Metalloenzymes required for glycan processing and morphological development in *Streptomyces lividans*

Amanda Kay Chaplin

A thesis submitted for the degree of PhD Biochemistry

School of Biological Sciences

University of Essex

January 2016

Abstract

Streptomyces are filamentous soil-dwelling bacteria with a complex life cycle. Elucidating the signals that regulate morphological change in microbes is of fundamental importance for biotechnology applications. In the case of *Streptomyces* a development switch occurs with the concomitant production of secondary metabolites, many of which have pharmaceutical properties. For the industrially used strain *S. lividans* this switch is dependent on the bioavailability of copper (Cu) in the environment. This thesis has explored the relationship between Cu-chaperones and a haem peroxidase, part of the *sco* operon, with the maturation of a Cu-containing oxidase, GlxA. In *S. lividans* the GlxA gene is part of the *cslA/glxA* operon that contains genes encoding putative enzymes involved in glycan processing. Both these gene clusters are highly conserved in streptomycetes.

In Chapter 2 the characterisation of GlxA is reported. It was found to be membrane-associated with a mononuclear Cu site and possess a Cys-Tyr redox cofactor capable of housing a protein radical, comparable to the fungal galactose oxidase (Gox). The tertiary structure of GlxA revealed a unique domain arrangement, atypical spectroscopic properties compared to Gox and a lack of enzymatic activity with classical Gox substrates. Generation of the *ΔglxA* null mutant was found to stall aerial hyphae development on solid media and dramatically change the morphology in liquid cultures. This was ascribed to the absence of the oxidation of a glycan by GlxA produced by CslA (a cellulose-synthase), required for morphogenesis on solid and liquid cultures. The molecular nature of this glycan is unknown.

A number of GlxA variants were created in Chapter 3 to elucidate the proteins unique spectroscopic properties. It was found that the second coordination sphere residue, Trp^{288}, plays a major role in tuning the electronic properties of the buried Cu site in GlxA. Its removal abolishes the Cys-Tyr radical and perturbs the spectroscopic properties such that they resemble Gox. Monoclonal antibodies were used to follow the maturation of GlxA through observing mobility differences on denaturing PAGE gels based on the presence or absence of the Cys-Tyr cross-link. X-ray crystallography provided structural insight into the maturation process. A surprising outcome of Chapter 3 was that upon removal of the cross-linking Cys^{121}, a new protein radical is formed as opposed to the expected abolition.

Chapter 4 addresses another surprising finding in that a putative haem peroxidase (DtpA), part of the *sco* operon, plays a role in GlxA maturation and in the Cu-dependent morphological development. DtpA is shown through enzymology and structural analysis to be a member of the dye-decolourising peroxidase (DyP) family. Crucially, it is shown that DtpA functions as a peroxidase in the presence of GlxA using the GlxA substrate, glycolaldehyde.

Synthesis and modification of the CslA/GlxA glycan will inevitably require degradation during the life cycle. As part of the *cslA-glxA* gene cluster are two genes encoding for putative polysaccharide degrading enzymes. One of these is a putative Cu lytic polysaccharide monooxygenase, *Sli*LPMO10E. Chapter 5 structurally characterises *Sli*LPMO10E and also investigates the kinetics of Cu-binding. The latter brings to the attention that LPMOs are able to bind Cu in two forms at a single site before relaxing into a final substrate active form. Importantly, *Sli*LPMO10E is found to be active only with chitin via a C1 sugar ring oxidation mechanism. This hints at the possibility that the glycan produced by CslA and modified by GlxA is chitin-like possessing N-acetyl glucosamine moieties. By combining the *in vitro* results from this thesis together with the *in vivo* results obtained through the duration of this work from collaborators at Leiden University an overall model of the Cu-dependent morphogenesis and glycan processing in the hyphal tips of *S. lividans* is presented.

Chapters 2 through to 5 focus on events that occur under Cu limitations, *i.e.* homeostasis. Chapter 6 extends on previous work that characterised the CsoR regulon. The Cu sensitive operon repressor (CsoR) protein determines the set point of Cu(I) concentration in the cytosol. Under Cu stress, Cu(I) binds to CsoR and de-represses genes under its transcriptional control. Chapter 6 explores the possibility of whether CopZ-like Cu-chaperones can traffic Cu(I) to the DNA-bound CsoR, resulting in the up-regulation of control systems to return the cell to homeostasis. Size-exclusion and EMSA studies showed that Cu(I) was transferred from CopZ to CsoR in a unidirectional manner. Re-analysis of previous RNA-seq data using the *S. lividans* genome as input, enabled for a more complete model for the CsoR regulon in *S. lividans* to be proposed.

Acknowledgements

Completing this PhD would not have been possible without support, guidance and encouragement from some very special people. Whilst my name may be alone on the front cover of this thesis a number of people deserve huge thanks, which include: patient friends, a caring supervisor, inspiring advisors, loving family and kind lab members.

Importantly, I would like to thank my family: Mum, Dad, Philip and Donna; my partner Michael and his family; and my grandparents, for their continual support, love, concern and strength during my PhD. Thank you for listening to me complain when I have been stressed (especially when experiments were not working), for making me cups of tea and eggs on toast, for encouraging me each and every day and for always telling me to just keep going. Above all else [thank you] for understanding how important my PhD is to me.

I want to thank present and past members of the ever changing and evolving lab 5.08: Badri, Sri and Katie for introducing me to the lab and helping me find my feet, Benedict for establishing that I do have a competitive streak, Yannis and Oliver for tea and lunch breaks that made my day enjoyable even during the most stressful times, Demet for patient and kind help with crystallography, as well as all other lab members that have made the lab such an great place to work (Dan, Harish, Nicola, Richard, Sam, Tadeo, Megan, Tatiana, Louise). I would also like to thank all members of lab 6.19 for their discussions, interest and help. I would also like to thank all the technical staff, without whose help my work would not have been possible: Miao, Victoria, Julie, Hannah, Amanda and Anna and everyone in the kitchen who has helped to get my ever-increasing quantities of media through the autoclave. I would also like to thank all the summer, undergraduate and masters students that I have had the pleasure to teach and help in the lab for making the lab such good fun. Special thanks to Penny, Katie and Oliver for being great friends and listening to me moan everyday. I would also like to thank all of my friends, especially Chelsea and Goob who have always been a constant support, listened to me after stressful days and always strengthened my confidence.

I would like to make a very special thank you to Mike Wilson. Thank you for all your help with stopped-flow experiments, discussions about life and science, the dreaded “maths” lessons, for always showing a huge interest in my work, and for always igniting my passion for science. Mike taught me how to question thoughts and to express and understand ideas, which I will take with me into my future scientific career.

A special mention to a friend no longer with us, Peter, thank you for all your help with my enzyme assays, your wealth of knowledge and incredible kindness with any question asked, you are missed very much.

Special thanks also to Dima, for all his help with EPR experiments and data analysis,

especially for having the patience and time to explain how to interpret my results. Thanks also to Mike Hough for all of his assistance with crystallography during crystal measurements and when solving and checking my structures. I would also like to thank our fantastic collaborators at Leiden University, especially, Erik, Marloes and Dennis for all the *in vivo* data presented in this thesis. Their hard work, discussions and input have allowed for a more rounded and complete understanding. I would also like to thank The University of York, especially Glyn Hemsworth for help with my mass spectrometry experiments.

Finally, I would like to thank Jonathan who above everyone else really made this PhD possible, who encouraged me from the beginning and never faltered, thank you. Thank you for guiding me from the start of my undergraduate degree as my personal tutor, sparking my interest in your research during my summer placement, allowing me to carry on during my final year project and really encouraging and believing in me to apply for a PhD. We have experienced many ups and downs, sharing happiness of success and depression of failure, but your selfless time and care were what always kept me going. Thank you for your advice, encouragement and patience, for allowing me to work independently and for trusting me. Your constant care, belief and stubborn ability to always keep going inspired me to do the same. It has been a real pleasure to work with you and I hope one day I can go on to be just as great a supervisor as you have been.

Lastly I would like to thank the University of Essex Silberrad Scholarship, which provided funding for this PhD. From the support of all these people I have developed as a more rounded scientist and person during my PhD and I hope that one day I can go on to inspire young scientists just as I have been inspired.

Table of Contents

Table of Figures

Table List

Abbreviations

$[Fe(CN)_6]^{3-}$ - Ferricyanide

$[Fe(CN)_6]^{4-}$ - Ferrocyanide

$[Ir(Cl)_6]^{3}$ - Iridium Chloride

AA - Auxiliary activity

ABTS - 2,2-azinobis (3-ethylbenthiazoline-6-sulfonic acid)

ALA - 5-aminolaevulinic acid

Ala - Alanine

APS - Ammonium persulphate

APX - Ascorbate peroxidase

AQ - Anthraquinone

Arg - Arginine

Asn - Asparagine

Asp - Aspartic acid

Atx1 - Anti-oxidant

Badan - 6-bromoacetyl-2-dimethylaminonaphthalene

BCA - Bicinchoninic acid

BCS - Bathocuproinedisulfonate

CAZy - Carbohydrate Active enzyme

C*c*O - Cytochrome *c* oxidase

C*c*P - Cytochrome *c* peroxidase

CD - Circular Dichroism

CFW - Calcofluor White

Cm - Chloramphenicol

CO - Carbon monoxide

CslA - Cellulose synthase-like protein

CsoR - Copper sensitive operon regulator

Cu - Copper

Cys - Cysteine

Da - Daltons

dd - double distilled

DMP - 2,6-dimethoxyphenol

DMSO - Dimethyl sulfoxide

dNTP - Deoxyribonucleotide triphosphate

DP - Degrees of polymerization

DTNB - 5,5'-dithiobis-(2-nitrobenzoic acid) or Ellman's reagent

DtpA - dye-type peroxidase A

DTT - Dithioreitol

DyP - Dye-decolourising peroxidase

ECuC - Extracellular Cu chaperone

EDTA - Ethylenediaminetetraacetic acid

E_m - Reduction potential midpoint

EMSA - Electrophoretic mobility shift assays

EPR - Electron paramagnetic resonance

GARAP - goat anti rabbit alkaline phosphatase

GH - Glycoside hydrolase

Glu - Glutamic acid

GlxA - Galactose oxidase A

Gly - Glycine

Gox - Galactose oxidase

H_2O_2- Hydrogen peroxide

HAH1 - Human ATX1 homologue 1

HCl - Hydrochloric acid

His - Histidine

HRP - Horseradish peroxidase

IPTG - Isopropyl β-D-1-thiogalactopyranoside

ITC - Isothermal titration calorimetry

Kan - Kanamycin

k_{cat}/V_{max} - Maximal turnover number

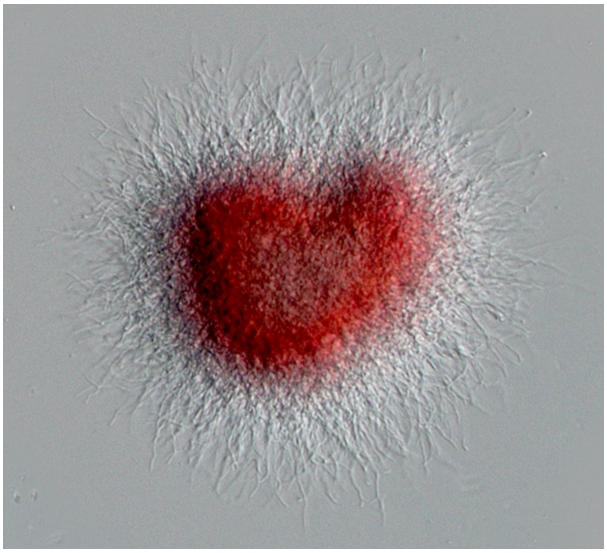
K_d - Dissociation constant

K_{eq} - Equation constant

K_m - Michaelis constant
Lac - Laccase
LB - Luria Bertani
LiP - lignin peroxidase
LMCT - ligand-to-metal charge transfer
LPMO - Lytic polysaccharide monooxygenase
LRET - Long range electron transfer
Lys - Lysine
MBD - Metal binding domain
MES - 2-(N-morpholino) ethanesulfonic acid
Met - Methionine
MnP - Manganese peroxidase
MOPS - 3-(N-morpholino) propanesulfonic acid
MRE - Mean Residue Ellipticity
MS - Mannitol salt
NaCl - Sodium chloride
NaOH - Sodium hydroxide
NBT/BCIP - Nitro Blue Tetrazolium/5-bromo-4-chloro-3-indolyl phosphate
Ni-NTA - Nickel-nitrilotriacetic acid
NRMSD - Normalised Root Mean Square Deviation
PCR - Polymerase Chain Reaction
PDB - Protein Data Bank
PEG - Polyethylene glycol
Phe - Phenylalanine
pK_a - Acid dissociation
Pro - Proline
QM/MM - Quantum mechanics/molecular mechanics
RB19 - Reactive blue 19
RPKM - Reads/kb of exon/million mapped reads
SCO - *Streptomyces coelicolor*
Sco - Synthesis of cytochrome *c* oxidase
SDS-PAGE - Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis
Ser - Serine
Sli - Streptomyces lividans
SOD - Superoxide dismutase
Tat - twin arginine translocation
TEMED - Tetramethylethylenediamine
Thr - Threonine
TMP - N,N,N',N'-tetramethyl-1,4-phenylenediamine
Tris - Tris(hydroxylmethyl)aminomethane
Trp - Tryptophan
TSB - Tryptic soya broth
Tyr - Tyrosine
UV - Ultra Violet
VA - Veratryl alcohol
Val - Valine
WT - Wild-Type
Zn - Zinc

Chapter One

INTRODUCTION



1.1. Streptomycetes

Streptomycetes are the largest genus of the phylum Actinobacteria and belong to the family Streptomycetaceae originally proposed by Waksman and Henrici in 1943 (Waksman and Henrici, 1943). The first true streptomycete appeared ~ 450 million years ago, eventually giving rise to today's *Streptomyces* species [1, 2]. They are Gram-positive, aerobic, bacteria predominately found in the soil existing ubiquitously in nature. Most members of streptomycetes display saprophytic, free-living existence, competing for resources with high numbers of microbial competitors in oligotrophic environments [3]. Streptomycetes grow as a mycelium of branching hyphal filaments, and reproduce in a mould-like manner by sending up aerial branches that turn into chains of spores. They contain a G + C rich genome and form multicellular colonies with a complex life cycle, resembling lower eukaryotes such as fungi [4, 5]. Their ability to grow in soil is facilitated by the formation of spores, which are resistant against low water and nutrient levels, assisting in their spread and persistence [6]. *Streptomyces* produce and secrete a large quantity of extracellular enzymes that can decompose polymers in dead plants, animals and fungal material facilitating a role of the bacteria in soil biodegradation [7]. However, streptomycetes are best known for their ability to produce a bewildering array of important secondary metabolites including antibacterial, immunosuppressive, anticancer, antihelmintic and antifungal agents [1, 8]. Since Waksman discovered streptomycin as the first therapeutically useful *Streptomyces* antibiotic ~ 70 years ago a large variety of compounds with pharmaceutical properties have been identified with some examples shown in Table 1.1 [8-10]. In addition, *Streptomyces* is a producer of commercially valuable proteins naturally secreted to adapt to a changing environment and to different nutritional sources [10, 11]. An example of the vast array of enzymes produced by actinomycetes is shown in Figure 1.1. In particular, they produce enzymes for the degradation of many organic polymers such as cellulose, chitin and lignin, which are needed for the production of second-generation biofuels [12-14]. These bacteria are therefore of vast importance to human welfare and biotechnology applications.

Table 1.1: Medical importance of some compounds produced from *Streptomyces*.

Target disease or organism	Medicine	Organism
Typhoid	Chloramphenicol	*S. venezuelae*
River blindness	Avermectin	*S. avermitilis*
Pathogens with transmissible penicillin resistance	Clavulanic acid	*S. clavuligerus*
Nematode and arthropod infestations	Ivermectin	*S. avermitilis*
HIV	Boromycin	*S. antibioticus*
Cancer	Daunomycin	*S. coeruleorubidus*
Cancer	Migrastatin	*S. platensis*
Cancer	Bleomycin	*S. verticillus*

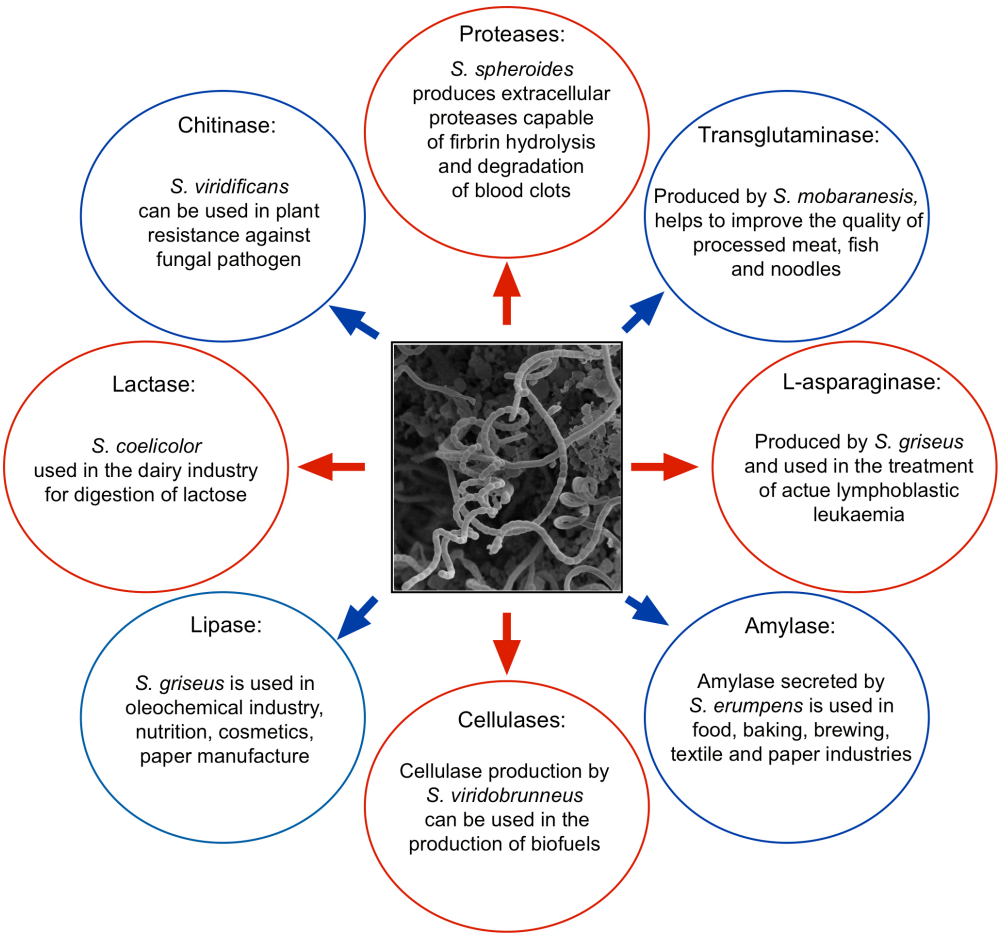


Figure 1.1: Applications of enzymes produced from *Streptomyces*. Image adapted from Sharma 2014 [15]. Applications of the enzymes were reported in [16-22].

1.2. Life cycle of *Streptomyces*

Streptomyces display a complex life cycle and are one of the model systems for bacteria morphological and physiological development [5, 23-25]. Growth of a colony begins from a spore, the reproductive form of the bacteria. Once this spore locates favorable conditions for germination, germ tubes are generated, initiating growth of the vegetative mycelium stage of development. Growth then continues through tip extension and branching resulting in a large network of vegetative hyphae [26] (Figure 1.2). The vegetative growth phase can be identified on solid agar media by the formation of red colonies originating from the antibiotic pigment undecylprodigiosin produced in some species (*S. coelicolor* and *S. lividans*) (Figure 1.2 and front page). As the colonies continue to grow, the older mycelia become densely populated leading to changes in the bacteria largely due to limitations in nutrients and physiological stress. During this period it becomes advantageous for the bacteria to export enzymes into the extracytoplasmic environment for degradation of insoluble nutrients in order to replenish growth. The bacteria also secrete antibiotics, to protect the solubilised material against other microorganism competitors [1, 27]. When nutrients become depleted, aerial mycelium is erected that feeds on the vegetative mycelium. Aerial hyphae grow via curling and tip elongation producing fluffy white colonies on solid agar media (Figure 1.1 and Figure 1.2) [28-31]. Transition from vegetative to aerial hyphae requires *bld* genes and two classes of surface-active molecules SapB (spore-associated protein B) and chaplins (*coelicolor* hydrophobic aerial proteins) [32-34]. Septa are then formed at regular intervals and multiple cell divisions take place to form cylindrical unigenomic pre-spore compartments. These compartments then mature and create spores, starting the cycle of development again. The spore walls become pigmented with a grey polyketide derived aromatic compound creating the dark grey appearance of the mycelium grown on solid media [5] (Figure 1.2).

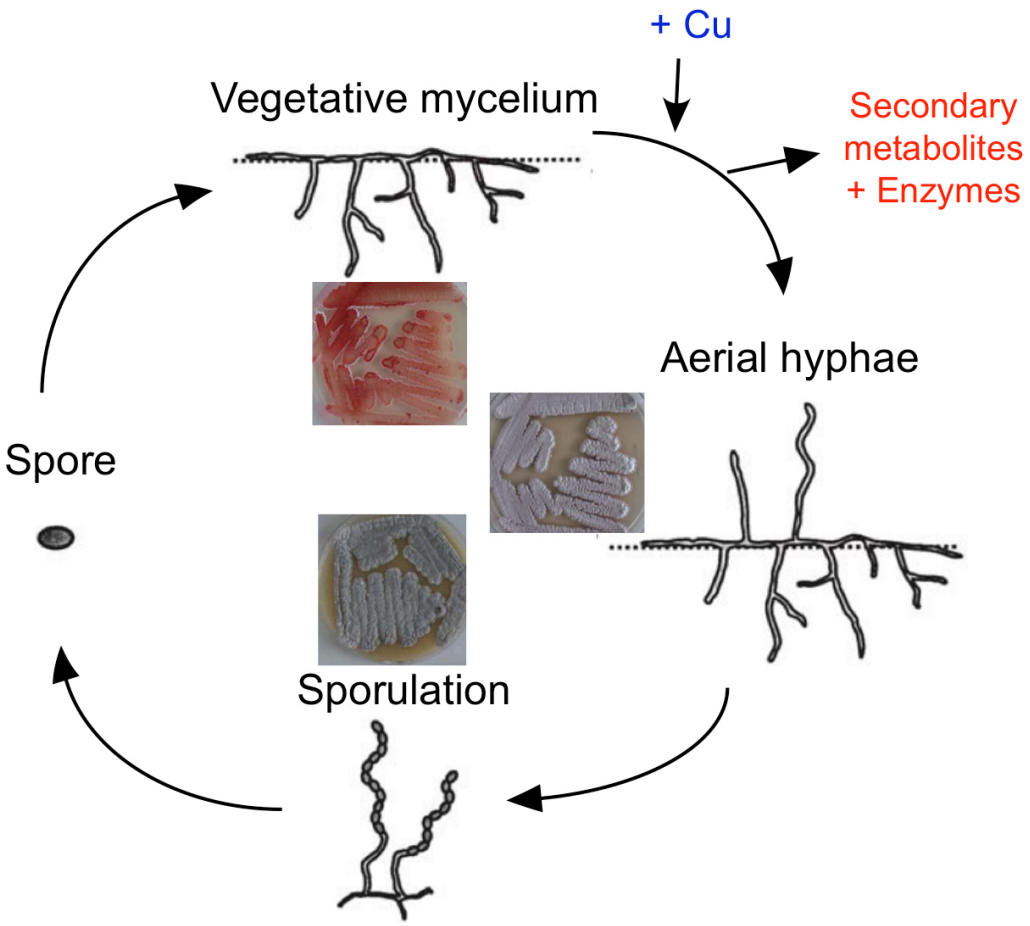


Figure 1.2: The life cycle of *Streptomyces*. Upon favourable conditions, spores germinate and branch into a network of vegetative mycelium. When nutrients become deplete, aerial mycelium is erected that feeds on the vegetative mycelium. Aerial hyphae grow in a curling motion, and then subsequent cell division occurs with the formation of cross walls at regular intervals. These compartments then mature and create spores, starting the cycle of development again. Inside the developmental life cycle, examples of colony formation of *S. lividans* on solid agar can be observed during the different developmental stages. Image adapted from Worrall and Vijgenboom 2010 [5].

1.3. Copper dependent morphogenesis

The switch between vegetative mycelium and aerial hyphae is concomitant with the production of a plethora of secondary metabolites, including antibiotic and anticancer agents, as well as many commercially valuable enzymes (Table 1.1 and Figure 1.1) [5, 8, 35, 36]. Onset of the morphological switch is associated with the production of secondary metabolites, indicating that these two processes share common regulatory elements (Figure 1.2). Copper (Cu) has been shown to have a crucial influence on the morphogenesis of *S. lividans* and in particular triggering the development switch between vegetative to aerial mycelium that coincides with the production of secondary metabolites [5, 24]. Increased Cu levels were shown to strongly induce aerial mycelium formation in *S. lividans* [24, 25, 37]. No such stimulation was observed with the addition of other metal salts such as, $FeCl_3$, $ZnCl_2$, $NiCl_2$, $CoCl_2$ or $MgCl_2$ [24]. In the presence of the Cu(I) specific chelator, Bathocuproinedisulfonate (BCS), inhibition of aerial hyphae development could be observed on solid medium. Since Cu plays a vital role in the transformation of vegetative to aerial mycelia, cuproproteins and cuproenzymes must have a role [5]. However, the identification and pathways involved have not yet been fully elucidated.

Recent *in vitro* and *in vivo* work in *S. lividans* has revealed an extracytoplasmic Cu-trafficking pathway that under limiting Cu bioavailability (*i.e.* homeostasis) is operable for the switch from vegetative to aerial mycelium to occur. This pathway involves two Cu chaperones, Sco (synthesis of cytochrome *c* oxidase) and ECuC (extracytoplasmic Cu chaperone) [38], which are part of the *sco* operon commonly found in other *Streptomyces* (Figure 1.3) [25, 37]. This operon also contains a putative metal exporter (SLI4212) implicated in transport of cuprous ions from the cytoplasm to ECuC [4, 5] and a gene referred to as *dtpA,* a putative secreted protein with a Tat (twin arginine translocation) signal sequence and a DyP (dye-decolourising)-type domain (DtpA; dye-type peroxidase A) (*SLI4211*) [25, 37, 39, 40]. DyPs are a new class of monohaem peroxidases that are found widely dispersed among bacteria and fungi, but their physiological role is unknown [41, 42]. Within this Cu-trafficking pathway, Sco receives a Cu(I) ion from ECuC (extracytoplasmic Cu chaperone) and subsequently delivers Cu to the Cu_A site of an aa_3-type cytochrome *c* oxidase (C*c*O) [38]. In the absence of ECuC and C*c*O, development is not impaired, but in the absence of Sco development stalls [25, 37, 38]. Whereas Δ*ecuc* and Δ*sco* significantly reduce C*c*O activity, these data indicate that morphological development is not linked to impaired C*c*O activity [37, 38]. Notably, the addition of exogenous Cu to the Δ*sco* mutant rescues development [37]. This implies that a branched Cu trafficking pathway is operable under Cu homeostasis conditions in *S. lividans*, whereby Sco is acting as Cu chaperone for C*c*O and possibly for other extracytoplasmic cupro-proteins/enzymes that trigger formation of aerial hyphae [37,

38]. The presence of a putative haem peroxidase in the *sco* operon could suggest that the branch point in the Sco Cu-trafficking pathway may involve a H_2O_2 producing enzyme.

