Storing copper in *Streptomyces lividans*: structural and biochemical properties of a copper storage protein

Megan Laura Straw

A thesis submitted for the degree of PhD Biochemistry

School of Biological Sciences

University of Essex

April 2019

Abstract

Copper (Cu) is essential to the growth and the morphological development of *Streptomyces lividans*. Understanding how Cu regulates the key development switches in this Gram-positive bacterium has been an area of extensive research. In particular, how Cu homeostasis and regulation are controlled and how metalation of enzymes important for morphological development is achieved have been previously investigated. To this end, this thesis reports the discovery of a cytosolic copper storage protein in *S. lividans* and offers new insight into intracellular Cu regulation, which is not under direct control of the Cu regulator protein CsoR (copper sensitive operon repressor). This copper storage protein belongs to a family of recently discovered cytosolic proteins known as Csp3. These members are exclusively found in the bacterial cytosol and comprise of a four-helix bundle that assemble into homotetramers and can bind between 70-80 Cu(I) ions through mainly Cys thiolate coordination. In Chapter 2 bioinformatic analyses reveals the phylogenetic distribution of Csp3 across Bacteria and Archaea and confirms the presence of Csp3 in *S. lividans*. Furthermore, the Csp3 in *S. lividans* is located in a gene environment that is sensitive to elevated Cu levels. Taxonomic distribution of these genes reveals a possible link to a novel transmembrane Cu export system that could facilitate removal of Cu from Csp3. X-ray structures of the apo and Cu(I) bound forms of the Csp3 from *S. lividans* have been determined and confirm a homotetramer assembly that can bind up to 80 Cu(I) ions (Chapter 3). The binding of Cu(I) ions in Csp3 is found to be cooperative with a Hill coefficient of 1.9 and Cu(I) can be transferred to Csp3 from a CopZ-like Cu(I) chaperone (Chapter 3). A *Δcsp3* null-mutant in *S. lividans* reveals that Csp3 is operable at high Cu levels and this suggests it acts to provide an additional level of protection against Cu toxicity once the CsoR system becomes saturated (Chapter 3). The mechanism of Cu(I)-loading to Csp3 has also been investigated through X-ray crystallography, site-directed mutagenesis and stopped-flow reaction kinetics using aqueous Cu(I) and Cu(I) chelated by a donor. A clear role for a His residue (His107) leading to the formation of a tetranuclear $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})]$ cluster is observed, followed by the loading of Cu(I) in a fluxional and dynamic manner (Chapters 4 and 5). Finally, over-expression studies of a putative transmembrane protein (SLI_RS17250) that is encoded by a neighbouring gene to the *S. lividans* Csp3 gene and could be part of a novel Cu export system, identified in Chapter 2, is described (Chapter 6).

Acknowledgements

like to thank lab 5.08 technician, Mrs Hannah Adams, technicians of lab 6.19 and past and present members of the lab I worked in, lab 5.08.

Most of all, I would like to thank my family. My parents, Chantal and Thomas, for their constant support. My mother, Chantal, for her boundless generosity, intuition and always being there for me – 'J'ai de la chance d'avoir une mère aussi dévouée et attentionnée que toi.' I thank my father, Thomas, for his advice, patience and encouragement. They have been an anchor throughout my life. I would like to thank my two older brothers, Laurent and Ivan for their brotherly affection and with whom I share many great memories and I would like to include a special mention for my sisters-in-law, Lauren and Penny. Most particularly, I would like to thank my little nephew, Roman, who has been a new joy to our family. He is perhaps the funniest and brightest little boy I've known. During the write-up of my thesis, he would come knocking at my bedroom door for us to hang out and play – a welcome distraction from my writing. I would also like to thank my grandparents, Susan and Albert for their support and well wishes. I would also like to thank my family in France for all their years of affection, help and support: to my grandmother, Jeannine; Pierre & Serge; Claude & Denis; Sylvie & Alain; Maximilienne & Claude; all my cousins and a special thought for Florence and André who, sadly, are no longer with us – 'Un grand merci à tout le monde!'

Even though, I was living far away from all of my family throughout my PhD, the thoughts and memories of them sustained my happiness. With regard to maintaining my sense of determination, I refer to a quote my father passed on to me, attributed to former U.S. president, Calvin Coolidge: 'Nothing in the world can take the place of Persistence. Talent will not; nothing is more common than unsuccessful men with talent. Genius will not; unrewarded genius is almost a proverb. Education will not; the world is full of educated derelicts. Persistence and determination alone are omnipotent. The slogan 'Press On' has solved and always will solve the problems of the human race.'

Table of Contents

Table of Figures

List of Tables

Abbreviations

Ag - Silver
AgNO$_3$ – Silver nitrate
Amp – Ampicillin

Ala – Alanine
Arg - Arginine
Asn – Asparagine

Asp – Aspartic acid
BCA – Bicinchoninic acid
BCIP/NBT - Nitro-blue tetrazolium/ 5-bromo-4-chloro-3'-indolyphosphate
BSA – Bovine serum albumin
BCS/BCDA – bathocuproine disulfonate
Ccsp – Cytosolic copper storage protein
CD - Circular dichroism
CNBr – Cyanogen bromide
CsoR – Copper sensitive operon regulator
Csp – Copper storage protein

Cu - Copper
CuCl – Copper chloride
Cys – Cysteine

Da – Dalton
DDM – n-Dodecyl-β-D-Maltoside
DTT – Dithioreitol
ECF – Energy coupling factor
ECuC – Extracellular Cu chaperone
GlxA – Galactose oxidase A
H$_2$O$_2$- Hydrogen peroxide
HCl – Hydrochloric acid
HEPES - 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HGT - Horizontal gene transfer
His – Histidine
IPTG – Isopropyl β-D-1-thiogalactopyranoside
Kan – Kanamycin
k_{Cu} – Copper dissociation rate constant
k_d - Dissociation constant
k_{obs} – Observed first-order rate constant
LB - Luria Bertani
LCA – Last common ancestor
LCP – Lipidic cubic phase
LGT - Lateral gene transfer

LMCT - ligand-to-metal charge transfer
Lys - Lysine
MALDI - Matrix-assisted laser desorption/ionisation
MALS – Multi-angle light scattering
Mbn - Methanobactin
MES - 2-(N-morpholino) ethanesulfonic acid
Met - Methionine
MgCl$_2$ – Magnesium chloride
MOPS - 3-(N-morpholino) propanesulfonic acid
MRE - Mean Residue Ellipticity
NaCl – Sodium chloride
Ni-NTA - Nickel-nitrilotriacetic acid
PDB - Protein Data Bank
PDC – Protein detergent complexes
PCR - Polymerase Chain Reaction
pMMO - Particulate methane monooxygenase
PMSF - Phenylmethylsulfonyl fluoride
Sco - Synthesis of cytochrome c oxidase
SDS-PAGE - Sodium dodecyl sulphate polyacrylamide gel electrophoresis
SEC – Size exclusion chromatography
Ser – Serine
SFM - Soya flour mannitol
Sli - Streptomyces lividans
sMMO - Soluble methane monooxygenase
Tat - Twin arginine translocation
TEMED - Tetramethylethylenediamine
TM - Transmembrane
TMPD - N,N,N',N'-tetramethyl-p-phenylenediamine
Tyr - Tyrosine
UV - Ultra violet
WT – Wild type

YT – Yeast tryptone

Chapter One

Introduction

1.1　An introduction to Copper

Copper (Cu) is a d-block transition metal with valuable properties to humanity that have been exploited for thousands of years. Cu derives from the Earth's crust and is one of the few metals that naturally occurs in the environment (25). The first recorded use of Cu dates back between 5th and 6th millennia B.C. as it was found in metallic form in nature thus smelting was not essential (26). The antimicrobial properties of Cu were evidenced through initial uses in treating burns, wounds, headaches and for hygiene by Aztec, Greek and Roman civilisations (26).

Cu can exist in one of two oxidation states, the cupric oxidised form (Cu(II)) or the cuprous reduced form (Cu(I)). It is this redox chemistry that can cause toxicity problems in all organisms. For example Cu can participate in Fenton type reactions (27) (Equation 1.1),

$$Cu^+ + H_2O_2 \rightarrow Cu^{2+} + OH^- + OH^{\cdot} \qquad 1.1$$
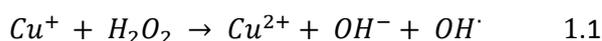
producing hydroxyl radicals that can be involved in several processes that lead to damaging oxidation of lipids and proteins (26, 28). Cu ions can take part in a cycle (Equations 1.2 & 1.3) that depletes sulfhydryls in proteins/peptides (e.g. cysteine, glutathione) (26). This cycle produces hydrogen peroxide whereby it can partake in the Fenton reaction (Equation 1.1), thus triggering further damage to the cell (26).

$$2\,Cu^{2+} + 2\,RSH \rightarrow 2\,Cu^+ + RSSR + 2H^+ \qquad 1.2$$

$$2\,Cu^+ + 2\,H^+ + O_2 \rightarrow 2\,Cu^{2+} + H_2O_2 \qquad 1.3$$
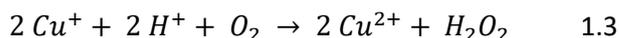
The amino acid residues that demonstrate preferential binding to the soft Cu(I) ion possess thiol and thioether groups and thus include cysteine and methionine residues (25). The hard Cu(II) ion can also coordinate with imidazole nitrogen and oxygen groups such as histidine and aspartic amino acids (25). Relative stabilities of metal-ion complexes as described in the Irving-Williams series (Mn(II) < Fe(II) < Co(II) < Ni(II) < Cu(II) > Zn(II)) suggests that Cu is capable of displacing certain metals from their binding sites in metalloproteins thus disrupting enzymatic processes (25, 29). An example of this includes iron-sulphur cluster proteins in which thiolate bonding occurs between sulphur and Cu(I) which is detrimental to these types of proteins (29-31).

The antibacterial properties of Cu are now recognised to have a role in the host-pathogen response in humans. The host innate immune system will employ several mechanisms to kill the invading pathogen. These include use of bactericidal toxins and to limit availability of nutrients to starve the pathogen (nutritional immunity) (30, 32). In addition, the host will expose the pathogen to Cu. There is growing research into how Cu toxicity is used by the human immune system to destroy invading pathogens; the "Cu burst" (33). The use of

increased concentrations of Cu is utilised in the host macrophage as it engulfs the pathogen. During phagocytosis, the use of plasma membrane CTR1 copper importers to transfer Cu to cytoplasmic ATOX Cu(I) chaperone is achieved. The Cu(I) is transferred to ATP7A copper pump which delivers Cu(I) to the phagosome where the pathogen lies (30, 32). The Cu undergoes Cu(I) Fenton chemistry to produce hydroxyl radicals which derive from hydrogen peroxide generated by superoxide produced by NADPH oxidase (NOX) (30, 32). Indeed, this defence system against invading pathogens seems efficient in the prevention of infection. However, there are existing Cu resistance systems in pathogens that act to combat Cu toxicity. Indeed, the environment of copper surfaces or salts could be fatal to most bacteria, but many have developed robust Cu tolerance genes as a counteractive measure (30, 34, 35). It is well known that this involves expression of one or more Cu-exporting ATPases that are often under the control of a Cu(I) metalloregulator. Examples of this type of system will be given later in this chapter for the non-pathogenic *Streptomyces lividans*, but the interesting aspect is that the same mechanism can be utilised in bacteria as part of their pathogenicity.

1.2 Importance of Cu as a cofactor in proteins

Cu is required by many proteins and enzymes to carry out electron-transport processes, oxygen activation, denitrification and many other functions (36, 37). To achieve these various functions, a variety of Cu active sites have evolved. Examples of such sites can be found in Fig. 1.1. An example of a denitrifying enzyme is nitrous oxide reductase which is a multicopper protein and is involved in reducing N_2O to produce dinitrogen and H_2O (38). The multicopper sites of this protein consists of a Cu_Z centre which is a tetranuclear centre and a dinuclear Cu_A centre that maintains either one or two sulphide bridges (39). Another important example of protein that requires Cu in order to function is cytochrome *c* oxidase (CcO) which utilises Fe and Cu as cofactors (40). CcO possesses a binuclear heme-Cu centre and is involved in cellular respiration by reducing molecular oxygen combined with proton pumping activity across either a bacterial membrane or eukaryotic mitochondrial membrane (40). There are four active sites in CcO which are metal binding that include Cu_A, heme *a*, heme *a3*, and Cu_B, (40). Indeed, CcO has remained as one of the most studied metalloenzymes and the characterisation of its dinuclear Cu_A site and mononuclear Cu_B sites has greatly aided the understanding of its role in the mitochondrial electron transport chain (41-44). Another example of a Cu metalloenzyme are laccases which belong to the class of blue multicopper oxidases. Laccases can oxidise a range of organic aromatic compounds in combination with reducing molecular oxygen to water (45). These proteins are found in both eukaryotes and prokaryotes but are most abundant in fungi (46). Laccase active sites contain four Cu ion
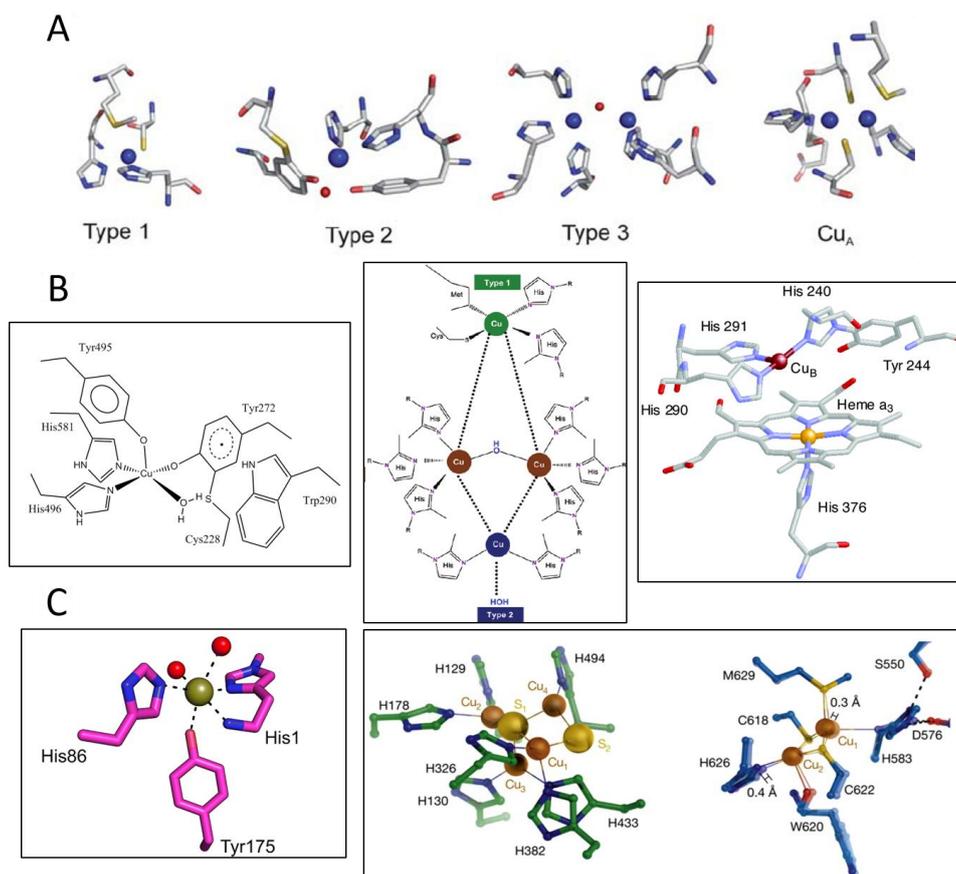
Figure 1.1 – *Examples of copper active sites in proteins (A) Various types of copper sites which includes Types 1-3 and a Cu_A site (image taken from (14)). (B) (From left to right) Copper active sites of the following proteins; Galactose oxidase displaying monocopper active site; Laccase with a binuclear type 3 copper centre, type 1 and 2 copper centres (image taken from (21)); Cytochrome c oxidase showing the binuclear heme-Cu centre (image taken from (22)). (C) (From left to right) Copper active sites of the following proteins; LPMO with a mononuclear Cu ion in 'His brace' binding site (image taken from (23)); Nitrous oxide reductase, left hand image shows catalytic Cu_Z site in the N-terminal domain whereas the right hand image displays the C-terminal domain Cu_A site (images taken from (24))*

binding sites (46). These binding sites include a mononuclear type 1 blue Cu site close to the substrate binding pocket, the second site is a trinuclear Cu site consisting of one type 2 Cu and a type 3 coupled dinuclear centre (47, 48). An example of a monocopper oxidase is the fungal protein galactose oxidase (49). This enzyme catalyses the oxidation of primary alcohols groups ($R-CH_2OH$) to aldehydes with a coupled reaction, involving a two-electron reduction of O_2 to produce H_2O_2 (49). Another class of proteins that utilises Cu as a cofactor are lytic polysaccharide monooxygenases (LPMOs). An LPMO possesses a mononuclear Cu ion in its active site which is coordinated in trigonal manner by a 'histidine brace' and the N-terminal amino group (50). The single Cu ion centre is involved in activating oxygen which leads to the process of oxidatively cleaving $\beta-1$-4 glycan linkages in crystalline polysaccharides, such as chitin, cellulose and starch (50).

1.3   Methods to determine Cu(I) affinity

The techniques required to analyse the Cu(I) content and behaviour within bacterial cuproteins are paramount. Such methods include the use of Cu(I) binding ligands as affinity standards such as 2,2' bicinchoninic acid (BCA) (Fig. 1.2). This compound has been extensively used in assays for determining
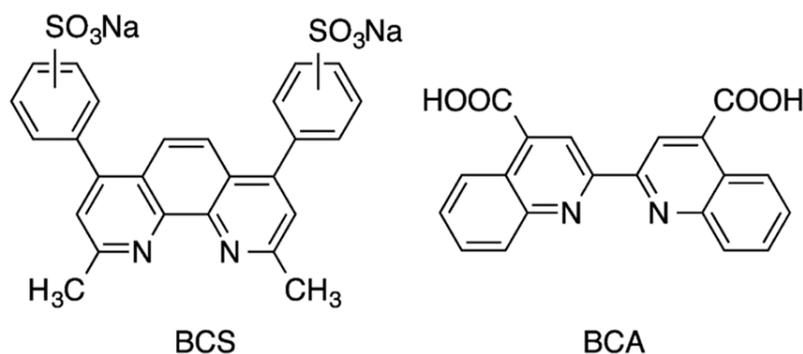
Figure 1.2 – (Image taken from paper by Bagchi et al. (13)). Compounds BCS and BCA.

protein concentrations since its discovery by Smith *et al*. in 1985 (51, 52). BCA facilitates two reactions in this assay; firstly, the biuret reaction at the peptide bonds and certain residues in an alkaline milieu which involves reducing Cu(II) ions to Cu(I) ions (51, 52). Secondly, two BCA molecules chelate with one Cu(I) ion to form a chromogenic purple complex which absorbs strongly at a wavelength of 562 nm (51, 52). Indeed, BCA is a commonly used ligand to determine Cu(I) concentration and has even been used to determine Cu(I) concentrations in biological serum samples due to its high sensitivity and specificity for Cu(I) (53). Another high affinity Cu(I) binding ligand is bathocuproine disulfonate hydrate (BCS) (Fig. 1.2) which also binds Cu(I) in a bidentate manner similar to BCA. Upon Cu(I) binding, the BCS-Cu(I) chromogenic complex is orange in colour and absorbs strongly at 483 nm (54). Xiao *et al*. have documented the issues of inconsistent metal binding constants being reported in the literature (54). These issues are caused by several factors including minimal control of pH and how this affects affinities of ligand probes and disregard of metal affinities of pH buffers (54). Following carefully laid out protocols, BCA and BCS have become the standard ligands to determine the Cu(I) affinities of proteins under competitive conditions.

Methods in quantifying Cu(I) concentration and binding affinities of cuproteins are ever evolving to improve their accuracy. For example a study by Bagchi *et al*. suggested high-affinity ligand stability constants are more reliably attained by carrying out competitive binding studies by using an affinity standard with a well-established solution chemistry and stability constant (13). In the study by Bagchi *et al*., they characterised three new monovalent Cu binding ligands that are water soluble; MCL-1, MCL-2 and MCL-3 (13). It was reported that

all three ligands form 1:1 stoichiometry with Cu(I) and have been used to predict binding affinity of an *Escherichia coli* metalloprotein, CusF (13).

1.4 Bacterial copper storage proteins

An example of a bacteria that require Cu for essential cell maintenance and biological reactions is methanotrophic mycobacteria (1). Vita *et al.* report that these bacteria require a substantial amount of Cu to oxidise methane. These prokaryotes can accomplish this by using Cu for their specialised membranes that store methane monooxygenase (1). These Gram-negative bacteria possess a regular Cu efflux system and aerobically oxidising methane to methanol involves two differing families of metalloenzymes which are particulate methane monooxygenase (pMMO) (55, 56) and soluble methane monooxygenase (sMMO) (55, 57). Both proteins are regulated by Cu and some methanotrophs possess both metalloenzymes (55). In particular, pMMO is located in intracytoplasmic membranes and is produced in the presence of Cu (55, 58, 59). It is noted that pMMO is a trimeric protein with three Cu sites (55, 56) and is widely expressed, with around a fifth of the cellular protein mass comprised of this protein (55, 60). Indeed, methanotrophs have provided much information in bacterial systems involving procurement of Cu.
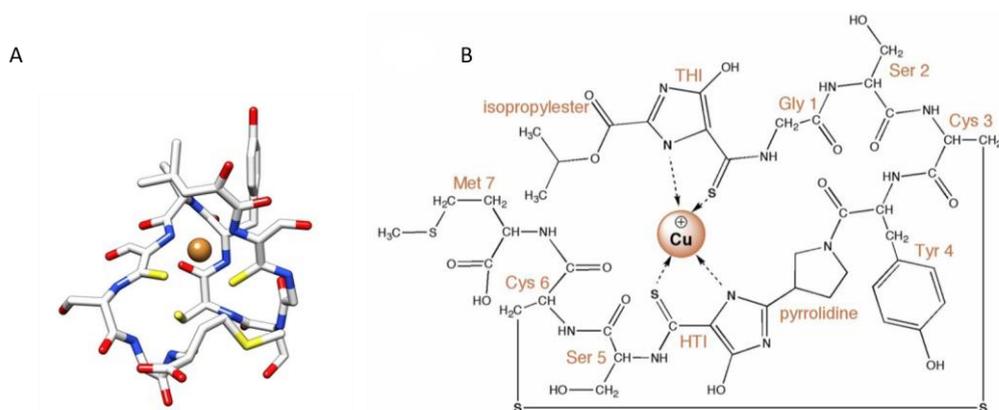
Figure 1.3 – *Structure of Mbn from M. trichosporium OB3b. (A) Stick representation of crystallographic structure of Mbn (PDB file 2XJH) (6). (B) Diagram of Cu-Mbn (image taken from (16))*

Methanotrophs store Cu in modified peptides known as methanobactins (Mbns), which aid in detoxifying the cytosol from high Cu concentrations (1, 55, 61). Mbns have a high affinity for Cu which they chelate through thioamide/enethiol moieties and paired nitrogen-containing hetero cycles (55, 62). These proteins are key members of Cu binding metallophore family known as chalkophores (55, 61). It is unsurprising to find methanotrophs in near-neutral pH, organic rich habitats such as lake sediments, peatlands and rice paddies due to the high Cu affinity of Mbn and their ability to remove Cu from such environments

(55, 63, 64). The first Mbn to be structurally characterised was from *Methylosinus trichosporium* OB3b in 2004 (55, 62). The structure of *M. trichosporium* Mbn (Fig. 1.3) revealed the heterocycles to be oxazolones (Oxa$_A$ and Oxa$_B$), as previous research had mis-assigned these as imidazolone rings (55, 65). Two absorption peaks are observed at 345 nm and 392 nm for both oxazolone rings in the apo-form and this absorption range is typical for oxazolone compounds (55, 66). In addition, Mbn can bind both Cu(I) and Cu(II) but in the absence of Cu, the oxazolone rings are susceptible to acid-catalysed hydrolysis (55, 67, 68). It has been found that Cu binding by these oxazolone rings is a reductive process, because even after Cu(II) binding, the final species is Cu(I)-Mbn (55, 69, 70). Mbn is capable of binding other metal ions but with lower affinity compared to Cu(I) (55, 71), for example Ni(II), Co(II), Ag(I), Pb(II), Fe(III) and Zn(II) (55, 71). There is speculation that Mbn interacts with pMMO but there is lack of evidence to prove this (55). It has been reported that pMMO activity is increased in the presence of apo-Mbn (55, 72). Though, this may be due to apo-Mbn binding inhibitory metals since pMMO becomes inactivated by excess amounts of zinc or copper (55, 73, 74).
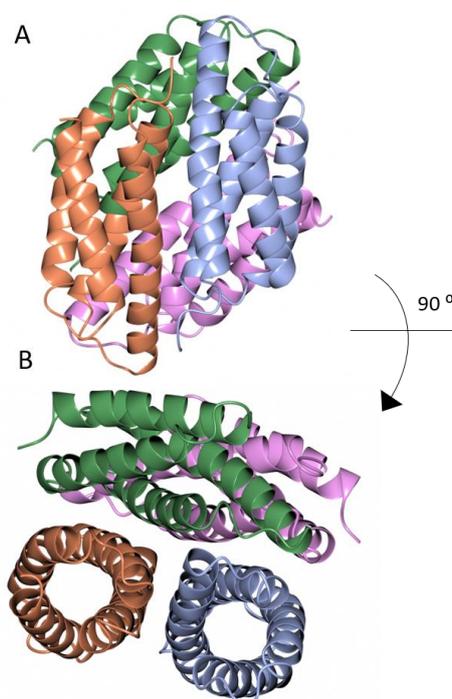
Figure 1.4 – X-ray crystallographic structure of apo MtCsp1 (1) (A) showing the tetramer arrangement of four coloured alpha helical bundles (B) and a 90˚ rotation of fig. (A)

Until recently is was a long-held view that prokaryotes do not possess any proteins that store Cu (1). Vita *et al*. discovered and isolated a soluble protein present in the methanotroph *Methylosinus trichosporium* OB3b that is capable of binding a high number of Cu(I) ions and

was named copper storage protein 1 (Csp1) (1). The Csp1 was found to contain 122 amino acids with a molecular mass of 12,591.4 Da (1) and shown crystallographically to be a four helix-bundle which assembles in to a homotetramer (Fig. 1.4). In the core of each four helix-bundle, 13 Cys residues are present (Fig. 1.5A). These residues are not involved in disulfide bonding and were shown to primarily bind Cu(I) ions (1) (Fig. 1.5B). Thus, Csp1 can bind up to 52 Cu(I) ions per tetramer assembly and contains a signal peptide that suggest it is a periplasmic protein (1). Signal peptides are 16-30 amino acids in length that are fused to newly synthesised proteins and are involved in transporting that protein within or outside the cell. The specific signal peptide found in Csp1 was a twin arginine translocation (Tat) peptide (1) involved in transporting proteins outside the cytosol. Vita *et al.* identified another homologous protein known as Csp2 from *M. trichosporium* OB3b which also possessed a signal peptide. Notably a third homologue was identified, called Csp3, but this homologue differed in that it did not possess a signal peptide and therefore it was inferred that Csp3 was a cytosolic protein (1). Bioinformatic analyses, revealed the Csp3 homologue is widespread in non-methanotrophic prokaryotes.
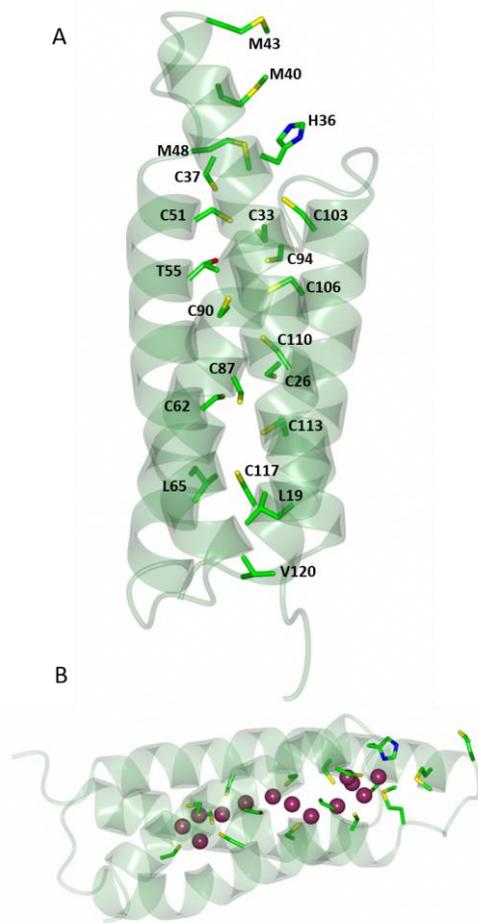
Figure 1.5 - X-ray crystallographic structure of MtCsp1 protomer (1, 2) (A) (Image taken from paper by Vita et al. (2)) Labelled amino acids include 13 Cys residues in the inner core of the helical bundle. The pore opening also includes residues Met40, Met43, Met48 and His36. The hydrophobic end of the bundle includes residues Leu19, Leu65 and Val120 (2) . (B) All 13 Cys residues are involved in Cu(I) binding and the protomer unit of MtCsp1 binds a total of 13 Cu(I) ions.

The X-ray structure of *Mt*Csp3 reveals 18 cysteine residues in the core of the four helix-bundles (1) which, like Csp1, can coordinate Cu(I) but with a higher capacity due to the additional cysteine residues, enabling 76 Cu(I) ions to bind (2, 75, 76). *In vivo* data suggested a role of these Csp3s in preventing Cu(I) toxicity within the bacterial cell (2, 75). The binding affinities for *Mt*Csp3 and *Bs*Csp3 are $(1.7 \pm 0.5) \times 10^{17}$ M^{-1} and $(1.5 \pm 0.4) \times 10^{17}$ M^{-1} respectively and unlike Csp1 do not display Cu(I) binding cooperativity (2, 75). Vita *et al*. have shown Cu(I) removal experiments between the Csp3s and the Cu(I) specific ligand BCS which has a higher Cu(I) affinity for Cu(I) that Csp3, and shows that Cu(I) removal is slow from these proteins, it was reported by Vita *et al*. that after 85 hours, ~85% Cu(I) was removed from *Bs*Csp3 by BCS whereas only ~20% was removed from *Mt*Csp3 in the same period of time (2, 75). This is a unique finding for these Csp3s whereas previous studies with *Mt*Csp1 (1) have shown rapid Cu(I) removal by this ligand (2). These Csp3s possess three His residues that participate in coordination of Cu(I) at one end of the four helix-bundle, which has been proposed to be the entry site for Cu(I) ions into the Cys core (2). These His residues (Fig. 1.6A) are highly conserved in Csp3s and are found in both *Mt*Csp3 and *Bs*Csp3 (2, 75). In contrast, *Mt*Csp1 has three Met residues at the opening of the bundle (1, 2) (Fig. 1.5A). Another structural feature of *Mt*Csp3 includes a coordinating Asn58 with Cu(I) shared with Cys13 (2, 76). An Asn residue at this position is unique to *Mt*Csp3 and an Asp residue is more frequent in Csp3s and is highly conserved. (2, 76). The structural aspects of Cu loading in Csp3s has been established in the study by Basle *et al*. (76). Whereby

X-ray crystallographic analyses display semi-Cu loading into these proteins (76). The complete filling of the protein core with metal ions is unique in Csps and until now was not seen in either engineered or naturally occurring four helix-bundles (76-81). In addition, the presence of thiolate Cu(I) clusters is uncommon in biological systems thus the discovery of Csps offers a new perspective on prokaryotic copper storage (76). In the study by Basle *et al.*,

MtCsp3 seems to demonstrate binding of partial occupancy Cu(I) ions at different sites (76). For instance, MtCsp3 loaded with ~2 molar equivalents of Cu(I) shows partial occupancies of four Cu(I) ions in the structure which are arranged in a symmetrical tetranuclear cluster (76). Additionally, MtCsp3 was subsequently loaded with ~9 equivalents of Cu(I) to give a structure displaying 18 Cu(I) ions present with partial occupancy. But the occupancy of these Cu(I) ions was higher compared to the structure loaded with ~2 Cu(I) equivalents giving an overall occupancy value of 8.2 versus 1.3 (76). Thus, it is apparent that the occupancy values of individual Cu(I) ions rises upon Cu(I) loading and appears to be a key feature of organothiolate-coordinated tetranuclear clusters (76). These formations are considered uncommon in biological systems due to their latent toxicity and complexity

Figure 1.6 – X-ray crystallographic structures of MtCsp3 (A) (Image taken from paper by Vita et al. (2)) Protomer of apo MtCsp3 showing 18 Cys residues and an N-terminal α -helix (α$_N$). The pore opening includes three His residues (His104, His108 and His110). The hydrophobic end of the protomer highlights amino acids Leu21, Leu83 and Leu87. (B) Fully Cu(I) loaded MtCsp3 protomer that binds a total of 18 Cu(I) ions.

but are often found in inorganic complexes (75). Despite this, various proteins have emerged to form these Cu(I) clusters due to the evolution of Cu(I) handling in both eukaryotes and prokaryotes (75).

1.5   *Streptomyces* development

The *Streptomyces* genus belonging to the phylum Actinobacteria, offers many possibilities in the domain of biotechnology where certain enzymes/proteins and metabolites can be used for industrial and pharmaceutical purposes. This type of work was also greatly aided by the advent of new DNA sequencing technologies at the turn of the 21st century, enabling for the complete sequence of *Streptomyces coelicolor* A3(2) to be determined (82). This bacterium is the best representative bacterium of the *Streptomyces* genus (82). The vast amount of proteins coded in the genome of *S. coelicolor* offers many research possibilities and allows to perhaps better understand certain pathways and regulatory functions in other similar organisms through comparison. For example, *S. coelicolor* harbours cognate signal sequences and other components to support the TAT (twin arginine transport) pathway as well as the Sec system (82). The understanding of pathways such as this may offer possibilities in manipulating other similar organisms for biotechnological/biopharmaceutical purposes. Also, Streptomycetes produce a range of antibiotics, anti-tumour agents, cytostatics, fungicides and other secondary metabolites that have pharmaceutical properties that are highly used in medicine and other industries (83). For instance, the antibiotic streptomycin was first discovered in *Streptomyces* by Waksman (83). In addition, this genus is often used for large scale production of enzymes (84).

The growth and morphology of Streptomycetes is unlike other typical prokaryotes; this genus almost mimics a fungal life cycle whereby spores germinate to firstly produce two germ tubes (14). These evolve into vegetative or substrate mycelia which grow downwards into their soil habitats to create an extended network of hyphae (Fig. 1.7). Once nutrients
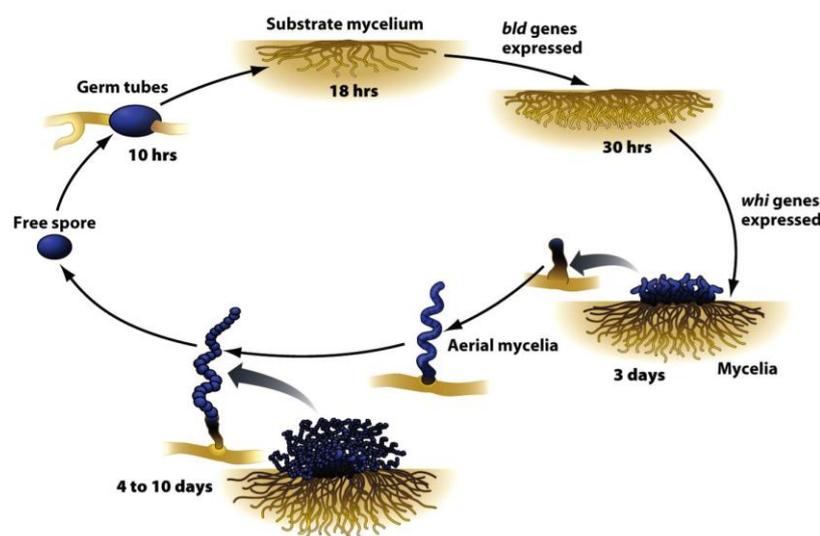
Figure 1.7 *– Life cycle of Streptomyces*

become scarce, substrate mycelium develops into aerial mycelium (76). In this stage of morphogenesis, aerial hyphae are divided through septation and sporulate (76). Indeed, this growth cycle of *Streptomyces* had challenged the classification system in the past; during the 1950s researchers considered this genus as mainly fungi (85), whereas others classified it as a group between bacteria and fungi due to its Gram positive cell wall characteristics and vulnerability to anti-bacterial antibiotics (85). Overall the most fascinating stage of this life cycle is aerial hyphae formation. In response to lack of nutrients and other signals, this stage produces secondary metabolites (as well as aerial hyphae) that include compounds that have been shown to have antibiotic and antitumor properties (14, 86). Thus, it is not surprising that this stage has been of great research interest to pharmaceutical industries. The aerial hyphae continue to grow out of the vegetative mycelium in the aqueous environment and develop into the air (86). The hyphae transform into extended chains of pre-spore compartments which is due to a type of regulated cell division (86). In this growth phase proteins from the SsgA-Like Protein (SALP) family regulate this complex cell division (14). These pre-spores eventually form thick spore walls and evolve into mature spores (86). Finally, these matured spores are released from the spore chains into the environment whereby this mycelium growth can be repeated (14) (Fig. 1.7).

It is believed that the function of many of the secondary metabolites produced during the switch between substrate and aerial mycelia within *Streptomyces* is to inhibit growth of competitive microorganisms (85, 87, 88). The genes that are responsible for antibiotic production has been discussed in a review by van Wezel *et al*. It has been discussed that gene clusters for antibiotic production are regulated and transduced by cluster-situated (transcriptional) regulators (CSRs) (88, 89). CSRs directly control the transcription of these genes that encode enzymes involved in producing these antibiotics. Examples of gene clusters regulated by CSRs include streptomycin antibiotic in *Streptomyces griseus* and actinorhodin in *S. coelicolor* (88). However, van Wezel *et al*. mention that these two antibiotic coding gene clusters are regulated by transcriptional regulators which are StrR (88, 90) and ActII-ORF4 (88, 91) , respectively. Indeed, there are existing examples of multiple CSRs that control biosynthetic genes, such as five transcription regulators that control expression of tylosin cluster in *Streptomyces fradiae* (88). It has been found that the production of antibiotics can be increased several-fold by cloning URAPs (ultimate (pathway-specific) regulator of antibiotic production) in high-copy number plasmids (88). URAPS are similar to CSRs and categorised as the final downstream regulators (88).

1.6 Cu bioavailability for *Streptomyces* development

For certain *Streptomyces* strains, a distinct dependence on the bioavailability of Cu to initiate the morphological development switch between substrate mycelia and aerial hyphae is known (86, 92). Indeed, Cu is needed for production of spores and aerial hyphae thus making it an important element to this genus. In order to exploit the valuable secondary metabolites produced in *Streptomyces*, an understanding of

R5 morphology
solid

0.1   1   10 µM Cu(II)
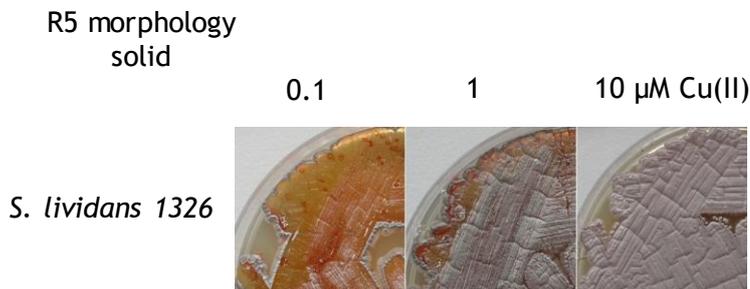
S. lividans 1326

Figure 1.8 – *S. lividans morphological development at increasing concentrations of Cu(II) on agar R5 medium (contains yeast extract and glucose as carbon source). Aerial hyphae(grey) and spores develop as the Cu(II) bioavailability increases.*

how Cu is used must firstly be obtained. Using *Streptomyces lividans* as an example, it was found that the vegetative growth phase in *S. lividans* can occur even at low Cu availability (14). The use of Cu is essential for the complete maturation of aerial hyphae and spores which has a distinct phenotype of a grey pigment colour (Fig. 1.8) The development of aerial mycelium is blocked in mutant strains of *Streptomyces* such as *S. lividans* with defective Tat (twin-arginine translocation) secretory pathway (14, 93). There are two pathways that occur in *Streptomyces*; Tat and Sec (secretion pathway), the Tat pathway involves secretion of folded proteins across the cytoplasmic membrane that contain cofactors including metal ions (14). The Sec pathway involves simply the secretion of unfolded proteins out of the cell and is was believed that *S. lividans* relied mainly on this pathway (84). But the Tat pathway was identified (84, 94) and is apparently used as a general pathway for the transport of a large number of Tat substrates and is not specific to enzymes with cofactors (14, 95). In the mutant Tat strains, morphological development can be re-established with increased copper supplementation, thus suggesting that secreted Cu proteins are essential for *Streptomyces* development (14).

The formation of aerial hyphae that transform into extended spore chains with cell division producing septa arranged into ladders (96-98) poses issues from a biotechnological perspective. *Streptomyces* tends to form 'pellets' or 'clumps' in liquid environments due to this mycelial growth (98). This very manner of growth is one of the main obstacles of using *Streptomyces* as an expression host. There has been recent research to overcome this issue as highlighted by van Dissel *et al.* which discusses a gene cluster target involved in mycelial aggregation which upon modification could lead to better performance of *Streptomyces* as

cell factories (98). It was stated by van Dissel *et al*. that the formation of mycelial structures occurs when growing Streptomycetes in submersion and overall forms pellets which on an industrial scale poses several issues including culture heterogeneity, mass-transfer impediments and slow development (98). van Dissel *et al*. aimed to address this issue of pellet formation by identifying the genes involved (98). This was achieved by firstly using a chemostat to grow and select over 100 generations of *S. lividans* to obtain a loose pellet forming mutant (PM01) and a non-pelleting strain (PM02) (98). Thus, with the use of various growth limiting substrates, PM02 was analysed in continuous culture for the plasmids' segregational stability which are plasmids *pIJ2* (99) and selected pSG5 plasmids reported in previous research (98, 100). Mutations were identified in both strains PM01 and PM02 – the non-pelleting mutation was discovered in a membrane protein gene (*matA* gene) that is co-transcribed with a bifunctional polysaccharide deacetylase/synthase gene (*matB* gene) (98). Both genes are needed for pellet formation and reverse engineering was performed to decipher an original molecular factor required for pellet development (98). A single point mutation was identified that was responsible for the phenotype of mycelial aggregation and this mutation was labelled *Mat* (98). Based on genetic complementation of strain PM02 studied it was found that the mutations in the *matA* gene was the main target of this morphogenesis (98). Also, it was found that a more dispersed mycelium with increased growth rate in *S. lividans* when deletion of the *mat* genes was performed (98).

Indeed, research behind this pellet formation in *Streptomyces* is essential to ever transform members of this genus into production hosts. The requirement to then grow *Streptomyces* in large bioreactors on an industrial scale would require the growth to produce a more fragmented mycelium and inhibit pellet formation altogether. The exact biochemical mechanism behind pellet formation has yet to be fully elucidated. However, proposed mechanisms in *S. lividans* have been offered.

This begins with the Sco operon discovered in *S. lividans* which contains three genes (*SLI4212, SLI4213, SLI4214*) that encode cuproteins that are essential to Cu dependent development (44, 101). Specifically, the Sco protein (*SLI4214*) has been shown to participate in the Cu driven development of *S. lividans* (44). Sco proteins are best known for the role in delivering Cu to the Cu_A site of CcO. However it was discovered that this development can occur in the absence of CcO and thus it was suggested that Sco participated in a branched Cu-trafficking pathway, whereby one branch delivered Cu to CcO and the second branch to a protein required for morphological development (44).

The paper by Petrus et al. discusses the cuproenzyme GlxA in S. lividans alongside the newly discovered DyP-type peroxidase (DtpA) and how these proteins function in a Cu regulated pathway in the hyphal tips of the aerial mycelium of S. lividans (12). At these hyphal tips, extracellular glycans are produced which are essential to pellet formation when S. lividans is grown in liquid culture (12, 92, 102). The role of GlxA in S. lividans involves a crucial Cu dependent morphological development in this bacterium. GlxA active site contains a redox cofactor that is a cross-linked Tyr-Cys along with a mononuclear Cu ion (103). According to Chaplin et al., GlxA is similar to fungal galactose oxidase (Gox) in terms of its Cu co-ordination geometry but different to Gox based on spectroscopic data (103). Complete loss of glycan development in the hyphal tips of S. lividans strains containing GlxA null-mutants grown in both liquid and solid culture has been reported (103). It was observed that the GlxA null-mutant did not respond to Cu supplementation and thus supports a hypothesis that GlxA is essential for development of an as yet unknown glycan (103). Thus, the GlxA null-mutant
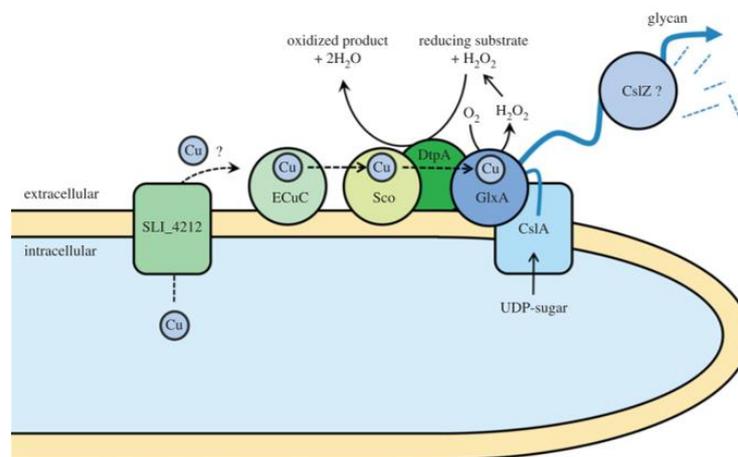
Figure 1.9 - *(Image taken from paper by Petrus et al. (12)) Illustrated pathway for hyphal tip formation in S. lividans. An uncharacterised transporter SLI_4212 is involved in transporting intracellular Cu to the extracellular space. Lipoprotein ECuC transfers Cu to Sco chaperone. DtpA is involved in changing Cu(I) to Cu(II) which is required for GlxA maturation. This is followed by Sco chaperone transferring Cu to GlxA. GlxA generates H_2O_2 during oxidation of its substrate which is possibly removed by DtpA. Cellulose synthase protein CslA cooperatively functions with GlxA to produce extracellular glycan which in turn could be possibly treated by endoglucanase CslZ (12)*

offers promise of creating the desired morphology of open mycelium in liquid culture; this brings possibilities of using S. lividans as an heterologous expression host of valued enzymes (103).

To place GlxA into the context of the Cu trafficking pathway involving DtpA (Fig. 1.9), the extracytoplasmic Cu chaperone (ECuC) transfers Cu ion to the extracellular Sco chaperone. The Sco chaperone delivers this Cu ion to GlxA which incorporates it into its active site to initiate enzyme activity (12). DtpA would act to removes hydrogen peroxide produced

by GlxA thus maintaining its stability (12). A cellulose synthase protein, deriving from family 2 of glycosyl transferases is encoded by the CslA (12, 104) and is proposed to act cooperatively with GlxA to produce an extracellular glycan (synthesizes a β(1,4)-glycan at hyphal tips) which could then be processed by CslZ (endoglucanase) (12). It was found that mutation of the CslA gene inhibits aerial growth and thus pellet formation when *S. lividans* is grown in a submerged environment (12, 105, 106).

The genes for *CslA, GlxA* and *CslZ* form a gene cluster in *S. lividans* and this operon can be found in most Streptomycetes (12). Deletion of either *glxA* or *cslA* inhibits morphological development in *S. lividans* leading to poor glycan formation and overall absence of pellets (12, 103, 107). As discussed in the paper by Chaplin *et al*., *in vivo* null mutants of *glxA* and *cslA* were created and by adding exogenous Cu(II) to these mutants, this does not restore *S. lividans* morphological growth (103). In addition, the lack of Cu inhibits aerial hyphae and pellet formation in the null-mutants *glxA, dtpA, sco and cslA* except for *cslZ* (12). It was found via Western analysis that *dtpA* and *sco* mutants inhibit GlxA development but with the addition of Cu, this morphology can be salvaged (12). The research by Petrus *et al.* provides a Cu pathway whereby GlxA and CslA are essential to the morphological development in *S. lividans* which functions with the Tat-secreted DtpA (12).

