KFPD-VAECKACGDSCKTCAEECRKISA-
Hyphomicrobium zavarzinii            KHAEHHEVCKACADSCAAFTKEAKKLSDA
Rhodopseudomonas palustris TIE-1     KHADHHATCKACADACKSCADACRSV---
Burkholderiales bacterium JOSHI_001  KHEKKHSQCKDCADACKACADDCRKVAA-
Dechlorosoma suillum PS              KHEKKHAECKACADSCAACLKECKAVAA-
Rhodocyclaceae bacterium R294        KHEKMHAECRQCAEACVACAKECRKVAA-
Gallionella capsiferriformans        KHAEKHQECKACGESCAACAKECKKLSA-
Kingella kingae                      AHASHHAECKACYESCLACIAECEKFAA-
Neisseria gonorrhoeae                EHSAHHAECKACYESCLDCIKECEKLAA-
Cardiobacterium hominis              AHAEHHAECKACFEACKECAEQCGKIAA-
Pseudomonas putida                   KHENEHPECKACADSCAACLKECKALAA-
Nevskia ramosa                       PHIEHHAICKACFDACESAIKAAQSI---
```

Figure 4.2 Alignment of the mature protein sequences (not including Tat-leader peptides) for the Csp1 protein homologues found in bacteria. The residues known to participate in Cu(I) binding in Csp1 are almost fully conserved in the homologous proteins and have been highlighted (Cys (yellow), His36 (green) and Met48 (magenta)). Cys-X-X-X-Cys motifs involved in Cu(I) binding in Csp1 are underlined. The abbreviated names of the organisms in the upper panel are shown in full in the lower panel of the alignment. The first organism of each bacterial group has been highlighted according to classification as alphaproteobacteria (grey), betaproteobacteria (light blue), gammaproteobacteria (green).

```
Csp3        --------------------------------------------MHVEAMIS--------KHPQ-----------ARGQTDRSLVQCVEMCFDCAQTCAAC-ADACLGE--DKVADLRHCIRLNLDCAE
R legum     -----------------------------------------------------MTMHHMSTEMKACIDNCLACYRECLSMAMGHCLELG--GEHTKPQHFKLMMACAE
B WSM1253   --------------------------MHAQEMIS--------THPQ----------VRGQTNDALIRCIEECYSCAQTCTSC-ADACLAE--NNVKSLTQCIRLNLDCAD
C BNC1      --------------------------MHVQEMIS--------AHPD----------VQGSTADRLLRCIEECYNCAQTCTAC-ADACLAE--SSVDMLKQCIRLNLDCAD
L barat.    --------------------------MNAREMIA--------THPA----------VKGNVNAALVACIDDCYACAQTCTAC-ADACLGE--EMVDQLRQCIRTDLDCAD
B abyss.    --------------------------MQIQPMIR--------THPD----------VRGGVNDSLVRAIEAAYGCAAVCHIC-ADACLAE--EMVKDLTQCIRLNLDCAD
R centen.   --------------------------MSLREMIA--------DHPD----------VRGDLNPALAACIEETLACAQVCTAC-ADACLAE--GMVAELRQCIRLNLDCAD
A brasil.   --------------------------MHAREMIS--------THPQ----------VRGNANDALIRCIEACYDCAQTCTTC-ADACLGE--GQLAELVQCIRLNMDCAD
P zucin.    --------------------------MHAQQMIS--------THPQ----------VRGSTNDSLVQCIEACYDCAQACTAC-ADACLGE--DAVRELIQCIRLNQDCAD
C Taiwan.   ----------------------------------------------MIRPTVQENAARYADCIAACNAAAAALKC-AAACLEEQ--DVRKMARCIALDMDCAG
R metall.   ----------------------------------------------MIRPTVQENFSRYADCIAACNAAAAACLKC-AAACLEEP--DTRKMTRCIALDMDCAG
N multif.   ---------------------------------------------MFLYTETDQNLQACIDACNHCYRTCLRMAMNHCLEAG--GKHVEADHLRLMMNCAE
T denitr.   -------------------------------------------MHTEFASCIKACDECAAACDHC-ATACLQEA--DPKPMARCIALDIDCAA
P aerug.    ---------------------MTR-AIND-PGN----------EDPGSLLETDADALLGGAAAQAPEERCRLAAQACIRACERY-LALCTESS--R-----EQRQHAGDCAD
A vinel.    -----------------------------------------------MNASMYESCIQACSDCAWSCETC-AASCLRED--DVQAMARCISLDMDCAD
C condim.   -----------------------------------------------MSLQYSECIEACYACATACDNC-AASCLKES--DLDMMRECIRLDMQCAS
H lutea     -----------------------------------------------MDLEKLKSCIDACNECATYCDYC-AVSCLQED--DVKMMAECIRLDMQCAQ
P agar.     -----------------------------------------------MSHNKFTSCIEACYKCATACDHC-STACLQEP--DPAKMADCIRADLDCAD
S typhim.   -----------------------------------------------MQQEHRECIEQCYECAAACDIC-ASSCLRED--NVEMMKHCIQLDMQCAA
Enterobac   -----------------------------------------------MLDEYKKCIEACYLCATACDNC-AASCLDEE--NLEMMRECIKLDMQCAN
L pneumo.   -----------------------------------------------MTHQQYDMCIKACQACLLECEHC-ANACLHEE--DCNDLARCISLDRDCAA
D Africa.   ---------------------MFRARDMLE--------HHK----------AAGRMDRESLVRCIEACYECAQTCAMC-ADACLAE--DEVKRLARCIRLDLDCMD
S elongat   ----------------------------------------------MMMTMMNPSMTDDMQMCMAACMDCMKTCMET-MGYCLKMG--GDYMDPMMMGMMRDCAE
Anabaena    --------------------------------------------MQECIQNCLDCHSICLNT-VTYCLQKG--GNHTEASHIRLLLDCAE
P PCC7327   -------------------------------------------MLLVSKEYQSSFDTAMYCVVECEHC-AKACMGDP--E---MLGCARTCLDCVE
B subtil.   ------------------------------------------MEQYSEACIEACIDCMKACNHC-FTKCLEES--VQHHLSGCIRLDRECAD
T marian.   -----------------------------------------MAVVHTDVQQYKKCIDACNSCMQACEQC-LTACLKEP--DAAQRGRCVQLLRDCAD
B lichen.   --------------------------------------------MKSGYEECIKACRECLEACNHC-FDKCLMEE--EAGMMAECIRLDRECAE
B sonor.    ---------------MDLQPF-VRFARFSR-RVLC--------------T------ESAEKGVFTVENVSYDECIKACQECLQLCNEC-FDKCLMEE--DASMMAECIRLDRECAD
V aren.     ----------------------------------------------MNRQYEECIKICMEYLEACNTC-FDSCLKED--DVKMMTKCIRYDRACAD
B atrophy   ----------------------------------------------MTQHAETCLEACVSCIEAFNTC-FSRCLIEQ--AHHDLSGCIRLDRECAD
B mojav.    -----------------------------------------MQQTIEACITCMEACNHC-FTKCLAES--AHHDLSGCIRLDRECAD
M halotol   ---------------------------------------------MNAAQKECLEACIECMKQCNQC-FNDCLQED--DVKMMAECIRLDRECAD
P FSL R5    ---------------------------------------------MTMSHEQYQKCIEECIRCMMVCNHC-YNACLDEE--NVAMMTKCIRLDRQCAD
S kuerl.    ---------------------------------------------MTKEQMRECIDECLKCMEVCNNC-YDACLQEE--DVAMMRDCIRLDRMCAD
B cereus    ----------------------------------------------MQNMYQACIEECLKCMEICNSC-YSACLQES--DVKMMVECIRLDRECAD
B coagul.   ----------------------------------------------MDKSYEECIKACQECMEACNQC-YSACLKED--NVKMMAGCIRMDRECAN
B 17376     ----------------------------------------------MVNQKYDCIQACIECIEACNVC-FDSCLKEE--NVKMMAECIRMDRECAE
Lysinibac   ----------------------------------------------MKMSYEECIKACLECMKECNKC-YHECLEED--DVKMLAECIRTLRECAD
B oceani.   ----------------------------------------------MYSQSFEECIQACLECMKACNVC-YDACLGEE--DVKMMAECIRLDRECAD
B neals.    ----------------------------------------------MNMNYEECIKACLECMKACNSC-FDACLNEE--DIKMMARCIRLDRECAD
```

183

```
B fordii    ----------------------------------------------------------------MNSTYEACIKACLECMEACNHC-FDACLKEE--DVQMMAECIRLDRECAD
B massil.   ----------------------------------------------------------------MPMTYEECIKACLECMEACNGC-YDACLKEE--DVKMMAECIRLDRECAD
E pavil.    ----------------------------------------------------------------MHRIYANELETLRDCVETCNYC-LESCLEEH--DVKMMVDCIRLDRECAA
C WN1359    ----------------------------------------------------------------MQEQLVKISEKMFECQRVCNEC-FDACLKED--HVQMMAECIRLDRECAD
S MJ3       ----------------------------------------------------------------MAEKQYENAMKALHECMEACNYC-FDSCLKED--DVKMMAGCIRLDRECAD
B vallis.   ----------------------------------------------------------------MAYEQYQSLVETLHDCMVACNHC-YDACLKEE--NENMMADCIRLDRECAD
S xylosus   ----------------------------------------------------------------MSYGKNRLLIQTLHECVEACNYC-FQACLKEE--NVKMMAECIRLDRECAD
P dongh.    ----------------------------------------------------------------MATEQHAELIKTLHDCAAACNHC-FDACLQED--DVKMMAQCIRLDRECAD
G borac.    ----------------------------------------------------------------MSHEKYKDLLETLHDCMTECNHC-FNACLQED--NVKMMAECIRLDRECAD
B cecemb.   ------------------------------------------MGV----------QR-------IRSKQEGFQLAHEQHQELLTALHECMEACNHC-FDACLKEE--NVKMMAECIRLDRECAD
B pseudo.   ----------------------------------------------------------------MSHEKYKSMITTLHECMTACNHC-YDSCLKED--HLEMMRECIRLDRECAD
B endoph.   ----------------------------------------------------------------MSHENYQSIIETLHECMSTCNHC-YDACLKED--DVKMMAQCIRLDRECAD
B isron.    ----------------------------------------------------------------MLHEQHQQLLQTLHECMAACNHC-FDACLKED--HVQMMTECIRLDRECAD
V halod.    ----------------------------------------------------------------MSHEQHQQLNVLHECMEACNHC-YNACLQED--DIKMMADCIRLDRECAD
O Ndiop     ----------------------------------------------------------------MSHEQHQELIQSLHECAVACNYC-YDACLKEE--DIKMMAECIRLDRECAD
S albus     ----------------------------------------------------------------MAHEQHQQLIQTLHECMEACNHC-YDACLREE--DVKMMAECIRLDRECAD
P TG2       ----------------------------------------------------------------MNSRYVECIDACQECLEACNVC-FDSCTREE--DVKMMAACIRLDRECAD
S newyor.   MAQRKVICLFGVQHIKKEKGFF-KEFLLFSF-KTTF-KGK----------IS------LIAKIAGGESMNTKYEECLKACLECLEACNVC-FDACLKED--DMKMMADCIRLDRECAD
A alba      ----------------------------------------MMVVSEMLR--------AYPA----------EMAKMDGDMLARCIEECYSCAQICTAC-ADACMAESDKVMATLRKCIRSCEDCAD
S polym.    ---------------------------------------MLVAEMQE--------AHPR----------PLVGIDADALAECIEACLACAQACTSC-ADACLHE--EMVADLRRCIRLNQDCAD
Nocardia    -------------------------------------MLE--------SHPH----------ASNETGTAELAACIEACFECAQTCTAC-ADACLGE--PSVAELVDCVRSDLDCAD
A Rue61a    -----------------------------MMAHHISAVIG--------AHPK----------GAGSLHKEKLAECIAACFECAQTCTAC-ADACLGE--DMVADLAVCIRTNLDCAD
L aquat.    -------------------------------MSAVEQMLD--------SYPA----------DLGDVDRQKLVACIEACIECAQTCTAC-ADACLSE--EMVADLVKCVRTDLDCAD
Nocardio    -------------------------------MHTVVQLLK--------TYPK----------DLGGIDTSQLTKTIQSLIECSQACTAC-ADACLSE--DGVADLTKCIRSNLDCAD
R rhodo.    --------------------------------MLE--------SYPK----------NVGEIDTGVLAACIEACIDCAQTCTAC-ADACLAE--DTVAELTACIRTDLDCAD
Brevibact   ------------------------------MTHITSMIE--------THPN----------DTTGLDVQKLADCIAACFECAQTCTAC-ADACLAE--DMVAELRNCIRLNLDCAD
Ornithin.   ------------------------------MTHVTSMLE--------TYPG----------DLGSVDTQKLAECIAACFECAQTCTAC-ADACLAE--DMVADLRECIRTNLDCAD
Citricoc.   ------------------------------MTHHVSSMLQ--------TYPK----------DLGAIDRQRLAECIEACFECAQTCTAC-ADACLAE--DMVADLRQCIRLNLDCAD
M luteus    ------------------------------MTHHVSAMLD--------TYPK----------NVGNIDRQKLAECIQACFECAQTCTAC-ADACLAE--DMVADLRQCIRLNLDCAD
K sedent.   ------------------------------MTHHVDAMLK--------THPQ----------GTGTIDQAKLAACIEACFECAQTCTAC-ADACLAE--DMVAELRDCIRTDLDCAD
I calvum    ------------------------------MHTVEAMLE--------TYPK----------DLGDVDRGRLAACIQACYECSQACTAC-ADACLSE--DMVADLTTCIRTNLDCAD
C halotol   ------------------------MEHDMTHHVRTMLD--------THPK----------DLGQIDKDKLAECIEACFECAQTCTAC-ADACLGE--DMVAELTTCIRLNLDCAD
A pavil.    -----------------MGTQ-LRMLKGEGTMTHHVETMLD--------TYPK----------DLGGVDKQALAACIEACFECAQTCTAC-ADACLAE--DMVAELTKCIRTNLDCAD
B parac.    ------------------------------MTHHVASMLQ--------TYPK----------DLGTIDQQKLAECIEACFECAQTCTAC-ADACLAE--DMVAELTQCIRLNQDCAD
K palust.   ------------------------------MTHHVSSMLQ--------TYPK----------DLGGIDTQKLAECIEACFECAQTCTAC-ADACLAE--DMVAELTQCIRLNLDCAD
A LLX17     ---------------------------MTQAKMLE--------AHPN----------DLSGVDRQKLAECIAACFECAQVCTAC-ADACLSE--DMVAELTDCIRTNLDCAD
N CF8       ---------------------------MTHDKMLD--------AYPQ----------DLGTIDRDKLSECIAACFECAQVCTAC-ADACLAE--EMVALLTTCIRTDLDCAD
C flavig.   ---------------------------MKTTQMLD--------TYPA----------TI-NLDRQLLARVIESLVACSQACTAC-ADACLSE--EMVADLRKCIRSNLDCAD
C gilvus    ---------------------------MKTTQMLQ--------TYPA----------EI-NLDRDLLARVIDALVECSQSCSAC-ADACLSE--DMVADLRKCIRTNLDCAD
S coelic.   ------------------------MPTTVNDLLR--------TYPA----------DLGGVDREAMARCIEECLRCAQACTAC-ADACLSE--PTVADLTKCIRTDMDCAD
M lupini    ------------------------MPSTTMPMLE--------TYPQ----------SI-NLDRAKLAATIDALNDCAQACTAC-ADACLSE--DMVADLTKCIRTDLDCAD
```

184

```
Mycobact. -----------------------------------MSTASKMLE---------TYPQ----------DLGGIDRAALASCIEACLECAQACTAC-ADACLGE--DSVQQLTTCVRTNLDCAD
N halotol -------------------------------------MVD--------TYPA---------SLGDVDREKLIRCIEECFVCAQACTAC-ADACLSE--DRVSELRKCIRANSDCAD
T fusca   ------------------------------MSVAGQMLE--------TYPQ---------SLGGVDQEKLRACIEACFECAQACTAC-ADACLSE--DRVAELTKCIRTDLDCAD
B saxobs. ------------------------------MTVAAQMLD--------TYPK---------DLGGVDKQKLIECIEACVECAQACTAC-ADACLSE--DMVAELTTCIRANTDCAD
M marinus------------------------------MSVAGRMLE--------AYPK---------DLGGVDRQKLQECIEACVECAQACTAC-ADACLSE--EMVAELATCIRTNLDCAD
L biflexa-------------------MNRKELLQKAG-MAVAVSGILSTLS-----AE------DHDHS-TAMPTAGKSKYAKAMMAAIHCQLSAEVC-LSHCITELGKGDKAMAACAASTREVIS
B baltica--------------------------------------------------------MKSLNSELQACIDNCNGCITAARIC-LDKHLGEP--D---MKKCHQLCLDCTT
A agaril.-------------------------------------------------MKNQKLIDALNKCISHCNYC-ADACLETD--NIKMMVDCIRTDRACAE
C Kuenen.---------------------------------------------MSSKQYEGCIEACNECVTACEHC-AASCLREQ--DIKTMVRCIELDRDCAD