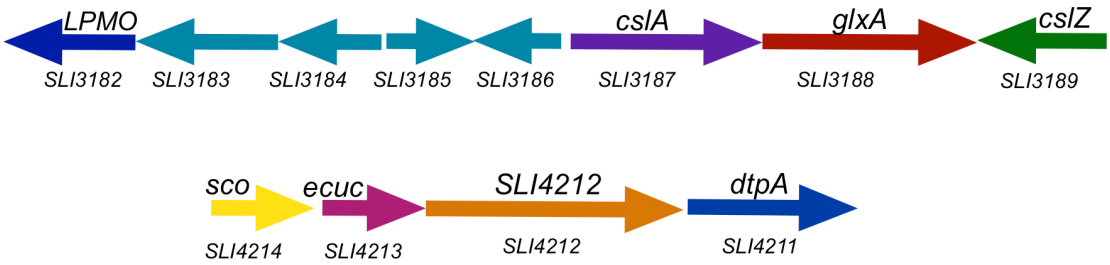


Figure 1.3: Genome environment of the *cslA-glxA* and the *sco-dtpA* locus in *S. lividans* [4]. Gene annotations are as follows: *SLI3182* secreted chitin binding protein (LPMO); *SLI3183* putative membrane protein; *SLI3184-SLI3186* are hypothetical proteins; *SLI3187* glycosyl transferase (CslA); *SLI3188* GlxA radical monocopper oxidase; *SLI3189* CslZ secreted endoglucanase; *SLI4214* Sco a secreted Cu chaperone, *SLI4213* ECuC a secreted Cu chaperone, *SLI4212* a putative Cu transport protein and *SLI4211* encodes DtpA, a dye-decolourising peroxidase A type.

1.4. Streptomycetes as a production host in biotechnology

The application of recombinant proteins in industrial and pharmaceutical processes has grown steadily in the last decades and has become an indispensable part of the manufacturing of many products. Recently, streptomycetes have gained a great deal of interest as a potential host due to their excellent ability to directly secrete proteins or enzymes into the culture broth [43-45]. *Streptomyces* is a promising bacterial host already used for the high-level production of several proteins. As a host it has the following advantages over other bacterial systems; 1) no literature reports of the formation of inclusion bodies (often a problem when using *E. coli*); 2) well suited for the expression of very G + C-rich genes without codon adaption [46]; 3) a high protein secretion efficiency, making it feasible to extract the target from the culture supernatant and thus minimising downstream bio-processing costs [44, 47]. *S. lividans* has been successfully used for the secretion of both eukaryotic and prokaryotic heterologous proteins with some displaying low expression levels in *E. coli* and *B. subtilis* efficiently produced in *Streptomyces* [43, 44, 46, 48, 49]. Despite the encouraging perspectives there are several constraints against the utilization of this organism on a larger industrial scale. From a host point of view, the main bottleneck is its mycelial growth, which results in rather dense clumps of mycelium in liquid cultures [50, 51]. This morphology strongly influences the overall efficiency of the fermentation, and the secretion and integrity of the desired final product. Furthermore, the very limited choice of stable vectors with strong promoters for expression of heterologous proteins is often a drawback for the selection of *Streptomyces* as host [52].

1.5. Submerged cultures of *Streptomyces*

Metabolite and enzyme production typically occurs in large bioreactors. Growth under these conditions differs to colony differentiation on surface cultures and is characterised by the formation of large, biofilm-like aggregates of mycelium, called pellets [50, 51]. Strains of *Streptomyces* differ in their complexity from dispersed mycelium, mycelial aggregates, or ball-shaped pellets of a dense network of mycelium [53, 54]. Pellet formation is due to the tight entanglement and subsequent interweaving of numerous branching hyphae [54, 55]. From an industrial perspective these mycelial aggregates are unfavourable principally due to mass-transfer problems, slow growth and culture heterogeneity [50]. Pellets prevent the effective transfer of gasses and nutrients to the centre, thus lowers the attainable product yield [56]. Classical programs for strain improvement are slow, labour intensive and involve large numbers of mutations [57]. Growth rate and enzyme production in *S. lividans* has been improved by expression of the actinomycete-specific cell

division factor SsgA (Sporulation of *Streptomyces griseus*), which controls the septum site localization and cell wall homeostasis, thus leads to increased fragmentation and smaller pellet size [58-61]. Importantly, this increased fragmentation is accompanied by a large increase in enzyme expression [58].

1.6. A role for glycans in pellet formation

Like in biofilms, formation and integrity of streptomycetes pellet aggregates in liquid cultures depend upon the synthesis of extracytoplasmic glycans [54, 62, 63]. Therefore, the growth of *Streptomyces* is dependent upon glycan synthesis. Recently, reverse engineering of a non-pelleting strain of *S. lividans* selected in a chemostat indicated a crucial role for the *mat* gene locus, putatively involved in synthesis of an extracytoplasmic glycan required for pellet formation [57]. Deletion of the *mat* genes leads to a dispersed mycelium with a 60 % increase in growth rate and productivity of *S. lividans* [57].

A second extracellular glycan, with a role in pellet formation has been identified in *S. coelicolor* and is produced by enzymes encoded by the *cslA-glxA* locus [64, 65]. CslA, which stands for cellulose synthase-like protein, plays multiple roles in growth and development; mutants lacking the *cslA* gene fail to form aerial hyphae and are affected in attachment to solid surfaces [65, 66]. Furthermore, CslA plays an important role in the morphology of mycelial clumps in liquid cultures [64, 65]. CslA is classified as a family 2 glycosyltransferase (GT2, Carbohydrate Active Enzymes Database, www.cazy.org) [67]. This family does include cellulose synthases but also chitin synthases and other polysaccharide synthases. The polymers produced by these synthases have many different functions, often directed at providing structural integrity.

The polysaccharide synthesized by CslA has not yet been characterized. Typically, genes involved in bacterial cellulose synthesis are organized in an operon, containing the *bcsABCD* genes [68]. However, *Streptomyces* lacks a c-di-GMP binding protein (BcsB), conserved in cellulose-producing organisms, essential for the synthesis of cellulose according to the currently accepted model [69]. In light of this it would suggest that the glycan produced by CslA has N-acetyl glucosamine functionalities (*i.e.* chitin-like). Interestingly, *cslA* is translationally coupled to *glxA*. GlxA stands for galactose oxidase (Gox) A, due to its weak sequence homology with the fungal Gox enzyme. From sequence alignment GlxA like Gox contains a 'kelch motif' a signature domain of this class of Cu oxidases. Gox is the founding member of the radical Cu oxidase family that catalyzes the oxidation of a broad range of D-isomers of primary alcohols to aldehydes, coupled to the reduction of dioxygen to H_2O_2 [70, 71]. Therefore, GlxA may function in a similar manner to Gox, oxidizing sugar molecules and producing H_2O_2. The *cslA-glxA* operon is conserved in all streptomycetes, with some species containing two copies, believed to be acquired via horizontal gene transfer from fungi [72]. Disruption of the *glxA* gene stalls development and abolishes pellet formation as seen in the *cslA* mutation [72]. *In situ* visualisation of GlxA revealed its

abundance at hyphal tips not dependent upon CslA, indicating that both GlxA and CslA functionally cooperate in the hyphal tip [72].

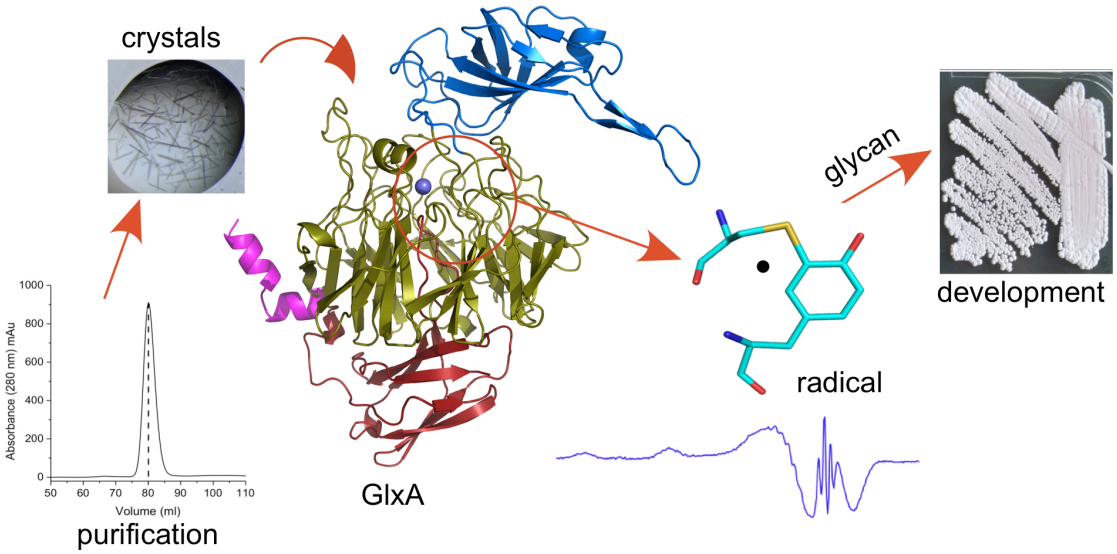
1.7. Focus of this thesis

This thesis is focused on metalloenzymes involved in the Cu-dependent morphological development in *S. lividans*. In particular a link between the Cu-trafficking *sco* operon and the *cslA/glxA* locus is explored. In *S. coelicolor* the *cslA* and *glxA* null-mutants have a 'bald' phenotype (*i.e.* no aerial hyphae formation) on solid media [72], which is similar to that observed for *Δsco* in *S. lividans* [37]. The latter can be rescued by addition of exogenous Cu(II) to the medium [37], whether this is the case for a Δ*glxA* is not known. Based on the weak sequence similarity to Gox, it is hypothesised that GlxA will be a cuproenzyme possessing oxidase activity. This has a number of potential implications for understanding Cu-dependent morphogenesis in *S. lividans*. The first being it will require Cu for function. Metallation of GlxA could proceed through the ECuC/Sco Cu trafficking pathway and may indeed be the second Sco target required to trigger aerial hyphae development [38]. Secondly, the translational coupling of *glxA* with *cslA* and the putative oxidase activity of GlxA could suggest a functional cooperation with CslA, in that GlxA modifies through an oxidation process the glycan synthesised. This modification would then enable hyphal tip development on solid substrate and also result in pellet formation in liquid cultures. Continuing with a Cu-oxidase role, H_2O_2 would be produced by GlxA upon interaction with a glycan substrate, which would require subsequent removal suggesting a possible role for DtpA. This scenario would provide an explanation for the presence of a putative extracytoplasmic haem peroxidase in the Cu requiring *sco* operon. Further analysis of the *cslA/glxA* genome environment in *S. lividans* reveals two annotated genes, *cslZ* and *LPMO*, that encode for a putative endoglucanase and a Cu-containing lytic polysaccharide monooxygenase (LPMO), respectively (Figure 1.3). This would suggest that the glycan produced by CslA could be further processed or degraded through hydrolytic (CslZ) or oxidative (LPMO) processes. In keeping with the metalloenzyme theme the LPMO has been characterised and has provided functional insight into the chemical properties of the glycan produced by the CslA.

The structural, spectroscopic and biochemical characterisation of GlxA, DtpA and the LPMO are presented in this thesis. Results obtained are integrated with *in vivo* data obtained from collaborators at the University of Leiden leading to the construction of a model for the Cu-dependent hyphal tip development in *S. lividans*.

Chapter Two

Structural and functional characterisation studies of GlxA from *Streptomyces lividans*



Synopsis: GlxA is determined to be a monomeric cuproenzyme with a unique structure that is required for glycan synthesis to initiate morphological development in *Streptomyces lividans*

Results from this Chapter have been published in:

Chaplin, A.K., Petrus, M.L.C., Mangiameli, G., Hough, M.A., Svistunenko, D.A., Nicholls, P., Claessen, D., Vijgenboom, E., Worrall, J.A.R. "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*" 2015 *Biochem. J.,* **469,** 433-44.

2.1. Introduction

Streptomycetes produce a rich and diverse range of clinically useful secondary metabolites during the morphological switch from vegetative to aerial mycelium, [26, 63]. They also hold promise as large-scale production hosts for the heterologous expression of commercially valuable proteins and enzymes due to their capability to directly secrete these products into the culture broth [43, 44]. *Streptomyces lividans* is a favoured industrial host, primarily due to its low level of endogenous proteolytic activity, which together with its secretion properties minimises downstream bioprocessing costs. *S. lividans* also displays a distinct dependence on the bioavailability of copper (Cu) in order to fully initiate morphological development [24, 25, 37]. Cu is an essential transition metal ion in biology due to its ability to cycle between the Cu(II) and Cu(I) redox states and thus acts as a co-factor for a wide range of proteins and enzymes.

The importance of Cu to the development program in *S. lividans* has been elucidated through recent *in vivo* and *in vitro* studies on two extracytoplasmic Cu chaperones, ECuC (extracytoplasmic Cu chaperone) and Sco (synthesis of cytochrome *c* oxidase) [37, 38]. These chaperones facilitate a Cu-trafficking pathway that involves Sco receiving a Cu(I) ion from ECuC via a ligand-exchange mechanism and subsequent transfer of the Cu ion to the Cu_A site of an aa_3-type cytochrome *c* oxidase (C*c*O). Under Cu limiting conditions genetic knockout studies have revealed morphological development in *S. lividans* proceeds in the absence of ECuC (*Δecuc*) and C*c*O (*Δcco*) but not in the absence of Sco (*Δsco*) [37, 38]. However, the null-mutants, *Δecuc* and *Δsco*, significantly reduce the activity of C*c*O illustrating that morphological development is not linked to a reduction in C*c*O activity [37]. Notably, morphogenesis of the *Δsco* mutant is restored by the addition of exogenous Cu to the medium, indicating the possibility that a branched Cu-trafficking pathway is operative [37], whereby Sco acts as a Cu chaperone for C*c*O but also for other(s) extracytoplasmic cupro-proteins/enzymes required to illicit morphological development [38]. A Cu-trafficking target, other than C*c*O, for either prokaryotic or eukaryotic Sco proteins is not known. Blundell *et al.,* have suggested the possibility in *S. lividans* that the *glxA* gene product is a potential Sco target.

In *S. lividans* the *SLI3188 (glxA)* gene product (GlxA) has weak amino acid sequence homology with the cuproenzyme galactose oxidase (Gox) [4, 5]. Fungal Gox is a secreted enzyme originally isolated from cultures of *Fusarium spp*., *Gibberella fujikuroi* [73], and *Fusarium graminarium* (previously identified as *Dactylium dendroides*) [71, 74-77]. Gox is the founding member of the radical Cu oxidase family that can catalyze the oxidation of a broad range of D-isomers of primary alcohols to aldehydes, coupled to the reduction of dioxygen (O_2) to hydrogen peroxide (H_2O_2) [70, 71]. Another closely related member of the

radical Cu oxidase family is glyoxal oxidase from *Phanerochaete chrysosporium* [78-80] with its role being to catalyze the oxidation of simple aldehydes to carboxylic acids [78]. Radical Cu oxidases combine two distinct redox centers, a mononuclear Cu ion capable of one electron redox cycling and a stable protein free radical enabling for the two-electron oxidation of specific substrates. In this family of enzymes the protein radical resides on a crosslink formed between one of the Cu-coordinating tyrosine residues and a cysteine, forming a tyrosyl-cysteine thioether bond (Cys-Tyr) [76, 81-85]. Studies with fungal Gox have shown that Cu is required to initiate the processing of the Cys-Tyr crosslink resulting in the active (mature) form of the enzyme [86, 87]. The physiological function of Gox remains unknown, but its broad substrate specificity implies that production of H_2O_2 may be important in the extracytoplasmic environment of fungi.

In *S. coelicolor* (> 95 % sequence homology with *S. lividans*) a gene homologous to *SLI3188* also annotated as *glxA (SCO2837)* is required for aerial hyphae development under osmotic stress conditions [72]. The *glxA* gene in *S. coelicolor* and *S. lividans* is the distal gene in an operon with *cslA*, which encodes the CslA protein belonging to family 2 glycosyltransferases [62, 65]. CslA synthesizes a β–(1-4)-glycan at the hyphal tips, thought to provide protection during ongoing cell wall remodeling [65]. Both CslA and GlxA in *S. coelicolor* are co-localised at the hyphal tips indicating a role in glycan processing for both these proteins [65, 72]. The *cslA-glxA* operon was probably acquired via horizontal gene transfer from fungi and is conserved among all streptomycetes with some having two copies [72]. It has been reported that deletion of *cslA* and *glxA* has marked effects on growth and development of *Streptomyces* [65, 72], signifying important roles of these enzymes within the life-cycle.

Typical growth of streptomycetes is characterized by large biofilm-like aggregates of mycelium, called pellets [50]. Formation and integrity of these pellets depends on the synthesis of extracellular glycans [54, 62, 63]. Deletion of *mat* genes involved in glycan synthesis creates a dispersed growing mycelium with an increase in growth rate and productivity by 60 % [57]. Interestingly the *cslA-glxA* genetic locus contains several other genes encoding enzymes implicated in glycan processing, including a Cu-containing lytic polysaccharide monooxygenase (*Sli*LPMO10E) and a putative endoglucanase called CslZ, possibly involved in glycan breakdown (Figure 2.1) [65, 72]. The biological function of GlxA has not yet been reported but the genetic environment suggests a role in glycan synthesis or modification.

Given that there is little functional or biochemical information on GlxA, the intention of this study was to characterize this putative enzyme through structural and functional studies to gain a greater understanding of its role within *S. lividans*. Various spectroscopic techniques have been employed to determine the physicochemical properties of GlxA and

how it compares to, Gox. Enzymatic activity and crystallographic data were analyzed alongside *in vivo* and molecular genetic studies to further elucidate its functional role. The findings highlight the unique nature of this cuproenzyme and its role in mediating Cu-dependent morphogenesis in *Streptomyces*.

Figure 2.1: Genome environment of the *cslA-glxA* locus in *S. lividans*. Gene annotations are as follows: *SLI3182* secreted chitin binding protein (*Sli*LPMO10E); *SLI3183* putative membrane protein *SLI3184-SLI3186* are hypothetical proteins; *SLI3187* glycosyl transferase (CslA); *SLI3188* GlxA radical monocopper oxidase; *SLI3189* CslZ secreted endoglucanase.

2.2. Experimental procedures

2.2.1. Mycelium washes, fractionation and GlxA detection

The following experimental procedures were carried out by collaborators at Leiden University (Dr Erik Vijgenboom, Dr Dennis Claessen and Marloes Petrus).

To detect the location of GlxA within the mycelium of *S. lividans* the bacteria was grown washed and fractions analysed using western blotting. *S. lividans* 1326 was grown in TSBS (Tryptic soy broth with 10 % sucrose) at 30 ^{o}C for 18 h, shaking at 200 rpm. The mycelium were harvested by centrifugation at 15,000 g for 10 min in 1.5 ml aliquots and stored at -80 ^{o}C until required. Mycelial pellets were first resuspended in 25 mM Tris/HCl, pH 7.5, 100 mM NaCl then centrifuged at 25,000 g with the supernatant kept as the 100 mM NaCl wash. The pellet was then resuspended in 25 mM Tris/HCl, pH 7.5, 1 M NaCl, incubated on ice for 30 min then centrifuged at 25,000 g with the supernatant kept as the 1 M NaCl wash. For mycelium extracts the pellets were resupsended in 25 mM Tris/HCl, pH 7.5 and sonicated (Bioruptor, 12 cycles: 30 s on, 30 s off). This suspended pellet was then centrifuged at 30,000 g to produce the soluble supernatant (S30) and insoluble pellet fractions (P30). The supernatant (S30) was then centrifuged again to remove all membrane particles and ribosomes at 100,000 g to produce a soluble supernatant (S100) and insoluble pellet (P100) fractions. Pellets P30 and P100 were then resuspended in buffer containing 1 % Triton X-100, incubated on ice for 30 min then centrifuged at 30,000 g to produce triton soluble (P30-TS) and insoluble fractions (P30-TP) then at 100,000 g to produce triton soluble (P100-TS) and insoluble (P100-TP) fractions. All these fractions were run on a 10 % SDS-PAGE gel and blotted to Hybond-P membranes for immune-detection of GlxA and EF-Tu1. The blots were incubated with the corresponding antibodies (GlxA from James W. Whittaker Oregon Health and Science University, USA and EF-Tu1 antibodies were raised in rabbits against *Streptomyces ramocissimus* EF-Tu1) [88] (10000 x diluted for GlxA and 5000 x diluted for EF-Tu1) in PBS and 5 % milk (FrisoLac Extra, FrieslandCampina) for 18 h with gentle rocking at 4 ^{o}C. The secondary antibody GARAP (goat anti rabbit alkaline phosphatase, Sigma) was used to detect the primary antibodies along with NBT/BCIP (Nitro Blue Tetrazolium/5-bromo-4-chloro-3-indolyl phosphate, Sigma) used as the substrate. Once signals were identified digital images were taken and the total intensities of the signals determined using ImageJ [89].

2.2.2. Cloning of GlxA

The *glxA* gene (*SLI3188*) with 200 flanking nucleotides at the 5' and 3' ends was amplified from *S. lividans* 1326 (*S. lividans* 66, stock number 1326 from the John Innes Centre) and cloned into a pUC19 vector. Topology of GlxA was predicted using TMHMM, TopPred2 and TMpred, which displayed the presence of a transmembrane region between residues 13-32 [90-92]. Therefore, all subsequent cloning steps were designed to amplify the DNA so that a truncated GlxA product beginning at residue 35 was produced. Primers were created containing NdeI and BamHI restriction sites (see below) to amplify the truncated *glxA* gene from the pUC19 plasmid containing the full *glxA* sequence (this was carried out by a summer project student Emma Blundell). Table 2.1 displays the composition of the reagents and Table 2.2 the PCR protocol that was used to create *glxA* (residues 35-645). The resulting PCR product was analysed on a 1 % agarose gel and gel extracted using a Qiagen gel extraction kit. It was then ligated into the NdeI and BamHI restriction sites of a pET28a (Kan^r) vector (Novagen) and transformed into *Escherichia coli* JM109 cells and plated on LB (Luria broth) kanamycin (Kan) (50 μg ml^{-1}) agar plates. Colonies were selected, digested using NdeI and BamHI, and sequenced to corroborate the correct sequence (Figure 2.2). The final construct was designated *pET3188* consisting of an N-terminal His_6-tag with a thrombin cleavage site between the His_6-tag and the N-terminal start codon, under the control of the T7 promoter.

Forward and reverse primers used for amplifying the truncated *SLI3188* from the pUC19 vector. The restriction sites for NdeI and BamHI are underlined.

GlxA - Forward

5′-ATACATATGAACGGGCCGTGGCTCTACCGCTTC-3′

GlxA - Reverse

5′-ATAGGATCCGGCGCTACGGAACTCGCACCCAC-3′

Table 2.1: Reagents used in the PCR protocol to amplify *glxA* (35-645 residues)

Reagent	Concentration	Volume (μl)
Plasmid DNA	N/A	2.5
Forward Primer	75 ng/μl	1.0
Reverse Primer	75 ng/μl	1.0
dNTP's	10 mM	2.5
10X Buffer Pfu + $MgSO_4$	-	5.0
DMSO	100 %	2.5
Sterile, deionized water	-	35.0
Pfu DNA polymerase	-	0.5
Total	-	50.0

Table 2.2: The PCR protocol used for the amplification of *glxA* (35-645 residues*)*. Numbers highlighted in bold were cycled 35 times.

Temperature (°C)	Time (min)
95	3
95	**1**
56	**1**
72	**2**
72	7.5
4	Final Extension

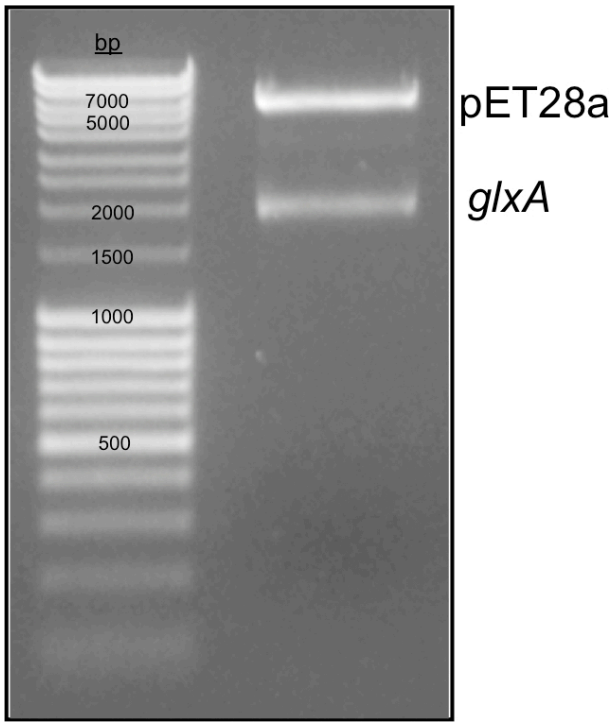


Figure 2.2: DNA agarose gel (1 % w/v) of *glxA (SLI3188)* pET28a construct. DNA consisting of the *pET3188* construct was digested using NdeI and BamHI restriction enzymes and corroborated against the DNA ladder (far left lane) (pET28a – 5,369 bp and *glxA* – 1,863 bp).

2.2.3. Over-expression and purification of GlxA

pET3188 encoding for the N-terminal truncated GlxA protein (residues 35-645) was transformed into chemically competent *E. coli* BL21 (DE3) cells using the heat-shock method. Cells containing the plasmid were selected by growing on 50 μg ml^{-1} Kan agar plates and single colonies transferred into 3 ml of LB medium (Melford) containing 50 μg ml^{-1} Kan. These were incubated overnight at 37 °C shaking at 180 rpm. Overnight pre-cultures were successively used to inoculate 750 ml of 2xYT medium in 2 L flasks and incubated at 37 °C, 220 rpm. At an OD_{600} of 0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG; Melford) was added to the cultures to a final concentration of 1 mM, the temperature then decreased to 25 ^{o}C and the cultures incubated for a further 16 h. Cells were harvested at 3,501 g for 20 min at 4 °C (using a Sorvall SLC6000 rotor) and the pellets re-suspended in buffer A (50 mM Tris/HCl, 500 mM NaCl, 20 mM imidazole, pH 8). The cell suspension was then lysed using an EmulsiFlex-C5 cell disrupter (Avestin) followed by centrifugation at 38,724 g for 20 min at 4 °C (using a Sorvall SS34 rotor). The clarified supernatant was loaded to a 5 ml Ni^{2+}-NTA Sepharose column (GE Healthcare) equilibrated with buffer A and the protein eluted with a linear imidazole gradient using buffer B (buffer A with 500 mM imidazole) using an ÄKTA purification system. A single peak at ~20 % buffer B was eluted from the column and fractions pooled and dialysed overnight at 4 °C against buffer C (50 mM Tris/HCl, 150 mM NaCl, pH 8). Following dialysis, the N-terminal His_6-tag was removed by incubating the protein at room temperature overnight in the presence of 125 U of thrombin (Sigma). The protein/thrombin mixture was then reapplied to the Ni^{2+}-NTA Sepharose column and the flow-through collected and concentrated at 4 ^{o}C using a centricon (vivaspin) with a 30 kDa cut-off. Concentrated protein was loaded to a S200 Superdex column (GE Healthcare) equilibrated with buffer C. A major peak eluted with at a retention volume of ~80 ml consistent with a monomer species with a mass in the region of 70 kDa, (actual mass of GlxA 67,456.8 Da). Samples from this peak were assessed for purity by 15 % SDS-PAGE (see section 2.2.4) and UV-visible spectroscopy. Protein was then concentrated and stored at 4 ^{o}C for future use.

2.2.4. Visualisation of GlxA by SDS-PAGE

Sodium dodecyl sulphate polyacrylamide gels (SDS-PAGE) (15 % w/v) were prepared using a mini-protean tetra cell BioRad system with the compositions shown in Table 2.3. Cell suspensions or purified protein were mixed with 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 ^{o}C for 10 min for cells and 3 min for pure protein. Samples were then loaded to the gel with a page-ruler protein marker

used to verify the size of the protein bands. Gels were run at 170 V, for ~1 h then 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 2.3: Components required for the construction of two 15 % SDS PAGE mini gels.