1.7   Cytosolic Cu control in *S. lividans*

As discussed above, Cu is needed for *S. lividans* morphological development, but it is essential to regulate the amount of Cu in the cytoplasm to prevent toxicity. This is achieved via various cytosolic Cu metallochaperones as they are essential for transporting Cu to areas within cells where Cu is required. Worrall *et al.* initially identified two operons in *S. coelicolor* that encoded for a CopZ-like Cu metallochaperone and a CopA-like P1-type ATPase transporter (14). Further work by Chaplin *et al.* revealed a total of four *copZ* and five *P-type ATPase* encoding genes in *S. lividans* (17). A CsoR operator sequence was found to precede the *copZ* genes (17, 108), indicating Cu regulation was under the control of a copper sensitive operon repressor (CsoR) metalloregulator. The CopZ Cu(I) binding motif was deduced as $MX_1CX_2X_3C$ with two Cys thiols which is typical of Atx1 type proteins (17, 109-111). Chaplin *et al*. discovered the K_D for two of the CopZ chaperones in *S. lividans* were 2.1×10^{-17} M and 3.7×10^{-18} M (17). The mechanism used by most bacteria to avoid Cu toxicity is by buffering the cytosol but if this method is insufficient then Cu sensors (metalloregulatory proteins) stimulate the expression of efflux system genes such as Cu exporting P_1-type ATPases that restore the cytosol to homeostasis (14). Overall, a mechanism of a Cu homeostatic pathway in *S. lividans* was deduced through RNA-seq, promoter probing and other methods (14, 17).

CsoR can regulate the Cu concentrations in *S. lividans* which involves transcriptional derepression of Cu efflux genes which include CopZ-like Cu chaperones and CopA-like P-type ATPases (17, 108). This is achieved by cuprous ions binding to apo-CsoR (17, 108, 112). This mechanism overall allows the CsoR regulon to maintain a set level of Cu ions within the bacterial cell under homeostatic and stress induced conditions (17, 108, 112) (Fig. 1.10). Research into cytosolic Cu regulation in *S. lividans* such as this can aid in the understanding of similar systems in other organisms. For instance, the work by Novoa-Aponte *et al.* describes Cu delivery to *P. aeruginosa* CueR transcriptional regulator (113-115) by cytoplasmic Cu chaperones CopZ1 and CopZ2 in both *in vitro* and *in vivo* studies (114). It was highlighted that CopZ2 had a greater role of Cu storage as it was more abundant and has higher Cu affinity due to a His residue in a Cu binding loop (MXCXHC) (114). This is relevant for CopZ chaperone found in *S. lividans* as deletion of this His residue affects Cu affinity (17, 114). It was stated by Novoa-Aponte *et al*. that these two CopZ chaperones operated in different metal pools *in vivo* (114). This is supported by the finding that CopZ1 is responsible for Cu delivery to CueR as CopZ1 has a higher K_D compared to that of CopZ2 (114). It was deduced that CopZ2 had a role in responding to Cu stress (114). Novoa-Aponte *et al*. described experiments involving *P. aeruginosa* being exposed to 0.5 mM Cu^{2+} whereby *copZ2* gene was induced in response to this Cu stress (114).
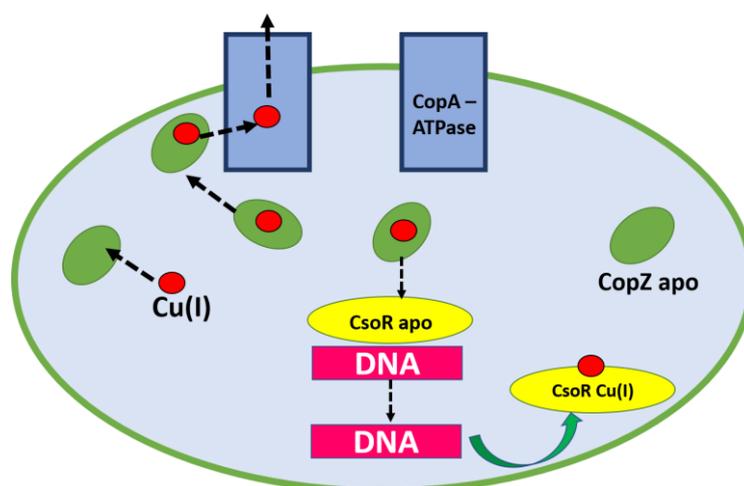
Figure 1.10 *- Schematic diagram displaying the function of CopZ-like copper chaperones in the cytosol of S. lividans when the cell is under Cu(I) stress. These CopZ chaperones contain Cys residues that can bind Cu(I) ions (17). The CsoR is also displayed, which regulates Cu concentrations in S. lividans through the transcriptional derepression of copper efflux genes. As shown in the diagram, this is achieved by CopZ transferring Cu ions to apo-CsoR bound to DNA whereby it dissociates itself (17).*

1.8   New concepts regarding Cu regulation in *Streptomyces*

The existence of Csp3 proteins challenges previous ideas of cytoplasmic Cu storage in prokaryotes and raises the question if other organisms possess these proteins. The presence

Figure 1.11 – *Chemical structure of SF2768 diisonitrile compound from S. thioluteus (5)*

of Csp3 in *P. aeruginosa* (PA2107) has been mentioned in the paper by Novoa-Aponte *et al*. (114). The discovery of Csp3 has led to suggestions that these proteins could act as virulence factors in

pathogenic bacteria (75). The ability of Csp3 to store large amounts of cytosolic Cu is ideal for pathogens that could possess homologues of these Csp3 proteins and utilise them in preventing Cu toxicity inflicted by the host immune system. But there is currently no evidence to justify this virulence factor aspect of Csp3s and still requires further investigation. Csp3 proteins could also provide a better understanding of the copper regulated pathways in *Streptomyces* and other non-pathogenic bacteria. Before the discovery of Csp3, the existence of cytosolic copper storage proteins remained unknown (1, 75). Indeed, to be able to find a homologous protein in *S. lividans* would be beneficial to the goal of utilising this organism as an expression host. Before this can be achieved, the understanding of cytoplasmic Cu regulation in *S. lividans* must be elucidated. Recent findings in how cytosolic Cu affects gene expression, secondary metabolism, spore germination and vegetative growth in *Streptomyces coelicolor* have been reported by González-Quiñónez *et al.* (116). It was found that cytosolic Cu secretion was managed by SCO2730/2731 copper chaperone/P-type ATPase export system during germination (116). The absence of this Cu export system was found to delay germination and sporulation in *S. coelicolor* but enhanced secondary metabolism by 40% (116) which increased production of industrially valuable secondary metabolites (116).

There are significant discoveries of Cu efflux systems in many prokaryotes but very few examples of Cu uptake. Cu uptake has been well studied in eukaryotic systems. A recent study in *Streptomyces* by Wang *et al*. (5) highlights the discovery of a Cu uptake system. Chalkophores have been identified in bacteria that lack Mbn. An example includes diisonitrile natural product, SF2768 (Fig. 1.11), identified in the Gram-positive bacterium *Streptomyces thioluteus* (5, 55). The research by Wang *et al*. (5) discusses the discovery of biosynthetic gene cluster by genome sequencing, labelled as putative nonribosomal peptide synthetase (NRPS) that included the *sfa* operon that produced the compound SF2768 (5). The SF2768 was found to bind extracellular Cu to produce copper-SF2768 complex and its chalkophore activity demonstrated transporting Cu into *S. thioluteus* (5). Through various biochemical *in*

vitro and *in vivo* experiments, the work by Wang *et al*. had uncovered a novel Cu uptake system (5). This included the characterisation of major facilitator superfamily (MFS) exporter operon that is responsible for exporting SF2768, specifically Orf12 transporter (5). SF2768 binds environmental Cu(II) in a reductive manner to produce Cu(I) (5). Wang *et al*. reported that the ABC transporter Orf19-21 transported the copper-SF2768 back within the cell whereby the bound Cu(I) is released and utilised for various cuproteins (5).

The biotechnological possibilities associated with utilising a bacterium such as *S. lividans* requires much consideration in research applications. The need to fully understand the inner regulatory pathways on a genetic, morphological and biochemical level are paramount. Thus, further research into each area is required and investigation into other microorganisms can aid in this task. For example, the discovery of Csps in *M. trichosporium* (1). Such proteins that can be found in *S. lividans* could provide the missing link in Cu regulation as to how exactly Cu is stored in this bacterium. Additionally, the work by Petrus *et al.* in formulating a model of a Cu regulated pathway in development of extracellular glycans in the hyphal tips of *S. lividans* could offer a solution to the issue of pellet formation when this prokaryote is grown in liquid culture (12). This also includes the research carried out by Chaplin *et al.* on the characterisation of LPMO AA10 found in *S. lividans* and its role in glycan formation (117). The overall aim to characterise each component and pathway involved in morphological development in *S. lividans* is crucial as this would allow modification of these components to exploit this bacterium for industrial purposes. Indeed, improving *S. lividans* as an expression host will open up new possibilities for the production of high value enzymes and compounds, which would greatly benefit the biotechnological industry.

1.9   Aim and scope of thesis

The aim of this thesis is to investigate cytoplasmic Cu regulation in *S. lividans*. *S. lividans* is shown to contain a Csp3 and this thesis will explore various aspects of this novel metalloprotein. This includes an in-depth phylogenetic analysis using bioinformatic approaches as well as structural, biochemical and functional characterisation of the protein. There will also be an investigation into the kinetics of Cu(I) loading into this non-methanotrophic Csp3 and initial expression studies of a transmembrane protein which could be a possible component in Cu homeostasis/regulation in *S. lividans* linked to the function of the Csp3 is explored.

Chapter Two

Phylogenetic characterisation and taxonomic distribution of a novel cluster of copper genes in *Streptomyces lividans*

2.1 Introduction

Streptomycetes are Gram-positive bacteria that are members of the Terrabacteria group which belongs to the greater phylum of Actinobacteria. *Streptomyces* is the greatest genus within this phylum and its members are predominantly soil-dwelling organisms. Many streptomycetes display a strong dependence on the bioavailability of copper (Cu) for their morphological development. As indicated in Chapter 1, the bioavailability of Cu governs the morphological switch between vegetative mycelium and aerial hyphae in *Streptomyces livida*ns and understanding this phenomena has been a target of research due to the many useful secondary metabolites produced during this switch (14). The Gram-positive bacterium obtains Cu through its natural surroundings in soil and how Cu is utilised to enable the development switch has been extensively investigated (12, 14, 17-19, 44, 101, 103, 108, 112, 118, 119). The Cu proteome of *S. lividans* contains an array of Cu chaperones (extracellular and cytosolic), a Cu metalloregulator belonging to the copper sensitive repressor (CsoR) family (120), that regulates efflux pumps and several classes of Cu enzymes, including lytic polysaccharide monooxygenases and Cu oxidases (14). Cu is strictly regulated within the cell due to its toxicity and *S. lividans* has developed tightly regulated efflux systems to efficiently remove Cu out of the cytosol when under Cu stress (17, 18, 108, 119). Extensive structural, biochemical and transcriptional characterisation has been carried out on the Cu regulatory systems in *S. lividans* (17, 18, 108, 119), which have been essential for understanding the salient mechanisms of cytosolic Cu handling.

Evolutionary analyses of biological systems, such as genes involved in Cu toxicity and regulation, requires investigating sequence data across the three domains of life; Archaea, Bacteria and Eukaryota. (121, 122). Eukaryotes are known to hold their chromosomes in the nucleus and possess organelles whereas prokaryotes maintain their DNA in a circular plasmid in the cytoplasm. An investigation into the key aspects of the Tree of Life was reviewed by Williams *et al*. (123), whereby recent advances in evolutionary biology suggest the Eurkaryota domain originated from Archaea (123). The rise of eukaryotes occurred through the possible fusion of a bacterial cell and an archaean and lateral gene transfer events occurred thereafter (123, 124). Horizontal gene transfer (HGT) or lateral gene transfer (LGT) involves genetic material moving across regular mating barriers between unrelated organisms. LGT is known to be a great contributor to influencing the evolution of genomes across the Tree of Life (125, 126).

Cu storage proteins (Csp) were first identified in the methanotroph *Methylosinus trichosporium OB3b* and have subsequently been found to store high quantities of cuprous Cu ions (1). Csp1 and Csp2 have Tat export sequences and are secreted to the periplasm and

have been considered to act as a Cu(I) store for particulate methane monooxygenase (pMMO) (1, 55, 127). A third Csp, Csp3, was identified in *M. trichosporium* and is distinct from Csp1/2 through the absence of a Tat signal sequence and therefore Csp3 remains in the cytosol (1). Following the discovery of *M. trichosporium OB3b* Csp3, BLAST searches revealed Csp3 homologs were present in non-methanotrophic bacteria (1).

This chapter reports the identification of a gene encoding for a cytosolic Csp in *S. lividans* (referred to in this chapter as Ccsp) and the identification of two genes in the Ccsp genomic environment that may serve to function alongside the Ccsp. The taxonomic occupancy of these three genes and their evolution across the Tree of Life using similarity searches and evolutionary reconstruction methods has been carried out.

2.2   Methods

Protein sequences were retrieved by using BLAST (128) to carry out similarity searches by using the online version in the National Center for Biotechnology Information webpage (129). The following sequences of these proteins from *Streptomyces lividans* 1326; Ccsp SLI_RS1725541 (Accession number: AIJ15215), SLI_RS17245 (Accession number: AIJ15217) and SLI_RS17250 (Accession number: AIJ15216) were utilised as queries. These searches were carried out against each major group within Bacteria, Archaea, and Eukaryota as listed in NCBI Taxonomy (130). In addition, as *S. lividans* belongs to the bacterial group Terrabacteria, a more in-depth search into this group was carried out. The default e-value threshold of 2E-02 was used in these BLAST searches. For each taxonomic group, the top 5 results (between 1 and 5 hits) were selected. Thereafter, MAFFT was used for multiple sequence alignment using the "Auto" strategy option (15). The online version of Gblocks (131) using the "less stringent" options was used to remove positions of ambiguous alignment. The Maximum Likelihood method was used to construct the phylogenetic trees by using the program FastTree2 (132); the WAG + Gamma evolutionary model of substitutions (133) and a combination of parameters that overall carried out a slow and accurate tree search (*-spr 4 -mlacc 2 -slownni -no2nd*). The Shimdodaira-Hasegawa (134) test was used to calculate the local support values.

2.3   Results

2.3.1   *Identification of a Csp gene in S. lividans and genes that could represent a putative Cu transport system*

The *S. lividans* 1326 genome was searched for genes encoding Csp members using *Mt*Csp3 (1) as input. A gene encoding a putative Csp was discovered between the genes *SLI_36*25 and *SLI_3626* that transcribes on the opposite strand and is thus not part of the *SLI_36*25/*SLI_3626* operon (19, 135). This gene was not originally annotated in the *S. lividans*

1326 sequence but was later given the locus-tag *SLI_RS1725541* in the *S. lividans* Genbank annotation (CM001889) (19, 136). The upstream and downstream genes were given new locus tags; *SLI_RS17250* (old tag *SLI_3625*) and *SLI_RS17260* (old tag *SLI_3626*) (19). The *SLI_RS1725541* gene translates into a protein sequence of 136 amino acids that is lacking a recognisable export signal sequence (19) and contains 18 Cys residues whereby 17 are in a CXXXC or CXXC motif (19). Sequence alignment with Csp3 proteins from other bacteria (Fig. 2.1) revealed strong sequence conservation of the Cys residues, and together with the absence of a signal peptide, this strongly suggested *S. lividans* 1326 possessed a Csp3 member. The protein was named Ccsp (cytosolic copper storage protein).
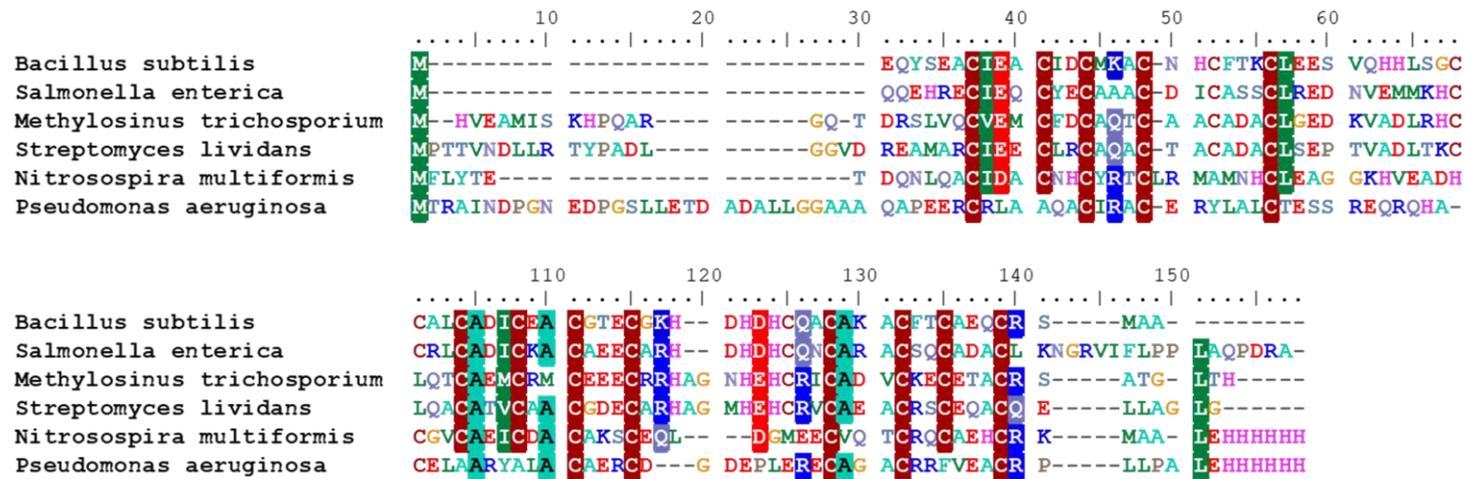
Figure 2.1 - *Multiple sequence alignment of Csp3 homologues made using MAFFT (15). Conserved amino acid residues highlighted using 80 % stringency. Cys residues are highlighted in dark red.*

Further analysis into the Ccsp genomic environment revealed two upstream genes predicted to encode a Na^+/H^+ antiporter (SLI_RS17245) and a protein that belongs to the domain of unknown function DUF4396 superfamily composed of 181 amino acids (SLI_RS17250) (Fig. 2.2 inset). A precorrin-8x methyl mutase is predicted to be encoded in the adjacent downstream gene (*SLI_RS17260*) (Fig. 2.2 inset). Further examination of SLI_RS17250 (DUF4396) protein sequence predicts a His-rich N-terminal sequence (13 His in total) and at least four transmembrane helices with the beginning of the first helix having a CXXXC motif. Additionally, no close homologues of SLI_RS17250 with a known structure were found in the BLAST search against the PDB and thus SLI_RS17250 could represent a distinct family of transporter. Intriguingly, some remote homology can be established

Figure 2.2 – Genomic environment of the Ccsp (inset). RNA-seq data represented as fold change of S. lividans gene expression after a Cu pulse (18). The genes include Ccsp, SLI_RS17245 (Na^+/H^+), SLI_RS17250 (DUF4396), CsoR efflux and 17260 (SLI_RS17260).

with the substrate binding S-subunits of the energy coupling factor family of micronutrient transporters, and the SLC11/NRAMP family, which are involved in transition metal ion transport. (137) (138, 139).

From a previous RNA-seq study with *S. lividans* 1326 grown under Cu stress (18), it was reported that SLI_RS17245 and SLI_RS17250 become upregulated 6-fold following a 30 min Cu pulse (400 μM) (18, 19). At the time of the initial RNA-seq report on global Cu regulation in *S. lividans*, the *Ccsp* gene had not been identified and re-analysis of the RNA-seq data (18) revealed that the Ccsp transcript is present at low-levels under homeostasis conditions but following the Cu pulse is up-regulated some 5-fold (Fig. 2.2). Moreover, the precorrin-8x methyl mutase gene, *SLI_RS17260*, is unaffected following the Cu pulse (Fig. 2.2) (18). The fold change of the transcript levels for Ccsp/DUF/Na^+/H^+ is on a similar magnitude as the CopZ/ATPase transcripts under regulatory control of the copper sensitive operon repressor protein (CsoR) (Fig. 2.2). Furthermore, RNA-seq data with a Δ*csoR* strain indicate that the Ccsp/DUF/Na^+/H^+ is not under the transcriptional control of CsoR (18). Thus, the *Ccsp* gene and the *SLI_RS17245* and *SLI_RS17250* genes are concomitantly sensitive to elevated

Cu levels and in a sense decoupled from CsoR control. This could hint at the possibility of a novel Cu export system whereby Ccsp could act as a donor to SLI_RS17250, which through the coupled action with a Na$^+$/H$^+$ antiporter (SLI_RS17245), moves Cu out of the cytosol.

2.3.2 *Taxonomic distribution of the putative Ccsp Cu transport system*

BLAST searches revealed that SLI_RS17245 was the most abundant across the Tree of Life with a higher number of BLAST hits compared to Ccsp and SLI_RS17250 (SLI_RS17245 giving 1,174 BLAST hits versus 621 and 277 BLAST hits for Ccsp and SLI_RS17250, respectively). However, Ccsp and SLI_RS17245 showed a considerable number of BLAST hits in Bacteria groups Terrabacteria, FCB and Proteobacteria, which for all three groups, gave over 100 hits for both Ccsp and SLI_RS17245. Overall, less than ten BLAST hits were found in most of the other bacterial groups. This may be due to sampling bias in NCBI for certain bacterial groups or may represent a genuine lower profusion of those protein sequences in some bacterial groups. The overall taxonomic distribution of the three proteins is displayed in Table 2.1.

The evolutionary lineage of the three protein sequences was next investigated (Ccsp, SLI_RS17245 and SLI_RS17250). From the taxonomic distribution shown in Table 2.1, it was found that 13 Bacterial groups contain at least one element of this putative Cu export system. There were 5 groups within Bacteria that possess all three components but only few species possessed all three: Acidobacteria, FCB group, Proteobacteria, Nitrospirae and Terrabacteria. Due to our interest in *Streptomyces* a deeper search for the three genes was performed in the Terrabacteria group. All the groups within Terrabacteria have at least one representative of the putative Cu resistance system, with the exception of Tenericutes. The four terrabacterian groups showing the highest number of BLAST hits are also the ones in which the three members of the system are present: Actinobacteria, Chloroflexi, Firmicutes, and Deinococcus-Thermus (Table 2.1). Regarding the three major groups of Archaea, Ccsp and SLI_RS17250 were not found in the DPANN group but were present (with a low number of hits) in the TACK group and the halophilic Euryarcheota (Table 2.1). Like the situation found in Bacteria many hits were found for SLI_RS17245 in the three archaean groups compared to Ccsp and SLI_RS17250. Finally, among eukaryotes only the fungi (Opisthokonta) show representatives for the three genes, with Ccsp also found in land plants (Table 2.1).

Table 2.1 - *Taxonomic distribution of the Ccsp, SLI_RS17245 and SLI_RS17250 in the Tree of Life*

	Ccsp	SLI_RS17250	SLI_RS17245
BACTERIA			
Acidobacteria	■	■	
Aquificae			■
Caldiserica			■
Chrysiogenetes			■
Deferribacteres			■
Dictyoglomi			
Elusimicrobia			
FCB group	■	■	
Fusobacteria			
Nitrospinae/Tectomicrobia group			■
Nitrospirae	■	■	■
Proteobacteria (purple bacteria)	■	■	■
PVC group	■		■
Rhodothermaeota			
Spirochaetes	■		■
Synergistetes			■
Terrabacteria group	■	■	
Thermodesulfobacteria			■
Thermotogae			
Terrabacteria group			
Actinobacteria	■	■	■
Armatimonadetes	■		■
Chloroflexi (green non-sulfur bacteria)	■	■	■
Cyanobacteria/Melainabacteria group	■		■
Deinococcus-Thermus	■	■	■
Firmicutes (Gram-positive bacteria)	■	■	■
Tenericutes			
ARCHAEA			
DPANN group			■

Euryarchaeota

TACK group

EUKARYOTA

Alveolata (alveolates)

Amoebozoa

Apusozoa

Breviatea

Centroheliozoa (centrohelids)

Cryptophyta (cryptomonads)

Euglenozoa

Fornicata

Glaucocystophyceae
(glaucocystophytes)

Haptophyceae (coccolithophorids)

Heterolobosea

Jakobida

Katablepharidophyta

Malawimonadidae

Opisthokonta

Oxymonadida (oxymonads)

Parabasalia (parabasalids)

Rhizaria

Rhodophyta (red algae)

Stramenopiles (heterokonts)

Viridiplantae (green plants)

Construction of phylogenetic trees from the three protein sequences (Ccsp, SLI_RS17245 and SLI_RS17250) reveal mainly bacterial branches with a low number of archaea and eukaryote twigs interspersed (Fig. 2.3, 2.4 & 2.5). Most nodes display high statistical supports; however, the evolutionary trees do not recover the monophyly of many bacterial groups due to the high degree of conservation of the three sequences. Their broad taxonomic distribution across the major Bacteria groups supports a hypothesis that these three sequences originated in the bacterial last common ancestor (LCA), followed by multiple gene losses in independent bacterial lineages. In contrast, the scarce presence of the three sequences in
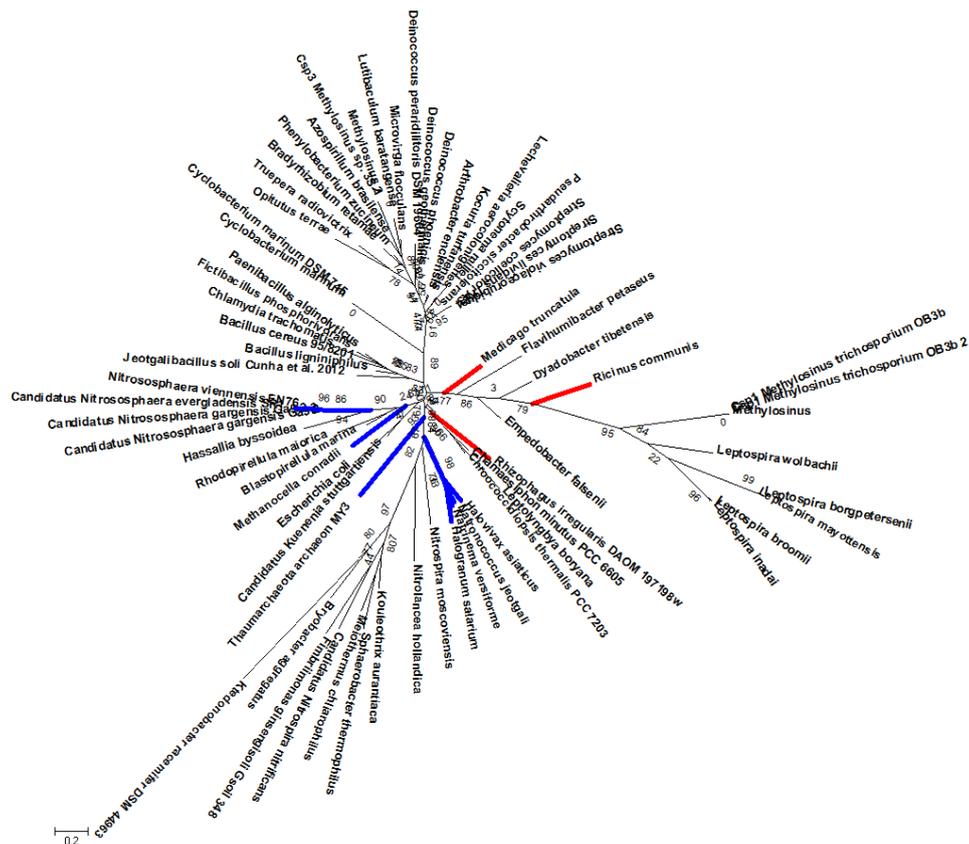
Figure 2.3 *– Spiral phylogenetic tree of all potential homologues of Ccsp. Maximum Likelihood method was used to construct the phylogenetic trees by using the program FastTree2 (21). Final construction of the spiral trees was carried out using MEGA7 (7). The branches highlighted in blue colour represent all species that belong to Archaea and all branches highlighted in red colour are species that are part of Eukaryota group*

different members of Archaea and Eukaryota, together with the lack of evolutionary relationships between their sequences, points to the absence of these sequences in the LCA of each of those domains and a most likely origin is multiple lateral gene transfer (LGT). Using the taxonomic occupancy together with the phylogenetic trees a reconstruction of the origin and evolutionary history of the three protein sequences can be created (Fig. 2.6). This illustrates that the bacterial Ccsp gene jumped twice to Euryarcheota, twice to the TACK, and

twice to eukaryotes (fungi and plants). SLI_RS17245 underwent at least five independent LGT towards Archaea, and three towards Eukaryota (Fig. 2.6). The SLI_RS17250 gene jumps the least, with two transfers to different Archaea, and two others to the eukaryotic fungi (Fig.
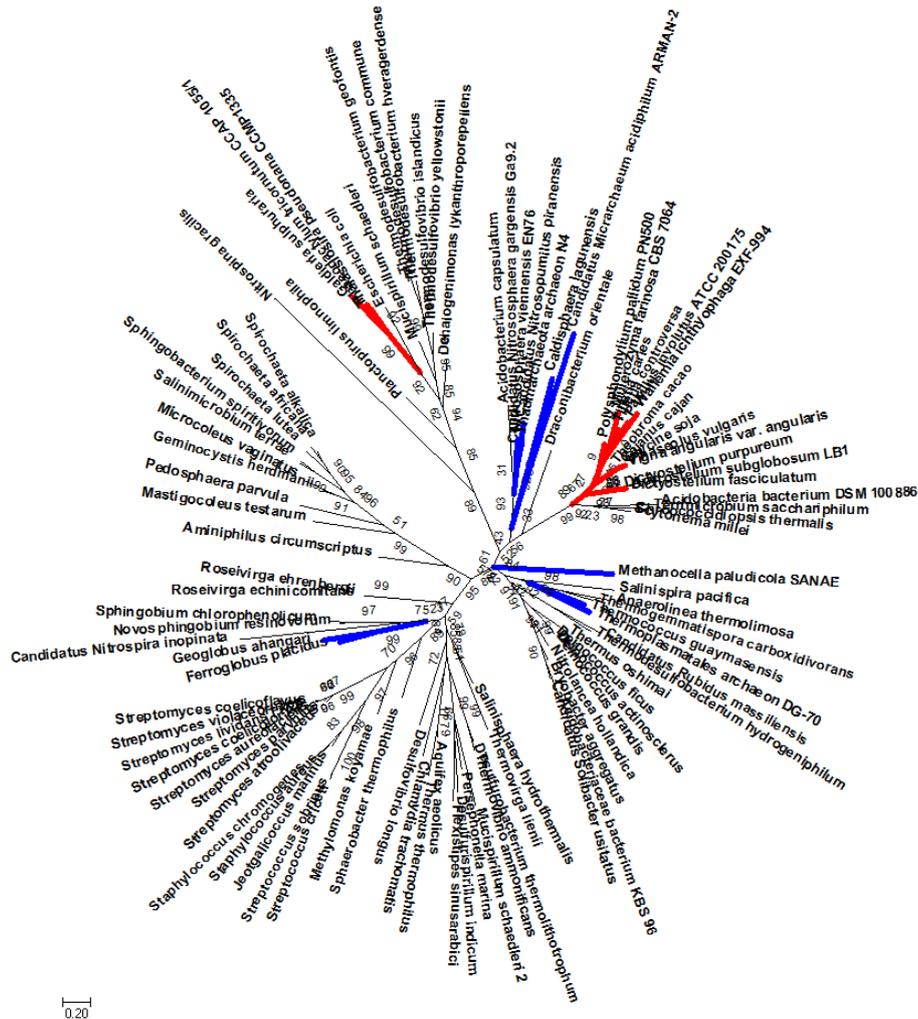
Figure 2.4 - *Spiral phylogenetic tree of all potential homologues of SLI_RS17245.*

2.6). It is important to note that the presence of the three genes in eukaryotes will be analysed further in the Discussion. It is not possible to infer from this data if there is coincidence in the timing of the transfers for the three genes. However, it is striking that the genome of three species of TACK archaeans (related to the genus *Nitrososphaera*, an ammonia oxidizing archaean) hold the three genes, suggesting they may have been transferred together in block. It must be noted that for all three genes, the data collected for the Eukaryota domain has been recently questioned as to whether these targets are in fact been mistaken for bacterial species (75). This will be explored further in the Discussion, but this issue has been considered for all data collected and the eukaryotic species can be identified in all phylogenetic trees (Figs. 2.3-2.5 and Appendix 1.1).
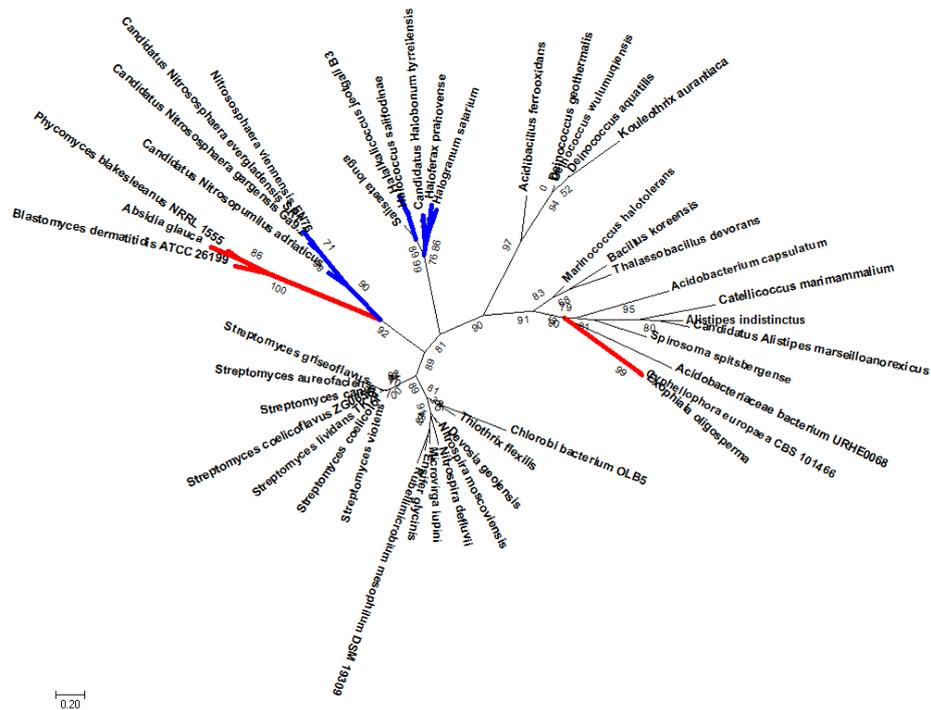
Figure 2.5 - *Spiral phylogenetic tree of all potential homologues of SLI_RS17250.*
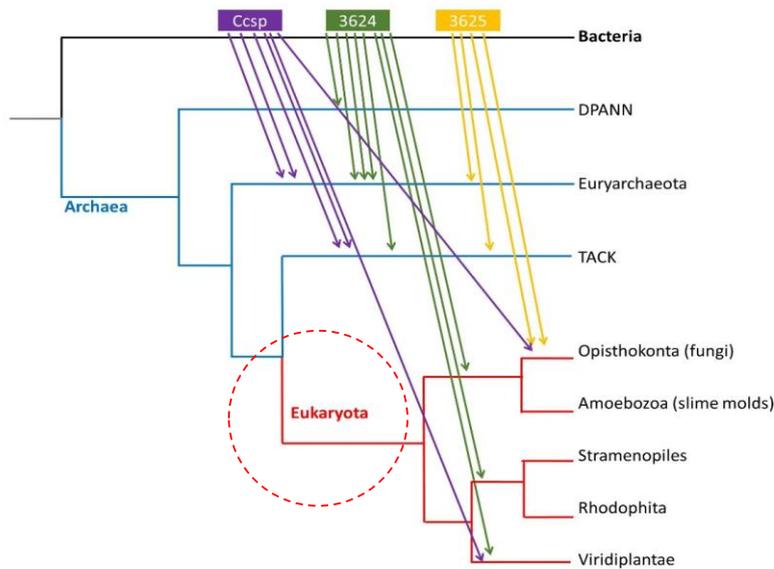
Figure 2.6 - *Evolution of the Ccsp, SLI_RS17245 (3624) and SLI_RS17250 (3625) in the Tree of Life. The occupancy and phylogenetic patterns point to multiple transfer events (indicated with arrows) from Bacteria to the other domains of the Tree of Life. Eukaryota has been circled in red due to issues in validating sample taxonomy (see Discussion).*

2.4   Discussion

2.4.1   Origin and Evolution of the copper storage system

It has been possible to determine the evolutionary history and origins of Ccsp, SLI_RS17245 and SLI_RS17250 by constructing the phylogenetic trees and their taxonomic occupation presented in this chapter. It has been hypothesised that all three genes derive from a bacterial last common ancestor (LCA) due to their wide distribution among the main Bacteria groups and were subjected to several gene losses in independent bacterial lineages. Although, it cannot be disregarded that possible LGT events occurred as there is a shortfall in resolution in parts of the phylogenetic trees. On the other hand, Ccsp, SLI_RS17245 and SLI_RS17250 are lacking in various members of Archaea and Eukaryota, along with inadequate evolutionary relationships between their sequences. Lest a great number of secondary gene losses occurred to explain their scarcity, this overall suggests an absence of these genes in the LCA of each of those domains. The likely explanation for this is due to multiple LGT events but with more genetic sampling this theory may change.

Figure 2.6 has helped by illustrating possible transfer patterns across all three phylogenetic trees as high levels of conservation in these genes have yielded some surprising phylogenetic patterns. Figure 2.6 initially shows the bacterial Ccsp gene jumped twice to Euryarcheota, twice to the TACK, and twice to eukaryotes (fungi and plants). As for SLI_RS17245, it is possible that at least five independent LGT events occurred towards Archaea and three towards Eukaryota. The SLI_RS17250 gene displayed the least amount of movement showing only two transfer events to various Archaea and two others to eukaryotic fungi. From this data alone, it is not possible to suggest any coincidence in the timing of the multiple transfers of the three genes. Interestingly, it was found that three species of TACK archaeans (related to the genus *Nitrososphaera*, an ammonia oxidizing archaean) possess all three genes in their genomes and suggests that, Ccsp, SLI_RS17245 and SLI_RS17250 have been transferred together simultaneously. For this to be validated, more information into these three genes structure and function within *Nitrososphaera* must be obtained to support this theory. Also, it has been observed that all three genes have jumped multiple times towards multicellular groups such as plants, slime molds, and fungi, as well as some unicellular algae.

Despite fascinating observations made from these phylogenetic analyses especially regarding transfer events of all three genes occurring in Eukaryota, it is essential to note that these eukaryotic targets could have been mistaken for bacterial species. It has been mentioned for Ccsp in particular, that due to a high sequence similarity between eukaryotic homologues and Ccsp raises this suspicion as stated by Dennison *et al*. in their review of Csps (75). As suggested by Dennison *et al*. this error is not surprising as many bacterial species are

soil dwelling or widespread in nature (75). Thus, contamination of samples with bacterial DNA is highly likely and gives this probable error in the phylogenetic analyses (75). The same issue could also be applied for SLI_RS17245 and SLI_RS17250 eukaryotic targets. The Eukaryota branch in Figure 2.6 has also been highlighted for this purpose, see also Appendix 1.1.

It was widely considered that the bacterial cytosol did not possess the machinery to store copper due to the absence of metabolic requirement and the toxicity of copper. This is unlike mammals, which possess various metallothioneins which are able to bind cadmium(II), copper(I) , zinc(II) or two of these metal species simultaneously (75, 140). The unexpected discovery and characterisation of cytosolic copper storage proteins known as Csp3s overturned the previous ideas about cytosolic copper in bacteria. This has now led to the discovery of Ccsp in *S. lividans* and has revealed its extensive taxonomic distribution across the Tree of Life (Figs. 2.3-2.5). Additionally, the phylogenetic distribution of SLI_RS17245 (Na^+/H^+ antiporter) and SLI_RS17250 (DUF4396) could suggest the existence of a copper cytosolic storage and regulation system that have not yet been characterised in other bacteria. It is well established that *S. lividans* requires copper as part of its development (19, 118). Stemming from this, previous studies in the extensive characterisation of *S. lividans* copper efflux and trafficking system, CsoR/CopZ/P_1-type ATPase, have yielded some interesting transcriptional responses to copper (17-19, 138, 139). Many genes other than CsoR/CopZ/ATPase efflux system responded to copper stress by becoming up- or down regulated (18). As mentioned previously in this chapter, these genes included the upregulation of *Ccsp*, *SLI_RS17245* and *SLI_RS17250*, not under control of CsoR (Fig. 2.2) (18, 19) and thus appear to be under the control of an unknown regulator. As suggested in the study by Dwarakanath *et al*., it is possible that multiple copper homeostatic mechanisms are simultaneously involved in regulating copper other than the CsoR regulon in *S. lividans* such as that described in redox homeostasis (18). Overall, these results present a potential new model for cytosolic copper storage and transport in *S. lividans* that requires further testing.

Chapter Three

Characterisation of a cytosolic copper storage protein from *Streptomyces lividans*

3.1   Introduction

Streptomyces lividans shows a distinct dependence on copper (Cu) for initiating a morphological switch from vegetative to aerial growth that simultaneously produces secondary metabolites (12, 44, 103, 141, 142). From a biotechnology perspective there is interest in using *S. lividans* as an industrial cell factory for the heterologous production of high value proteins and enzymes for processing biomass waste, diagnostic, therapeutic and agricultural uses (143). Indeed, the *Streptomyces* genus contains some important antibiotics and industrially useful enzymes that would prove beneficial if exploited correctly. The bioavailability of metal ions in microbial growth cultures is known to be important for optimizing batch processes, and to a certain extent this has been shown for Cu bioavailability in *S. lividans* and as such can impact on growth morphology in submerged (liquid) cultures (103). Thus a thorough understanding of how Cu is utilized in the host, i.e. in 'correctly' metalating secreted nascent apo-enzymes or proteins (37, 144, 145) in Cu resistance mechanisms (146) and in Cu trafficking pathways is important for creating re-engineered strains for optimized and improved growth.

Linked to understanding Cu handling in *S. lividans* is the discovery of a cytosolic copper storage (Ccsp) protein (Fig. 3.1) that has been described in Chapter 2. This Chapter describes initial structural and biochemical characterization of Ccsp and its ability to bind Cu(I). In addition, a Δ*ccsp* null-mutant in *S. livdians* has been constructed by collaborators at Leiden University, The Netherlands, and its effect on growth and morphology investigated under increasing exogenous Cu concentrations. Cu(I) trafficking has also been investigated, and using size-exclusion chromatography evidence for a *S. lividans* Cu(I) metallochaperone, CopZ, being able to traffic Cu(I) to Ccsp is presented.

SLI_RS1725541

```
MPTTVNDLLRTYPADLGGVDREAMARCIEECLRCAQACTACADACLSEPTVADLTKCIRT
DMDCADVCTATAAVLSRHTGYDANVTRAVLQACATVCAACGDECARHAGMHEHCRVCAEA
CRSCEQACQELLAGLG
```

Figure 3.1– Amino acid sequence of SLI_RS1725541 (Ccsp) derived from S. lividans. The 18 Cys residues have been highlighted in yellow.

3.2   Materials and Methods

3.2.1   Cloning of SLI_RS1725541 (Ccsp)

The *SLI_RS1725541* gene was amplified from *S. lividans* genomic DNA and restricted into a pUC19 vector provided by Dr Erik Vijgenboom (Leiden University, The Netherlands). The primers shown in Fig. 3.2 were used to amplify the Ccsp gene from the pUC19 vector for subsequent ligation into a pET28a vector (Novagen) using the NdeI and HindIII restriction

sites to create a N-terminal (His)$_6$-tagged Ccsp construct. The PCR cycle and reagents used are reported in Tables 3.1 and 3.2, respectively.

Forward primer

NdeI-F 28-mer $_{Tm}$ ~ 63 °C (5 mM MgSO$_4$ salt)

5'-TCAACATATGCCCACCACCGTCAACGAC-3'

Reverse primer

HindIII-R 30-mer $_{Tm}$ ~ 62 °C (5 mM MgSO$_4$ salt)

5'-AGTTAAGCTTGCATGCCTGCAGGTCGACTC-3'

Figure 3.2 *- Forward and reverse primers used to amplify Ccsp gene via PCR. NdeI and HindIII sites underlined.*

Table 3.1 *– The reagents and volumes used to amplify SLI_RS1725541 (Ccsp) gene from the plasmid pUC19*.

Reagent	Concentration	Volume (µl)
SLI3625_A_pTZ19R plasmid DNA	N/A	2.5
Forward Primer	-	1.0
Reverse Primer	-	1.0
dNTP's	10 mM	2.5
10 x Buffer Pfu + MgSO$_4$	-	5.0
DMSO	100 %	2.5
Sterile, deionized water	-	35.0
Pfu DNA polymerase	-	0.5
Total	-	50.0

Table 3.2 *- PCR cycles – bold text indicates these steps were repeated 35 times and other steps only once.*

Temperature (^{O}C)	Time (minutes)
95	3.0
95	**1.0**
62 (Annealing temperature)	**1.0**
72	**2.0**
72	7.5 (Final Extension)
4	Finish

The amplified product was electrophoresed on an agarose gel, extracted and gel purified. The NdeI-HindIII gene insert was ligated using T4 DNA Ligase (Thermo Scientific) into a pET28a plasmid, the vector was used to transform *Escherichia coli* XLI-Blue competent cells and transformants selected for DNA sequencing following a mini-prep (Thermo Scientific).

3.2.2 *Over-expression of Ccsp*

The Ccsp_pET28a construct was used to transform chemically competent *E. coli* BL21 (DE3) cells for overexpression of the Ccsp protein. Overnight cultures were prepared by selecting individual colonies from the transformation plate and inoculating 3 ml of 2xYT medium (Melford) containing 50 mg ml^{-1} Kan. These cultures were incubated at 37 $^{\circ}$C with shaking at 225 rpm for 16 h and then used to inoculate 750 ml of 2xYT media in 2 L flasks. Shaking continued at 37 $^{\circ}$C and 220 rpm until an optical density at 600 nm (OD$_{600}$) of 0.6 was reached followed by induction with IPTG (Melford) to a final concentration of 1 mM. At this point the temperature was reduced to 25 $^{\circ}$C and the flasks were grown for a further 16 h. The cells were harvested by centrifugation at 3,501 g (F8-6x1000y rotor) using a Sorvall Evolution RC Superspeed centrifuge for 20 min at 4 $^{\circ}$C. The resulting pellets were re-suspended in Buffer A (50 mM Tris/HCl pH 7.5, 500 mM NaCl, 20 mM imidazole) followed by addition of 1 µl of 1 M MgCl$_2$ to every 1 ml of cell suspension with stirring at 4 $^{\circ}$C for 35 min. The cell suspension was passed through an EmulsiFlex-C5 cell disrupter (Avestin), equilibrated with Buffer A, to lyse the cells. The cell lysis was then centrifuged at 38,724 g (SS-34 rotor) using a Sorvall RC-5 centrifuge for 25 min at 4 $^{\circ}$C. The cell pellet was discarded, and the supernatant was applied to a 5 ml His-trap FF Ni-NTA column (GE Healthcare) to bind the (His)$_6$-tagged Ccsp protein and then attached to an AKTA-Purifier and washed with Buffer A. A linear imidazole gradient generated using Buffer B (Buffer A with 500 mM Imidazole) to elute the bound Ccsp. The fractions from the Ni-NTA column were pooled and dialysed overnight at 4 $^{\circ}$C in Buffer C (50 mM Tris/HCl pH 7, 100 mM NaCl). The dialysate was clarified by centrifugation at 38,724 g (SS-34 rotor) using a Sorvall RC-5 centrifuge for 10 min and then concentrated at 4 $^{\circ}$C, using centricon (Vivaspin) with 30 kDa cut-off. The N-terminal (His)$_6$-tag was removed by incubating the protein at room temperature overnight with 125 KU of thrombin (Sigma) and then re-applied a Ni-NTA column in the same manner as previously described. The flow-through was collected and concentrated to 2 ml in a centricon (30 kDa cut-off) and injected onto a 120 ml Sephadex G75 column (GE Healthcare), equilibrated in Buffer C. Selected fractions from the major elution peak monitored at 280 nm were examined for purity on 15 % SDS-PAGE (Appendix 5). Samples deemed of good purity were pooled, concentrated, flash frozen in liquid nitrogen and stored at -80 $^{\circ}$C for further use.

3.2.3   *Over-expression and purification of CopZ-3079*
The overexpression of CopZ-3079 was carried out in chemically competent *E.coli* BL21 (DE3) cells and purified using the method described by Chaplin *et al.* (17).

3.2.4   *Preparation of Cu(I) and Ag(I) solutions and titration of apo-Ccsp*

Solid CuCl (Sigma) was dissolved in 10 mM HCl and 500 mM NaCl and diluted with 10 mM MOPS pH 7.5, 150 mM NaCl in an anaerobic chamber (Don Whitley Scientific [O_2] < 2 ppm). The Cu(I) concentration was determined spectrophotometrically using a Cary 60 UV-visible spectrophotometer (Varian) at 20 ^{o}C through step-wise addition of the stock CuCl solution into a known concentration of the Cu(I) specific bidentate chelator bicinchoninic acid (BCA; Sigma). Formation of the [Cu(BCA)$_2$]$^{3-}$ complex was monitored at 562 nm and the concentration determined using an extinction coefficient (ε) of 7,900 M^{-1} cm^{-1} (147). A stock solution of AgNO$_3$ was diluted to 1 mM using MOPS buffer and then used for titration to apo-Ccsp. An ε at 280 nm of 4,105 M^{-1} cm^{-1} was used to determine Ccsp concentration spectroscopically as predicted by Protparam (https://web.expasy.org/protparam). Apo-Ccsp (4-10 μM) samples for titration with either Cu(I) or Ag(I) were sealed in an anaerobic quartz cuvette (Hellma) and absorbance changes, following addition of the monovalent metal ions monitored between 350 and 200 nm.