D perar. -------------------------MTNPMTQPLQGMLE---------THPQ--------AGQGNLDQQALLECLAACFECAQVCTSC-ADACLGE--QNLDMLRRCIRLNLDCAD
C flavus -------------------------------------------MDPEKYQACIEACHACVTACENC-AASCLEED--DVEMMVGCIELDRSCAD


Csp3                                          ICVAAGSIASRAAGTEESILRTMLQTCAEMCRMCEEECRRHAGNHEHCRICADVCKECETACRSATGLTH----
Rhizobium leguminosarum bv.viciae 3841        ICRTSAHFMLIGSEH----HKHVCRECAEICGQCAEDCERIGD----MQSCVDACRRCADSCRKMAA-------
Bradyrhizobium sp. WSM1253                    ICNITGRISTRRTGSDEEMIRRMLDTCAAACRLCGEECEKHAKMHEHCRICAESCRRCMNACEDAGRSMVH---
Chelativorans sp. BNC1                        ICAVTGAVASRRTGRNADVMRAMIEVCEQACRTCGEECHMHADQHEHCRICAEACMSCADACREAVADVH----
Lutibaculum baratangense                      VCAATGAVATRRTGSDETIIRAMLEACATACRTCADECGKHASMHEHCRICAEACRRCEESCRKALASMA----
Brevundimonas abyssalis                       VCLATAGIAARRTGSNEALIKRMLETCAEACADCAVECEKHADMHEHCRICAEECRRCEQACRDAAATISPSQH
Rhodospirillum centenum SW                    VCAAAGAVATRRTGGNVPVIRALLDACALACRRCGEECARHAGAHAHCRLCADSCRRCEAACREALESLPA---
Azospirillum brasilense Sp245                 VCTATGSVATRRSGSNEAVIRAMLDACATACRLCAEECERHAGMHQHCRICAEACRTCEQACRQARQTLSH---
Phenylobacterium zucineum HLK1                ICAAAGAIASRRTGSNPEAIRAVLAACQDACRVCAEECERHASMHEHCRICAESCRRCEQACQAALQTAH----
Cupriavidus taiwanensis                       IAQLAASYMLRNSEF----APLVCEDCAEVCKWCKEECERHD-A-EHCQECARACAVCMEQCLKMTA-------
Ralstonia metallidurans                       IANLAASYMLRNSEF----APLVCEDCAEVCKWCKEECERYD-H-WHCQECAKACAACMEMCLKMTA-------
Nitrosospira multiformis pdb code 3LMF        ICQTSLNFMLSGSRF----SPKVCGVCAEICDACAKSCEQLDG---MEECVQTCRQCAEHCRKMAA-------
Thiobacillus denitrificans ATCC 25259         ICRLASGYMARGSEF----ARRMCAICAEVCEACGAECAKHQ-H-DHCQECAQACRRCAEECRRMAA-------
Pseudomonas aeruginosa pdb code 3KAW          LCRLAALLLERRSPW----APAACELAARYALACAERCDGD--E-PLERECAGACRRFVEACRPLLPA-----
Azotobacter vinelandii                        FCRMAATLMARGSDH----APAFCRQCAQVCRACAEECARHE-A-GHCRRCAQACRACAEECERMAA------
Cronobacter condimenti                        LCRLAAQFMALDSEF----ARALCQICADICQKCGETCGKHQ-A-DHCQACSRACLHCAQVCRSMN-------
Halomonas lutea                               ICRLAASFMAQDSEY----MKEICRLCADICRKCGEECAKHE-M-DHCQQCAKACMRCAEECQAMA-------
Pseudoalteromonas agarivorans                 ICRLAAAAMARGSEH----AKQICRLCADICEACGELCAKHE-P-DHCQECARACKSCADECRKMAA------
Salmonela typhimurium                         ICRLAAQFMALESEY----SQKLCRLCADICKACAEECARHD-H-DHCQNCARACSQCADACLKMAA------
Enterobacteriaceae                            ICRLAAQFMTLNSGS----AQDLCRLCADVCQKCGDECGKHE-H-DHCQDCSSACHHCAEQCRKMAA-------
Legionella pneumophila                        ICALAIEMMARNSPF----AKEICALCAKICRACGDECSKHQHM-EHCQRCAKACYQCAEACEKMA-------
Desulfovibrio africanus                       ICVATGNALSRQVGVEPAIQRAQVEACMEACRVCADECDKHAKMHEHCRICSEMCDTCKNACKVLLGNLEAKAA
Synechococcus elongatus PCC7942               MCQTCMNMMMCGSEF----IASICKLCSEVCMKCAECCSGMIDD-EMMMSCSASCRACADACMKMCPA------
Anabaena                                      ICETSANFMLRTSEL----HTRTCGTCAEVCERCAQNCDRFGDD-AQMKACTDMCRRCAESCRRMSMATA----
Pleurocapsa sp. PCC 7327                       TCRTIAIYMVRGSRF----IPHLAKACAEICDACAKECEKH-KD-EHCQKCARACRQAAEEYRKIAGVAAARA-
Bacillus subtilis                             ICALAVKAMQTDSPF----MKEICALCADICEACGTECGKH-DH-DHCQACAKACFTCAEQCRSMAA-------
Thermaerobacter marianensis DSM 12885         ICALASRVMSRGSDF----AGAICRVCAEICEACAQECGRFQ-D-EHCQECARECRACAEECRRMAA-------
Bacillus licheniformis                        MCGYAIQAMTRNSPY----AEDICQLCAKVCEACGNECSQH-KH-DHCQFCAESCFACAEACRKMAS-------
```

185

```
Bacillus sonorensis                          ICHAAVQAMTRNSPY----AEDICLLCAKVCEACGNECSFH-KH-SHCQVCAESCFRCAEACRNMAS------
Viridibacillus arenosi FSL R5-213            ACMFAIQAMVSNSPF----IDEICNLCADICERCAEECAKH-NH-EHCQQCARTCQKCADACRNMSA------
Bacillus atrophaeus                          MCELAAKAMQTDSPF----MKQICALCADVCEACGTECQKH-DH-GHCQACAKACIACAAECRKMAA------
Bacillus mojavensis                          ICALAVKAMQTDSPF----MKEICALCADICEACGTECEKH-DH-DHCQACAKSCFACAELCRSMAA------
Marinococcus halotolerans                    ICALTAKAIQSNSPL----MKETAQLCAEACKACGDECAKH-DH-QHCKDCADACYRCEKACRELAAS-----
Paenibacillus sp. FSL R5-808                  ICAFAAHAMSTNSVY----AREICSICADVCEECGTECKMH-DV-KHCQECAEACFRCAEACRAMAA------
Saccharibacillus kuerlensis                  MCEFAARAMTQNSPY----AAQICALCAQICEDCGNECAKH-DM-EHCQRCAEQCKRCAEVCRKMAA------
Bacillus cereus                              ICALAAKSMSSNSPF----AKEICQLCAKICEACGNECKKH-EH-QHCKECADACFRCAEACKLMAS------
Bacillus coagulans                           ICALALQAMQSNSPF----VKQICNLCAEICEACGEECRKY-SH-DHCQKCAEACLRCAEACRKMA-------
Bacillus sp. 17376                           ICGMAVTAMQTNSHF----VNQICGLCAEICEACGNECKKH-DH-DHCQKCAEACFKCAEECRKMAS------
Lysinibacillus                               ICAFSAKSMQSGSPF----VKQICQLCAEICQACGDVCKKHEHH-KHCQECAETCYRCAEVCKKMAV------
Bacillus oceanisediminis                     ICAYAAKALQSNSPF----AKQICILCAEICEACGNECKKH-DH-GHCKRCAEACFMCAEECRKMAS------
Bacillus nealsonii                           ICSLAAKAMQSNSPL----VKEICLLCADICEKCGEECKKHSHH-EHCLQCAESCFKCAEFCRKMVS------
Bacillus fordii                              ICAFAAKALQTDSPF----AKEICQLCADICEECGKECKKH-HH-SHCQKCAEACLKCAEICRNMAA------
Bacillus massiliosenegalensis                MCAFAAKAMQSNSPF----AKQICQLCAEICEACGNACKSHEHH-EHCRECAESCFRCAEICRKMAA------
Exiguobacterium pavilionensis                ICSFLAEAMTRDSAF----VPELARACAVVCKACAQECEKH-KH-AHCQECARVCVECASMCDRLAA------
Carnobacterium sp. WN1359                    ICAFASTAITRESAL----STDLIALCATICQACGNECEKH-EH-DHCQKCAKVCLECAELCRSYA-------
Salimicrobium sp. MJ3                         MCGYLEAAISRNSPY----ISELASVCAKICDDCAEECAKH-DH-DHCQKCAEACRKCAEECRKIA-------
Bacillus vallismortis                        ICAYVAQSIVRGTPF----VSELTQACAAICEACGNECKKH-DK-EHCQDCAEACFSCAEACKQAV-------
Staphylococcus xylosus                       MCTFLERELTIDSPF----AYDLAEICSKICEACGNECQKH-EH-DHCQECAKSCFKCAKACKDIA-------
Planococcus donghaensis                      ICAYLEHAITRNSPF----VAELAKACATICDACAEECSKH-DH-DHCKKCADACAKCAEACRNVA-------
Gracilibacillus boraciitolerans JCM 21714    ICGYLEAAISRNSPF----IKELAAVCTKLCTTCNEECKKH-DH-DHCQKCADACFKCADACKKVT-------
Bhargavaea cecembensis                       ICGFLEAAISRNSPF----ISQLAQVCAEVCEACGEECRRH-DH-EHCKKCAEACMKCAEACRNVA-------
Bacillus pseudofirmus                        ICSYFEQALVRGTPY----VSELAALCAKICEDCGNECKKH-DH-DHCQKCAEACFKCAEECRKLAA------
Bacillus endophyticus                        ICAYLEQALGRGTPF----VPELAQVCADICEACGTECKKH-DH-DHCQKCAEVCFKCAEECKKLAS------
Bacillus isronensis                          ICNYLEQAISRGTPF----ISELAAVCAKICEACGNECKKH-DH-DHCQKCAEACFKCAEACKNVA-------
Virgibacillus halodenitrificans              ICGYLEQAISRGTPF----VSELASVCAKICEACGNECKKH-DH-DHCQACAEACLKCAEECKKVA-------
Oceanobacillus sp. Ndiop                      ICSYLEQALSRNTPF----VSELASVCAKICEACGNECKKH-DH-NDiopCQKCADACFKCAEACRSVA-------
Salinicoccus albus                           ICAFLEQALTRNSPF----SSDLAAVCAKVCEACGNECQKH-DH-DHCQKCADACFKCAEACKEIA-------
Paenisporosarcina sp. TG20                   ACSYAIQAMTRNSPF----TKEICDLCATTCDRCGDECTQH-DH-DHCQRCAEACRKCAEACRAMIA------
Sporosarcina newyorkensis                    ICAFAAQTMTRNSPF----TKQILELCAEVCERCAEECAKH-DH-DHCQRCAESCRKCAEACRQMVA------
Actinopolymorpha alba                        VCDTTGRILSRQTGYDPAMTRTILEACMAMCRACAEECERHAGRHEHCRICAEACRRCEQACSQLMGALR----
Sporichthya polymorpha                       LCDALARILTRGPDTAPGLLRAAVRACADACRMCGDECATHAEMHEHCRICAEACRRCEQACLALNLG------
Nocardia ATCC 202099                         VCETTGRVLSRRTAPDQDVIRALLETCALACKRCGDQCAEHADHHEHCRTCAEACRRCERACRDLLTTLR----
Arthrobacter sp. Rue61a                      ICAATGNILTRLTGANTELTREALQACRTACAVCAEECERHAGMHEHCRICAEACRRCENACAALLESMV----
Leifsonia    aquatica                        ICASTANVLSRHTGYDANVTRAALESCRAACGDACADECDKHASMHEHCRICTDVCRRCERACADLIAGL-----
Nocardioidaceae                              ACSATANILSRHTGYDANITRAFLEACATACKACGDECARHADMHEHCRVCADACRACEQACRDLLVKLG----
Rhodococcus rhodochrous                      LCAATARILSRQTGNNAAVTKAVLEACATACRVCGDECDRHSSHHEHCRICAEACRRCEQACRDLLAALD----
Brevibacterium sp. JC43                      LCAATGSILSRRTGQNLATVKAALEACRTACAECAAECEKHADMHEHCRVCAEACRRCEQACADLLAAIG----
Ornithinimicrobium Pekingese                 VCAATGAVLTRQTGANAETVRAMLEACRTACKACGDECADHADMHEHCRVCADACRRCEAACEALLSSMG----
Citricoccus sp. CH26A                        LCAVTGRVLSRQTGTNAETVRALLEACRAACKGCGDECASHADMHEHCAEACRRCEQACADLLATLG----
```

186

```
Micrococcus luteus NCTC 2665                              VCAATGRMLSRQTGNNVETTRALLEACRAACKTCGDECESHAQMHEHCKVCAEACRRCEQACADLLVTLG----
Kytococcus sedentarius DSM 20547                          ICAATGAVLTRQTGENTAVLRAQLEACRTACSACAEDCEQHAGMHEHCKVCAEACRRCEQACADLLAALG----
Intrasporangium calvum DSM 43043                          VCATTGAVLSRQTGHDANLARAVLEACAAACKACGDECERHASMHEHCRVCAESCRRCEQSCRDLLATLA----
Corynebacterium halotolerans DSM 44683                    ICDVTGRVLSRQTGWDVNLIRSVLETCRAACQACGEECARHADMHEHCKVCAEACRRCEQACAELLATLA----
Agrococcus pavilionensis                                  VCETTGRVLSRQTGSDAAVIRALLEACRAACQACAEECERHAEMHESHCRVCAEACRRCERACADLLATLG----
Brachybacterium paraconglomeratum                         VCETTGRVLSRQTGKNDALNRALLEVCQTACQSCAEECEKHAGMHEHCKVCAEACRRCEQACAELLASIG----
Kocuria palustris                                         LCETTGRVLSRQTGNNIDVNRAALETCRTACRSCAEECEKHVDMHEHCRVCAEACRRCEIACADLLASLD----
Actinobacterium LLX17                                     ICVATGKALSRLSGYGTNVTRVFLEACAEACKACGDECAQHAEMHEHCKICAESCRRCEEACRALIATL-----
Nocardioides sp. CF8                                       ICLATGNALSRNTGRNTDLTRALLEACAVACRTCADECAQHAEMHEHCRICAEVCRRCEEACRSLLAAL-----
Cellulomonas flavigena DSM 20109                          SCAATARILSRHTGYDANITRAHLEACIAACRACGDECEQHAGMHEHCRICAEACRDCETACAELLAAIR----
Cellvibrio gilvus ATCC 13127                              SCATTARILSRHTGYDANITRAHLQACIAACRACGDECAQHAQMHEHCRICAEACRTCEAACTELLNAIG----
Streptomyces coelicolor A3(2)                             VCTATAAVLSRHTGYDANVTRAVLQACATVCAACGDECARHAGMHEHCRVCAEACRSCEQACQELLAGLG----
Micromonospora lupini                                     ICTTTARVLSRHTGYDANLTRSLLEACATACLSCGDECARHADMHEHCRVCADACRACERACRDLLATMS----
Mycobacterium sp. JLS                                     VCAVTGQVLSRHTGYDATLTRVVLEACATACRSCGDECDRHAGHHEHCRICAESCRRCEQACRQLLTSLG----
Nocardiopsis halotolerans                                 ICLTTGRVLSRHTGYDANITRAVLQACMQACKSCGDECASHADRHDHCRICADACRRCEQACEDLLAALG----
Thermobifida fusca YX                                     ICETTGRVLSRHTGYDANLTRAILEACAEACKVCADECARHADHHEHCRVCAEACRRCEMVCRELLASL-----
Blastococcus saxobsidens DD2                               ICDTTGRVLSRHTGYDANLTRAVLEACAAACKSCGDECGSHAEMHEHCRVCAEACRRCEAACRDLISSLG----
Modestobacter marinus                                     VCDATARVLSRHTGYDANLTRAVLEACAAACKTCGDECTQHAEMHEHCRVCAESCRRCEAACRELLSSLG----
Leptospira biflexa serovar Patoc strain Patoc 1 (Ames)    LCDSFVKLASQNSSF----TKKLANLCVEVCEACAKECDKHANHHAVCKECRDSCLACAKELKKV--------
Belliella baltica DSM 15883                               LCAACVQMMASQSDY----SSRVCAICADLCNACADECDKF-DS-EACKQCAEKCRQCAETCKKMAA------
Aquimarina agarilytica                                    ICATTVKLLAMNSAF----AKAMVEKCHEICNQCADECSKH-DH-QHCKDCADACKKCADACKAYLA------
Candidatus kuenenia stuttgartiensis                       ICSLAAQFMSRGSSF----AAKLCALCAEICQACGDECAKYK-T-EHCQQCAKACHKCAEECRKMAAIPSA---
Deinococcus peraridilitoris DSM 19664                     VCDATGRVLTRCTQPDMNVVRTQLQACLAACEACGAECEQHAEHHQHCAICAESCRRCADACRNLLSGISA---
Chthoniobacteer flavus                                    ICALAAREMARDSDF----AVRVCAICAEVCEACGTECGRHK-M-DHCQHCAAACRRCAKLCLEMSQAHSSQPL
```

Figure 4. 3 Alignment of the protein sequences of the Csp3 protein homologues identified in bacteria. The residues known to participate in Cu(I) binding in Csp3 have been highlighted (Cys (yellow), Asn58 (magenta), His110 (green)). Cys-X-X-X-Cys motifs involved in Cu(I) binding in Csp3 are underlined. The abbreviated names of the organisms in the upper panel are shown in full in the lower panel of the alignment. The first organism of each bacterial group has been highlighted according to its classification as alphaproteobacteria (grey), betaproteobacteria (light blue), gammaproteobacteria (green), deltaproteobacteria (cyan), cyanobacteria (yellow), firmicutes (magenta), actinobacteria (red), spirochaetes (dark red), bacteroidetes (dark grey), planctomycetes (dark yellow), deinococcus/thermus (dark green) and verrucomicrobia (purple) .