Chemical	Separating (μl)	Stacking (μl)
Water	2300	2250
30 % Acrylamide/Bisacrylamide 29:1	5000	666
1.5 M Tris/HCl pH 8.8	2500	0
0.5 M Tris/HCl pH 6.8	0	1000
10 % SDS	100	40
10 % APS (Ammonium persulphate)	100	40
TEMED (Tetramethylethylenediamine)	5	4

2.2.5. UV-visible spectroscopy

Concentrations of GlxA were determined using a Varian Cary 50 UV-visible spectrophotometer with a 1 cm path-length Quartz cuvette (Hellma) at 20 ^{o}C. The absorbance at 280 nm was measured and concentrations calculated using the Beer-Lambert law with an extinction coefficient (ε) of 78,730 M^{-1} cm^{-1} at 280 nm for GlxA determined using ProtParam ExPASy [93]. Various forms of GlxA were prepared by either the addition of a > 50-fold excess of ferricyanide ($[Fe(CN)_6]^{3-}$) (Sigma) or irdium chloride ($[Ir(Cl)_6]^{3-}$) (Acros) or the addition of excess sodium dithionite ($Na_2S_2O_4$) (Sigma) followed by removal of either oxidant of reductant using a PD-10 desalting column (GE Healthcare). For Cu(II) binding to GlxA a stock of $CuSO_4$ was prepared and diluted as required. Cu(II) was then added to GlxA at a 1:1 equivalent in 50 mM Tris, 150 mM NaCl, pH 8 and left for ~12 h with a scan taken every 5 minutes between 1000-200 nm using a Varian Cary 50 spectrophotometer. Excess Cu was then removed using a PD-10 column and the UV-visible spectra recorded and further experiments conducted.

2.2.6. Circular Dichroism spectroscopy

GlxA samples used for circular dichroism (CD) analysis were exchanged into 10 mM KPi, 50 mM KF, pH 7 using a PD-10 desalting column (GE Healthcare). Far UV-CD spectra were recorded between 260 and 175 nm and visible CD between 800 and 300 nm at 20 ^{o}C measured on an Applied Photophysics Chirascan CD spectrophotometer equipped with a thermostatic cell holder controlled with a Peltier system (Leatherhead, UK).

2.2.7. EPR spectroscopy

Cu(II)-GlxA was analysed by measuring the electron paramagnetic resonance (EPR) of the protein as isolated (85 μM) in a mixed buffer system of 10 mM Tris, potassium acetate, MES, MOPS and 200 mM potassium chloride with pH varying from 7 to 10. Oxidant, potassium ferricyanide ($K_3[Fe(CN)_6]$) and reductant, $Na_2S_2O_4$ (Sigma) were added in excess to the protein (120 μM) followed by removal using a PD-10 desalting column before measurement. Wilmad SQ EPR tubes were filled with 250 μl of the corresponding GlxA solutions, frozen in methanol, kept on dry ice then wiped and transferred to liquid nitrogen. EPR spectra were measured at 10 K (pH effects) and 40 K (redox centre) on a Bruker EMX EPR spectrophotometer (X-band) at a modulation frequency of 100 kHz. A spherical high-quality Bruker resonator SP9703 and an Oxford instruments liquid helium system were used to measure the low temperature EPR spectra. The EPR spectra of the blank samples (frozen water) were subtracted from the EPR spectra of the protein samples to eliminate the baseline caused by the resonator walls, quartz insert or quartz EPR tube. The baseline was corrected by subtraction of a polynomial line drawn through a set of points randomly chosen on the baseline. Digitising of a published EPR spectrum was performed using Un-Scan-It, v.6, Silk Scientific.

2.2.8. Crystallisation and structure determination of GlxA

Cu(II)-GlxA (15 mg ml^{-1}) was first used to screen a range of crystallisation conditions using an ARI-Gryphon 96-well crystallisation robot. Initial crystal hits were discovered predominately in the PEG suite (Qiagen). Scaling-up and optimisation of GlxA crystals from the initial hits was carried out in 24-well plates using the hanging drop vapour diffusion method at 18 ^{o}C. 1 μl of protein solution at 15 mg ml^{-1} was mixed with an equal volume of reservoir solution containing 20 % polyethylene glycol (PEG) (Sigma) 20 K, 0.1 M sodium acetate, pH 4. Long needle crystals suitable for diffraction grew within 48 h. Single crystals were transferred into a cryoprotectant solution containing 20 % PEG 20 K, 0.1 M sodium acetate, pH 4 and 20 % glycerol, prior to flash cooling to 100 K by plunging into liquid nitrogen. Crystallographic data was measured to 1.77 Å resolution at the I03 beamline, Diamond Light Source using an X-ray wavelength of 0.979 Å and a Pilatus 6-M-F detector (Dectris).

The diffraction data obtained from the GlxA crystals were indexed using iMosflm [94] and scaled and merged using aimless [95] in the CCP4i suite. A model from Chainsaw [96], using the last common atom function was created using the PDB entry 2EIE as the alignment model. This model was used for molecular replacement using Molrep giving a solution for one of the two molecules in the asymmetric unit that was predicted from the Matthews coefficient. ARPwARP [97] and Refmac5 [98] were used to trace and build the

initial electron density and PhaserMR [99] to determine the second, less well-ordered molecule in the asymmetric unit. Models were rebuilt between refinement cycles in Coot [100] and Refmac5 and riding hydrogen atoms added when refinement of the protein atoms had converged. The final structure was validated using the Molprobity server, Coot and the JCSG Quality Control Server [101-105]. A summary of data and refinement statistics and the quality indicators for the structure are given in Table 2.4 Coordinates and structure factors were deposited in the RCSB Protein Data Bank with accession number 4UNM.

Table 2.4: X-ray data processing and refinement parameters of GlxA. The GlxA crystal structure was in space group $P2_1$ with unit cell parameters 50.4, 126.6, 107.6 Å, 90, 91.1, 90 º. Values in parentheses refer to the outermost resolution shell (1.80-1.77 Å).

Resolution (Å)	53.3-1.77
Unique reflections	130533 (21898)
Completeness (%)	99.7 (99.6)
Redundancy	3.6 (3.4)
R_{merge} (%)	0.064 (0.633)
Mn (I/Sd)	10.5 (2.0)
Wilson B factor ($Å^2$)	20.5
R_{cryst}	0.188
R_{free}	0.228
RMSD bond lengths (Å)	0.011
RMSD bond angles (^{o})	1.41
ESU based on ML (Å)	0.096
Ramachandran favoured (%)	95.9
PDB accession code	4UNM

2.2.9. Computation of tunnels in the GlxA structure

CAVER analyst 1.0 was used for the identification of tunnels in GlxA using molecule B containing the long β-hairpin loop [106]. The tunnels were calculated using the tunnel computation tool within CAVER defined with a starting point using the first coordination sphere Cu binding residues (Tyr^{289}, Tyr^{501}, His^{502}, His^{589}). Default settings for tunnel analysis were used, including the minimum probe radius set to 0.9 Å.

2.2.10. Enzymatic activity of GlxA

Enzyme assays were carried out using a variety of putative substrates. These included the following monosaccharides; D-galactose, D-glucose, D-sucrose, D-fructose, N-acetyl-D-glucosamine, and D-glucuronic acid; the disaccharides, D-lactose and D-cellobiose; and the aldehydes and primary alcohols glycolaldehyde, glyoxal and glycerol (all purchased from Sigma). Catalytic turnover of GlxA with these substrates was measured using a horseradish peroxidase (HRP)-coupled assay. The ability of GlxA to function as an oxidase would lead to the production of H_2O_2, which would then be detected by HRP (Sigma) leading to the subsequent oxidation of guaiacol (Sigma). Samples were prepared in 3 ml quartz cuvettes containing 0.1 M sodium phosphate, pH 7.4, 1 mM guaiacol, 1 µl HRP (10 mg ml^{-1}), 20 - 30 µM GlxA and varying concentrations of the individual substrate. The oxidation of guaiacol was monitored at 470 nm using a Hewlett-Packard 8453 diode-array spectrophotometer scanning between 190 and 1100 nm and thermostatted at 25 ^{o}C. Time courses of the absorbance increase at 470 nm at varying substrate concentrations allowed the initial rate ($\Delta A_{470\,nm}/t$) to be calculated in which ΔA_{470} is the absorbance change at 470 nm and t is the time in seconds. Plots of turnover rate constant (k, s^{-1}) versus substrate concentration were then constructed, whereby k which is the initial rate normalised to the enzyme concentration calculated from $((\Delta A_{470}/\varepsilon_{gc})/t)/[GlxA])$ where $\Delta A_{470}/t$ is the initial rate at 470 nm upon guaiacol oxidation, ε_{gc} is the extinction coefficient of the guaiacol oxidation product taken as 5.57 mM^{-1} cm^{-1}, and [GlxA] is the total millimolar concentration of GlxA in the reaction. The average of triplicate experiments was plotted and the apparent affinity, turnover number and catalytic efficiency were estimated by non-linear least-squares fitting to the Michaelis–Menten model (see Equation 2.1) (whereby, k is the reaction rate normalized to enzyme concentration, [S] is the substrate concentration, k_{cat} is the maximum rate and K_m is the substrate concentration at which the reaction rate is half of V_{max} (maximal velocity)).

$$k = \frac{k_{cat}[S]}{K_m + [S]} \qquad \text{Equation 2.1}$$

2.2.11. Construction of Δsco, ΔcslA and ΔglxA null mutants and complementation plasmids

The following experimental procedures were carried out by collaborators at Leiden University (Dr Erik Vijgenboom, Dr Dennis Claessen and Marloes Petrus).

Genes *SLI4214, SLI3187* and *SLI3188* encoding Sco, CslA and GlxA, respectively, were deleted in *S. lividans* 1326 in a two-step process using the CRE-lox system as has previously been described in Blundell *et al.* [37]. In the Δ*sco* mutant nucleotides + 33 to + 490, and in Δ*glxA* nucleotides + 60 to + 1916 relative to the start codon of *SLI4214* or *SLI3188* respectively were replaced by a 62 nt scar of the loxP recombination site including two XbaI sites. In the Δ*cslA* mutant, nucleotides + 79 to + 1827 relative to the start of *SLI3187* were replaced, consistent with the one described by Xu *et al.* [65]. For the complementation plasmids the pGlxA plasmid contains the *glxA* ORF (*SLI3188*) and the promoter region (-545 to -1) upstream of *cslA* (*SLI3187*) and the pSco plasmid contains the *sco* ORF (*SLI4214*).

2.2.12. Examination of growth morphology

The following experimental procedures were carried out by collaborators at Leiden University (Dr Erik Vijgenboom, Dr Dennis Claessen and Marloes Petrus).

The morphology of *Streptomyces* was monitored by its growth on solid R5 agar medium (a complex medium containing yeast extract and glucose as a carbon source) and by growth in liquid cultures both supplemented with Cu as required (concentrations depicted in Figure legends). Morphology in liquid media was determined following growth in shaking flasks with coils containing tryptic soya broth (TSB) supplemented with 10 % sucrose and Cu as indicated. All solid and liquid-grown cultures were incubated at 30 ^{o}C. After 24 h, mycelium morphology was recorded by light microscopy with a Zeiss Standard 25 microscope and digital pictures taken with an AxioCam camera linked to AxioVision software. All spore stocks were attained from cultures grown on MS (Mannitol salt) agar plates stored in 20 % glycerol at -20 ^{o}C until use.

2.2.13. Glycan production analysis in WT, ΔcslA and ΔglxA in S. lividans

The following experimental procedures were carried out by collaborators at Leiden University (Dr Erik Vijgenboom, Dr Dennis Claessen and Marloes Petrus).

To investigate the synthesis of glycans or their attachment at hyphal tips images of stained mycelia were taken in WT, *ΔcslA* and *ΔglxA* of *S. lividans*. Glycan production was measured in eight-well microscopy plates (Lab-TEK II Chambered Coverglass) containing 500 ml of mNMMP medium (2 g of $(NH_4)_2SO_4$, 5 g of Difco casamino acids, 0.6 g of $MgSO_4 \cdot 7H_2O$, 50 g of PEG 6000, 1 ml of minor elements solution, 15 mM NaH_2PO_4, K_2HPO_4buffer (pH 6.8), 0.5% glucose per liter supplemented with 25 mM mannitol) following growth of the mycelium after 24 h at 30 ^oC in standing liquid cultures [107]. The mycelium were stained by the addition of 50 μl of calcofluor white (CFW) solution (Remel Bactidrop) to each well, which binds to β-(1-4) glycans. The mycelium were incubated for 5-20 min then analysed with a laser-scanning confocal microscope (Zeiss LSM5 Exciter/Axio Observer) by excitation with a 405 nm laser, a 405/488 nm beamsplitter and a bandpass emission filter of 420-480 nm. Images were shown as fluorescence and phase contrast brightfield adjusted for brightness and contrast using Image J [89].

2.3. Results

2.3.1. Amino acid sequence alignment of S. lividans GlxA

Analysis of the genome of *S. lividans* has revealed a number of genes that encode for putative extracytoplasmic cuproenzymes [4, 5]. The GlxA amino acid sequence was aligned with the GlxA homolog (SCO2837) from *S. coelicolor* and the secreted fungal Gox from *F. graminearum* using Clustal Omega (Figure 2.3) [108]. GlxA from *S. lividans* shares 27.3 % identity to Gox from *F. graminearum,* but contains only eight amino acid differences compared to GlxA from *S. coelicolor.* Therefore, *S. lividans* GlxA is almost identical to that from *S. coelicolor* and has very weak sequence homology with Gox. However, the putative ligands to the Cu ion (two Tyr and two His residues) and the Cys residue cross-linking to one of the Cu coordinating Tyr residues are conserved between both GlxA species and Gox (Figure 2.3).

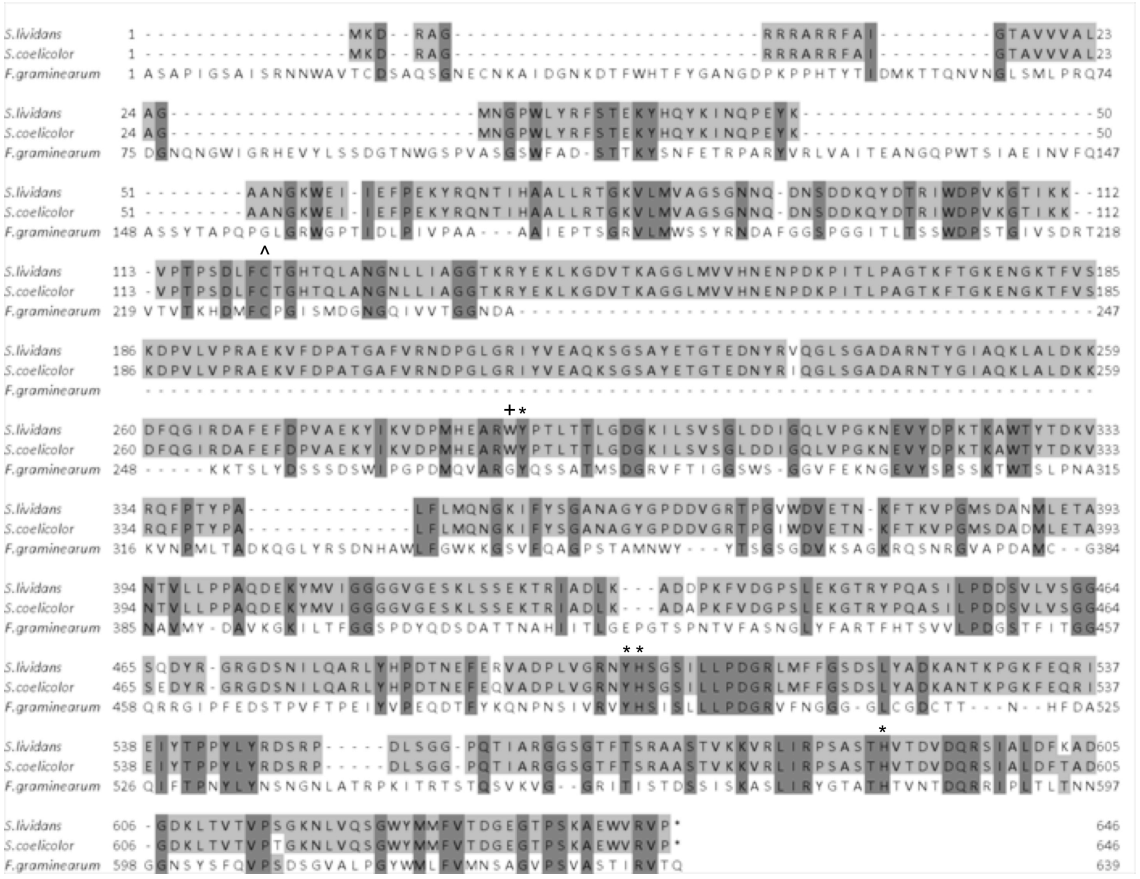


Figure 2.3: Clustal Omega sequence alignment of *S. lividans* GlxA with *S. coelicolor* GlxA and Gox from *F. graminearum* [108]. Sequence conservation is indicated in dark blue to light grey respectively, made using Jalview [109]. The * symbol indicates the sequence position of ligands coordinating to the Cu ion (two His and two Tyr) the ^ symbol shows the Cys residue that forms the Cys-Tyr crosslink and + the location of the second coordination sphere stacking Trp residue.

2.3.2. GlxA is associated with the membrane

To determine the location of GlxA within *S. lividans* the soluble and insoluble portions of the mycelium following various washes was investigated through western blotting. After sonication and centrifugation, examination of the fractions identified a small amount of GlxA in the soluble protein portions (S100) but the majority found in the insoluble pellets (P30 and P100) (Figure 2.4.A and 2.4.B). This differs to the cytoplasmic marker EF-Tu1 in which the majority is identified in the soluble fraction (S100) (Figure 2.4.A and 2.4.B). When the insoluble pellet was resuspended in buffer containing 1 % Triton X-100 a large amount of GlxA becomes solubilized, indicating that GlxA is not covalently bound to the membrane. When the mycelium were washed with salt GlxA was not removed and remained in the pellet fraction (Figure 2.4.C). Small amounts of EF-Tu1 were detected in the soluble fractions possibly due to some lysis. GlxA is therefore shown to be firmly associated with the mycelium, although not via a covalent bond.

In silico analysis of *S. lividans* GlxA predicts residues 1-11 to be a N-terminal signal peptide [110], with a weak signal peptidase cleavage site between residues 11-12, followed by a transmembrane helix from residues 12-32, which could function as a membrane anchor. Topology modelling also predicted the presence of a transmembrane region (13-32) (section 2.2.2). This was also the case for the homologous GlxA from *S. coelicolor* secreted by a tat-dependent pathway [111]. Therefore, to create a soluble form of GlxA an N-terminal truncated construct consisting of residues 35-645 (Δ1-34) was prepared.

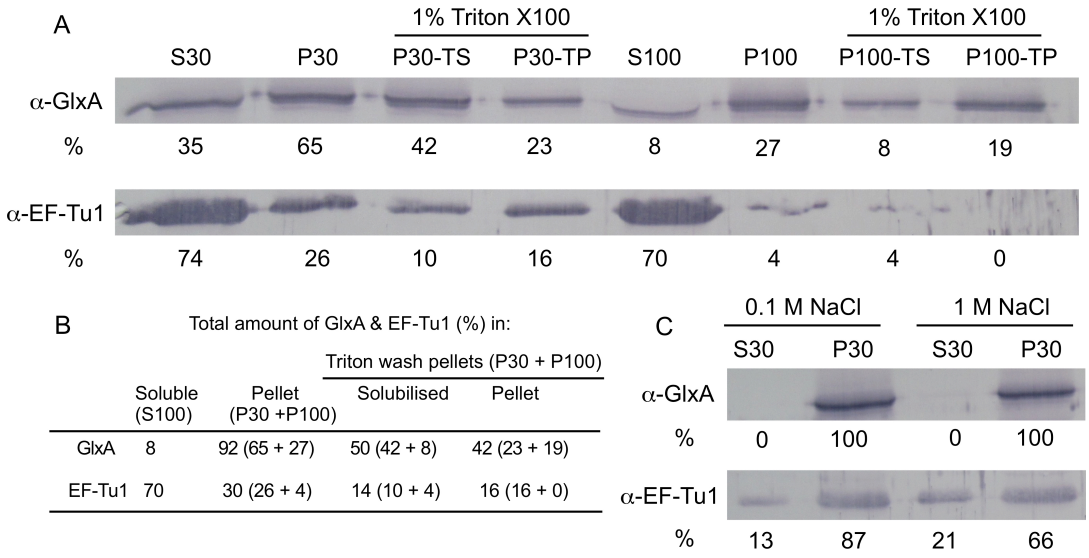


Figure 2.4: GlxA location in the mycelium determined by Western blotting. A) Detection of GlxA and EF-Tu1 in the various mycelium fractions. Soluble (S30) and insoluble fractions (P30). S30 was centrifuged to remove all membrane particles and ribosomes to give the S100 and P100 fractions. Pellets (P30 and P100) were resuspended in buffer containing 1 % Triton X-100, and centrifuged, to separate triton soluble (P30-TS, P100-TS) and insoluble (P30-TP, P100-TP) fractions (See experimental 2.2.1). The band intensities are expressed as percentages relative to the total amount of GlxA or EF-Tu1 in the S30 plus P30 fraction, of which the intensity was set at 100 %. B) The Table presents the total soluble (S100), 1 % triton soluble (P30-TS + P100-TS) and insoluble fraction (P30-TP + P100-TP) of GlxA and EF-Tu1 expressed in percentage according to (A). C) NaCl washes of the intact mycelium. The signals detected in the S30 plus P30 samples of the 100 mM NaCl wash was set at 100 %.

2.3.3. Over-expression and purification of GlxA

GlxA (residues 35-645) was expressed in *E. coli* in 2xYT media and induced with IPTG. Over-expression was detected using 15 % SDS-PAGE gels (see experimental 2.2.4) with samples taken before and after induction. A single band can be visualised between 70-80 kDa corresponding to GlxA (69 kDa before His_6-tag removal), this band increases in intensity over growth time, until harvesting ~ 16 h after induction (Figure 2.5.A). The overall yield of GlxA expressed in *E. coli* is ~ 45 mg L^{-1}. GlxA was purified by first using a Ni^{2+}-NTA sepharose column producing a single broad peak at 280 nm at ~ 40 ml elution volume and ~ 30 % buffer B (Figure 2.5.B). After His_6-tag removal using thrombin, GlxA was further purified by application to a S200 size exclusion column. This produced a clear single peak with an elution volume of ~ 80 ml consistent with a monomeric protein species (Figure 2.5.C). Fractions eluting from the major peak were analysed on a 15 % SDS-PAGE gel (see experimental 2.2.4) displaying a single pure band running at ~ 75 kDa consistent with the approximate molecular weight of GlxA (67 kDa after His_6-tag removal) (Figure 2.5.D).

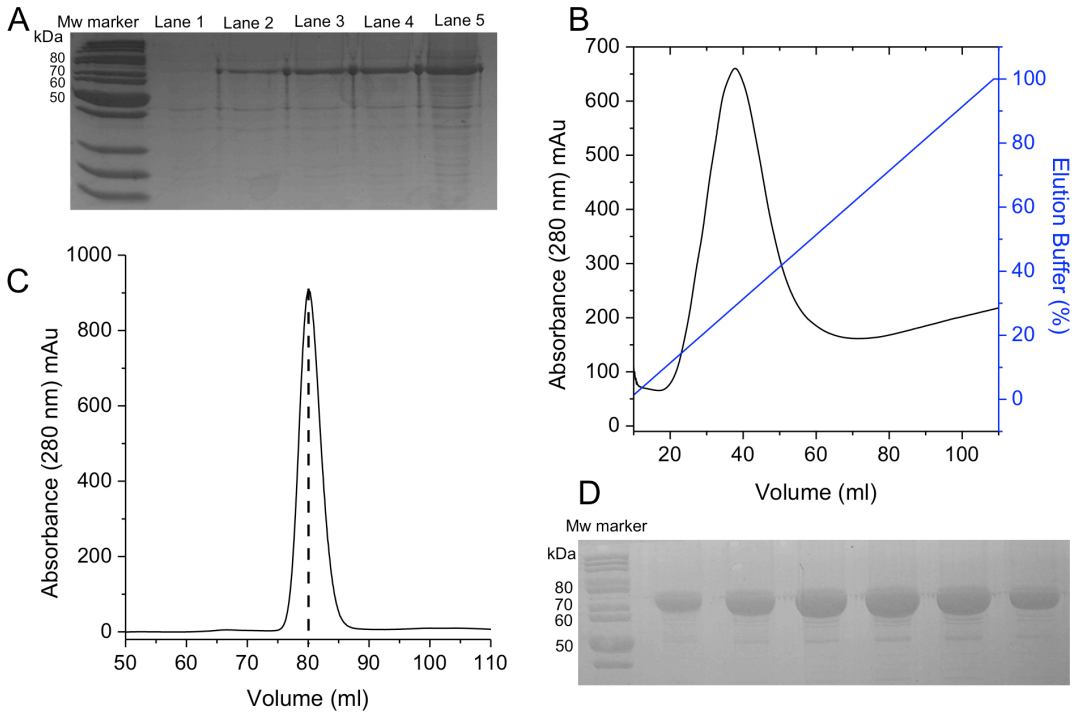


Figure 2.5: Over-expression and purification of GlxA. A) A coomassie stained 15 % SDS-PAGE gel of GlxA over-expression in *E. coli*. Lane 1: GlxA cell growth fraction before induction with IPTG. Lane 2: Cell pellet after 1 h of growth after IPTG induction. Lane 3: 2 h after induction. Lane 4: 3 h after induction. Lane 5: 18 h after cell induction with IPTG. A single protein band between 70-80 kDa can be observed. B) Ni^{2+}-NTA profile of GlxA showing an absorbance peak (mAu) at 280 nm in black at ~30 % buffer B shown in blue. C) S200 size-exclusion Sephadex column absorbance 280 nm (mAu) profile of GlxA eluting at ~ 80 ml in 20 mM NaPi, 100 mM NaCl, pH 7. D) Coomassie stained 15 % SDS-PAGE gel analysis of GlxA fractions from an S200 size-exclusion Sephadex column. A single protein band between 70-80 kDa can be identified as GlxA (after cleavage of the His_6-tag).

2.3.4. UV-visible spectroscopic analysis of the Cu site in GlxA

UV-visible spectroscopy was used to monitor spectral changes upon addition of Cu, oxidants or reductants to GlxA. As isolated, GlxA is blue/grey in colour and gives rise to specific spectral transitions, comprising of a broad, low intensity band in the visible region λ_{max} ~ 577 nm and two distinct shoulders at ~ 362 and 320 nm (Figure 2.6). No change in these transitions was observed over the pH range 4-8. Addition of $Cu(II)SO_4$ to GlxA results in only minor absorbance increases at λ_{max} ~ 577 nm and slight shifts in the shoulders over a ~ 12 h period (Figure 2.6.A) indicating that the protein as isolated is Cu(II) loaded. The absorbance spectrum differs to that of Cu(II)-Gox which has weak intensity bands at λ_{max} 441 nm due to phenolate (Tyr) to Cu(II) ligand-to-metal charge transfer (LMCT) and at λ_{max} 630 nm arising from mixed Cu ligand field transition and LMCT [81, 83]. Cu(II)-GlxA from *S. coelicolor* also differs in absorbance spectrum to both Cu(II)-Gox and *S. lividans* Cu(II)-GlxA with λ_{max} ~ 570 nm and λ_{max} ~ 800 nm [111]. Upon addition of the oxidants, $[Fe(CN)_6]^{3-}$ or $[Ir(Cl)_6]^{3-}$ to Cu(II)-GlxA, the absorbance spectrum is not altered (Figure 2.6.B). In contrast, Gox and *S. coelicolor* GlxA when fully oxidised to the Cu(II)-Cys-Tyr• form are reported to be an intense green colour and produce distinct spectral features with an absorption band λ_{max} 445 nm and a broad feature in the near infrared region with λ_{max} ~ 810 nm [81, 83]. The intense peak in the visible region arises from overlapping LMCT between Tyr and Cu π to π^* ring transitions and the red peak at ~ 810 nm has been assigned to a mixture of LMCT and charge resonance excitation between aromatic systems [81, 83]. In addition, a shoulder can be observed in the near UV at about 320 nm and weaker absorption in the 500-700 nm region [81]. Addition of $Na_2S_2O_4$ to GlxA, however, bleaches all absorption features indicative of a Cu(I) spectrum implying that the Cu is redox active and can cycle between Cu(II) and Cu(I) (Figure 2.6.B). Removal of $Na_2S_2O_4$ and subsequent exposure to air results in a return of the initial GlxA spectrum. Therefore, the spectrum of GlxA is unique and does not resemble either the one electron oxidised (Cu(II)) or the two electron oxidised (Cu(II)-Tyr•) forms of fungal Gox or GlxA from *S. coelicolor*.