3.2.5   *Preparation of Cu(I)-loaded Ccsp*

Cu(I)-bound Ccsp was prepared anaerobically by diluting apo-Ccsp to a concentration of ~75 μM and incubating 25 equivalents of a stock CuCl solution. The Ccsp-CuCl mix was left to equilibrate for > 15 min in an anaerobic chamber and then applied to a PD10 column equilibrated in MOPS buffer to remove unbound Cu(I). Cu(I)-bound Ccsp samples were taken out of the anaerobic chamber and concentrated to the desired concentration in a centricon (30 kDa cut-off) at 4 ^{o}C. Cu(I)-Ccsp samples prepared in this way were used for analytical gel-filtration (10/300 GL G75 Superdex column (GE Healthcare)) equilibrated with 10 mM MOPS pH 7.5, 150 mM NaCl, and far UV-CD spectroscopy carried out on Applied Photophysics Chirascan CD spectrophotometer (Leatherhead, UK). The ellipticity (mdeg) was converted to mean residue ellipticity (deg.cm^{2}dmol.res^{-1}) using equation 3.1,

$$MRE = \frac{E}{10(Pc)l} \qquad 3.1$$

where *MRE* is mean residue ellipticity, *E* is the ellipticity in mdeg, *P* is the number of peptide bonds (number of residues -1), *c* is the molar concentration and *l* is the pathlength in cm.

3.2.6   *Determination of apparent Cu(I) binding constants*

Samples of apo-Ccsp (4-10 μM) in 10 mM MOPS pH 7.5, 150 mM NaCl were incubated under anaerobic conditions with various concentrations of BCA (50-1000 μM) and increasing amounts of Cu(I) were added to each sample. Each Ccsp-BCA sample was prepared as individual solutions in Eppendorf tubes whereby each tube contained a constant [Apo Ccsp] and [L] (ligand – BCA). After the addition of increasing Cu(I) concentrations, each BCA series

were left to equilibrate for 2 to 16 h and spectrophotometric measurements were taken to quantify the $[CuL_2]^{3-}$ BCA complex as function of Cu(I):Ccsp ratio.

3.2.7 X-ray Crystallography

Crystals of Ccsp suitable for X-ray diffraction were grown using the hanging drop vapor diffusion method at 18 °C following discovery of initial crystal hits in commercial screens using 96 well plates dispensed using an ARI Gryphon crystallization robot. For optimisation of the initial crystal hits, apo- and Cu(I)-loaded Ccsp, 1 μl of protein solution at a concentration of 15 mg/ml were mixed with an equal volume of reservoir solution which for apo-Ccsp contained 1.4 M ammonium sulphate, 0.1 M HEPES pH 7.0 and for Cu(I)-loaded contained 1.4 M ammonium sulphate, 0.1 M MES pH 6.0. Crystals of apo and Cu(I)-loaded were transferred to a cryoprotectant solution consisting of 40 % w/v sucrose, and flash cooled by plunging into liquid nitrogen. Apo-Ccsp crystals were measured at the Diamond Light Source on beamline I02 using an X-ray wavelength of 0.979 Å and a Pilatus 6M-F detector. Cu(I)-loaded crystals were measured at the ESRF on beamline ID29 using a Pilatus 6M detector and an X-ray wavelength of 0.976 Å. All data were indexed using XDS (148) and scaled and merged using Aimless (149) in the CCP4 suite with the CCP4i2 interface. The apo-Ccsp structure was solved by molecular replacement in PHASER (150) using the PDB-ID 3lmf as the search model. Automated model building was carried out using the Buccaneer pipeline (151) followed by cycles of model building in Coot (152) and refinement in Refmac5 (153). Riding hydrogen atoms were added when refinement of the protein atoms had converged. The final model of apo-Ccsp was used as the search model for Cu(I)-Ccsp molecular replacement. The Cu(I)-Ccsp data were twinned and twin refinement against intensities was performed in Refmac5 together with TLS refinement. An anomalous map for validation of Cu atom positions was generated using PHASER (150) in the CCP4i2 interface from a separate dataset measured at a wavelength of 1.368 Å. Structures were validated using the Molprobity server (154) the JCSG Quality Control Server and tools within Coot (152). Structural superpositions were carried out using GESAMT in CCP4i2 (155). Coordinates and structure factors were deposited in the RCSB Protein Data Bank. A summary of data, refinement statistics and the quality indicators for the structures are given in Table 3.3.

3.2.8 Cu(I) transfer experiments

The CopZ-3079 protein to be used in Cu(I) transfer experiments between Ccsp was prepared in an anaerobic chamber in 10 mM MOPS pH 7.5, 150 mM NaCl and 2 mM DTT. Following over-night incubation, DTT was removed by passing twice down a PD10 column equilibrated in 10 mM MOPS pH 7.5, 150 mM NaCl. Cu(I)-CopZ-3079 was then prepared by addition of a

stock concentration of CuCl and excess Cu removed by passing down a PD-10 column. Apo and Cu(I)-loaded CopZ were loaded to a 10/300 GL G75 Superdex column (GE Healthcare) equilibrated with 10 mM MOPS pH 7.5, 150 mM NaCl, using a 100 µl loop and the respective elution profiles recorded. For Cu(I)-transfer experiments, stoichiometric equivalents of Cu(I) were loaded to either CopZ or Ccsp, under anaerobic conditions, before mixing samples in a molar ratio followed by incubation of up to 3 h and then loading to the mixture to 10/300 GL G75 Superdex column. Stock protein concentrations were in the range of 1-3.5 mM.

3.2.9   Generation of the Δccsp mutant of S. lividans

The Ccsp mutant (Δ*ccsp*) was constructed according to the protocol described by Blundell *et al*. (44) from the parental strain *S. lividans* 1326 (John Innes Institute collection). The *ccsp* open reading frame (ORF) was replaced by a 62 nt scar of the lox recombination site including two XbaI sites. The mutant was analysed by PCR with genomic DNA as template to confirm the loss of the *ccsp* gene. For complementation of the Δ*ccsp* mutant the *ccsp* ORF with 150 bp upstream was cloned as an EcoRI-HindIII fragment in the moderate copy number plasmid pHJL401 (156) and designated pCcsp-1.

3.2.10   Growth morphology of S. lividans

Soya flour mannitol (SFM) plates were used to grow fresh spores which were extracted and diluted to the desired concentration in sterile water. These spores were spotted in 10 µl drops containing 10^3 spores and left to dry in a flow cabinet before incubation at 30 °C for 6 days. Standard petri dishes (diameter 9 cm) containing the indicated agar medium or 24 well plates with 1.8 ml agar medium per well were used for spore spotting. Cu(II) citrate (Sigma-Aldrich) was used as the Cu source and diluted to the desired concentration. 2 x 10^6 spores were used to inoculate Bennett's glucose medium and liquid R5 medium in 125 ml baffled flasks and incubated with shaking (160 rpm) for 32 h. Cu was added as Cu(II)citrate to the desired final concentration along with BCDA (bathocuproinedisulfonic acid; Sigma-Aldrich) which was added to a final concentration of 50 mM. After 32 h, 2 ml samples were collected in duplicate in pre-weighed Eppendorf tubes. The mycelium was collected by centrifugation, the pellets dried for 48 h at 98 °C and dry weight of all biomass collected was deduced using an analytical balance.

3.2.11   Cytochrome c oxidase activity

The *in vivo* CcO activity was carried out using TMPD (Sigma-Aldrich) as substrate (44, 157, 158). DNA or Bennett's glucose agar was used for spotting strains (10 ml containing 1000 spores) which was incubated at 30 °C for 24 h. The mycelium spots were fixed by using a light

spray of 0.3% (w/v) agarose in water proceeded overlaying this with 10 ml of 25 mM sodium phosphate pH 7.4 solution containing 20% ethanol, 0.6% agarose, 1% sodium deoxycholate and 10 mg TMPD. Digital images were taken every 30 seconds for 5-10 minutes to record CcO activity. The average pixel intensities of the indophenol blue stained mycelium were calculated using IMAGEJ software (159).

3.3   Results

3.3.1   Construction of an expression construct for Ccsp

Amplification of the Ccsp DNA from the pUC19 vector using the primers reported in Fig. 3.2 was successful based on the size of the visualised PCR product (~500 bp) on an agarose gel (Fig. 3.3). This band was excised, gel purified and subjected to restriction digest with the NdeI and HindIII enzymes before ligation to a pET28a vector cut with the same enzymes. Transformants following ligation were checked for the correct insert by performing a restriction digest with the enzymes NdeI and HindIII (Fig. 3.3) and DNA sequencing confirmed the correct sequence.
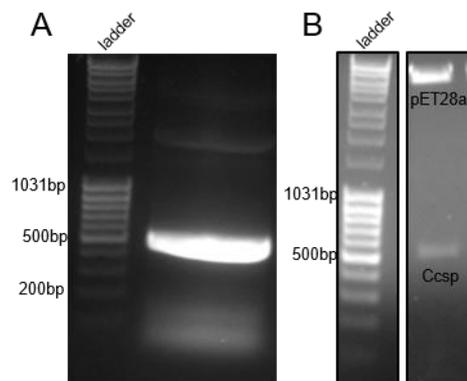
Figure 3.3 – *(A) Agarose gel showing PCR product of amplified Ccsp gene (470 bp). (B) Agarose gel confirming the Ccsp insert ligated into pET28a plasmid as a NdeI/HindIII fragment.*

3.3.2 Ccsp purification and confirmation of a homotetramer in solution

The Ccsp-pET28a construct was over-expressed in *E. coli* and following cell lysis and loading to a Ni-NTA Sepharose column eluted as a single broad peak starting at ~30 % buffer B (Fig. 3.4A). SDS-PAGE analyses of the fractions eluting from the Ni-NTA column under this peak showed a major band running at ~15 kDa (Fig. 3.4C). After thrombin digestion to remove the (His)$_6$-tag, Ccsp ran at a lower molecular weight on SDS-PAGE (Fig. 3.4D) and was further purified on a G75 size-exclusion column resulting in a major peak eluting at ~58 ml. Based on the column calibration this volume is consistent with Ccsp existing as a higher-order
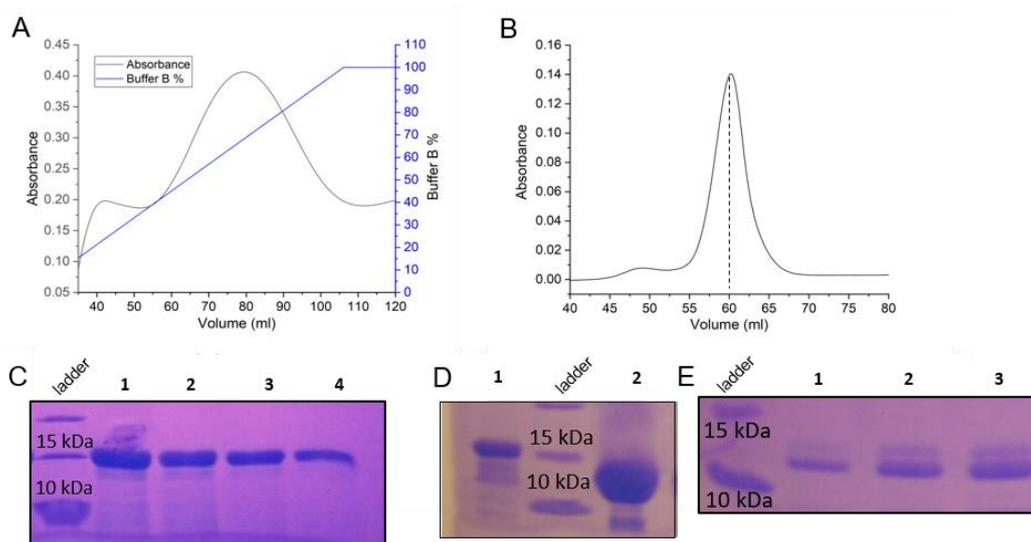
Figure 3.4 – (*A) Elution profile of Ccsp (His)$_6$-tag on a 5 ml His-trap FF Ni-NTA column, blue line represents Buffer B gradient. (B) Gel filtration elution profile of Ccsp on a 120 ml Sephadex G75 column. SDS-PAGE gel analysis. (C) Coomassie stained 15 % SDS-PAGE analysis of fractions from the Ni-NTA column showing a strong band running at ~15 kDa. (D) Cleavage of (His)$_6$-tag, lane 1 before addition of thrombin, lane 2 flow-through from Ni-NTA column. (E) G75 column fractions.*

assembly, with the elution volume suggestive of a homotetramer (Fig. 3.4B). Denaturing ESI-MS (carried out by Dr. Jason Crack at the University of East Anglia) gave a mass of 14,604.6 Da as expected for a Ccsp protomer following cleavage of the (His)$_6$-tag. Native ESI-MS gave a mass of 58,418.16 Da (14,604 Da x 4), thus corroborating the observation from the gel-filtration profile that Ccsp existed in solution as a higher-order assembly most consistent with a homotetramer. Furthermore, the native ESI-MS studies were consistent with the absence of Cu bound to the purified Ccsp. Therefore, the purification protocol used not only leads to the successful production of Ccsp for further *in vitro* studies but also delivers the apo-form of the protein.

3.3.3 Absorption and far-UV CD spectroscopy of apo-Ccsp

The UV-vis spectrum of apo-Ccsp revealed no absorption transitions in the visible region (350-800 nm). In the UV-region (200-350 nm) a distinct peak at 280 nm is observed (Fig. 3.5A),

which likely arises from the presence of two Tyr residues present in the Ccsp sequence (Fig. 3.1). The far UV-CD spectrum of apo-Ccsp is shown in Fig. 3.5B and has the characteristic minima at 222 and 208 nm that are typical signatures of α-helical secondary structure, confirming that the isolated apo-Ccsp is folded following purification.
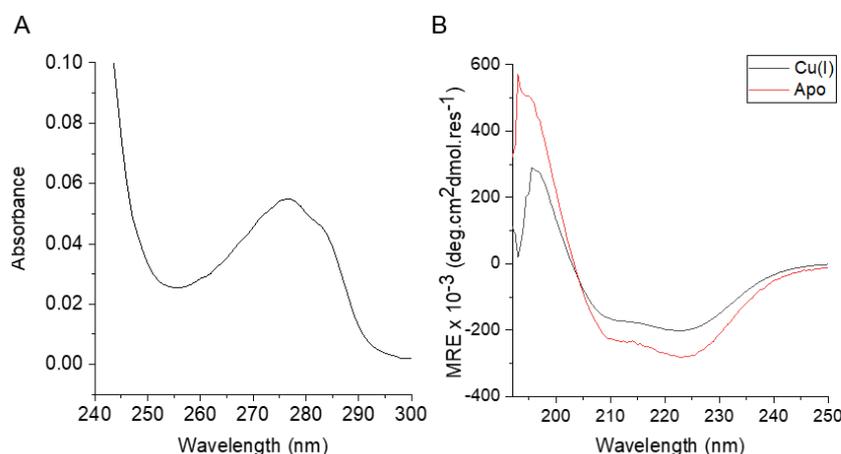
Figure 3.5 – (A) UV spectroscopy profile of apo-Ccsp. (B) Far UV-CD spectra of Apo-Ccsp (~3 μM) at 20 ^{O}C, pH 7 and of Cu(I)-Ccsp (~2 μM), at 20 ^{O}C, pH 7.5.

3.3.4   X-ray crystal structure of apo-Ccsp

Initial crystal hits for apo-Ccsp obtained from screening against commercial crystallisation screens were optimised to obtain larger crystals for X-ray diffraction experiments. The optimised conditions for apo-Ccsp produced large colourless crystals after one week (Fig. 3.6)
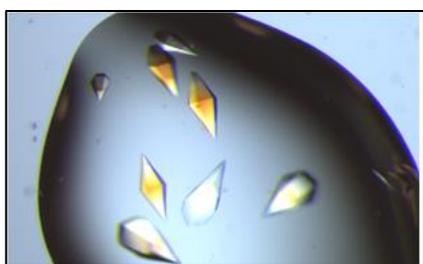
Figure 3.6 – Examples of **a**po-Ccsp crystals in 1.4 M Ammonium sulphate, 0.1 M HEPES pH 7.

and these were selected for diffraction studies at Diamond Light Source. The X-ray structure of Ccsp was determined by molecular replacement to a resolution of 1.34 Å (Table 3.3). Four Ccsp protomers (Chains A to D) were identified in the crystallographic asymmetric unit (Fig. 3.7A), with unbroken electron density observed for residues 17–136 in chain A, 20–135 in chain B, 19–135 in chain C and 16–136 in chain D. Thus, in all four protomers electron density corresponding to residues 1–15 in the sequence was not observed. Each protomer is made-up of four α-helices arranged to form a four-helix bundle motif (Fig. 3.7A). Chains A, C and D together with a symmetry related molecule create the functional homotetramer quaternary structure (Fig. 3.7B). The core of each four helix-bundle protomer creates a solvent shielded pore or channel that is lined with the 18 Cys residues

(Fig. 3.7B), none of which participate in disulfide bonds or exhibit any modifications.
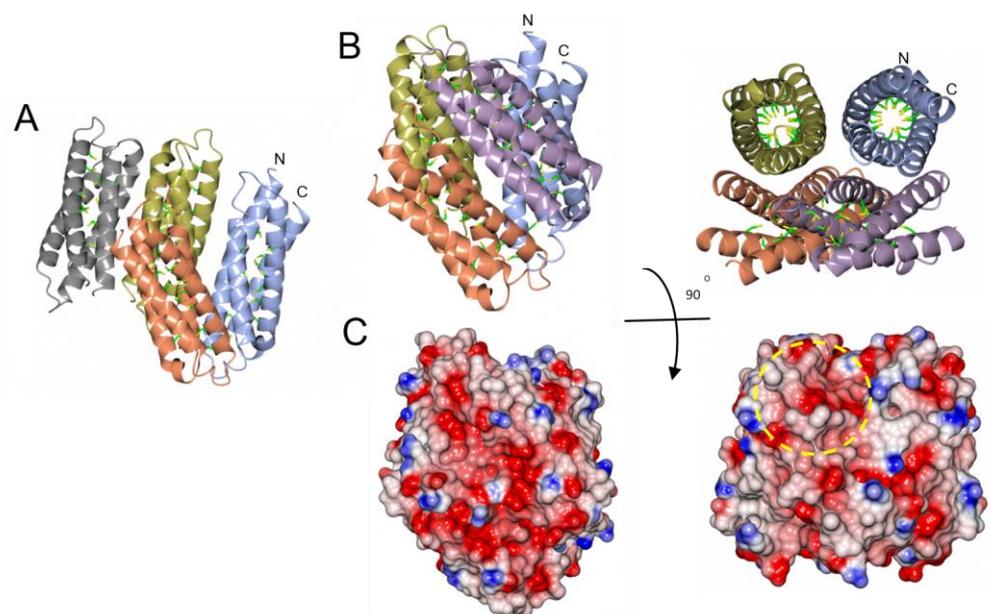
Figure 3.7 – Cartoon representation of X-ray crystallographic structure of apo-Ccsp. (A) Arrangement of Ccsp protomers in the asymmetric unit. Colour coding as follows: gold chain A, grey chain B, blue chain C and orange chain D. The N and C-termini are labelled in chain C. (B) Biological quaternary structure of the homotetramer of Ccsp. Chain B is omitted and the symmetry related Ccsp protomer required to form the biologically relevant unit is shown in lilac. (B) The 18 Cys residues are shown as green sticks with the Sγ atom coloured yellow. (C) Electrostatic surface potential of Ccsp in the same orientations as in (B) and the yellow dashed circle indicates the asymmetry in charge distribution at the opposite ends of the pore openings. Image was created in CCP4i2 using the graphics program CCP4mg.

Generation of the electrostatic surface potential of the quaternary homotetramer reveals large stretches of negative charge spanning essentially the length of the protomer interaction site (Fig. 3.7C). Notably, from the surface representation it is apparent that there an asymmetry of negative charge between the two ends of the pore opening (Fig. 3.7C) and may have consequences for Cu(I) loading.

Table 3.3 *– Crystallographic data processing and refinement statistics for apo-Ccsp. Values in parenthesis refer to the outermost resolution cell.*

Structure	Apo-Ccsp
Space group	$P6_122$
Unit cell (Å)	93.6, 93.6, 213.4
Resolution (Å)	80.3–1.34
Unique reflections	123 798
Mn (I/SD)	18.0 (0.8)
$CC_{1/2}$	0.999 (0.764)
Completeness (%)	100 (99.1)
Redundancy	18.2 (12.6)
R_{cryst}	0.157
R_{free}	0.192
RMS dev. bond lengths (Å)	0.015
RMS dev. bond angles (°)	1.63
Ramachandran favoured (%)	99.8
PDB accession code	6EI0

3.3.5   Ccsp can bind Cu(I) and Ag(I) ions

Ccsp is able to bind both Cu(I) and Ag(I) Group 11 monovalent ions. Under anaerobic conditions, Cu(I) ions were added to Ccsp which resulted in the appearance of prominent absorbance bands in the UV-region of the absorption spectrum (Fig. 3.8A). These bands signify (Cys)Sγ→Cu(I) ligand to metal charge transfer (LMCT) bands and increase concomitantly with increasing Cu(I) concentrations until reaching a saturation point indicative of a stoichiometry of ~18-20 Cu(I) ions bound per Ccsp protomer (Fig. 3.8B). This indicates that up to 80 Cu(I) ions may be bound in a homotetramer of Ccsp. Stoichiometric loading of Ccsp with Cu(I) ions to create the holo-Ccsp resulted in a far UV-CD spectrum that was not significantly different from apo-Ccsp, demonstrating that bound Cu(I) ions do not grossly alter the secondary structure (Fig. 3.5B). Addition of Ag(I) ions to Ccsp also displayed changes in the UV-region of the absorption spectrum, which most likely represent (Cys)Sγ→Ag(I) LMCT bands (Fig. 3.9A). The saturation point was less well defined compared to that of Cu(I) but seems to reveal a stoichiometry of ~15 Ag(I) ions bound per Ccsp protomer (Fig. 3.9B).
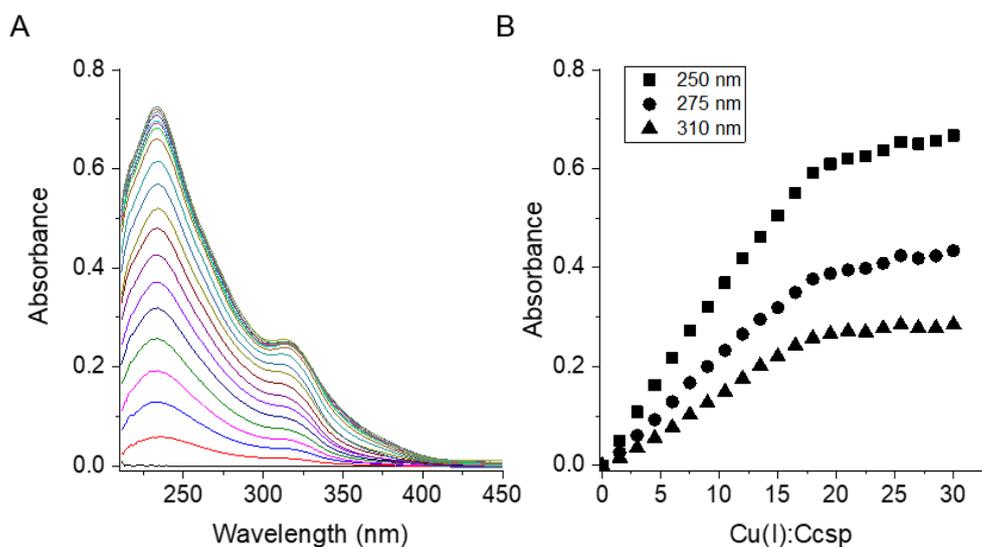
Figure 3.8 *– (A) Changes in the UV-vis difference spectra baselined with apo-Ccsp (5.6 µM) on titration with 1.5 equivalents of Cu(I). (B) Selected wavelengths plotted against Cu(I):Ccsp ratio derived from the UV-vis difference spectrum in (A). Experiments were conducted at 20 ºC, pH 7.5.*
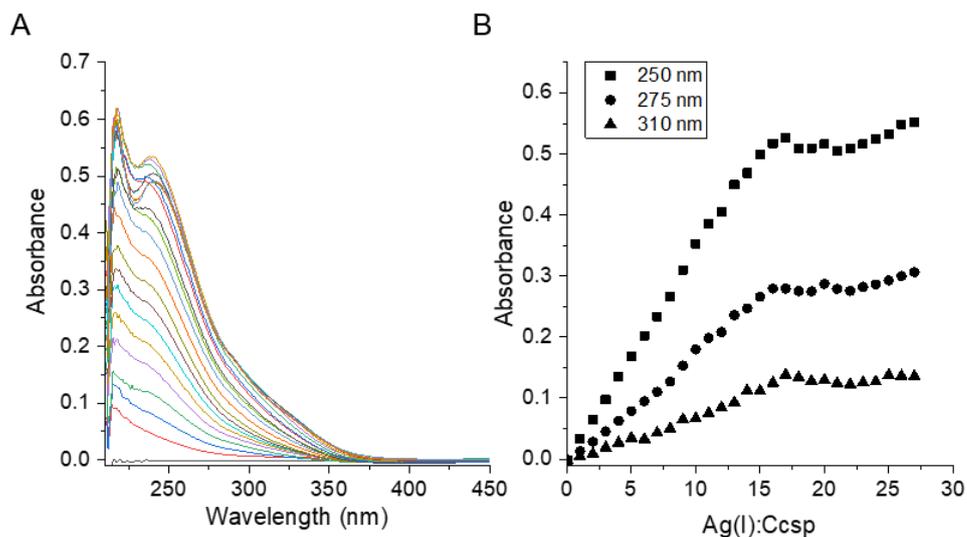
Figure 3.9 *- (A) Changes in the UV-vis difference spectra baselined with apo-Ccsp (5 µM) on titration with 1 equivalent of Ag(I) (B) Selected wavelengths plotted against Ag(I):Ccsp ratio derived from the UV-vis difference spectrum in A. Experiments were conducted at 20 ºC, pH 7.5.*
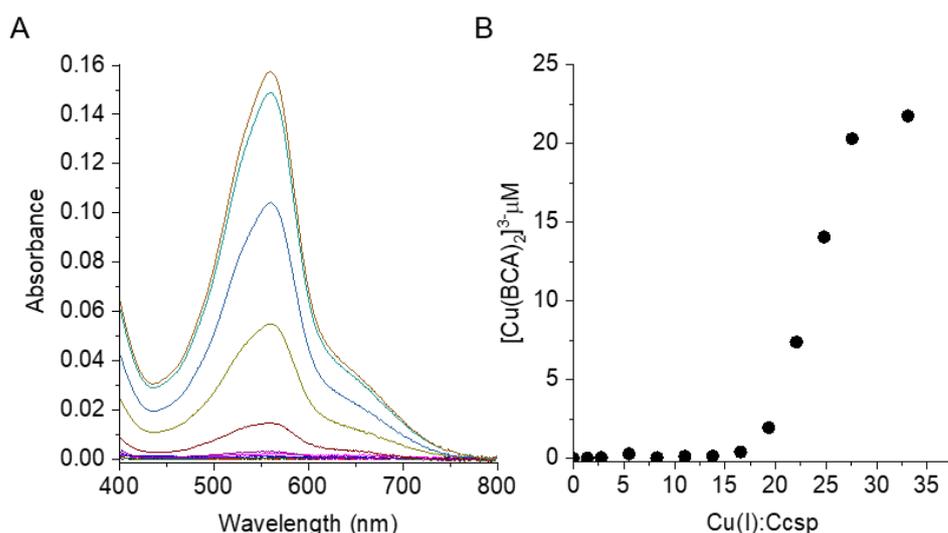
Figure 3.10 – (A) Changes in the UV-vis spectra monitoring the formation of the [Cu(BCA)$_2$]$^{3-}$ complex in the presence of BCA (50 μM) and apo-Ccsp (3.7 μM) on titration with 1.35 Cu(I) equivalents. (B) Absorbance data at wavelength 562nm in A plotted against Cu(I):Ccsp ratio. Experiments were conducted at 20 °C, pH 7.5.

3.3.6   Determination of a Cu(I) affinity of Ccsp via competitive Cu(I) titrations with BCA

An estimation of Cu(I) binding affinities for cuproteins is possible through competition experiments using high affinity chromogenic Cu(I) bidentate ligands such as BCA (13, 160). At low BCA concentrations (50-100 μM) Ccsp binds all Cu(I) ions until > 15 Cu(I) equivalents have been added and the formation of the [Cu(BCA)$_2$]$^{3-}$ complex occurs (Fig. 3.10A and B). Upon increasing the BCA concentration (250-1000 μM) competition for the titrated Cu(I) ions between Ccsp and BCA is now observed (Fig. 3.11A), with a maximum Cu(I) occupancy in a Ccsp protomer estimated to be 15 Cu(I) equivalents. All titration experiments at the different BCA concentrations were carried out in triplicate. An approximate Cu(I) binding affinity was determined at BCA concentrations of ≥ 250 μM in two ways. The below equilibria are present under the experimental conditions employed.

$$2L_f + Cu_f^+ \rightleftharpoons Cu(L)_2 \qquad K_{BCA} = \frac{[L_f]^2 \, [Cu_f^+]}{[Cu(L)_2]}$$

$$S_f + Cu_f^+ \rightleftharpoons Cu^+S \qquad K_{Cu} = \frac{[S_f] \, [Cu_f^+]}{[Cu^+S]}$$
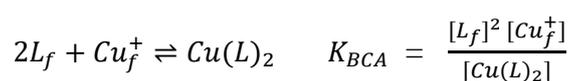
where L$_f$ = free BCA ligand and S$_f$ = sites on Ccsp that are unoccupied with Cu(I), Cu^+_f is free Cu and K_{BCA} and K_{Cu} are equilibrium dissociation constants for the affinities of Cu(I) for BCA (BCA formation constant β$_2$ = 10$^{17.7}$ M^{-2}) (161) and Ccsp, respectively. Based on the above equilibria the [Cu^+_f] is given by

$$[Cu_f^+] = \frac{K_{BCA}[Cu(L)_2]}{[L_f]^2} = \frac{K_{Cu}[Cu^+S]}{[S_f]}$$

which can be rearranged to solve for K_{Cu}

$$K_{Cu} = \frac{K_{BCA}[Cu(L)_2]\left([S_t] - [Cu_t^+] + [Cu(L)_2]\right)}{\left([L_t] - 2[Cu(L)_2]\right)^2 \left([Cu_t^+] - [Cu(L)_2]\right)} \qquad (3.2)$$

where [S_t] is the total concentration of sites occupied in Ccsp, [Cu_t^+] is the total concentration of Cu(I) added and [L_t] is the total concentration of BCA in the experiment. The derivation of equation 3.2 can be found in Appendix 2. The maximum Cu(I) occupancy for Ccsp in the presence of BCA was estimated to be 15 Cu(I) equivalents. In addition, the K_{Cu} can be determined by calculating the [Cu_f^+] using equation 3.3

$$[Cu_f^+] = \frac{[Cu(L)_2]}{[L^*]^2 \beta_2} \qquad (3.3)$$

where [L^*] = [L_t] − 2[$Cu(L)_2$] and β_2 is the affinity of BCA for Cu(I). Plots of [Cu_f^+] against the fractional Cu(I) occupancy (YCu^+) of Ccsp at a given [BCA] were best fitted to a nonlinear form of the Hill equation (3.4) to yield a K_{Cu} value and a Hill coefficient (n).

$$YCu^+ = \frac{[Cu_f^+]^n}{K_{Cu}^n + [Cu_f^+]^n} \qquad (3.4)$$

Using equation 3.2 apparent Cu(I) dissociation constants (K_{Cu}) for each titration of Cu(I) into Ccsp at BCA concentrations of 250, 500 and 1000 µM can be determined, with an average K_{Cu} = 3.3 ± 1.3 x 10^{-17} M (error given is a standard error from the replicates at each BCA concentration). Alternatively, the data in Fig. 3.11A at BCA concentrations ≥ 250 µM can be used to calculate the [Cu_{free}^+] using equation 3.3. Plots of fractional occupancy of Cu(I) sites versus the [Cu_{free}^+] at two set BCA concentrations are illustrated in Fig. 3.11B. The data clearly show a sigmoidal dependence and given that the system is at equilibrium then this implies cooperative of Cu(I) binding. Therefore, these data have been fitted accordingly using a

nonlinear form of the Hill equation (eq. 3.4) (Fig. 3.11B). From triplicate experiments with set BCA concentrations ranging between 250-1000 μM an average K_{Cu} = 2.9 ± 0.2 x 10^{-17} M and a Hill coefficient, n, = 1.9 ± 0.2 are determined. Thus, Cu(I) binding to *S. lividans* Ccsp appears to be a cooperative process with a binding affinity in line with a role in sequestering and storing cytosolic Cu(I) ions.
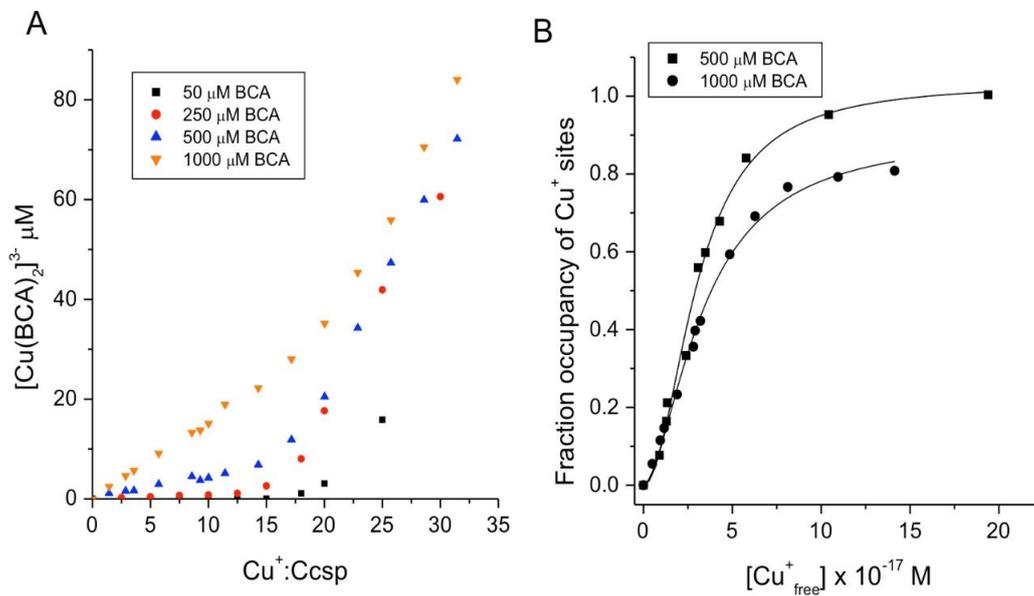
Figure 3.11 - *Ccsp Cu(I) affinity (A) Plots of [Cu(BCA)$_2$]$^{3-}$ concentration versus the Cu$^+$:Ccsp concentration ratio upon titrating Cu(I) ions into Ccsp (5-10 μM) in the presence of increasing BCA concentrations. Competitive Cu(I) binding occurs with increasing BCA concentrations. (B) Plots of fractional occupancy of Cu(I) binding sites in Ccsp at varying Cu$^+_{free}$ concentrations determined from the data in (A) at 500 and 1000 μM BCA concentrations. The data points are fitted with non-linear form of the Hill equation as shown in equation 3.4 to give a K_{Cu} = 3.0 ± 0.1 x 10^{-17} M and a Hill coefficient, n = 2.0 ± 0.2 at 500 μM BCA and a K_{Cu} = 3.4 ± 0.2 x 10^{-17} M and n = 1.7 ± 0.1 at 1000 μM BCA. All experiments were performed in 10 mM MOPS pH 7.5, 150 mM NaCl.*

3.3.7 The X-ray crystal structure of Cu(I)-loaded Ccsp

Examples of Cu(I)-loaded Ccsp crystals from optimisation screens are shown in Fig. 3.12. The structure of the Cu(I)-loaded Ccsp was determined to 1.5 Å resolution by X-ray crystallography with one Ccsp protomer found in the asymmetric unit. Statistics and refinement details are given in Table 3.4. Unbroken electron density was observed for residues 16–136 with strong anomalous scattering attributed to the presence of bound Cu(I) ions observed (Fig. 3.13A). The anomalous electron density, plotted from a dataset measured at a wavelength of 1.368 Å, reveals 20 Cu(I) ions can be coordinated in the core of the Ccsp four-helix bundle (Fig. 3.13A), thus corroborating the

Figure 3.12 – *Examples of Cu(I)-Ccsp crystals in 1.4 M ammonium sulphate, 0.1 M MES pH 6.*

estimate from the titration data (Fig. 3.8A). Therefore, the Ccsp homotetramer has the capacity to bind a total of 80 Cu(I) ions. Inspection of the Cu(I)–Ccsp structure reveals that Cu(I) ions 1 to 13 all have bis- cysteinate coordination with (Cys)–Sγ–Cu(I) bond distances of between 2.0 and 2.3 Å and each Cys residue bridging two different Cu(I) ions (Fig. 3.13B). Out of these 13 Cu(I) ions seven (Cu2, Cu4, Cu5, Cu7, Cu10, Cu12 and Cu13) are coordinated by CXXXC motifs, with the remainder coordinated by Cys residues that are on different helices of the bundle. No Cu(I) ions are coordinated by CXXC motifs. The bis-cysteinate coordination pattern is broken at Cu(I) ion 14 which has a third coordinate bond from the Oδ1 atom (2.2 Å) of Asp61 (Fig. 3.13B and C). Similarly, the Cu(I) ion 15 has a coordinate bond with the Oδ2 atom of Asp61 (2.1 Å) as well as thiolate coordination from Cys104, which also participates in coordination with the Cu(I) ions 13 and 14. (Fig. 3.13B and C). It is possible that the Cu(I) ion 15 is further coordinated by Cys41 and Cys57 (the latter being the only Cys residue not in either a CXXC or CXXXC motif), to create a distorted tetrahedral coordination geometry, however we note that the (Cys)–Sγ–Cu(I) bond distances of 2.5 and 2.7 Å (Fig. 3.13C, indicated with blue arrows), respectively, are longer than for other thiolate Cu(I) interactions. The remaining 5 Cu(I) ions, 16 to 20, cluster beyond Cu(I) ion 15 towards the entrance of the pore (Fig. 3.13B and C) and if a coordinate bond from Cys41 and Cys57 to Cu(I) ion 15 is absent it may be considered as a separate cluster. None of these remaining Cu(I) ions are coordinated in CXXXC motifs. Cu(I) ions 16 and 18 have bis-cysteinate coordination (Fig. 3.13C), with the Cu(I) ions 17 and 19 having bis-cysteinate coordination as well as ligation from the Nδ of His113 (1.9 Å) and His107 (2.1 Å), respectively. Finally, Cu(I) ion 20 is

coordinated by Cys114 and the Nδ of His111 (2.2 Å). Cys114 is the only other Cys residue to participate in coordination with three different Cu(I) ions (Fig. 3.13C). Finally, a total of nine Cu(I)–Cu(I) interactions are identified with distances between 2.5 and 2.8 Å.
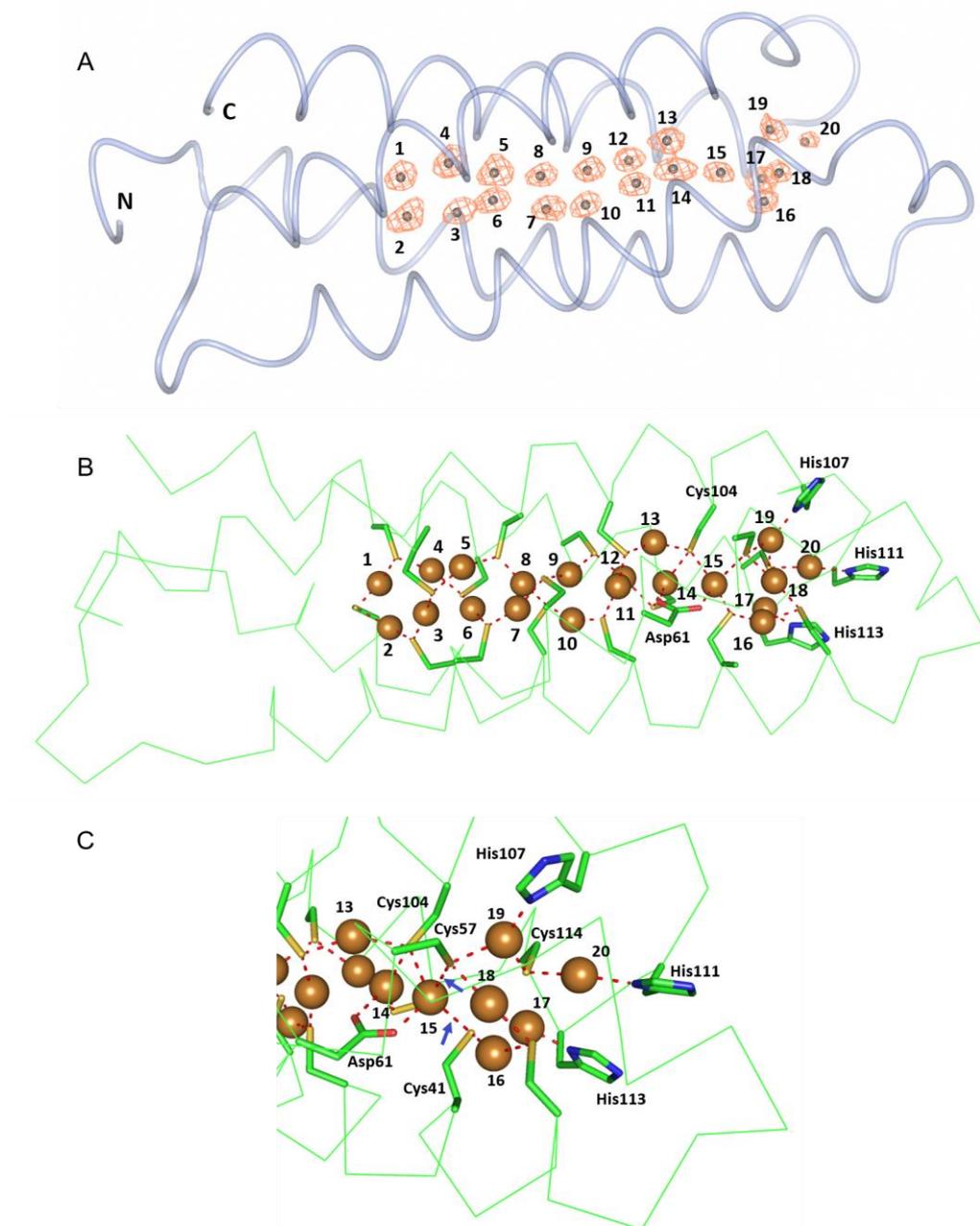
Figure 3.13 - X-ray crystal structure of Cu(I) loaded-Ccsp. (A) Worm representation of Cu(I)–Ccsp with the anomalous electron density map (orange) contoured at 5σ. Twenty Cu(I) ions have been modelled into the density and labelled 1 to 20 starting at the N and C-termini (B) Ribbon representation of a Ccsp protomer and coordination bonds (red dashed lines) to Cu(I) ions (brown spheres) labelled 1 to 20, from Sγ(Cys) Oδ(Asp) and Nδ(His) atoms. (C) Close-up of the His coordinating pore opening. The blue arrows indicate the (Cys)–Sγ–Cu(I) bond distances of 2.5 and 2.7 Å in a distorted tetrahedral coordination geometry.

Table 3.4 - *Crystallographic data processing and refinement statistics for Cu(I)-Ccsp. Values in parenthesis refer to the outermost resolution cell.*

Structure	Cu(I) - Ccsp
Space group	I222
Unit cell (Å)	62.1, 64.1, 66.0
Resolution (Å)	45.2–1.50
Unique reflections	20 701
Mn (I/SD)	9.7 (4.5)
$CC_{1/2}$	0.992 (0.958)
Completeness (%)	96.4 (96.6)
Redundancy	3.6 (3.5)
R_{cryst}	0.206
R_{free}	0.231
RMS dev. bond lengths (Å)	0.020
RMS dev. bond angles (°)	2.05
Ramachandran favoured (%)	99.2
PDB accession code	6EK9

3.3.8 *S. lividans Ccsp is required for growth under extreme Cu conditions in vivo*

To investigate the role of Ccsp in the Cu-dependent morphological development a Δ*ccsp* mutant strain of *S. lividans* was created by Dr Erik Vijgenboom (Leiden University) and *in vivo* growth and morphology studies investigated. It has been previously demonstrated that WT *S. lividans* is highly tolerant to increased Cu levels, which differentiates with various growth media (44, 141). This is illustrated in Fig. 3.14, using defined medium with either mannitol or glucose. The WT strain, grown on glucose, tolerated growth with Cu at concentrations past 5 mM, but, as shown in Fig. 3.14A, the presence of aerial mycelium becomes scarce at Cu concentrations above 1 mM. This differs for the WT strain grown on mannitol (Fig. 3.14B), with Cu tolerance much reduced (~ 5-fold). The WT shows a higher Cu tolerance overall on complex media such as R5 or Bennets-glucose media but is again media dependent (Fig. 3.15). On all media tested the Δ*ccsp* mutant consistently shows a weaker tolerance to Cu compared to the wild-type. This can be observed for glucose supplemented media (Fig. 3.14A) whereby growth is weakened at 500 µM Cu whereby the growth on mannitol (Fig. 3.14B) is inhibited at 200 µM Cu. However, when the Δ*ccsp* mutant is grown at low Cu concentrations, the growth does not differentiate from the WT. This clearly demonstrates that Ccsp is needed for growth at higher Cu concentrations but is not required for development at lower Cu concentrations. The pCcsp-1 represents the *ccsp* gene inserted into a plasmid whereby it is transcriptionally controlled by its own promoter. The Δ*ccsp* mutant becomes restored in growth to the same level as the WT strain with pCcsp-1, in all media tested (Fig. 3.14 A-B). In liquid media, a similar pattern to that observed on solid media is
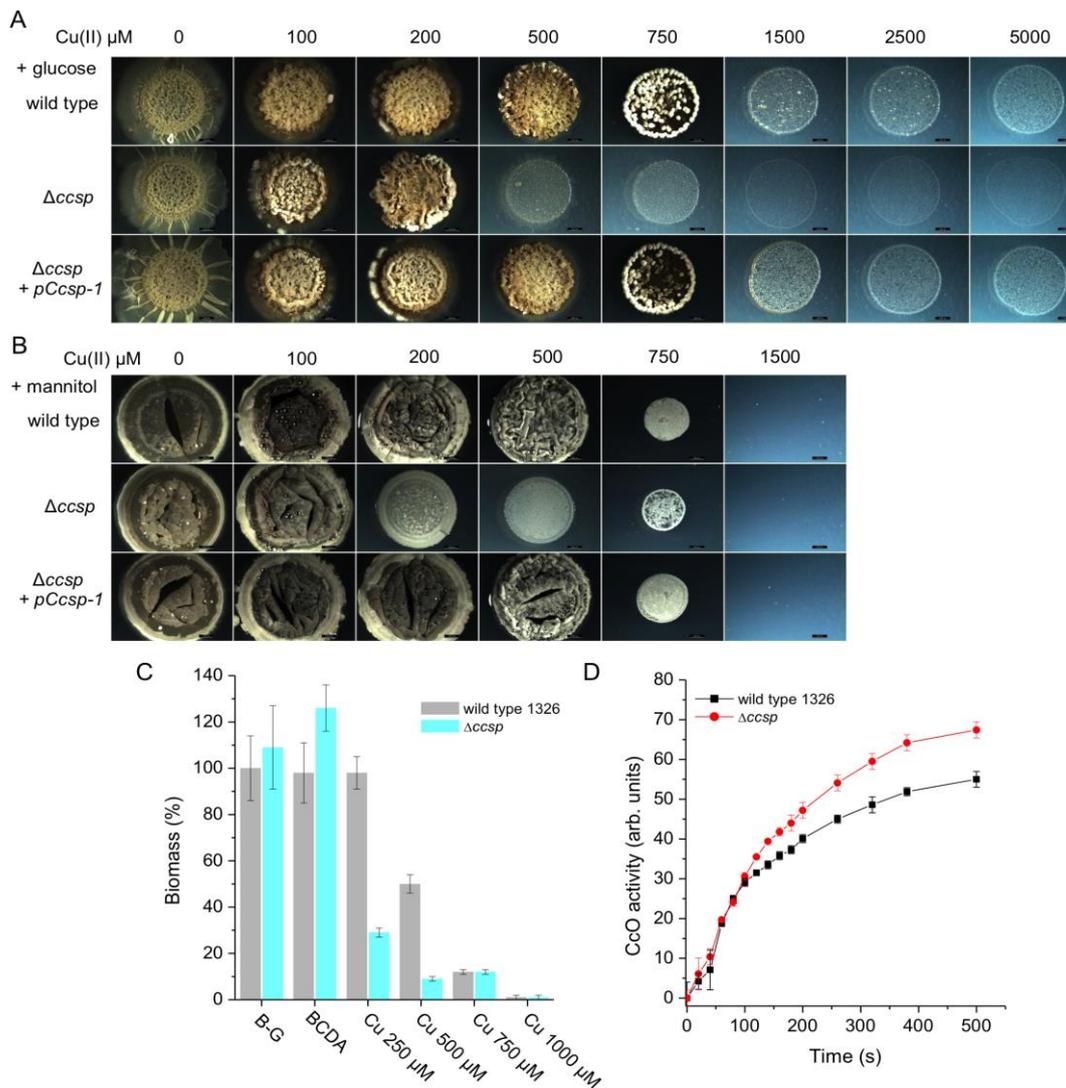
Figure 3.14 - *The effect of Ccsp on the growth and development of S. lividans at 30 °C. Cu tolerance after 6 days growth of the wild type parent strain, the Δccsp mutant strain and the Δccsp mutant strain complemented with the pCcsp-1 plasmid on defined agar media with (A) glucose and (B) mannitol as the sole carbon source. Cu(II) concentrations as indicated. All images are the same magnification with a scale bar of 2 mm. (C) Biomass production after 32 h in liquid Bennetts-glucose (B-G) cultures for the wild type and the Δccsp mutant strain in the presence of the Cu chelator BCDA and various concentrations of Cu(II) citrate. The dry weight biomass of the wild type strain in the B-G culture was set at 100%. (D) CcO oxidase activity at 24 h growth on B-G agar detected by the TMPD assay. Average pixel intensity of the indophenol blue stained mycelium was calculated using ImageJ software38 and expressed in arbitrary units. Experiments were carried out in triplicate.*

found (Fig. 3.14C) with a reduced tolerance for Cu in terms of the biomass produced in the Δ*ccsp* mutant as illustrated for Bennetts-glucose medium (Fig. 3.14C).