4.4   Discussion

4.4.1   Csp1 and Csp3 protein homologues are present in MOB

The identification of Csp1 and Csp3 protein homologues in the genomes of a range of MOB is consistent, at first inspection, with their high copper requirements as pMMO, the main methane oxidising enzyme, is copper-dependent (2, 6). It appears, therefore, that the presence of a protein involved in copper storage that can possibly supply copper to pMMO is reasonable. However, the presence of Csp1 and/or Csp3 protein homologues in MOB does not appear to follow a clear pattern, while Csp1/Csp3 protein homologues are completely absent from other MOB strains. Representatives of the alphaproteobacterial *Methylosinus* and *Methylocystis* genera (*M. trichosporium* OB3b, *Methylosinus* sp. LW4 and the *Methylocystis* strains Rockwell (ATCC49242) and SC2) (8, 11, 29,) encode both Csp1 and Csp3 protein homologues (Figure 4.1). However, the presence of both proteins cannot be regarded as a general feature of these genera as *Methylocystis rosea* only encodes a Csp1 protein, while *Methylocystis parvus* OBBP does not appear to encode any of the two proteins.

The absence of sMMO encoding genes from the above *Methylocystis* strains (8, 11, 17, 61) could imply higher requirements in copper due to the sole use of pMMO for methane oxidation, compared to switchover MOB that also express the iron-dependent sMMO. The duplicate copies of the pMMO-expressing operon *pmoCAB*, present in MOB genomes (9, 11, 17), need to be expressed simultaneously in the case of *Methylocystis* strain SC2 for maximal growth (20), while the second type of pMMO (*pmoCAB2* operon) expressed by this organism is also present in most strains of the *Methylocystis* and the *Methylosinus* genera (21). The increased use of pMMO for methane oxidation, combined with the presence of different types of pMMO, could suggest that these alphaproteobacterial MOB use Csp1 and the cytosolic Csp3 to meet the copper requirement of the different types of pMMO (20, 22). This possibility is also supported by the significantly different kinetics of Csp1 and Csp3 in Cu(I) release that could be relevant with supplying copper to pMMOs with different methane oxidation kinetics (20). Although this scenario does not explain the complete absence of such homologues in the case of *Methylocystis parvus* OBBP, the absence of both Csp1 or Csp3 protein homologues from *Methylocella silvestris* BL2, the only MOB that does not encode pMMO (13), is consistent with the hypothesis of Csp1 and Csp3 being copper storage proteins for methane oxidation by pMMO.

Csp1 protein homologues were also present in the betaproteobacterium *Methyloversatilis* sp. NVD (5), and also in the methylotrophic alphaproteobacteria *Methylobacterium extroquens* DM4 and *Methylobacterium populi* BJ001. *Methylobacterium populi* BJ001 can grow on methane but also on methanol or methylamine, and also on two-, three- or four-carbon compounds (23). *Methylobacterium extroquens* DM4 on the other hand is unique in its ability to grow on dichloromethane (54). The gammaproteobacteria *Methylobacter tundripaludum* SV96 and *Methylomicrobium album* BG8 were found to encode only Csp3 protein homologues and notably, the genomes of these two organisms have the pMMO-encoding operon *pmoCAB*, but also the *pxmABC* operon, which is thought to encode a copper membrane MMO of unknown substrate specificity (10, 15, 22).

The factors that determine the distribution of Csp1 and/or Csp3 protein homologues in MOB are unclear. Csp1, containing a predicted Tat-leader peptide, is exported folded from the cytosol, where it possibly acquires copper, while Csp3 is expected to be cytosolic, as it does not contain a Tat-leader peptide (1). At high copper to biomass ratios MOB develop intra-cytoplasmic membranes that house pMMO (48) and two different patterns have been observed for this process. MOB in the gammaproteobacteria class develop intra-cytoplasmic membranes throughout the cell, as opposed to alphaproteobacterial MOB where the intra-cytoplasmic membranes are arranged parallel to the cytoplasmic membrane (6, 48). The nature of the intra-cytoplasmic membranes is not understood in terms of whether they form closed vesicles, as is thought for the cyanobacterial thylakoid (43, 47), or are invaginations of the cytoplasmic membrane, however, these differences in the intra-cytoplasmic membrane formation may be related with the preference of MOB for Csp1 or Csp3.

The presence of Csp1 and/or Csp3 protein homologues does not appear to be related with mb production either as *Methylocystis parvus* OBBP, which encoded the precursor peptide MbnA for mb synthesis (27), does not encode Csp1 or Csp3 homologues, in contrast with *Methylomicrobium album* BG8, not known to produce mb, which encodes a Csp3 protein homologue. While Csp1 and Csp3 have been hypothesised to store copper for methane oxidation, the presence of Csp1 and Csp3 protein homologues in a range of MOB with metabolic versatility, such as the strains of the *Methylocystis* genus (8, 11) and the methylotrophic *Methylobacterium* strains (23, 54), as well as the absence

of protein homologues from other MOB, indicates a broader role of Csp1 and Csp3 in copper homeostasis.

4.4.2 Csp1 and Csp3 protein homologues are distributed among bacterial phyla

The distribution of Csp1 and Csp3 protein homologues in a range of organisms representing the major bacterial phyla, (notably in proteobacteria, cyanobacteria actinobacteria, firmicutes, deinococcus/thermus, verrucomicrobia, planctomycetes, bacteroidetes and spirochaetes) (Figure 4.3) (55, 56) implies a broader role of these proteins in copper homeostasis, other than copper storage for potential supply to pMMO in MOB (Figure 4.2 and 4.3). Bioinformatics studies revealed that the vast majority of bacteria consist of copper-utilising organisms, in the sense that they have at least one copper-dependent protein (61, 62). The role of a copper storage protein in bacteria may involve supplying metal to a copper-dependent protein. This could explain the presence of Csp1 or Csp3 in the case of saprotrophic bacteria, rhizobia and denitrifying bacteria with copper requirement for enzymes such as nitrosocyanin, ammonia monooxygenase and copper nitrite reductase (31, 34-36, 44, 57). Another example is cyanobacteria, where plastocyanin is located in the thylakoid and acquires copper from the cytoplasm, which is delivered to the thylakoid through PacS (43-46). Moreover, competition for metal acquisition on the host-pathogen interface is a common route for pathogenicity and also for host-mediated defence against the pathogen (63-65). The presence therefore of a copper storing protein in bacterial pathogens may be critical for survival during infection for these organisms.

While Csp1 is expected to be exported from the cytosol, which may be meaningful in MOB, in which pMMO is localised in the intra-cytoplasmic membranes, the majority of bacteria listed encode Csp3 protein homologues. Based on the wider distribution of the cytosolic Csp3 in bacteria, including organisms from cyanobacteria, actinobacteria and firmicutes, which are thought to be among the most ancient bacterial phyla (62, 66, 67), it is plausible that Csp3 evolved earlier in time than Csp1. The distribution of Csp3 in different bacterial phyla, without however being conserved throughout the organisms of each phylum, may be a result of horizontal gene transfer or gene loss events, which are the easiest way organisms employ to adapt to the environment (31, 66). Csp3 either evolved early in time and, subsequently, the Csp3-encoding gene was lost by organisms that developed to entirely export copper from the cytosol or, in the case Csp3 appeared later in evolution, it was acquired by bacteria with a need for cytosolic copper storage.

The low sequence identity (18%) that Csp1 shares with Csp3 implies Csp1 may have evolved independently in bacteria in which copper storage is needed outside the cytosol (Figure 4.2), whether this means the periplasmic space or potentially other subcellular compartments. The possibility of the evolution of Csp3 earlier in time than Csp1 is also consistent with the suggested common bacterial ancestor being a Gram-positive organism, as a large number of the Csp3 protein homologues are encoded by Gram-positive bacteria (66).

Regardless of the evolutionary events that led to the appearance of Csp1 and Csp3 in bacteria, the presence of Csp3 protein homologues in the cytosol of a large number of bacteria challenges the current notion that bacteria exclusively expel copper from the cytosol (39-42), and implies there may be a need for copper in the bacterial cytosol that has been overlooked until now. The Gram-positive *B. subtilis* senses Cu(I) through CsoR which is thought to regulate the CopZ metallochaperone and the CopA transporter through which Cu(I) efflux is achieved (69-71). A copper laccase (CotA) located on the surface the spore coat of *B. subtilis* has some similarities to multicopper oxidases (72) (Figure 4.4). The presence of Csp3 inside the cytosol of *B. subtilis* that has no known copper requirement inside the cell, potentially implies the existence of unknown cytosolic copper enzymes. This idea is in contrast with the present model according to which copper is exported from the bacterial cytoplasm. Csp1 and Csp3 belong in a novel family of copper storage proteins, that have a well determined structure in the metal-free form, and store Cu(I) ions inside the core of a 4-helix bundle. The presence of Csp3 in the cytosol of a large number of bacteria may imply the existence of copper-utilising proteins in the bacterial cytoplasm, yet to be identified, in which case the current model for bacterial copper homeostasis needs to be redefined.
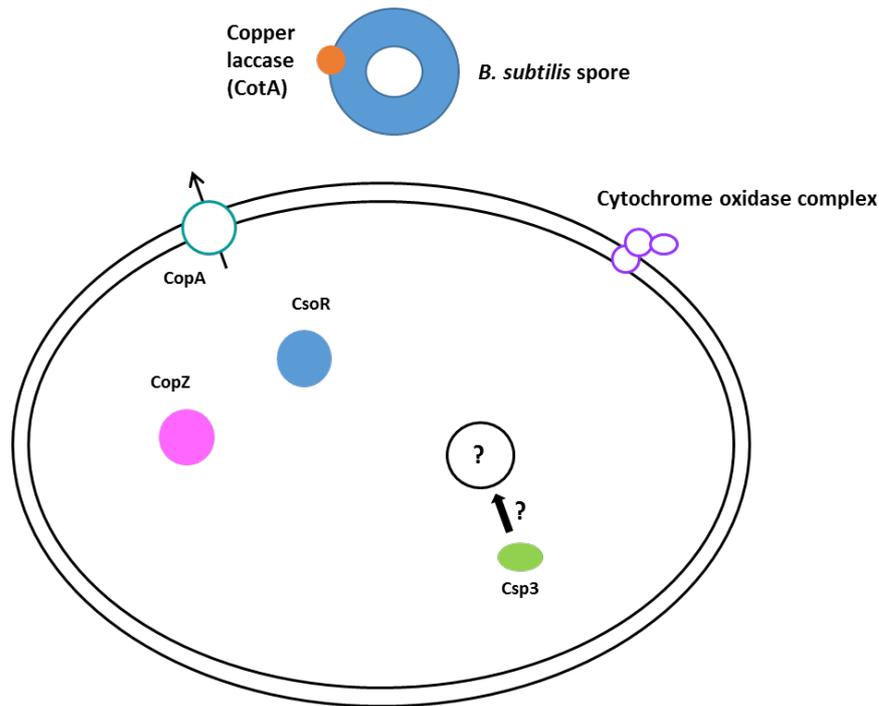
Figure 4.4 Schematic representation of the Gram-positive *B. subtilis*, where the Cu(I) sensor CsoR regulates the expression of the metallochaperone CopZ and the Cu(I) transporter CopA, for Cu(I) export from the cell. Copper requiring proteins in *B. subtilis* are located outside the cytosol. Such examples are the cytochrome *caa₃* complex and the copper laccase (CotA), which is located on the surface of the spore coat of *B. subtilis.* The presence of Csp3 in the cytoplasm of *B.subtilis* may serve as a mechanism to prevent copper-induced toxicity or to provide copper to an unknown cytosolic copper enzyme.

4.5   References

1. Palmer T., Berks B. C., Moving folded proteins across the bacterial cell membrane, Microbiology, 2003, 149, 547-556

2. Balasubramanian R., Kenney G. E., Rosenzweig A. C., Dual pathways for copper uptake by methanotrophic bacteria, J. Biol. Chem., **2011**, 286, 37313-37319

3. Sutherland D. E. K., Stillman M.J., The 'magic' numbers of metallothionein, Metallomics, **2011**, 3, 444-463

4. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

5. Kalyuzhnaya M G., De Marco P., Bowerman S., Pacheco C. C., Lara J. C., Lidstrom M. E., Chistoserdova L., Methyloversatilis universalis gen. nov., sp. nov., a novel taxon within the Betaproteobacteria represented by three methylotrophic isolates, Int. J. Syst. Evol. Micr., **2006**, 56, 2517-2522

6. Hanson R. S., Hanson T. E., Methanotrophic bacteria, Microbiol. Rev., **1996**, 60, 439-471

7. Op den Camp H. J. M., Islam T, Stott M. B., Harhangi H. R., Hynes A., Schouten S., Jetten M. S. M., Birkeland N-K., Pol1 A., Dunfield P. F., Environmental, genomic and taxonomic perspectives on methanotrophic Verrucomicrobia, Environ. Microbiol. Rep., **2009**, 1, 293-306

8. Stein L. Y., Bringel F., Dispirito A. A., Han S., Jetten M. S., Kalyuzhnaya M. G., Kits K. D., Klotz M. G., Op den Camp H. J., Semrau J. D., Vuilleumier S., Bruce D., Cheng J. F., Davenport K. W., Goodwin L., Han S., Hauser L., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Genome sequence of the methanotrophic alphaproteobacterium Methylocystis sp. Strain Rockwell (ATCC49242), J. Bacteriol., **2011**, 193, 2668-2669

9. Ward N., Larsen Ø., Sakwa J., Bruseth L., Khouri H., Durkin A. S., Dimitrov G., Jiang L., Scanlan D., Kang K. H., Lewis M., Nelson K. E., Methé B., Wu M., Heidelberg J. F., Paulsen I. T., Fouts D., Ravel J., Tettelin H., Ren Q., Read T., DeBoy R. T., Seshadri R., Salzberg S. L., Jensen H. B., Birkeland N. K., Nelson W. C., Dodson R. J., Grindhaug S. H., Holt I., Eidhammer I., Jonasen I., Vanaken S., Utterback T., Feldblyum T. V., Fraser C. M., Lillehaug J. R., Eisen J. A., Genomic insights into methanotrophy: the complete genome sequence of Methylococcus capsulatus (Bath), PLoS Biol., **2004**, 2, 1616-1628

10. Kits K. D., Kalyuzhnaya M. G., Klotz M. G., Jetten M. S., Op den Camp H. J., Vuilleumier S., Bringel F., Dispirito A. A., Murrell J. C., Bruce D., Cheng J. F., Copeland A., Goodwin L., Hauser L., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Stein L. Y., Genome sequence of the obligate gammaproteobacterial methanotroph Methylomicrobium album strain BG8, Genome Announc., **2013**, 1

11. Dam B., Dam S., Kube M., Reinhardt R., Liesack W., Complete genome sequence of Methylocystis sp. Strain SC2, an aerobic methanotroph with high-affinity methane oxidation potential, J. Bacteriol., **2012**, 194, 6008-6009

12. Boden R., Cunliffe M., Scanlan J., Moussard H., Kits K. D., Klotz M. G., Jetten M. S., Vuilleumier S., Mikhailova N., Teshima H., Tapia R., Kyprides N., Ivanova N., Pagani I., Cheng J. F., Goodwin L., Han C., Hauser L., Land M. L., Lapidus A., Lucas S., Pitluck S., Woyke T., Stein L., Murrell J. C., Complete genome sequence of the aerobic marine methanotroph Methylomonas methanica MC09, J. Bacteriol., **2011**, 193, 7001-7002

13. Chen Y., Crombie A., Rahman M. T., Debysh S. N., Liesack W., Stott M. B., Alam M., Theisen A. R., Murrell J. C., Dunfield P. F., Complete genome sequence of the aerobic facultative methanotroph Methylocella silvestris BL2, J. Bacteriol., **2010**, 192, 3840-3841

14. Vuilleumier S., Khmelenina V. N., Bringel F., Reshetnikov A. S., Lajus A., Mangenot S., Op den Camp H. J., Jetten M. S., DiSpirito A. A., Dunfield P. F., Klotz M. G., Semrau J. D., Stein L. Y., Barbe V., Médigue C., Trotsenko Y. A., Kalyuzhnaya M. G., Genome sequence for the haloalkaliphilic methanotrophic bacterium Methylomicrobium alcaliphilum 20Z, J. Bacteriol., **2012**, 194, 551-55

15. Svenning M. M., Hestnes A. G., Wartiainen I., Stein L. Y., Klotz M. G., Kalyuzhnaya M. G., Spang A., Bringel F., Vuilleumier S., Lajus A., Médigue C., Bruce D. C., Cheng J. F., Goodwin L., Ivanova N., Han J., Han C. S., Hauser L., Held B., Land M. L., Lapidus A., Lucas S., Nolan M., Pitluck S., Woyke T., Genome sequence of the Arctic methanotroph Methylobacter tundripaludum SV96, J. Bacteriol., **2011**, 193, 6418-6419

16. Kittichotirat W., Good N. M., Hall R., Bringel F., Lajus A., Médigue C., Smalley N. E., Beck D., Bumgarner R., Vuilleumier S., Kalyuzhnaya M. G., Genome sequence of Methyloversatilis universalis FAM5 T, a methylotrophic representative of the order *Rhodocyclales*, J. Bacteriol., **2011**, 193, 4541-4542

17. del Cerro C., García J. M., Rojas A., Tortajada M., Ramón D., Galán B., Prieto M. A., García J. L., Genome Sequence of the Methanotrophic Poly-Hydroxybutyrate Producer Methylocystis parvus OBBP, J. Bacteriol., **2012**, 194, 5709-5710

18. Theisen A. R., Ali M. H., Radayewski S., Dumont M. G., Dunfield P. F., McDonald I. R., Debysh S. N., Miguez C. B., Murrell J. C., Regulation of methane oxidation in the facultative methanotroph Methylocella silvestris BL2, Mol. Microbiol., **2005**, 58, 682-692

19. Debysh S. N., Knief C., Dunfield P. F., Methylocella species are facultatively methanotrophic, J. Bact., **2005**, 187, 4665-4670

20. Baani M., Liesack W., Two isoenzymes of particulate methane monooxygenase with different methane oxidation kinetics are found in Methylocystis sp. Strain SC2, P. Natl. Ac. Sci., **2008**, 105, 10203-10208

21. Tchawa Yimga M., Dunfield P. F., Ricke P., Heyer J., Liesack W., Wide distribution of a novel pmoA-like gene copy among Type II methanotrophs and its expression in Methylocystis strain SC2, Appl. Environ. Microb., **2003**, 69, 5593-5602