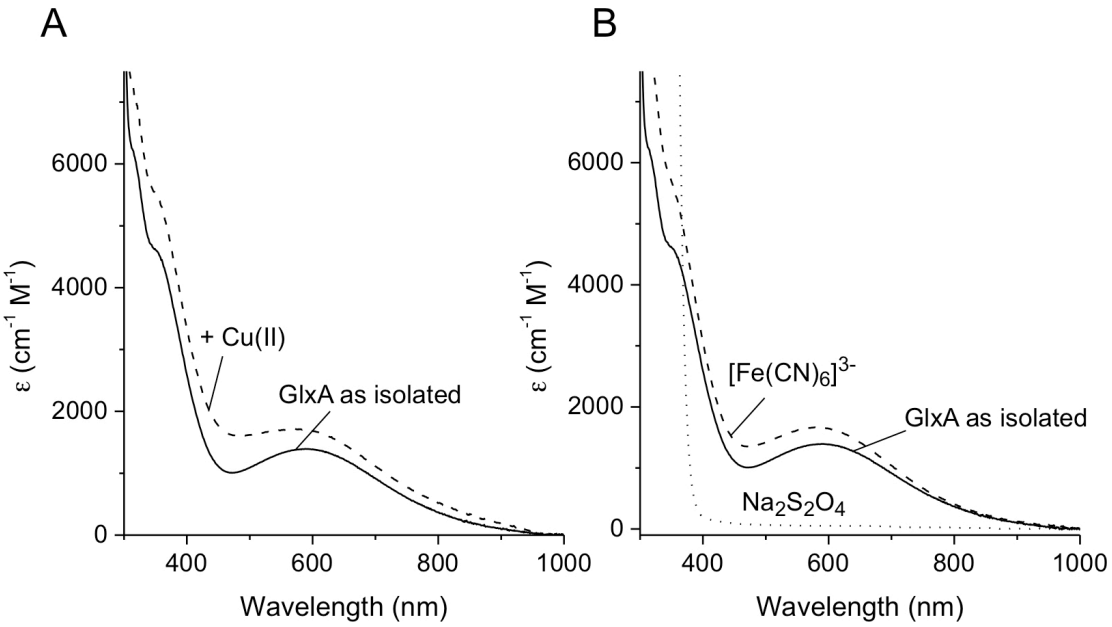


Figure 2.6: UV-visible absorbance spectra of GlxA. A) GlxA as isolated and after the addition of 1 equivalent Cu(II)SO_4 and incubated for ~ 12 h. B) The effect of oxidizing and reducing agents on the UV-vis spectrum of GlxA as isolated, after the addition of 50-fold excess $[Fe(CN)_6]^{3-}$ and after the addition of excess $Na_2S_2O_4$. GlxA concentration was 80 μM in 50 mM Tris/HCl, 150 mM NaCl, pH 8 and 20 ^oC.

2.3.5. Secondary structure examination of GlxA

Secondary structure prediction of GlxA was determined *in silico* using Pro-origami. This creates a 2D protein topology diagram of structures with complex folds [112]. It predicted GlxA to consist of a largely β-sheet fold with the presence of 51 β-sheets (and 4 α helices) in comparison to Jpred4 (another secondary structure prediction program), predicting 41 β-sheets (Figure 2.7.A) [113]. The topology map also predicts the orientation of the secondary structural elements. From analysis of the Pro-origami output GlxA begins with a central core of β-sheets encompassing the majority of the sequence. This then extends up into a domain containing six antiparallel β-sheets and the map then ends in a domain below the central fold containing 7 β-sheets with one of these sheets extending up into the central domain. Far UV-CD of GlxA confirms that the protein is folded and also displays features consistent with predictions that the structure largely consists of β-sheets revealed by the minima at 200 nm and maxima at ~190 nm (Figure 2.7.B). Visible CD for GlxA as isolated revealed a spectrum with minima at 350 nm and maxima at ~ 320 and 650 nm noticeably different to Gox but still confirms the presence of a coordinated Cu(II) ion (Figure 2.7.C). Visible CD of Gox displays significant differences between fully oxidised (Cu(II)-Tyr•), as isolated (Cu(II)) and reduced (Cu(I)) with minor similarities to peaks identified in GlxA [81]. No noticeable alterations were identified after treatment of GlxA with the oxidant $[Fe(CN)_6]^{3-}$ in contrast to Gox (Figure 2.7.C).

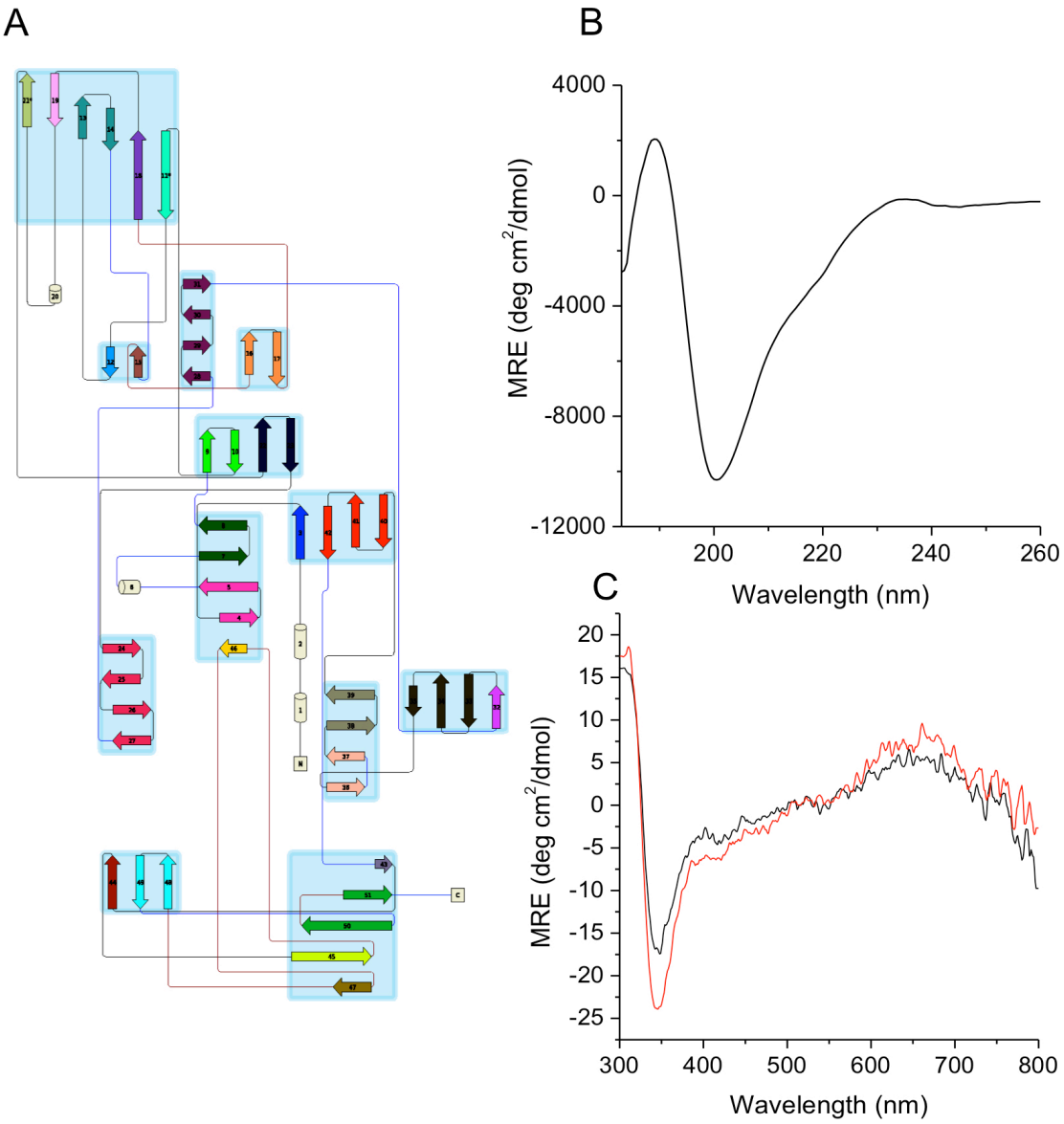


Figure 2.7: Secondary structural analysis and CD spectroscopy of *S. lividans* GlxA. A) The secondary structure topology map of GlxA showing 51 β-sheets and the predicted 2D assembly (Pro-origami) [112]. B) Far UV-CD of isolated GlxA (2 µM) recorded between 260 and 180 nm at 20 ^{o}C. C) Visible CD of GlxA (142 µM) measured between 800 and 300 nm 20 ^{o}C. Black is GlxA as isolated and red is after the addition of > 50 fold excess $[Fe(CN)_6]^{3-}$. All CD spectroscopy was carried out in 10 mM KPi, 50 mM KF, pH 7.

2.3.6. Electron Paramagnetic Resonance spectroscopy of Cu(II)-GlxA

The Cu(II)-GlxA EPR spectrum at pH values between 7-10 display features consistent with an axial Cu(II) g-tensor ($g_x = g_y > g_z$), $g_{\parallel} = 2.182$ and $g_{\perp} = 2.05$ (Figure 2.8). The $g_{\parallel}$ (parallel) is split into a quartet 1:1:1:1 due to the electron spin (S = ½) interacting with I = 3/2 ^{63}Cu nuclear spin. The strong line on the right from the indicated $g_{\perp}$ (perpendicular) region is not a principal g-factor component but an "overshoot" line, [114] this occurs for specific orientations of the Cu(II) complexes, due to particular relationships between the values of $g_{\parallel}$, $g_{\perp}$ and the anisotropic components of the hyperfine interaction of the electron spin (S = ½) with the Cu nuclear spin (I = 3/2), $A_{\parallel}$ and $A_{\perp}$. The spectrum line shape of GlxA is similar to that of Cu(II)-Gox [86], which also has an axial Cu(II) g-tensor [115-117]. However, the GlxA spectrum displays a more clearly resolved overshoot line, likely to be a consequence of a smaller $g_{\parallel}$ (2.182 as compared to 2.21-2.23 in Gox). The $g_{\perp}$ component displays a hyperfine interaction of the electron spin (S = ½) with the Cu(II) ligands, which is best seen at pH 7 and disappears at higher pH values (Figure 2.8, inset). For Cu(II) with electron spin S = ½, interacting with two equivalent (equidistant) nitrogen nuclei spins I = 1, such as two His ligands, a quintet structure of lines 1:2:3:2:1 with a distance between the lines typical for a N hyperfine interaction (14 - 18 G) should be observed as illustrated in Gox [118, 119]. However, at pH 7 for GlxA there are seven lines with relative intensities of 1:4:8:10:8:4:1, not five lines, indicating that an additional hyperfine interaction is present, which is not observed in the EPR spectra of Gox (Figure 2.8) [116, 117].

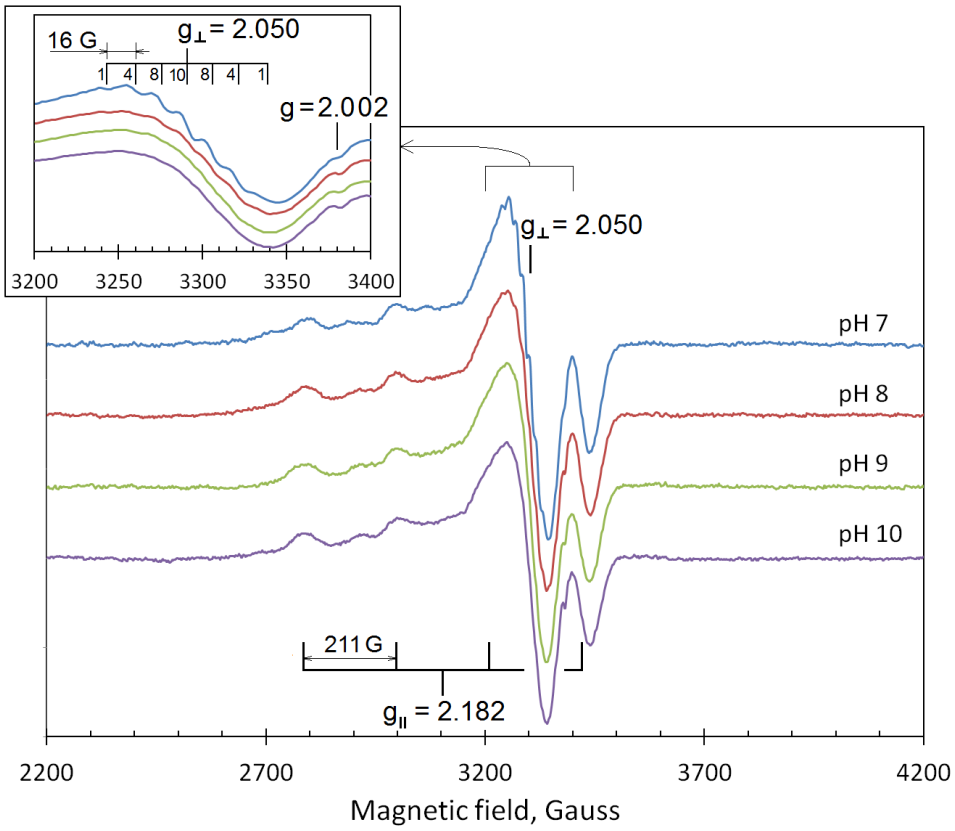


Figure 2.8: Cu(II)-GlxA EPR spectra (85 μM) at various pH values. The $g_{\|}$ component shows four components 1:1:1:1 separated by A_Z^{Cu} = 211 G. *inset* the $g_{\perp}$ region of the spectrum at the four pH values, which displays 7 components of relative intensities 1:4:8:10:8:4:1 separated by 16 G better seen at the lower pH values. Instrumental conditions: temperature 10 K, microwave frequency ν_{MW} = 9.47 GHz, microwave power P = 3.18 mW, modulation frequency ν_M = 100 kHz, modulation amplitude A_M = 5 G, time constant τ = 82 ms, scan rate V = 22.6 G s^{-1}, number of scans per spectrum NS = 1. GlxA samples were prepared in a mixed buffer system consisting of 10 mM each of Tris, potassium acetate, MES, MOPS and 200 mM KCl with the pH adjusted to the specific pH value.

2.3.7. Crystallisation of GlxA

In concentrated form GlxA is a distinctive blue/grey colour (Figure 2.9.A). Long needle-like crystals of GlxA grew within 48 h in 24-well plates (Figure 2.9.B). Single GlxA crystals were used for X-ray measurements after being transferred into a cryoprotectant and flash cooled to 100 K. Crystallographic data was measured to 1.77 Å resolution, at Diamond Light Source. GlxA crystals were also crushed and dissolved in ddH_2O and loading buffer and run on a 15 % SDS-PAGE gel (see experiment section 2.2.4). Figure 2.9.C displays the GlxA crystals illustrating a single band ~75 kDa corresponding to purified mature GlxA.

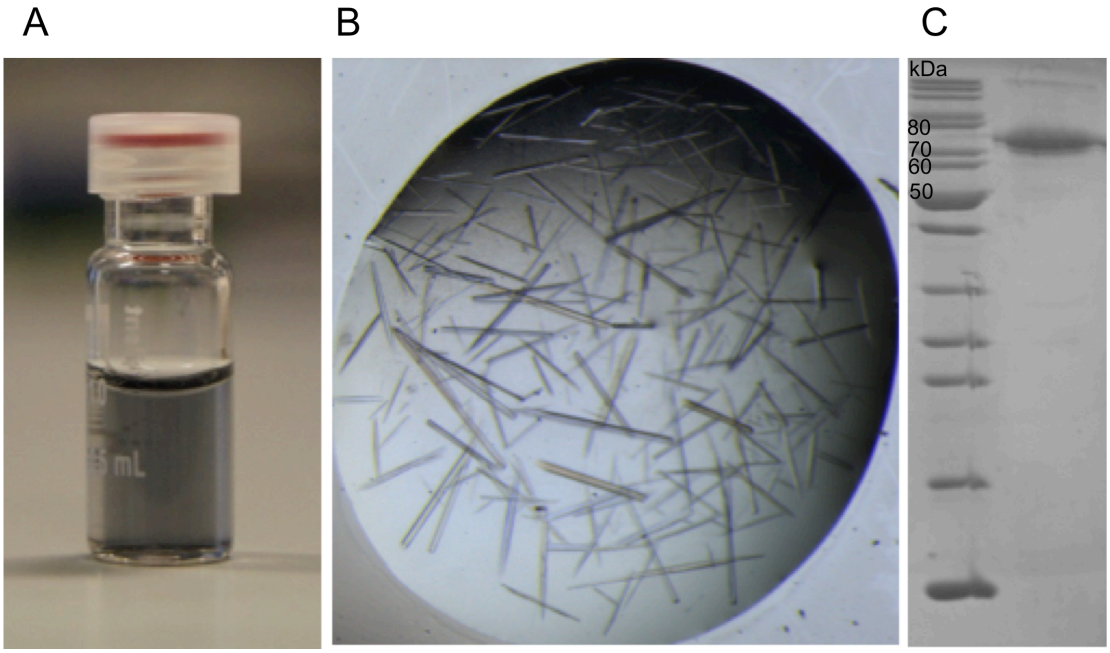


Figure 2.9: Crystallisation of GlxA. A) A vial of concentrated GlxA displaying the blue-grey colour of the protein. B) Needle-like crystals of GlxA in 20 % PEG 20 K, 0.1 M sodium acetate, pH 4. C) Crystals of GlxA dissolved in 20 mM NaPi, 100 mM NaCl, pH 7 buffer and ran on a 15 % SDS-PAGE gel showing a single band between 70-80 kDa.

2.3.8. X-ray crystal structure of GlxA

The X-ray crystal structure of GlxA was determined to 1.77 Å resolution with a space group of $P2_1$ with two molecules identified in the asymmetric unit. The orientation of molecule A (yellow) and molecule B (green) can be seen in Figure 2.10.A. Molecule A contains a higher quality electron density with significantly lower B–factors in which residues 38-645 were modelled. However, a disordered region is present between residues 198-206 creating a long β-hairpin loop (not visible in the electron density). Molecule B, however, contains a lower quality electron density but the main chain is complete between 37-645 residues (including the long β-hairpin loop not present in molecule A) (Figure 2.10.A). The overall structure of GlxA (67 kDa) (Figure 2.11) consists of an essentially β-sheet fold with three clear domains. The first domain is preceded by two short N-terminal α-helices (α1 and α2; Figure 2.11 shown in pink) with the α1 helix orientated approximately perpendicular to the α2 helix. Domain 1 (Figure 2.11, gold) is present after the two helices and consists of 7 Kelch motifs (blades) forming a large β-propeller structure. The blades of the propeller comprise four anti-parallel β-sheets (*a*, *b*, *c*, *d*) with each sheet connected through variable length loops (Figure 2.11). Blades 1 and 2 contain inserts in the loops connecting sheet *b* to sheet *c*, in blade 1 this is an α-helix (α3) and in blade 2 they are connected via domain 2 (Figure 2.11). Domain 2 (Figure 2.11, blue) consists of 10 β-sheets and 1 short α-helix (α4) with the CATH (class, architecture, topology and homologous superfamily) database [120] classifying it as mainly beta with a β-barrel design. The β-barrel core covers approximately half of domain 1, thus occluding the Cu site. The domain is formed by β-sheets 1, 3, 4, 8, 9 and 10 with a long β-hairpin loop (residues 194-210) between β-sheets, 6 and 7, which protrudes from the core of the β-barrel into the solvent (Figure 2.11). This long β-hairpin loop is only present in molecule B (Figure 2.10.A) (and not in A) due to the orientation of the two molecules in the asymmetric unit, molecule A can act to stabilise molecule B through polar interactions (Figure 2.10.B). Domain 2 in GlxA is novel compared to known structures determined using PDBeFold, which identified no structures within the PDB with a high structural homology, (the highest Q-score obtained was only 0.14 for a skich domain of NDP52 (3vvv)). Domain 3 (Figure 2.11, red) begins after the seventh Kelch motif orientated below domain 1. This domain forms a β-sandwich fold consisting of 8 β-sheets and one short α-helix and displays strong structural homology to domain 3 of Gox (1.18 Å RMSD from superimposition). Within domain 3 there is a loop (residues 583-590) connecting β-sheets 3 and 4, which extends upwards into domain 1 and provides one of the ligands (His^{589}) to the Cu ion. This ligand was predicted prior to X-ray structural information using Pro-origami *vide supra* (section 2.3.5) (Figure 2.7.A). Electrostatic surface representations of GlxA (pI 8.2) show a somewhat disperse distribution of charge across the

surface of the protein (Figure 2.12). However a large negative (red) area can be observed around the collar connecting domains 1 and 2, inferring the possibility of a substrate entry/exit site (Figure 2.12).

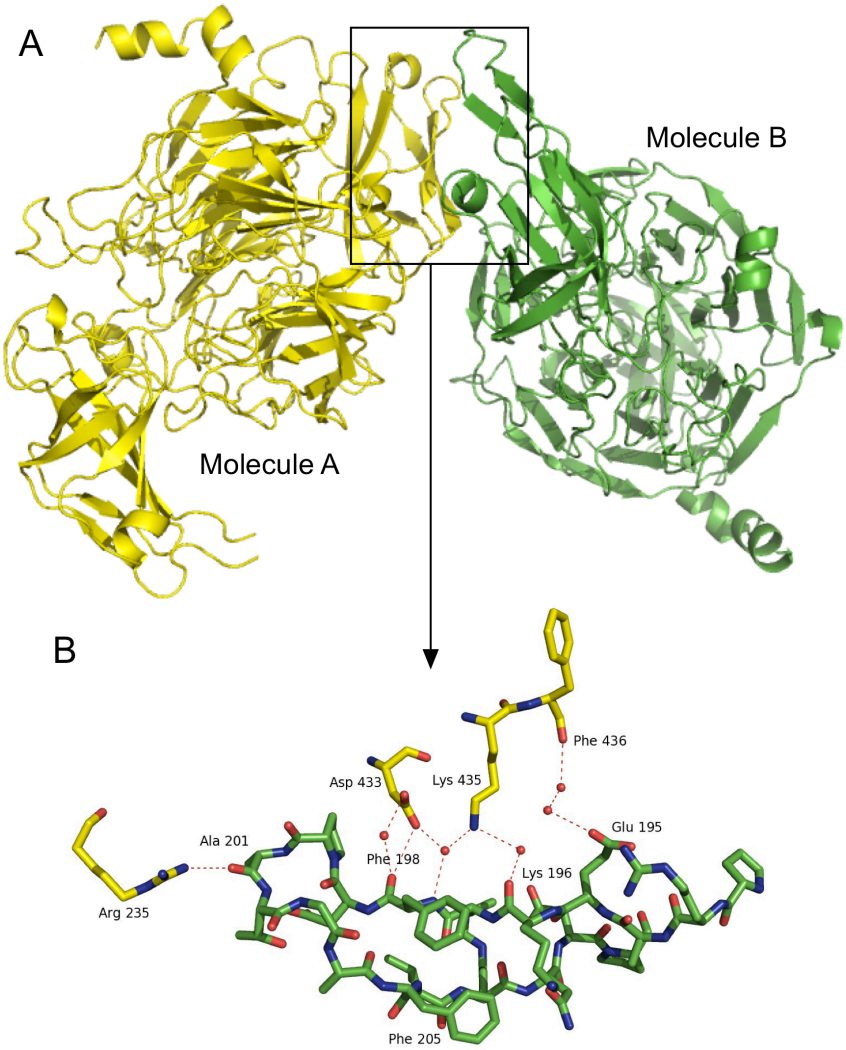


Figure 2.10: X-ray crystal packing of *S. lividans* GlxA. A) The arrangement of the two molecules of GlxA present in the crystallographic asymmetric unit. Molecule A is shown in yellow and molecule B in green. Molecule A does not display density for the β-hairpin loop (disordered region between residues 198 and 206), which is observed in molecule B and is therefore omitted from the structure. B) Stabilizing crystal-packing interactions of the β-hairpin loop in molecule B of the GlxA structure. Interactions with adjacent molecules are largely solvent-mediated with a small number of direct interactions between amino acids. Note that an additional linkage between the amide nitrogen of Phe^{205} and Lys^{323} of an adjacent molecule, via bridging water 461 is omitted from the figure for clarity. Images were made in PyMol.

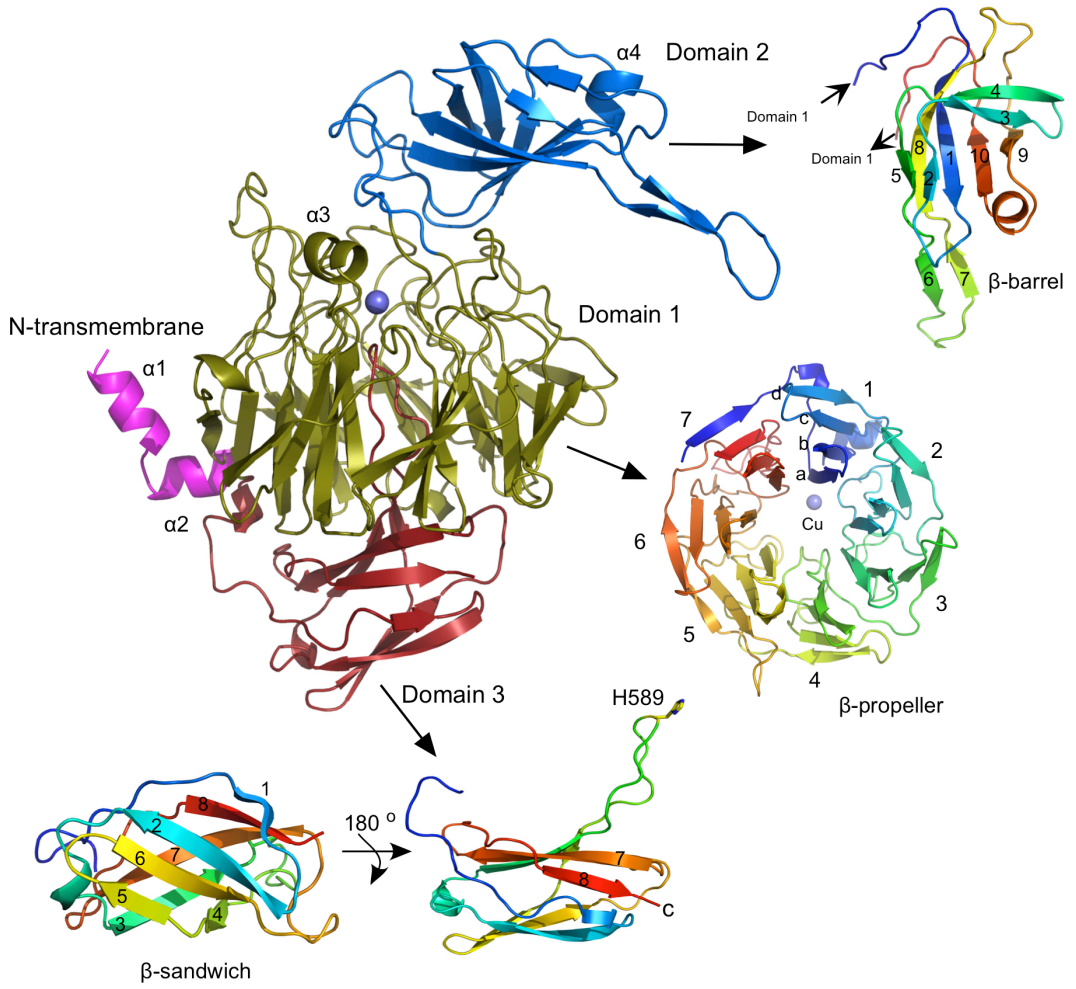


Figure 2.11: X-ray crystal structure of *S. lividans* GlxA. Cartoon representation of Cu-GlxA with the three domains forming the tertiary structure displayed and coloured gold (domain 1), blue (domain 2) and red (domain 3). The 4 short α-helices present in the structure are labelled with the N-terminal α1 and α2 coloured pink. Each domain is shown individually and in more detail. Domain 1 is labelled as a β-propeller domain with the 7 blades numbered and β-strands indicated (a-d). Domain 2 forms a β-barrel fold and β-strands are labelled. Domain 3 creates a β-sandwich fold with one of coordinating Cu residues (His^{589}) labelled and β-strands numbered. The Cu atom is represented as a purple sphere. Images were prepared in PyMol and CCP4MG [121].