Under Cu homeostasis, enzymes requiring Cu for their activity such as CcO or GlxA in *S. lividans* obtain their Cu from a Cu trafficking pathway that involves at least two Cu metallochaperones, ECuC and Sco (12, 44, 101, 103). To investigate whether Ccsp influences this extracellular Cu trafficking pathway, the activity of CcO under low exogenous Cu
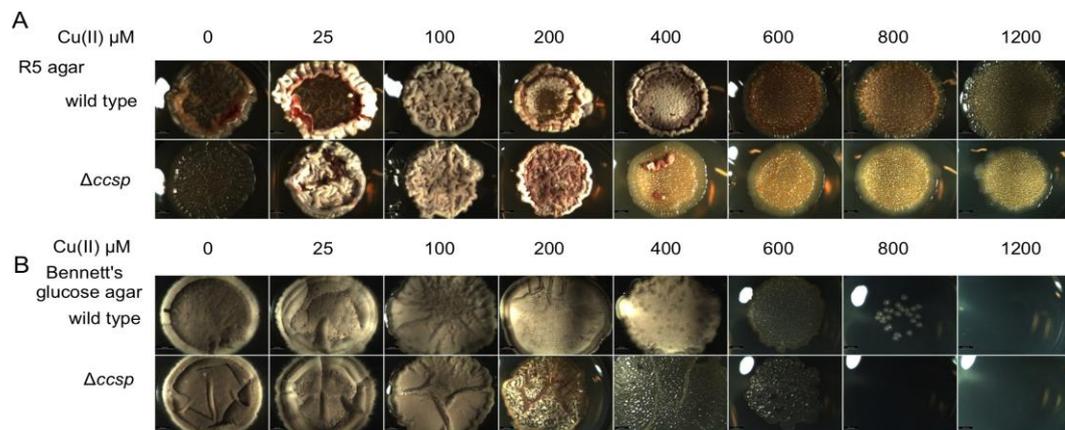
Figure 3.15- *The effect of Ccsp on the growth and development of S. lividans at 30 °C. Cu tolerance after 6 days growth of the wild type parent strain and the Δccsp mutant strain on R5 agar media (A) and Bennett's glucose media (B). Cu(II) concentrations as indicated. All images are the same magnification with a scale bar of 2 mm.*

concentrations was determined in the WT strain and the Δ*ccsp* mutant on various media. As illustrated on Bennetts-glucose agar (Fig. 3.14D) the CcO activity of the Δ*ccsp* mutant is identical to the WT, therefore demonstrating that Ccsp is not participating in the Cu trafficking pathway for maturation of CcO.

3.3.9 Analytical gel filtration chromatography - Cu(I) transfer between CopZ3079 and Ccsp

The movement and trafficking of Cu in the cytosol of *S. lividans* under homeostasis and stress has been shown to involve CopZ-like Cu metallochaperones (17, 18). CopZ-like Cu metallochaperones are involved in Cu(I) homeostasis in the cytosol of *S. lividans* (17, 18). These Atx-1 homologues possess a βαββαβ-fold and a MXCXXC metal binding motif that utilizes the two Cys residues for bis-cysteinate Cu(I) ion coordination (162). Therefore, it was investigated *in vitro* whether a CopZ plays a role in Cu(I) trafficking either to and/or from Ccsp. Transient ligand-exchange mechanisms occur in metal trafficking from donor to acceptor *in vitro* to allow ease of metal transfer ensuring that no unbound metal is present in solution (163, 164). Thus, by simply mixing the donor and acceptor in solution, transfer can occur. To test this with Ccsp and CopZ a method using gel-filtration was devised. CopZ-3079 (hereafter CopZ) in the apo monomeric form (M_r ~ 8 kDa) displays an elution peak on a size-exclusion column at ~12.5 ml (Fig. 3.16A). At Cu(I):CopZ ratios higher than 1:1 a shift in the elution profile (~11 ml) is observed, which based on the column calibration is consistent with the presence of a dimer species. Binding of Cu(I) to *Bacillus subtilis* CopZ has been shown to be a very complex process with initial binding of Cu(I) resulting in dimerization to form a $Cu_n^+(CopZ)_2$ species which has the capacity to bind three further Cu(I) ions at the monomer interface to create a $Cu_4^+(CopZ)_2$ species as the addition of stoichiometric Cu(I) ions increase (165-167). In the present work with *S. lividans* CopZ, it has not been determined how many

Cu(I) ions, n, are at the monomer interface ($Cu_n^+(CopZ)_2$). However, based on the size-exclusion chromatography profile, a protein dimer species is predominately formed at > 1 Cu(I)/CopZ. Thus, if Cu were to be transferred from the CopZ then the peak at the higher elution volume would be observed.
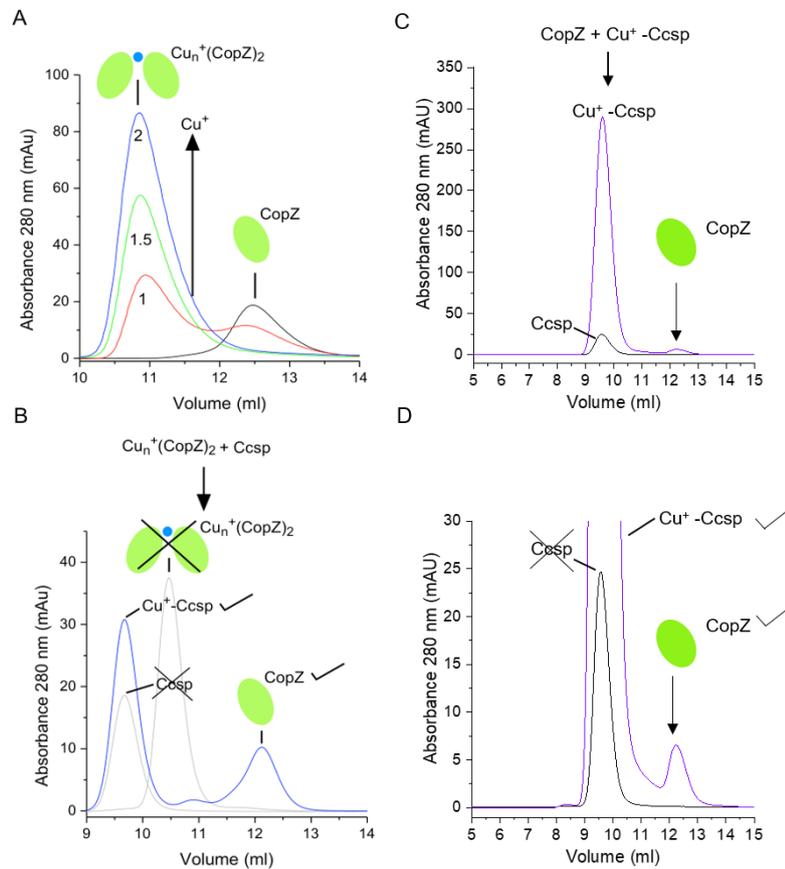
Figure 3.16 - Cu trafficking from CopZ to Ccsp. (A) (the data in chromatogram (A) were provided by Dr A.K. Chaplin) Size exclusion elution profiles of S. lividans CopZ prior to the addition of Cu(I) (black line) and post-incubation with 1, 1.5 and 2 equivalents of Cu(I). In the absence of bound Cu(I), CopZ elutes as a monomer (CopZ), with a dimer species, Cu_n^+–$(CopZ)_2$, predominantly formed in the presence of >1Cu(I) equivalents. (B) A size exclusion experiment whereby Cu_n^+–$(CopZ)_2$ is mixed with Ccsp and the resulting products indicated in the blue elution profile. The grey elution profiles indicate where the starting samples would elute if no Cu(I) transfer had occurred. (C) - Analytical gel filtration chromatogram for apo-Ccsp (black line) and 1:1 ratio of Ccsp–Cu(I) (loaded ~15 equivalents of Cu(I)) and apo-CopZ3079 (purple line). All samples and column were prepared in 10 mM MOPS pH 7.5, 150 mM NaCl. (D) Is a magnified image of the peaks in fig. (C) showing the apo-CopZ3079 indicated with the arrow, in the 1:1 mixture at ~12.1 ml.

The profile in Fig. 3.16B reveals transfer of Cu(I) from Cu(I)-CopZ to apo-Ccsp in a 1:1 ratio. In the blue coloured elution profile, mixing $Cu_n^+(CopZ)_2$ and apo-Ccsp (Fig. 3.16B) results in the presence of the elution peak for the monomeric apo-CopZ (~12.1 ml) and therefore supporting the notion that Cu(I) has been transferred. The same experiment was repeated but carried out in reverse by loading the Ccsp with Cu(I), and mixing with apo-CopZ with incubation periods of > 1 h (Fig. 3.16 C & D). Formation of the higher-order CopZ peak was

not observed suggesting that Cu(I) transfer from Ccsp does not occur under these conditions. But this may be due to a lack of sensitivity of analytical gel filtration as possible transfer of <1 Cu(I) ion could have occurred. This amount of Cu(I) transferred would be insufficient to cause dimerization of CopZ thus making detection of transfer difficult through gel filtration alone. Taken together these results support the notion that *in vitro* CopZ can transfer Cu(I) ions to Ccsp, but the possible release of Cu(I) from Ccsp to CopZ is not initially observed under the conditions employed.

3.4 Discussion

The bacterial cytosol has no known metabolic requirement for Cu and therefore Cu storage proteins had been thought not to exist. The discovery of Ccsp from *S. lividans* and extensive distribution of cytosolic Csp3 members in the Tree of Life as shown in Chapter 2 offers the tantalising possibility that new layers of Cu resistance or a possible cytosolic requirement for Cu exists amongst many bacteria that are yet to be fully understood. The subsequent analyses of Ccsp has shown many similarities in structure and stoichiometry of Cu(I) binding compared to the first discovered members of the Csp family reported by Vita *et al*, 2015 (1). For instance, *Mt*Csp3 contains 18 Cys residues in the core of each protomer unit and exists as a homotetramer in solution (1, 2, 75, 168) as is the case here with Ccsp.

The K_{Cu} value determined for the Ccsp lies within the range 10^{-17} to 10^{-18} M for cuproproteins with bis-cysteinate Cu(I) coordination (160). However, we note that our K_{Cu} values using BCA as the affinity probe are at the lower end to those determined for other Csp3 members (3.1 x 10^{-17} M *vs* 6 x10^{-18} M, average for *Mt*Csp3 and *Bs*Csp3 (2)). In addition, Ccsp demonstrates cooperative binding with an average Hill coefficient (*n*) value of 1.9. Cooperativity is also seen in *Mt*Csp1 (1) but surprisingly this is not so for *Mt*Csp3 and *Bs*Csp3 (2). Though concerning the cooperativity of Ccsp, it must be considered that the [Cu$^+_{free}$] values are very low (Fig. 3.11B) and calculated indirectly from binding to BCA. Regardless, concluding that cooperativity occurs in Ccsp, then it can be suggested that the mechanistic basis for this phenomenological result may lie in the observations by Dennison and co-workers that Cu(I) clusters of the type [Cu$_4$(S-Cys)$_4$] are thermodynamically favoured (76).

The X-ray crystal structures of apo-Ccsp and Cu(I)-loaded Ccsp (Fig. 3.7 and 3.13) offer further valuable insight into the structural characteristics of Csp3 members. These cytosolic proteins can store large amounts of Cu(I) and most of the Cu is bound by Cys residues. No evidence for disulfide bond formation is observed and a maximum of 20 Cu(I) ions are found to be coordinated in a Ccsp protomer. Therefore, Ccsp has the highest Cu(I) binding capacity of any Csp3 member so far characterised (2). Furthermore, Ccsp is the first

Csp3 member to reveal a unique coordination role of a highly-conserved Asp residue (Asn in *Mt*Csp3 (2)). This residue is positioned at the end of the Cys cage that harnesses Cu(I) ions 1 to 13, which are all coordinated by two Cys ligands (Fig. 3.13A & B). Asp61 is situated such that it breaks this coordination trend by providing coordination to Cu(I) ions 14 and 15 through its Oδ1 and Oδ2 carboxylate atoms, respectively. The Asp61 is homologous to Asp42 in *Bs*Csp3 whereas *Mt*Csp3 has an Asn at the same position (2). The five remaining Cu(I) ions form a distinct cluster and is coordinated by three conserved His residues and four Cys residues. The three His residues clearly guard the more solvent accessible end of the four-helix bundle core and in addition to their coordination role may also play a part in partner recognition to assist Cu(I) loading to Ccsp.

CopZ Cu chaperones are mainly involved in a Cu stress response but have been shown to help buffer the bacterial cytosol (17, 18, 119, 163). The possibility of a CopZ being involved in Cu transfer with Ccsp was investigated through size-exclusion chromatography (Fig. 3.16). Whilst transfer does occur it proceeds only in one direction from a Cu_n^+–$(CopZ)_2$ species to apo-Ccsp (Fig. 3.16B). Under homeostasis conditions *S. lividans* has three CopZ/ATPase couples, expressed at a basal level, that serve to buffer the Cu(I) concentration in the cytosol (17, 18). One of these couples is not characterised but does act under Cu stress whereas the other two are under direct transcriptional control of CsoR (18). A fourth couple is also present in the genome and controlled by CsoR but is not constitutively present under Cu homeostasis (17). It has been well established *in vitro* that CsoR accepts Cu(I) from two CopZ chaperones in a unidirectional way (17). This would induce transcription of three of the couples *in vivo* that are under CsoR influence (17). In addition, the fourth couple not regulated by CsoR would create a substantial defence against Cu stress (17). *In vitro,* two of the CopZ proteins have been determined to transfer Cu(I) to the CsoR in a unidirectional manner (17). Therefore *in vivo* this would induce transcription of the three CopZ/ATPases couples under CsoR control, which together with the fourth non-CsoR controlled couple creates a high capacity for Cu resistance under Cu stress. The ability of CopZ to traffic Cu(I) to CsoR and P_1-type ATPases indicates a characteristic promiscuity in these small chaperone proteins for off-loading their metal cargo. This indiscriminate nature is also observed for CopZ with Ccsp, where Cu(I) is transferred from the Cu_n^+–$(CopZ)_2$ species to Ccsp as evidenced by the reformation of the apo-CopZ monomer (Fig. 3.16B). In contrast, under stoichiometric conditions, apo-CopZ is unable to remove Cu(I) from Cu(I)–Ccsp in a physiologically meaningful time frame as has also reported for *Bs*Csp3 with its cognate *Bs*CopZ (2). A K_{Cu} value of 3.9×10^{-18} M has been determined for this particular *S. lividans* CopZ used in this

work under the same buffer conditions as employed in the transfer experiments with Ccsp (17). This value is lower than the K_{Cu} value determined for Ccsp and thus from a thermodynamic perspective the transfer can be viewed as being unfavourable. However, other factors such as the tuning of the reactivity of the ligands involved in the transfer of Cu(I) from donor to acceptor are important considerations and have been shown to influence the directionality of transfer. Such an example has been reported for the transfer of Cu(I) from CopZ to the CsoR, where the Cys residues in the CXXXC motif of CopZ have been optimised to favour Cu release to the acceptor (CsoR) (17).

The *in vivo* data show that Ccsp is not required for Cu homeostasis at Cu concentrations up to several hundred μM depending on the medium used (Fig. 3.14 & 15). This observation is further demonstrated by the CcO activity data, which precludes a downstream role for Ccsp in supplying Cu(I) to the extracytoplasmic environment to be utilised by the Cu-chaperones Sco and ECuC for metalation of CcO and GlxA (12, 44, 101). The CopZ/ATPase couples appear to maintain the control of cytoplasmic Cu levels. However, as exogenous Cu concentrations rise above 200 μM, a clear phenotype for Ccsp is observed though with a range limit that is strongly medium dependent (Fig. 3.14). Re-analysed RNA-seq data (18) reveals transcription of *ccsp* is up-regulated 5-fold in liquid defined medium supplemented with 400 μM Cu and thus fits with the phenotype in Fig. 3.14 showing that Ccsp becomes essential for growth and development in the 200–500 μM Cu range. Importantly, in contrast to three out of the four CopZ/ATPase couples, the Ccsp expression is not under the control of CsoR as demonstrated by the absence of a consensus CsoR binding site in the *ccsp* promoter region and expression induction in the *csoR* mutant (18). This all suggests that a second layer of Cu responsive transcription is operating on top of the CsoR regulon in *S. lividans* and becomes operative at more extreme Cu concentrations to express *ccsp*.

In conclusion, the results confirm the discovery of Ccsp in *S. lividans*; the crystallography data strongly supports a tetramer of alpha-helical bundles capable of binding up to 20 Cu(I) ions with cooperativity of Cu binding. The overall interaction of Ccsp with a CopZ in Cu(I) transfer and the *in vivo* data also enforces the idea that when the CsoR efflux system becomes saturated during Cu stress, Ccsp accepts Cu from CopZ and stores it as the environment returns to homeostasis. It remains unclear how the release of Cu from Ccsp occurs and what further actions take place with this excess Cu. It is a possibility to consider that non methanotrophic bacteria may possess Cu scavenging systems, like methanobactins in methanotrophs. In support of this, a diisonitrile compound produced from a non-ribosomal

peptide synthetase in *S. thioluteus* has recently been reported and shown to have a chalkophore function (i.e. Cu-import into the cytosol) (5). This discovery could suggest that chalkophores are more widespread than originally considered and an interplay with Ccsps in non-methanotrophic bacteria is a possibility and certainly requires further investigation.

Chapter Four

Visualising Cu(I) loading to *Sl*Csp3 and the effect of His and Cys mutations

4.1 Introduction

Structural studies of non-methanotrophic Csp3 members reveal structural homology with *Mt*Csp3, (2, 19) but only the Csp3 from the Gram-positive bacterium *S. lividans*, *Sl*Csp3 (previously referred to as Ccsp), has been structurally characterised with Cu(I) bound as described in Chapter 3, revealing up to 20 Cu(I) ions can bind per protomer (80 per homotetramer) (19). In addition to the extra Cys residues lining the four helix-bundle core in Csp3s compared to *Mt*Csp1/2, a set of three His residues are present at one end of the four helix-bundle, which together with nearby Cys residues also participate in Cu(I) coordination (2, 19). The His end of the four helix-bundle is considered from structural insights to be the loading and leaving point for Cu(I), as access to the Cys core from the opposite end of the bundle is prevented by hydrophobic side chains (2, 19). However, no experimental evidence has been reported to prove this. Further Cu(I) coordination arises from the $O^{\delta 1}$ atom of an Asn residue (58 in *Mt*Csp3) (2) and the $O^{\delta 1}$ and $O^{\delta 2}$ atoms of an Asp residue (61 in *Sl*Csp3) (19). The Asp and the Asn residues are highly conserved across Csp3 species and are structurally positioned to create a crossing point, dividing the Cu(I) ions participating in His/Cys coordination at the mouth of the Cys core and those coordinated solely through bis-cysteinate coordination in the Cys core.

Insights into Cu(I)-loading of the *Mt*Csp3 have been obtained through X-ray crystallography studies (76). Structures determined at various stoichiometric loadings of Cu(I) reveal the existence of initial tetranuclear Cu-thiolate clusters $[Cu_4(\mu_2\text{-S-Cys})_4]$ in the Cys core (Cu sites 3 to 14) of the four helix-bundle (76). As more Cu(I) is loaded, the tetranuclear clusters considered as 'intermediates' evolve into the final Cu(I) clusters (76). Thus, the formation of tetranuclear clusters is considered a driving force for acquisition and safe storage of Cu(I) by Csp3 members (76, 169).

In this chapter, Cu(I)-loading to *Sl*Csp3 has been investigated by using a combination of X-ray crystallography and site-directed mutagenesis (170). It has been discovered that at low Cu(I) loading, a tetranuclear $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})]$ cluster is first formed in the His entrance of the four helix-bundle. As more Cu(I) is loaded, the Cys core of the four helix-bundle of *Sl*Csp3 together with Cu sites in the His entrance become occupied to varying extents but no evidence for the formation of tetranuclear clusters in the Cys core is observed, consistent with a highly fluxional process of Cu(I) binding. *Sl*Csp3 protein variants have been created in which the His residues at the hydrophilic entrance of the Cys core have been changed to Ala and two Cys variants also in the His entrance have been mutated to Ser, to begin to build a picture of how these residues influence Cu(I) loading and binding.

4.2 Methods

4.2.1 Site-directed mutagenesis, over-expression and purification

To create the H107A, H111A, H113A, C41S, and C57S, *Sl*Csp3 site-directed variants, the Quikchange (Stratagene) site-directed mutagenesis strategy was carried out and forward and reverse mutagenic primers were designed and synthesised (Sigma-Aldrich) with the respective nucleotide change(s) (Table 4.1). A double His variant, H107A/H111A, was also constructed and this variant was created by taking the H107A construct and then using the H111A primers (Table 4.1) for site-directed mutagenesis PCR (Tables 4.2 and 4.3). Mutant clones were corroborated for the presence of the desired mutation by DNA sequencing (Source Bioscience). Recombinant production in *Escherichia coli* BL21(DE3) cells and purification of wild-type (WT) *Sl*Csp3 and the His variants was carried out as described in

Table 4.1 - *Mutagenic forward and reverse primers used for site directed mutagenesis to create the SlCsp3 variants. The codons highlighted show where the mutation was made and the nucleotides in uppercase show this single nucleotide change.*

Variant	Forward primer	Reverse primer	T_m (˚C)
H107A	5'-GTGCGCCCGGgcCGCCGGCATG-3'	5'-CATGCCGGCGgcCCGGGCGCAC-3'	74
H111A	5' – CACGCCGGCATGgcCGAGCACTGCC- 3'	5' – GGCAGTGCTCGgcCATGCCGGCGTG – 3'	73
H113A	5' -CATGCACGAGgcCTGCCGGGTC- 3'	5' –GACCCGGCAGgcCTCGTGCATG- 3'	67
C41S	5' GTGCACCGCGAgcGCCGACGCC 3'	5' GGCGTCGGCgcTCGCGGTGCAC 3'	72
C57S	5' GATCTGACCAAGAgcATCCGCACCG 3'	5' CGGTGCGGATgcTCTTGGTCAGATC 3'	63

Table 4.2 – *The reagents and volumes used for site directed mutagenesis.*

Reagent	Concentration	Volume (µl)
Plasmid DNA	25 ng/ µl	1.0
Forward Primer	75 ng/ µl	1.0
Reverse Primer	75 ng/ µl	1.0
dNTPs	10 mM	0.6
PFU Buffer	10 x	3.0
PFU Turbo Polymerase	-	0.6
DMSO	5 %	1.5
Sterile ddH2O	-	21.3
Total		30.0

Chapter 3. Far UV-circular dichroism spectroscopy using a Chirascan CD spectrometer (Applied Photophysics) was used to assess whether the proteins were folded. All proteins once purified were stored at -20 °C until required.

Table 4.3 *– QuickChange site-directed mutagenesis protocol*

Temperature (°C)	Time (minutes)
95	3
95	0.5
58 (Annealing temperature)	1
68	13
68	8 (Extension)
10	Hold

4.2.2 X-ray crystallography

Under anaerobic conditions WT *Sl*Csp3 (2000 μM) was incubated with 5, 10 and 25 molar equivalents of CuCl, and to the *Sl*Csp3 variants (1500 to 3900 μM) 25 molar equivalents of CuCl was added. Unbound Cu was removed by passing samples through a PD-10 column (Generon) and concentrated to ~ 10-15 mg/ml. Optimisation of crystallisation conditions for the Cu(I)-loaded samples was carried out by screening around 1.4 to 1.6 ammonium sulfate, 0.1 M MES pH 6.0 by mixing equal (1 μl) volumes of protein and reservoir solution. Crystals were transferred to a cryoprotectant solution consisting of 40 % w/v sucrose and precipitant, and flash cooled by plunging into liquid nitrogen. WT Cu(I)-loaded *Sl*Csp3 crystals were measured at the Swiss Light Source on beamline XS10A using an X-ray wavelength of 1.33 Å and a Pilatus 6M-F detector. Crystals of the His variants were measured at the Diamond Light Source on beamline I04 using an X-ray wavelength of 0.979 Å and a Pilatus 6M-F detector. All data were indexed using XDS (148) and scaled and merged using Aimless (149) in the CCP4 suite with the CCP4i2 interface. Structures were solved by molecular replacement in MOLREP using the apo-*Sl*Csp3 structure (PDB ID 6EI0) as the search model. Cycles of model building in Coot (152) and refinement in Refmac5 (153) were carried out and riding hydrogen atoms were added when refinement of the protein atoms had converged. For all data sets anomalous difference maps for validation of Cu(I) atom positions was generated using PHASER (150) in the CCP4i2 interface. Structures for His variants and partially Cu(I) loaded *Sl*Csp3 samples were validated using the Molprobity server (154) the JCSG Quality Control Server and tools within Coot (152). Coordinates and structure factors were deposited in the RCSB Protein Data Bank. A summary of data, refinement statistics and the quality indicators for the structures are given in Tables 4.4 & 4.5.

Table 4.4 - Crystallographic data processing and refinement statistics for the partial Cu(I) loaded forms of SlCsp3 used in this work. Values in parenthesis refer to the outermost resolution shell.

Structure	Ccsp-5Cu	Ccsp-10Cu
Space group	$P6_122$	$P6_122$
Unit cell (Å)	93.1, 93.1, 212.3	93.4, 93.4, 216.0
Resolution (Å)	75.4-1.50	75.8-1.90
Unique reflections	85587 (3492)	44544 (2831)
Mn (I/SD)	19.5 (1.3)	12.8 (1.2)
$CC_{1/2}$	0.99 (0.45)	0.99 (0.46)
Completeness (%)	98.0 (82.5)	99.7 (100)
Redundancy	8.0 (8.7)	9.1 (9.2)
R_{cryst}	0.195	0.218
R_{free}	0.214	0.255
RMS dev. bond lengths (Å)	0.013	0.015
RMS dev. bond angles (º)	1.47	1.64
Ramachandran favoured (%)	99.8	98.2
PDB accession code	6Q58	6Q6B

Table 4.5 - Crystallographic data processing and refinement statistics for the variants of SlCsp3 used in this work. Values in parenthesis refer to the outermost resolution shell.

Structure	C41S	C57S	H111A	H113A	H107A/H11A
Space group	I222	I222	$P2_12_12$	I222	I222
Unit cell (Å)	62.2, 64.2, 65.1	63.4, 64.7, 66.6	65.3, 62.2, 65.2	62.2, 65.1, 65.3	62.2, 64.0, 65.7
Resolution (Å)	44.70-1.49	46.45-2.11	65.3-1.20	32.7-1.30	44.6-1.20
Unique reflections	21446 (1039)	7385 (408)	86182 (3562)	33353 (1526)	42048 (1551)
Mn (I/SD)	12.9 (1.9)	16.8 (2.1)	17.6 (1.2)	23.7 (1.4)	18.6 (1.0)
$CC_{1/2}$	0.99 (0.58)	0.99 (0.62)	0.99 (0.61)	0.99 (0.59)	0.99 (0.50)
Completeness (%)	98.8 (97.4)	90.0 (99.0)	98.1 (81.6)	96.7 (89.2)	96.9 (73.1)
Redundancy	7.6 (7.9)	4.1 (4.4)	6.3 (3.3)	7.3 (7.3)	7.1 (4.1)
R_{cryst}	0.229	0.236	0.182	0.161	0.168
R_{free}	0.242	0.295	0.194	0.187	0.192
RMS dev. bond lengths (Å)	0.06	0.20	0.0064	0.0052	0.0079
RMS dev. bond angles (°)	2.44	10.23	1.40	1.33	1.48
Ramachandran favoured (%)	98.3	79.6	100.0	96.7	100.0
PDB accession code	-	-	6QYB	6QVH	6R01

4.3   Results

4.3.1   Crystal formation and conditions

Crystallisation trials with low Cu(I) concentrations of WT *Sl*Csp3 (5, 10 Cu(I) equivalents and 25 Cu(I)-equivalents) to all *Sl*Csp3 variants were carried out. The conditions used for these trials were 1.2 – 1.9 M ammonium sulphate with 0.1 M HEPES pH 7 or 0.1 M MES pH 6. The variants that were successfully crystallised were H111A, H113A, H107A/H111A, C41S and C57S. Crystals also
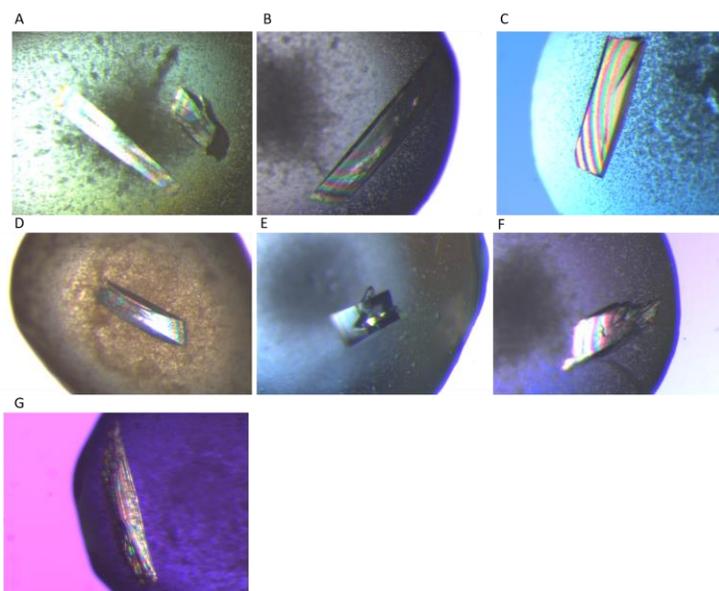
Figure 4.1 – Images of crystals of SlCsp3 variants and partial Cu(I) loaded proteins. (A) 5 Cu(I) (B) 10 Cu(I) (C) H111A (D) H113A (E) H107A/H111A (F) C41S (G) C57S.

grew of partially loaded *Sl*Csp3, with 5 and 10 Cu(I) molar equivalents. The morphology of these crystals can be seen in Fig. 4.1 whereby they were often large block crystals but occasionally some crystals were smaller fragmented cube like structures.

4.3.2   Definition of Cu(I) sites and their grouping within SlCsp3

From previous structural characterisation of Csp3 members, modes of Cu(I) binding within the protein have been discussed and defined (2, 19, 169). Prior to reporting the results from the present study, Cu(I) sites, cores and coordination are briefly defined. For *Sl*Csp3 the positions of the 20 Cu(I) ions of the fully Cu(I)-loaded form are shown in Fig. 4.2A and can be subdivided into outer and inner cores (dashed red lines Fig. 4.2) that incorporate the Cu(I) sites, 1-14 (inner), and 15-20 (outer) (Fig. 4.2A) (19). Within these two cores Cu(I) coordination can be divided into three groups based on differences in coordination environment (169). In group I, Cu(I) ions are coordinated by two Cys thiolates on the same helix in a CXXXC motif, in group II, by two Cys thiolates on different helices of the four helix-bundle and in group III, by Cys thiolates and other residues e.g. $N^{\delta 1}$ or O^{δ} atoms from His or Asp, respectively. The Cu(I) ions in *Sl*Csp3 assigned to each of these groups are reported in the legend to Fig. 4.5.

4.3.3   Polynuclear Cu(I) clusters form in the outer core at low Cu(I) loading

The X-ray structure of *Sl*Csp3 incubated with 5 Cu(I)-equivalents was found to contain four protomers in the asymmetric unit in a similar manner to apo-*Sl*Csp3 (19). Anomalous electron-density map features were evident in each protomer, which were used, together with strong peaks in the *2Fo-Fc* map to assign the location of bound Cu(I) ions (Fig. 4.2B). Of the four protomers, one displayed anomalous electron-density peaks consistent with the presence of three Cu(I) ions with the remaining protomers consistent with the presence of four Cu(I) ions (Fig. 4.2B). The Cu(I) ions are coordinated by the $N^{\delta 1}$ atom of His113 and the thiolates of Cys41, Cys57, Cys104 and Cys114 (Fig. 4.2C). One of the Cu(I) ions, occupies the same binding site as Cu17 in the fully Cu(I)-loaded *Sl*Csp3 and has an identical group III coordination sphere of Cys114/Cys57/His113 (19). In contrast, two Cu(I) ions, although occupying similar positional locations to Cu15 and Cu18 in the fully Cu(I)-loaded *Sl*Csp3 structure, they are distinct in that they have an altered coordination sphere (19). In this respect, Cu15 (group III coordination) is not positioned close enough to the $O^{\delta 2}$ atom of Asp61 to fulfil the requirements for a coordinate bond and is now assigned as group II coordination and Cu18 (group II coordination) no longer coordinates to Cys45 but instead, occupies a position enabling coordination by Cys114, but remains group II coordination (Fig. 4.2C). To recognise these differences compared to the fully Cu(I)-loaded structures these Cu(I) sites have been designated Cu15* and Cu18*. A fourth Cu(I) ion is present at a site, which is absent in the fully Cu(I)-loaded *Sl*Csp3, and is designated as a non-cognate Cu(I) binding site with group II coordination (Fig. 4.2C green circle).

Group II and III dominate the coordination chemistry of the Cu(I) ions in the outer core of the four helix-bundle at low Cu(I) loading (Fig. 4.3A), creating a negatively charged trinuclear $[Cu_3(\mu_2\text{-S-Cys})_2(\text{S-Cys})_2(N^{\delta 1}\text{-His})]^-$ cluster and a neutral tetranuclear $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})]$ cluster (Fig. 4.3A). The latter is symmetrical in that all Cys thiolates are bridging (μ_2-S-Cys) a Cu(I) ion, whereas in the trinuclear cluster this symmetry is broken as two Cys thiolates display monodentate Cu(I) coordination (Fig. 4.3A). The Cu-$N^{\delta 1}$(His113) bond distance is 2.1 Å in both CuS clusters and the Cu-Sy(Cys) bond distances range between 1.9-2.2 Å. In addition, interactions (2.5-3.1 Å) between Cu(I) ions within the $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-}His)]$ cluster are observed. From the anomalous electron-density peaks, it is apparent in the trinuclear cluster that Cu15* exhibits weaker electron-density, which is attributed to a lower
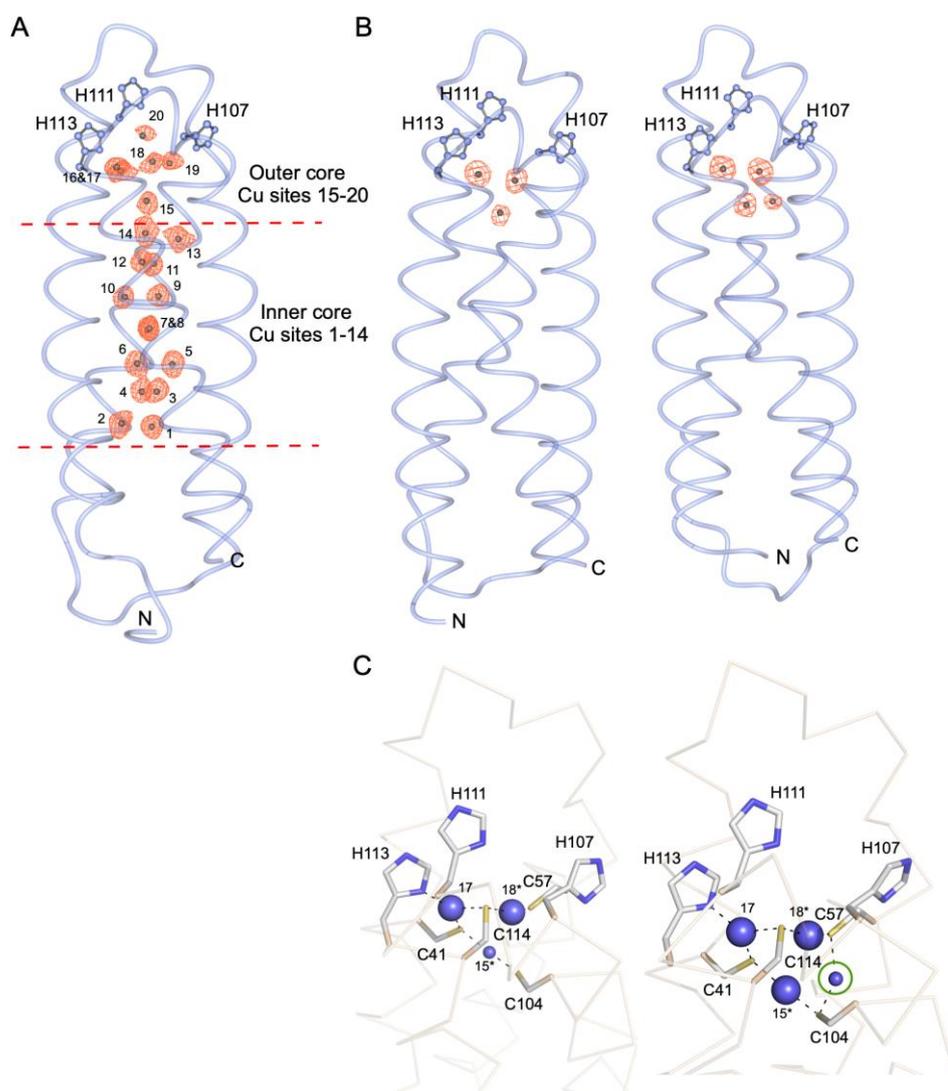
Figure 4.2 *- Location of Cu(I) sites in fully and partially Cu(I)-loaded SlCsp3 determined by X-ray crystallography. The location of bound Cu(I) ions are inferred by the anomalous electron-density peaks shown in orange mesh and contoured at 5σ. A) Fully Cu(I)-loaded structure with the location and the number of Cu(I) ions found in the inner and outer cores indicated by dashed red lines (PDB 6EK9) (19). B & C) 5 Cu(I)-equivalent structures. B) Two protomers showing three and four Cu(I) bounds, respectively. C) Coordination chemistries found in the outer core of the two promoters in (B), with Cu(I) ions represented in blue spheres and coordinate bonds as dashed lines. The smaller spheres indicate partial occupancy based on the anomalous electron-density peaks, and the green circle indicates a non-cognate site with a Cu(I) ion bound.*

Cu(I) occupancy relative to the other Cu(I) coordination sites (Fig. 4.2B). Notably, in the $[Cu_4(\mu_2\text{-}S\text{-}Cys)_4(N^{\delta 1}\text{-}His)]$ cluster, the anomalous electron-density peak for Cu15* is consistent with a higher occupancy, whereas the anomalous electron-density peak for the non-cognate Cu(I) ion required to form the tetranuclear cluster has lower occupancy (Fig. 4.2B). This could imply that the $[Cu_4(\mu_2\text{-}S\text{-}Cys)_4(N^{\delta 1}\text{-}His)]$ cluster is formed in a sequential manner, whereby Cu17 and Cu18* are bound first, followed by Cu15* (trinuclear) and finally binding to the non-cognate site to create $\mu_2\text{-}S\text{-}Cys57$ and $\mu_2\text{-}S\text{-}Cys104$ coordination (Fig. 4.3A).
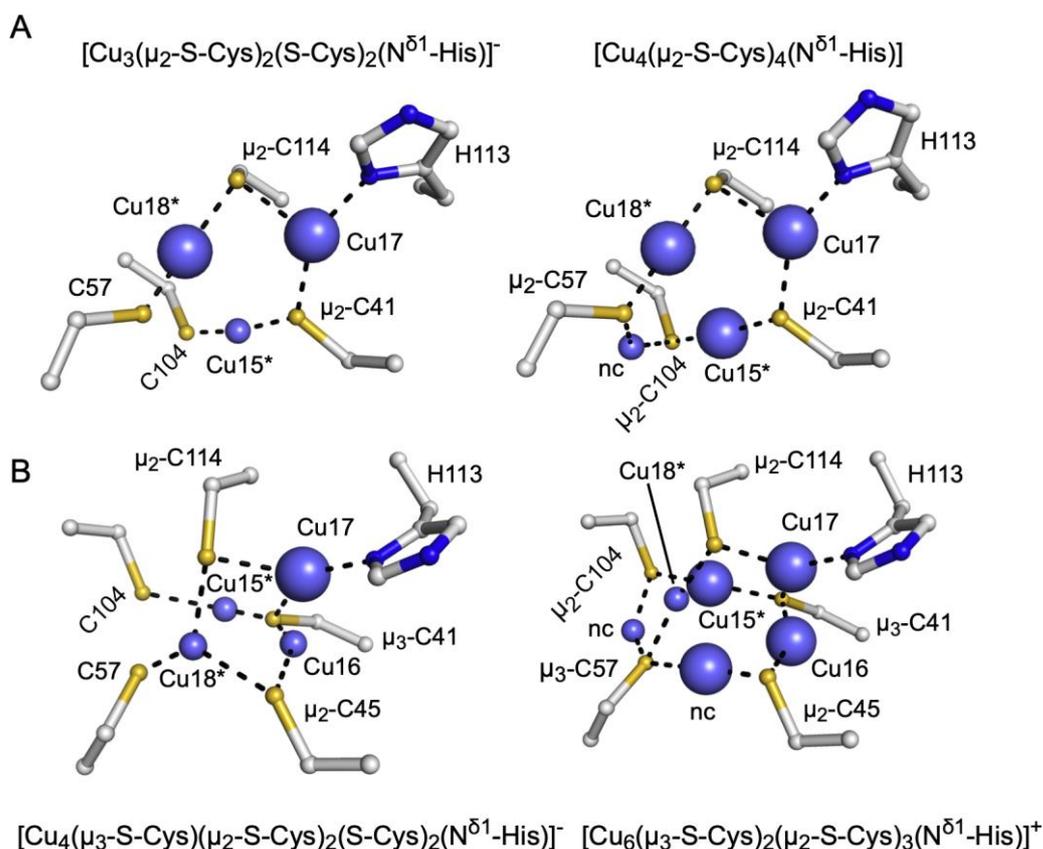
Figure 4.3 - *Polynuclear Cu(I) clusters present in the outer core of SlCsp3. The clusters present in the 5 Cu(I)-equivalent structure (A) and the 10 Cu(I)-equivalent structure (B). The Cu(I) ions (blue spheres) bound at non-cognate sites are labelled, nc, and the smaller blue spheres indicate a lower occupancy as determined form the anomalous electron-density peaks.*

4.3.4   Cu(I) ions fill the inner core in a dynamic and fluxional manner

In the X-ray structure for *Sl*Csp3 incubated with 10 Cu(I)-equivalents, surprisingly, the anomalous electron-density maps revealed more sites occupied than Cu(I) equivalents added (Fig. 4.4), indicating that some of the sites are not fully occupied. In protomer A, fourteen Cu(I) ions have been modelled into the anomalous electron-density map, with ten of these positioned in the inner core and four located in the outer core (Fig. 4.4). All Cu(I) ions observed in protomer A are occupying cognate sites *i.e.* found in the fully Cu(I)-loaded *Sl*Csp3 structure, but the anomalous electron-density peaks for Cu(I) ions 11, 12, 15*, 16 and 18* (Fig. 4.4), indicates reduced occupancy. Notably, no anomalous electron-density peaks are observed for sites 1, 2 and 4 in the inner core, which is also the case in the other protomers that make up the crystallographic asymmetric unit (Fig. 4.4). For protomer B, the anomalous electron-density map is once more consistent with the presence of fourteen Cu(I) ions. However, Cu12 and Cu16 are absent, and anomalous electron-density peaks are present for two non-cognate sites filled with Cu(I) ions, and therefore distribution of ten Cu(I) ions in the inner core and four in the outer core as found in protomer A is maintained. For protomers C

and D (the latter not shown), eighteen Cu(I) ions are observed, with Cu12, Cu13 and Cu16 present, along with additional anomalous electron-density peaks located between Cu15* and 18* (green circle Fig. 4.4). Furthermore, Cu6, Cu8, and Cu14 have anomalous electron-density peaks consistent with lower occupancy relative to Cu(I) ions in other sites. These observations of different distributions and occupancies of Cu(I) ions within the inner and outer cores reflect a clear fluxionality of site occupancies during Cu(I) loading (Fig. 4.4).

An overview of the coordination chemistries of the Cu(I) ions in protomers A, B and C, together with the fully Cu(I)-loaded SlCsp3 are illustrated in Fig. 4.5. In common with the 5 Cu(I)-equivalent structure, only His113 out of the three His residues at the mouth to the
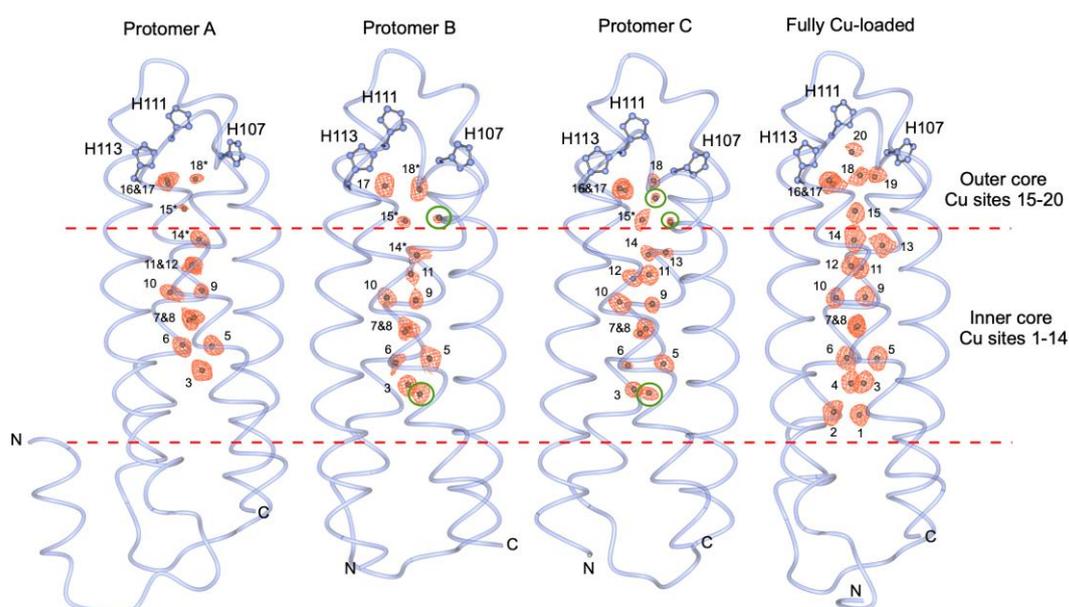
Figure 4.4 - X-ray structures of SlCsp3 with 10 Cu(I)-equivalents added. Protomers A, B and C are represented with the fully Cu(I)-loaded structure shown for comparison. Anomalous electron-density for the Cu(I) ions is depicted in orange mesh and contoured at 5σ. Green circles indicate the location of non-cognate sites containing a Cu(I) ion. In protomer A, electron-density is present in the asymmetric unit that enables for additional residues to be modelled at the N-termini, which now starts at residue 7, as opposed to residue 15 in the other protomers in the crystallographic asymmetric unit.

outer core is found to participate in coordination chemistry (Fig. 4.5). Furthermore, Cu15 is once more not coordinating to the $O^{\delta 1}$ atom of Asp61 (Cu15*), but the $O^{\delta 1}$ atom of Asp61 does maintain coordination to Cu14, albeit with a longer than average bond length of 2.8 Å compared to 2.2 Å in the fully Cu(I)-loaded structure. Notably, when Cu13 is absent (protomers A and B), Cu14 adopts a distorted tetrahedral coordination geometry (Cu14*) through coordination by Cys100, normally reserved for Cu13 (Fig. 4.5).