22. Tavormina P. L., Orphan V. J., Kalyuzhnaya M. G., Jetten M. S. M., Klotz M. G., A novel family of functional operons encoding methane/ammonia monooxygenase-related proteins in gammaproteobacterial methanotrophs, Envron. Microbiol. Rep., **2011**, 3, 91-100

23. Van Aken B., Peres C. M., Lafferty Doty S., Yoon J. M., Schnoor J. L., Methylobacterium populi sp. Nov., a novel aerobic, pink-pigmented, facultatively methylotrophic, methane-utilising bacterium isolated from poplar trees (*Populus deltoides x nigra* DN34), Int. J. Syst. Evol. Micr., **2004**, 54, 1191-1196

24. Scanlan J., Dumont M. J., Murrell J. C., Involvement of MmoR and MmoG in transcriptional activation of soluble methane monooxygenase genes in Methylosinus trichosporium OB3b, FEMS Microbiol. Lett., **2009**, 301, 181-187

25. Stein L. Y., Arp D. J., Berube P. M., Chain P. S. G., Hauser L., Jetten M. S. M., Klotz M. G., Larimer F. W., Norton J. M., Op den Camp H. J. M., Shin M., Wei X., Whole-genome analysis of the ammonia-oxidising bacterium Nitrosomonas eutropha C91: implications for niche adaptation, Environ. Microbiol., **2007**, 9, 2993-3007

26. Krentz B. D., Mulheron H. J., Semrau J. D., DiSpirito A. A., Bandow N. L., Haft D. H., Vuilleumier S., Murrell J. C., McEllistrem M. T., Hartsel S. C., Gallagher

W. H., A comparison of methanobactins from Methylosinus trichosporium OB3b and *Methylocystis* strain SB2 predicts methanobactins are synthesized from diverse peptide precursors modified to create a common core for binding and reducing copper ions, Biochemistry, **2010**, 49, 10117-10130

27. Kenney G. C., Rosenzweig A. C., Genome mining for methanobactins, BMC Biology, **2013**, 11

28. Campbell M. A, Nyerges G., Kozlowski J. A., Poret-Peterson A. T., Stein L. Y., Klotz M. G., Model of the molecular basis for hydroxylamine oxidation and nitrous oxide production in methanotrophic bacteria, FEMS Microbiol. Lett., **2011**, 322, 82-89

29. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., **2010**, 192, 6497-6498

30. Auman A. J., Speake C. C., Lidstrom M. E., *nif*H sequences and nitrogen fixation in Type I and Type II methanotrophs, Appl. Environ. Microbiol, **2001**, 67, 4009-4016

31. Klotz M. G., Stein L. Y., Nitrifier genomics and evolution of the nitrogen cycle, FEMS Microbiol. Lett., **2008**, 278, 146-156

32. Nyerges G., Stein L. Y., Ammonia cometabolism and product inhibition vary considerably among species of methanotrophic bacteria, FEMS Microbiol. Lett., **2009**, 297, 131-136

33. Elmore B. O., Bergmann D. J., Klotz M. G., Hooper A. B., Cytochromes P460 and *c'*-beta; a new family of high-spin cytochromes *c*, FEBS Lett., **2007**, 581, 911-916

34. Klotz M. G., Schmid M. C., Strous M., Op den Camp H. J. M., Jetten M. S. M., Hooper A. B., Evolution of an octahaem cytochrome c protein family that is key to aerobic and anaerobic ammonia oxidation by bacteria, Environ. Microbiol., **2008**, 10, 3150-3163

35. Jason J., Cantera L., Stein L. Y., Molecular diversity of nitrite reductase genes (*nirK*) in nitrifying bacteria, Environ. Microbiol., **2007**, 9, 765-776

36. Raymond J., Siefert J. L., Staples C. R., Blankenship R. E., The natural history of nitrogen fixation, Mol. Biol. Evol., **2004**, 21, 541-554

37. Stein L. Y., Klotz M. G., Nitrifying and denitrifying pathways of methanotrophic bacteria, Bioc. Soc. Trans., **2011**, 39, 1826-1831

38. Ridge P. G., Zhang Y., Gladyshev V. N., Comparative genomic analyses of copper transporters and cuproproteomes reveal evolutionary dynamics of copper utilisation and its link to oxygen, PLoS ONE, **2008**, 1

39. Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V., Mondragón A., Molecular basis of the metal- ion selectivity and zeptomolar sensitivity by CueR, Science, **2003**, 301, 1383-1387

40. Radford D. S., Kihlken M. A., Borrelly G. P. M., Harwood C. R., Le Brun N. E., Cavet J. S., CopZ from *Bacillus subtilis* interacts in vivo with a copper exporting CPx-type ATPase CopA, FEMS Microbiol. Lett., **2003**, 220, 105-112

41. Outten F. W., Huffman D. L., Hale J. A., O'Halloran T. V., The independent cue and cus systems confer copper tolerance during aerobic and anaerobic growth in Escherichia coli, J. Biol. Chem., **2001**, 276, 30670-30677

42. Singh S. K., Roberts S. A., McDevitt S. F., Weichsel A., Wildner G. F., Grass G. B., Rensing C., Montfort W. R., Crystal structures of multicopper oxidase CueO bound to copper (I) and silver (I): functional role of a methionine-rich sequence, J. Biol. Chem., **2011**, 286, 37849-37857

43. Tottey S., Harvie D. R., Robinson N. J., Understanding how cells allocate metals using metal sensors and metallochaperones, Acc. Chem. Res, **2005**, 38, 775-783

44. Badarau A., Firbank S. J., McCarthy A. A., Banfield M. J., Dennison C., Visualizing the metal-binding versatility of copper trafficking sites, Biochemistry, **2010**, 49, 7798-7810

45. Banci L., Bertini I., Ciofi-Baffoni S., Su X. C., Borrelly G. P., Robinson N. J., Solution structures of a cyanobacterial metallochaperone: Insight into an atypical copper-binding motif, JBC, 2004

46. Borrelly G. P., Blindauer C. A., Schmid R., Butler C. S., Coper C. E., Harvey I., Sadler P. J., Robinson N. J., A novel copper site in a cyanobacterial chaperone, Biochem. J., 2004

47. Van de Meene A. M. L., Hohmann-Marriott M. F., Vermaas W. F. J., Roberson R. W., The three-dimensional structure of the cyanobacterium Synechocystis sp. PCC 6803, Arch. Micobiol., **2006**, 184, 292-270

48. Davies S. L., Whittenbury R., Fine structure of methane and other hydrocarbon-utilising bacteria, J. Gen. Microbiol., **1970**, 61, 227-232

49. Tucker J. D., Siebert C. A., Escalante M., Adams P. G., Olsen J. D., Otto C., Stokes D. L., Hunter C. N., Membrane invagination in Rhodobacter sphaeroides is initiated at curved regions of the cytoplasmic membrane, then forms both budded and fully detached spherical vesicles, Mol. Microbiol., **2010**, 76, 833-847

50. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T. L., Software BLAST+: architecture and applications, BMC Bioinformatics, **2009**, 10, 421

51. Petersen T. N., Brunak S., von Heijne G., Nielsen H., SignalP 4.0: discriminating signal peptides from transmembrane regions**,** Nature Methods, **2011**, 8, 785-786

52. Bendtsen J. D., Nielsen H., Widdick D., Palmer T., Brunak S., Prediction of twin-arginine signal peptides, BMC Bioinformatics, **2005**, 6

53. Sievers F., Wilm A., Dineen D. G., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D., Higgins D., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol., **2011**, 7

54. Muller E. E. L., Hourcade E., Louhichi-Jelail Y., Hammann P., Vuilleumier S., Bringel F., Functional genomics of dichloromethane utilisation in Methylobacterium extroquens DM4, Environ. Microbiol, **2011**, 13, 2518-2535

55. Sayers E. W., Barrett T., Benson D. A., Bryant S. H., Canese K., Chetvernin V., Church D. M., DiCuccio M., Edgar R, Federhen S, Feolo M, Geer LY, Helmberg W, Kapustin Y, Landsman. D, Lipman D. J., Madden T. L., Maglott D. R., Miller V., Mizrachi I., Ostell J., Pruitt K. D., Schuler G. D., Sequeira E., Sherry S. T., Shumway M., Sirotkin K., Souvorov A., Starchenko G., Tatusova T. A., Wagner L., Yaschenko E., Ye J., Database resources of the National Center for Biotechnology Information, **2009**, Nucleic Acids Res., Database issue D5-15

56. Benson D. A., Karsch-Mizrachi I., Lipman D. J., Ostell J., Sayers E. W., GenBank, Nucleic Acids Res., **2009**, Database issue D26-31

57. Moreno R., Rojo F., The contribution of proteomics to the unveiling of the survival strategies used by Pseudomonas putida in changing and hostile environments, Proteomics, **2013**, 13, 2822-2830

58. Cànovas D., Cases I., de Lorenzo V., Heavy metal tolerans and metal homeostasis in Pseudomonas putida as revealed by complete genome analysis, Environ. Microbiol., **2003**, 5, 1242-1256

59. Deshpande C. N., McGrath A. P., Font J., Guilfoyle A. P., Maher M. J., Jormakka M., Structure of an atypical FeoB G-domain reveals a putative domain-swapped dimer, Struc. Comm., **2013**, 69, 399-404

60. Glöckner F. O., Babenzien H-D., Amann R., Phylogeny and identification in situ of Nevskia ramosa, Appl. Environ. Microbiol., **1998**, 64, 1895-1901

61. Andreini C., Banci L., Bertini I., Rosato A., Occurrence of copper proteins through the three domains of life: a bioinformatics approach, J. Proteome Res., **2008**, 7, 209-216

62. Ridge P. G., Zhang Y., Gladyshev V. N., Comparative genomic analyses of copper transporters and cuproproteomes reveal evolutionary dynamics of copper utilisation and its link to oxygen, PLoS ONE, **2008**, 1

63. Rowland J. L., Niederweis M., Resistance mechanisms of Mycobacterium tuberculosis against phagosomal copper overload, Tuberculosis, 2012, 92, 202-210

64. Chaturvedi J. P., Henderson J. P., Pathogenic adaptations to host-derived antibacterial copper, Front. Cell. Infect. Microbiol., **2014**, 4, 1-12

65. Chaturvedi J. P., Hung C. S., Giblin D. E., Urushidani S., Austin A. M., Dinauer M. C., Henderson J. P., Cupric yersiniabactin is a virulence-associated superoxide dismutase mimic, ACS Chem. Biol., **2014**, ACS Chem. Biol., 9, 551-561

66. Ciccarelli F. D., Doerks T., von Mering C., Creevey C. J., Snel B., Bork P., Toward automatic reconstruction of a highly resolved tree of life, Science, **2006**, 311, 1283-1286

67. Ventura M., Canchaya C., Tauch A., Chandra G., Fitzgerald G. F., Chater K. F., van Sinderen D., Genomics of actinobacteria: tracing the evolutionary history of an ancient phylum, Microbiol. Mol. Biol. Rev., **2007**, 71, 495-548

68. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

69. Ma Z., Cowart D. M., Scott R. A., Giedroc DP, Molecular insights into the metal selectivity of the Cu(I)-sensing repressor CsoR from *Bacillus subtilis*, Biochemistry, **2009**, 48, 3325-3334

70. Kihlken M. A., Leech A. P., Le Brun N. E., Copper-mediated dimerization of CopZ, a predicted copper chaperone from Bacillus subtilis, Biochem. J., **2002**, 368, 729-739

71. Radford D. S., Kihlken M. A., Borrelly G. P., Harwood C. R., LeBrun N. E., Cavet J. S., CopZ from Bacillus subtilis interacts in vivo with a copper exporting CPx-type ATPase CopA, FEMS Microbiol. Lett., 2003, 220, 105-112

72. Hullo M-F., Moszer i., Danchin A., Martin-Verstraete I., CotA from Bacillus subtilis is a copper-dependent laccase, J. Bacteriol., **2001**, 183, 5426-5430

CHAPTER 5:

Discussion

5.1 Identification of Csp1 from an abundance of soluble copper pools in *M. trichosporium* OB3b

MOB have an unusually high requirement in copper due to the use of pMMO, the main methane oxidising enzyme with copper-dependent activity that is believed to contain copper in its active sites (1-3). In order to satisfy the high need for copper, MOB possess a unique copper uptake system which includes mb, a small peptide secreted in the environment for copper sequestration (1, 2). Mb is well-characterised in terms of its copper binding stoichiometry and affinity (4, 5) and has been visualised inside the cytoplasm of MOB (6), where it is believed to mediate the copper-regulated switchover between the expression of sMMO, the soluble iron-dependent form of MMO, and pMMO (5, 7). pMMO is localised in intra-cytoplasmic membranes that form at high copper to biomass ratios (1, 2), however it is unclear whether these membranes are invaginations of the cytoplasmic membrane or form separate subcellular compartments (8). The mechanism through which pMMO acquires the copper necessary for enzyme activity is unknown, especially since MOB possess protein machinery dedicated to efflux of Cu(I) from the cytosol, such as the efflux pump CopA, the regulator CueR and the chaperone CopZ, all of which are well-characterised in other bacteria (9-16). It is possible that protein components, yet to be identified, participate in the delivery of copper from mb to pMMO and it is this question that the current project aimed to investigate.

In an attempt to identify soluble copper proteins to which mb may deliver copper, a metalloproteomics approach was used to analyse the soluble extract of *M. trichosporium* OB3b, a 'switchover' organism capable of expressing both sMMO and pMMO. The abundance of soluble copper pools visualised from the cell profiles of *M. trichosporium* OB3b was surprising, seen as the only known copper proteins present in the organism are CopA, CopZ and CueR (17). The most abundant soluble copper pool eluted consistently in the 200 mM NaCl fraction from anion-exchange chromatography (Figures 2.2 B, 2.6 A, 2.8 and 2.12) and the copper peak from further resolution on a size-exclusion column contained a large number of proteins (Figures 2.7 and 2.9). While soluble copper pools were present in all the anion-exchange fractions obtained (Figure 2.6), the better resolution of the copper peak eluting in 200 mM NaCl appeared to be promising for subsequent optimisation of the experiment. The improved resolution of the protein content of this soluble copper pool, achieved by introducing a linear NaCl

gradient in the anion-exchange chromatography (Figures 2.8 and 2.12), resulted in the identification of Csp1. The identity of the protein was also confirmed in an optimised experiment, performed on sMMO-inactive (therefore pMMO-active) *M. trichosporium* OB3b, from which a Csp1-containing fraction of higher purity was obtained and was estimated to contain significant amounts of copper (Figure 2.14).

The isolation and identification of copper-loaded Csp1 from a copper-containing fraction, as the only protein band with intensity correlating to the copper concentration profile, in addition to the fact that *M. trichosporium* OB3b was grown under high levels of copper and was tested for sMMO inactivity, led to the hypothesis of the protein being involved in copper homeostasis. The copper-responsive expression of Csp1 is also in agreement with this hypothesis (Figure 2.15). While Csp1 was identified from the native organism as a copper protein, bioinformatics revealed the presence of two more protein homologues in *M. trichosporium* OB3b, Csp2 and Csp3. Notably, Csp1 and Csp2 contain a predicted Tat-leader peptide, responsible for the export of folded proteins from the cytosol (18), while Csp3 does not have a Tat-leader and is therefore expected to be cytosolic.

5.2 Csp1 and Csp3 represent a novel family of copper storage proteins that store Cu(I) inside a 4-helix bundle

The *in vitro* characterisation of recombinant Csp1 and Csp3 revealed the high Cu(I) binding capacity of these proteins, at 13 and 18 equivalents of Cu(I) per monomer, respectively, that is consistent with the number of Cys residues for each protein. Both Csp1 and Csp3 are tetramers in solution, further supporting their high Cu(I) binding capacity, which translates in 56 and 72 Cu(I) ions, respectively. The defining feature of these proteins is that Cu(I) is stored inside a 4-helix bundle, and this protein fold is highly ordered in the apo-forms of the proteins. This property, supported by CD and crystallographic data showing Cu(I) binding does not effect the secondary structure of the protein (Figures 3.3, 3.4, 3.23 and 3.24), differentiates Csp1 and Csp3 from MTs, where metal binding drives protein folding around the metal clusters (19, 20). In another family of iron-storage proteins, ferritins, iron is deposited as FeO(OH) in the center of a shell formed by 24 4-helix bundles, but notably not inside the core of the 4-helix bundle (21-23). The crystal structures for fully loaded $Cu(I)_{13}$-Csp1 and $Cu(I)_{18}$-Csp3 show the Cu(I) ions span the core of the 4-helix bundle with all the Cys residues of each protein participating in coordination of the metal (Figures 3.27 and 3.32). Notably, in many

cases Cys residues from Cys-X-X-X-Cys motifs are involved in binding of the same Cu(I) site. Taken together, the established secondary structure of the protein in its apo-form and the binding of Cu(I) inside the core of a 4-helix bundle, establish that Csp1 and Csp3 belong to a novel family of copper storage proteins.