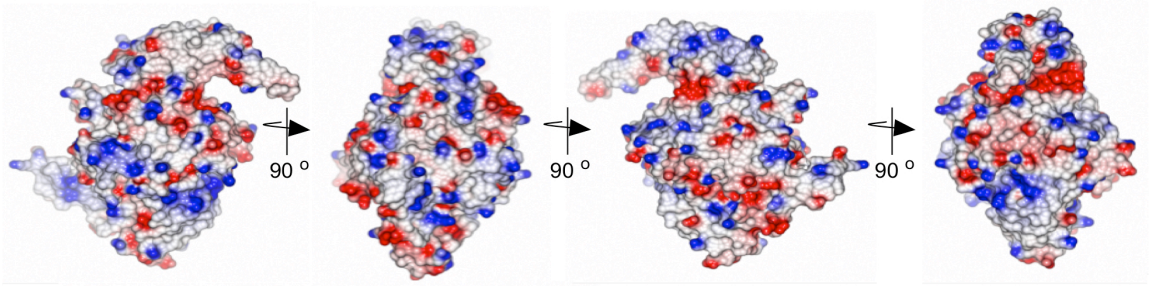


Figure 2.12: Electrostatic surface representations of *S. lividans* GlxA (molecule B) in 4 orientations, with 90 ° rotations displayed. Positive and negative charges indicated in blue and red, respectively. Images made in CCP4MG [121].

2.3.9. The Cu site of GlxA

A Cu ion could be modelled into a clear 16 σ peak in the σ-weighted F_o-F_c difference map displayed in the electron density of both GlxA molecules within the asymmetric unit. The Cu ion lies off-centre towards blade 1 of domain 1 coordinated with square pyramidal geometry with the Oγ of Tyr^{501} acting as the axial ligand and the equatorial coordination positions occupied by the Oγ of Tyr^{289}, the Nε2 of His^{502} and His^{589} and a H_2O molecule (Figure 2.13). Gox consists of identical Cu coordinating ligands (Tyr^{272} Tyr^{495} His^{496} His^{581}) with a coordinating acetate rather than a water molecule. A comparison of bond lengths from the ligands to the Cu ion for GlxA and Gox are reported in Table 2.5. GlxA also shows unbroken electron density between the side chain Sγ atom of Cys^{121} and the Cε1 ring atom of Tyr^{289} giving evidence of a post-translational modification, consisting of a Cys-Tyr crosslink with a bond length of 1.9 Å (2.0 Å in chain B) (Table 2.5 and Figure 2.13). Furthermore, the *cis* orientation of the Cys^{121} side chain and the orientation of the Cβ atom lying in the same plane as the phenoxyl ring of Tyr^{289}, is seen in Gox (Tyr^{272}-Cys^{228}) [76], and thus indicates a stereo-chemical and structural conservation of this Cys-Tyr crosslink between species. Adjacent to Tyr^{289} is Trp^{288} in GlxA, which creates an important second coordination sphere residue. The side chain of Trp^{288} has its indole ring orientated so that the benzene ring is π-π stacking with the phenoxyl ring of Tyr^{289} with distances between atoms in the rings between 3.5 and 4 Å (Figure 2.13 and 2.3 shown by + symbol). This second coordination sphere residue is also present in Gox (Trp^{290}) with an altered orientation and has been reported for Gox to play a critical role in enzyme catalysis, substrate binding and radical properties [122, 123].

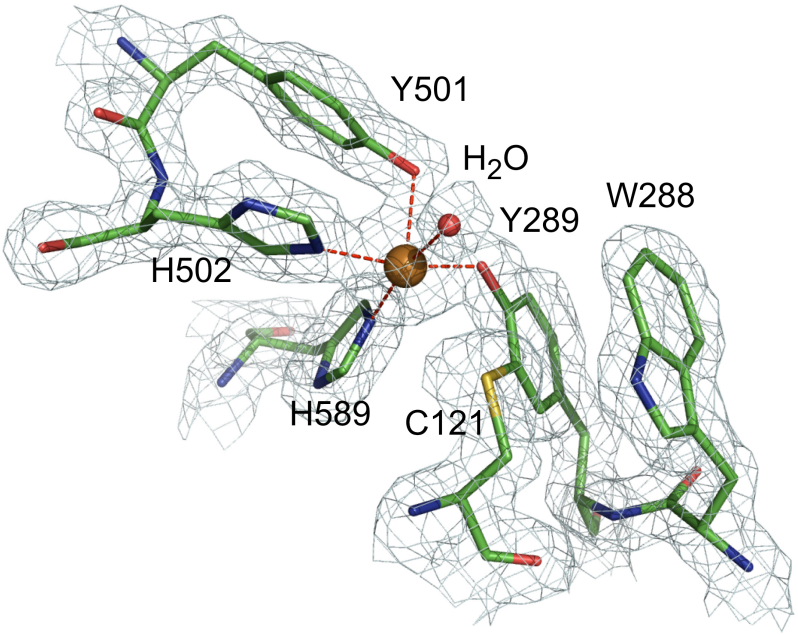


Figure 2.13: 2 F_o-F_c electron-density map contoured at 2 σ of the Cu site in *S. lividans* GlxA. The Cu ion is represented as a brown sphere with the ligands to the Cu ion depicted as sticks and the equatorially coordinating H_2O molecule as a red sphere. All coordinating residues and the water molecule labelled. Image made in PyMol.

Table 2.5: Bond lengths of the Cu sites (monomer A and B) in *S. livdians* GlxA and *F. graminearum* Gox (PDB 1GOF) [76].

GlxA (A/B)	Distance (Å)(A/B)	Gox	Distance (Å)
Cu-Tyr^{289} OH	1.84/1.97	Cu-Tyr^{272} OH	1.93
Cu-Tyr^{501} OH	2.24/2.14	Cu-Tyr^{495} OH	2.69
Cu-His^{502} Nε2	2.13/2.18	Cu-His^{496} Nε2	2.11
Cu-His^{589} Nε2	2.18/2.23	Cu-His^{581} Nε2	2.14
Cu-H_2O	2.44/2.50	*Cu-acetate	*2.26
Tyr^{289}-Cys^{121}	1.92/1.97	Cys-Tyr	1.83

*1GOF crystallised in acetate buffer pH 4.5 (GlxA acetate buffer pH 4), with an acetate molecule found in place of the H_2O molecule in GlxA. In a Gox structure (1GOG) [76] crystallised in the absence of acetate buffer at pH 7 a H_2O molecule is found with a bond length to the Cu ion of 2.81 Å.

2.3.10. Putative substrate entry sites and substrate binding pocket in GlxA

The Cu active site of GlxA is significantly buried within the protein by loops from domain 2 and therefore entry of a substrate is likely to be via a tunnel. CAVER is a computational program used to identify tunnels within proteins by defining them as void pathways [106, 124]. In GlxA the tunnels were measured from the Cu site (cavity inside) to the protein surface into the solvent. CAVER calculated the presence of 3 main tunnels in GlxA, each with different lengths (Figure 2.14.A). Tunnel I is the shortest of the 3 with a distance to the surface of ~ 10 Å and an opening of ~ 8 x 8 Å. No H_2O molecules are present in this tunnel (Figure 2.14.B). The entrance lies between the α3 helix on the loop connecting sheets *a* and *b* in blade 1 and a well-ordered loop in blade 7 of domain 1. The opening is large enough to allow pyranose carbohydrates such as D-galactose and also C2 substituted pyranose carbohydrates such as N-acetyl-glucosamine to enter. Tunnel II is the next longest with a distance to the surface of ~ 25 Å comprising a network of 6 hydrogen-bonded (H-bonded) H_2O molecules. The opening to the tunnel is just below the β-hairpin loop of domain 2 (Figure 2.14), with an opening, of ~ 9 x 7 Å also large enough for pyranose carbohydrates to pass (Figure 2.14.B). Tunnel III spans the largest area with a distance from the Cu ion to the surface of ~ 70 Å and lined with a continuous network of well-ordered H_2O molecules (13 in total). Tunnel III covers a large area of domain 1 beginning below the equatorial His ligand and leading to a surface entrance formed by residues Ala^{128}, Gly^{298}, Pro^{322} and Ser^{622}.

Above the Cu coordinating Tyr^{501} a pocket can be identified and is predicted to be a putative substrate-binding site. The pocket is accessible by either Tunnel I or II and is occupied with several well ordered H_2O molecules, all forming an extensive H-bond network, including an interaction with the Tyr^{501} ligand (Figure 2.14.C). The pocket is formed from 6 amino acid side chains (Tyr^{501}, Tyr^{289}, Tyr^{359}, Lys^{253}, Tyr^{448}, Tyr^{468}) creating a completely polar environment with some residues H-bonding with the H_2O network (Figure 2.14.C). Tunnel I residues Asn^{88} and Arg^{469} are also capable of H-bond interactions. This pocket is ideal for accommodating a pyranose carbohydrate substrate for catalytic turnover by GlxA.

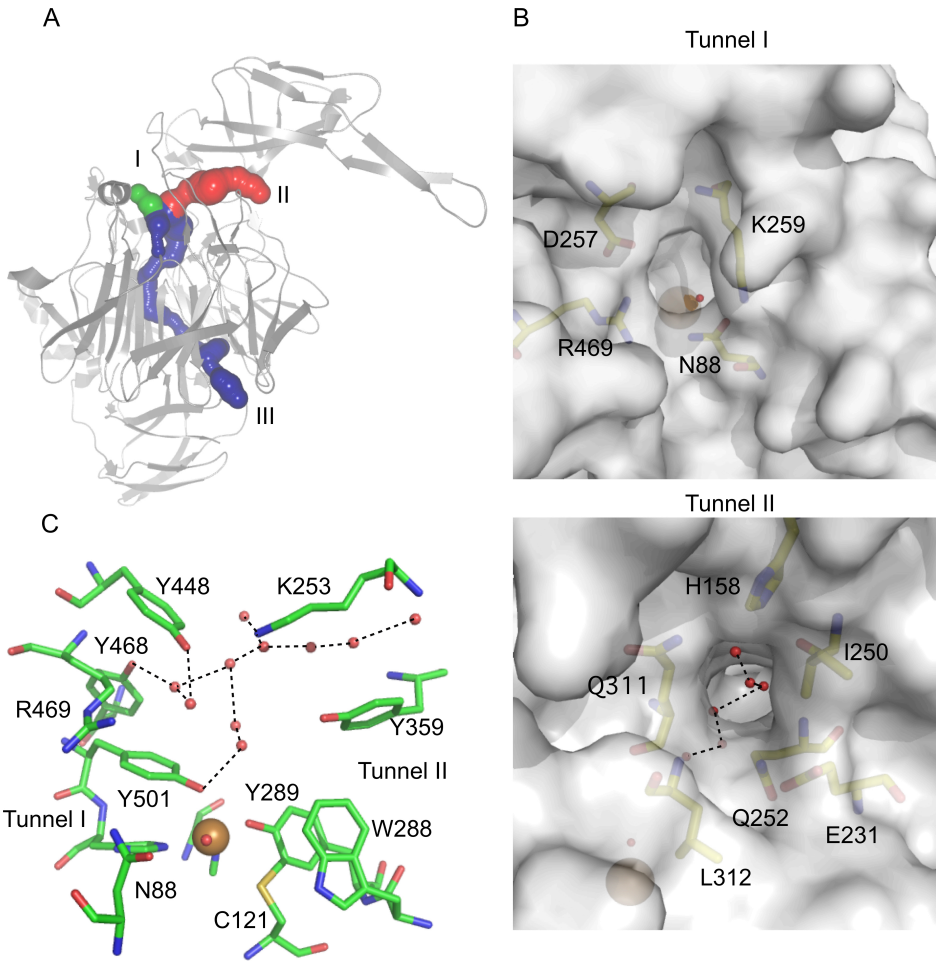


Figure 2.14: Substrate access channels and the binding pocket in *S. lividans* GlxA. A) The position of 3 surface-to-Cu site tunnels in GlxA identified, tunnel I in green, tunnel II in red and tunnel III in blue identified using the programme CAVER [106]. B) Partial transparent surface views of the openings to tunnels I and II. The amino acids forming the openings are labelled and shown as sticks, H_2O molecules are depicted as red spheres and the Cu ion as a brown sphere. In tunnel I the equatorially Cu coordinated H_2O molecule is visible, and tunnel II illustrates the H-bonded H_2O network leading from the surface to the Cu site. C) The putative substrate-binding pocket in GlxA with Cu ion present as a brown sphere, water molecules visible as red spheres and residues labelled and shown in green.

2.3.11. A Cys-Tyr triplet radical is present in GlxA

For GlxA to perform a two-electron oxidation of a substrate the Cys-Tyr crosslink must be capable of harbouring a radical species, as is the case in Gox. Addition of the oxidants, $[Fe(CN)_6]^{3-}$ or $[Ir(Cl)_6]^{3-}$, to Cu(II)-GlxA does not affect the UV-visible absorbance spectra (Figure 2.15) however, a change in the EPR spectrum is observed after addition of $[Fe(CN)_6]^{3-}$ (Figure 2.15.B.C.E.F). A 3 component free radical EPR spectrum is detected which is comparable to that seen for Gox [82], glyoxal oxidase [125] and the GlxA from *S. coelicolor* [111] (Figure 2.15.G). This radical is distinctive and reveals the presence of the Cys-Tyr• free radical solely existing to this class of radical Cu oxidases. A change in EPR spectrum is also seen following addition of $Na_2S_2O_4$, reducing Cu(II) to Cu(I), removing any detectable EPR signal (Figure 2.15.D). These observations signify the ability of GlxA to access three oxidation states (reduced (Cu(I)), semi-reduced (Cu(II)) and fully oxidised (Cu(II)-Cys-Tyr•)) as is the case for Gox.

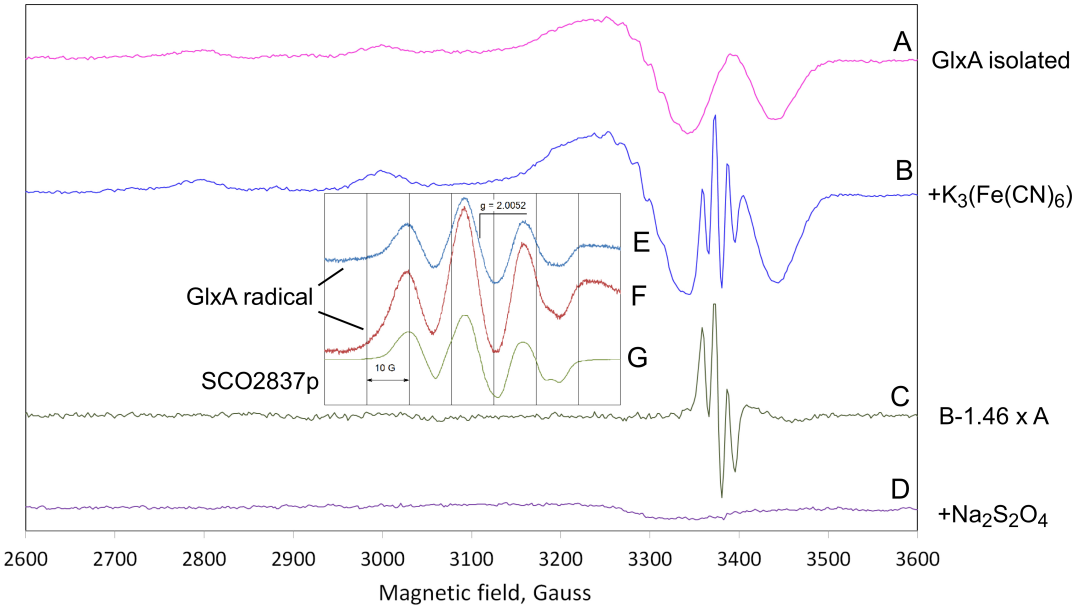


Figure 2.15: EPR spectra of Cys-Tyr radicals present in *S. lividans* GlxA and GlxA from *S. coelicolor*. A) GlxA from *S. lividans* as isolated, B) after 50-fold excess $[Fe(CN)_6]^{3-}$ treatment, C) difference spectrum (B-1.46 x A) showing the free radical EPR signal in its pure form (the coefficient 1.46 was found empirically to minimise input of other EPR signals to the difference spectrum), D) Isolated GlxA after treatment with $Na_2S_2O_4$. *Inset* - the same sample as the one in B was used to measure the free radical EPR signal in greater detail using two different microwave power levels, 50 mW (E) and 3.18 mW (F). G) The EPR spectrum of the radical GlxA from *S. coelicolor* [111]. All spectra were recorded at 40 K using 120 μM of GlxA in a mixed buffer system consisting of 10 mM each of Tris, potassium acetate, MES, MOPS and 200 mM KCl, pH 7. The instrumental conditions for A, B and D, were the same as Figure 2.8. Whereas for spectra E and F the conditions consisted of a modulation amplitude A_M = 3 G, scan rate V = 0.596 G s^{-1}. Spectrum G is a digitised image from [111], aligned with the GlxA spectrum on the basis of g-factors, therefore the magnetic field axis is not indicated.

2.3.12. GlxA enzymatic activity and its significance

The identification of a Cys-Tyr crosslink in the X-ray structure of GlxA and the confirmation from EPR spectroscopy that a radical is formed (Figure 2.15) (See section 2.3.11) indicates that GlxA is most likely capable of carrying out a two-electron oxidation of substrates. A number of putative substrates were measured for their activity with GlxA using a coupled peroxidase assay followed by the oxidation of guaiacol (Figure 2.16.A) (see Experimental section 2.2.10). Only four substrates showed any significant activity, glycolaldehyde, D-galactose, D-glucose and glycerol as indicated from the absorbance versus time courses in Figure 2.16, and initial rates calculated. For glycolaldehyde ($C_2H_4O_2$) the smallest molecule to contain both an alcohol and aldehyde the $A_{470\,nm}$ increases as is the case for the other substrates but then decays, possibly a result of glyoxal (product of glycolaldehyde oxidation) reacting with the oxidation product of guaiacol (tetra-guaiacol) (Figure 2.16.B). For D-galactose, D-glucose and glycerol the initial rates gave a linear substrate concentration relationship Figure 2.17.A. From the slopes of these plots, second order rate constants were determined for D-galactose; 8.4 x 10^{-3} $M^{-1}s^{-1}$, D-glucose; 8.8 x 10^{-3} $M^{-1}s^{-1}$ and glycerol; 1.7 x 10^{-2} $M^{-1}s^{-1}$. Glycolaldehyde gave the highest activity and turnover rates following Michaelis-Menten kinetics (Figure 2.17.B), enabling a K_m of 115 mM and a maximal aerobic turnover number (k_{cat}) of 0.14 s^{-1} to be determined. A V_{max}/K_m of 1.22 $M^{-1}s^{-1}$ was calculated for glycolaldehyde, which is 100-times faster when directly compared with the two monosaccharides and glycerol. No activity was detected for the C2 modified position monosaccharide N-acetyl-D-glucosamine or for fructose, a five membered ring monosaccharide. D-glucuronic acid also gave no activity showing that a C1 glycosidic group is not effective as a reductant when the C6 is modified (COOH compared with CH_2OH). No activity was also noted for disaccharides D-lactose and D-cellobiose. Also no activity was detected for the aldehyde glyoxal the substrate of glyoxal oxidase, which contrasts with Gox also reported to convert aldehydes to their resultant carboxylates. Therefore, based on the highest activity being with the simple alcohol glycolaldehyde GlxA could be considered a putative alcohol oxidase.

Another possibility for the physiological function of GlxA could be in the production of H_2O_2. This could be significant for development rather than the substrate product produced by GlxA. The effect of H_2O_2 on the development in WT and Δ*glxA* mutant in *S. lividans* was monitored by addition of sub-lethal concentrations of H_2O_2 to young cultures (Figure 2.18). It can be seen that addition of H_2O_2 did not stimulate development in either WT or Δ*glxA* mutant and infers the importance of the substrate product and not H_2O_2 production in development (Figure 2.18).

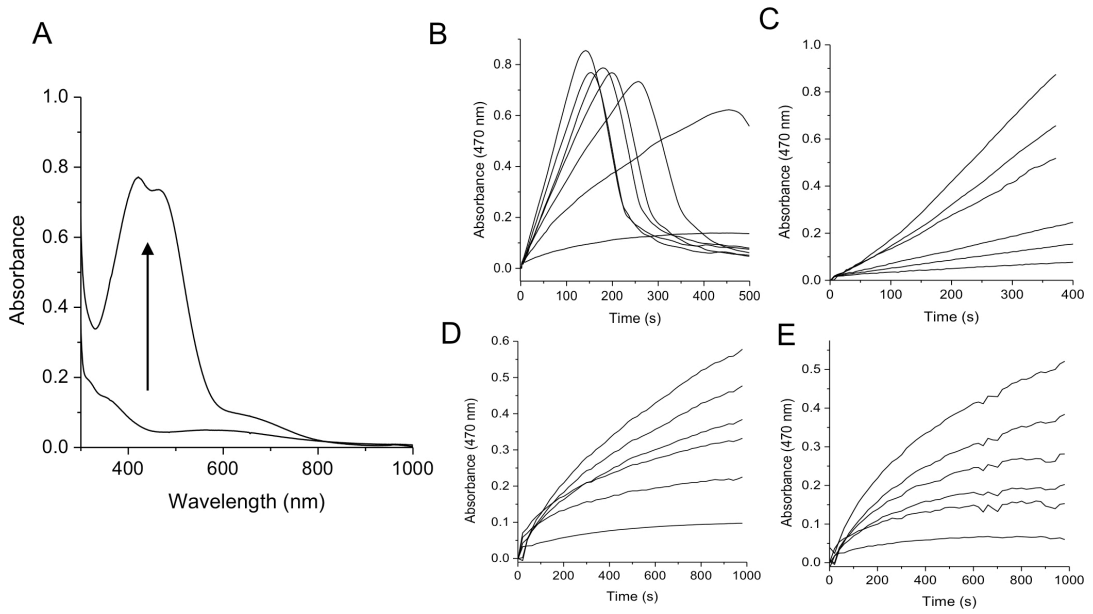


Figure 2.16: Monitoring guaiacol oxidation of H_2O_2 produced by *S. lividans* GlxA. A) An example of the visible region of the absorption spectrum of guaiacol oxidation during enzymatic turnover of glycolaldehyde by GlxA. The arrow indicates the directionality of the absorbance increase at 470 nm as GlxA produces H_2O_2 in the reaction. B-E) Catalytic turnover of GlxA showing absorbance versus time traces at 470 nm. B) With glycolaldehyde concentrations of, 185, 155, 125, 95, 65, 35, 5 mM with 20 μM GlxA. C) With glycerol D) with D-galactose. E) With D-glucose all with concentrations of, 100, 200, 300, 400, 500 and, 600 mM with 30 μM GlxA for all three substrates. Assays were all carried out in 0.1 M NaPi, pH 7.4 at 25 ^{o}C.

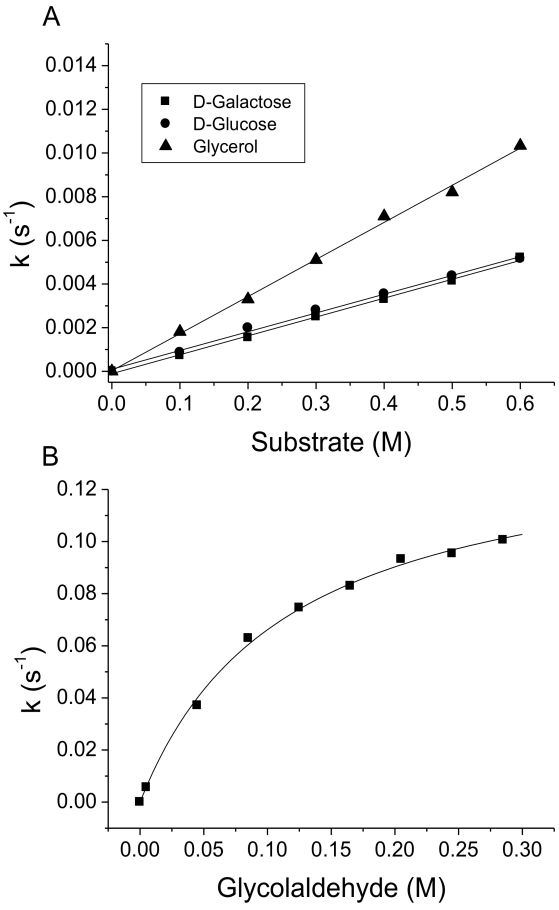


Figure 2.17: Enzyme activity of *S. lividans* GlxA. A and B) Plots of turnover rate constants (k) of GlxA with four different substrates (25 ^{o}C). A) Enzyme activity of GlxA with glycerol, D-galactose and D-glucose with a line of best fit through the data points enabling a second-order rate constant ($M^{-1}s^{-1}$) to be determined. B) Enzyme activity of GlxA with glycolaldehyde as a substrate with the data points fitted to the Michaelis-Menten equation (Equation 2.1) to yield a K_m value and turnover rate (k_{cat}). Assays were all carried out in 0.1 M NaPi, pH 7.4.

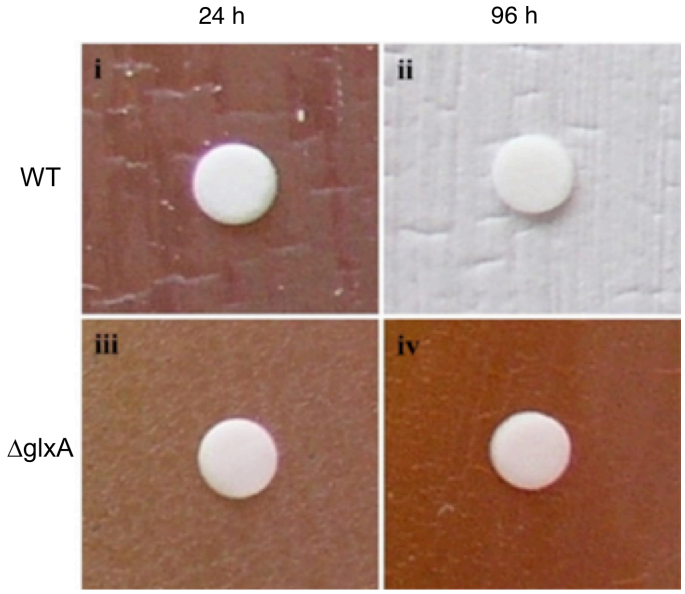


Figure 2.18: The effect of H_2O_2 addition on the development in WT and Δ*glxA* mutant in *S. lividans*. WT and Δ*glxA* mutant spores were inoculated onto R5 agar plates and incubated for 24 h at 30 ^oC producing confluent growth. (i and iii) WT and Δ*glxA* mutant 24 h and (ii and iv) WT and Δ*glxA* mutant 96 h after a paper disk containing 10 μl of 10 μM H_2O_2 were applied on top of the young mycelium and incubation continued. No stimulation of development was seen after 24 h for both WT and mutant (red/brown vegetative mycelium) and only the WT showed full uniform development (white aerial hyphae) after 96 h (ii) but the development was even and was therefore independent of where H_2O_2 was applied. The Δ*glxA* mutant did not produce aerial hyphae or spores at all (iv). Therefore H_2O_2 cannot induce development in either the WT or the *glxA* null-mutant.