Polynuclear Cu(I)-thiolate clusters dominated by group II and III coordination are observed in the outer core for all protomers of the 10 Cu(I)-equivalent structure (Fig. 4.3B). In protomer B, the same neutral symmetrical tetranuclear [$Cu_4(\mu_2$-S-Cys)$_4(N^{\delta 1}$-His)] cluster as seen in the 5 Cu(I)-equivalent structure is present (Fig. 4.3A and 4.5). However, a different

tetranuclear cluster is found in protomer A, where Cys45 acts as a bridging (μ_2-S) ligand to Cu16 and Cu18*, and Cys41 as a μ_3-S ligand to Cu15*, 16 and 17 and Cys57 and Cys104 as monodentate ligands to create an asymmetric negatively charged [Cu$_4$(μ_3-S-Cys)(μ_2-S-Cys)$_2$(S-Cys)$_2$(N$^{\delta1}$-His)]$^-$ cluster (Fig. 4.3B). Notably the Cu18* position is shifted in this cluster in respect to the other clusters and becomes three coordinate (Fig. 4.3B). In protomer C, the two polynuclear clusters observed in protomers A and B combine, together with a second non-cognate Cu(I) ion (green circle Fig. 4.5) to form a positively charged hexanuclear [Cu$_6$(μ_3-S-Cys)$_2$(μ_2-S-Cys)$_3$(N$^{\delta1}$-His)]$^+$ cluster (Fig. 4.3B). The second non-cognate Cu(I) ion in this cluster is coordinated by Cys45 and Cys57, now making the latter a μ_3-S ligand (Fig. 4.3B). Beyond Cu14 and into the inner core no evidence of polynuclear CuS clusters is observed and Cu(I) ions are coordinated in their respective group 1 or group 2 coordination. Therefore, these polynuclear Cu(I)-thiolate clusters confined to the outer core serve to illustrate the coordinative flexibility inherent within the group II and group III sites and how these can adapt to increase cluster size whilst retaining either four or five Cys thiolates as ligands (Fig. 4.3).

Protomer A

Protomer B

Protomer C
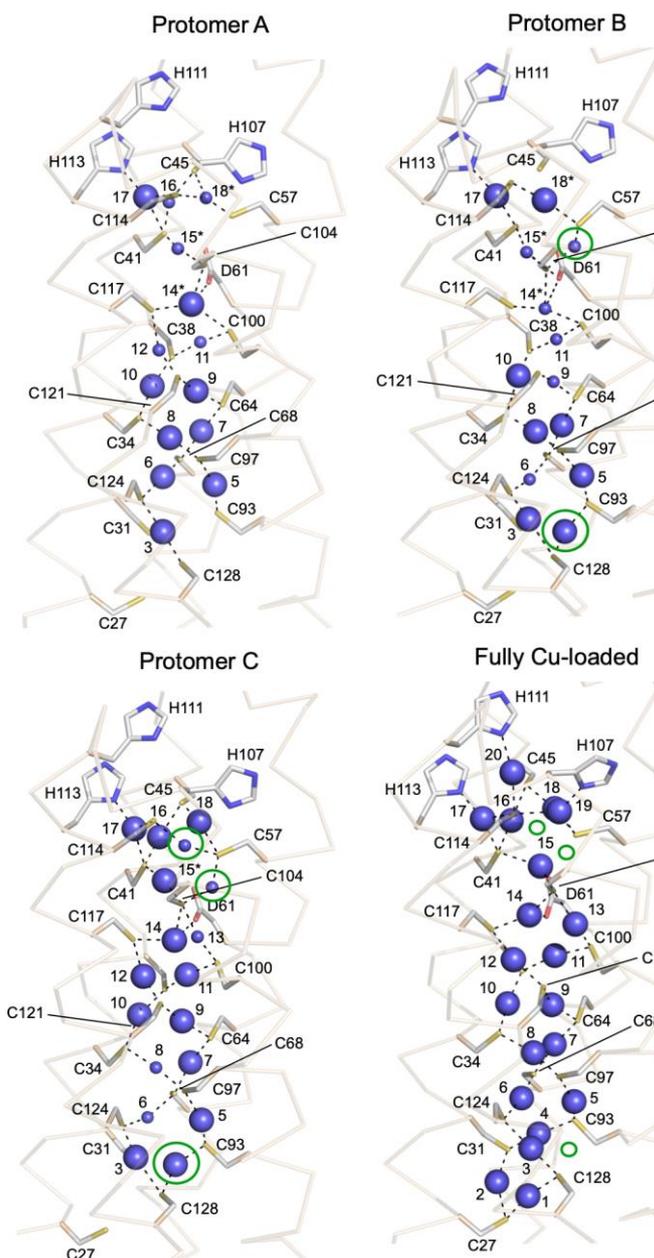
Fully Cu-loaded

Figure 4.5 - *Coordination chemistry of SlCsp3 with 10 Cu(I)-equivalents added. Protomers A, B and C are represented with the fully Cu(I)-loaded structure shown for comparison. Coordinate bonds to the Cu(I) ions (blue spheres) from Cys-S$^\gamma$, His-N$^{\delta 1}$ and Asp-O$^\delta$ atoms indicated by dashed lines. The smaller spheres indicate partial occupancy based on the anomalous electron-density peaks. Green circles indicate the location of non-cognate sites filled with a Cu(I) ion. The following Cu(I) ions belong to group I coordination, Cu2, Cu4, Cu5, Cu7, Cu10, Cu13 and Cu16; group II coordination, Cu1, Cu3, Cu6, Cu8, Cu9, Cu11, Cu12 and Cu18; group III 14, 15, 17, 19 and 20.*

4.3.5   Structural effects of the entrance His residues on Cu(I) ion coordination

The three His residues positioned at the solvent exposed entrance to the Cys lined core in *Sl*Csp3 were tested for their contribution to Cu(I) binding and loading. Each His was individually replaced with Ala to create the H107A, H111A and H113A variants. Furthermore, from inspection of the X-ray structures of the partially Cu-loaded *Sl*Csp3, His113 was consistently found to coordinate a Cu ion (Figs. 4.2 and 4.5). It was therefore hypothesized that His107 and His111 could be involved in initial Cu capture and facilitate transfer to His113. Therefore, the double His-variant, H107A/H111A, was constructed. All variants were purified in the apo-state and CD spectroscopy indicated the mutations caused no significant effect to the protein fold in solution.

To visualise the effect on the Cu(I) coordination chemistry on removing the His residues, X-ray crystallography studies were

carried out. X-ray structures of the H111A, H113A and the H107A/H111A variants after loading with 25 Cu(I)-equivalents were determined to the resolutions reported in Table 4.5,

and the positioning of Cu(I) ions inferred through creation of anomalous electron-density maps (Fig. 4.6). For the H111A variant, two *Sl*Csp3 protomers were identified in the crystallographic asymmetric, whereas only one *Sl*Csp3 protomer was found in the H113A and H107A/H111A variant structures. Suitable diffraction quality crystals for the H107A variant were not obtained. From the anomalous electron-density maps (Fig. 4.6), it is apparent that for all variants the inner core housing Cu(I) ions 1-14 retains identical coordination chemistry to the WT *Sl*Csp3, however, some variation is observed in the outer core (Fig. 4.6). For all variants, an additional Cu(I) ion is present that is absent in the fully Cu(I)-loaded WT *Sl*Csp3 structure, and is located adjacent to Cu15 and coordinated by the $O^{\delta 1}$ of Asp61 and S^{γ} of Cys 57 (Fig. 4.6 green circle). In the H111A and H107A/H11A structures both Cu19 and 20 are notable by their absence, implying that H111 and H107 are important for initial Cu loading. Removal of His113 results in the absence of Cu17 (Fig. 4.6), but would appear to release steric constraints enabling for the side chain of His111 to adopt an alternative conformation to that observed in the WT *Sl*Csp3 structure resulting in facile coordination to Cu20. It is noted that in one of the H111A protomers a further additional Cu(I) ion is present that shares the coordination of the $N^{\delta 1}$ atom of His113 with Cu17, with further coordination by the S^{γ} of Cys41. This brings the total number of Cu ions bound to 20 in one H111A protomer (19 in the other), eighteen regular sites and two additional sites, and underscores the adaptability and flexibility of Cu(I) coordination sites within these proteins.
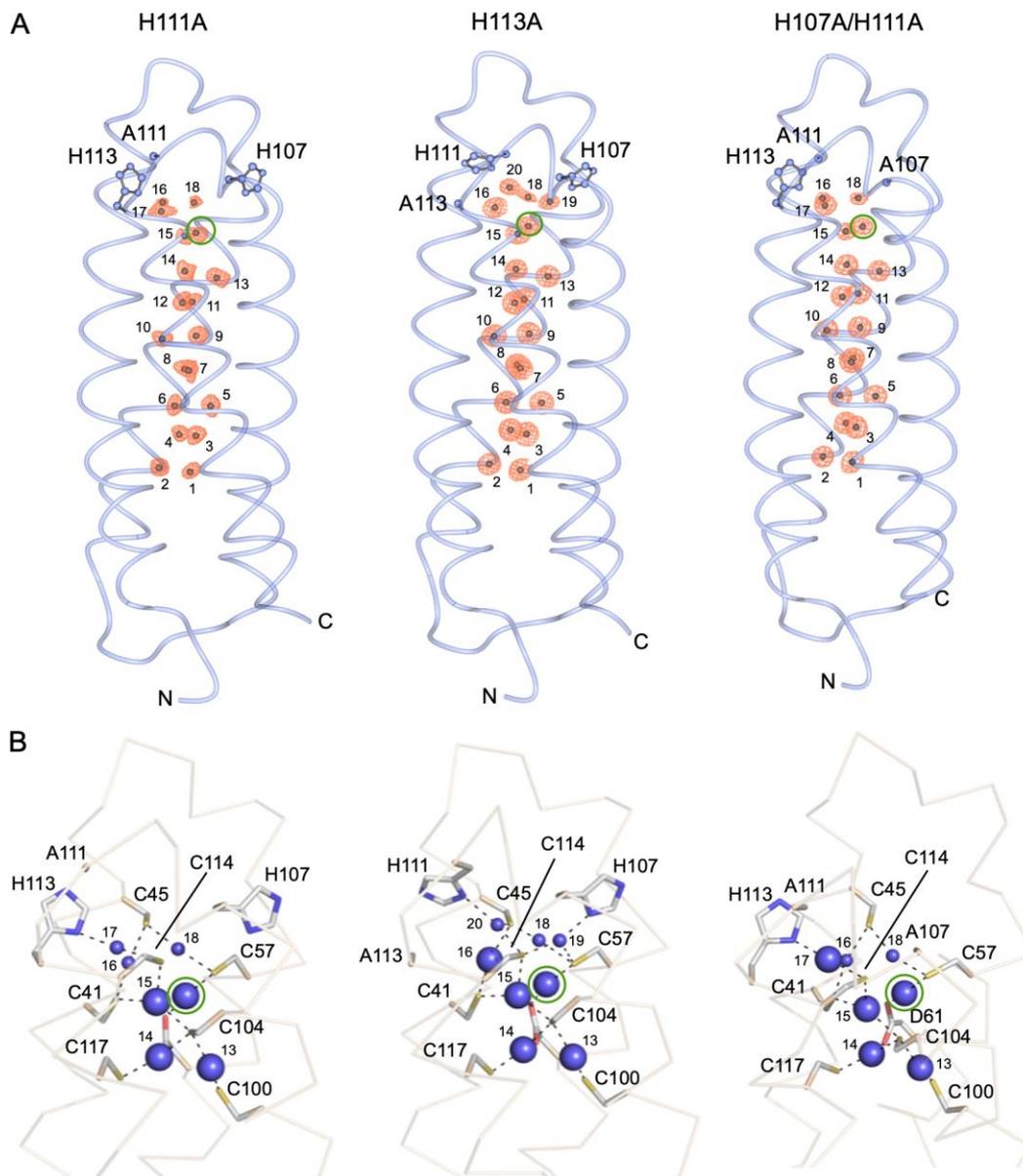
Figure 4.6 *- X-ray structures of the His variants of SlCsp3 fully Cu(I)-loaded. A) Worm representations with the His and Ala side chains of each variant shown in ball-and-stick representation and the N and C-terminus labelled. The anomalous electron density for the Cu(I) ions is shown in orange mesh and contoured at 5σ. B) Close-up of the coordination chemistry of the Cu(I) ions (blue spheres) in the outer core with coordinate bonds indicated by dashed lines. The smaller spheres indicate partial occupancy, based on the anomalous electron-density and the green circles indicate the location of non-cognate sites not present in the fully Cu(I)-loaded structure.*

4.3.6 Cu(I) arrangement in SlCsp3 Cys variants

To further understand Cu(I) loading into *Sl*Csp3, the C41S and C57S variants, were also created to assess whether these Cys variants would impact Cu(I) arrangements further into the inner core of *Sl*Csp3. As observed with the His variants, anomalous electron-density maps were produced for the two Cys variant structures (Fig. 4.7). Similar Cu(I) ion arrangements were observed in the Cys variants whereby the inner core of Cu(I) ions 1-14 maintained similar coordination chemistry as WT *Sl*Csp3. The outer core showed differences in Cu(I) ion arrangement (Fig 4.7A). For both Cys variants, one Cu(I) was absent compared to the fully loaded Cu(I) WT SlCsp3 structure thus giving a total of 19 Cu(I) ions (Fig. 4.7). The outer core Cu(I) arrangement in C41S variant shows Cu17 to be absent (Fig. 4.7A, green circle). Cu16 seems to take the place of Cu17 by coordinating with the $N^{\delta 1}$ of His113 giving a bond length of 2.3 Å (Fig. 4.7B). In addition, Cu15 is observed to coordinate with S^{γ}(Cys) of C57 in the C41S variant structure giving a bond length of 2.3 Å and also newly coordinates with S^{γ}(Cys) of C104, giving a bond length of 2.0 Å. Whereas in the WT Cu(I)-*Sl*Csp3 structure, this Cu15 is coordinated to S^{γ}(Cys) of C41.

All Cu(I) ions surrounding the C41S site maintain full occupancy and adjust to the loss of Cys41 by positioning themselves slightly closer to the adjacent Cu(I) binding residues (Fig. 4.7B). However, the positioning of these outer core Cu(I) ions does not differ greatly compared to the WT Cu(I)-*Sl*Csp3 structure. The structure of C57S variant demonstrates similar characteristics. Again, the absence of one Cu(I) is highlighted in the C57S structure as

shown in Fig. 4.7A and B (red and green circles) thus giving a total of 19 Cu(I) ions. The inner core (Cu(I) ions 1-14) maintain same positioning as WT *Sl*CSp3. The outer core displays similar patterns in Cu(I) positioning as C41S variant. The Cu(I) ion missing from this structure is Cu18 (Fig. 4.7B) and slight re-positioning of outer core Cu(I) ions is observed to compensate for the loss of C57 binding site. But again, this repositioning of Cu(I) ions is only slight and does not differ greatly to the WT Cu(I)-*Sl*Csp3 structure thus all Cu(I) ions still maintain coordination with their corresponding ligands.

Figure 4.7 *- X-ray structures of the Cys variants of SlCsp3 fully Cu(I)-loaded. A) Worm representations with the His side chains of each variant shown in ball-and-stick representation and the N and C-terminus labelled. The anomalous electron density for the Cu(I) ions is shown in orange mesh and contoured at 5σ with a red circle and green sphere showing the absence of one Cu(I) ion. B) Close-up of the coordination chemistry of the Cu(I) ions (blue spheres) in the outer core with coordinate bonds indicated by dashed lines. Based on the anomalous electron-density and the green circles indicate the location of Cu(I) ions not present in the fully Cu(I)-loaded structure.*

4.4   Discussion

At low Cu(I) to protein ratios, Cu(I) sites 15*, 17 and 18* are loaded with variable occupancy and a trinuclear $[Cu_3(\mu_2\text{-S-Cys})_2 (\text{S-Cys})_2(N^{\delta 1}\text{-His})]$ and a tetranuclear $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})]$ cluster, with the fourth Cu(I) occupying a non-cognate site (a binding site that is not occupied in the fully Cu(I)-loaded *Sl*Csp3) is visualised in the outer core (Fig. 4.3A). Cu(I) sites 19 and 20 remain empty but they must participate in Cu(I)-loading. It is notable from the structural studies of partially loaded *Sl*Csp3 that Cu(I) occupancy at sites 19 and 20 is only observed in the fully Cu(I)-loaded structure, and in the case of the His variants, sites 19 and 20 are only occupied when His107 and His111 are together present (Fig. 4.6B). Thus, Cu(I) sites 19 and 20 may be considered as transient loading sites or entrance sites to the outer core. Dennison and co-workers using X-crystallography also visualised an initial $[Cu_4(\mu_2\text{-S-Cys})_4]$ cluster forming, but in the inner core of *Mt*Csp3 under low Cu(I) loading conditions (76). However, in *Sl*Csp3, the formation of $[Cu_4(\mu_2\text{-S-Cys})_4]$ cluster is observed not in the inner core but in the outer core.

At low Cu(I) loadings (five Cu(I) equivalents), binding sites 1-14 remain empty, implying an energetic barrier to the distribution of Cu(I) beyond the initial tetranuclear cluster formed in the His entrance must exist. Thus, the $[Cu_4(\mu_2\text{-S-Cys})_4]$ cluster must be thermodynamically more favoured than occupancy of the sites available in the inner core, at least under low Cu(I) to protein ratios. On addition of 10 Cu(I)-equivalents, sites within the inner core become occupied, with it noted that variability in site occupancy between different chains of the homotetramer exist, implying flexibility and fluxionality of Cu(I) in the protein during loading. Furthermore, sites 1, 2 and 4 display similar traits to sites 19 and 20 in that occupancy is only observed when Ccsp becomes fully Cu(I) loaded (Figs. 4.4 and 4.5).

The Cys variant structures upon full Cu(I)-loading maintain an undisturbed inner core which was identical to the WT Cu(I) *Sl*Csp3 (Fig. 4.7). The outer core however did show slight differences in Cu(I) ion arrangement. In particular, Cu15 seemed to adjust in the absence of either Cys41 or Cys57 by coordinating with either one of these Cys ligands that was still present. In each structure, there was an absence of one Cu(I) ion, in the C41S structure, the absent Cu(I) ion was Cu17 whereby in the C57S structure, the absent Cu(I) ion was Cu18. The overall impact that changing Cys41 and Cys57 had on the Cu(I) ion arrangement of the outer core was minimal. The remaining outer core Cu(I) ions showed only small variances in positioning but maintained analogous coordination symmetry as the WT Cu(I)-*Sl*Csp3 structure.

The combination of the Cu(I)-loaded variant structures and partial loaded structures of *Sl*Csp3 will aid in understanding the kinetics of Cu(I)-loading to *Sl*Csp3. Chapter 5 describes

stopped flow kinetic studies, which together with the structures presented in this chapter, help define the mechanism of Cu(I) loading, with particular focus of the three His variants (H111A, H113A and H107A/H111A).

Chapter Five

Studying the kinetics of Cu(I) loading and Histidine coordination in *Sl*Csp3

5.1 Introduction

To circumvent the potentially toxic effects of aqueous Cu(I) ions in the bacterial cytoplasm, Cu(I) is moved between sites in a defined and safe manner. The accepted view of cellular Cu(I)-trafficking *i.e.* transfer of Cu(I) from a donor to an acceptor, is that it comprises of a ligand-substitution process at the inorganic centre, a fundamental process of inorganic reaction mechanisms (171). In this manner, a Cu(I)-bound donor (protein or low molecular weight ligand) associates with an acceptor, enabling intrusion into the donor coordination sphere of a ligand from the acceptor, ensuring that Cu(I) remains at all time coordinated between donor and acceptor and facilitating rapid Cu(I) exchange (111, 172-176). Therefore, for Csp3 members to act as a cytosolic Cu(I) store, cuprous ions will unquestionably need to be delivered by a donor. The nature of the donor that acts to deliver Cu(I) to Csp3 members *in vivo* is unknown, but as shown in Chapter 3, a Cu(I)-chaperone protein, CopZ, that coordinates a solvent exposed Cu(I) ion through thiolate coordination in a CXXC motif can readily transfer Cu(I) *in vitro* to *Sl*Csp3 (19).

Recent insights into Cu(I)-loading of *Mt*Csp3 have been obtained through X-ray crystallography studies (76). Structures determined at various Cu(I) to protein ratios reveal the existence of initial tetranuclear Cu(I)-thiolate clusters, [$Cu_4(\mu_2\text{-S-Cys})_4$], located in the Cys core (Cu sites 3 to 14) of the four helix-bundle (76). As more Cu(I) is loaded the tetranuclear clusters considered as 'intermediates' evolve into the final Cu(I) coordinated states (76). Thus, the formation of tetranuclear clusters is considered a driving force for acquisition and safe initial storage of Cu(I) by Csp3 members (76, 169).

This Chapter investigates stopped-flow kinetics of Cu(I) loading to WT *Sl*Csp3 and His variants (170). This work complements the structural studies described in Chapter 4, where the formation of various intermediate clusters including a tetranuclear [$Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})$] cluster in the outer core is discussed. These data reveal that at low Cu(I) loadings, polynuclear Cu(I) clusters form exclusively in the His entrance of the four helix-bundle. As more Cu(I) ions are loaded, Cu(I) sites become occupied to varying extents in the Cys core. Kinetic studies using the Cu(I) bicinchoninic acid complex ([$Cu(BCA)_2]^{3-}$) as a donor, reveals rapid uptake by *Sl*Csp3 of two Cu(I) ions within the first few seconds of the reaction time course, followed by additional slower phases. The role of the His residues lining one end of the four helix-bundle in Cu(I) loading have been determined, offering the first experimental kinetic evidence that Cu(I) loads at the His end. From these data, a model of Cu(I) binding to *Sl*Csp3 is proposed and discussed. In addition, data obtained from split chamber tandem cuvette studies is also presented to further demonstrate the Cu(I) acceptor/donor interactions between CopZ chaperone, CopZ3079 and WT *Sl*Csp3.

5.2   Methods

5.2.1   *Preparation of proteins, Cu(I) solutions and complexes*

Apo-*Sl*Csp3 proteins were exchanged into 10 mM MOPS pH 7.5, 150 mM NaCl. For Cu(I)-titrations and stopped-flow kinetics samples were prepared together with CuCl solutions in an anaerobic chamber (DW Scientific $[O_2]$ < 2 ppm). Solid CuCl (Sigma-Aldrich) was dissolved in 10 mM HCl and 500 mM NaCl and diluted with 10 mM MOPS pH 7.5, 150 mM NaCl. The Cu(I) concentration was determined spectrophotometrically using a Cary 60 UV-visible spectrophotometer (Varian) thermostatted at 20 ^{o}C through step-wise addition of the stock CuCl solution into a known concentration of the Cu(I) specific bidentate chelator bicinchoninic acid (BCA; Sigma-Aldrich). Formation of the $[Cu(BCA)_2]^{3-}$ complex was monitored using absorption spectroscopy by following the increase in absorbance at 562 nm on addition of Cu(I) and the concentration determined using an (Ɛ) = 7,900 M^{-1} cm^{-1} (147).

5.2.2   *Stopped-flow absorption spectroscopy*

An Applied Photophysics (Leatherhead, UK) stopped-flow spectrophotometer operating in absorbance mode using either a photomultiplier capture system or diode array and thermostatted to 20 ^{o}C was employed to monitor the kinetics of Cu(I)-loading to *Sl*Csp3 and the His variants. Anaerobic buffers were prepared by repeated exposure to vacuum followed by equilibration with oxygen free argon. Buffers were taken into glass syringes equipped with coupling tubes allowing dilution of the anaerobic Cu(I) solutions without exposure to oxygen. Protein solutions were prepared by similar cycles of gentle degassing and equilibration with oxygen free argon. The stopped-flow apparatus was washed through with anaerobic buffer prior to introduction of the reactants (protein and Cu(I)) under study. This procedure permits reactions to be studied at oxygen concentrations of 2 µM or below.   The appropriate extinction coefficient (accounting for the slit-width used in the stopped-flow experiments and wavelength discrimination in the diode array) for bleaching the $[Cu(BCA)_2]^{3-}$ complex on Cu(I) removal was determined by mixing a known concentration of the $[Cu(BCA)_2]^{3-}$ complex (200 µM) with an excess of protein and monitoring full bleaching of the absorption band at 562 nm using the diode array. The value determined for $Ɛ_{562nm}$ = 7,200 M^{-1} cm^{-1}. This value constitutes ~ 90 % of the literature value (147).

5.2.3   *UV-visible spectroscopy Cu(I) transfer studies*

Apo-CopZ3079 protein (3 mM) was prepared in an anaerobic chamber (DW Scientific $[O_2]$ < 2 ppm) in 10 mM MOPS pH 7.5, 150 mM NaCl and 2 mM DTT in a total volume of 2 ml and incubated over-night at room temperature. This CopZ3079 was applied twice to a 25 ml desalting PD10 column equilibrated in the experimental buffer to remove DTT. If required, the CopZ3079 was loaded with an excess of Cu(I) (2-5 molar equivalents) then applied to a

PD10 column to remove any unbound Cu(I). Apo-*Sl*Csp3 protein was exchanged into the experimental buffer in the anaerobic chamber. The proteins were then added in equal volume to each compartment of a glass split chamber tandem cuvette (0.5 – 1 ml). A baseline scan measurement scan was taken using a Cary 60 UV-visible spectrophotometer (Varian) thermostatted at 20 °C before mixing of the two proteins. Then both proteins were manually mixed together and a measurement scan was taken in the first ~ 10 seconds after mixing. Subsequent scan measurements were taken at set time intervals to observe for possible Cu(I) transfer.

5.3   Results

Addition of Cu(I) to the apo-state of the Cys variants, single and double His-variants under anaerobic conditions, led to the appearance of absorbance bands in the UV-spectrum that have previously been attributed in the WT protein to arise from (Cys)Sγ→Cu(I) ligand to metal charge transfer (LMCT) bands (1, 19). For all variants the absorbance bands in the UV-region of the spectrum increase concomitantly with the Cu(I):*Sl*Csp3 ratio (Fig. 5.1A and Appendix 3). A saturation point coinciding with a stoichiometry of ~18-20 Cu(I) ions bound per protomer (Fig. 5.1B) was observed as also noted for WT *Sl*Csp3 and thus, the Cys variants, single and double His-variants do not prevent Cu(I) loading.
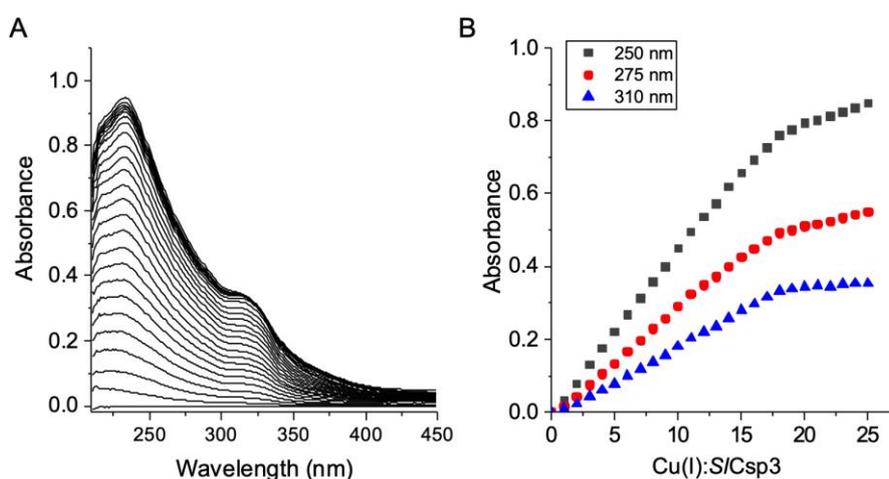
Figure 5.1 *- Cu(I) titration to the SlCsp3 H107A/H111A variant. A) UV-vis difference spectrum upon titration of a stock solution of CuCl to 5.6 µM of the protein revealing the appearance of (Cys)Sy→Cu(I) LMCT bands. B) Plots of absorbance versus the Cu(I):SlCsp3 concentration ratio at selected wavelengths taken from (A). A break point in the absorbance is reached at ~18-20 Cu(I) equivalents. Experiments were performed at 20 °C in 10 mM MOPS pH 7.5, 150 mM NaCl.*

5.3.1   Aqueous Cu(I) can rapidly fill SlCsp3 binding sites

On rapidly mixing Cu(I) with *Sl*Csp3 at sub- to super-stoichiometries with respect to the Cu(I) binding sites within *Sl*Csp3 under anaerobic conditions, optical transitions in the UV region of the absorption spectrum were observed, consistent with previously reported static

titrations (19). The reaction time courses for the transition at 280 nm are shown in Fig. 5.2A. A rapid increase in absorbance is observed within the first 2 seconds followed by slower processes (Figs. 5.2A and B). The amplitudes of the fast processes at 280 and 310 nm (0-2 seconds of the reaction) show distinct dependences on the Cu(I) concentration (Fig. 5.2C), consistent with the titration of *Sl*Csp3 with Cu(I), *i.e.* below stoichiometric Cu(I) concentrations the amplitude increases linearly, indicating high affinity binding. Thereafter, (at super-stoichiometries) the amplitude plateaus, as expected for saturation of all available sites and the intersection of these two titration phases indicates a stoichiometry of Cu(I) binding of ~90 µM Cu(I) (Fig. 5.2C), consistent with 20 Cu(I) sites per protomer (19). Therefore, from the stopped-flow data, *Sl*Csp3 becomes fully Cu(I) loaded within 2 seconds.
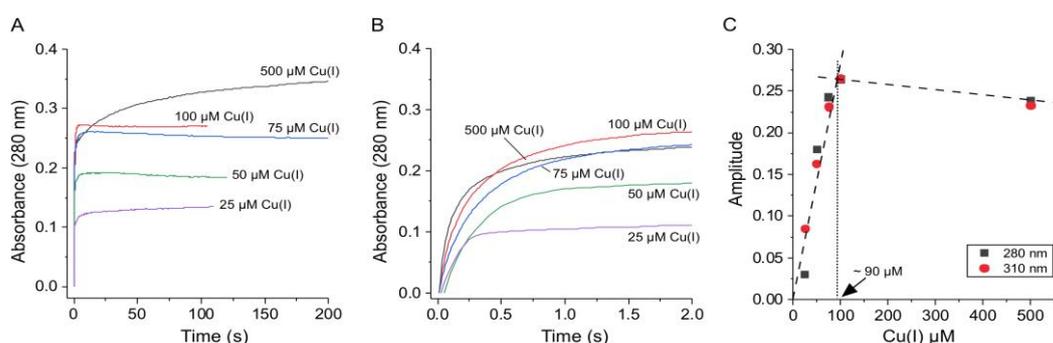
Figure 5.2 - *Kinetics of aqueous Cu(I) loading to SlCsp3. A) and B) Stopped-flow reaction time-courses monitored at 280 nm on mixing increasing concentrations of CuCl under anaerobic conditions. A fast phase between 0-2 seconds is observed (B), followed by slower phases over timescales up to 200 s (A). C) The amplitudes of the fast phase (between 0-2 seconds) observed at 280 nm and 310 nm plotted against increasing concentrations of Cu(I). The intersection of the two dashed lines is indicated revealing a saturating Cu(I) concentration of ~ 90 µM. Experiments were performed at 20 ᵒC in 10 mM MOPS pH 7.5, 150 mM NaCl with 4.5 µM of SlCsp3 after mixing. Solutions of Cu(I) at known concentration were obtained from dilution of a 1 mM stock CuCl solution under anaerobic conditions.*

The rates of the loading processes seen in the time courses in Fig. 5.2A and B, display some Cu(I) concentration dependence, but do not conform, when analysed as sums of one or two exponentials, to second-order rate processes. This is to be expected given the complex physical chemistry of Cu(I) loading at a single site and then transferring through the protomer. For full loading to occur within 2 seconds, the individual binding sites within the protomer although having a high intrinsic affinity for Cu(I), are able to pass the Cu(I) between sites suggesting an internal ligand-exchange mechanism is operating through the protomer. Thus, these data imply that the half-life for Cu(I) dissociation from any site within the binding tube is << 2 seconds. The slower phases seen in Fig. 5.2A could not be assigned, but were variable in rate and amplitude, and may result from either non-specific binding or metal-induced protein-protein interactions and are not further discussed. In addition, the rate constants associated with every time course reported in this chapter have not been included.

These values were often varying and inconsistent during plotting and fitting for further analysis. This could be due to the complex mechanisms potentially occurring in Cu(I) transfer and binding with *Sl*Csp3 thus are not further mentioned.

5.3.2 Cu(I) is loaded to SlCsp3 from a donor in multiple phases

Loading of aqueous Cu(I) to *Sl*Csp3 is unlikely to occur *in vivo*. Trafficking of Cu(I) within the bacterial cytosol involves ligand-exchange between the metal containing donor and an acceptor (171), ensuring that the Cu(I) ion remains coordinated at all times. To investigate such a process, the Cu(I) nitrogen donor BCA has been used to monitor the kinetics of Cu(I)
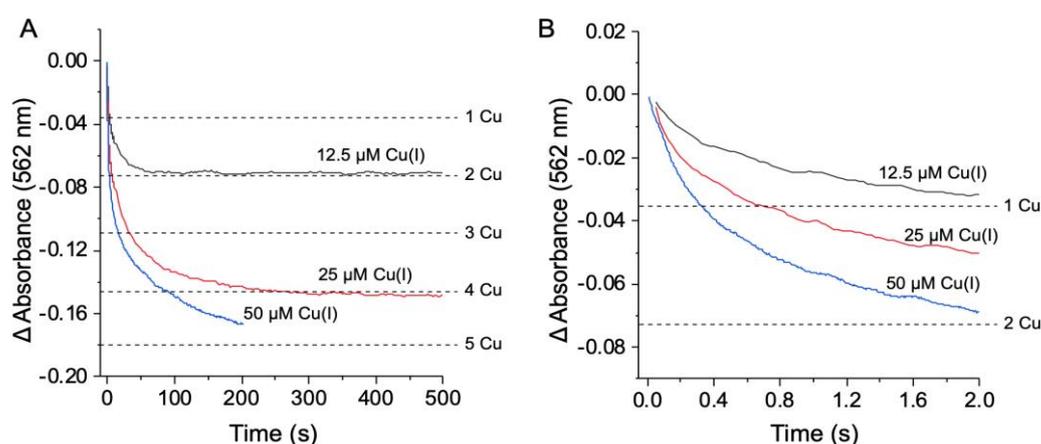
Figure 5.3 - Kinetics of Cu(I) loading to SlCsp3 from the $[Cu(BCA)_2]^{3-}$ complex. A) and B) stopped-flow reaction time-courses monitored at 562 nm on mixing Sl*Csp3 (5 µM) with increasing concentrations of Cu(I) chelated in the $[Cu(BCA)_2]^{3-}$ complex revealing an initial fast phase (B) followed by slower phases (A). Dashed lines indicate the expected absorbance changes for removal of Cu(I) equivalents from the $[Cu(BCA)_2]^{3-}$. Up to 50 µM $[Cu(BCA)_2]^{3-}$ complex the kinetics could be monitored satisfactorily for the first 200 seconds, thereafter, at longer times and at higher $[Cu(BCA)_2]^{3-}$ concentrations, interactions between $[Cu(BCA)_2]^{3-}$ complexes themselves perturbed the spectra and made it impossible to analyse confidently the data in terms of Cu(I) transfer to SlCsp3. Experiments were performed at 20 °C in 10 mM MOPS pH 7.5, 150 mM NaCl.

loading to *Sl*Csp3. In Chapter 3 it was shown that *Sl*Csp3 can remove Cu(I) from the $[Cu(BCA)_2]^{3-}$ complex (19). On rapid mixing of *Sl*Csp3 with $[Cu(BCA)_2]^{3-}$ in the stopped-flow spectrophotometer, bleaching of the complex absorbance band centred at 562 nm was observed. The kinetics of this process are shown Figs. 5.3A & B. A rapid decrease in absorbance at 562 nm in Fig. 5.3B, illustrates Cu(I) transfer occurs within the first 2 seconds of the time monitored. Given, the high affinity of BCA for Cu(I) (implying a vanishingly small dissociation rate constant), the transfer cannot proceed with aqueous Cu(I) as an intermediate and thus must proceed via complex formation between *Sl*Csp3 and $[Cu(BCA)_2]^{3-}$. The amplitude of the absorbance change is seen to be $[Cu(BCA)_2]^{3-}$ concentration dependent (Fig. 5.3A). Using an $\varepsilon_{562\ nm}$ = 7,200 M^{-1} cm^{-1} for the $[Cu(BCA)_2]^{3-}$ complex as determined for monitoring in the stopped-flow spectrophotometer (see 5.2.2 Methods), it is indicated on Fig. 5.3A the expected absorbance changes for filling up to four sites in *Sl*Csp3 (5 µM) with

Cu(I). Within the first 2 seconds, between 1 and 2 Cu(I) ions are delivered depending on the concentration of the [Cu(BCA)$_2$]$^{3-}$ complex (Fig. 5.3B). Over a longer time-period (Fig. 5.3A), 80-85 % of the Cu(I) is removed from the BCA complex, e.g. 10 µM Cu(I) of the 12.5 µM Cu(I) available. The discrepancy of the total Cu(I) available and the Cu(I)-loaded may arise either from a much slower delivery over a longer time-period that is not observed or in the experimental error of solution concentrations. Analysis of the kinetic processes showed them to comprise of at least three-exponential phases. This may be expected given that Cu(I) ions must be delivered from the [Cu(BCA)$_2$]$^{3-}$ complex one at a time and thus involve i) complex formation, ii) Cu(I) transfer and iii) dissociation of the free BCA. Furthermore, sequential loading of Cu(I) implies Cu passing from site-to-site within the *Sl*Csp3 protomer as inferred from aqueous Cu(I)-loading. Nevertheless, it is observed that the most rapid phase is most easily interpreted as transfer of a Cu(I) ion to a first coordination site in *Sl*Csp3.

5.3.3   His107 is important for initial Cu(I) entry

To probe further the mechanism of the initial Cu(I)-loading to *Sl*Csp3, stopped-flow kinetics of the His-variants with the [Cu(BCA)$_2$]$^{3-}$ complex as the donor was carried out. Figure 5.4 compares the loading of Cu(I) to the WT *Sl*Csp3 and the His variants with the expected optical density change for filling one site with Cu(I) indicated. Over the first 2 seconds, rapid transfer is observed (Fig. 5.4A) and it is apparent that the H111A and H113A variants are essentially indistinguishable from the WT *Sl*Csp3, whereas the H107A and the H107A/H111A double variant clearly show that entry of Cu(I) is perturbed (Fig. 5.4A). This implies that H107A has a major role in initial Cu(I) transfer. Over a longer time-period, during which ~ 80% of the available Cu(I) can be delivered to the WT *Sl*Csp3, further effects of the His-variants may be
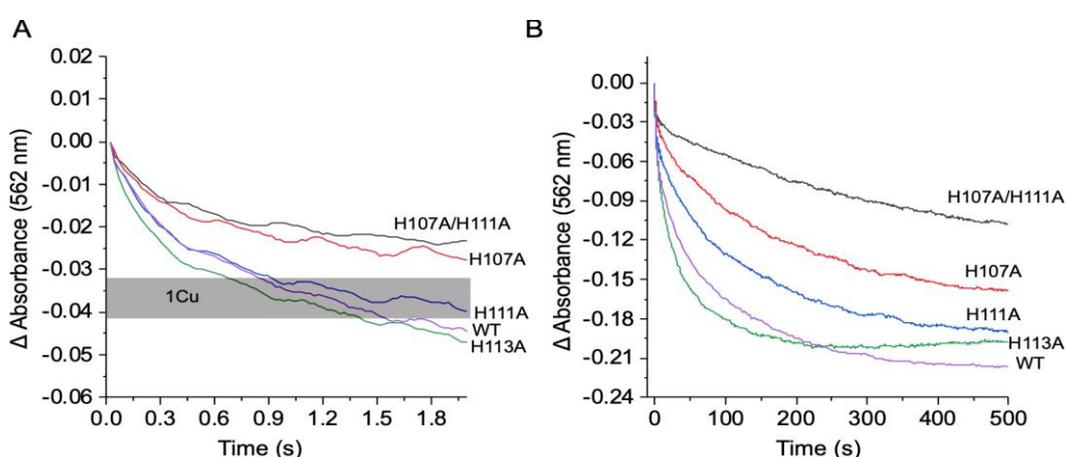
Figure 5.4 *- Kinetics of Cu(I) loading to the SlCsp3 His variants from the [Cu(BCA)$_2$]$^{3-}$ complex. A) and B) stopped-flow reaction time-courses monitored at 562 nm on mixing SlCsp3 (4.5 µM) and the His variants (5.5-5.8 µM) with 50 µM [Cu(BCA)$_2$]$^{3-}$ complex revealing an initial fast phase A) followed by slower phases B). Shaded area in (A) indicates the expected absorbance changes for removal of one Cu(I) equivalents from the [Cu(BCA)$_2$]$^{3-}$ complex based on the variation of protein concentrations used. Experiments were performed at 20 ^{o}C in 10 mM MOPS pH 7.5, 150 mM NaCl.*

discerned (Fig. 5.4B). The H107A variant identified as important in initial entry, slows subsequent loading (Fig. 5.4B), whereas the H111A variant, while not effecting the initial entry, does further slow Cu(I) entry (Fig. 5.4B). Of note, the double variant which is similar to H107A in affecting the initial entry, has a significant effect on subsequent loading (Fig. 5.4B). Examination of the structure in Fig. 4.5 and these data, indicate H107 and H111 are important ligands to the first two Cu(I) sites (19 and 20). The H113A variant affects neither the initial entry nor subsequent loading of up to at least four Cu(I) ions per protomer, suggesting that the removal of this coordinating His does not impair binding to the site or impair transfer from the initial entry sites to other available sites.

5.3.4   Cu(I) transfer between WT *Sl*Csp3 and CopZ3079

The studies of Cu(I) transfer between cytosolic Cu(I) chaperone, CopZ3079 and WT *Sl*Csp3 have been previously described in Chapter 3 (19). These *in vitro* experiments involving gel filtration revealed a clear distinction between the elution profiles of Cu(I)-CopZ dimer species (Cu^+–$(CopZ)_2$) and the apo CopZ monomer species. This allowed visual clarification of Cu(I) transfer between these two proteins; Cu^+–$(CopZ)_2$ donating Cu(I) to apo *Sl*Csp3 thus reforming the monomer CopZ species (19). This outcome indicated the possibility of these two cytosolic proteins interacting *in vivo* through transient Cu(I) exchange mechanisms (163). To further investigate these interactions and support the previous work, UV-vis spectrophotometry was carried out to observe Cu(I) transfer using a split chamber tandem cuvette. The results of these experiments are shown in Fig. 5.5. Upon mixing Cu^+–$(CopZ)_2$ with apo WT *Sl*Csp3, LMCT bands can be observed at wavelengths 310 nm – 360 nm (Fig. 5.5A). Though these LMCT bands are not shown in their entirety, as absorbance at wavelengths lower than 300 nm were not possible due to using a glass cuvette in these studies. Regardless, this outcome clearly shows Cu(I) transfer from donor CopZ to acceptor protein *Sl*Csp3. The time recorded in Fig. 5.5B showing a plateau in Cu(I) transfer was reached after 10-15 minutes. Fig. 5.5C displays the same experiment but involved mixing of WT Cu(I)-*Sl*Csp3 with a ~28-fold excess of apo CopZ. It is seen in the UV-vis spectrum (Fig. 5.5C) that a drastic increase in the LMCT bands demonstrates the CopZ removing Cu(I) ions from Cu(I)-*Sl*Csp3. The plots shown in Fig. 5.5D further demonstrate this instant and rapid Cu(I) transfer as a plateau was reached after mixing both proteins (~10 seconds). This outcome is a direct result of the presence of excess CopZ chaperone that outcompetes *Sl*Csp3 for Cu(I) due to this protein concentration difference (Cu(I)-*Sl*Csp3 (6.28 µM) and apo CopZ (176 µM).
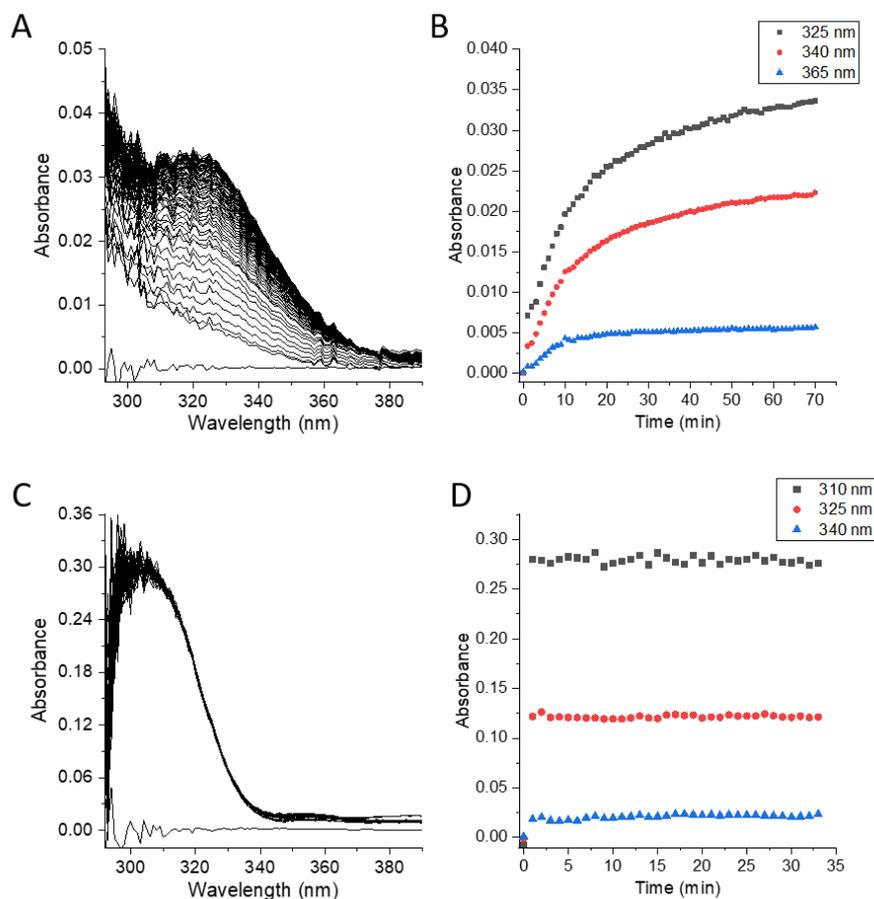
Figure 5.5 *– Cu(I) transfer between CopZ3079 and SlCsp3. (A) UV-vis spectrum of split cuvette measurement of Cu(I) loaded CopZ3079 (88.13 µM) and apo SlCsp3 (7.6 µM). A measurement of the baseline was taken which is before mixing these two proteins. After mixing, measurements were taken every 1 minute for 70 minutes (B) Plots of absorbance versus time (minutes) at selected wavelengths taken from (A). A plateau in Cu(I) transfer between the two proteins was reached after 10-15 minutes. (C) UV-vis spectrum of split cuvette measurement of Cu(I) loaded SlCsp3 (6.28 µM) and apo CopZ3079 (176 µM). A measurement of the baseline was taken which is before mixing these two proteins. After mixing, measurements were taken every 1 minute for 33 minutes. (D) Plots of absorbance versus time (minutes) at selected wavelengths taken from (C). An almost instant plateau was reached in Cu(I) transfer after mixing (~10 seconds). Experiments were performed at 20 °C in 10 mM MOPS pH 7.5, 150 mM NaCl.*

5.4   Discussion

The capacity to bind multiple Cu(I) ions through predominately thiolate coordination chemistries to protect bacteria against potential toxicity is an inherent feature of the recently discovered Csp3 members (169). Understanding the kinetic and thermodynamic intricacies associated with Cu(I) loading to these proteins is particularly challenging considering the number of potential binding sites (cognate and non-cognate) that Cu(I) thiolate chemistry can impose. The results from these kinetic studies are rationalised below and a model of Cu(I) binding in terms of the relative energy of binding to the distinct Cu(I) sites in *Sl*Csp3 has been constructed.

An initial binding complex is not directly observed when using either aqueous Cu(I) or the $[Cu(BCA)_2]^{3-}$ complex as a donor to load *Sl*Csp3 as implied by the lack of second-order binding kinetics (*i.e.* no linear dependence of k_{obs} on Cu(I) concentration). Such a complex must be present in the initial binding step and is a situation reminiscent of Cu(I) loading to other proteins which have been studied using stopped-flow spectroscopy, whereby rapid metal ion binding may be inferred but not directly observed (112, 118, 177). It seems reasonable to suggest that this complex involves initial His binding to Cu(I), which is expected to be optically silent as proposed earlier (112).