Except for the different Cu(I) binding capacity, directly related to the number of Cys residues they contain, Csp1 and Csp3 differ in the kinetics, and therefore most likely in the mechanism of Cu(I) release. This property may indicate different physiological roles, which could be related to the different cellular localization of Csp1 (exported from the cytosol) and Csp3 (cytosolic). Csp1, has been shown to bind tightly at least 10 Cu(I) ions (Figure 3.9) and readily releases effectively all Cu(I) within an hour in the presence of an excess of BCS (Figure 2.17), while for Csp3, that has an average Cu(I) binding affinity at the range of 10^{17} M^{-1} (Figures 3.14 and 3.16), this reaction is extremely slow (Figure 2.18). The same trend in Cu(I) transfer from Cu(I)-Csp1 and Cu(I)-Csp3 was also observed in the presence of mb, which has a Cu(I) affinity at the range of 10^{21} M^{-1} (4, 5). Apo-mb removes all Cu(I) from Cu(I)-Csp1 within an hour (Figure 3.19), while the same reaction with Cu(I)-Csp3 takes over two weeks (Figure 3.21). On the other hand, Cu(I) transfer from Cu(I)-mb towards either apo-Csp1 or Csp3 does not occur, indicating that apo-Csp1 or apo-Csp3 do not mediate copper transport from internalized Cu(I)-mb to pMMO (Figures 3.20 and 3.22). It is however possible that apo-mb, also localised in the cytoplasm after uptake (6), extracts Cu(I) from Cu(I)-loaded Csp1 or Csp3, in order to provide the metal to pMMO.

5.3   Csp1 stores Cu(I) for pMMO activity

The main target for copper delivery in *M. trichosporium* OB3b is pMMO is housed in intra-cytoplasmic membranes (2, 6). The exact nature of the intra-cytoplasmic membranes, where pMMO is located, is not clear (8). The intra-cytoplasmic membranes may either be invaginations of the cytoplasmic membrane or form separate vesicles in the cytoplasm. These two possibilities are directly relevant to the localisation of Csp1 and Csp2, both of which are exported from the cytosol (due to the presence of predicted Tat-leader peptide in their sequence). If the the two proteins have discrete physiological roles, Csp1 may be exported to the intra-cytoplasmic membranes where pMMO is located to act as a copper storage/supply protein for pMMO, and Csp2 may be exported for copper storage to the periplasm, or vice versa (Figure 5.1). Alternatively, Csp1 and Csp2 may be redundant and both proteins may be exported from the cytosol to store

copper for use by pMMO. Based on the fast rate of Cu(I) exchange between Cu(I)-Csp1 and apo-mb, it is possible that apo-mb extracts Cu(I) from Csp1 and supplies the metal to pMMO.
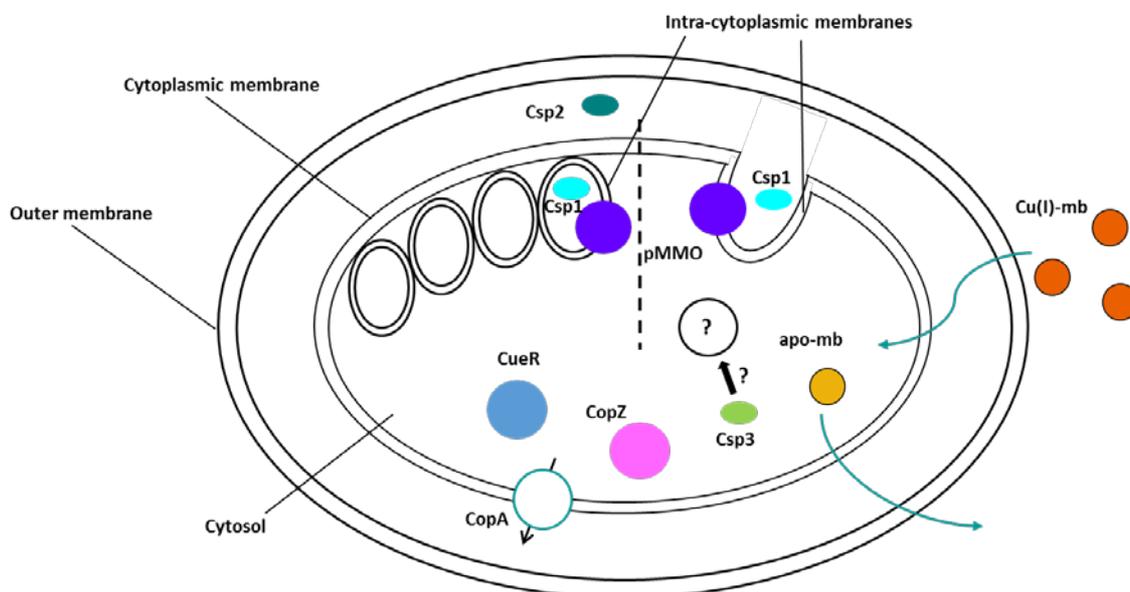
Figure 5.1 Schematic representation of the suggested localisation and function of Csp1, Csp2 and Csp3 in a MOB. Cu(I)-mb mediates copper uptake for the needs of the copper-containing pMMO and is thought to be internalised in the cytoplasm. pMMO is housed in intra-cytoplasmic membranes which may form separate subcellular compartments (shown on the left side of the dashed line), or may form as a result of invaginations of the cytoplasmic membrane (right side of dashed line). Csp1 may be exported from the cytosol to the intra-cytoplasmic membranes to provide copper storage for pMMO, while Csp2 stores copper in the periplasm, or vice versa. Csp3 is a cytosolic copper storage protein that acts either for copper sequestration to prevent copper-induced toxicity or provides copper to an unknown cytosolic copper enzyme.

The involvement of Csp1 and/or Csp2 in supplying copper to pMMO is supported by experiments performed with the double deletion mutant of *M. trichosporium* OB3b lacking the genes responsible for the expression of Csp1 and Csp2 (Δcsp1/csp2) (Prof. Christopher Dennison, Prof. Colin Murrell and co-workers, pending publication). The rate at which the organism switched from using pMMO to using the iron-dependent sMMO, when the cells were starved from copper, was compared between the wild type

and Δ*csp1/csp2 M. trichosporium* OB3b. Δ*csp1/csp2* was found to switch to using sMMO for methane oxidation faster, compared to the wild-type cells. This result indicates that Csp1 and Csp2, which likely acquire Cu(I) in the reducing environment of the cytosol before export, provide a store of Cu(I) that is available to maintain pMMO activity. When the availability of copper in the environment is low, *M. trichosporium* OB3b is able to use pMMO for longer, before switching to sMMO, by supplying the enzyme with copper from Csp1 and /or Csp2. Csp3 on the other hand, remains in the cytosol where it either sequesters copper acting as a defence mechanism against copper-induced toxicity, similar to the role of MTs (19, 24), or it functions as a copper storage/supply protein for the needs of copper enzymes yet to be identified. Notably, some of the soluble copper pools visualised in cell profiles of *M. trichosporium* OB3b (Figures 2.2 and 2.6), may represent cytosolic proteins which would justify a requirement for cytosolic copper that has been overlooked until now.

5.4   Future work

 The identification of native Csp1 as a copper protein from *M. trichosporium* OB3b and the initial characterisation of recombinant Csp1 and Csp3, homologues of which are widespread among bacteria, open a new chapter in bacterial copper homeostasis and point towards further research. While Csp1 was identified from *M. trichosporium* OB3b grown under high copper concentrations, a detailed study of the conditions that induce Csp1 expression is missing, as is the case for Csp2 and Csp3 that were not isolated from the native organism. The experiment comparing cell profiles of sMMO-active and sMMO-inactive *M. trichosporium* OB3b indicates that Csp1 expression depends on the copper concentration of the growth media (Figure 2.15). Csp2 and Csp3 were not identified from the cell profiles of *M. trichosporium* OB3b, either because they were not present in the soluble fraction of the organism under the growth conditions used or due to poor resolution possibly combined low abundance. In that respect reverse transcription quantitatve PCR could be used to monitor the mRNA transcript levels of Csp1, Csp2 and Csp3 from cultures exposed to varying copper concentrations, in order to determine the optimal copper conditions under which the proteins are expressed.

 A comparison of cell profiles of *M. trichosporium* OB3b grown at the optimal copper condition, determined from the previous experiment, and the organism grown in the absence of copper may provide information on the soluble copper pool corresponding to Csp2 or Csp3. Alternatively, cell profiles of the wild type *M. trichosporium* OB3b and

the deletion mutant (Δ*csp2* or Δ*csp3*) grown under the same optimal copper condition would also confirm the soluble copper pool corresponding to Csp2 or Csp3. In the absence of copper from the growth media, this soluble copper pool would be expected to decrease in abundance, along with the intensity of the protein band corresponding Csp2 or Csp3. Subsequent optimisation of the metalloproteomic experiment to improve the resolution of the soluble copper pool of interest would then result in the identification of Csp2 or Csp3 from native *M. trichosporium* OB3b. Studying the conditions under which Csp1, Csp2 and Csp3 are expressed in *M. trichosporium* OB3b would be a first step towards understanding how MOB respond to copper stress.

Csp1 and Csp2 are exported from the cytosol and are potentially redundant in their physiological role. Both proteins have 13 Cys residues and, it would be reasonable to assume that also Csp2 binds 13 Cu(I) ions. Investigating the phenotypes of single deletion mutants Δ*csp1* and Δ*csp2* would provide information on the redundancy of the physiological role of Csp1 and Csp2 in providing Cu(I) for pMMO activity. Moreover, studying the phenotype of the Δ*csp3* mutant and a comparison of the profile of soluble copper pools to those of the wild type organism, may shed light on the cytosolic copper protein that Csp3 stores Cu(I) for. Additionally, fluorescent labelling of the proteins would enable us to visualise them inside the organism, and shed more light on whether their physiological roles are distinct.

Further characterisation if Csp1 and Csp3, but also Csp2, i*n vitro* will progress our understanding of these proteins. While an average Cu(I) binding affinity has been determined for Csp3, the crystal structure of Cu(I)$_2$-Csp3 indicates the Cu(I) sites with the highest affinity, situated in the center of the molecule (Figure 3.27). An Asn is involved in Cu(I) coordination in one of these sites and possibly regulates the mechanism of Cu(I) binding/release. By using site-directed mutagenesis on Csp1 and Csp3, aiming to replace the residues we know to be involved in Cu(I) binding, the Cu(I) affinity of each site could be determined. Studying the average Cu(I) affinity of the molecule in each case, would also offer insight in whether Cu(I) binding occurs synergistically. Obtaining crystal structures for the Cu(I)-loaded mutants would additionally help understand how Cu(I) binding occurs and is regulated and potentially explain why fluorescence, indicative of Cu(I) cluster formation in Csp3 (20), peaks when the protein is semi-loaded and subsequently decreases (Figures 3.10 and 3.11). Finally, *in vitro* characterisation of Csp1/ Csp3 protein homologues from other

organisms, combined as for example recombinant Csp3 from Gram-positive *Bacillus subtilis*, combined with metalloproteomics experiments, will help understand the role of a cytosolic copper storage protein.

5.5 References

1. Hanson R. S., Hanson T. E., Methanotrophic bacteria, Microbiol. Rev., **1996**, 60, 439-471

2. Hakemian A. S., Rosenzweig A. C., The biochemistry of methane oxidation, Annu. Rev. Biochem., **2007**. 76, 223–241

3. Culpepper M. A., Rosenzweig A. C., Architecture and active site of particulate methane monooxygenase, Crit. Rev. Biochem. Mol. Biol., **2012**, 47, 483-492

4. El Ghazouani A., Baslé A., Firbank S. J., Knapp C. W., Gray J., Graham D. W., Dennison C., Copper-binding properties and structures of methanobactins from Methylosinus trichosporium OB3b, Inorg. Chem., **2011**, 50, 1378-1391

5. El Ghazouani A., Baslé A., Gray J., Graham D. W., Firbank S. J., Dennison C., Variations in methanobactin structure influence copper utilization by methane-oxidising bacteria, Proc. Natl. Acad. Sci., **2012**, 109, 8400-8404

6. Balasubramanian R., Kenney G. E., Rosenzweig A. C., Dual pathways for copper uptake by methanotrophic bacteria, J. Biol. Chem., **2011**, 286, 37313-37319

7. Semrau J. D., Jagadevan S., DiSpirito A. A., Khalifa A., Scanlan J., Bergman B. H., Freemeler B. C., Baral B. S., Bandow N. L., Vorobev A., Haft D. H., Vuilleumier S., Murrell J. C., Methanobactin and MmoD work in concert to act as the 'copper-switch' in methanotrophs, Environ. Microbiol., **2013**, 15, 3077-3086

8. Davies S. L., Whittenbury R., Fine structure of methane and other hydrocarbon-utilising bacteria, J. Gen. Microbiol., **1970**, 61, 227-232

9. Changela A., Chen K., Xue Y., Holschen J., Outten C. E., O'Halloran T. V., Mondragón A., Molecular basis of the metal- ion selectivity and zeptomolar sensitivity by CueR, Science, **2003**, 301, 1383-1387

10. Singh S. K., Roberts S. A., McDevitt S. F., Weichsel A., Wildner G. F., Grass G. B., Rensing C., Montfort W. R., Crystal structures of multicopper oxidase CueO bound to copper (I) and silver (I): functional role of a methionine-rich sequence, J. Biol. Chem., **2011**, 286, 37849-37857

11. Radford D. S., Kihlken M. A., Borrelly G. P. M., Harwood C. R., Le Brun N. E., Cavet J. S., CopZ from *Bacillus subtilis* interacts in vivo with a copper exporting CPx-type ATPase CopA, FEMS Microbiol. Lett., **2003**, 220, 105-112

12. Badarau A., Dennison C., Thermodynamics of copper and zinc distribution in the cyanobacterium Synechocystis PCC 6803, Proc. Natl. Acad. Sci., **2011**, 108, 13007-13012

13. Zhou L., Singleton C., LeBrun N. E., High Cu(I) and low proton affinities of the CXXC motif of Bacillus subtilis CopZ, Biochemistry, **2008**, 413, 459-465

14. Xiao Z., Brose J., Schimo S., Ackland S. M., La Fontaine S., Wedd A. G., Unification of the copper(I) binding affinities of the metallo-chaperones Atx1, Atox1, and related proteins: detection probes and affinity standards, J. Biol. Chem., **2011**, 286, 11047-11055

15. Xiao Z., Loughlin F., George G. N., Howlett G. J., Wedd A. G., C-Terminal Domain of the Membrane Copper Transporter Ctr1 from Saccharomyces cerevisiae Binds Four Cu(I) Ions as a Cuprous-Thiolate Polynuclear Cluster: Sub-femtomolar Cu(I) Affinity of Three Proteins Involved in Copper Trafficking, J. Am. Chem. Soc., **2004**, 126, 3081-3090

16. Badarau A., Dennison C., Copper trafficking mechanism of CXXC-containing domains: Insight from the pH dependence of their Cu(I) affinities, J. Am. Chem. Soc., **2011**, 133, 2983-2988

17. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., **2010**, 192, 6497-6498

18. Palmer T., Berks B. C., Moving folded proteins across the bacterial cell membrane, Microbiology, **2003**, 149, 547-556

19. Sutherland D. E. K., Stillman M.J., The 'magic' numbers of metallothionein, Metallomics, **2011**, 3, 444-463

20. Otvos J. D., Armitage I. M., Structure of the metal clusters in rabbit liver metallothionein, Proc. Natl. Acad. Sci., **1980**, 77, 7094-7098

21. Yang X., Le Brun N. E., Thomson A. J., Moore G. R., Chasteen N. D., The iron oxidation and hydrolysis chemistry of Escherichia coli bacterioferritin,

Biochemistry, **2000**, 39, 4915-4923

22. Baaghil S., Lewin A., Moore G. R., Le Brun N. E., Core formation in Escherichia coli bacterioferritin requires a functional ferroxidase center, Biochemistry, **2003**, 42, 14047-14056

23. Harrison P. M., The structure and function of ferritin, **1986**, Biochem. Educ., 14, 154-162

24. Gold B., Deng H., Bryk R., Vargas D., Eliezer D., Roberts J., Jiang X., Nathan C., Identification of a copper-binding metallothionein in pathogenic mycobacteria, Nat. Chem. Biol., **2008**, 4, 609-616

CHAPTER 6:

Materials and methods

6.1 Buffer solutions

All buffer solutions were prepared in Milli-Q water filtered through a Millipore 'Simplicity' water purification system with a sensitivity of 18 MΩ cm. For all solutions pH was adjusted with either HCl (BDH, AnalR) or NaOH (Sigma).

6.1.1 4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid buffer

4-(2-Hydroxyethyl)piperazine-1-ethanesulfonic acid (Hepes) (VWR) buffer was used for washing pellets of *M. trichosporium* OB3b, for cell profiles, storage of purified recombinant protein, and analytical gel-filtration chromatography. pH was adjusted at the range of 7.0 to 8.0.

6.1.2 Tris(hydroxymethyl)aminomethane buffer

Tris(hydroxymethyl)aminomethane (Tris) (Sigma) buffer was used at a pH range of 7.0 to 8.5. This buffer was used for electrophoresis and purification of recombinant protein.

6.1.3 2-(N-morpholino)ethanesulfonic acid buffer

2-(N-morpholino)ethanesulfonic acid (Mes) (Sigma) buffer was used for Cu(I) titrations in the presence of BCA at pH 6.5.