2.3.13. Morphology of S. lividans 1326 WT, Δsco and ΔglxA mutants

Deletion of *glxA* (*SLI3188*) and the *sco* (*SLI4214)* genes create the Δ*glxA* and Δ*sco* null-strains enabling a comparison of the development profiles with the WT on solid R5 medium (Figure 2.19.A). For WT, morphological development can continue and aerial hyphae and spores can develop with the mycelium changing to white/grey in colour (vegetative mycelium gives a red/yellow colour arising from the production of the pigment undecylprodigiosin). It is apparent that for both Δ*sco* and Δ*glxA* development is stalled in the vegetative phase under low Cu concentrations creating a bald phenotype (red/yellow appearance). However, whereas development can be rescued upon addition of Cu to the Δ*sco* mutant, this is not the case for Δ*glxA* and development remains stalled in the vegetative mycelium state (Figure 2.19.A). This is consistent with previous work in *S. coelicolor* [72]. Complemented mutants Δ*sco* + pSco and Δ*glxA* + pGlxA display aerial mycelium and spores and therefore retain the morphology of the WT (Figure 2.19.A).

In liquid cultures, differences in Δ*sco* and Δ*glxA* compared to the WT can be observed. For WT, growth is characterised by the formation of compact mycelium pellets (Figure 2.19.B). Whereas, in the Δ*sco* mutant the morphology changes in the absence of

exogenous Cu(II) with mycelial clumps being replaced with an open, less dense, mycelium (Figure 2.19.B). By inoculating the Δ*sco* mutant with Cu(II) in the media, the mycelium returns to the compact pellet morphology observed for the WT (Figure 2.19.C). For Δ*glxA* this open mycelium morphology is observed both in the absence and presence of exogenous Cu(II) (Figure 2.19.B). These observations of solid and liquid culture growth provide initial support that Sco may be required as a Cu chaperone for GlxA and that importantly GlxA has a key functional role in the Cu-dependent morphological differentiation of *S. lividans*.

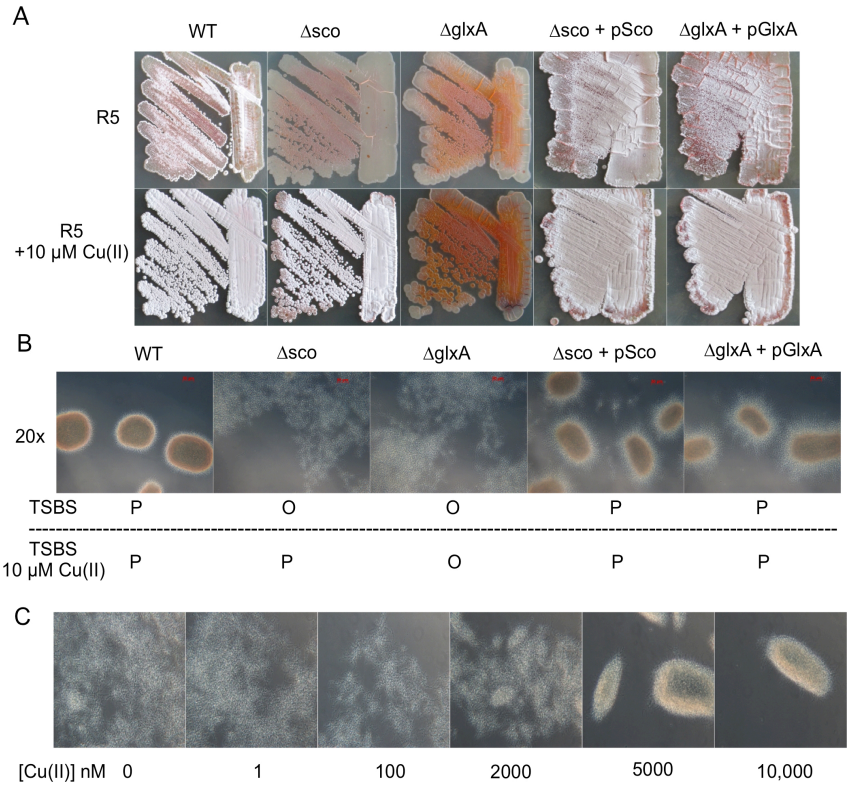


Figure 2.19: Morphology of *S. lividans* 1326 WT, Δ*sco* and Δ*glxA* mutants and complemented mutants (Δ*sco* + pSco, Δ*glxA* + pGlxA) mutants on solid and in liquid medium. A) Respective strains plated on R5 medium (top row) and R5 medium supplemented with 10 μM Cu(II) (bottom row). Vegetative mycelium gives a red/yellow colour arising from the production of the pigment undecylprodigiosin and if development continues to aerial hyphae and spores the mycelium changes colour to white/grey. WT and complemented mutants (Δ*sco* + pSco, Δ*glxA* + pGlxA) display aerial mycelium and spores as well as the Δ*sco* mutant with addition of Cu(II) in contrast to the Δ*glxA* mutant which shows an increase in red pigment production. B) TSB medium supplemented with 10 % sucrose (TSBS) inoculated with spores (10^6 mL^{-1}) of the respective strain. After 24 h, mycelium morphology was recorded by light microscopy, with WT producing pellets (P) while both mutants produce an open mycelium (O). In TSBS medium supplemented with 10 μM Cu(II) at inoculation, Δ*sco* reverts to WT morphology (P), with the Δ*glxA* mutant unchanged (O). The complemented mutants show morphology very similar to that of WT. C). The increase in Cu(II) concentration (0-10,000 nM) whereby the Δ*sco* mutant shifts from O to P. All microscope images are recorded with the same magnification.

2.3.14. GlxA is required for morphogenesis and glycan synthesis at hyphal tips

As described in section 2.3.13 GlxA is required for morphogenesis in *S. lividans* identified both on solid agar and liquid cultures. Comparison of *ΔglxA* to the null mutant of *cslA* (*SLI3188)* encoding a cellulose-synthase protein (Figure 2.20.A), identified that both show identical morphologies, with a bald phenotype on solid agar and disperse growth in liquid cultures as previously reported in *S. coelicolor* [65, 72]. The bald phenotype on solid agar for *ΔglxA* and *ΔcslA* cannot be rescued with addition of exogenous Cu(II) (Figure 2.20.A) as shown for *Δsco* in Figure 2.19.A. Both mutants also display an open mycelium in liquid cultures in comparison to the compact pellets seen in the WT (Figure 2.20.A). RNA-seq analysis of the transcriptomes of *glxA* and *cslA* mutants has revealed that genes involved in osmoprotection are up-regulated, indicating that these mutants suffer from osmotic stress as also revealed by Liman *et al* [72].

To identify the effects of *glxA* and *cslA* mutants on glycan synthesis and deposition at hyphal tips the mycelium of *S. lividans* were stained with CFW which binds to β-(1-4) glycans. In both the mutants the β-(1-4) glycans were no longer stained in contrast with those of the WT strain (Figure 2.20.B). It is now clearly visible that CslA and GlxA from *S. lividans* are required for glycan synthesis or attachment of it to the hyphal tips as has been implicated in *S. coelicolor* [65, 72].

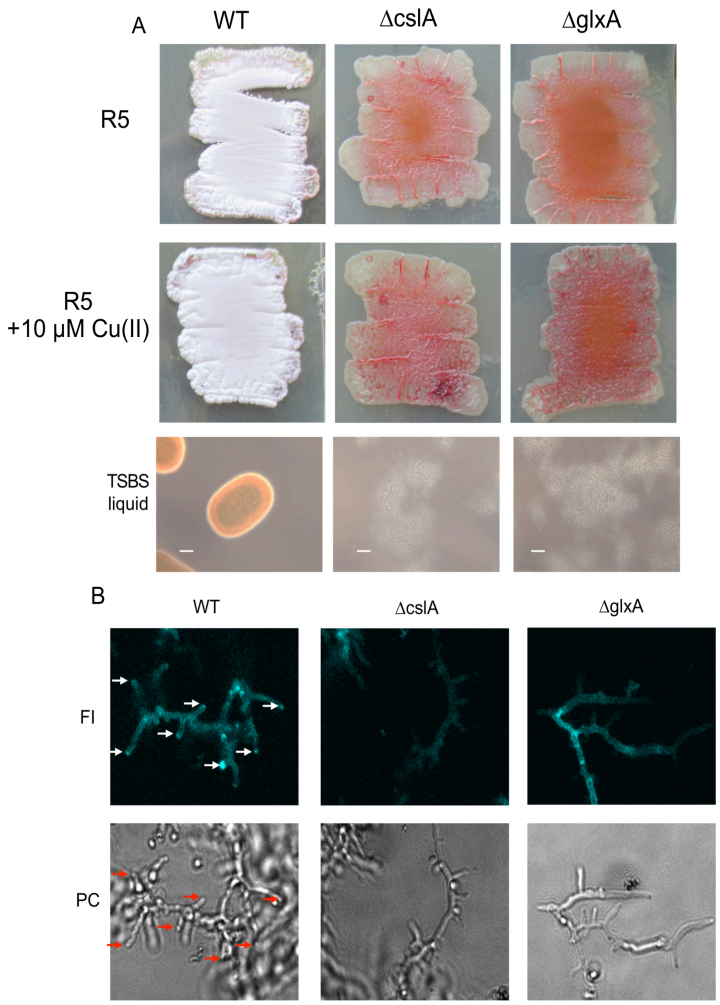


Figure 2.20: Deletion of *cslA* or *glxA* in *S. lividans* stalls morphological development arrests pellet formation and prevents glycan deposition. A) Growth of WT, *ΔcslA* and *ΔglxA* on solid R5 medium and pellet formation or dispersed growth in TSBS liquid cultures after 24 h. Scale-bar (white line) is 100 µm. B) Calcofluor white staining in 24 h mNMMP standing cultures grown at 30 ^{o}C shown in fluorescence image (FI) and phase contrast brightfield (PC). Arrows indicate tip staining that is present in WT cultures but not in *ΔcslA* and *ΔglxA* mutants. Analyzed with a laser-scanning confocal microscopy (Zeiss LSM5 Exciter/Axio Observer), excitation with a 405 nm laser, a 405/488 nm beamsplitter and a bandpass emission filter of 420-480 nm. Images were adjusted for brightness and contrast using ImageJ [89].

2.4. Discussion

2.4.1. Spectroscopic properties of GlxA are distinctive

Spectroscopic properties of the Cu site in GlxA including, UV-visible absorbance, visible CD and EPR, reveal significant differences to the fungal homologue Gox. The distinct number of differences is remarkable considering the first coordination sphere Cu ligands are identical. EPR spectroscopy of Cu(II)-GlxA and Cu(II)-Gox both display the presence of axial Cu(II) g-tensors, however, differences arise in the $g_{\perp}$ component with hyperfine interactions arising from the electron spin (S = ½) with the Cu(II) ligands. In the Cu(II)-Gox EPR spectrum the $g_{\perp}$ component gives rise to a quintet hyperfine splitting with relative intensities 1:2:3:2:1 and distances consistent with two His ligands (14-18 G) (I = 1) corresponding to His^{496} and His^{581}. However, the Cu(II)-GlxA spectrum displays seven lines with relative intensities 1:4:8:10:8:4:1 with a 16 G splitting (Figure 2.8) [116, 117]. It is difficult to explain the direct cause of the additional hyperfine splitting but it is apparent that differences in the electronic properties of the Cu sites of GlxA and Gox must exist. The hyperfine splitting is best seen at pH 7 for GlxA and disappears at higher pH values, likely to be caused by a deprotonation event on a surrounding residue(s) perturbing the geometry and electronic transitions of the Cu site.

Visible CD of GlxA also illustrates significant differences to Gox. GlxA displays no alteration in the spectrum after addition of $[Fe(CN)_6]^{3-}$, which differs to Gox as it displays significant changes with different oxidation states of the enzyme. GlxA displays spectrum with minima at 350 nm and maxima at ~ 320 and 650 nm, this compares to active Gox with minima at approximately, 560, 440 and 350 nm and maxima at 650, 500, 380, and 320 nm (Figure 2.7.C) [81]. Some of these spectral transitions coincide with GlxA, (maxima at 320 nm and 650 nm and minima at 350 nm) however, many of the peaks identified in Gox are absent from GlxA thus again indicating alterations in the electronic properties of the Cu sites.

In agreement with EPR and visible CD spectroscopy the UV-visible absorbance spectrum of GlxA differs significantly to that reported for Gox and also for the homologous GlxA from *S. coelicolor*. As isolated *S. lividans* GlxA displays a distinct spectrum with a low intensity broad peak at λ_{max} ~ 577 nm and two distinct shoulders at ~ 362 and 320 nm (Figure 2.6), differing to native Gox with peaks at λ_{max} 441 nm and λ_{max} 630 nm [81, 83]. Nevertheless, what is more noticeable is the lack of perturbation in the GlxA spectrum. Changes can only be observed upon addition of the reductant, $Na_2S_2O_4$, which bleaches all spectral features in the visible region owing to the reduction of Cu(II) to Cu(I) (Figure 2.6.B), also confirmed by EPR (Figure 2.15.D). Whereas, addition of excess Cu(II) or oxidants display no significant affect on the optical spectrum (Figure 2.6.B). In contrast, Gox and GlxA from *S. coelicolor* produce an intense green colour with peaks at 445 and 810 nm when fully oxidised [81, 83, 111]. *S. coelicolor* GlxA contains only 8 amino acid differences to GlxA but the absorbance spectrum is indistinguishable from Gox (Figure 2.21.A) [111]. It is

unlikely that the spectral differences can be accounted for by the change in these residues, as they are a significant distance away from the Cu ion with the closest residue, Asn^{388}, ~ 22 Å away (Figure 2.21.B). *S. coelicolor* GlxA was over expressed in *Pichia pastoris* thus differing to the expression of *S. lividans* GlxA in *E. coli* [111]. The difference in expression systems and the change in residues could account for an alteration in the folding or variation in spatial positions of residues surrounding the Cu active site, resulting in modifications in the electronic transitions, thus affecting the spectroscopy. Alternatively, the spectroscopic characterisation by Whittaker and Whittaker arises from their preparations from *P. pastoris* having some background Gox present.

For Gox, clear absorbance spectral changes exist between the Cu(II) and the Cu(II)-Cys-Tyr• oxidation states (*vide supra*). A second coordination sphere residue, Trp^{290}, has been reported to be important for radical stability and in substrate binding and recognition in Gox. Trp^{290} has been proposed to provide an extended aromatic system with its indole ring stacking over the Cys-Tyr bond, stabilising the delocalisation of the radical (Figure 2.22.B) [76, 123]. In GlxA the equivalent residue, Trp^{288}, no longer has its indole ring stacking with the Cys-Tyr cross-link, but instead has the benzene ring component π - π stacking with the Tyr^{289} Cu ligand (Figure 2.22.A). The Trp^{290} residue in Gox has been shown to tune the redox potential of the Cu(II)/Cu(II)-Cys-Tyr• couple [126], but substitutions to Phe, His or Gly have surprisingly little effect on the absorbance spectra of either the Cu(II) or Cu(II)-Cys-Tyr• forms [123]. Nevertheless it could be possible that the unique nature of the Trp residue with the π-π ring of Trp^{288} stacking with Tyr^{289} influences the electronic transitions of the Cu and the radical resulting in the absence of 'classical' Gox features in the absorbance, EPR and CD spectrum of GlxA.

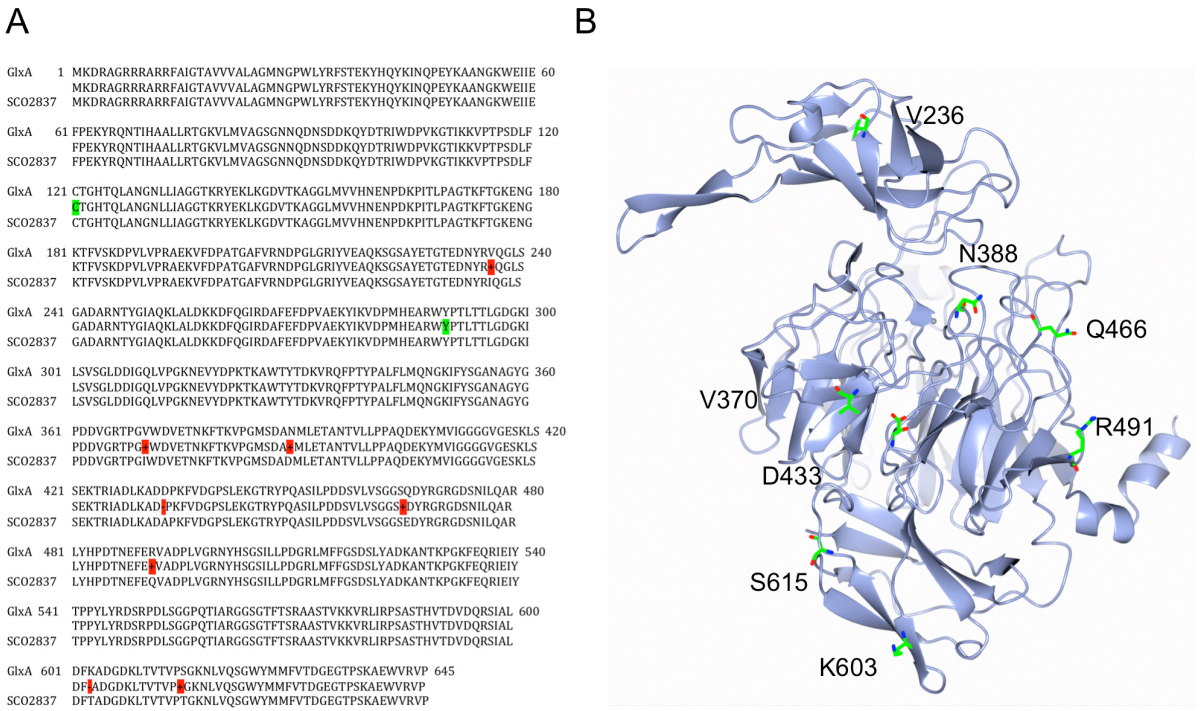


Figure 2.21: Comparison of GlxA from *S. coelicolor*. A) Amino acid sequence alignment of GlxA from *S. lividans* with *S. coelicolor* GlxA (SCO2837), with Cys[121] and Tyr[289] forming the Cys-Tyr crosslink shown in green and the amino acid changes between the two proteins shown in red. B) Structure of *S. lividans* GlxA chain B with the amino acids differing to *S. coelicolor* GlxA shown as sticks and labelled. Image made in CCP4MG [121].

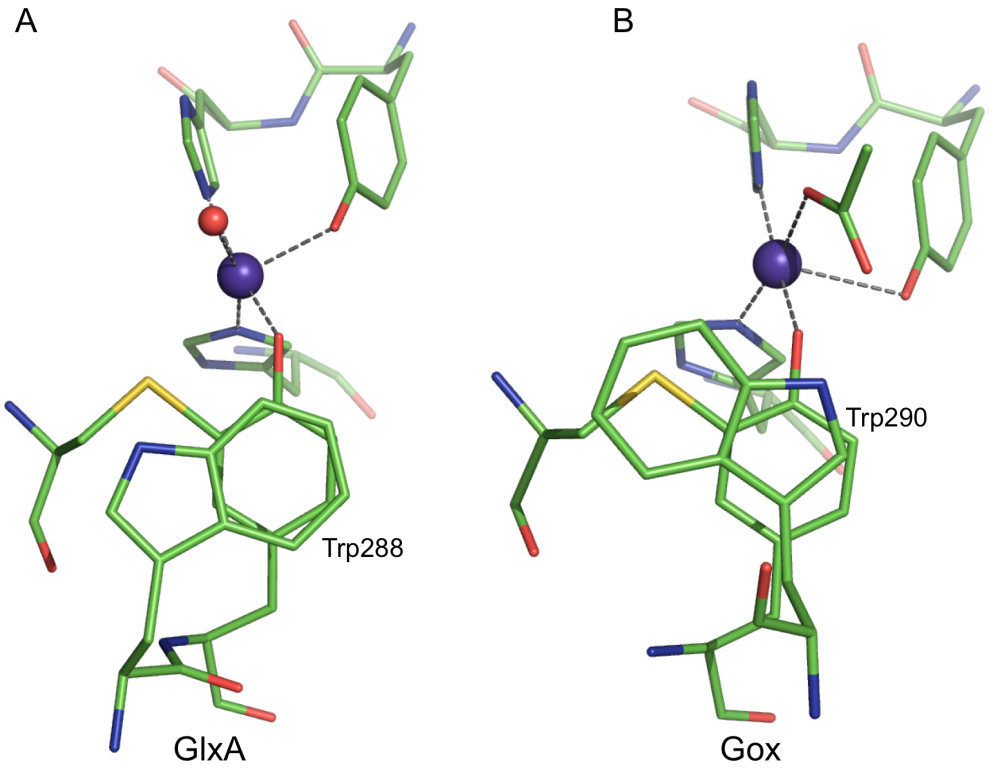


Figure 2.22: Comparison of the orientation of the stacking tryptophan residue in *S. lividans* GlxA and *F. graminarium* Gox (1GOF) [76]. A) An above view of the active site of GlxA showing the Cu ligands and the stacking tryptophan Trp^{288}. B) An above view of the active site of Gox with all Cu ligands shown and the stacking tryptophan Trp^{290}. The Cu ion is shown as a purple sphere. Images were made in PyMol.

Although the addition of oxidants has no affect on the UV-visible spectrum of GlxA, following $[Fe(CN)_6]^{3-}$ addition, the EPR spectrum reveals a three component radical signal assigned to the Cys-Tyr cofactor (Figure 2.15). This radical is almost identical to that identified in *S. coelicolor* GlxA and Gox (Figure 2.15) [111], illustrating that this radical is still present in GlxA even if UV-visible spectroscopy does not reveal it. However, in Gox and *S. coelicolor* GlxA the Cu(II) EPR signal disappears upon treatment with an oxidant, however, this is not the case in GlxA with both the Cu(II) and radical signal both present. This must be due to either a change in the Cu site so that the resulting locations of the Cu(II) ion and the radical are such that they result in a loss of coupling in the EPR spectrum. A less likely alternative, could be that GlxA consists of a mixture of species with approximately half of the population bound to Cu(II) and half lacking the metal ion, but producing a radical species.

2.4.2. The biologically relevant substrate of GlxA is unknown

GlxA has the ability to generate three oxidation states including, the reduced Cu(I) state, the semi-reduced Cu(II) form and the fully oxidised Cu(II) Cys-Tyr radical shown through EPR (Figure 2.15), therefore, suggesting an enzymatic function in *S. lividans*. However, GlxA displayed little activity with the substrates tested displaying nothing comparable to Gox (D-glucose was an exception). D-galactose is the recognised monosaccharide substrate for Gox with a k_{cat}/K_m of 44.6 $M^{-1}s^{-1}$ [122], whereas, under aerobic conditions a k_{cat}/K_m of 10^{-3} $M^{-1}s^{-1}$ for GlxA was determined significantly lower and more comparable to the activity of Gox with D-glucose (Figure 2.17.A) [127]. The substrate which gave the highest turnover for GlxA was glycolaldehyde displaying a K_m value (115 mM) comparable to Gox for D-galactose (70-80 mM), but a k_{cat}/K_m value > 4-orders of magnitude lower (Figure 2.17.B). Glycolaldehyde turnover with Gox has been reported to be 75 % of the D-galactose rate [128, 129]. These substantial differences in enzymatic turnover between GlxA and Gox must be accounted for by the distinct structural modifications. Although the activity identified in GlxA is distinct to that reported for Gox it is not dissimilar to that reported for GlxA from *S. coelicolor*. Simple alcohols supported the highest turnover for *S. coelicolor* GlxA with glycolaldehyde displaying the highest with a K_m value of 200 mM nearly identical GlxA from *S. lividans* (K_m 115 mM) [111]. The rates for *S. coelicolor* GlxA were reported as percentages relative to the activity of glycolaldehyde, (set at 100 %) and thus a direct comparison of rates cannot be fully examined [111]. *S. coelicolor* GlxA also displayed some activity for glycerol and negligible activity for glyoxal as also identified for GlxA from *S. lividans* [111]. It is therefore apparent that although *S. coelicolor* GlxA displayed spectroscopic properties distinct from GlxA the enzymatic activity is comparable. Therefore, the structure of this new radical Cu oxidase must be essential in determining the nature of a specific biologically relevant substrate(s).

Furthermore, observations from the molecular genetics and CFW staining indicate that GlxA is directly involved in the production and localization of a hyphal tip glycan (Figure 2.20). This is consistent with the report of *S. coelicolor* GlxA being localised at the hyphal tip [65, 72]. CslA is a cellulose synthase-like enzyme encoded by the translationally coupled gene upstream of *glxA* and co-localised with GlxA at the hyphal tip. CslA has also been reported to be required for synthesis of the glycan in *S. lividans*. This glycan accumulates at apical sites during vegetative growth and is absent in *ΔcslA* [62, 65, 66] and also in the *ΔglxA* mutant (Figure 2.20.B). The absence of this glycan leads to a change in morphology observed on solid and liquid cultures of both *ΔcslA* and *ΔglxA* (Figure 2.20.A). Both mutants show identical morphologies, with a bald phenotype on solid agar and disperse growth in liquid cultures as previously reported in *S. coelicolor* (Figure 2.20.A) [65, 72]. The nature of this glycan has not yet been identified but it could be that GlxA modifies it through a two-electron oxidation process once synthesised by CslA. From the structure of *S. lividans* GlxA determined in this Chapter, it could be envisaged that the glycan produced by CslA feeds down the tunnels identified in the GlxA structure (Figure 2.14.A), become modified (oxidised), and then leave through another tunnel.

The open morphology displayed in liquid cultures of *ΔcslA*, *ΔglxA* and also *Δsco* (Figure 2.19 and 2.20) null-mutants could be beneficial for improving the use of *S. lividans* as an enzyme production host in biotechnology. It has been shown that a more fragmented and therefore also more open mycelium growth in *Streptomyces* increases enzyme production several fold [50, 58]. This is due to pelleted growth restricting the efficient transfer of nutrients and gasses to the centre and also by the rate at which they can be formed [56, 57]. The *Δsco* null-mutant shows a gradual increase in pellet formation with addition of Cu(II) therefore creating a way to control the growth morphology of *S. lividans* (Figure 2.19.C). The observation that *Δsco* can resume normal growth with addition of Cu(II) (Figure 2.19) whereas GlxA cannot links GlxA to the Cu dependency of *S. lividans* and infers the possibility that GlxA acquires Cu from Sco.