A model to describe Cu(I) binding in terms of the relative energy of binding to the distinct Cu(I) sites is depicted in Fig. 5.6 and illustrates the relative stability of the complexes
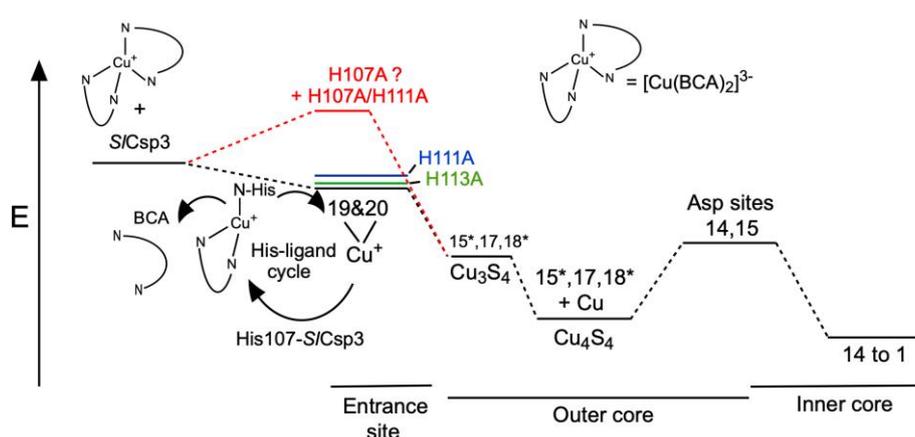
Figure 5.6 *- Scheme to illustrate Cu(I)-loading from the $[Cu(BCA)_2]^{3-}$ complex to the Cu(I) sites in SlCsp3. The relative stabilities of the sites are depicted relative to an arbitrary energy scale. The extent of loading of Cu(I) is in accordance with the thermodynamic stabilities of the sites, whilst the kinetics of loading are controlled by ligand-exchange mechanisms between adjacent sites. See main text for details.*

of Cu(I) in the sites described. When Cu(I) transfer is from the $[Cu(BCA)_2]^{3-}$ complex, this must occur via a ligand-substitution mechanism to account for the relatively fast transfer (Fig. 5.3). As the tetrahedral $[Cu(BCA)_2]^{3-}$ complex is coordinatively saturated we suggest that on mixing with *Sl*Csp3 a heteroleptic complex forms, consisting of $[Cu(BCA)]^{-}$ and a nitrogen from a His residue with the net loss of a BCA ligand (Fig. 5.6). This His residue can be identified as His107 based on the initial phase of the kinetic time course (Fig. 5.4A). Following this step, an energetically favourable transfer to sites 19 and 20 will occur, which structurally would involve both His and Cys coordination (group III). Sites 19 and 20 are transiently filled and Cu(I) is transferred to sites 15*, 17 and 18* causing release of His107 from coordination and enabling for further Cu(I) capture from the donor (His-ligand cycle, Fig. 5.6). The trinuclear cluster that is visualised in one of the protomers at 5 Cu(I)-equivalents may therefore be considered as a higher energy cluster on the way to forming the more thermodynamically stable tetranuclear cluster (Figs. 5.6 and 4.3A).

Figure 5.4B shows that over a longer reaction time course, both His107 and His111 slow Cu(I) loading beyond sites 19 and 20 and their effects are additive. This may be accounted for by reference to the crystal structure (Fig. 4.6) that indicates both His residues are involved in binding and stabilising Cu(I) in sites 19 and 20. The removal of these residues may lead to a decrease in the Cu(I) loading to these sites and hence to a slowing of loading to the interior of the protein as this must occur from Cu(I) populated at sites 19 and 20. Furthermore, as His107 is important for the ligand-exchange mechanism of passing Cu(I) from BCA, it may be expected that this residue has a marked effect, because in order for up to four Cu(I) ions to load as illustrated in Fig. 5.4B, the $[Cu(BCA)_2]^-$-His-SlCsp3 complex must form, transfer Cu(I) and dissociate after each Cu(I) is donated. Thus, for four Cu(I) ions loaded, the His-ligand cycle must occur four times and as each loading requires His107 for efficient transfer its absence would have a significant effect on the loading and could account for the observed kinetics (Fig. 5.4B).

At low Cu(I) stoichiometry, Cu(I) transfer beyond site 15* is not observed. Thus, the formation of the $[Cu_4(\mu_2\text{-S-Cys})_4(N^{\delta 1}\text{-His})]$ cluster creates an unfavourable barrier to loading into the inner core (Fig. 5.6). A reason for this may be that to move Cu(I) into the inner core via the binding sites 15 and 14, interaction with Asp61 must occur (Asp sites Fig. 5.6). A hard ligand/soft metal interaction is less favoured and of lower affinity than the Cu(I) thiolate coordination dominating the tetranuclear cluster. However, as more Cu(I) is loaded into the outer core this barrier can be overcome, leading to transfer into the inner core becoming favourable. This accounts for the observation that at 10 Cu(I)-equivalents Cu(I) ions are present in both cores (Fig. 4.4).

The model depicted in Fig. 5.6, also gives insight into how SlCsp3 achieves very high affinity for Cu(I) while retaining rapid kinetic transfer. A mathematical analysis of a multi-site protein (up to 20 sites) is complex and requires solutions of at least quartic equations. However, without resorting to such complexity it is possible to see that the dissociation constant (K_d) for SlCsp3, defined as the free Cu(I) concentration in solution at equilibrium with the protein at half saturation, arises from the K_d of the initial Cu(I) capture site, which is likely to be of the order 10^{-5} M (typical for Cu(I) His interaction (112, 118)) divided by a function of the multiple of the equilibrium constants for transfer from the capture complex to sites 19 and 20, sites 15*, 17 and 18* and so on. Overall this multiple will be large and positive, resulting in the extremely small K_d value (10^{-17} M) measured for SlCsp3 (19).

The *in vitro* Cu(I) transfer studies between CopZ3079 and WT SlCsp3 presented in this chapter further support work presented in Chapter 3 (19). The analytical gel filtration

studies shown in Chapter 3 demonstrated Cu$^+$–(CopZ)$_2$ is able to transfer Cu to apo *Sl*Csp3 (19). These interactions are not uncommon for metalloproteins and normally involves a transient ligand exchange mechanism (163). Indeed, the split cuvette method yielded results that Cu(I) transfer is unidirectional thus Cu(I) is only transferred from Cu$^+$–(CopZ)$_2$ to apo *Sl*Csp3 (Fig. 5.5A & B). This exchange does not occur vice versa using equal protein concentrations in the split cuvette method as equilibrium does not favour this but further experiments were carried out using an ~28-fold excess of apo CopZ in the presence of Cu(I)-*Sl*Csp3 (Fig. 5.5C & D). This demonstrated an almost instant removal of Cu(I) ions from *Sl*Csp3 by apo CopZ after mixing (~10 seconds). The split cuvette experiments support *in vitro* that Cu(I) transfer does occur between these two proteins and under the experimental conditions in Fig. 5.5A & B, transfer reaches a plateau after 10-15 minutes.

Future work could involve testing the Cys and His *Sl*Csp3 variants for Cu(I) transfer with CopZ and further examining rates of Cu(I) transfer. For instance, Dennison and co-workers presented similar studies between a Csp3 from *Bacillus subtilis* and a CopZ from the same organism (2, 75). It was determined that Cu(I) removal from *Bs*Csp3 by *Bs*CopZ was a slow process whereby ~40% of Cu(I) was removed from *Bs*Csp3 by *Bs*CopZ in 64 hours (2, 75). This was also observed for Cu(I) removal from *Methylosinus trichosporium* OB3b *Mt*Csp3 by its physiological partner, Mbtin, which has a high affinity for Cu(I) (2, 6, 13). It was found in a study by Vita *et al*. that Mbtin removes all Cu(I) from *Mt*Csp3 in ~15 days (2). This highlights the thermodynamically unfavourable direction of Cu(I) transfer for these Csp3s by their physiological partners.

In conclusion, this study offers the first insight into the kinetics of Cu(I) loading to a non-methanotrophic Csp3 member and in particular highlights that the His residues at the hydrophilic mouth of the outer core are the entrance sites of Cu(I) loading. Furthermore, it illustrates that efficient kinetic transfer occurs from an organic Cu(I) nitrogen donor to *Sl*Csp3. *In vivo* potential Cu(I) donors to *Sl*Csp3 could involve Cu chaperone proteins that utilise digonal bis-cystientate coordination (111, 176, 178) and thus the kinetics of Cu(I) transfer to a Csp3 may differ from a thiol Cu(I) donor compared to a nitrogen donor. However, chalkophores, natural products produced by bacteria that chelate and transport copper (16, 62, 179, 180) predominately utilise nitrogen as a Cu(I) donor, as is the case here with the [Cu(BCA)$_2$]$^{3-}$ complex. Until recently chalkophores have been associated with methanotrophic bacteria (62). Interestingly, diisonitrile derivatives with nitrogen and/or oxygen as Cu(I) donor, with 1:2 (Cu(I):ligand) stoichiometry, have been discovered in *Streptomyces thioluteus* and identified to play a role in Cu uptake mechanisms (5). This poses

the question of whether similar compounds serve as donors to Csp3s, as opposed to thiol donors such as CopZ Cu chaperones. Thus, the kinetic studies of Cu(I) transfer to *Sl*Csp3 utilizing a nitrogen Cu(I) donor in the guise of the $[Cu(BCA)_2]^{3-}$ complex and the discovery of role for a His residue in facilitating the Cu(I) transfer bear significance on potential biological events. Structural studies further indicate that the driving force to sequester Cu(I) and prevent toxicity by Csp3 members is through the initial formation of tetranuclear Cu(I)-thiolate clusters (76, 169). Contrary to cluster formation in the inner core as revealed in *Mt*Csp3, a tetranuclear cluster forms in the outer core of *Sl*Csp3 and its presence affects the loading of the inner core at low Cu(I) loading.

Chapter Six

Developing an over-expression system for the integral membrane protein SLI_RS17250 (DUF4396) from *Streptomyces lividans*

6.1   Introduction

The expression and characterisation of membrane proteins continues to be one of the most challenging areas in protein structural biology (181, 182). This is due to their hydrophobic nature as they are found in the mosaic lipid bilayer of the cell with occasionally extra domains outside the bilayer (183). This makes it difficult to express these proteins with an added difficulty being that they are often unstable in solution without a suitable detergent present (183-185). Indeed, the structural biology of this class of proteins is highly sought after as they represent more than 60% of drug targets (186, 187). In addition, most membrane proteins are present in low amounts in membranes and the need to over-express them efficiently is
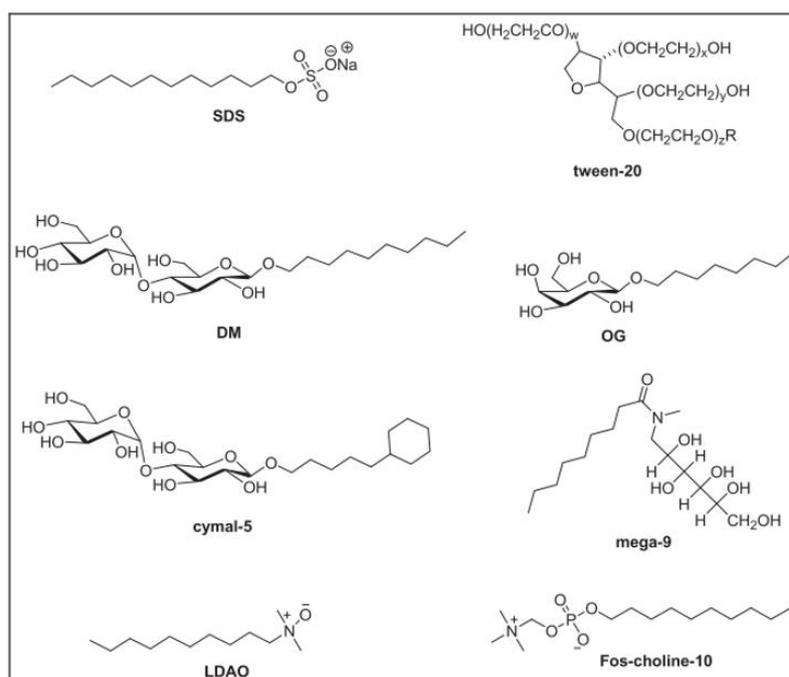
Figure 6.1 – *Chemical structures of detergents used in membrane protein solubilisation. Sodium dodecyl sulfate (SDS); polyoxyethylene sorbitan monolaurate (Tween20); n-decyl-β-ᴅ-maltoside (DM); n-octyl-β-ᴅ-glucoside (OG); 5-cyclohexyl-1-pentyl-β-ᴅ-maltoside (Cymal-5); nonanoyl-N-methylglucamide (Mega-9); lauryldimethylamine-N-oxide (LDAO); decylphosphocholine (Fos-choline-10). Image taken from (8).*

paramount. Furthermore, the issues that arise from expressing cloned constructs of membrane proteins in any system is protein aggregation in the cytoplasm therefore obtaining high amounts of protein is extremely challenging (185). The use of detergents in membrane protein purification is a pre-requisite as they comprise of a hydrophobic chain and a polar head group and are amphipathic overall (185), thus mimicking the natural membrane lipid environment. Some examples of commonly used detergents are shown in Fig. 6.1.

The detergent molecules in Fig. 6.1 have physicochemical properties that are similar to the phospholipids in the mosaic lipid bilayer of cell membranes. They form micelles in aqueous

solution and thus form around the membrane protein which maintains their stability and solubilises these proteins (185). However, despite the wide use of detergents, their effectiveness in maintaining membrane protein structure and activity has often proven to be inadequate. The structure of some cell membranes contains various other lipids and membrane proteins have adapted to exist in such environments whereby the membrane provides this lipidic complexity that ensures proper protein folding which leads to enhanced membrane protein activity (188). Due to this, detergents alone are insufficient in replicating this environment. The use of detergents also poses the fundamental question of which detergent to use? The lack of a single solution to solubilise any membrane protein leads to endless experiments of testing a great number of detergents which is time consuming (188). Thus, a new technology had been developed to replicate the physically diverse membrane bilayer (188-190) known as styrene maleic acid lipid particles (SMALPs) (188, 191, 192). This method involves directly removing proteins from the membrane into SMALPs (188). SMALPs maintains a structure of an outer layer of styrene maleic acid co-polymers (SMA) (188, 193) supporting a central lipid bilayer (188). This assembly is maintained by intercalating hydrophobic styrene groups between the acyl chains in the bilayer with solvent exposed maleic groups (188). This novel bilayer capsule allows to extract proteins by maintaining the lipid formation and physical characteristics of the protein's native membrane environment (188, 194). This method is an improvement of the detergent-based method but does have limitations. The SMALP nominal maximal diameter is ~15 nm (188) which requires that protein molecular mass cannot be more than ~400 kDa (188, 193).

Indeed, the use of SMALPs limits the types of membrane proteins that can be isolated. To overcome issues in heterologous membrane protein expression, it is also paramount to understand their behaviour *in vivo*. The functions of membrane proteins range from acting as ion channels, ATPases, transporters, electron carriers and more (20). Integral inner membrane proteins are either β-barrel or α-helical bundles (20). The α-helical membrane proteins can exist monomerically or as oligomers and possess membrane spanning helices that are either re-entrant, curved, straight or kinked (20). In prokaryotes, membrane protein expression consists of targeting the protein to the correct membrane location, topogenesis to properly integrate the membrane protein which could involve structural folding of loops or multiple subunits brought together to form an oligomeric membrane protein complex
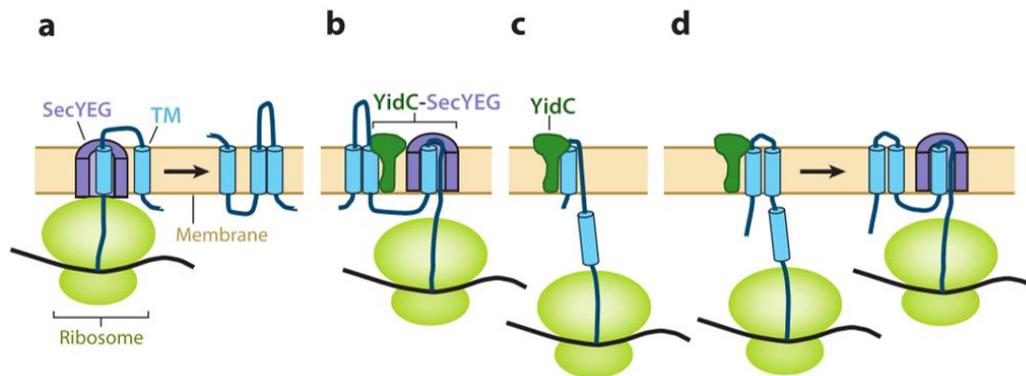
Figure 6.2 - *Membrane protein insertion diagram and the function of translocases. (a) Transmembrane (TM) fragments are partitioned into the lipid phase by SecYEG (b) The insertion and integration of TM subunits is mediated by SecYEG and YidC. TM segments are inserted into the Sec channel and before entering the lipid phase, these segments are transferred to YidC. TM segments are regulated by YidC whereby multiple fragments can dock onto YidC for assembly whereby they are released as a fully fabricated α-helix bundle. (c) YidC functions independently as an insertase. (d) SecYEG and YidC function in unison for membrane protein insertion. In this diagram, YidC insertase mediates the insertion of the N-terminal region of the protein whereas the insertion of the C-terminus is controlled by SecYEG. Diagram taken from (20).*

(20). This process is referred to as membrane protein biogenesis which has been widely studied in *E. coli* (195) and involves a series of complex steps (Fig. 6.2). The first step involves targeting the membrane protein which takes place before or during protein synthesis (20). A signal recognition particle (SRP) plays a role by interacting with the newly synthesised membrane protein as it appears from the ribosome and forming an SRP-ribosome-nascent chain complex (20). This complex is transported to a receptor, FtsY which is stationed on the membrane (20). The complex becomes disassembled, a process that is energetically supported by GTP hydrolysis by binding to FtsY and SRP and overall releases the membrane protein (20). The FtsY breaks away from GFP and SRP and interacts with the SecYEG translocation channel to transport the membrane protein (20). The next step involves

insertion of the membrane protein into the membrane which is aided by Sec translocase (20). This translocase protein consists of SecA which a peripheral membrane component and two membrane embedded complexes known as SecDFyajC and SecYEG (20). The role of each subunit is essential to the membrane protein integration process. SecA is a motor ATPase which translocate membrane proteins by moving their polypeptide chains through SecYEG which is a protein channel (20).

The topology and insertion of membrane proteins into the membrane bilayer is influenced by their amino acid composition. The possession of positively charged residues in the cytoplasmic loops of bacterial membrane proteins is a positive-inside characteristic (20). This configuration is found in systems such as the endoplasmic reticulum (20, 196), chloroplasts' thylakoid (20, 197) and mitochondrial membranes (20, 198). Overall, the positioning of the transmembrane region is dictated by the contiguous positively charged residues that are found in the cytoplasm – this is the positive-inside law (20).

6.1.1 A putative transmembrane protein from S. lividans

In conjunction with the up-regulation of the *slcsp3* gene under Cu stress in *S. lividans* the neighbouring gene, *SLI_RS17250*, was also found to be upregulated to a similar level and expresses a putative transmembrane protein, SLI_RS17250 which belongs to the functionally uncharacterised DUF4396 family. The amino acid sequence for SLI_RS17250 is shown in Fig. 6.3.

```
MDHSTHHSAP  EDDGPGHPHG  HGHLHGPGHP  HGHVHVHGTT  WATAMQATLH
CLTGCAIGEI  LGMVIGTALM  WGNVPTMVLA  IALAFVFGYS  LTLFAVVRAG

VSMKAAIKVA  LAADTVSIAV  MELVDNGIIA  LVPGAMEAHL  SDGLFWYALL
GGFAVAFVIT  TPVNKWMIGR  GKGHAVVHAY  H
```

Figure 6.3 – Full amino acid sequence of SLI_RS17250 derived from S. lividans.

This chapter is aimed at firstly creating *Escherichia coli* over-expression constructs to produce SLI_RS17250, test over-expression and assess whether SLI_RS17250 can be characterised *in vitro*. The function of SLI_RS17250 is entirely unknown, but this protein is part of a developing theory of a new Cu resistance system that also involves a translationally coupled Na^+/H^+ antiporter (SLI_RS17245) and a cytosolic copper storage protein (*Sl*Csp3) (see Chapters 3-5). Initially it was believed that the only Cu homeostatic mechanism present in *S. lividans* was CsoR/CopZ/ATPase system (17, 18, 108, 112). But the structural and biochemical data collected for *Sl*Csp3 as well as *in vivo* studies have suggested a second layer of Cu resistance when excess levels of Cu seem too overwhelming for the CsoR system to cope.

6.2   Methods and Materials

6.2.1   Cloning procedure for SLI_RS17250

The *SLI_RS17250* gene (568 bp) was amplified from *S. lividans* 1326 and ligated into plasmid

pTZ19R with NdeI-BamHI restriction sites. Three constructs were made to test

overexpression in *E. coli*. The first construct was prepared by ligating the *SLI_RS17250* gene

Table 6.1 - *Forward and reverse primers used to amplify SLI_RS17250 gene via PCR to insert into plasmids pET21a and pET26b. Restriction sites underlined.*

Plasmid and restriction sites	Forward primers	T_m (C)	Reverse primers	T_m (C)
pET21a (NdeI, XhoI)	5'-GAATT <u>CATATG</u> GACCACAGCACGCACCACTC-3'	61	5'-AATAA <u>CTCGAG</u> GTGGTAGGCGTGGACGACGG-3'	62
pET26b (BamHI, XhoI)	5'-GAATT <u>GGATCC</u>G ATGGACCACAGCACGCACCACTC-3'	64	5'-AATAT <u>CTCGAG</u> GTGGTAGGCGTGGACGACGG-3'	62

into a pET21a plasmid (Novagen) with the original restriction sites (NdeI-BamHI) to ensure

the absence of any purification tag. Primers for PCR were designed for the next two

constructs (Table 6.1). The second pET21a construct was designed with NdeI-XhoI restriction

sites, so that on insertion a C-terminal (His)$_6$-tag was present. The pET26b (Novagen)

construct was designed with BamHI-XhoI restriction sites, so that the construct would

possess a C-terminal (His)$_6$-tag. The PCR conditions and reagents used to make both

constructs are shown in Tables 6.2 and 6.3. The amplified products were ligated using T4 DNA

Ligase (Thermo Scientific), the plasmids were used to transform *Escherichia coli* XLI-Blue

competent cells and transformants selected for DNA sequencing following a mini-prep

(Thermo Scientific). Sequencing data are reported in Appendix 4.

Table 6.2 - *PCR cycles – bold text indicates these steps were repeated 35 times and other steps only once.*

Temperature (^{O}C)	Time (minutes)
95	3.0
95	**1.0**
60 (Annealing temperature)	**1.0**
72	**2.0**
72	7.5 (Final Extension)
4	Finish

Table 6.3 – *The reagents and volumes used to amplify SLI_RS17250 gene*

Reagent	Concentration	Volume (μl)
Plasmid DNA	N/A	2.5
Forward Primer	-	1.0
Reverse Primer	-	1.0
dNTP's	10 mM	2.5
10 x Buffer Pfu + MgSO$_4$	-	5.0
Sterile, deionized water	-	39.0
Pfu DNA polymerase	-	0.5
Total	-	50.0

6.2.2 Protein expression and purification

The *SLI_RS17250*_pET21a construct with a C-terminal (His)$_6$-tag was used to transform BL21-C43 *E. coli* cells for overexpression studies. Overnight cultures were made by selecting individual colonies from the transformation plate and inoculating 3 ml of LB (Luria Broth) medium (Melford) containing 100 mg ml^{-1} Amp. These cultures were incubated at 37 °C with shaking at 225 rpm for 16 h. Then 100 mg ml^{-1} Amp was added to 1 L of LB media in 2 L flasks and these were inoculated with 3 ml of the pre-culture, followed by further incubation at 37 °C with shaking (220 rpm) until an optical density at 600 nm (OD$_{600}$) of 0.6 was reached. At this point IPTG (Melford) was added a to final concentration of 0.5 mM and temperature reduced to 25 °C and shaking continued for 16 h. Cells were harvested by centrifugation at 3,501 g (F8-6x1000y rotor) using a Sorvall Evolution RC Superspeed centrifuge for 20 min at 4 °C. The resulting pellets were re-suspended in lysis buffer (50 mM Tris/HCl pH 7.5, 300 mM NaCl, 10% glycerol). To the resuspension, 1 μl of 1 M MgCl$_2$ was added to every 1 ml and stirred at 4 °C for 35 min. The cell suspension was passed through an EmulsiFlex-C5 cell disrupter (Avestin), equilibrated with lysis buffer, to lyse the cells. The lysate was centrifuged at 38,724 g (SS-34 rotor) using a Sorvall RC-5 centrifuge for 20 min at 4 °C. The resulting supernatant was subjected to ultracentrifugation at 108, 864 g (JA 30.50 Ti rotor), for 2 h at 4 °C in Beckman Coulter Avanti JXN-30 centrifuge. For the solubilisation of the membrane protein, the resulting pellet was resuspended in 25ml 50 mM Tris/HCl pH 7.5, 1 mM DTT, 10 mM PMSF and n-Dodecyl-β-D-Maltoside (DDM) detergent was added to a final concentration of 1% (v/v). The solubilisation step was carried out at room temperature with endo rotation for 2h. The resuspension was then centrifuged at 108, 864 g (JA 30.50 Ti rotor), in a Beckman Coulter Avanti JXN-30 centrifuge for 15 min at 4 °C. The soluble fraction was collected and applied to two 1 ml HiFliQ Ni-NTA columns (Generon) attached together equilibrated in Buffer A (50 mM Tris/HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 0.05% DDM) to bind the (His)$_6$-tagged SLI_RS17250 protein. The Ni-NTA column was attached to an AKTA-Prime and

washed extensively with Buffer A followed by elution with 100% Buffer B (Buffer A with 500 mM imidazole).

6.2.3   Western Blot
For Western Blot analysis, samples were prepared using cracking buffer (Appendix 5) and were separated by SDS-PAGE (Appendix 5). The samples in the SDS-PAGE were transferred onto a nitrocellulose membrane by electrophoresis at 66 Volts for 66 min. A blocking agent was applied to the membrane, 3% BSA (Thermo Fischer) and incubated at room temperature for 1 h. Then, the primary antibody was added, Anti (His)$_6$-tag antibody (100 µg at 0.1 mg/ml) - alkaline phosphatase (Abcam), to 3% BSA buffer in a 1:5000 dilution which was applied to the membrane. The antibody was incubated with the membrane for 1 h at room temperature. Finally, the membrane was developed chromogenically by adding one BCIP/NBT tablet (Sigma Fast) dissolved in 10 ml sterile water and incubated for 15 min at room temperature.

6.2.4   Mass Spectrometry - MALDI
Matrix-assisted laser desorption/ionisation (MALDI) mass spectrometry was carried out. This involved excising a gel fragment from an SDS-PAGE gel which were of the correct predicted molecular mass (18.92 kDa, predicted by Expasy (www.expasy.com)). The gel slice was subsequently treated and reduced, followed by digestion into peptides using protease trypsin. An Orbitrap Elite MS/MS spectrometer equipped with liquid chromatography device (nano-flow LC system with reverse phase column) was used to separate peptides, determine their size and amino acid sequence. These sequences are compared against those in a protein database to confirm identity of the protein. All MALDI analyses and sample preparation were carried out at University of Birmingham, Advanced Mass Spectrometry Facility Services (199).

6.3   Results
The phylogenetic distribution of SLI_RS17250 was discussed previously in Chapter 2 whereby various species possessed potential homologues of this membrane protein across the Tree of Life. It was deduced that SLI_RS17250 originated from a last common ancestor (LCA) in bacteria as many homologous proteins were found in prokaryotes. It is worth noting there were fewer species found that possessed homologues of SLI_RS17250 compared to phylogenetic analyses of Ccsp and SLI_RS17245 (the Na$^+$/H$^+$ antiporter). Also, the lack of known homologous structures suggests SLI_RS17250 to be a unique transporter type protein.

6.3.1   Cloning of expression constructs of SLI_RS17250

The first expression construct involved ligation of the *SLI_RS17250* gene insert into pET21a
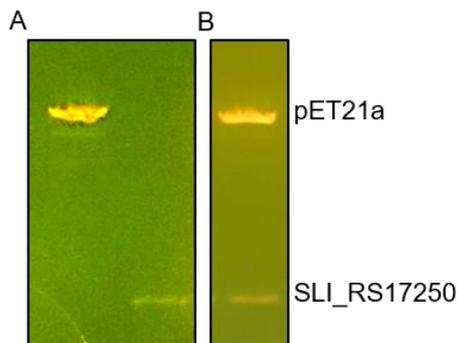
Figure 6.4 – *(A) Agarose gel image of restriction digested SLI_RS17250 insert (568 bp) and pET21a plasmid. (B) Agarose gel confirming the gene insert ligated into pET21a plasmid as a NdeI/BamHI fragment.*

plasmid (Fig. 6.4). Restriction digest (NdeI-BamHI) of both components was carried out beforehand to ensure successful ligation (Fig. 6.4). The production of two remaining expression constructs involved amplification of the *SLI_RS17250* DNA from the pTZ19R vector using the primers reported in Table 6.1 was successful based on the size of the visualised PCR products (~568 bp) on an agarose gel (Fig. 6.5A). These bands were excised, gel purified and subjected to restriction digest depending

on the plasmid. For ligation into a pET21a vector, restriction digest was carried out with NdeI and XhoI enzymes before ligation into this plasmid. For ligation into a pET26b vector, restriction digest was carried out with BamHI and XhoI enzymes before ligation into this plasmid. Transformants following ligation were checked for the correct insert by performing a restriction digest with the corresponding restriction enzymes (Fig. 6.5B & C) and DNA sequencing confirmed the correct sequence (Appendix 4).
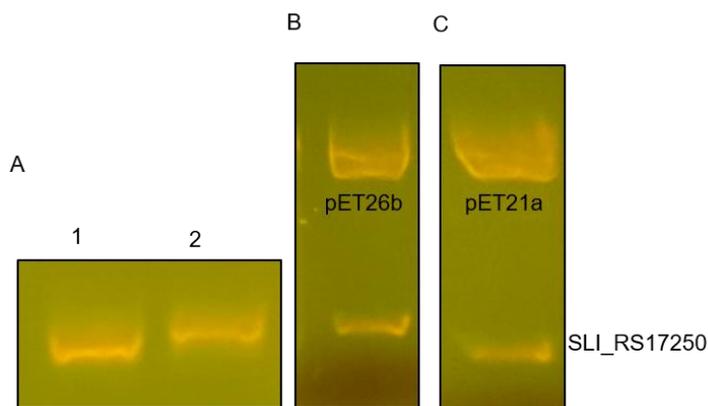
Figure 6.5 – (A) Agarose gel image of PCR products of amplified SLI_RS17250 DNA; lane (1) corresponds to gene insert for ligation into pET26b plasmid whereas lane (2) is gene insert targeted for insertion into pET21a vector. (B) and (C) displays agarose gels confirming the gene insert ligated into both pET26b (BamHI/XhoI) and pET21a (NdeI/XhoI).

6.3.2 Over-expression of the C-terminal His-tagged SLI_RS17250

The expression and purification of SLI_RS17250 involved a crucial step of membrane protein solubilisation using the detergent DDM (Fig. 6.6). DDM is a commonly used non-ionic detergent that is considered more gentle than other detergents and effective at preserving protein activity (200). The concentration of detergent added was 1 % (v/v) and observations would be made by 15 % SDS-PAGE analysis if this concentration was effective in solubilising SLI_RS17250. The success in detergent solubilisation of membrane proteins would be apparent if the protein has not precipitated out of solution. The purpose of utilising the detergent would be to prevent this and to hopefully visualise a protein band of correct molecular weight on the SDS-PAGE, in the detergent solubilised fractions. The SDS-PAGE image in Fig. 6.7 revealed a distinct band between 15-20 kDa, and thus in the area of the predicted mass (18.92 kDa) of the C-terminal (His)$_6$-tagged SLI_RS17250 protein. The solubilisation of the membrane protein appeared to be successful as a clear band can be seen in post solubilisation supernatant (Fig. 6.7). However, the expression band is also seen in the post solubilisation pellet suggesting that not the entirety of the protein has been solubilised (Fig. 6.7). The IMAC purification whereby 100 % elution buffer was applied yielded a distinct peak after ~15 ml had passed through the column with an absorbance of ~100 mAU (Fig. 6.8A). The 1 ml fractions were collected within this peak and were used for further analysis by SDS-PAGE whereby the expression band can be seen in all fractions collected (Fig. 6.8B). All SDS-PAGE samples were prepared by gentle heating
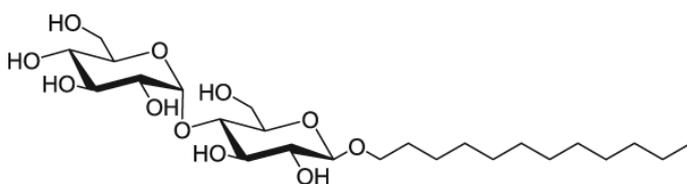
Figure 6.6 – Chemical structure of DDM

in a 42 °C water-bath for ~30 min. It is noteworthy that in all SDS-PAGE, distinct high molecular weight bands were visible at the top of the gels (Fig. 6.7 & 6.8 (orange arrows)). This likely indicates the SLI_RS17250 protein did not fully enter the gels which is a common issue for analysing membrane proteins using this method. This is often due to aggregation because of hydrophobicity of membrane proteins (201) but also their behaviour in the presence of SDS detergent (202).
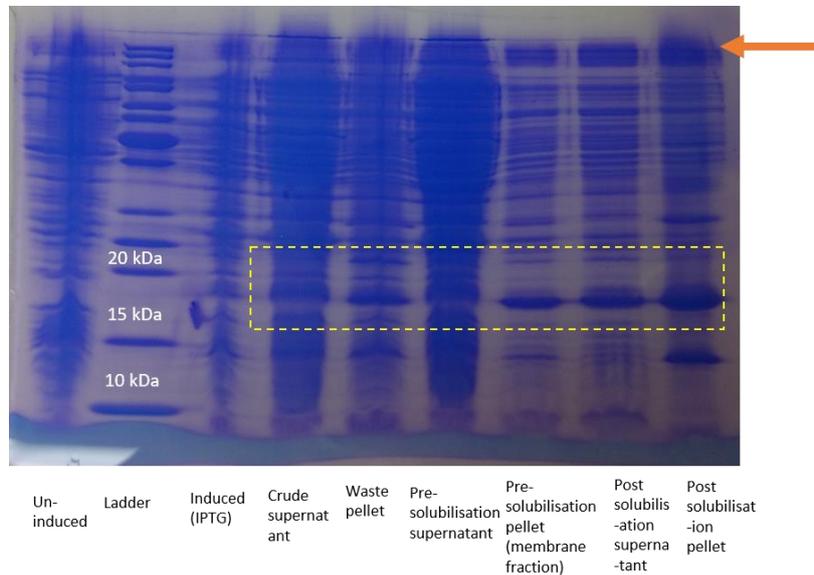
Figure 6.7 - *Coomassie stained 15 % SDS-PAGE analysis of fractions from the protein purification including solubilisation step with 1% DDM detergent. Yellow dashed box highlights a strong over-expression band running between 15-20 kDa consistent with that of the C-terminal (His)$_6$-tagged SLI_RS17250 protein. The orange arrow indicates high molecular weight bands at the top of SDS-PAGE.*
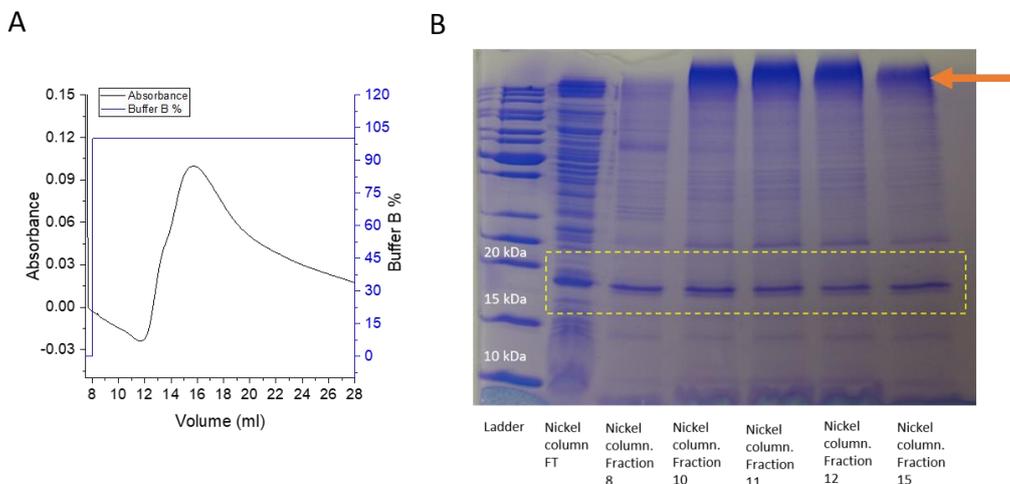
Figure 6.8 – *(A) Elution profile of (His)$_6$-tagged SLI_RS17250 protein on a 1 ml His-trap FF Ni-NTA column, blue line represents Buffer B gradient (B) Coomassie stained 15 % SDS-PAGE analysis of fractions from the Ni-NTA column. Yellow dashed box highlights visible bands running between 15-20 kDa. The orange arrow indicates high molecular weight bands at the top of SDS-PAGE.*

113

6.3.3 Western blot analysis is consistent with the presence of a His-tagged species

A Western blot analysis was carried out to verify whether the clear over-expression band and the fractions eluted and collected from the Ni-NTA column were from SLI_RS17250. To check this, the simplest way was to detect the presence of the C-terminal $(His)_6$-tagged that is fused to the SLI_RS17250 protein (Fig. 6.9). As shown in Fig. 6.9A, each lane of the immunoblot shows the detection of a band using the anti $(His)_6$-tag antibodies that migrates at ~ 20 kDa.
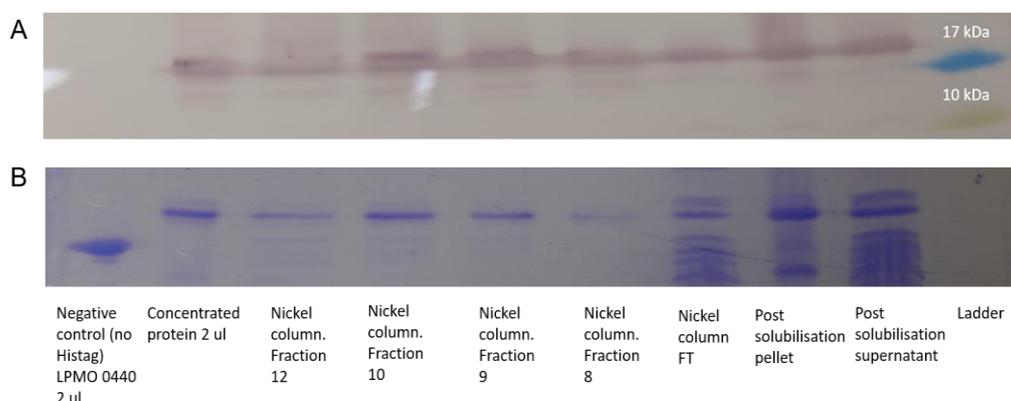
Figure 6.9 – (A) Western blot which shows positive $(His)_6$ tag signal and (B) corresponding SDS-PAGE gel used for blot of SLI_RS17250 protein fractions.

Based on a molecular weight of 18.92 kDa, predicted by Expasy (www.expasy.com) the band is consistent with where the solubilised $(His)_6$-tagged SLI_RS17250 should migrate. Staining the immunoblot with Coomassie blue clearly indicates the presence of a protein band (Fig. 6.9B), with the negative protein control, an LPMO from *S. lividans,* of similar molecular weight, but without a $(His)_6$-tag now also visible (Fig. 6.9B). Indeed, these bands show a positive presence of $(His)_6$-tagged protein in the post solubilisation supernatant and IMAC purification fractions. Despite this, there was also protein detected in the post solubilisation pellet which suggests significant protein precipitation after the detergent solubilisation step.

6.3.4 Proteomic analysis is not conclusive

The results of the trypsin digest MALDI experiment carried out at the University of Birmingham did not yield convincing results (Table 6.4). Analysis of the sample provided gave a mixture of peptides belonging to two species of *Streptomycetes: S. lividans* and *Streptomyces griseochromogenes* (Table 6.4). However, the results from mass spectrometry did not detect any matching peptides to the SLI_RS17250 protein sequence. From the amino-acid sequence of SLI_RS17250 shown in Fig. 6.3, it is apparent that the sequence lacks a sufficient number of Lys and Arg residues. In total there are four Lys and two Arg residues. Trypsin has specific protease activity at the carboxy side of these two positively charged amino acids which involves hydrolysing at the C-termini of these two amino acids during

proteolysis. MALDI involves separation of tryptic cleaved peptides by liquid chromatography which is combined with a mass spectrometer whereby the peptide masses are predicted. The identification of these peptides is carried out by MS/MS which is compared against a protein database. For the mass spectrometry method, a peptide fragment size of 6-20 amino acids is preferable for detection. Peptides that are larger or smaller than this range could be too difficult to detect by MS/MS (203, 204). The protein sequence of SLI_RS17250 was analysed via online webserver PeptideCutter (11) for specific trypsin cleavage sites (Fig. 6.10). This highlighted the issue of the size of resulting peptide fragments that were not in the ideal mass range.

```
      MDHSTHHSAPEDDGPGHPHGHGHLHGPGHPHGHVHVHGTTWATAMQATLHCLTGCAIGEI
  1   ---------+---------+---------+---------+---------+---------+   60


                                                  Tryps
                                    Tryps Tryps    |
                                      |     |  |
                                      |     |  |
      LGMVIGTALMWGNVPTMVLAIALAFVFGYSLTLFAVVRAGVSMKAAIKVALAADTVSIAV
  61  ---------+---------+---------+---------+---------+---------+  120


                                               Tryps
                                          Tryps  |
                                    Tryps    |   |
                                      |      |   |
                                      |      |   |
      MELVDNGIIALVPGAMEAHLSDGLFWYALLGGFAVAFVITTPVNKWMIGRGKGHAVVHAY
 121  ---------+---------+---------+---------+---------+---------+  180


      H
 181  -    181
```
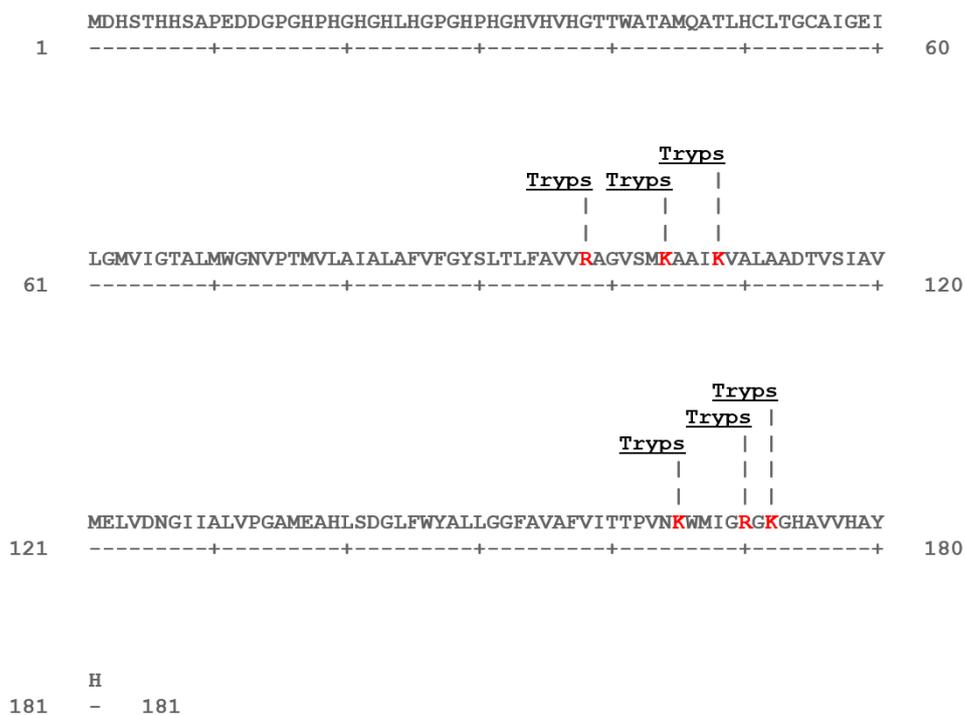
Figure 6.10 *– Amino acid sequence of SLI_RS17250 analysis using PeptideCutter (11). The corresponding sites that are cleaved by Trypsin are highlighted (red coloured amino acids). There is a total of six cleavage sites that include four Lys residues and two Arg residues. These amino acid position numbers are the following; 98, 104, 108, 165, 170 and 172. Cleavage by trypsin occurs on the right side (C-terminal direction) of the marked amino acid.*

Table 6.4 – *Peptides detected for SLI_RS17250 in MALDI with trypsin digest*

Accession	Description	Sequence
A0A076LWT5	Beta-lactamase OS=Streptomyces lividans TK24 GN=SLIV_00225 PE=4 SV=1 - [A0A076LWT5_STRLI]	AVLTAGMASALVGFtRGPtAASAPASR
D6EHV2	Transcriptional regulatory protein OS=Streptomyces lividans TK24 GN=SLIV_36875 PE=4 SV=1 - [D6EHV2_STRLI]	GVHDLLDDEPDItVVGEAATVEqALVR
A0A076M013	Hydrolase OS=Streptomyces lividans TK24 GN=SLIV_12985 PE=4 SV=1 - [A0A076M013_STRLI]	AVNGLAFALPMLGAARGAAAVWTSWTAGKLAGPtGqnAVSSqDR
D6ECN0	Uncharacterized protein OS=Streptomyces lividans TK24 GN=SLIV_14160 PE=4 SV=1 - [D6ECN0_STRLI]	VRTVAPGTPLFDATAGGmGLtGVILtATLRLQPVETALmSVDTER
M1GT27	Uncharacterized protein OS=Streptomyces griseochromogenes GN=ex-bls-5 PE=4 SV=1 - [M1GT27_9ACTN]	yStSRSySENHAFRLSSQMSIR
A0A076LZX1	Uncharacterized protein OS=Streptomyces lividans TK24 GN=SLIV_06460 PE=4 SV=1 - [A0A076LZX1_STRLI]	TQAyFERRtQEGK
A0A076MGH7	Uncharacterized protein OS=Streptomyces lividans TK24 GN=SLIV_28780 PE=4 SV=1 - [A0A076MGH7_STRLI]	AAVAVSATSAtVtGTGSR

6.3.5   Bioinformatic topology prediction of SLI_RS17250

To gain some insight into the topology of SLI_RS17250 a web server TOPCONS (http://topcons.net/) (3, 4) was used. TOPCONS utilises a consensus approach whereby topology predictions are taken from various different predictor programs (Fig. 6.11A) to form a topology profile that is inputted into TOPCONS Hidden Markov Model (3, 4) to produce a final topology result (Fig. 6.11B) (3, 4). The FASTA amino acid sequence of SLI_RS17250 (Fig. 6.3) was used as input for the web server topology predictor. TOPCONS was able to identify four transmembrane (TM) helices in SLI_RS17250 in the following amino acid positions: 49-69 (TM1), 75-95 (TM2), 111-131 (TM3) and 143-163 (TM4) (Fig. 6.11). TM1 and TM3 are predicted to span in the membrane from the cell cytoplasm towards the extracellular environment (IN -> OUT) whereas TM2 and TM4 helices are presumed to be positioned (OUT-

> IN) in the membrane thus spanning from the extracellular environment towards the cell cytoplasm (Fig. 6.11B). This result is, in majority, supported by the other predictor programs (3, 4) showing the same outcome and particularly showing negative ΔG values TM2 and TM4 suggests a high probability that this is the true topology for these regions (Fig. 6.11A). This is reflected in the final TOPCONS prediction in Fig. 6.11B whereby the reliability score for the positions of most of these TM helices is close to 1 except for TM3, lacking a reliability score. Despite this, the topology of other regions of SLI_RS17250 remain uncertain due to positive ΔG values (Fig. 6.11A) and lacking reliability scores (Fig. 6.11B). TOPCONS analysis shows majority of regions in SLI_RS17250 are predicted mainly to lie 'inside' the membrane, closer to the cytoplasmic space (Fig. 6.11B and C). TOPCONS predicts smaller regions to lie 'outside' the membrane which include amino acid positions 70-74 and 132-142 (Fig. 6.11B). This prediction is also supported by the following programs Philius, PolyPhobius and SCAMPI (Fig. 6.11A) (3, 4). But other predictor programs (OCTOPUS, SPOCTOPUS) (3, 4) reveal different results to this, differences in TM helices topology as well as a lack of TM3 in their consensus (Fig. 6.11A). This is insightful as the reliability score for TM3 in Fig. 6.11B reflects this uncertainty. Moreover, the overall positive ΔG scores for the non-TM helical regions in SLI_RS17250 reveal the unlikelihood of all predicted topologies (Fig. 6.11A). Overall the topology and amino acid position predictions of TM1, TM2 and TM4 in the final TOPCONS consensus result (Fig. 6.11B) show high reliability scores thus suggests a potential topology whereas existence and topology of TM3 remains unresolved (Fig. 6.11C).
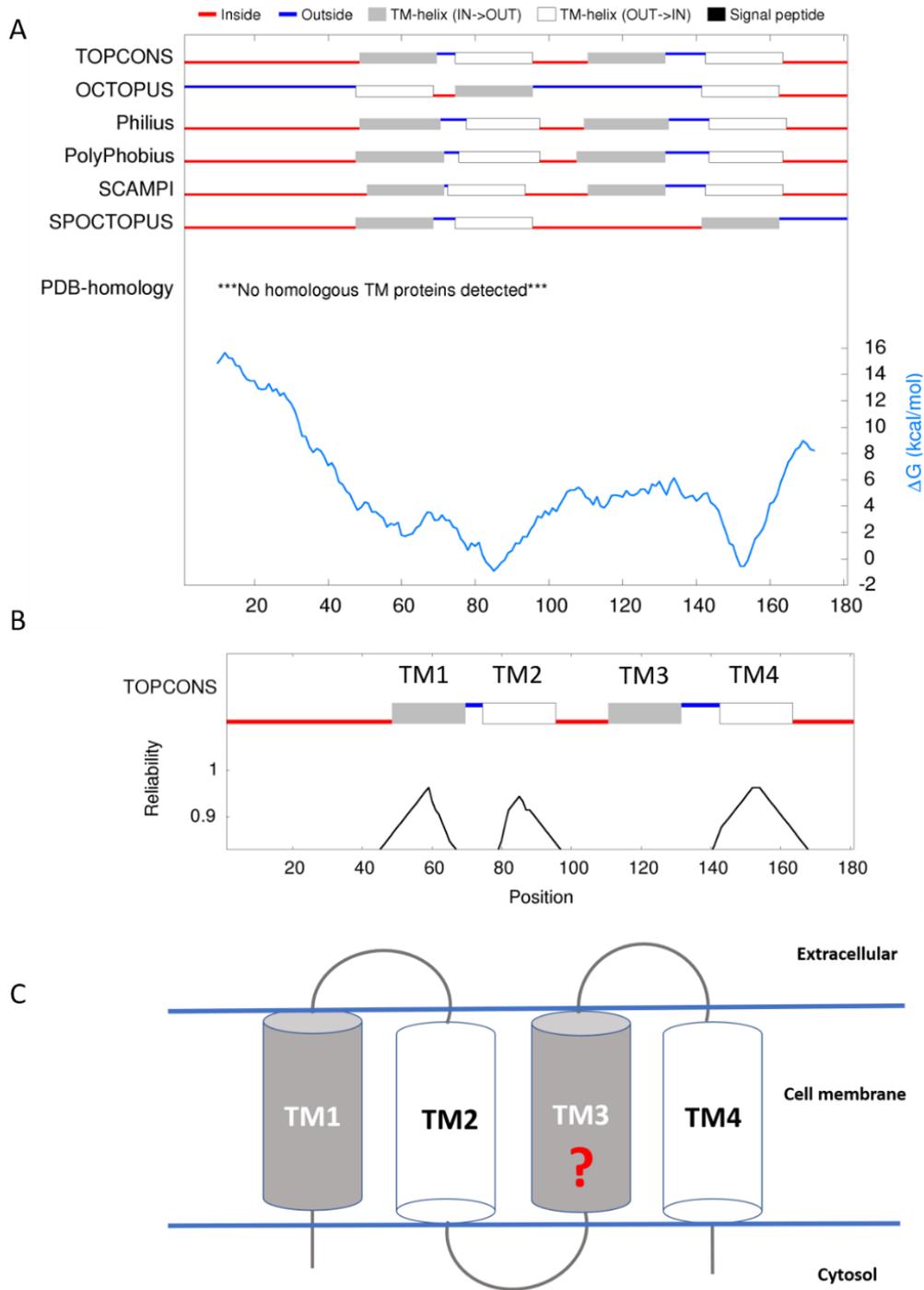
Figure 6.11 - *The amino acid sequence is analysed using webserver TOPCONS (3, 4) for certain regions that may be situated nearer the cell cytoplasm (inside) or towards the extracellular space (outside) (A) The figure also predicts amino acids that may form transmembrane (TM) helices and their orientation in the membrane based on the 'positive inside' law (9, 10). These predictions are further quantified for reliability with the ΔG (kcal/mol). TOPCONS prediction revealed four TM helices for SLI_RS17250, in the following amino acid positions 49-69 (TM1), 75-95 (TM2), 111-131 (TM3) and 143-163 (TM4). These TM helices have been labelled in (B) and are applicable for the same helices in (A). The figure in (B) reveals the final consensus result for topology based on the data in figure (A) showing reliability score on the y axis for topologies of certain amino acids. Figure (C) is a topology model of SLI_RS17250 based on the results shown in figure (B).*

6.4 Discussion

The expression of membrane proteins remains an ongoing challenge. 20-30% of all genes in pro- and eukaryotes encode membrane proteins (195) but there is a low abundance of these proteins naturally to extract for *in vitro* studies (195). The expression of helical bundle membrane proteins in prokaryotic systems has proven to be more difficult versus expression of β-barrel membrane proteins. *E. coli* is the most utilised bacterial expression host and β-barrel membrane proteins can be synthesised as inclusion bodies in *E. coli* which can be easily extracted and refolded structurally (195, 205). In contrast, helical bundle membrane proteins are rarely extracted from inclusion bodies successfully (195). Helical bundle membrane proteins require to be properly inserted into the membrane followed by detergent extraction (195). Indeed, this has limited the amount of published structures available of helical bundle membrane proteins. There is a vast amount of methods available for improving expression yields of membrane proteins. This includes the engineering of various strains of *E. coli* to improve its capabilities in expressing membrane proteins. In addition, the existence of *E. coli*-based cell-free protein production systems prevent the possible toxicity of membrane protein expression (195). The development of new *E. coli* strains includes C41(DE3) and C43(DE3) strains, also known as the Walker strains (195, 206, 207). The Walker strains (C43 used in the present study) derive from the commonly used *E. coli* strain BL21 (DE3) in which its protein expression mechanism has been well established. Protein expression in BL21 (DE3) involves a mutant lacUV5 promoter which regulates expression of bacteriophage T7 RNA polymerase (T7 RNAP) and is superior to the wild type lac promoter (195, 208). T7 RNAP transcribes the gene of interest and is more rapid than *E. coli* RNAP (195). This expression system overall produces more protein due to more mRNA being made (195). The issues in membrane protein expression in *E. coli* derive from this fact as over-expression of membrane encoding genes can cause the Sec-translocon to become saturated during biogenesis (Fig. 6.2) (195). The Walker strains can cope with toxic membrane protein expression due to mutations in the lacUV5 promoter (195, 207) but these strains do not demonstrate improved expression yields for every membrane protein (195, 206, 207). The mutations in the lacUV5 promoter lead to a lower expression of T7 RNAP compared to BL21(DE3) thus leading to lower amounts of mRNA being produced thus does not limit the activity of the Sec-translocon of embedding proteins into the membrane (195, 207, 209, 210).