6.1.4 N-tris(Hydroxymethyl)methyl-3-aminopropanesulfonic acid sodium-potassium salt buffer

N-tris(Hydroxymethyl)methyl-3-aminopropanesulfonic acid sodium-potassium salt buffer (Sigma) was used for Cu(I) titrations in the presence of BCA at pH 8.5.

6.1.5 Phosphate buffer

Phosphate buffer was used to buffer solutions at pH 6.0-8.0. The buffer was prepared by mixing appropriate volumes of 1M K_2HPO_4 and KH_2PO_4 (Fluka). Dibasic and monobasic phosphates were combined in the volumes shown in the table 5.1

pH	K$_2$HPO$_4$ (mL)	KH$_2$PO$_4$ (mL)
6.0	13.2	86.8
6.2	19.2	80.8
6.4	27.8	72.2
6.6	38.1	61.9
6.8	49.7	50.3
7.0	61.5	38.5
7.2	71.7	28.3
7.4	80.2	19.8
7.6	86.6	13.4
7.8	90.8	9.2
8.0	94.0	6.0

Table 5.1 Volumes of 1 M K$_2$HPO$_4$ and KH$_2$PO$_4$ required for the preparation of 100 ml phosphate buffer at 1 M and at pH range of 6.0 to 8.0.

6.2 Growth media

All growth media were prepared in deionised water and sterilised by autoclaving at 121 ˚C for 30 mins.

6.2.1 Luria-Bertani medium

Luria-Bertani (LB) medium consisted of 10 g/L tryptone (Melford), 10 g/L NaCl (Sigma) and 5g/L yeast extract (Melford). The LB-agar solid medium consisted of LB supplemented with 15 g/L agar (Melford). The medium was used to grow *E. coli* strains for molecular cloning and protein expression, typically in the presence of the appropriate antibiotic (100 μg/ml ampicillin or 50 μg/ml kanamycin).

6.2.2 Nitrate minimal salts medium

Nitrate minimal salts (NMS) medium was prepared from stock solutions (100-fold for A), B) and D) and 1000-fold for C)) containing the chemicals summarised in Table 5.2. NMS medium was prepared in 1 L bottles by appropriately mixing phosphate buffer pH 7 (solution A)), salts solution (B) and metals solution (C). The iron solution (D) was filter-sterilised directly into the culture (1).

Final solutions used for NMS media			
A) Phosphate Buffer pH 7	B) Salts Solution	C) Metals Solution	D) Iron Solution
3.9 mM KH_2PO_4 6 mM Na_2HPO_4	10 mM $NaNO_3$ 0.3 mM $MgSO_4$ 60 mM $CaCl_2$ 1 mM K_2SO_4	0.5 µM Na_2MoO_4 1.7 µM $MnSO_4$ 0.4 µM $CoCl_2$ 2 µM $ZnSO_4$ 1 µM KI 2 µM H_3BO_3	10 µM $FeSO_4$

Table 5.1 Concentrations of chemicals used for NMS media

6.3   Culturing *M. trichosporium* OB3b

6.3.1   Starting cultures of *M. trichosporium* OB3b from glycerol socks

M. trichosporium OB3b pre-cultures were started by diluting a glycerol stock into approximately 30 ml of NMS medium, supplemented with 10 µM Fe(II), in a 100 ml sealed vile saturated with methane. The vile was stored horizontally in a shaking incubator (120 rpm) at 27 ºC. When the culture started growing, it was transferred in a sealed 1 L bottle containing approximately 300 ml of NMS medium supplemented with 10 µM Fe(II) and saturated with methane, which was stored as mentioned above. In order to start a culture in a fermentor, the content of a 1 L culture bottle was added in the fermentor already containing NMS media (2 L or 4 L for the 3 and 5 L fermentors, respectively). Stirring was set at 250 rpm and 1 µM Cu(II) was routinely added in order to accelerate growth. A filtered mixture of methane (BOC, research grade) and air was supplied to the culture with flow rates of 10-20 ml/min and 1 L/min respectively (1). The growth of the cells was monitored by measuring the optical density at 600 nm (OD_{600nm}).

6.3.2 Culturing *M. trichosporium* OB3b for cell profiles

For the small scale ($\leq$ 2 L) cell profiles with copper and silver, *M. trichosporium* OB3b was cultured at 27 ℃ in a 5 L fermentor (stirred at 250rpm), in NMS media supplemented with 2 μM Cu(II) and 10 μM Fe(II) (culture A). When culture A reached an OD_{600} ~0.8, 1.6 L of the culture A was transferred to each of two smaller fermentors (3 L) and supplemented with 10 μM Fe(II) plus 1 μM Cu(II) (culture A1) or 1 μM Ag(I) (culture A2), respectively. Cultures A1 and A2 were incubated for 24 h and 1 L of each ferementor was harvested at ~OD_{600} 1.3. Cells were harvested by centrifugation (9000 g, 15 mins, 10 ℃) (JLA 16,250 rotor, Beckman Coulter) and the pellets were washed by repeated resuspension/centrifugation (5000 g, 15 min, 10 ℃) cycles. The pelleted cells were first resuspended in 20 mM (4-(2-hydroxymethyl)-1-piperazineethanesulfonic acid (Hepes) pH 7.5 plus 100 mM NaCl, then centrifuged. The subsequent wash involved incubation with 20 mM Hepes pH 7.5 plus 100 mM NaCl plus 10 mM ethylenediaminetetraacetic acid (EDTA) for 10 min at room temperature, followed by further centrifugation. Finally, the EDTA solution was removed by washing again with 20 mM Hepes pH 7.5 plus 100 mM NaCl and the pelleted cells were stored at -20 ℃ (2).

For the cell profile of *M. trichosporium* OB3b exposed to Ag(I)-mb a culture of *M. trichosporium* OB3b was grown in a 3 L fermentor and 1 μM Ag(I)-mb was added when the culture reached an OD_{600} 0.8 (culture B). 1 L of culture B was harvested at OD_{600} 1.3, and the cell pellet was washed. In order to grow sMMO-active *M. trichosporium* OB3b, copper was omitted from the growth medium and the cells were assayed for sMMO activity (see below). Two cultures of *M. trichosporium* OB3b were grown in 3 L fermentors supplemented with 10 μM Fe(II) and either 5 μM Cu(II) (culture C) or no copper (culture D), for the sMMO-active cell profile, and were incubated for 24 h. 2 L of each of the cultures C and D were harvested at a ~OD_{600} 2 and were tested for sMMO activity. The cell profile of sMMO-active *M. trichosporium* OB3b cells was compared to the cell profile of the organism grown under 5 μM Cu(II), which was found to be sMMO-inactive. For the large scale ($\geq$10 L) cell profiles of *M. trichosporium* OB3 the organism was grown in NMS media supplemented with 10 μM Fe(II) and high Cu(II) concentration. In the case of the 10 L cell profile of *M. trichosporium* OB3b, pellets were combined from 6 L of culture, grown at 5 μM Cu(II), and 4 L grown at 1 μM Cu(II). The 10 L cell profile resulted in the identification of a

novel copper protein from *M. trichosporium* OB3b and subsequently a 16 L cell profile, where all pellets were from cultures grown at 5 μM Cu(II), was performed aiming to optimise the purification of this protein. In all cases cells were collected by centrifugation and pellets were washed.

6.4 sMMO activity assay

M. trichosporium OB3b cultures were tested for sMMO activity using the o-dianisidine/naphthalene spectrophotometric assay. A 2 ml sample of culture was harvested and the cells were pelleted by centrifugation (12000 g, 10 mins) on a desktop centrifuge (Eppendorf centrifuge 5415D). The pellet was resuspended in 2 ml of 10 mM phosphate buffer pH 7 plus 10 mM sodium formate and an appropriate dilution was made to achieve an OD_{600} 0.3-0.7 in 2 ml. A scoop of crushed naphthalene crystals was added and the solution was vigorously shaken to initiate the reaction. Finally, 50 μl of o-dianisidine dye (5 mg/ml stock) were added to the solution and the absorbance at 528 nm, corresponding to the formation of napthol, was monitored for 30-40 mins. The activity was normalised to the dry weight of cells (OD_{600}=1 corresponds to 430 mg/L cells for *M. trichosporium* OB3b). sMMO activity in [ng of napthol min^{-1} mg^{-1} cells] was calculated from the measured ΔA_{528nm}/min taking into account that ΔA_{528nm}=1 corresponds to the formation of 2.15 10^7 ng napthol (3, 4).

6.5 Purification and quantification of methanobactin

The cell-free medium isolated from *M. trichosporium* OB3b cultures was filtered and loaded onto a C18 cartridge (Sep-pak Plus, Waters) from which the mb-containing crude extract was eluted in acetonitrile (HPLC grade) and lyophilized overnight. In order to purify through a semi-preparative column (C18-300, 250x10 mm, Hichrom) using reversed phase HPLC, the crude extract was resuspended in 10 mM ammonium acetate pH 7 plus 10 mM EDTA buffer and filtered (0.45 μm filter). The buffers used for HPLC were 10 mM ammonium acetate pH 7 (buffer A) and 10 mM ammonium acetate in 80% acetonitrile (buffer B). A linear gradient from 5% to 40% buffer B was run over 40 minutes. The column was connected to an Agilent 1100 Series instrument equipped with a diode array detector and the flow rate was 2 ml/min. The purified samples collected from the HPLC were lyophilized overnight and stored at -20 ºC (1). Purified mb forms were lyophilized overnight and dried mb was resuspended in 20 mM Hepes pH 7.5. Apo- and Cu(I)- mb were quantified by UV-VIS spectroscopy using ε_{340}= 21.8 mM^{-1} cm^{-1} and ε_{290}= 16.4 mM^{-1} cm^{-1}, respectively (1).

6.6 Ag(I) binding to mb

A working solution of $AgNO_3$ at 1.32 mM was prepared in 20 mM Hepes pH 7 from a 10 mM stock solution and titrated into apo-fl-mb (13.2 μM) in 20 mM Hepes pH 7. To test the replacement of Cu(I) by Ag(I) in Cu-mb, a Cu-mb sample (8 μM) was prepared anaerobically in 20 mM Hepes pH 7 and incubated with a stoichiometric amount of Ag(I) in the presence of 40 μM BCS. The replacement of Cu(I) by Ag(I) was followed by monitoring the absorbance at 483 nm until equilibrium was reached. The Ag(I) titration and the Ag(I) replacement experiment were performed in an anaerobic chamber (Belle technology, $O_2 < 2ppm$) and anaerobic 10 mm path length quartz cuvettes (Hellma), sealed with a gas-tight septum, were used for taking samples outside the chamber. A gas-tight syringe (Hamilton) was used for titrations.

6.7 Purification and identification of metalloproteins from *M. trichosporium* OB3b

6.7.1 Anaerobic lysing of cells and quantification of soluble protein

Based on a modified method previously used (2), frozen cell pellets were resuspended in 20 mM Hepes pH 8.8 and lysed by freeze grinding in liquid nitrogen. The lysate was allowed to thaw in an anaerobic chamber (Belle technology) ($O_2 < 1$ ppm) and sealed anaerobically prior to ultracentrifugation (160,000 g, 10 ℃, 1 h) (SW40 Ti rotor, Beckman Coulter). The supernatant, consisting of the soluble fraction of the cells, was recovered anaerobically and the concentration of total protein (mg/ml) in the sample was determined by Bradford assay. Bovine serum albumin (BSA) standards (0-20 μg/ml) were prepared from a 20 μg/ml BSA stock and 100 μl of each standard was used in a well of a 96-well microplate. Serial dilutions of the soluble protein sample were made (1:10, 1:100 and 1:1000 in a final volume of 100 μl) and 100 μl of Bradford reagent (Thermo Scientific) was added to all wells at the same time. The protein content of the samples was determined by the absorbance at 595 nm monitored on a plate reader (Thermo Labreader, Multiskan Ascent).

6.7.2 Anion exchange chromatography

For all the small scale profiles ($\leq$ 2 L), the anion exchange chromatography was performed entirely inside the anaerobic chamber. The soluble fraction of the cells was loaded on a 1 ml Hi Trap Q HP column, equilibrated in 20 mM Hepes pH 8.8, with a peristaltic pump at 0.8 ml/min flow rate. The loaded column was washed with 20 mM

Hepes pH 8.8 and proteins were eluted in 1 ml fractions of the same buffer containing 100 mM, 200 mM, 300 mM, 400 mM and 1 M NaCl. For the large scale profiles ($\geq$ 10 L), the soluble protein extract was loaded on a 5 ml column of the same type that was then sealed. Proteins were eluted aerobically in 0.5 min fractions with two linear NaCl gradients, either using a homemade gradient mixing system (10 L cell profile), performed anaerobically, or using the ÄKTA purification system (GE Healthcare) (16 L cell profile) at 2 ml/min flow rate. All buffers were degassed and purged with nitrogen during elution. The NaCl gradients ranged from 0-500 mM NaCl, over 80 fractions, followed by 500 mM-1 M NaCl, over 20 fractions, in the case of the 10 L profile, and 0-250 mM NaCl, over 80 fractions, followed by 250 mM-1 M NaCl, over 40 fractions, in the case of the 16 L profile. The metal content of anion fractions from the large scale profiles was quantified by inductively coupled plasma mass spectrometry (ICP-MS) (section 6.7.4) and the proteins present in metal-containing fractions were visualised by sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) (section 6.11.2).

6.7.3   Size exclusion chromatography

For the 1 L profiles, a 200 μl aliquot from each anion exchange fraction was resolved on a size exclusion (SW3000, Tosoh Bioscience) HPLC column, pre-equilibrated and run at a flow rate of 0.5 ml/min in 5 mM Hepes pH 7.5 plus 50 mM NaCl. 35 x 1 min fractions were collected from each injection and he first 5 fractions (devoid of protein and metal) were discarded. A solution of 5 mM Hepes pH 7.5 plus 50 mM NaCl plus 10 mM EDTA was injected on the column between each sample to remove any metal contamination, followed by re-equilibration (2).

In the case of the large scale profiles only selected copper-containing fractions (200 μl aliquots of each) were further purified by gel filtration chromatography on either a Sephadex G100 (Sigma), packed in a glass econo-column (dimensions 1 x 20 cm, Bio-Rad) and used inside the anaerobic chamber, or aerobically on a S75 10/300 GL (GE Healthcare) column using a degassed buffer that was purged with nitrogen during elution. Both columns were equilibrated and run in 20 mM Hepes pH 7.5 at flow rates of 0.35 ml/min or 0.8 ml/min, respectively. Superdex S75 was also used for the purification of selected anion fractions from the sMMO-active and 5 μM Cu^{2+} 2 L profiles. The Superdex S75 column was calibrated using albumin (67 kDa), ovalbumin (43 kDa), chymoptrypsinogen (25 kDa) and ribonuclease (13 kDa). Elution fractions

were analysed first by ICP-MS and copper-containing peaks were investigated by SDS-PAGE.

In order to assess where mb would elute from the columns used for anion-exchange and size-exclusion chromatography in cell profiles of *M. trichosporium* OB3b, control experiments with apo-, Cu(I)- and Ag(I)-mb were performed. The three forms of mb (20-40 µM) were loaded onto an anion-exchange (HiTrap Q HP, 1 ml) or a size-exclusion (SW3000) column at concentrations. Elution from the anion-exchange column was performed with 20 mM Hepes pH 8.8 containing 100, 200, 300, 400 and 1000 mM NaCl. Size-exclusion chromatography was performed in 5 mM Hepes pH 7.5 plus 50 mM NaCl. UV-VIS spectra of the eluted fractions from both columns were recorded in order to test whether they contained mb.

6.7.4 Metal analysis by ICP-MS

For the small scale profiles 300 µl aliquots of the size exclusion fractions were each diluted in 2.7 ml of 2.5% HNO_3, containing an internal standard (20 ppb Co or Pt), and the metal content was determined by ICP-MS. Standard solutions (0-50 ppb Ag, Mn, Fe, Cu and Zn), prepared with identical buffer/acid composition to the samples for analysis, were used for calibration (Thermo Electron Corp., X-Series) (2). Mass ions (^{55}Mn, ^{57}Fe, ^{59}Co, ^{63}Cu, ^{66}Zn, ^{107}Ag and ^{195}Pt) were each measured 100 times using the peak-jump method (20 ms dwell time, 3 channels, 0.02 AMU separation), in triplicate for each fraction, and counts were compared with the metal standards to calculate metal concentrations. In the large scale profiles 250 µl aliquots of each of the anion or the size exclusion fractions were diluted into 4.750 ml of 2.5% HNO_3, containing 20 ppb Co as internal standard, and the metal content was determined by ICP-MS as above with some modifications (30 ms dwell time, 5 channels, 0.02 AMU separation) . Data were converted to µM and plotted as surface representations in SigmaPlot 11.0 with the *z*-axis representing metal concentration.

6.7.5 Protein identification

Proteins whose intensity on SDS-PAGE gels correlated with metal concentration in profiles were excised from gels and digested with trypsin (25 ng/µl) in 50 mM ammonium bicarbonate pH 8 plus 5 mM $CaCl_2$ (5). Peptide digests were resuspended in 0.1% aqueous trifluoroacetic acid and desalted using C18 ZipTips (Millipore) according to the supplied manufacturer's protocol (available from Millipore.com). Desalted

peptides were then resuspended in 0.1% formic acid and injected onto a 180μm X 20mm C18 trapping column (Waters) in-line with a 75 μm X 100mm C18 capillary column (Waters) connected to a NanoAcquity ultra-performance liquid chromatography system (UPLC) (Waters). The buffers used for UPLC were 0.1% aqueous formic acid (buffer A) and 0.1% formic acid in acetonitrile (buffer B). A linear gradient from 1% to 50% buffer B was run over 30 minutes at a flow rate of 0.3 μl/min. During the run peptides eluting from the capillary column were analysed by nano-LC/MS/MS in positive ion mode on a Finnigan linear trap quadrupole Fourier transform (FT) mass spectrometer (ThermoElectron). Eluate was sprayed using uncoated 10 μm internal diameter SilicaTips (New Objective) at a spray voltage of ~2 kV. The mass spectrometer performed survey MS scans over the mass range m/z = 300 – 1500 in data-dependent mode. MS/MS data was acquired with a FT-MS resolution setting of 50,000 (at m/z = 400) and the five top ions in the parent scan were then automatically subjected to MS/MS in the linear ion trap region of the instrument. Following acquisition the data file was searched using the Mascot MS/MS ion search tool (Matrix Science) against the NCBI non- redundant protein sequence database of alphaproteobacteria in order to identify the parent protein(s) from which the detected MS/MS peptides originated.