2.4.3. GlxA has a unique tertiary structure for a radical Cu oxidase

The unique spectroscopic and enzymatic properties of GlxA may also be influenced by the difference in the tertiary structure of GlxA compared to Gox, with a more buried Cu site clearly apparent. These structural differences can be observed in Figure 2.23. The domain arrangement in Gox is such that domains 1 and 2 contribute to a flat surface comprising a readily accessible Cu site and substrate-binding pocket (Figure 2.23.B and C). GlxA however, does not consist of a flat surface due to the presence of longer loops from domain 1 and the orientation of domain 2 creating an occlusion of the active site (Figure 2.23.A). Domain 2 is a comparatively novel structure and thus its presence in GlxA could expose a unique functional role for the protein. It may be envisaged that this domain can either act by

changing the substrate access channels creating more specific entry sites, or that it can move upon interacting with a protein partner, opening the protein up thus creating a structure more alike to Gox. First coordination sphere Cu residues for GlxA and Gox positions are almost identical (Figure 2.24) with an acetate ion present in Gox compared to a H_2O molecule in GlxA. Thus, this further proves the readily accessible Cu site in Gox compared to GlxA as both proteins were crystallised in conditions containing acetate but no acetate ion was found near the active site of GlxA (Figure 2.24). As mentioned above (2.4.1) a prominent difference between the Cu sites is the orientation of the Trp residue with Trp^{288} in GlxA having its indole ring flipped relative to Trp^{290} in Gox, making this orientation less favourable from a distance perspective for a substrate H-bond interaction (Figure 2.24). Additionally, the substrate binding pockets differ significantly between Gox and GlxA with only the Cu ligands conserved (Figure 2.24.B). Gox consists of an asymmetric polar/apolar substrate-binding pocket with residues Arg^{330}, Gln^{406} and the Nε1 atom of Trp^{290} providing H-bonds to the canonical D-galactose substrate, and the opposite side of the pocket formed by the aromatic residues, Phe^{194} and Phe^{464} (Figure 2.24.B) [76]. The pocket of GlxA differs significantly, with the exception of the Cu ligands no other residues are identical and an alteration in the spatial arrangement of these residues is observed (Figure 2.24). Residues in the pocket including Tyr^{359}, Tyr^{468} and Arg^{469}, could create a strong H-bond network with the substrate, with absence of the two Phe in Gox altering the properties of the site (Figure 2.24.B). These differences illustrate that despite the Cu coordination being similar, the substrate pockets of Gox and GlxA contain no conserved features. Therefore, the substrate of Gox (D-galactose) is unlikely to display the same high turnover kinetics with GlxA, as it is unlikely to be optimally accommodated or positioned in the pocket (section 2.4.2).

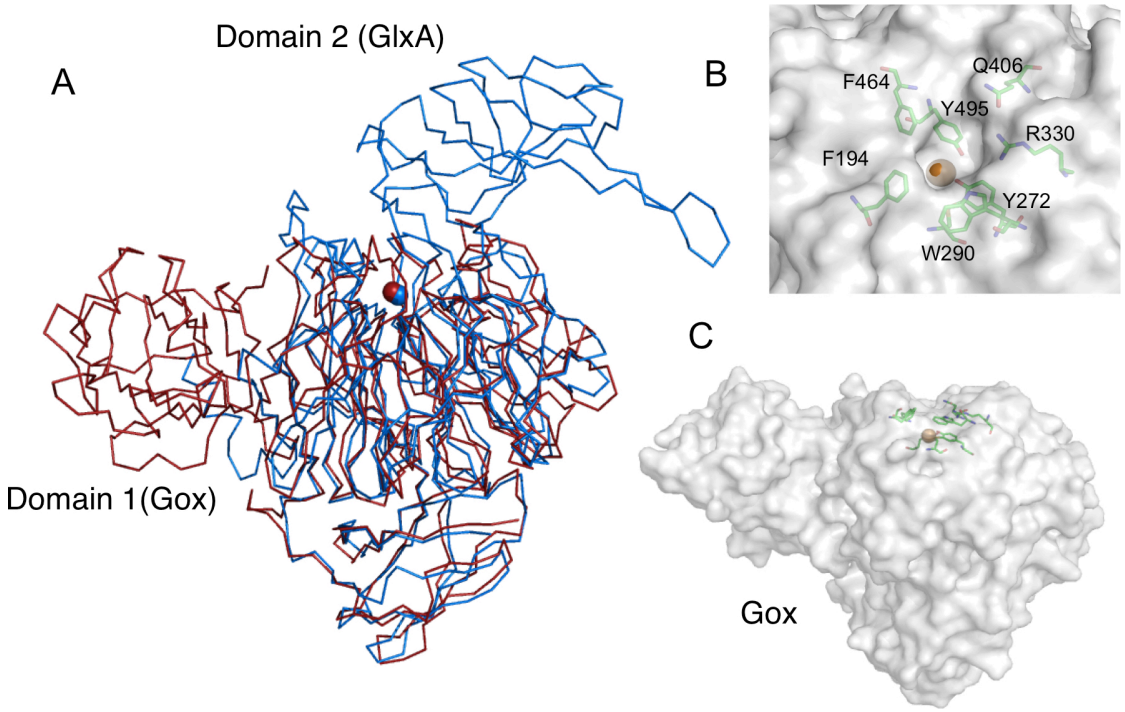


Figure 2.23: Comparison of the *S. lividans* GlxA and *F. graminarium* Gox (1GOF) structures [76]. A) An overall structural comparison of GlxA (blue) with Gox (red) shown by a ribbon representation with a superposition of the Cα atoms. The Cu atoms are shown as spheres (Red in Gox and blue in GlxA). B) The substrate-binding pocket of Gox looking down the 7-fold symmetry axis towards domain 3. Location of the surface exposed Tyr^{495} Cu ligand and the stacking Trp^{290} as well as residues forming the substrate pocket are shown in sticks, with the solvent exposed Cu atom represented as a brown sphere. C) Surface representation of Gox, with location of the Cu ion shown as a brown sphere and residues in the substrate pocket indicated in sticks. Images created in PyMol.

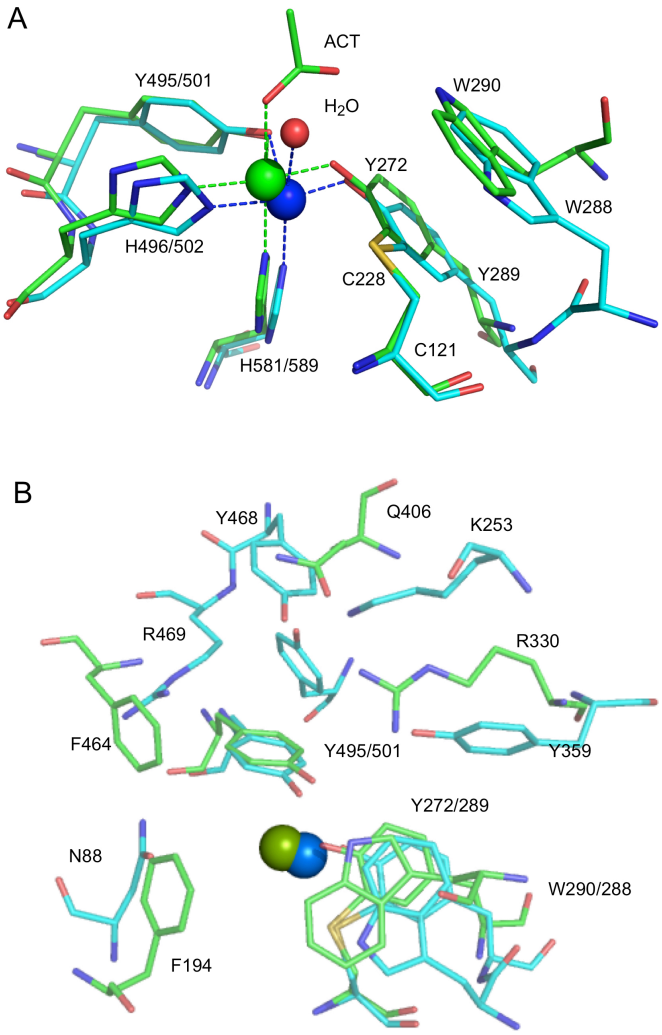


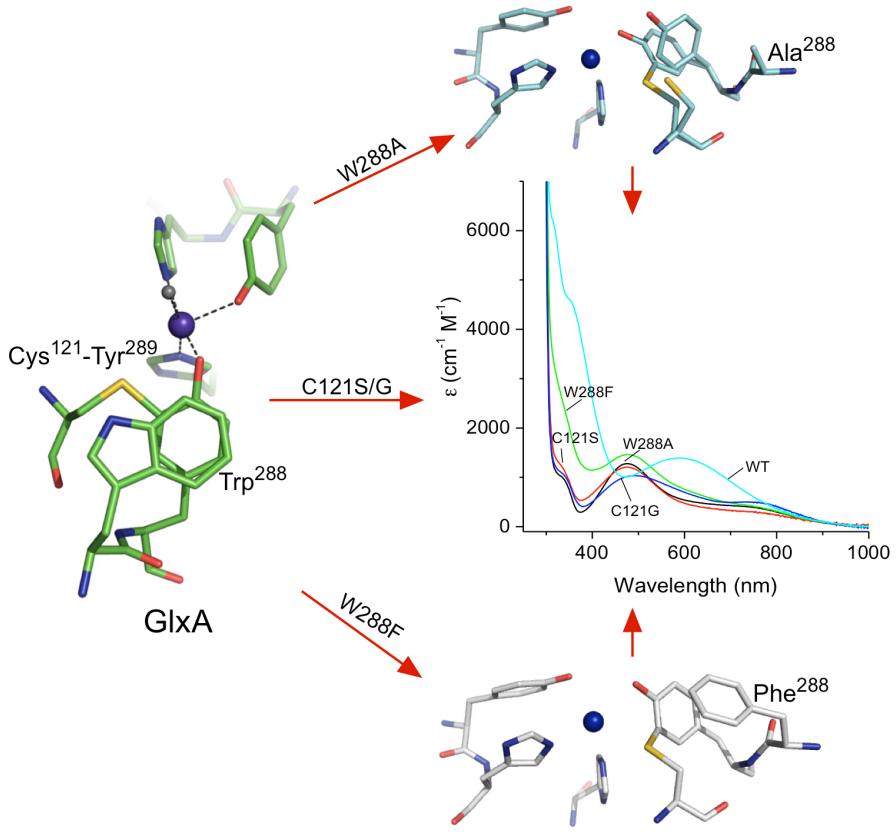
Figure 2.24: Comparison of the active sites of *S. lividans* GlxA and *F. graminarium* Gox (1GOF) [76]. A) Superposition of the Cu sites of GlxA (blue) and Gox (green). Cu atoms (blue-GlxA and green-Gox) and H_2O (red) are shown as spheres and residues depicted as sticks and labelled with the corresponding sequence number. B) Overlay of the substrate pockets of GlxA (blue) and Gox (green). With Cu shown as a sphere (blue-GlxA and green-Gox) and residues depicted as sticks and labelled. Images were made in PyMol.

2.4.4. Conclusions

In this Chapter it has been shown that GlxA is a membrane-associated cuproenzyme with a mononuclear Cu site and possess in addition to a redox active Cu site a Cys-Tyr redox cofactor. It is however apparent that GlxA is distinct from the fungal Gox through its unique tertiary structure, atypical spectroscopic properties and a lack of enzymatic activity with classical Gox substrates. The *ΔglxA* null mutant stalls aerial hyphae development on solid media, lacks the hyphal tip glycan and dramatically changes the morphology in liquid cultures. Thus, this data implies GlxA has a significant role in *S. lividans* and is a new structural and functional member of the radical Cu oxidase family. The next step is to investigate GlxA further through mutational variants to gain an insight into the unique nature of this cuproenzyme.

Chapter Three

Characterisation of the cross-linking cysteine and stacking tryptophan in GlxA from *Streptomyces lividans*



***Synopsis*:** GlxA is a novel mononuclear copper oxidase from *Streptomyces lividans*. Removal of the cross-linking cysteine (Cys^{121}) and a second coordination sphere 'stacking' tryptophan (Trp^{288}) imposes significant effects on its unique electronic and structural characteristics.

3.1. Introduction

Post-translational modifications creating new intrinsic cofactors enable a more diverse range of mechanisms in enzyme catalysis [130]. The first broad class of post-translational modifications includes tyrosyl, glycyl, thiyl and tryptophanyl protein radicals [131-133]. The second class includes larger alterations comprising either a covalent modification of an amino acid, peptide cleavage or modification, or formation of a new bond [87, 134]. These post-translational modifications can be formed by either the participation of accessory enzymes or by a self-processing pathway [87]. The resulting post-translational modification may be necessary for additional structural stability, functionality or to provide increased reactivity [87, 134]. Tyr residues appear most frequently in post-translational modifications in the formation of cross-links and quinones [135]. Metal centres are also typically coupled to these quinone cofactors and induce catalytic activity in many well-studied enzymes. These include, amine oxidases, which contain a copper (Cu) site and an active Tyr which is converted by dioxygen (O_2) and Cu to 2, 4, 5-trihydroxyphenylalanine quinone [136-139] and lysyl oxidases, which contain a catalytically active lysyltyrosylquinone [140]. KatG is a haem containing catalase peroxidase enzyme which contains a novel Met-Tyr-Trp crosslink [141]. This post-translational modification forms during a turnover reaction with H_2O_2 in a oxidative and autocatalytic reaction via Compound I enabling katG to possess catalase activity [142, 143]. Cytochrome *c* oxidase contains a Tyr-His bond at the Cu_B site with the His residue acting as a ligand to the Cu ion, creating a tyrosyl radical at this position, enabling enzyme catalysis [144-146]. A covalent bond between a Tyr and a Cys has been observed in the crystal structure of *Mycobacterium tuberculosis* sulphite reductase (NirA), which also plays an essential role in catalysis [147]. Likewise a Cys-Tyr cross-link has also been shown in the structures of mouse and rat cysteine dioxygenase, neighbouring a non-haem iron (Fe) centre and are proposed to play a role in enzymatic turnover [148, 149]. Furthermore, a Cys-Tyr cross-link is also present in glyoxal oxidase, containing a mononuclear Cu site. These two redox centres enable the catalysis of aldehyde oxidation to their corresponding carboxylic acid [80, 125].

A Cys-Tyr (Cys^{228} and Tyr^{272}) thioether cross-link has also been identified in galactose oxidase (Gox) from *Fusarium graminearum* (PDB entry 1GOG) [76]. This was the first X-ray structure to identify a post-translational modification involving one of the metal centre coordinating residues (Tyr^{272}) [76]. This post-translational modification is proposed to occur via an autocatalytic mechanism requiring Cu and an electron acceptor [87, 150]. Whittaker and Whittaker reported [86] that a pre-mature form of Gox generates the Cys-Tyr cross-link aerobically with Cu(I) and Cu(II), but processing was much faster with Cu(I), concluding that *in vivo* processing involves Cu(I) [86]. Extensive investigations have been carried out on the maturation of this post-translational modification with mature (Cys-Tyr cross-link) and pre-mature (no cross-link) proteins possessing different migration patterns on

SDS-PAGE. Mature Gox containing the Cys-Tyr cross-link travels faster than pre-mature Gox, thus making observations of Gox maturation easily detected [87, 150].

Oxidation of mature Gox creates a unique triplet radical species identified through EPR spectroscopy and is situated on the Cys-Tyr co-factor. This thioether cross-link lowers the redox potential to ~ 400 mV from ~1000 mV of an unmodified Tyr enabling generation of a radical species under ambient biological conditions [81, 82, 131, 133, 151]. This enables Gox to oxidise galactose to galacto-hexodialdose via a two-electron oxidation process, coupled to the reduction of O_2 to hydrogen peroxide (H_2O_2) [70, 71, 76, 84].

It is apparent that the environment surrounding protein-based radicals plays an essential role in modulating the reactivity and properties of the radical species [152-155]. In *Escherichia coli* ribonucleotide reductase, a dinuclear Fe containing metalloenzyme, contains a Tyr residue buried deep within a hydrophobic environment that can form a tyrosyl radical [156, 157]. Mutations of surrounding hydrophobic residues in the dinuclear Fe second coordination sphere substantially lower the radical stability compared to the wild-type (WT) protein [156, 157]. Extensive investigations have also been carried out on the importance of second sphere coordination resides affecting substrate binding in superoxide dismutase [158], substrate affinity and catalytic activity of phenylalanine hydroxylase [159, 160], reduction potential in the $[4Fe\text{-}4S]^{2+/+}$ cluster of ferredoxin I [161] and the enzymatic activity of lipoxygenase-1 [162].

In Gox, Trp^{290} is a second coordination sphere residue that creates a 'stacking' interaction directly on top of the Cys-Tyr cross-link and has been implicated in being important for radical behavior and enzyme catalysis [122]. Studies on Trp^{290} Gox variants including, W290F, W290G and W290H signify the important role of this residue [122, 123]. Enzymatic turnover of Gox for D-galactose is considerably lower in all three variants with the K_m for W290H similar to WT but significantly higher in the W290G and W290F variants suggesting a role of Trp^{290} in substrate binding [123]. Trp^{290} also modulates the redox potential of the Cys-Tyr radical in Gox, with W290H and W290G displaying an unstable radical species with higher redox potentials than determined for the WT (430 mV). This contrasts to the W290F Gox variant, which displays similar radical characteristics to the WT [122, 123, 126, 151, 163]. X-ray crystal structures of these Gox variants illustrate that Trp^{290} restricts access of the Cys-Tyr radical to the surrounding solvent with W290G being significantly more solvent exposed than W290F [123]. Trp^{290} also modulates the reactivity of the catalytic base in Gox, Tyr^{495}, which abstracts a proton from the bound alcohol substrate leading to oxidation and removal of the product from the Cu ion [83, 123, 164]. Therefore, Trp^{290} appears to be a fundamentally important second sphere coordination residue in Gox playing a key role in enzyme catalysis, substrate binding, redox potential/radical stability and controlling the reactivity of the axial Tyr^{495} residue [123].

In Chapter 2 the structural and functional characterisation of GlxA from *Streptomyces lividans* was discussed. GlxA is a membrane-associated mononuclear Cu

oxidase that shares similarities to Gox. It has a Cu coordination sphere identical to Gox and a Cys-Tyr cross-link (Figure 3.1.A). This cross-link possesses the ability to house a triplet radical species as determined by EPR spectroscopy, resembling that identified in Gox. The presence of the Cu ion and the Cys-Tyr radical enables GlxA to perform two-electron oxidation of substrates. However, unlike Gox the physiological substrate of GlxA has not yet been determined with low activity observed for D-galactose and the highest activity determined for glycolaldehyde, the smallest molecule to contain an alcohol and aldehyde group. GlxA also contains the second coordination sphere 'stacking' Trp residue (Trp^{288}) (Trp^{290} in Gox). However, the orientation of this Trp residue shown to be crucial for functional properties of Gox differs in GlxA (Figure 3.1). Trp^{290} in Gox has its indole ring stacking over the Cys-Tyr bond (Figure 3.1.B), whereas in GlxA Trp^{288} no longer has its indole ring stacking with the Cys-Tyr cross-link, but instead has the benzene ring component π - π stacking with the Tyr^{289} Cu ligand (Figure 3.1.C). The location of these two important residues, Cys^{121} and Trp^{288} are shown by a ^ and + symbol, respectively, within the sequence of GlxA from *S. lividans* in Figure 2.3 in Chapter 2.

The X-ray structure of GlxA revealed a new fold topology for this family of mononuclear Cu oxidases, with a Cu site buried deep within the protein in contrast to Gox. Absorbance and EPR spectroscopy revealed the electronic properties of the Cu(II) site differ significantly to those reported for Gox, despite the first sphere coordination ligands being identical. In this Chapter variants of Cys^{121} and Trp^{288} in GlxA have been constructed, over-expressed and purified. Four variants were created, C121G, C121S, W288A and W288F and through the use of UV-visible and EPR spectroscopy, X-ray crystallography, stopped-flow absorbance kinetics, western blotting and enzymatic activity an evaluation of the role these residues play in regulating the spectroscopic properties, Cu(II) binding, maturation, and functional properties of GlxA was determined.

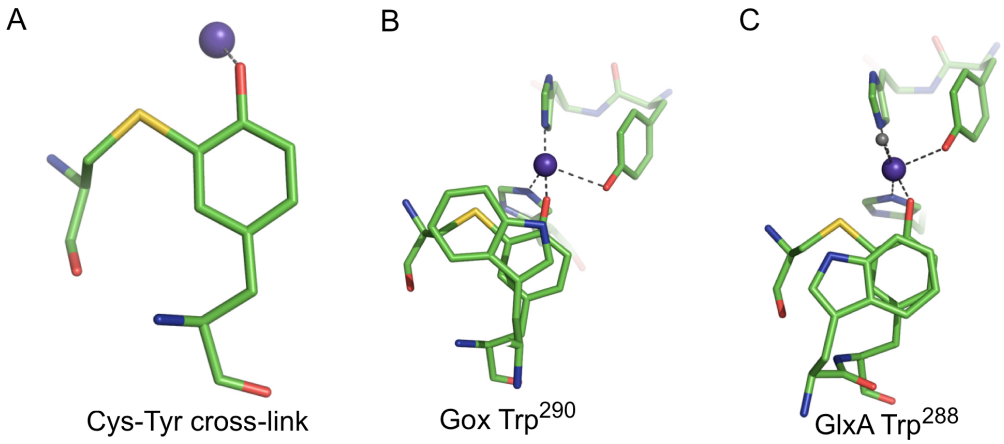


Figure 3.1: Cys and Trp environments within GlxA and Gox. A) Cys-Tyr cross-link identified in the structure of GlxA (PDB entry 4UNM) B) Stacking Trp^{290} in Gox from *F. graminearum* (PDB entry 1GOF). C) Stacking Trp^{288} in GlxA (PDB entry 4UNM). The Cu ion is shown as a purple sphere and a coordinating H_2O molecule in (C) as a grey sphere.

3.2. Experimental procedures

3.2.1. Site-directed mutagenesis of GlxA variants

Quikchange mutagenesis was used to create the C121G, C121S, W288A and W288F variants of GlxA. The mutagenic primers used to introduce the desired mutations are shown in Table 3.1 with 75 ng/µl used in the reaction mixture. The pET28a plasmid containing the WT GlxA gene (*pET3188*) (starting at residue 35) was used as the template DNA (15 ng/µl) in a total PCR reaction volume of 30 µl consisting of; 0.6 µl Pfu Turbo polymerase (Stratagene), 3 µl of 10 x Pfu buffer (Stratagene), 1 µl of each specific primer and plasmid DNA (*pET3188*), 1.8 µl DMSO (100 %), 0.6 µl dNTP's (10 mM) and 21 µl deionised water. The following program was used to carry out the PCR: (i) 95 °C for 3 min, (ii) 95 °C for 30 s, (iii) 58 °C for 1 min, (iv) 68 °C for 13 min, (v) repeat steps (ii-iv) 15 times, 72 °C for 8 min and a 10 °C hold. PCR products were checked on 1 % agarose gels before DpnI digestion (Figure 3.2.A). After DpnI digestion PCR products were transformed into JM109 cells followed by overnight cultures and mini-preps. Mutant plasmids were checked by digestion with NdeI and BamHI and then sequenced to confirm that the intended nucleotide changes had successfully been introduced (Figure 3.2.B)

Table 3.1: Mutagenic primer pairs used to generate C121G, C121S, W288A and W288F GlxA mutations. Underlined are the triplet codons that are changed from the original sequence to generate the respective GlxA mutation.

Primer	Length	GC %	Tm (^{o}C)
C121G-F 5′-GTCCGACCTGTTCGGCACCGGACACA-3′ C121G-R 5′-TGTGTCCGGTGCCGAACAGGTCGGAC-3′	26	65	61
C121S-F 5′-CCGACCTGTTCTCCACCGGACACACGC-3′ C121S-R 5′-GCGTGTGTCCGGTGGAGAACAGGTCGG -3′	27	66	62
W288A-F 5′-CACGAGGCCCGCGCGTACCCGACGC-3′ W288A-R 5′-GCGTCGGGTACGCGCGGGCCTCGTG-3′	25	80	66
W288F-F 5′-CACGAGGCCCGCTTCTACCCGACGC-3′ W288F-R 5′-GCGTCGGGTAGAAGCGGGCCTCGTG-3′	25	72	62

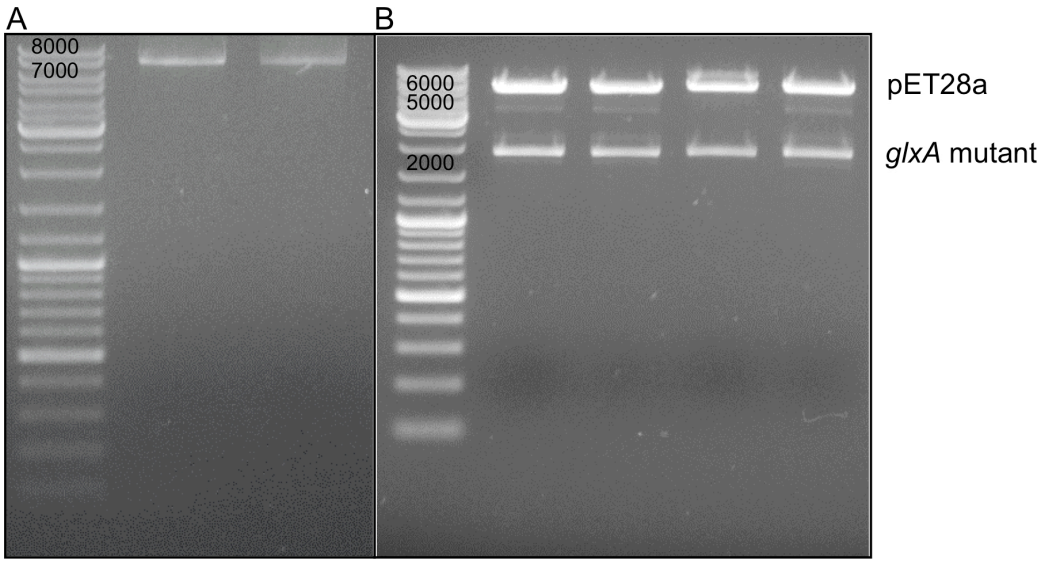


Figure 3.2: 1 % agarose gels of PCR products and restriction digests of *glxA (SLI3188)* mutants. A) PCR product of the *pET3188* plasmid containing the *glxA* mutants C121G and C121S. B) Restriction digest of the *pET3188* mutant constructs, C121G, C121S, W288F and W288A digested using NdeI and BamHI restriction enzymes and corroborated against the DNA ladder (far left lane) (pET28a – 5,369 bp and *glxA* – 1,863 bp).

3.2.2. Over-expression and purification of GlxA variants

The C121G, C121S, W288A and W288F GlxA variants were over-expressed in *E. coli* and purified as previously reported for WT GlxA (Chapter 2 section 2.2.3) (Chaplin *et al*., 2015). However, Cu(II)SO_4 (Sigma) addition was omitted before and after lysis so as to ensure Cu starved conditions during variant preparations.

3.2.3. Apo WT-GlxA preparation by autoinducing media

Using methods previously outlined (Chapter 2 section 2.2.3) (Chaplin *et al*., 2015) to over-express and purify WT GlxA consistently yielded Cu bound protein even without addition of exogenous Cu. It was therefore necessary to adopt a new expression procedure to obtain completely Cu free WT enzyme. The use of autoinduction media to improve expression levels of the fungal Gox enzyme has previously been reported (Deacon and McPherson 2011). This method was originally developed by Studier in 2005 and removes the need for IPTG used as an inducing agent [165]. All plastic and glassware were soaked in 0.1 M EDTA and rinsed extensively with doubly deionised water prior to use to ensure essentially metal free conditions. GlxA was over-expressed in *E. coli* strain BL21 (DE3) starting from a pre-culture (3 ml LB, 3 µl Kan (50 µg ml^{-1})) grown at 37 ºC for 6 h with shaking at 200 rpm. 1 ml of this culture was subsequently used to inoculate 400 ml (in 2 L

baffled flasks) of autoinducing medium designated as 8ZY-4LAC-SUC (Deacon and McPherson 2011) with the contents shown below in Table 3.2 and containing Kan (50 μg ml^{-1}). Cultures were grown at 180 rpm, 25 ºC for 48 h. Cells were harvested via centrifugation (10,000 g**,** 10 min, 4 °C) then resuspended in buffer A (50 mM Tris/HCl, 500 mM NaCl, 20 mM imidazole, pH 8) and purified as previously described for Cu(II)-GlxA ensuring the presence of 2 mM EDTA during dialysis (Chapter 2 section 2.2.3) [166].

Table 3.2: Autoinduction media used to produce apo-GlxA. The following solutions were prepared and sterilised as described by Studier [165]. 8ZY: Yeast extract (40 gL^{-1}) and Tryptone (80 gL^{-1}), 50 x LAC: glycerol (25 %), glucose (2.5 %), and α-lactose monohydrate (10 %), 20 x NPSC: NH_4Cl (1 M), Na_2SO_4 (0.5 M), KH_2PO_4 (0.5 M), Na_2HPO_4 (0.5 M).