The method described for the expression of SLI_RS17250 still requires some improvements. Indeed, a crucial component is the concentration of detergent used, in this case DDM. DDM is non-ionic and considered one of most utilised detergents due its

119

preservation of protein activity and low UV-Vis absorptivity (200). An excess of detergent is used in the primary solubilisation step of re-forming the membrane protein during protein expression and purification (185). Many studies investigate the interaction of detergents with membrane proteins and are essential in protein crystallisation. The key in resolubilising the membrane protein is using a suitable concentration of detergent as too high amounts could denature the protein or, while too little will cause aggregation/precipitation (200, 211, 212). This can be observed in Fig. 6.7 as a distinct band of the protein of interest can be seen in 'Post solubilisation pellet' lane. This is an indication that perhaps not enough detergent was used to fully resolubilise SLI_RS17250 and has aggregated in the waste pellet.

The SDS-PAGE results initially confirmed the expression of SLI_RS17250 which displayed a band of the correct size of around ~18.92 kDa (Figs. 6.7 & 6.8). This was further confirmed by Western blot (Fig. 6.9) using an Anti-(His)$_6$ tag antibody. However, this positive confirmation by these two methods does reveal some issues in membrane protein expression. The hydrophobicity of membrane proteins causes aggregation thus leading to no separation on SDS-PAGE and minimal transfer from gel to blotting membrane for Western blots (201). This can be observed in Figs. 6.7 & 6.8 whereby much aggregation of protein is seen at the top of the SDS-PAGE. This also makes it difficult to quantify the amount of protein present. This aggregation may also be due to the behaviour of membrane proteins in the presence of SDS as it has been reported that SDS may not fully denature some membrane proteins (202). However, success in this area have been seen in other studies with using smaller transmembrane proteins that have fewer alpha helices (201, 213). Further confirmation about the identity of the ~18 kDa protein band was sought out through MALDI with trypsin digest. But the results from this analysis were not informative and the peptides (Table 6.4) did not match with the SLI_RS17250 protein sequence (Fig. 6.3). This again highlights the difficulties associated with transmembrane protein expression and the need for further research into this area. Further consideration into using MALDI again for proteomic analysis of SLI_RS17250 could involve utilising a different protease. An ideal proteolytic compound for future use on SLI_RS17250 could be cyanogen bromide (CNBr) as this compound can cleave the C-terminus of methionine. It is observable in the amino acid sequence of SLI_RS17250 (Fig. 6.3) there are an ideal number of Met residues that CNBr could act upon and produce ideal peptides for MALDI mass spectrometry. As highlighted in Fig. 6.10, trypsin cleaves SLI_RS17250 peptides that are not suitable for MALDI.

It would be interesting to determine the topology of SLI_RS17250. It is well known that transmembrane (TM) proteins embed themselves in the lipid bilayer in a specific

orientation. For instance, predictions can be made on whether their C- and N- termini are positioned outside of the cytosol (out) or pointing within the cell (in) (214). This is the case for alpha helical TMs whereby their helices are specifically orientated within the membrane (214). The attempt to predict the topology of SLI_RS17250 was carried out as shown in Fig. 6.11 Indeed, the use of TOPCONS (3, 4) revealed an interesting outcome whereby four TM helices were identified in this novel protein. These were appropriately labelled TM1, TM2, TM3 and TM4. TM1 and TM3 were predicted to possess an 'IN->OUT' topology whereas TM2 and TM4 were forecast to have a 'OUT-> IN' position in the membrane (Fig. 6.11B). Although, the topology prediction of TM3 remains questionable due to low reliability score (Fig. 6.11B) which this final consensus could be influenced by topology predictions of OCTOPUS and SPOCTOPUS (Fig. 6.11A). The topology of the non-helical regions of SLI_RS17250 remain vague due to conflicting topology predictions, low reliability scores and ΔG values of these results (Fig. 6.11). It is noteworthy that no matching homology structures were found in PDB (Fig. 6.11A). To determine topology, there are few techniques that can achieve this which are mainly computational simulations and very few biochemical analyses exist (25). There are protein tags that have been proven to determine C-terminal orientation such as green fluorescent protein (GFP) (214-217). GFP allows monitoring of the TM's integrity as the GFP moiety will only fluoresce if the TM is folded properly and inserted into the membrane (201). This allows visualisation of the GFP-TM complex in intact cells and using in-gel fluorescence in SDS-PAGE gels as well as size-exclusion chromatography combined with a fluorescence detector (201, 212). But fewer methods exist to determine the orientation of the N-terminus (214). This is due to the N-terminus becoming first inserted into the plasma membrane during TM protein biogenesis in *E. coli* (214). Moreover, computer simulations rely on deposited TM protein structures which there are very few in the Protein Database. In addition, the predictive algorithms still rely on experimental evidence which without it, could result in incorrect predictions about the orientation and/or alignment of these helices (214).

The work presented in this chapter provides a method of expressing and purifying SLI_RS17250 and appears to be soluble as seen in SDS-PAGE and Western blot despite lacking firm confirmation by MALDI. This provides a foundation for further work to be carried out with SLI_RS17250 in order to decipher its structural characteristics and possibly its function *in vivo* which remains unknown. Additional purification of SLI_RS17250 is still required such as size exclusion chromatography. Indeed, this would offer more information concerning its molecular mass but can prove problematic due to presence of detergent. In the study by Gimpl *et al*., it is discussed that the addition of detergent to the membrane protein would

form protein/detergent complexes (PDCs) (212). If high amounts of detergent are added, then these PDCs would be present with detergent monomers and other micelles which could affect the composition of these PDCs (212). The method of size exclusion chromatography (SEC) is useful in determining oligomeric state and homogeneity of soluble proteins but predicting these aspects of PDCs proves challenging (212). This is since the presence of detergent alters a protein's elution performance which SEC is highly dependent on and can only provide an estimate on molecular mass, homogeneity and stability (212, 218). Thus, in order to improve understanding of PDCs conformation, an additional technique can be applied to SEC known as dynamic light scattering (DLS) which identifies aggregation in PDCs and offers information about their hydrodynamic radius (212, 219, 220). DLS has the drawback of being unable to distinguish between the effects of either protein or detergent to the overall hydrodynamic behaviour (212). However, other techniques exist that combined with SEC can better determine the overall conformation of PDCs. For instance, SEC-MALS (multi-angle light scattering) can differentiate and calculate molar masses of desired protein and the detergent micelles. SEC-MALS can achieve this through three main detectors which include ultraviolet (UV), MALS and refractive index (RI). The combination of these three components provides an effective method for separation and detection of different oligomeric species. SEC is crucial for the separation of these species though elution in SEC does not correspond to the molar masses of the species of interest. It is only important that these oligomers are separated from one another before passing through the MALS detector.

Further work with SLI_RS17250 would include deciphering the X-ray crystal structure of this novel TM protein. The challenges associated with crystallisation of TM proteins have consequently led to very few solved structures being published. Since the first published structure in 1985 (221) there has been a gradual increase in this area but less than 1% of all entries in Protein Data Bank (PDB) consist of membrane protein structures (183). It is well established that the reasons for this include TM proteins are hydrophobic and are instable without presence of detergents. Additionally, the presence of the detergent affects the process of protein crystallisation (183). Thus, methods have been developed to overcome these issues which includes the use of lipidic cubic phase (LCP) crystallisation. LCP crystallisation (or *in meso* crystallisation) involves crystallising proteins in lipidic mesophases (222). The method of LCP involves two main steps. Firstly, lipid is mixed with the protein of interest followed by a lipidic cubic phase forming which creates a highly viscous gel like substance (223). Moreover, crystallisation can be improved by the addition of precipitants or

other enhancers (223). LCP provides an environment similar to the lipid bilayer which membrane proteins are naturally found thus facilitating their nucleation and crystal growth. The LCP method was first used to obtain a high-resolution structure of bacteriorhodopsin (bR) by Landau and Rosenbusch in 1997 (224). Thereafter, some of the most important membrane proteins used as drug targets include human G protein-coupled receptors (GPCRs) family (222, 225-229) which members have been successfully crystallised through LCP. However, as instrumental LCP has been in membrane protein crystallisation there were initial difficulties that included mixing and handling of the highly viscous cubic phase material (223). This set drawbacks in crystallisation especially when testing a high number of screening conditions. Thus, recent improvements in LCP include the development of automated instruments that facilitated mixing of protein and lipid to produce the cubic mesophase (223). Indeed, continuous research into membrane protein crystallisation will undoubtedly produce more solved structures in the future.

Copper (Cu) is essential for the morphological development of *S. lividans*. Despite the detailed characterisation of the transcriptional response of the CsoR/CopZ/P_1-type ATPase trafficking and efflux system, it is possible that this is not the only Cu homeostatic mechanism. What has been reported in transcriptional studies (RNA-Seq) during Cu stress were other genes that become up- and down-regulated (18). This data carried out in Δ*csoR* and thus not regulated transcriptionally by CsoR (18) showed a 6-fold upregulation of SLI_RS17250 (DUF4396) (Fig. 2.2). Thus, revealing a possible new Cu resistance system but the exact function of DUF4396 remains unknown. The amino acid sequence of this novel protein was searched in BLAST against the PDB but did not yield any close homologues to SLI_RS17250 sequence that have been structurally characterised. It is possible that this transmembrane protein could be a member of a different protein family. There are however, some small similarities in the substrate binding S-subunits of the energy coupling factor (ECF) family which are micronutrient transporters (137). There are also connections to the SLC11/NRAMP family which have a role in the transport of transition metal ions (138, 139). Indeed, further investigation is required to establish if SLI_RS17250 is involved in metal ion transport and if it interacts with *Sl*Csp3.

Chapter Seven

7.1 Conclusion

The intention of utilising *Streptomyces lividans* as an expression host on an industrial scale would involve growing large quantities of cell cultures in bioreactors. The overall importance of this is to be able to characterise each component involved in the overall development of *S. lividans*. By deciphering the biochemical pathways and proteins that catalyse them, steps can be taken to modify these pathways to limit the production of aerial hyphae when *S. lividans* is cultivated in liquid culture (12, 101, 103). This would inhibit the formation of pellets which otherwise would severely reduce the production of useful secondary metabolites (e.g. antibiotics) (12, 14). The importance of copper (Cu) to *S. lividans* development has been extensively discussed in this thesis, including the role of certain metalloproteins that handle Cu.

The newly discovered *Sl*Csp3 has been a main focus of this thesis. Indeed, the discovery of cytosolic storage proteins that store between 70-80 Cu(I) ions was unprecedented due to the belief that prokaryotes did not have a metabolic requirement for Cu in their cytosol and thus safeguard against the risk of Cu toxicity (1, 2, 19, 75). The first discovered Cu storage protein was Csp1, which is located in the bacterial periplasm and found in methanotroph *Methylosinus trichosporium* OB3b (1). Subsequently, other Csps have been characterised and cytosolic Csps have been associated with the family of Csp3 type protein (2, 75, 76). *Sl*Csp3 is the first non-methanotrophic member of the Csp3 family to have been structurally characterised loaded with Cu(I) (19). *Sl*Csp3 shares structural similarities with other Csp3 members including a ~50% similarity with *Mt*Csp3 (169). The Cu(I) binding affinity of *Sl*Csp3 reported in Chapter 3 is in line with those reported for other Csp3s such as *Bs*Csp3 and *Mt*Csp3 (2, 19) but is the only member to demonstrate cooperative binding with the exception of *Mt*Csp1 (19, 75). The mechanistic basis for this is most likely due to the [$Cu_4(\mu_2$-S-Cys)$_4$] clusters being thermodynamically more favoured (19, 76). This was confirmed under low Cu(I) loadings of *Sl*Csp3 whereby sites in the inner core remained empty (Chapter 4). In addition, the formation of a tetranuclear [$Cu_4(\mu_2$-S-Cys)$_4(N^{\delta 1}$-His)] cluster was formed in the outer core of *Sl*Csp3 which contradicts findings of cluster formation in the inner core of *Mt*Csp3 (75, 169).

Indeed, structural studies of Cu(I) loading in *Sl*Csp3 builds an elaborate picture of various Cu(I) cluster formations. Nonetheless, there is an importance of understanding the mechanism of Cu(I) capture and loading. This was investigated in Chapter 5 whereby kinetic studies of Cu(I) loading of the hydrophilic entry site (outer core) of *Sl*Csp3 were carried out.

The data revealed the importance of His107 in the formation of transient ligand-exchange complex with $[Cu(BCA)_2]^{3-}$ and the role of this residue in transferring Cu(I) to other sites within *Sl*Csp3. The formation of a binding complex in *Sl*Csp3 was not observed with free Cu(I) or $[Cu(BCA)_2]^{3-}$ due to no linear dependence of k_{obs} on Cu(I) concentration. The rapid metal ion binding and formation of transient ligand complexes is not uncommon in metalloproteins involved in metal transfer (163, 230) such as Cu(I) transfer studies carried out between CopZ3079 and *Sl*Csp3 (Chapters 3 and 5). The overall outcome of the crystallographic and kinetic studies described in Chapters 4 & 5 was formulating a thermodynamic model to illustrate Cu(I) binding in terms of relative energy and the relative stabilities of Cu(I) binding sites in *Sl*Csp3 (Fig. 5.6).

Future investigations may include continuing kinetic studies with *Sl*Csp3 to determine the mechanism of Cu(I) loading to the inner core. Indeed, this process of Cu(I) ions shuttling into the main core of the *Sl*Csp3 helix remains uncertain and further enquiry is required to decipher this. This can prove challenging due to the number of potential binding sites (cognate and non-cognate) but is possible this could involve an internal ligand-exchange mechanism as Cu(I) ions pass from site to site within the Cys core. Further studies for this could involve site-directed mutagenesis of multiple Cys residues in the inner core of the protomer simultaneously and observing the kinetics of Cu(I) binding of these variants compared to wild-type *Sl*Csp3.

As discussed in Chapter 6, the expression of SLI_RS17250 (DUF4396) is deemed possible. Further work would include structural characterisation of this putative transmembrane protein. Indeed, traditional methods of crystallising proteins via X-ray crystallography have been insufficient in crystallising membrane proteins (183). Thus, other methods for this would include lipidic cubic phase (LCP) crystallisation as described in Chapter 6. The overall function of SLI_RS17250 remains unknown thus to obtain a structure would aid to infer its purpose. In addition, SLI_RS17250 requires a Na^+/H^+ antiporter to regulate the ion gradient across the cell membrane which is SLI_RS17245. This antiporter has not been studied in this thesis but future projects could include attempts to express this protein separately or in a co-expression system with SLI_RS17250 to obtain a soluble amount for structure determination, biochemical analyses and *in vivo* tests would be desirable.

Previous transcriptional RNA-seq studies illustrated a 6-fold up-regulation of *SLI_RS17250* when under Cu stress (18, 19) (Fig. 2.2). It seems to not be purely coincidental that under Cu stress, a clear growth phenotype of *S. lividans* was observed from in *in vivo* studies (Chapter 3). This was demonstrated with a knockout mutant of *Sl*Csp3 in the presence

of a high concentration of Cu (> 200 μM and with media dependence) and it could also be inferred that *Sl*Csp3 appears not to participate in the Cu trafficking pathway for CcO and GlxA metalation (Chapter 3). Moreover, *Sl*Csp3 expression is not under the control of CsoR as demonstrated by the absence of a consensus CsoR binding site in the *slcsp3* promoter region and expression induction in the *csoR* mutant (18). Indeed, these findings strongly suggest a second layer of Cu resistance in *S. lividans* that becomes active once the CsoR/CopZ/ATPase system becomes saturated (Chapter 3) (19).

Whilst the up-regulation of *slcsp3* may well be an act of last resort to survive, there remains several lines of future enquiry. One of these concerns whether *Sl*Csp3 is simply acting in a storage capacity, taking delivery of Cu(I) from CopZ when the CsoR regulon becomes saturated. If this is the case then on returning to homeostasis there will be a large store of Cu in the cytosol for which its requirement is presently unclear. Future studies include testing if *Sl*Csp3 could act as a donor to SLI_RS17250 (DUF4396), which through the coupled action with a Na^+/H^+ antiporter (SLI_RS17245), moves Cu(I) out of the cytosol. The taxonomic distribution discussed in Chapter 2 revealed that all three genes are present in representatives of five Bacteria groups. It is noted that one of these Bacterial groups is the proteobacteria, which possess numerous pathogens. It may therefore be that a role in the pathogen for *Sl*Csp3 and the export system is to harness the host derived bactericidal Cu during infection (32, 75, 231, 232). The overall function of Csps to act as virulence factors in pathogens remains unverified and requires further study. Clearly there is a need for further structural and functional investigation of the DUF4396 domain to be able to establish whether a role in transporting metal ions, such as Cu(I) in the present case, is possible and if there is an interplay with *Sl*Csp3. Finally, it will be paramount to discover the identity of the Cu sensitive regulator that acts to suppress the expression of *Sl*Csp3 under homeostasis conditions and to determine whether it is the same regulator that governs the Cu responsive control of SLI_RS17245 and SLI_RS17250.

References

1.   Vita N, Platsaki S, Basle A, Allen SJ, Paterson NG, Crombie AT, et al. A four-helix bundle stores copper for methane oxidation. Nature. 2015;525(7567):140-3.

2.   Vita N, Landolfi G, Basle A, Platsaki S, Lee J, Waldron KJ, et al. Bacterial cytosolic proteins with a high capacity for Cu(I) that protect against copper toxicity. Sci Rep. 2016;6:39065.

3.   Bernsel A, Viklund H, Hennerdal A, Elofsson A. TOPCONS: consensus prediction of membrane protein topology. Nucleic Acids Res. 2009;37(Web Server issue):W465-8.

4.   Tsirigos KD, Peters C, Shu N, Kall L, Elofsson A. The TOPCONS web server for consensus prediction of membrane protein topology and signal peptides. Nucleic Acids Res. 2015;43(W1):W401-7.

5.   Wang L, Zhu M, Zhang Q, Zhang X, Yang P, Liu Z, et al. Diisonitrile Natural Product SF2768 Functions As a Chalkophore That Mediates Copper Acquisition in *Streptomyces thioluteus*. ACS Chem Biol. 2017;12(12):3067-75.

6.   El Ghazouani A, Basle A, Firbank SJ, Knapp CW, Gray J, Graham DW, et al. Copper-binding properties and structures of methanobactins from *Methylosinus trichosporium OB3b*. Inorganic chemistry. 2011;50(4):1378-91.

7.   Kumar S, Stecher G, Tamura K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Mol Biol Evol. 2016;33(7):1870-4.

8.   Arachea BT, Sun Z, Potente N, Malik R, Isailovic D, Viola RE. Detergent selection for enhanced extraction of membrane proteins. Protein Expr Purif. 2012;86(1):12-20.

9.   Bernsel A, Viklund H, Falk J, Lindahl E, von Heijne G, Elofsson A. Prediction of membrane-protein topology from first principles. Proc Natl Acad Sci U S A. 2008;105(20):7177-81.

10.   von Heijne G. Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol. 1992;225(2):487-94.

11.   Gasteiger E. HC, Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. Protein Identification and Analysis Tools on the ExPASy Server. (In) John M Walker (ed): The Proteomics Protocols Handbook, Humana Press Copyright Humana Press; 2005. p. pp. 571-607.

12.   Petrus MLC, Vijgenboom E, Chaplin AK, Worrall JAR, van Wezel GP, Claessen D. The DyP-type peroxidase DtpA is a Tat-substrate required for GlxA maturation and morphogenesis in *Streptomyces*. Open Biology. 2016;6(1).

13.   Bagchi P, Morgan MT, Bacsa J, Fahrni CJ. Robust affinity standards for Cu(I) biochemistry. J Am Chem Soc. 2013;135(49):18549-59.

14.   Worrall JA, Vijgenboom E. Copper mining in *Streptomyces*: enzymes, natural products and development. Nat Prod Rep. 2010;27(5):742-56.

15.   Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 2002;30(14):3059-66.

16.   Balasubramanian R, Rosenzweig AC. Copper methanobactin: a molecule whose time has come. Curr Opin Chem Biol. 2008;12(2):245-9.

17.   Chaplin AK, Tan BG, Vijgenboom E, Worrall JA. Copper trafficking in the CsoR regulon of *Streptomyces lividans*. Metallomics. 2015;7(1):145-55.

18.   Dwarakanath S, Chaplin AK, Hough MA, Rigali S, Vijgenboom E, Worrall JA. Response to copper stress in *Streptomyces lividans* extends beyond genes under direct control of a copper-sensitive operon repressor protein (CsoR). J Biol Chem. 2012;287(21):17833-47.

19.   Straw ML, Chaplin AK, Hough MA, Paps J, Bavro VN, Wilson MT, et al. A cytosolic copper storage protein provides a second level of copper tolerance in *Streptomyces lividans*. Metallomics. 2018;10(1):180-93.

20.   Dalbey RE, Wang P, Kuhn A. Assembly of bacterial inner membrane proteins. Annu Rev Biochem. 2011;80(1):161-87.

21.   Boeckx T, Winters A, Webb J, Kingston-Smith A. Polyphenol oxidase in leaves: Is there any significance to the chloroplastic localization?2015.

22.   Kim E, Helton ME, Wasser IM, Karlin KD, Lu S, Huang HW, et al. Superoxo, mu-peroxo, and mu-oxo complexes from heme/O2 and heme-Cu/O2 reactivity: copper ligand influences in cytochrome c oxidase models. Proc Natl Acad Sci U S A. 2003;100(7):3623-8.

23.   Eijsink VGH, Petrovic D, Forsberg Z, Mekasha S, Rohr AK, Varnai A, et al. On the functional characterization of lytic polysaccharide monooxygenases (LPMOs). Biotechnol Biofuels. 2019;12(1):58.

24.   Ward B. Chapter 11 - Bacterial Energy Metabolism. In: Tang Y-W, Sussman M, Liu D, Poxton I, Schwartzman J, editors. Molecular Medical Microbiology (Second Edition). Boston: Academic Press; 2015. p. 201-33.

25.   Festa RA, Thiele DJ. Copper: an essential metal in biology. Curr Biol. 2011;21(21):R877-83.

26.   Grass G, Rensing C, Solioz M. Metallic copper as an antimicrobial surface. Appl Environ Microbiol. 2011;77(5):1541-7.

27.   Robinson NJ, Winge DR. Copper metallochaperones. Annu Rev Biochem. 2010;79:537-62.

28.   Yoshida Y, Furuta S, Niki E. Effects of metal chelating agents on the oxidation of lipids induced by copper and iron. Biochim Biophys Acta. 1993;1210(1):81-8.

29.   Macomber L, Imlay JA. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc Natl Acad Sci U S A. 2009;106(20):8344-9.

30.   Hodgkinson V, Petris MJ. Copper homeostasis at the host-pathogen interface. J Biol Chem. 2012;287(17):13549-55.

31.   MacKenzie EL, Iwasaki K, Tsuji Y. Intracellular iron transport and storage: from molecular mechanisms to health implications. Antioxid Redox Signal. 2008;10(6):997-1030.

32.   Fu Y, Chang FM, Giedroc DP. Copper transport and trafficking at the host-bacterial pathogen interface. Acc Chem Res. 2014;47(12):3605-13.

33.   Ladomersky E, Petris MJ. Copper tolerance and virulence in bacteria. Metallomics. 2015;7(6):957-64.

34.   Elguindi J, Wagner J, Rensing C. Genes involved in copper resistance influence survival of *Pseudomonas aeruginosa* on copper surfaces. Journal of applied microbiology. 2009;106(5):1448-55.

35.   Espirito Santo C, Taudte N, Nies DH, Grass G. Contribution of copper ion resistance to survival of *Escherichia coli* on metallic copper surfaces. Appl Environ Microbiol. 2008;74(4):977-86.

36.   Huffman DL, O'Halloran TV. Function, structure, and mechanism of intracellular copper trafficking proteins. Annu Rev Biochem. 2001;70(1):677-701.

37.   Waldron KJ, Rutherford JC, Ford D, Robinson NJ. Metalloproteins and metal sensing. Nature. 2009;460(7257):823-30.

38.   Coyle CL, Zumft WG, Kroneck PM, Korner H, Jakob W. Nitrous oxide reductase from denitrifying *Pseudomonas perfectomarina*. Purification and properties of a novel multicopper enzyme. Eur J Biochem. 1985;153(3):459-67.

39.   Pauleta SR, Dell'Acqua S, Moura I. Nitrous oxide reductase. Coordination Chemistry Reviews. 2013;257(2):332-49.

40.   Yoshikawa S, Muramoto K, Shinzawa-Itoh K. Proton-pumping mechanism of cytochrome C oxidase. Annu Rev Biophys. 2011;40(1):205-23.

41.   Glerum DM, Shtanko A, Tzagoloff A. SCO1 and SCO2 act as high copy suppressors of a mitochondrial copper recruitment defect in Saccharomyces cerevisiae. J Biol Chem. 1996;271(34):20531-5.

42.   Nittis T, George GN, Winge DR. Yeast Sco1, a protein essential for cytochrome c oxidase function is a Cu(I)-binding protein. J Biol Chem. 2001;276(45):42520-6.

43.   Rentzsch A, Krummeck-Weiss G, Hofer A, Bartuschka A, Ostermann K, Rodel G. Mitochondrial copper metabolism in yeast: mutational analysis of Sco1p involved in the biogenesis of cytochrome c oxidase. Curr Genet. 1999;35(2):103-8.

44.   Blundell KL, Wilson MT, Svistunenko DA, Vijgenboom E, Worrall JA. Morphological development and cytochrome c oxidase activity in *Streptomyces lividans* are dependent on the action of a copper bound Sco protein. Open Biol. 2013;3(1):120163.

45.   Rivera-Hoyos CM, Morales-Álvarez ED, Poutou-Piñales RA, Pedroza-Rodríguez AM, RodrÍguez-Vázquez R, Delgado-Boada JM. Fungal laccases. Fungal Biology Reviews. 2013;27(3-4):67-82.

46.   Solomon EI, Heppner DE, Johnston EM, Ginsbach JW, Cirera J, Qayyum M, et al. Copper active sites in biology. Chem Rev. 2014;114(7):3659-853.

47.   Hakulinen N, Kiiskinen LL, Kruus K, Saloheimo M, Paananen A, Koivula A, et al. Crystal structure of a laccase from *Melanocarpus albomyces* with an intact trinuclear copper site. Nat Struct Biol. 2002;9(8):601-5.

48.   Rodgers CJ, Blanford CF, Giddens SR, Skamnioti P, Armstrong FA, Gurr SJ. Designer laccases: a vogue for high-potential fungal enzymes? Trends Biotechnol. 2010;28(2):63-72.

49.   Whittaker MM, Whittaker JW. The active site of galactose oxidase. J Biol Chem. 1988;263(13):6074-80.

50.   Hemsworth GR, Johnston EM, Davies GJ, Walton PH. Lytic Polysaccharide Monooxygenases in Biomass Conversion. Trends Biotechnol. 2015;33(12):747-61.

51.   Smith PK, Krohn RI, Hermanson GT, Mallia AK, Gartner FH, Provenzano MD, et al. Measurement of protein using bicinchoninic acid. Analytical Biochemistry. 1985;150(1):76-85.

52.   Huang T, Long M, Huo B. Competitive Binding to Cuprous Ions of Protein and BCA in the Bicinchoninic Acid Protein Assay. Open Biomed Eng J. 2010;4:271-8.

53.   Brenner AJ, Harris ED. A quantitative test for copper using bicinchoninic acid. Anal Biochem. 1995;226(1):80-4.

54.   Xiao Z, Wedd AG. The challenges of determining metal-protein affinities. Nat Prod Rep. 2010;27(5):768-89.

55.   Kenney GE, Rosenzweig AC. Chalkophores. Annu Rev Biochem. 2018;87(1):645-76.

56.   Lieberman RL, Rosenzweig AC. Crystal structure of a membrane-bound metalloenzyme that catalyses the biological oxidation of methane. Nature. 2005;434(7030):177-82.

57.   Rosenzweig AC, Frederick CA, Lippard SJ, Nordlund P. Crystal structure of a bacterial non-haem iron hydroxylase that catalyses the biological oxidation of methane. Nature. 1993;366(6455):537-43.

58.   Nielsen AK, Gerdes K, Murrell JC. Copper-dependent reciprocal transcriptional regulation of methane monooxygenase genes in *Methylococcus capsulatus* and *Methylosinus trichosporium*. Mol Microbiol. 1997;25(2):399-409.

59.   Semrau JD, DiSpirito AA, Yoon S. Methanotrophs and copper. FEMS Microbiol Rev. 2010;34(4):496-531.

60.   Choi DW, Kunz RC, Boyd ES, Semrau JD, Antholine WE, Han JI, et al. The membrane-associated methane monooxygenase (pMMO) and pMMO-NADH:quinone oxidoreductase complex from *Methylococcus capsulatus Bath*. J Bacteriol. 2003;185(19):5755-64.

61.   Dassama LM, Kenney GE, Rosenzweig AC. Methanobactins: from genome to function. Metallomics. 2017;9(1):7-20.

62.   Kim HJ, Graham DW, DiSpirito AA, Alterman MA, Galeva N, Larive CK, et al. Methanobactin, a copper-acquisition compound from methane-oxidizing bacteria. Science. 2004;305(5690):1612-5.

63.   Ho A, Kerckhof FM, Luke C, Reim A, Krause S, Boon N, et al. Conceptualizing functional traits and ecological characteristics of methane-oxidizing bacteria as life strategies. Environmental microbiology reports. 2013;5(3):335-45.

64.   Pesch ML, Hoffmann M, Christl I, Kraemer SM, Kretzschmar R. Competitive ligand exchange between Cu-humic acid complexes and methanobactin. Geobiology. 2013;11(1):44-54.

65.   Behling LA, Hartsel SC, Lewis DE, DiSpirito AA, Choi DW, Masterson LR, et al. NMR, mass spectrometry and chemical evidence reveal a different chemical structure for methanobactin that contains oxazolone rings. J Am Chem Soc. 2008;130(38):12604-5.

66.   Blanco-Lomas M, Funes-Ardoiz I, Campos PJ, Sampedro D. Oxazolone-Based Photoswitches: Synthesis and Properties. European Journal of Organic Chemistry. 2013;2013(29):6611-8.

67.   Kim HJ, Galeva N, Larive CK, Alterman M, Graham DW. Purification and physical-chemical properties of methanobactin: a chalkophore from *Methylosinus trichosporium* OB3b. Biochemistry. 2005;44(13):5140-8.

68.   Krentz BD, Mulheron HJ, Semrau JD, Dispirito AA, Bandow NL, Haft DH, et al. 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(47):10117-30.

69.   Choi DW, Antholine WE, Do YS, Semrau JD, Kisting CJ, Kunz RC, et al. Effect of methanobactin on the activity and electron paramagnetic resonance spectra of the membrane-associated methane monooxygenase in *Methylococcus capsulatus Bath.* Microbiology. 2005;151(Pt 10):3417-26.

70.   Hakemian AS, Tinberg CE, Kondapalli KC, Telser J, Hoffman BM, Stemmler TL, et al. The copper chelator methanobactin from *Methylosinus trichosporium* OB3b binds copper(I). J Am Chem Soc. 2005;127(49):17142-3.

71.   Choi DW, Do YS, Zea CJ, McEllistrem MT, Lee SW, Semrau JD, et al. Spectral and thermodynamic properties of Ag(I), Au(III), Cd(II), Co(II), Fe(III), Hg(II), Mn(II), Ni(II), Pb(II), U(IV), and Zn(II) binding by methanobactin from *Methylosinus trichosporium* OB3b. J Inorg Biochem. 2006;100(12):2150-61.

72.   Choi DW, Semrau JD, Antholine WE, Hartsel SC, Anderson RC, Carey JN, et al. Oxidase, superoxide dismutase, and hydrogen peroxide reductase activities of methanobactin from types I and II methanotrophs. J Inorg Biochem. 2008;102(8):1571-80.

73.   Balasubramanian R, Smith SM, Rawat S, Yatsunyk LA, Stemmler TL, Rosenzweig AC. Oxidation of methane by a biological dicopper centre. Nature. 2010;465(7294):115-9.

74.   Sirajuddin S, Barupala D, Helling S, Marcus K, Stemmler TL, Rosenzweig AC. Effects of zinc on particulate methane monooxygenase activity and structure. J Biol Chem. 2014;289(31):21782-94.

75.   Dennison C, David S, Lee J. Bacterial copper storage proteins. J Biol Chem. 2018;293(13):4616-27.

76.   Basle A, Platsaki S, Dennison C. Visualizing Biological Copper Storage: The Importance of Thiolate-Coordinated Tetranuclear Clusters. Angew Chem Int Ed Engl. 2017;56(30):8697-700.

77.   Chino M, Maglio O, Nastri F, Pavone V, DeGrado WF, Lombardi A. Artificial Diiron Enzymes with a De Novo Designed Four-Helix Bundle Structure. Eur J Inorg Chem. 2015;2015(21):3371-90.

78.   Handel TM, Williams SA, DeGrado WF. Metal ion-dependent modulation of the dynamics of a designed protein. Science. 1993;261(5123):879-85.

79.   Koder RL, Anderson JLR, Solomon LA, Reddy KS, Moser CC, Dutton PL. Design and engineering of an O2 transport protein. Nature. 2009;458:305.

80.   Robertson DE, Farid RS, Moser CC, Urbauer JL, Mulholland SE, Pidikiti R, et al. Design and synthesis of multi-haem proteins. Nature. 1994;368(6470):425-32.

81.   Schnepf R, Horth 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(10):2186-95.

82.   Bentley SD, Chater KF, Cerdeno-Tarraga AM, Challis GL, Thomson NR, James KD, et al. Complete genome sequence of the model actinomycete *Streptomyces coelicolo*r A3(2). Nature. 2002;417(6885):141-7.

83.   Chater KF, Biro S, Lee KJ, Palmer T, Schrempf H. The complex extracellular biology of *Streptomyces*. FEMS Microbiol Rev. 2010;34(2):171-98.

84.   Vrancken K, Anne J. Secretory production of recombinant proteins by *Streptomyces*. Future Microbiol. 2009;4(2):181-8.

85.   Hopwood DA. Forty years of genetics with *Streptomyces*: from *in vivo* through i*n vitro* to *in silico*. Microbiology. 1999;145 (Pt 9)(9):2183-202.

86.   Flardh K, Buttner MJ. *Streptomyces* morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol. 2009;7(1):36-49.

87.   Hopwood D. Antibiotics actions, origins, resistance. AMER ASSOC ADVANCEMENT SCIENCE 1200 NEW YORK AVE, NW, WASHINGTON, DC 20005 USA; 2003.

88.   van Wezel GP, McDowall KJ. The regulation of the secondary metabolism of *Streptomyces*: new links and experimental advances. Nat Prod Rep. 2011;28(7):1311-33.

89.   Huang J, Shi J, Molle V, Sohlberg B, Weaver D, Bibb MJ, et al. Cross-regulation among disparate antibiotic biosynthetic pathways of *Streptomyces coelicolor*. Mol Microbiol. 2005;58(5):1276-87.

90.   Vujaklija D, Horinouchi S, Beppu T. Detection of an A-factor-responsive protein that binds to the upstream activation sequence of strR, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. J Bacteriol. 1993;175(9):2652-61.

91.   Gramajo HC, Takano E, Bibb MJ. Stationary-phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. Molecular Microbiology. 1993;7(6):837-45.

92.   Claessen D, Rozen DE, Kuipers OP, Sogaard-Andersen L, van Wezel GP. Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. Nat Rev Microbiol. 2014;12(2):115-24.

93.   Schaerlaekens K, Van Mellaert L, Lammertyn E, Geukens N, Anne J. The importance of the Tat-dependent protein secretion pathway in *Streptomyces* as revealed by phenotypic changes in tat deletion mutants and genome analysis. Microbiology. 2004;150(Pt 1):21-31.

94.   Berks BC, Palmer T, Sargent F. The Tat protein translocation pathway and its role in microbial physiology. Advances in Microbial Physiology. Volume 47: Academic Press; 2003. p. 187-254.

95.   Widdick DA, Dilks K, Chandra G, Bottrill A, Naldrett M, Pohlschroder M, et al. The twin-arginine translocation pathway is a major route of protein export in *Streptomyces coelicolor*. Proc Natl Acad Sci U S A. 2006;103(47):17927-32.

96.   Jakimowicz D, van Wezel GP. Cell division and DNA segregation in *Streptomyces*: how to build a septum in the middle of nowhere? Mol Microbiol. 2012;85(3):393-404.

97.   McCormick JR. Cell division is dispensable but not irrelevant in *Streptomyces*. Curr Opin Microbiol. 2009;12(6):689-98.

98.   van Dissel D, Claessen D, Roth M, van Wezel GP. A novel locus for mycelial aggregation forms a gateway to improved *Streptomyces* cell factories. Microb Cell Fact. 2015;14(1):44.

99.   Roth M, Noack D, Geuther R. Maintenance of the recombinant plasmid pIJ2 in chemostat cultures of *Streptomyces lividans* 66 (pIJ2). Journal of Basic Microbiology. 1985;25(4):265-71.

100.   Roth M, Hoffmeier C, Geuther R, Muth G, Wohlleben W. Segregational stability of pSG5-derived vector plasmids in continuous cultures of *Streptomyces lividans* 66. Biotechnology Letters. 1994;16(12):1225-30.

101.   Blundell KL, Hough MA, Vijgenboom E, Worrall JA. Structural and mechanistic insights into an extracytoplasmic copper trafficking pathway in *Streptomyces lividans*. Biochem J. 2014;459(3):525-38.

102.   Petrus ML, Claessen D. Pivotal roles for *Streptomyces* cell surface polymers in morphological differentiation, attachment and mycelial architecture. Antonie Van Leeuwenhoek. 2014;106(1):127-39.

103.   Chaplin AK, Petrus ML, Mangiameli G, Hough MA, Svistunenko DA, Nicholls P, et al. GlxA is a new structural member of the radical copper oxidase family and is required for glycan deposition at hyphal tips and morphogenesis of *Streptomyces lividans*. Biochem J. 2015;469(3):433-44.

104.   Coutinho PM, Deleury E, Davies GJ, Henrissat B. An evolving hierarchical family classification for glycosyltransferases. J Mol Biol. 2003;328(2):307-17.

105.   van Veluw GJ, Petrus ML, Gubbens J, de Graaf R, de Jong IP, van Wezel GP, et al. Analysis of two distinct mycelial populations in liquid-grown *Streptomyces* cultures using a flow cytometry-based proteomics approach. Appl Microbiol Biotechnol. 2012;96(5):1301-12.

106.   Xu H, Chater KF, Deng Z, Tao M. A cellulose synthase-like protein involved in hyphal tip growth and morphological differentiation in streptomyces. J Bacteriol. 2008;190(14):4971-8.

107.   Liman R, Facey PD, van Keulen G, Dyson PJ, Del Sol R. A laterally acquired galactose oxidase-like gene is required for aerial development during osmotic stress in *Streptomyces coelicolor*. PLoS One. 2013;8(1):e54112.

108.   Porto TV, Hough MA, Worrall JA. Structural insights into conformational switching in the copper metalloregulator CsoR from *Streptomyces lividans*. Acta Crystallogr D Biol Crystallogr. 2015;71(Pt 9):1872-8.

109.   Hung IH, Casareno RL, Labesse G, Mathews FS, Gitlin JD. HAH1 is a copper-binding protein with distinct amino acid residues mediating copper homeostasis and antioxidant defense. J Biol Chem. 1998;273(3):1749-54.

110.   Klomp LW, Lin SJ, Yuan DS, Klausner RD, Culotta VC, Gitlin JD. Identification and functional expression of HAH1, a novel human gene involved in copper homeostasis. J Biol Chem. 1997;272(14):9221-6.

111.   Wernimont AK, Huffman DL, Lamb AL, O'Halloran TV, Rosenzweig AC. Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. Nat Struct Biol. 2000;7(9):766-71.

112.   Porto TV, Wilson MT, Worrall JA. Copper and nickel bind via two distinct kinetic mechanisms to a CsoR metalloregulator. Dalton Trans. 2015;44(46):20176-85.

113.   Thaden JT, Lory S, Gardner TS. Quorum-sensing regulation of a copper toxicity system in *Pseudomonas aeruginosa.* J Bacteriol. 2010;192(10):2557-68.

114.   Novoa-Aponte L, Ramirez D, Arguello JM. The interplay of the metallosensor CueR with two distinct CopZ chaperones defines copper homeostasis in *Pseudomonas aeruginosa.* J Biol Chem. 2019;294(13):4934-45.

115.   Quintana J, Novoa-Aponte L, Arguello JM. Copper homeostasis networks in the bacterium *Pseudomonas aeruginosa*. J Biol Chem. 2017;292(38):15691-704.
116.   Gonzalez-Quinonez N, Corte-Rodriguez M, Alvarez-Fernandez-Garcia R, Rioseras B, Lopez-Garcia MT, Fernandez-Garcia G, et al. Cytosolic copper is a major modulator of germination, development and secondary metabolism in *Streptomyces coelicolor*. Sci Rep. 2019;9(1):4214.
117.   Chaplin AK, Wilson MT, Hough MA, Svistunenko DA, Hemsworth GR, Walton PH, et al. Heterogeneity in the Histidine-brace Copper Coordination Sphere in Auxiliary Activity Family 10 (AA10) Lytic Polysaccharide Monooxygenases. J Biol Chem. 2016;291(24):12838-50.
118.   Blundell KL, Wilson MT, Vijgenboom E, Worrall JA. The role of the Cys-X-X-X-Cys motif on the kinetics of cupric ion loading to the *Streptomyces lividans* Sco protein. Dalton Trans. 2013;42(29):10608-16.
119.   Tan BG, Vijgenboom E, Worrall JA. Conformational and thermodynamic hallmarks of DNA operator site specificity in the copper sensitive operon repressor from *Streptomyces lividans*. Nucleic Acids Res. 2014;42(2):1326-40.
120.   Liu T, Ramesh A, Ma Z, Ward SK, Zhang L, George GN, et al. CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. Nat Chem Biol. 2007;3(1):60-8.
121.   Woese CR, Fox GE. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc Natl Acad Sci U S A. 1977;74(11):5088-90.
122.   Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87(12):4576-9.
123.   Williams TA, Foster PG, Nye TM, Cox CJ, Embley TM. A congruent phylogenomic signal places eukaryotes within the Archaea. Proc Biol Sci. 2012;279(1749):4870-9.
124.   Williams TA, Foster PG, Cox CJ, Embley TM. An archaeal origin of eukaryotes supports only two primary domains of life. Nature. 2013;504(7479):231-6.
125.   Eisen JA. Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. Current Opinion in Genetics & Development. 2000;10(6):606-11.
126.   Keeling PJ, Palmer JD. Horizontal gene transfer in eukaryotic evolution. Nat Rev Genet. 2008;9(8):605-18.
127.   Hakemian AS, Rosenzweig AC. The biochemistry of methane oxidation. Annu Rev Biochem. 2007;76(1):223-41.
128.   Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403-10.
129.   Johnson M, Zaretskaya I, Raytselis Y, Merezhuk Y, McGinnis S, Madden TL. NCBI BLAST: a better web interface. Nucleic Acids Res. 2008;36(Web Server issue):W5-9.
130.   Sayers EW, Barrett T, Benson DA, Bryant SH, Canese K, Chetvernin V, et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Res. 2009;37(Database issue):D5-15.
131.   Castresana J. Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis. Molecular Biology and Evolution. 2000;17(4):540-52.
132.   Price MN, Dehal PS, Arkin AP. FastTree 2--approximately maximum-likelihood trees for large alignments. PLoS One. 2010;5(3):e9490.
133.   Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol. 2001;18(5):691-9.