6.8 Bioinformatics

In order to search for protein homologues of a novel copper protein (Csp1) identified from one of the cell profiles of *M. trichosporium* OB3b, the protein sequence of Csp1 was BLAST searched against the *M. trichosporium* OB3b genome (6), using the NCBI protein blast tool (7). Only hits with an expect value (E) lower or equal to 10^{-3} were considered as reliable. SignalP was used to check for signal peptides (8) and Clustal Omega (9) was used for aligning protein sequences. Protein homologues whose metal-free structures have been solved were identified by blast searching the full sequence of the novel copper protein from *M. trichosporium* OB3b on the pdb database. Protein homologues of Csp1 and Csp3 were identified by BLAST searching the protein sequences of Csp1, Csp2 or Csp3 against MOB genomes using the NCBI protein-protein blast tool (7). Homologues of Csp1 and Csp3 in other bacteria were identified by BLAST search was repeated against the non-redundant database for bacteria using the NCBI protein-protein blast tool (7). In all cases the size of the protein sequence (number or residues), the number of conserved Cys residues and an expect value (E)

lower or equal to 10^{-3} were used as criteria for the identification of true homologues within the list of all the generated hits. SignalP (8) and TatP (10) were used to identify Tat-leader peptides. All protein sequences were aligned using Clustal Omega (9).

6.9 Manipulation of *Escherichia coli* strains

6.9.1 Strains

E. coli strains (JM101 and BL21 (DE3)) for growing cultures were maintained on LB-agar plates with an appropriate antibiotic at 4 ˚C. For long-term storage, stocks of the strains were prepared using 1.5 ml of pelleted culture in 500 µl of 50% glycerol plus 1% (w/v) Bacto peptone. The stocks were stored at -80 ˚C.

6.9.2 *E. coli* competent cells

A single colony of *E. coli* was picked from a plate and grown overnight in 5 ml LB media, shaking (250 rpm) at 37 ˚C. The next day a 1:100 dilution of the overnight culture was made into 10 ml LB and that culture was grown for approximately 2 hours, at the same conditions, up to an OD_{600}~0.2-0.4. Cells were harvested by centrifugation at 3300 rpm for 10 min and the pellet was kept on ice and resuspended in 1 ml ice cold TSS buffer (85% (v/v) LB, 10% (w/v) polyethylene glycol (PEG) 8000, 50 mM $MgCl_2$ and 5% (v/v) dimethyl sulfoxide (DMSO) at pH 6.5. Competent cells were stored at -80 ˚C.

6.9.3 *E. coli* transformation

Typically, 50 µl of *E. coli* competent cells were incubated on ice, in a sterile eppendorf tube, with 2 µl of plasmid DNA for 15-30 min. Cells were heat shocked at 42 ˚C for 60-90 sec and, subsequently, 500 µl of LB was added. The cells were allowed to recover shaking (250 rpm) at 37 ˚C for 1 hour. 100 µl of the transformation mix was plated on LB-agar (1.5% agar), containing the appropriate antibiotic, and grown overnight at 37 ˚C.

6.10 Molecular cloning

6.10.1 Genomic DNA extraction from *M. trichosporium* OB3b

Genomic DNA was isolated from cultures of *M. trichosporium* OB3b using a GenElute Bacterial Genomic DNA kit (Sigma) according to the manufacturer's protocol.

6.10.2 Polymerase chain reaction

Polymerase chain reactions (PCR) were typically prepared in 0.5 ml eppendorf tubes containing around 20-50 ng of template DNA, 0.4 mM deoxyribonucleotide triphosphates (dNTPs) mix (Promega), 0.4 μM (130-184 ng) of forward and reverse DNA primers (Sigma), 1 μl (2 units) of Phusion polymerase (New England Biolabs), 10 μl of 5x Phusion buffer made up to 50 μl volume by addition of sterile Milli-Q water. The PCR samples were initially incubated for 5 min at 95 ºC, followed by 30 cycles consisting of incubations of 1 min at 95 ºC, 1 min at 60 ºC and 1.2 min at 72 ºC. Finally the samples were incubated for 10 min at 72 ºC.

6.10.3 Extraction of DNA from *Escherichia coli*

Plasmid DNA was isolated from bacterial cultures using a GenElute Plasmid Mini- and Midi-Prep kit (Sigma) according to the manufacturer's protocol. The purity of the DNA was assessed by the relative absorption at 260 nm and 280 nm, where the A260/A280 ratio of approximately 1.8 indicated DNA with little protein contamination.

6.10.4 Digestion of DNA using restriction endonucleases

Plasmid and PCR-amplified DNA was digested using NdeI and NcoI restriction enzymes (New England Biolabs) according to the manufacturer's protocol. Typically DNA was incubated at 37 ˚C with the appropriate endonucleases for 1-2 h. Prior to the ligation reaction the digested DNA was analysed by agarose gel electrophoresis and purified.

6.10.5 Isolation of DNA from agarose gel

The target bands were excised from the agarose gel using a scalpel and isolated from the gel using a GenElute Extraction Kit (Sigma) according to the manufacturer's protocol.

6.10.6 DNA ligation

Ligation of the digested with restriction endonucleases vector and insert DNA was achieved using a T4 ligase (Fermentas) according to the manufacturer's protocol. Typically, the digested and purified vector was mixed with digested and purified DNA insert in a T4 DNA ligase buffer and incubated at room temperature with a T4 DNA ligase overnight. 2-5 μl of ligation mixture was used for subsequent transformation in *E. coli* JM101 competent cells. Successful ligation was verified by the digestion of the

ligated plasmid with the appropriate restriction endonuclease, followed by agarose gel electrophoresis and sequencing of the DNA insert.

6.10.7 DNA sequencing
The DNA sequence of the amplified genes in plasmid DNA was verified using the DNA sequencing service of Beckman-Coulter Genomics.

6.10.8 Determination of DNA concentration
The concentration of DNA was determined using a nanodrop reader.

6.11 Electrophoresis

6.11.1 Agarose gel electrophoresis
Agarose (1.2% (w/v), Melford) was dissolved in 50 ml of Tris-Acetate-EDTA (TAE) buffer (40 mM Tris pH 8.0, 40 mM acetate and 1 mM EDTA) by boiling in a microwave. The liquidified agarose was allowed to cool to 50-55 ˚C before mixing with GelRed nucleic acid stain (Biotium), according to the manufacturer's instructions, and pouring into a glass cast (Biorad). The gel was allowed to set at RT and was then transferred into a gel tank containing TAE buffer and loaded with DNA samples containing DNA loading buffer (0.4% (w/v) bromophenol blue, 5% glycerol). Electrophoresis was performed at a voltage of 100 V for approximately 45 mins, using a 1 kb DNA ladder (Promega) as a marker. Typically, DNA fragments were visualised under UV irradiation (Biorad Gel Doc 1000). When DNA fragments were required for further experiments (DNA ligation), the agarose gel was illuminated using a UV transilluminator (UV Tec).

6.11.2 Sodium dodecyl sulphate-polyacrylamide electrophoresis
Proteins were visualized by SDS-PAGE using gels consisting of 5% stacking gel (5% of 37.5:1 (2.6%C) 40% Acrylamide-Bis, 0.125 M Tris pH 6.8, 0.1% SDS, above 0.1% ammonium persulfate (APS) and 0.1% N, N, N', N' tetramethylethylenediamine (TEMED) and 15% (15% 37.5:1 (2.6%C) 40% Acrylamide-Bis, 0.375 M tris(hydroxymethyl)aminomethane (Tris) pH 8.8, 0.1% SDS, 0.1% ammonium persulfate (APS) and 0.1% N, N, N', N' tetramethylethylenediamine (TEMED)) or 18% resolving gel. For 18% gels, the same procedure as is followed using a higher percentage of Acrylamide-Bis (18%). Protein-containing fractions were mixed with loading dye (5-fold stock) (500 mM Tris pH 6.8, 500 mM dithiohreitol (DTT), 5% SDS,

50% glycerol and 0.5% bromophenol blue) and were heated at 95 ˚C for 10 min to denature the protein and centrifuged at 10000 rpm for 1 min. The sample was loaded on the gel SDS-PAGE was performed at 150 V for approximately 1.5 h, until the dye ran off the gel. The running buffer contains 0.125 M Tris, 1.25 M glycine, 0.5% SDS pH 8.3. A broad range marker (Bio-Rad) was used to determine the molecular weight of the proteins. This included myosin (205 kDa), β-galactosidase (116.5 kDa), bovine serum albumin (80 kDa), ovalbumin (49 kDa), carbonic anhydrase (32 kDa), trypsin inhibitor (27.5 kDa), lysozyme (18.5 kDa) and aprotinin (6.5 kDa). In te case of cell profiles, gels were were stained with Oriole fluorescent gel stain (Bio-Rad) using the manufacturer's protocol and proteins were visualised under UV light. Gels for recombinant protein purification were stained for 30 min in a 10% acetic acid, 40% methanol, brilliant blue solution and destained for 30 min in 10% acetic acid, 40% methanol.

6.12 Small scale protein over-expression

In order to determine the optimal conditions for protein over-expression, small-scale expression trials were performed. Typically, 10 ml of LB containing an appropriate antibiotic was inoculated with a single colony of *E.coli* BL21 transformed with the expression plasmid and grown overnight at 37 ˚C, shaking at 250 rpm. The next day a 100-fold dilution into 200 ml of LB containing an appropriate antibiotic was made and incubated for approximately 2 h, until an OD_{600} of approximately 0.5. Protein over-expression was induced by adding isopropyl-β-D-1-thiogalactopyranoside (IPTG). The growing parameters were optimised were growing time (2-24 h) and IPTG concentration (0.1 mM or 1 mM). Routinely, 10 ml samples were collected from the culture over time |(0, 2, 4, 6, 24 h), sonicated and centrifuged for 10 mins at 5000 g. 100 μl samples of sonicated cultures (representing all expressed proteins) and 100 μl of supernatant after the samples had been centrifuged (representing the soluble proteins expressed) were taken and lysed in SDS-PAGE loading dye. The samples were analysed by SDS-PAGE to identify the expression levels of the soluble proteins.

6.13 Large scale Csp1 and Csp3 protein over-expression and purification in *E. coli* BL21

6.13.1 Over-expression

A single colony of freshly transformed *E. coli* BL21 with pET29a containing the Csp1 or Csp3 insert was inoculated from a plate and grown during the day in 10 ml LB,

shaking (250 rpm) at 37 ˚C. After 8 hours a 1:100 dillution of the day culture was made into 50 ml of LB, and the culture was grown overnight under the same conditions. The next morning a 1:100 dillution of the overnight culture was made in eight 0.5 L conical flasks, containing preheated LB, and the cultures were grown up to an OD_{600}~0.6. At that point, 1 mM of isopropyl β-D-thiogalactopyranoside (IPTG) was added to each flask and the cultures were allowed to grow for another 6 h, for Csp1 over-expression. In the case of Csp3, over-expression was induced by or 0.1 mM IPTG and the cultures were grown for 24 h. Cells were harvested by centrifugation at 9000 g for 30 min and pellets were stored at -20 ˚C.

6.13.2 Purification

Cell pellets were thawed on ice and resuspended in approximately 40 ml of 20 mM tris(hydroxymethyl)aminomethane (Tris) pH 8.5 plus 1 mM dithiothreitol (DTT) (10 ml buffer/1 L cell culture). The resuspended pellets were lysed by sonication on ice (10 sets of 60 s pulse and 60 s interval) and the soluble fraction was separated from the cell pellet by centrifugation at 40000 g for 30 min. The soluble fraction was diluted five-fold with 20 mM Tris pH 8.5 plus 1 mM DTT to a final volume of approximately 200 ml and loaded on a Hi-Trap Q HP anion exchange column (5 ml), already equilibrated in the same buffer at 4 ˚C, using a peristaltic pump at approximately 3 ml/min. Protein was eluted from the column with a gradient of 0-300 mM NaCl in the same buffer over 36 2.5min fractions (5ml) and at a flow rate of 2 ml/min, while protein elution was monitored at 280 nm, 260 nm and 240 nm. The protein content of the elution fractions was visualized by SDS-PAGE. Recombinant Csp proteins run on the gels at a molecular weight of approxiamately 14.4 kDa. Fractions containing recombinant Csp proteins were combined, diluted ten-fold in 10 mM Tris pH 7.5 for Csp1, or pH 7 for Csp3, plus 1 mM DTT and loaded on a Hi Trap Q HP anion exchange column (5 ml) as before. Protein was eluted with a gradient of 0-200 mM NaCl in the same buffer, over 60 1.5 min fractions (3 ml) at 2 ml/min. As previously, elution fractions were run on 18% SDS-PAGE gels to visualize their protein content. Csp-containing fractions were combined and exchanged with ultra-filtration (Amicon stirred cell with a 10 kDa molecular mass cut-off membrane) into 20 mM Hepes pH 7.5. A final purification step was performed on a Superdex S75 10/30 GL gel filtration column in 20 mM Hepes pH 7.5 plus 200 mM NaCl and the absorbance was monitored at 280 nm for Csp1, or 240 nm and 420 nm for Csp3. The yield of pure protein is approximately 15-20 mg/L cell

culture. The intact mass of the recombinant protein was verified by matrix assisted laser desorption ionization- time of flight spectrometry (MALDI-TOF).

6.14 Intact protein mass determination

The intact mass of the purified proteins was determined by matrix assisted laser desorption ionisation (MALDI) time of flight (TOF) mass spectrometry. The protein sample was analysed using a Voyager DE-STR MALDI-TOF mass spectrometer (Applied Biosystems Inc., Framingham, MA, USA). The instrument used a nitrogen laser at 337 nm and was operated in linear mode at an accelerating voltage of 20 kV. The sample (typically ~10 μM) was diluted 1:5 with sinapinic acid matrix [10 mg/ml, 50% acetonitrile / 0.1% trifluoroacetic acid(aq)] and then 2 μl of the mix is plated onto the target. The intact protein mass spectrum was acquired over the mass range of interest (10-25 kDa), calibrated with an appropriate standard, typically comprising of insulin, thioredoxin and apomyoglobin (calibration mix 3, ABSciex). Typically 50 - 250 laser shots are acquired per spectrum. Spectral noise removal and smoothing was carried out using Data Explorer software (ABSciex) to generate the final output spectra.

6.15 Atomic Absorption Spectroscopy

The copper and zinc content of the purified proteins was determined by Atomic Absorption Spectroscopy (AAS) on a Thermo Electran M Series AA spectrometer. Protein samples at approximately 30 μM were prepared in 2% nitric acid and were measured against calibration copper or zinc standards (0-1 ppm) (Fluka).

6.16 Protein quantification

The concentration of apo-Csp proteins was determined by thiol quantification using dithiobis(2-nitrobenzoic acid) (DTNB, Ellman's reagent) (11). The proteins (<1 μM) were unfolded in an anaerobic chamber (Belle technology) in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing 8 M Urea and 1 mM EDTA and the thiol concentration was determined by adding 0.5 mM DTNB, from a stock solution (10 mM) prepared in 100 mM phosphate pH 8 plus 1 mM EDTA. The samples were transferred in anaerobic cuvettes (quartz, Hellma) and the absorbance at 412 nm was monitored on a λ35 spectrophotometer (Perkin-Elmer) for 30 mins. After the reaction reached equilibrium a spectrum of the sample was taken and the thiol concentration was determined using an extinction coefficient of $\varepsilon_{412nm}=14.15$ mM^{-1} cm^{-1}. Csp1 and Csp3 concentrations were quantified assuming 13 and 18 thiols respectively. The concentration of apo-Csp

proteins was also determined by Bradford assay using the Coomassie Plus protein assay kit (Thermo Scientific). For the calibration curve BSA standards were made at concentrations of 0-1000 μg/ml and all samples were prepared by diluting 10 μl of protein in 1 ml of Coomassie reagent. The samples were incubated for 15 mins and the absorbance was measured at 595 nm. The protein concentrations determined by this method were lower than those determined by DTNB assay by approximately 15% and 35% for Csp1 and Csp3, respectively. Based on this, DTNB assay was routinely used for quantifying apo-Csp proteins.

The Bradford assay was used routinely for the quantification of Cu(I)-Csp1 without correcting for the discrepancy between this method and the DTNB assay. Since the discrepancy between the protein concentration determined by DTNB and Bradford assay was larger in the case of Csp3, a different method was used for quantifying Cu(I)-Csp3. Cu(I) was extracted anaerobically from the protein (<1 μM) by bathocuproine disulfonate (BCS) (2.5 mM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing 6.8 M guanidine hydrochloride (see section 6.21). After the removal of Cu(I), the concentration of the free thiols was determined by adding 0.5 mM DTNB in the same cuvette. The absorbance at 412 nm was corrected using a blank sample containing the same amount of $[Cu(BCS)_2]^{3-}$ and DTNB as the protein sample.

6.17 Far-UV Circular Dichroism spectroscopy

Far UV (180-250 nm) CD spectra for folded proteins (0.3-0.7 mg/ml) in 100 mM phosphate pH 8 were recorded aerobically on a JASCO J-810 spectrometer at 20 °C (12, 13), using a 0.2 mm path length quartz cuvette. CD was also used to monitor the rate of protein unfolding for Csp3 (0.15-0.4 mg/ml) in 20 mM Hepes pH 7.5 plus 200 mM NaCl containing 8 M urea or 6.8 M guanidine hydrochloride. Protein was unfolded in the anaerobic chamber and samples were taken out for CD spectra to be recorded. In all cases, the ellipticity [θ] (in degrees) was converted to mean residue ellipticity (mre) (14) using the equation:

$$[\theta]_{mrw} = MRW*\theta/ (10*d*c)$$

where MRW is the mean residue weight given by the equation MRW=M/(N-1) (M is the molecular weight of the polypeptide in Da and N is the number of amino acids in the polypeptide), θ is the observed ellipticity, d is the pathlength of the cuvette (in cm) and c is the protein concentration (in g/ml). The α-helical content of the proteins was

determined from the mean residue ellipticity (mre) at 222 nm (15) using the following equation:

$$[([\theta]_{222}-3000)/(-36,000-3000)]*100$$

The α-helical content calculated from the experimental CD data was compared to that given from the crystal structures of the proteins when processed on the STRIDE interface (16).

6.18 Analytical gel-filtration chromatography

Analytical gel filtration was routinely performed on a Superdex 75 GL 10/300 column (GE healthcare) equilibrated in 20 mM Hepes pH 7.5 plus 200 mM NaCl and at 0.8 ml/min flow rate. The column was calibrated using Blue Dextran (2000 kDa), albumin (67 kDa), ovalbumin (43 kDa), chymoptrypsinogen (25 kDa) and ribonuclease (13 kDa) in the same buffer (17, 18). Apo- and Cu(I)-loaded Csp samples were injected on the column using a 200 μl loop at a range of concentrations at 5-100 μM.

6.19 Cu(I) binding to the Csp1 and Csp3 and Csp3 Cu(I) affinity determination

All experiments were performed in an anaerobic chamber (Belle technology, $O_2<$ 2ppm) and anaerobic 10 mm path length quartz cuvettes (Hellma), sealed with a gas-tight septum, were used for taking samples outside the chamber. A gas-tight syringe (Hamilton) was used for titrations.

6.19.1 Preparation of Cu(I) stock solution

A Cu(I) stock solution (50 mM) was prepared from $[Cu(CH_3CN)_4]PF_6$ (Sigma) in 100% anhydrous acetonitrile. The stock solution was diluted appropriately into 20 mM Hepes pH 7.5 plus 200 mM NaCl immediately before use. The copper concentration was determined with BCS (Sigma), a chromophoric ligand for Cu(I), and an $\varepsilon_{483nm}=12.5$ mM^{-1} cm^{-1} was used for quantification of $[Cu(BCS)_2]^{3-}$ (12, 19).

6.19.2 Preparation of Cu(I)-Csp1 and Cu(I)-Csp3 and determination of Cu(I) binding stoichiometries

Fully loaded Cu(I)-Csp1 and Cu(I)-Csp3 were prepared by addition of either 13 or 18 molar equivalents of Cu(I) to apo-Csp1 or apo-Csp3, respectively. The Cu(I) binding stoichiometries were estimated by titrating Cu(I) into apo-Csp1 (15 μM) or Csp3 (5 μM) in 20 mM Hepes pH 7.5 plus 200 mM NaCl. For Csp3 the experiment was also repeated in the same buffer in the absence of NaCl (12).

6.19.3 Calculation of Cu(I) binding affinity for Csp3

In order to verify the stoichiometry of tightly bound Cu(I) to Csp1 or Csp3, and also estimate the Cu(I) affinity in the case of Csp3, Cu(I) additions into Csp1 (11 μM) or Csp3 (2.5 μM) were also performed in the presence (100, 250, 500, 1000 μM) of bathocuproine disulfonic acid (BCS) ($log\beta_2$=20.8) or bicinchoninic acid (BCA) ($log\beta_2$=17.7) (20) (100, 250, 500, 1000 μM), using ε_{483nm}=12.5 mM^{-1} cm^{-1} and ε_{562nm}=7.7 mM^{-1} cm^{-1} for the quantification of $[Cu(BCS)_2]^{3-}$ and $[Cu(BCA)_2]^{3-}$, respectively (12, 19). Stock solutions of BCA and BCS (50 mM) were prepared anaerobically in 20 mM Hepes pH 7.5 plus 200 mM NaCl. In the case of Csp3, protein solutions containing the appropriate concentration BCA or BCS were prepared and aliquoted in anaerobic quartz cuvettes into which Cu(I) additions corresponding to increasing Cu(I), by 3 molar equivalents to protein per cuvette (for BCA) and by 6 molar equivalents to protein per cuvette (for BCS), were made separately. According to an earlier experiment, in which the equilibration time was tested over time for Cu(I) additions (made by 4 molar equivalents to protein per cuvette) into apo-Csp3 (2.5 μM) in the presence of 100 μM BCA, samples were found to reach equilibration after 24 h. Based on this, therefore, the samples containing Csp3, with various BCS or BCA concentrations, in which Cu(I) additions were made were incubated in anaerobic cuvettes inside the anaerobic chamber, and equilibration was monitored by UV-VIS from 24 h onward. For Csp3 the experiment was also repeated for 100 μM BCA in 20 mM Mes pH 6.5 or 20 mM Taps pH 8.5, both containing 200 mM NaCl. In order to estimate the average Cu(I) affinity of Csp3 the Hill equation (21) was used to non-linearly fit the data for Cu(I) additions into Csp3 in the presence of 1000 μM BCA or 100 μM BCS:

$$Y=[Cu(I)]^n / (K_{Cu}{}^n + [Cu(I)]^n)$$

Where Y represents the fractional occupancy of the protein, [Cu(I)] is he concentration of free copper, K_{Cu} is the dissociation constant, n is the Hill coefficient.

6.20 Fluorescence

Emission spectra were recorded on a Cary Eclipse fluorimeter (Varian) using a 10 mm path length quartz cuvette and by exciting at 300 nm and monitoring the emission in the 400-700 nm range. The instrument was set at high voltage and 5 mm slits were used for excitation and emission with an emission filter at 430-1100 nm (19).

6.21 Cu(I) removal from Cu(I)-Csp1 and Cu(I)-Csp3 by BCS

In order to remove Cu(I) from fully loaded Cu(I)-Csp samples, protein (1-1.2 μM) was mixed anaerobically with 2.5 mM BCS, from a 50 mM stock, with all solutions in 20 mM Hepes pH 7.5 plus 200 mM NaCl. The absorbance of the sample at 483 nm was monitored over time and a spectrum of the sample was taken after the reaction reached equilibrium. The concentration of $[Cu(BCS)_2]^{3-}$ was determined using an extinction coefficient of ε_{483nm}=12.5 mM^{-1} cm^{-1}. The experiment was also performed in the presence of 8 M urea or 6.8 M guanidine hydrochloride.

6.22 Copper exchange experiments

6.22.1 Copper exchange between Cu(I)-Csp1 or Cu(I)-Csp3 and apo-mb

Copper exchange was studied by anaerobically mixing fully loaded Cu(I)$_{13}$-Csp1 or Cu(I)$_{18}$-Csp3 with apo-mb at an appropriate excess in 20 mM Hepes pH 7.5 plus 200 mM NaCl. In the case of Csp1, the samples for the experiment were prepared by mixing 1.2 μM Cu(I)-Csp1 (from a stock of fully loaded Csp1 (45 μM)) with 15.4 μM, 32.3 μM or 46.1 μM apo-mb and incubating for a total of 16 h. A UV-VIS spectrum of apo-mb was run before mixing with Cu(I)-Csp1 and, subsequently, the reaction was monitored at 394 nm for 3.5 h after of Cu(I)-Csp1 was added. Another spectrum of the samples was taken after the reactions reached equilibrium (3.5 h) and then at 16 h. In the case of Csp3, samples were prepared by mixing 0.95 μM Cu(I)-Csp3 with 17 μM or 34 μM apo-mb in anaerobic cuvettes. The samples were monitored over time at 1h, 2h, 4h, 20h, 3 days, 6 days, 8 days, 13 days and 15 days. Control samples containing the same concentration of Cu(I) (17 μM) and apo-mb (17 or 34 μM) were also prepared and monitored over time. For all the above experiments the amount of Cu-mb formed was quantified using the difference between the extinction coefficient values at 394 nm for apo-mb and Cu-mb, which corresponds to $\Delta\varepsilon_{394\ nm}$= 7.2 mM^{-1} cm^{-1}. All samples were stored anaerobically between measurements.

6.22.2 Copper exchange between apo-Csp1 or apo-Csp3 and Cu(I)-mb

For copper exchange between apo-Csp proteins and Cu-mb, Csp proteins were concentrated using a centrifugal device with a 10 kDa membrane (Vivaspin 500). Samples were prepared by anaerobically mixing Cu-mb with an excess of apo-Csp in 20 mM Hepes pH 7.5 plus 200 mM NaCl and the experiment was monitored by UV-VIS. A spectrum of the Cu-mb sample was taken before the addition of apo-Csp and then

spectra were taken directly after mixing and at 10 min, 1 h, 3 h and 20 h. For Csp1 72.5 μM or 234 μM of apo-Csp1 were mixed with 2.6 μM Cu-mb and for Csp3, 98 μM or 243 μM of apo-Csp3 were mixed with 2.6 μM Cu-mb. All samples were stored anaerobically between measurements.

6.23 Crystallisation trials and diffraction data collection

Crystallisation screens using 96 well plates were set with a crystallisation robot and the plates were stored at 20 °C either aerobically for apo-Csp1 and apo-Csp3 or in the anaerobic chamber (where they were sealed) for Cu(I)-Csp1 and Cu(I)-Csp3. Screens using 24 well plates were set manually under aerobic conditions for apo-proteins or anaerobically for Cu(I)-Csp proteins and stored as already mentioned. The 96 well and 24 well plates contained 80 μl and 500 μl of reservoir solution, respectively. Protein was mixed with well solution at a 1:1 and 2:1 protein to reservoir solution ratio using 100 nl or 300 nl protein drops for the 96 well screens and 1 μl protein drops for the and 24 well screens. The sitting drop and the hanging drop method of vapour diffusion were used for the 96 well and the 24 well screen, respectively.

6.23.1 Preparation of apo-Csp1 and apo-Csp3 samples for crystal trials

Typically, apo-Csp proteins in 20 mM Hepes pH 7.5 were concentrated using a spinning centrifugal device (vivaspin 500, 10 kDa membrane). A 100-fold dilution of the final sample was checked by Bradford assay for protein concentration and the sample was diluted appropriately at a concentration of 20 mg/ml. The protein samples were set in 96 well plates containing the JCSG+ and Midas commercial screens (Molecular dimensions) for Csp1 and Csp3, respectively.

6.23.2 Preparation of Cu(I)-Csp1 and Cu(I)-Csp3 for crystal trials

Typically, apo-Csp1 (70-75 μM) was incubated with 12-14 molar equivalents of Cu(I) in 20 mM Hepes pH 7.5 plus 200 mM NaCl to give Cu(I)-Csp1 (refered to as $Cu(I)_{13}$-Csp1 for crystal structure). The sample was concentrated using a spinning centrifugal device (vivaspin 500, 3 kDa membrane) and a dilution of the final sample was made in order to assay the protein and copper concentration. The concentration of the Cu(I)-Csp1 protein was determined by Bradford assay and Cu(I) concentration was assayed by AAS and by Cu(I) removal by BCS (section 6.15 and 6.21). Cu(I)-Csp3 samples were prepared as above by loading apo-Csp3 (53-70 μM) with 2, 8, 16 or 18 molar equivalents of Cu(I) (to give $Cu(I)_2$-Csp3, $Cu(I)_8$-Csp3, $Cu(I)_{16}$-Csp3 and $Cu(I)_{18}$-Csp3,

respectively), and protein and Cu(I) concentration was assayed as above. The protein samples were set on 96 and 24 well plates, respectively, containing the Morpheus and JCSG+ commercial screens (Molecular dimensions) for Cu(I)-Csp1 and Cu(I)-Csp3, respectively. Diffraction data were collected at 100 K at the Diamond Light Source, U.K.. Data collection, model building and refinement were performed by Dr. Arnaud Baslé and the phase was determined using single-wavelength anomalous dispersion for copper. In the case of Cu(I)-Csp1 the oxidation state of copper was verified by X-ray absorption near edge spectroscopy (XANES), performed by Dr Neil Patterson,DLS).

6.24 References

1. El Ghazouani A., Baslé A., Firbank S. J., Knapp C. W., Gray J., Graham D. W., Dennison C., Copper-binding properties and structures of methanobactins from Methylosinus trichosporium OB3b, Inorg. Chem., **2011**, 50, 1378-1391

2. Tottey, S., Waldron, K. J., Firbank, S. J., Reale, B., Bessant, C., Katsuko, S., Cheek, T. R., Gray, J., Banfield, M. J., Dennison, C., Robinson, N. J., Protein-folding location can regulate manganese-binding versus copper- or zinc-binding, Nature, **2008**, 455, 1138-1142

3. El Ghazouani A., Baslé A., Gray J., Graham D. W., Firbank S. J., Dennison C., Variations in methanobactin structure influence copper utilization by methane-oxidising bacteria, Proc. Natl. Acad. Sci., **2012**, 109, 8400-8404

4. DiSpirito A. A., Zahn J. A., Graham D. W., Kim H. J., Larive C. K., Derrick T. S., Cox C. D., Taylor A., Copper binding compounds from Methylosinus trichosporium OB3b, J. Bacteriol., **1998**, 180, 3606-3613

5. Hellman U., Wernstedt C., Gonez J., Heldin C. H., Improvement of an 'in-gel' digestion procedure for the micropreparation of internal protein fragments for amino-acid sequencing, Anal. Biochem., **1995**, 224, 451-455

6. Stein L. Y., Yoon S., Semrau J. D., Dispirito A. A., Crombie A., Murrell J. C., Vuilleumier S., Kalyuzhnaya M. G., Op den Camp H. J., Bringel F., Bruce D., Cheng J. F., Copeland A., Goodwin L., Han S., Hauser L., Jetten M. S., Lajus A., Land M. L., Lapidus A., Lucas S., Médigue C., Pitluck S., Woyke T., Zeytun A., Klotz M. G., Genome sequence of the obligate methanotroph Methylosinus trichosporium OB3b, J. Bacteriol., **2010**, 192, 6497-6498

7. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T. L., Software BLAST+: architecture and applications, BMC Bioinformatics, **2009**, 10, 421

8. Petersen T. N., Brunak S., von Heijne G., Nielsen H., SignalP 4.0: discriminating signal peptides from transmembrane regions, Nature Methods, **2011**, 8, 785-786

9. Sievers F., Wilm A., Dineen D. G., Gibson T. J., Karplus K., Li W., Lopez R., McWilliam H., Remmert M., Söding J., Thompson J. D., Higgins D., Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol., **2011**, 7

10. Bendtsen J. D., Nielsen H., Widdick D., Palmer T., Brunak S., Prediction of twin-arginine signal peptides, BMC Bioinformatics, **2005**, 6

11. Riddles P. W., Blakeley R. L., Zerner B., Ellman's reagent: 5,5′-dithiobis(2-nitrobenzoic acid) - a re-examination, Anal. Biochem., **1979**, 94, 75–81

12. Badarau A., Dennison C., Thermodynamics of copper and zinc distribution in the cyanobacterium Synechocystis PCC 6803, Proc. Natl. Acad. Sci., **2011**, 108, 13007-13012

13. Allen S., Badarau A., Dennison C., The influence of protein folding on the copper affinities of trafficking and target sites, Dalton Trans., **2013**, 42, 3233–3239

14. Kelly S. M., Jess T. J., Price N. C., How to study proteins by circular dichroism, Biochim. Biophys. Acta, **2005**, 1751, 119-139

15. Barrow C. J., Yasuda A., Kenny P. T. M., Zagorski M. G., Solution conformations and aggregational properties of synthetic amyloid β-peptides of Alzheimer's disease- Analysis of Circular Dichroism Spectra, J. Mol. Biol., **1992**, 225, 1075-1093

16. Frishman D., Argos P., Knowledge-based protein secondary structure assignment, Proteins, **1995**, 23, 566-579

17. Badarau A., Firbank S. J., McCarthy A. A., Banfield M. J., Dennison C., Visualising the metal-binding versatility of copper trafficking sites, Biochemistry, **2010**, 49, 7798-7810

18. Allen S., Badarau A., Dennison C., Cu(I) affinities of the domain 1 and 3 sites in the human metallochaperone for Cu,Zn-superoxide dismutase, Biochemistry, **2012**, 51, 1439–1448

19. Badarau A., Firbank S. J., McCarthy A. A., Banfield M. J., Dennison C., Visualising the metal-binding versatility of copper trafficking sites, Biochemistry, **2010**, 49, 7798-7810

20. Bagchi P., Morgan M. T., Basca J., Fahrni C. J., Robust affinity standards for Cu(I) biochemistry, J. Am. Chem. Soc., **2013**, 135, 18549-18559

21. Banci L., Bertini I., Ciofi-Baffoni S., Kozyreva T., Zovo K., Palumaa P., Affinity gradients drive copper to cellular destinations, Nature, **2010**, 465, 645-648