134.   Shimodaira H, Hasegawa M. Multiple Comparisons of Log-Likelihoods with Applications to Phylogenetic Inference. Molecular Biology and Evolution. 1999;16(8):1114-6.

135.   Cruz-Morales P, Vijgenboom E, Iruegas-Bocardo F, Girard G, Yanez-Guerra LA, Ramos-Aboites HE, et al. The genome sequence of *Streptomyces lividans* 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. Genome Biol Evol. 2013;5(6):1165-75.

136.   Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. Nucleic Acids Res. 2013;41(Database issue):D36-42.

137.   Slotboom DJ. Structural and mechanistic insights into prokaryotic energy-coupling factor transporters. Nat Rev Microbiol. 2014;12(2):79-87.

138.   Ehrnstorfer IA, Geertsma ER, Pardon E, Steyaert J, Dutzler R. Crystal structure of a SLC11 (NRAMP) transporter reveals the basis for transition-metal ion transport. Nat Struct Mol Biol. 2014;21(11):990-6.

139.   Ehrnstorfer IA, Manatschal C, Arnold FM, Laederach J, Dutzler R. Structural and mechanistic basis of proton-coupled metal ion transport in the SLC11/NRAMP family. Nat Commun. 2017;8:14033.

140.   Sutherland D, Stillman M. The "magic numbers" of metallothionein2011. 444-63 p.

141.   Fujimoto M, Yamada A, Kurosawa J, Kawata A, Beppu T, Takano H, et al. Pleiotropic role of the Sco1/SenC family copper chaperone in the physiology of *Streptomyces*. Microb Biotechnol. 2012;5(4):477-88.

142.   Keijser BJ, van Wezel GP, Canters GW, Kieser T, Vijgenboom E. The ram-dependence of *Streptomyces lividans* differentiation is bypassed by copper. Journal of molecular microbiology and biotechnology. 2000;2(4):565-74.

143.   Anne J, Maldonado B, Van Impe J, Van Mellaert L, Bernaerts K. Recombinant protein production and streptomycetes. J Biotechnol. 2012;158(4):159-67.

144.   Foster AW, Osman D, Robinson NJ. Metal preferences and metallation. J Biol Chem. 2014;289(41):28095-103.

145.   Waldron KJ, Robinson NJ. How do bacterial cells ensure that metalloproteins get the correct metal? Nat Rev Micro. 2009;7(1):25-35.

146.   Ma Z, Jacobsen FE, Giedroc DP. Coordination chemistry of bacterial metal transport and sensing. Chem Rev. 2009;109(10):4644-81.

147.   Xiao Z, Donnelly PS, Zimmermann M, Wedd AG. Transfer of copper between bis(thiosemicarbazone) ligands and intracellular copper-binding proteins. insights into mechanisms of copper uptake and hypoxia selectivity. Inorganic chemistry. 2008;47(10):4338-47.

148.   Kabsch W. Integration, scaling, space-group assignment and post-refinement. Acta Crystallogr D Biol Crystallogr. 2010;66(Pt 2):133-44.

149.   Evans PR, Murshudov GN. How good are my data and what is the resolution? Acta Crystallogr D Biol Crystallogr. 2013;69(Pt 7):1204-14.

150.   McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007;40(Pt 4):658-74.

151.   Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr. 2006;62(Pt 9):1002-11.

152.   Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr. 2004;60(Pt 12 Pt 1):2126-32.

153.   Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr. 1997;53(Pt 3):240-55.

154.   Davis IW, Leaver-Fay A, Chen VB, Block JN, Kapral GJ, Wang X, et al. MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. 2007;35(Web Server issue):W375-83.

155.   Krissinel E. Enhanced fold recognition using efficient short fragment clustering. Journal of molecular biochemistry. 2012;1(2):76-85.

156.   Larson JL, Hershberger CL. The minimal replicon of a streptomycete plasmid produces an ultrahigh level of plasmid DNA. Plasmid. 1986;15(3):199-209.

157.   Green GN, Gennis RB. Isolation and characterization of an *Escherichia coli* mutant lacking cytochrome d terminal oxidase. J Bacteriol. 1983;154(3):1269-75.

158.   Mueller JP, Taber HW. Isolation and sequence of ctaA, a gene required for cytochrome aa3 biosynthesis and sporulation in *Bacillus subtilis*. J Bacteriol. 1989;171(9):4967-78.

159.   Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. Nat Methods. 2012;9(7):671-5.

160.   Xiao Z, Brose J, Schimo S, Ackland SM, La Fontaine S, Wedd AG. 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(13):11047-55.

161.   Morgan MT, Bagchi P, Fahrni CJ. Designed to dissolve: suppression of colloidal aggregation of Cu(I)-selective fluorescent probes in aqueous buffer and in-gel detection of a metallochaperone. J Am Chem Soc. 2011;133(40):15906-9.

162.   Rosenzweig AC, Huffman DL, Hou MY, Wernimont AK, Pufahl RA, O'Halloran TV. Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. Structure. 1999;7(6):605-17.

163.   Kay KL, Zhou L, Tenori L, Bradley JM, Singleton C, Kihlken MA, et al. Kinetic analysis of copper transfer from a chaperone to its target protein mediated by complex formation. Chem Commun (Camb). 2017;53(8):1397-400.

164.   Cobine PA, George GN, Jones CE, Wickramasinghe WA, Solioz M, Dameron CT. Copper transfer from the Cu(I) chaperone, CopZ, to the repressor, Zn(II)CopY: metal coordination environments and protein interactions. Biochemistry. 2002;41(18):5822-9.

165.   Hearnshaw S, West C, Singleton C, Zhou L, Kihlken MA, Strange RW, et al. A tetranuclear Cu(I) cluster in the metallochaperone protein CopZ. Biochemistry. 2009;48(40):9324-6.

166.   Kay KL, Hamilton CJ, Le Brun NE. Mass spectrometry of *B. subtilis* CopZ: Cu(i)-binding and interactions with bacillithiol. Metallomics. 2016;8(7):709-19.

167.   Kihlken MA, Leech AP, Le Brun NE. Copper-mediated dimerization of CopZ, a predicted copper chaperone from *Bacillus subtilis*. Biochem J. 2002;368(Pt 3):729-39.

168.   Banci L, Bertini I, Ciofi-Baffoni S, Kozyreva T, Zovo K, Palumaa P. Affinity gradients drive copper to cellular destinations. Nature. 2010;465(7298):645-8.

169.   Dennison C. The Coordination Chemistry of Copper Uptake and Storage for Methane Oxidation. Chemistry. 2019;25(1):74-86.

170.   Straw ML, Hough MA, Wilson MT, Worrall JAR. A histidine residue and a tetranuclear cuprous-thiolate cluster dominate the copper loading landscape of a copper storage protein from *Streptomyces lividans*. Chemistry. 2019;0(ja).

171.   Richens DT. Ligand substitution reactions at inorganic centers. Chem Rev. 2005;105(6):1961-2002.

172.   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(9):2983-8.

173.   Banci L, Bertini I, Cantini F, Felli IC, Gonnelli L, Hadjiliadis N, et al. The Atx1-Ccc2 complex is a metal-mediated protein-protein interaction. Nature Chemical Biology. 2006;2:367.

174.   Rae TD, Schmidt PJ, Pufahl RA, Culotta VC, O'Halloran TV. Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. Science. 1999;284(5415):805-8.

175. Boal AK, Rosenzweig AC. Structural biology of copper trafficking. Chem Rev. 2009;109(10):4760-79.

176. Huffman DL, O'Halloran TV. Energetics of copper trafficking between the Atx1 metallochaperone and the intracellular copper transporter, Ccc2. J Biol Chem. 2000;275(25):18611-4.

177. Chaplin AK, Svistunenko DA, Hough MA, Wilson MT, Vijgenboom E, Worrall JA. Active-site maturation and activity of the copper-radical oxidase GlxA are governed by a tryptophan residue. Biochem J. 2017;474(5):809-25.

178. Arnesano F, Banci L, Bertini I, Huffman DL, O'Halloran TV. Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. Biochemistry. 2001;40(6):1528-39.

179. DiSpirito AA, Semrau JD, Murrell JC, Gallagher WH, Dennison C, Vuilleumier S. Methanobactin and the Link between Copper and Bacterial Methane Oxidation. Microbiol Mol Biol Rev. 2016;80(2):387-409.

180. Yoon S, Kraemer SM, Dispirito AA, Semrau JD. An assay for screening microbial cultures for chalkophore production. Environmental microbiology reports. 2010;2(2):295-303.

181. He Y, Wang K, Yan N. The recombinant expression systems for structure determination of eukaryotic membrane proteins. Protein Cell. 2014;5(9):658-72.

182. Sachse R, Dondapati SK, Fenz SF, Schmidt T, Kubick S. Membrane protein synthesis in cell-free systems: from bio-mimetic systems to bio-membranes. FEBS Lett. 2014;588(17):2774-81.

183. Carpenter EP, Beis K, Cameron AD, Iwata S. Overcoming the challenges of membrane protein crystallography. Curr Opin Struct Biol. 2008;18(5):581-6.

184. le Maire M, Champeil P, Moller JV. Interaction of membrane proteins and lipids with solubilizing detergents. Biochim Biophys Acta. 2000;1508(1-2):86-111.

185. Seddon AM, Curnow P, Booth PJ. Membrane proteins, lipids and detergents: not just a soap opera. Biochim Biophys Acta. 2004;1666(1-2):105-17.

186. Focke PJ, Hein C, Hoffmann B, Matulef K, Bernhard F, Dotsch V, et al. Combining in Vitro Folding with Cell Free Protein Synthesis for Membrane Protein Expression. Biochemistry. 2016;55(30):4212-9.

187. Yildirim MA, Goh KI, Cusick ME, Barabasi AL, Vidal M. Drug-target network. Nat Biotechnol. 2007;25(10):1119-26.

188. Lee SC, Knowles TJ, Postis VL, Jamshad M, Parslow RA, Lin YP, et al. A method for detergent-free isolation of membrane proteins in their local lipid environment. Nat Protoc. 2016;11(7):1149-62.

189. Charalambous K, Miller D, Curnow P, Booth PJ. Lipid bilayer composition influences small multidrug transporters. BMC Biochem. 2008;9:31.

190. Debnath DK, Basaiawmoit RV, Nielsen KL, Otzen DE. The role of membrane properties in Mistic folding and dimerisation. Protein Eng Des Sel. 2011;24(1-2):89-97.

191. Jamshad M, Charlton J, Lin YP, Routledge SJ, Bawa Z, Knowles TJ, et al. G-protein coupled receptor solubilization and purification for biophysical analysis and functional studies, in the total absence of detergent. Bioscience reports. 2015;35(2).

192. Knowles TJ, Finka R, Smith C, Lin YP, Dafforn T, Overduin M. Membrane proteins solubilized intact in lipid containing nanoparticles bounded by styrene maleic acid copolymer. J Am Chem Soc. 2009;131(22):7484-5.

193. Jamshad M, Grimard V, Idini I, Knowles TJ, Dowle MR, Schofield N, et al. Structural analysis of a nanoparticle containing a lipid bilayer used for detergent-free extraction of membrane proteins. Nano Res. 2015;8(3):774-89.

194.   Scheidelaar S, Koorengevel MC, Pardo JD, Meeldijk JD, Breukink E, Killian JA. Molecular model for the solubilization of membranes into nanodisks by styrene maleic Acid copolymers. Biophysical journal. 2015;108(2):279-90.

195.   Schlegel S, Hjelm A, Baumgarten T, Vikstrom D, de Gier JW. Bacterial-based membrane protein production. Biochim Biophys Acta. 2014;1843(8):1739-49.

196.   Wallin E, von Heijne G. Genome-wide analysis of integral membrane proteins from eubacterial, archaean, and eukaryotic organisms. Protein Sci. 1998;7(4):1029-38.

197.   Gavel Y, Steppuhn J, Herrmann R, von Heijne G. The 'positive-inside rule' applies to thylakoid membrane proteins. FEBS Letters. 1991;282(1):41-6.

198.   Gavel Y, von Heijne G. The distribution of charged amino acids in mitochondrial inner-membrane proteins suggests different modes of membrane integration for nuclearly and mitochondrially encoded proteins. Eur J Biochem. 1992;205(3):1207-15.

199.   Birmingham Uo. Protein/Peptide Identification 2018 [updated 2018. Available from: https://www.birmingham.ac.uk/facilities/advanced-mass-spectrometry/services/protein-peptide-identification.aspx.

200.   Prive GG. Detergents for the stabilization and crystallization of membrane proteins. Methods. 2007;41(4):388-97.

201.   Wagner S, Bader ML, Drew D, de Gier JW. Rationalizing membrane protein overexpression. Trends Biotechnol. 2006;24(8):364-71.

202.   Rath A, Glibowicka M, Nadeau VG, Chen G, Deber CM. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. Proc Natl Acad Sci U S A. 2009;106(6):1760-5.

203.   Schaiberger AM, Moss JA. Optimized sample preparation for MALDI mass spectrometry analysis of protected synthetic peptides. J Am Soc Mass Spectrom. 2008;19(4):614-9.

204.   Baldwin MA. Protein identification by mass spectrometry: issues to be considered. Mol Cell Proteomics. 2004;3(1):1-9.

205.   Bannwarth M, Schulz GE. The expression of outer membrane proteins for crystallization. Biochim Biophys Acta. 2003;1610(1):37-45.

206.   Miroux B, Walker JE. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. J Mol Biol. 1996;260(3):289-98.

207.   Wagner S, Klepsch MM, Schlegel S, Appel A, Draheim R, Tarry M, et al. Tuning *Escherichia coli* for membrane protein overexpression. Proc Natl Acad Sci U S A. 2008;105(38):14371-6.

208.   Arditti RR, Scaife JG, Beckwith JR. The nature of mutants in the lac promoter region. J Mol Biol. 1968;38(3):421-6.

209.   Schlegel S, Lofblom J, Lee C, Hjelm A, Klepsch M, Strous M, et al. Optimizing membrane protein overexpression in the *Escherichia coli* strain Lemo21(DE3). J Mol Biol. 2012;423(4):648-59.

210.   Wagner S, Baars L, Ytterberg AJ, Klussmeier A, Wagner CS, Nord O, et al. Consequences of membrane protein overexpression in *Escherichia coli*. Mol Cell Proteomics. 2007;6(9):1527-50.

211.   Byrne B. Pichia pastoris as an expression host for membrane protein structural biology. Curr Opin Struct Biol. 2015;32:9-17.

212.   Gimpl K, Klement J, Keller S. Characterising protein/detergent complexes by triple-detection size-exclusion chromatography. Biol Proced Online. 2016;18:4.

213.   Korepanova A, Gao FP, Hua Y, Qin H, Nakamoto RK, Cross TA. Cloning and expression of multiple integral membrane proteins from *Mycobacterium tuberculosis* in *Escherichia coli*. Protein Sci. 2005;14(1):148-58.

214.   Lee CH, Chou CC, Hsu MF, Wang AH. Determining the N-terminal orientations of recombinant transmembrane proteins in the *Escherichia coli* plasma membrane. Sci Rep. 2015;5:15086.

215.   Daley DO, Rapp M, Granseth E, Melen K, Drew D, von Heijne G. Global topology analysis of the *Escherichia coli* inner membrane proteome. Science. 2005;308(5726):1321-3.

216.   Lee H, Kim H. Membrane topology of transmembrane proteins: determinants and experimental tools. Biochem Biophys Res Commun. 2014;453(2):268-76.

217.   van Geest M, Lolkema JS. Membrane topology and insertion of membrane proteins: search for topogenic signals. Microbiol Mol Biol Rev. 2000;64(1):13-33.

218.   Kunji ER, Harding M, Butler PJ, Akamine P. Determination of the molecular mass and dimensions of membrane proteins by size exclusion chromatography. Methods. 2008;46(2):62-72.

219.   Gast K, Johannes Modler A. Studying protein folding and aggregation by LASER light scattering2005.

220.   Murphy RM. Static and dynamic light scattering of biological macromolecules: what can we learn? Curr Opin Biotechnol. 1997;8(1):25-30.

221.   Deisenhofer J, Epp O, Miki K, Huber R, Michel H. Structure of the protein subunits in the photosynthetic reaction centre of *Rhodopseudomonas viridis* at 3A resolution. Nature. 1985;318(6047):618-24.

222.   Cherezov V. Lipidic cubic phase technologies for membrane protein structural studies. Curr Opin Struct Biol. 2011;21(4):559-66.

223.   Gaisford W, Schertler G, Edwards P. mosquito® LCP: Making membrane protein crystallization accessible to the research scientist. Nature Methods. 2011;8(6):i-ii.

224.   Landau EM, Rosenbusch JP. Lipidic cubic phases: a novel concept for the crystallization of membrane proteins. Proc Natl Acad Sci U S A. 1996;93(25):14532-5.

225.   Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, et al. High-resolution crystal structure of an engineered human beta2-adrenergic G protein-coupled receptor. Science. 2007;318(5854):1258-65.

226.   Chien EY, Liu W, Zhao Q, Katritch V, Han GW, Hanson MA, et al. Structure of the human dopamine D3 receptor in complex with a D2/D3 selective antagonist. Science. 2010;330(6007):1091-5.

227.   Jaakola VP, Griffith MT, Hanson MA, Cherezov V, Chien EY, Lane JR, et al. The 2.6 angstrom crystal structure of a human A2A adenosine receptor bound to an antagonist. Science. 2008;322(5905):1211-7.

228.   Shimamura T, Shiroishi M, Weyand S, Tsujimoto H, Winter G, Katritch V, et al. Structure of the human histamine H1 receptor complex with doxepin. Nature. 2011;475(7354):65-70.

229.   Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. Science. 2010;330(6007):1066-71.

230.   Banci L, Bertini I, Calderone V, Della-Malva N, Felli IC, Neri S, et al. Copper(I)-mediated protein-protein interactions result from suboptimal interaction surfaces. Biochem J. 2009;422(1):37-42.

231.   Koh EI, Henderson JP. Microbial Copper-binding Siderophores at the Host-Pathogen Interface. J Biol Chem. 2015;290(31):18967-74.

232.   Arguello JM, Raimunda D, Gonzalez-Guerrero M. Metal transport across biomembranes: emerging models for a distinct chemistry. J Biol Chem. 2012;287(17):13510-7.
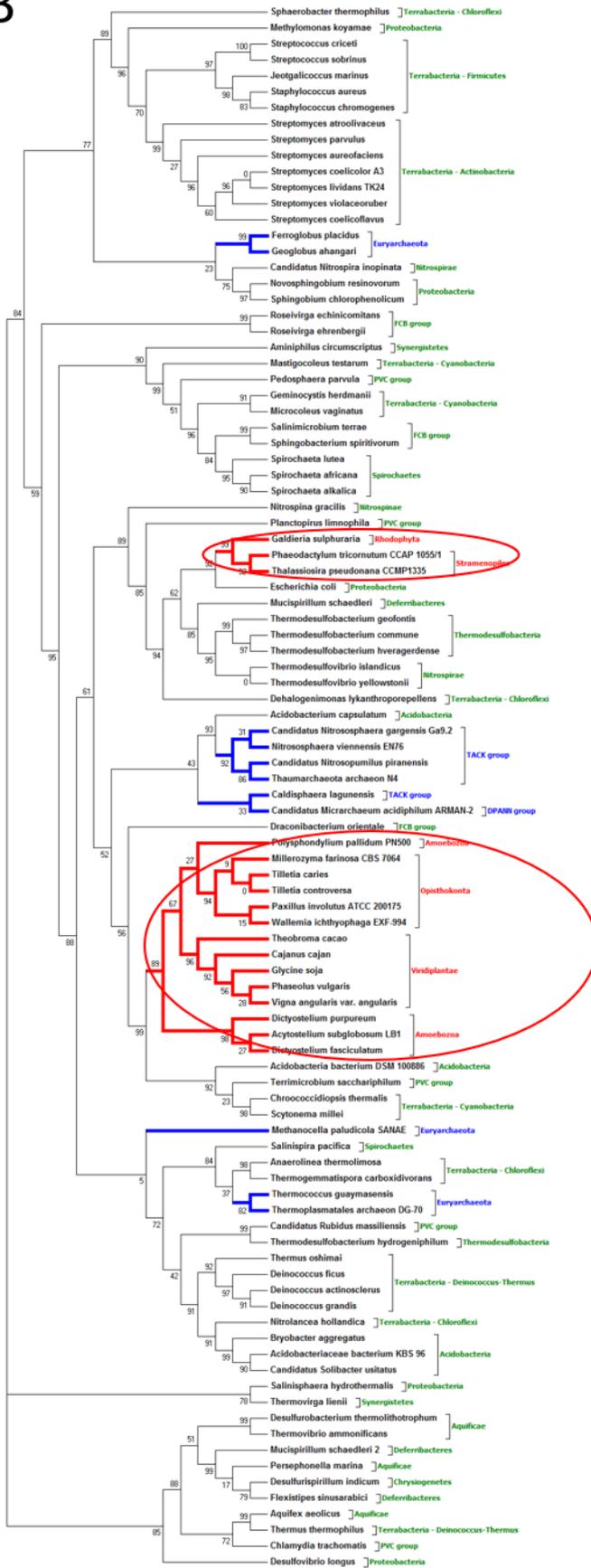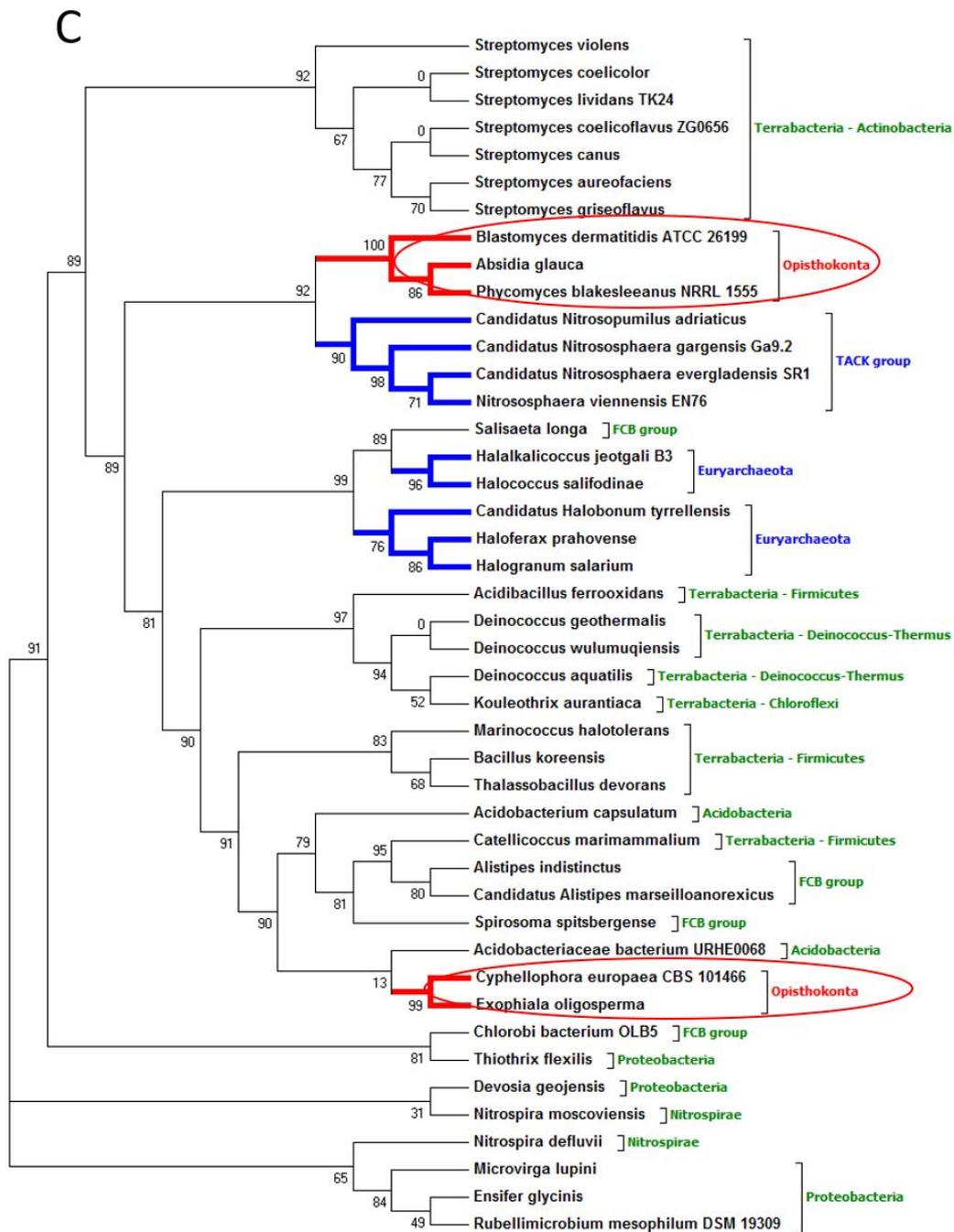
Appendix

Appendix 1

A

B

Sphaerobacter thermophilus `Terrabacteria - Chloroflexi`
Methylomonas koyamae `Proteobacteria`
Streptococcus criceti
Streptococcus sobrinus
Jeotgalicoccus marinus `Terrabacteria - Firmicutes`
Staphylococcus aureus
Staphylococcus chromogenes
Streptomyces atroolivaceus
Streptomyces parvulus
Streptomyces aureofaciens
Streptomyces coelicolor A3 `Terrabacteria - Actinobacteria`
Streptomyces lividans TK24
Streptomyces violaceoruber
Streptomyces coelicoflavus
Ferroglobus placidus `Euryarchaeota`
Geoglobus ahangari
Candidatus Nitrospira inopinata `Nitrospirae`
Novosphingobium resinovorum `Proteobacteria`
Sphingobium chlorophenolicum
Roseivirga echinicomitans `FCB group`
Roseivirga ehrenbergii
Aminiphilus circumscriptus `Synergistetes`
Mastigocoleus testarum `Terrabacteria - Cyanobacteria`
Pedosphaera parvula `PVC group`
Geminocystis herdmanii `Terrabacteria - Cyanobacteria`
Microcoleus vaginatus
Salinimicrobium terrae `FCB group`
Sphingobacterium spiritivorum
Spirochaeta lutea
Spirochaeta africana `Spirochaetes`
Spirochaeta alkalica
Nitrospina gracilis `Nitrospinae`
Planctopirus limnophila `PVC group`
Galdieria sulphuraria `Rhodophyta`
Phaeodactylum tricornutum CCAP 1055/1 `Stramenopila`
Thalassiosira pseudonana CCMP1335
Escherichia coli `Proteobacteria`
Mucispirillum schaedleri `Deferribacteres`
Thermodesulfobacterium geofontis
Thermodesulfobacterium commune `Thermodesulfobacteria`
Thermodesulfobacterium hveragerdense
Thermodesulfovibrio islandicus `Nitrospirae`
Thermodesulfovibrio yellowstonii
Dehalogenimonas lykanthroporepellens `Terrabacteria - Chloroflexi`
Acidobacterium capsulatum `Acidobacteria`
Candidatus Nitrososphaera gargensis Ga9.2
Nitrososphaera viennensis EN76
Candidatus Nitrosopumilus piranensis `TACK group`
Thaumarchaeota archaeon N4
Caldisphaera lagunensis `TACK group`
Candidatus Micrarchaeum acidiphilum ARMAN-2 `DPANN group`
Draconibacterium orientale `FCB group`
Polysphondylium pallidum PN500 `Amoebozoa`
Millerozyma farinosa CBS 7064
Tilletia caries
Tilletia controversa `Opisthokonta`
Paxillus involutus ATCC 200175
Wallemia ichthyophaga EXF-994
Theobroma cacao
Cajanus cajan
Glycine soja `Viridiplantae`
Phaseolus vulgaris
Vigna angularis var. angularis
Dictyostelium purpureum
Acytostelium subglobosum LB1 `Amoebozoa`
Dictyostelium fasciculatum
Acidobacteria bacterium DSM 100886 `Acidobacteria`
Terrimicrobium sacchariphilum `PVC group`
Chroococcidiopsis thermalis `Terrabacteria - Cyanobacteria`
Scytonema millei
Methanocella paludicola SANAE `Euryarchaeota`
Salinispira pacifica `Spirochaetes`
Anaerolinea thermolimosa
Thermogemmatispora carboxidivorans `Terrabacteria - Chloroflexi`
Thermococcus guaymasensis `Euryarchaeota`
Thermoplasmatales archaeon DG-70
Candidatus Rubidus massiliensis `PVC group`
Thermodesulfobacterium hydrogeniphilum `Thermodesulfobacteria`
Thermus oshimai
Deinococcus ficus
Deinococcus actinosclerus `Terrabacteria - Deinococcus-Thermus`
Deinococcus grandis
Nitrolancea hollandica `Terrabacteria - Chloroflexi`
Bryobacter aggregatus
Acidobacteriaceae bacterium KBS 96 `Acidobacteria`
Candidatus Solibacter usitatus
Salinisphaera hydrothermalis `Proteobacteria`
Thermovirga lienii `Synergistetes`
Desulfurobacterium thermolithotrophum `Aquificae`
Thermovibrio ammonificans
Mucispirillum schaedleri 2 `Deferribacteres`
Persephonella marina `Aquificae`
Desulfurispirillum indicum `Chrysiogenetes`
Flexistipes sinusarabici `Deferribacteres`
Aquifex aeolicus `Aquificae`
Thermus thermophilus `Terrabacteria - Deinococcus-Thermus`
Chlamydia trachomatis `PVC group`
Desulfovibrio longus `Proteobacteria`

140

Appendix 1.1 - Phylogenetic trees of maximum likelihood **(A)** *Sl*Csp3, **(B)** SLI_RS17245 and **(C)** SLI_RS17250. The Bootstrap values are indicated at each node of all trees and the phyla of the three main kingdoms have been colour coded (green – Bacteria; red – Eukaryotes; blue – Archaea) – the branches of Eukaryotes have been highlighted also.

Appendix 2

Appendix 2.1: Derivation of Equation 3.2

L_t = total concentration of BCA added

L_f = free BCA ligand

$Cu(L)_2$ = BCA-Cu(I) complex

Cu_t^+ = total concentration of Cu(I) added

Cu_f^+ = free (unbound) Cu(I)

CuP = Ccsp Cu(I) bound protein

S_t = total concentration of copper binding sites in Ccsp

S_f = sites on Ccsp that are unoccupied with Cu(I)

Cu^+S = binding sites in Ccsp occupied by Cu(I)

P_t = total concentration of Ccsp protein

P = apo Ccsp protein

$[L_t] = [L_f] + 2[Cu(L)_2]$

$[Cu_t^+] = [CuP] + [Cu(L)_2]$

$[S_t] = [S_f] + [Cu^+S]$

$[P_t] = [P] + [CuP]$

$$2L_f + Cu_f^+ \longleftrightarrow Cu(L)_2 \qquad\qquad K_{BCA} = \frac{[L_f]^2\,[Cu_f^+]}{[Cu(L)_2]}$$

$$S_f + Cu_f^+ \longleftrightarrow Cu^+S \qquad\qquad K_{Cu} = \frac{[S_f][Cu_f^+]}{[Cu^+S]}$$

$$[Cu_f^+] = \frac{K_{BCA}\,[Cu(L)_2]}{[L_f]^2} = \frac{K_{Cu}\,[Cu^+S]}{[S_f]}$$
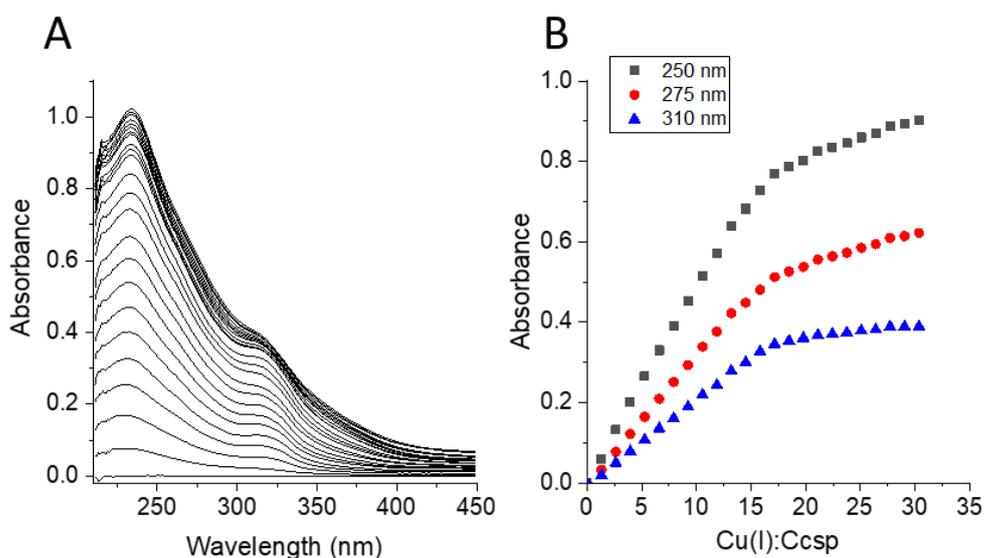
$[L_f]^2 = ([L_t] - 2\,[Cu(L)_2])^2$

$[S_f] = [S_t] - [Cu^+S] + [Cu(L)_2]$

$[Cu^+S] = [Cu_t^+] - [Cu(L)_2]$

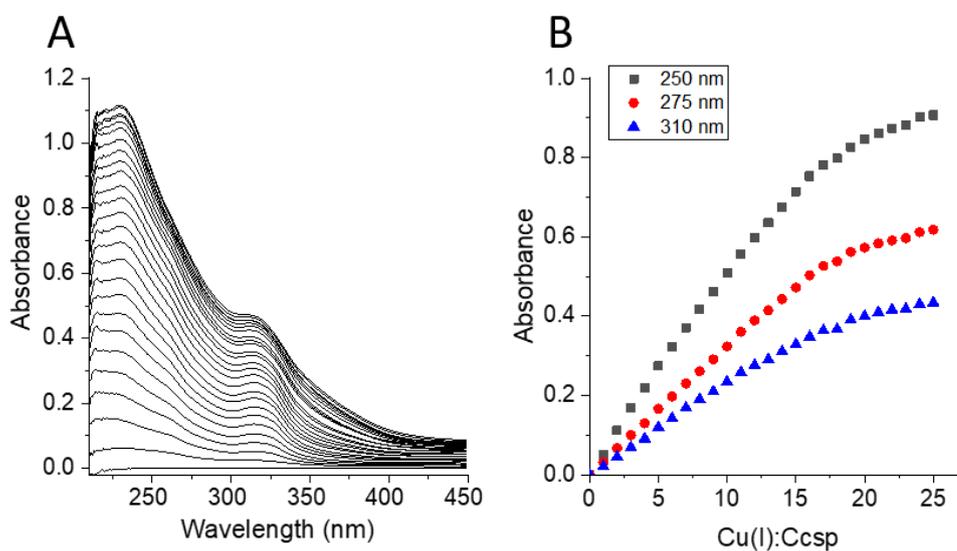$$K_{Cu} = \frac{K_{BCA}\,[Cu(L)_2]}{[L_f]^2} \times \frac{[S_t]}{[Cu^+S]}$$
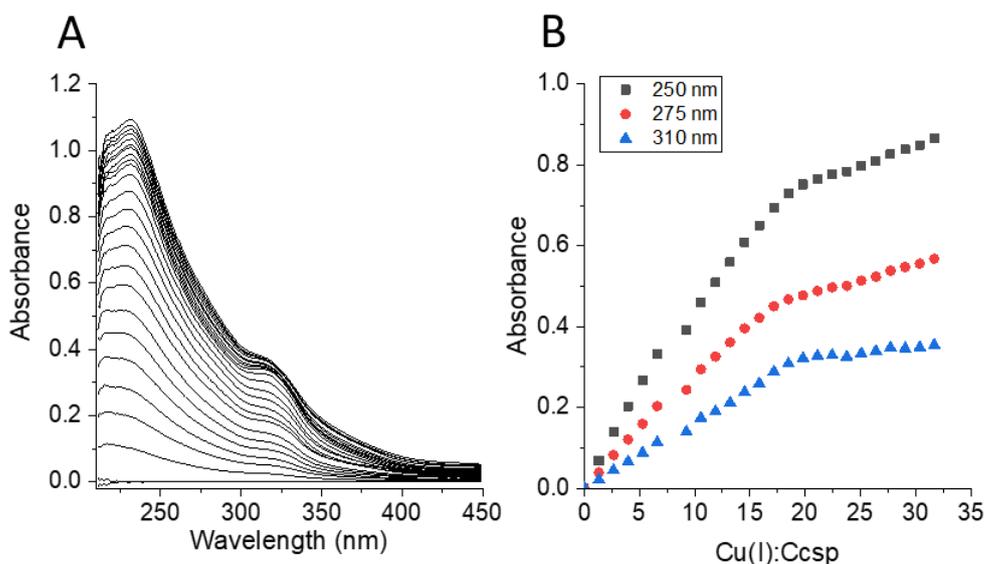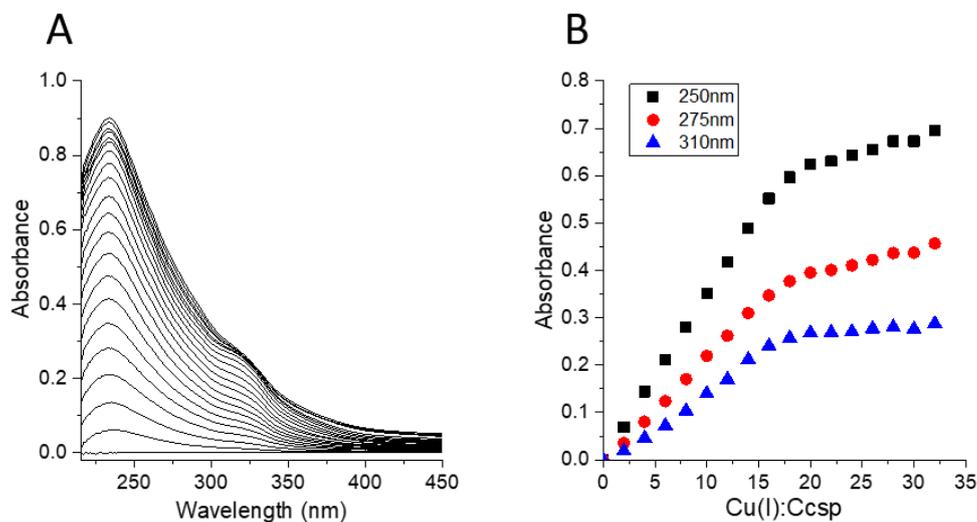
H107A

H111A

Appendix 3.1 - Cu(I) titration to the *Sl*Csp3 His variants H107A and H111A. A) UV-vis difference spectra upon titration of a stock solution of CuCl to 5-8 µM of the proteins revealing the appearance of (Cys)Sy→Cu(I) LMCT bands. B) Plots of absorbance versus the Cu(I):*Sl*Csp3 (Ccsp) concentration ratio at selected wavelengths taken from (A). A break point in the absorbance is reached at ~18-20 Cu(I) equivalents. Experiments were performed at 20 ^{o}C in 10 mM MOPS pH 7.5, 150 mM NaCl.
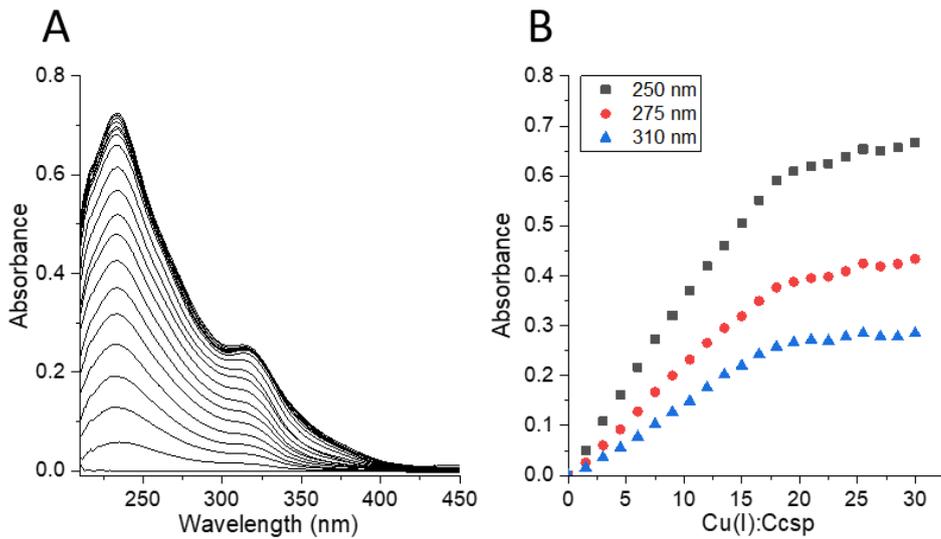
H113A

C41S

Appendix 3.2 - Cu(I) titration to the *Sl*Csp3 variants H113A and C41S. A) UV-vis difference spectra upon titration of a stock solution of CuCl to 5-8 µM of the proteins revealing the appearance of (Cys)Sy→Cu(I) LMCT bands. B) Plots of absorbance versus the Cu(I):*Sl*Csp3 (Ccsp)concentration ratio at selected wavelengths taken from (A). A break point in the absorbance is reached at ~18-20 Cu(I) equivalents. Experiments were performed at 20 °C in 10 mM MOPS pH 7.5, 150 mM NaCl.

C57S

Appendix 3.3 - Cu(I) titration to the *Sl*Csp3 variant C57S. A) UV-vis difference spectrum upon titration of a stock solution of CuCl to 5-8 µM of the protein revealing the appearance of (Cys)Sγ→Cu(I) LMCT bands. B) Plots of absorbance versus the Cu(I):*Sl*Csp3 (Ccsp) concentration ratio at selected wavelengths taken from (A). A break point in the absorbance is reached at ~18-20 Cu(I) equivalents. Experiments were performed at 20 °C in 10 mM MOPS pH 7.5, 150 mM NaCl.

Appendix 4

SLI_RS17250_pET21a (C-terminal (His)₆-tag)

CATATGGACCACAGCACGCACCACTCCGCCCCCGAAGACGACGGGCCCGGACACCCCCACGGGCACGGG
CACCTCCACGGGCCCGGACACCCTCACGGGCACGTGCACGTGCACGGCACCACCTGGGCGACCGCCATGC
AGGCGACGCTGCACTGCCTCACCGGGTGCGCCATCGGCGAGATCCTCGGCATGGTCATCGGAACCGCGCT
GATGTGGGGCAACGTGCCGACCATGGTGCTGGCCATCGCCCTGGCCTTCGTCTTCGGCTACTCCCTCACCC
TCTTCGCGGTCGTGCGCGCCGGGGTGTCCATGAAGGCCGCGATCAAAGTGGCGCTGGCCGCCGACACCG
TCTCCATCGCGGTGATGGAGCTGGTCGACAACGGAATCATCGCCCTGGTCCCCGGCGCCATGGAGGCGCA
CCTGTCGGACGGGCTGTTCTGGTACGCCCTGCTCGGCGGCTTCGCCGTGGCGTTCGTGATCACCACGCCG
GTCAACAAGTGGATGATCGGTCGCGGCAAGGGCCACGCCGTCGTCCACGCCTACCACCTCGA<mark>GCACCACC</mark>
<mark>ACCACCACCAC</mark>TGA

SLI_RS17250_pET21a

CATATGGACCACAGCACGCACCACTCCGCCCCCGAAGACGACGGGCCCGGACACCCCCACGGGCACGGG
CACCTCCACGGGCCCGGACACCCTCACGGGCACGTGCACGTGCACGGCACCACCTGGGCGACCGCCATGC
AGGCGACGCTGCACTGCCTCACCGGGTGCGCCATCGGCGAGATCCTCGGCATGGTCATCGGAACCGCGCT
GATGTGGGGCAACGTGCCGACCATGGTGCTGGCCATCGCCCTGGCCTTCGTCTTCGGCTACTCCCTCACCC
TCTTCGCGGTCGTGCGCGCCGGGGTGTCCATGAAGGCCGCGATCAAAGTGGCGCTGGCCGCCGACACCG
TCTCCATCGCGGTGATGGAGCTGGTCGACAACGGAATCATCGCCCTGGTCCCCGGCGCCATGGAGGCGCA
CCTGTCGGACGGGCTGTTCTGGTACGCCCTGCTCGGCGGCTTCGCCGTGGCGTTCGTGATCACCACGCCG
GTCAACAAGTGGATGATCGGTCGCGGCAAGGGCCACGCCGTCGTCCACGCCTACCACTGA

SLI_RS17250_pET26b (C-terminal (His)₆-tag)

CATATG<u>AAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCTGCTCCTCGCTGCCCAGCCGGCGATGGCCAT
GGATATCGGAATTAATTCGGATCCGATG</u>GACCACAGCACGCACCACTCCGCCCCCGAAGACGACGGGCCC
GGACACCCCCACGGGCACGGGCACCTCCACGGGCCCGGACACCCTCACGGGCACGTGCACGTGCACGGC
ACCACCTGGGCGACCGCCATGCAGGCGACGCTGCACTGCCTCACCGGGTGCGCCATCGGCGAGATCCTCG
GCATGGTCATCGGAACCGCGCTGATGTGGGGCAACGTGCCGACCATGGTGCTGGCCATCGCCCTGGCCTT
CGTCTTCGGCTACTCCCTCACCCTCTTCGCGGTCGTGCGCGCCGGGGTGTCCATGAAGGCCGCGATCAAAG
TGGCGCTGGCCGCCGACACCGTCTCCATCGCGGTGATGGAGCTGGTCGACAACGGAATCATCGCCCTGGT
CCCCGGCGCCATGGAGGCGCACCTGTCGGACGGGCTGTTCTGGTACGCCCTGCTCGGCGGCTTCGCCGTG
GCGTTCGTGATCACCACGCCGGTCAACAAGTGGATGATCGGTCGCGGCAAGGGCCACGCCGTCGTCCACG
CCTACCACCTCGA<mark>GCACCACCACCACCACCAC</mark>TGA

Appendix 4.1 - DNA sequencing of three *SLI_RS17250* (DUF4396) plasmid constructs. Nucleotide bases highlighted in yellow represent (His)₆ -tag sequences and nucleotide bases highlighted in red show stop codons. Nucleotide bases that are underlined show the pelB leader sequence.

Appendix 5

Appendix 5.1 - *SDS-PAGE protocol*

The mini-protean tetra cell BioRad system was used to prepare the sodium dodecyl sulphate polyacrylamide gels (SDS-PAGE) (15 % w/v), see gel compositions in Table 4.1. The samples were prepared in cracking/loading buffer (1 % SDS, 25 % glycerol, 50 mM Tris/HCl pH 6.8, 0.05 % bromophenol blue and deionised water with a few crystals of DTT (dithiothreitol) added fresh) and heated at 95 $^{\circ}$C for 10 minutes for cells and 3 minutes for pure protein. The size of protein bands was checked by loading a page-ruler protein marker (Fischer scientific) into the gel. The running conditions of the gel were 140 Volts for ~1 hour and afterwards stained with Coomassie brilliant blue (455 ml ethanol, 90 ml acetic acid, 455 ml water, 2.5 g Coomassie brilliant blue) for 30 min then destained (450 ml ethanol, 450 ml water, 100 ml acetic acid) for 1-2 h until bands were clearly visible.

Table 5.1 – Composition of 15 % w/v SDS-PAGE gels

Material	15 % SDS gel (Resolving)	Stacking gel
0.5 M Tris (Fisher scientific) pH 6.8	-	1000 µL
1.5 M Tris pH 8.8	2500 µL	-
dd H$_2$0	2300 µL	2250 µL
Acrylamide (Sigma Aldrich)	5000 µL	666 µL
10 % (w/v) SDS	100 µL	40 µL
10 % (w/v) APS (Fisher scientific)	100 µL	40 µL
TEMED (Sigma Aldrich)	10 µL	5 µL