	Volume (ml) added per L of medium				
	8ZY	50 x LAC	20 x NPSC	1 M $MgSO_4$	0.5 M Sodium Succinate
8ZY-4LAC-SUC	868	80	50	2	50

3.2.4. UV-visible and circular dichroism spectroscopy of GlxA variants

Concentrations of the GlxA variants were determined using a Varian Cary 50 UV-visible spectrophotometer with a 1 cm path-length Quartz cuvette (Hellma) at 20 ^oC. The absorbance at 280 nm was measured and concentrations calculated using the beer-lambert law with an extinction coefficient (ε) of 78,730 $M^{-1}cm^{-1}$ for WT-GlxA, C121G and C121S and 73,230 $M^{-1}cm^{-1}$ for W288A and W288F determined using ProtParam ExPASy [93]. The proteins were oxidised by the addition of either a > 50-fold excess of ferricyanide ($[Fe(CN)_6]^{3-}$) (Sigma) or iridium chloride ($[Ir(Cl)_6]^{3-}$) (Acros) followed by removal of the oxidant using a PD-10 desalting column (GE Healthcare).

Visible circular dichroism (CD) spectroscopy was used to monitor the changes between the Cu(II) loaded GlxA variants. All GlxA samples for CD analysis were exchanged into 10 mM KPi, 50 mM KF, pH 7 using a PD-10 desalting column (GE Healthcare). Visible CD were recorded between 800 and 300 nm at 20 ^oC on an Applied Photophysics Chirascan CD spectrophotometer equipped with a thermostatic cell holder controlled with a Peltier system (Leatherhead, UK).

3.2.5. Cu(II) binding to the GlxA variants

For Cu(II) binding to the GlxA variants a stock of $Cu(II)SO_4$ were prepared and diluted as required. Cu(II) aliquots were titrated into C121S and W288F (50 μM) with 3 min incubation times before the absorbance spectrum was recorded. Cu(II) binding to C121G and W288A (200 μM) were carried out by the addition of 1-equivalent Cu(II) to the proteins and the absorbance spectrum monitored every 2 or 5 min, respectively, until no further change was recorded. All spectra were base-lined with buffer and recorded in 50 mM Tris/HCl, 150 mM NaCl, pH 8.

3.2.6. Stopped-flow transient kinetics of Cu(II) binding to the GlxA variants

Transient-state kinetics of Cu(II) binding to the GlxA variants were performed using a SX20 stopped-flow spectrophotometer (Applied photophysics) equipped with a diode array unit or single wavelength photomultiplier, thermostatted at 25 ^{o}C with a Peltier system. GlxA variants (17 μM after mixing) were prepared in 50 mM Tris/HCl, 150 mM NaCl, pH 8 and rapidly mixed with excess $Cu(II)SO_4$ solutions. Spectra were recorded upon mixing using the photo-diode array multi-wavelength unit to observe overall spectral transitions. Reactions were then monitored using single wavelengths of 280 nm for C121G, C121S, W288A and W288F variants in addition to 480 nm for C121S and W288F variants. All kinetic data were analysed using the Applied Photophysics ProKinetist software. Data were fitted using single exponential equations, Equation 3.1 was used to obtain k_{obs} values at different [Cu(II)] for the 280 nm data and Equation 3.2 to obtain k_{obs} values for the 480 nm data (whereby A is the amplitude, k is the observed rate constant, t is time and C is a constant). The obtained rates were an average of three shots and were plotted against the [Cu(II)]. Second-order rate constants for initial Cu(II) binding at 280 nm could be calculated from the linear fits with the standard error in the data indicated. Equation 3.3 is derived in Appendix 1.1, and is used to determine the k_2, the first order rate constant for the rearrangement step following the initial Cu(II) binding in Scheme 1, and K_d the equilibrium dissociation constant for the first binding event in Scheme 1 (see Results).

$$y = A\,(\exp^{-kt}) + C \qquad \text{Equation 3.1}$$

$$y = A\,(1-\exp^{-kt}) + C \qquad \text{Equation 3.2}$$

$$k_{obs} = \frac{(k_2 * [Cu(II)])}{(K_d + [Cu(II)])} \qquad \text{Equation 3.3}$$

3.2.7. EPR spectroscopy of GlxA variants

All four GlxA variants for EPR analysis were prepared in a mixed buffer system of 10 mM Tris/HCl, potassium acetate, MES, MOPS and 200 mM potassium chloride pH 7. Samples (120 μM) were prepared after the addition of 1-equivalent $Cu(II)SO_4$ and after the addition of excess $K_3[Fe(CN)_6]$ or $[Ir(Cl)_6]^{3-}$ followed by removal of excess Cu(II) or oxidant using a PD-10 column before measurement. Wilmad SQ EPR tubes were filled with 250 μl of the corresponding GlxA variant solutions, frozen in methanol, kept on dry ice then wiped and transferred to liquid nitrogen. EPR spectra were measured at 40 K on a Bruker EMX EPR spectrophotometer (X-band) at a modulation frequency of 100 kHz. A spherical high-quality Bruker resonator SP9703 and an Oxford instruments liquid helium system were used to measure the low temperature EPR spectra. The EPR spectra of the blank samples (frozen water) were subtracted from the EPR spectra of the protein samples to eliminate the baseline caused by the resonator walls, quartz insert or quartz EPR tube. The baseline was corrected by subtraction of a polynomial line drawn through a set of points randomly chosen on the baseline.

3.2.8. EPR Cu(II) simulations

Spectral simulations were carried out using SimFonia v. 1.26 (Bruker Analytik, GmbH). The g_z value and its hyperfine splitting constant $|A_{Cu2+,z}|$ of the electronic spin S = 1/2 interacting with the Cu nucleus I = 3/2 were determined directly from the spectral parallel components. To maintain consistency in the simulation procedure while varying the minimal number of parameters, we imposed a number of constraints on the choice of the other simulation parameters. The A_{Cu2+} was assumed axial, the x and y values were set to 1/10 (or close to that) of the z value (measured directly from the spectra). All A-tensors were assumed collinear with the g-tensor (all Euler angles were assumed zero). The two Cu(II) coordinating nitrogen's were assumed to produce isotropic hyperfine interactions (A = 14 G and A = 15 G for Sim2 of C121G) with the electronic spin. Following previous observation for tyrosyl radical EPR spectra simulation (Svistunenko and Cooper, 2004) individual line widths were assumed to be linearly dependent on the g-values, with the higher the g_i, the higher the ΔH_i value (where i is either x, y or z component) [167]. All these constraints effectively leave just two numbers in the simulation of each spectrum – g_x and g_y to be varied.

3.2.9. Enzymatic activity assays

Enzymatic activity for the GlxA variants was performed using glycolaldehyde (Sigma) as described for WT GlxA in Chapter 2 section 2.2.10.

3.2.10. Crystallisation and structure determination of the GlxA variants

W288A and W288F as isolated (15 mg ml^{-1}) were first used to screen a range of crystallisation conditions using an ARI-Gryphon 96-well crystallisation robot. Initial crystal hits were discovered predominately in the PEG suite (Qiagen) as seen for WT GlxA. Scaling-up and optimisation from the initial hits were carried out in 24-well plates using the hanging drop vapour diffusion method at 18 ^{o}C. 1 µl of protein solution at 15 mg ml^{-1} was mixed with an equal volume of reservoir solution containing 20 % 20 K polyethylene glycol (PEG) (Sigma), 0.1 M sodium acetate, pH 4. Long needle crystals suitable for diffraction grew within 48 h. Single crystals were transferred into a cryoprotectant solution containing the reservoir solution and 20 % glycerol for 30 s for W288F and 20 % ethylene glycol for 30 s for W288A. W288A crystals were soaked in a cryoprotectant containing the reservoir solution, 20 % ethylene glycol and 10 mM $Cu(II)SO_4$ for 5 min prior to flash cooling the crystals to 100 K by plunging into liquid nitrogen. Crystallographic data were measured to 1.92, 1.49 and 2.3 Å resolution for W288A, Cu(II)-soaked W288A and W288F, respectively at the I03 beamline, Diamond Light Source using an X-ray wavelength of 0.979 Å and a Pilatus 6-M-F detector (Dectris).

Diffraction data obtained from W288A, Cu(II)-soaked W288A and W288F were processed automatically in XIA2 and scaled and merged using aimless [95] in the CCP4i suite. The structures were solved by molecular replacement using Molrep with the WT GlxA structure as input (PDB entry 4UNM) [166]. This produced a solution for two molecules in the asymmetric unit for W288A and W288F and only one molecule in the asymmetric unit for Cu(II)-soaked W288A that was predicted from the Matthews coefficient. The final structures were refined using Refmac5 [98], with model building between refinement cycles in Coot [100]. Riding hydrogen atoms were added when refinement of the protein atoms had converged. The GlxA structures were then validated using the Molprobity server [101], the JCSG Quality Control Server and tools within Coot [100]. A summary of data refinement statistics and the quality indicators for the structures are given in Table 3.3.

Table 3.3: X-ray data processing and refinement parameters for Trp^{288} GlxA variants. W288A is in the space group $P2_1$ with unit cell parameters of 50.5, 126.3, 107 Å, 90, 91.2, 90 °, Cu(II)-soaked W288A is in the space group $P2_12_12_1$ with unit cell parameters of 70.5, 80.9, 140.3 Å, 90, 90, 90 ° and W288F in the space group of $P2_1$ with unit cell parameters of 50.7, 126.8, 107.5 Å, 90, 91.2, 90 °. Values in parentheses refer to the outermost resolution shell.

	W288A	Cu(II)-soaked W288A	W288F
Resolution (Å)	35.6-1.92	40.5-1.49	50.6-2.3
Unique reflections	99551 (4825)	130534 (6400)	59208 (4591)
Completeness (%)	97.6 (95)	99.5 (99.5)	98.4 (99.1)
Redundancy	3.6 (3.6)	3.7 (3.8)	4.0 (4.0)
R_{merge} (%)	0.063 (0.640)	0.08 (0.661)	0.094 (0.533)
Mn (I/sd)	11.7 (2.0)	7.6 (2.0)	7.8 (2.0)
Wilson B factor ($Å^2$)	31.3	14.1	37.6
R_{cryst}	0.19	0.162	0.19
R_{free}	0.25	0.186	0.27
RMSD Bond lengths (Å)	0.016	0.028	0.018
RMSD Bond angles (°)	1.77	2.57	1.90
ESU based on ML (Å)	0.122	0.036	0.19
Ramachandran favoured (%)	95.3	95.8	94

3.2.11. Western blotting analysis

To determine the formation and maturation of the Cys^{121}-Tyr^{289} bond, Cu(II)-loaded WT GlxA, apo-WT GlxA, the C121G variant, apo- and Cu(II)-loaded W288A and W288F variants were monitored by western blotting. Purified proteins were diluted using loading buffer (1 % SDS, 25 % Glycerol, 50 mM Tris/HCl pH 6.8, 0.05 % bromophenol blue and deionised water) and run on a 7.5 % SDS-PAGE gel at 100 V for ~ 2 h. While the SDS-PAGE gel was running transfer pads and Whatmann filter paper were soaked for ~ 2 h in transfer buffer (0.5 X Towbin (10 X Towbin: 30 g Tris + 144 g Glycine, pH 8.3 per liter) + 10 % MeOH) and PVDF membrane was soaked for 10-30 s in 100 % MeOH followed by ~ 2 h in transfer buffer. Following running of the SDS-PAGE gel transfer to the PVDF membrane was carried out for 60 min at 200 mA (150-250 V) in transfer buffer under ice-cold conditions. Membranes were then washed (protein side up) with 25 ml PBS (PBS 10 X per liter: NaCl 80 g, KCl 2 g, Na_2HPO_4 14.4 g, KH_2PO_4 2.4 g, 800 ml H_2O, pH 7.4) then incubated for ~ 3-4 h in PBS plus 5 % milk whilst rocking at room temperature. Each membrane were then incubated overnight at 4 ºC in 20 ml PBS plus 5 % milk (fresh) with primary GlxA antibodies (1/20,000 diluted) with gentle rocking. To remove any non-specifically bound antibodies membranes were washed 3-5 times in 25 ml PBS for 15-20 min. Membranes were then washed for 5 min at room temperature in 25 ml TBST (TBST 10 X per liter: 12 g Tris, 88 g NaCl, 5 ml Tween, pH 7.4) followed by incubation at room temperature with gentle rocking in 25 ml TBST plus 2.5 µl GARAP (goat anti rabbit alkaline phosphatase; Sigma) used as the secondary antibody. After which the membranes were washed 4-5 times in 25 ml TBST at room temperature (to remove any unbound GARAP), washed briefly in 25 ml AP-buffer (AP-buffer 1 X per liter: 100 mM Tris/HCl 12.1 g, 150 mM NaCl 8.7 g, 1 mM $MgCl_2.6H_2O$ 0.2 g, pH 9) then incubated (without rocking) with 5 ml AP-buffer + 30 µl NBT (Nitro Blue Tetrazolium 50 mg/ml in 70 % DMF) + 37.5 µl BCIP (5-bromo-4-chloro-3-indolyl phosphate 20 mg/ml in 100 % DMF) used as the substrates. As soon as expected signals were visible the membranes were washed with water then dried and imaged.

3.3. Results

3.3.1. Over-expression and purification of apo WT and apo-GlxA variants

Over-expression of WT GlxA under Cu starved conditions as described for the GlxA variants failed to yield apo-protein (experimental section 3.2.2). It was therefore necessary to use autoinducing media in order to obtain Cu free WT GlxA (apo-GlxA) (see experimental procedures section 3.2.3). Purification of Cu free WT GlxA produced elution profiles and SDS-PAGE gels identical to GlxA purified not using autoinducing media.

GlxA variants, W288A, W288F, C121S and C121G (residues 35-645) were successfully cloned using a site directed mutagenesis approach and over-expressed in *E. coli* under Cu starved conditions. The overall yield of the GlxA variants ranged between 30-55 mg L^{-1}, comparable to WT (~ 45 mg L^{-1}). The proteins were purified in a similar manner to Cu(II)-WT GlxA protein but without any addition of $Cu(II)SO_4$ before and after lysis (see experimental procedures section 3.2.2). All GlxA variants produced clear single peaks on a S200 size-exclusion column, with elution volumes of ~ 80 ml consistent with monomeric species as observed for WT (Figure 3.3). Fractions eluting from the major peak were analysed on 15 % SDS-PAGE gels (see Chapter 2, experimental 2.2.4) and all displayed a major band running at ~ 75 kDa consistent with the approximate molecular weight of GlxA (67 kDa after His_6-tag removal) (Figure 3.3).

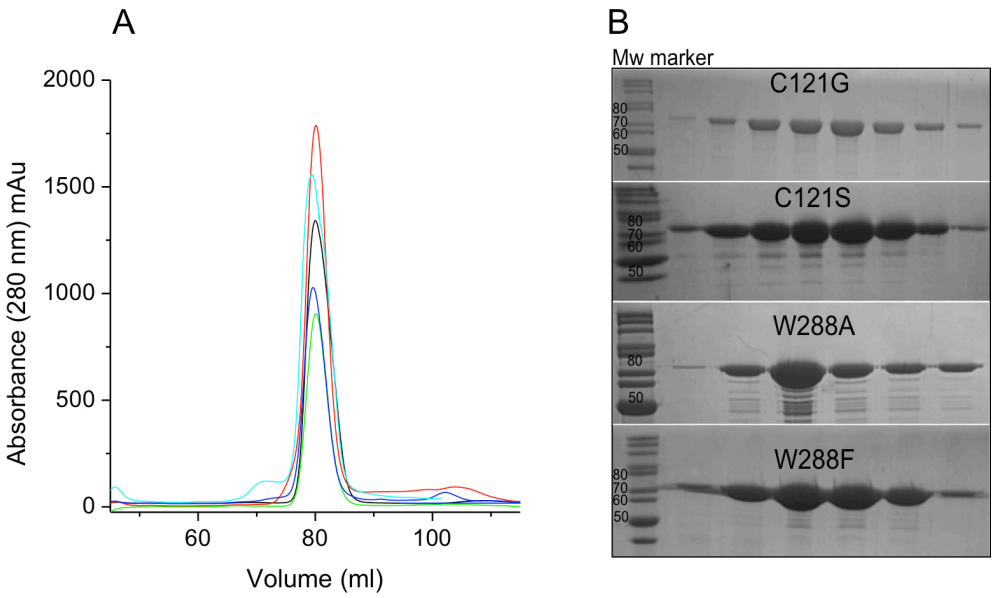


Figure 3.3: Purification of *S. lividans* GlxA variants, C121G, C121S, W288A and W288F. A) S200 Sephadex column absorbance 280 nm (mAu) profile of the WT and GlxA variants; C121G (red), C121S (cyan), W288A (black), W288F (blue) and WT (green), all eluting at ~ 80 ml in 20 mM NaPi, 100 mM NaCl, pH 7. B) Coomassie stained 15 % SDS-PAGE gels of fractions from the S200 Sephadex column. A major protein band at ~75 kDa is identified as the pure GlxA variant.

3.3.2. Cu(II) binding to the GlxA variants

As described in Chapter 2 WT Cu(II)-GlxA has a blue-grey colour with a UV-visible absorbance spectrum consisting of a broad, low intensity band in the visible region (λ_{max} ~ 577 nm) and two distinct shoulders (λ_{max} ~ 362 and 320 nm) (Figure 2.6 Chapter 2). The absorbance spectrum remains unchanged upon addition of Cu(II) or oxidants. This differs from the Cu(II)-Gox absorbance spectrum that has weak intensity bands at λ_{max} 441 nm and at 630 nm, with the fully oxidised Cu(II)-Cys-Tyr• form having characteristic absorption bands at λ_{max} 445 and 810 nm [81, 83].

The GlxA variants W288A, W288F, C121S and C121G were determined to be relatively Cu-free following purification, judged by either the absence or weak absorbance of the features characteristic of the Cu(II)-WT-GlxA spectrum. Initial addition of Cu(II) to each variant were monitored by UV-visible spectroscopy, where it was found that the four variants bound Cu(II) at different rates. W288F and C121S bind Cu(II) with a faster rate than W288A and C121G, with the final Cu(II) spectrum appearing within ~ 200 s. It was therefore possible to perform a Cu(II) titration with W288F and C121S (Figure 3.4). Upon aliquot additions of Cu(II) the absorbance at ~ 480 nm increases in a linear fashion until a break point is reached at a 1:1 stoichiometry for both C121S (Figure 3.4.A and B) and W288F (Figure 3.4.C and D). In contrast, W288A and C121G display much slower Cu(II) binding. The absorbance at ~ 480 nm of both variants was monitored over a period of time following the addition of 1 equivalent of Cu(II). A rate constant for the formation of the 480 nm band could be determined by fitting to a single exponential function (Equation 3.2) to the data shown in Figure 3.5. Rate constants of 6.18 (± 0.2) x 10^{-4} s^{-1} and 3.7 (± 0.1) x 10^{-5} s^{-1} were determined for C121G (Figure 3.5.B) and W288A (Figure 3.5.D), respectively.

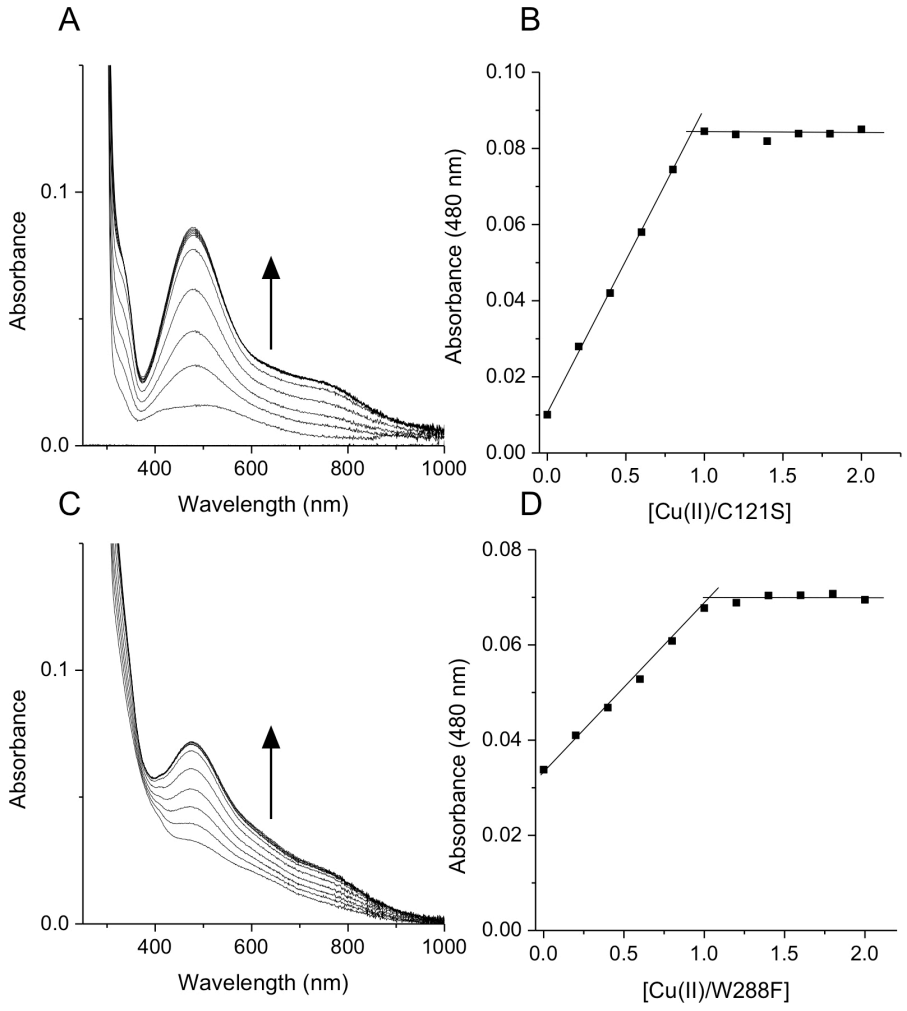


Figure 3.4: Cu(II) titrations to C121S and W288F GlxA variants monitored by UV-visible absorbance changes at 20 ^{o}C in 50 mM Tris/HCl, 150 mM NaCl, pH 8. A) The absorbance spectrum of C121S (50 μM) with increasing aliquots of 10 mM Cu(II)SO_4. B) The absorbance changes observed for C121S at 480 nm plotted as a function of [Cu(II)/C121S]. C) The absorbance spectrum of W288F (50 μM) with increasing aliquots of 10 mM Cu(II)SO_4. D) The absorbance changes observed for W288F at 480 nm plotted as a function of [Cu(II)/W288F]. The lines through the data points in (B) are (D) are to indicate the stoichiometry of the binding. All spectra were baselined with buffer.

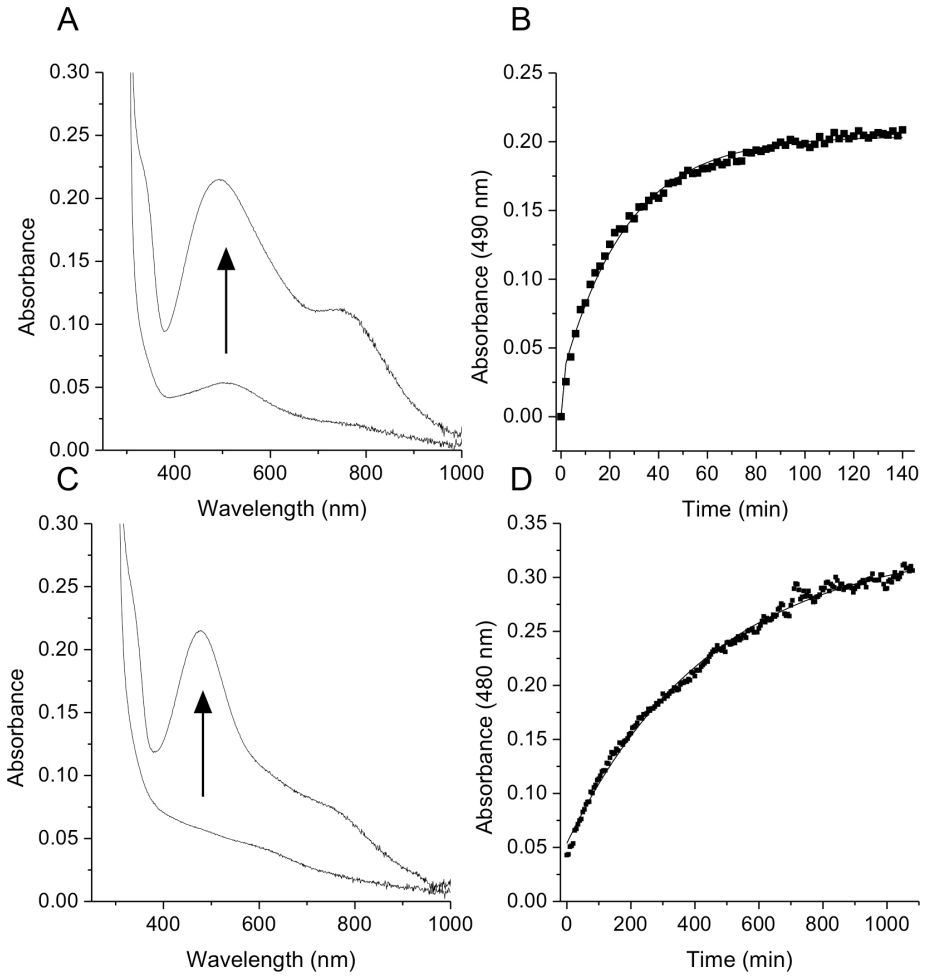


Figure 3.5: Cu(II) binding to C121G and W288A GlxA variants monitored by UV-visible absorbance changes at 20 ^{o}C in 50 mM Tris/HCl, 150 mM NaCl, pH 8. A) The absorbance spectrum of C121G (200 μM) before and after the addition of 1 equivalent of Cu(II)SO_4. B) The absorbance change observed for C121G at 490 nm as a function time (min). C) The absorbance spectrum of W288A (200 μM) before and after the addition of 1 equivalent Cu(II)SO_4. D) The absorbance changes observed for W288A at 480 nm plotted as a function time (min). The data points in B) and D) were fitted to Equation 3.2. All spectra were baselined with buffer.

3.3.4. UV-visible spectrum of Cu(II)-loaded GlxA variants

Comparison of the UV-visible absorption spectrum of each variant to that of the WT Cu(II)-GlxA was possible following stoichiometric addition of Cu(II) to each variant (Figure 3.6). Interestingly, in the Cu(II) form each variant has a pink colour in the buffer used, and a UV-visible spectrum more like that of Cu(II)-Gox rather than WT GlxA (Figure 3.6). All four mutants consist of a weak shifted narrower absorbance peak compared to WT GlxA (λ_{max} ~ 577 nm) at ~ 480 nm (λ_{max} ~ 477 nm for W288A, λ_{max} ~ 475 nm for C121S, λ_{max} ~ 479 nm W288F and λ_{max} ~ 490 nm C121G) (Figure 3.6). They also exhibit a very broad low intensity band at ~ 750 nm. W288A, C121S and C121G contain a single shoulder at λ_{max} ~ 340 nm in contrast to the two distinct shoulders in the WT protein (Figure 3.6). W288F does not contain a distinct shoulder and instead displays a more gradual increase in this region (Figure 3.6). These spectral transitions share similarities to Cu(II)-Gox and to the Gox Cu(II)-Cys-Tyr• state. As reported for WT GlxA all four variants displayed no change in the absorbance spectra upon addition of an oxidant to the apo- or Cu(II)-loaded forms.

3.3.5. Visible CD of GlxA variants

Visible CD of the Cu(II)-GlxA variants also expose differences compared to the WT. WT GlxA (as isolated) revealed a spectrum with a trough at 350 nm and two peaks at ~ 320 and 650 nm, notably different to Gox but still confirming the presence of a coordinated Cu(II) ion (Figure 3.6.C). The GlxA variants all show peaks at ~760, 510, 400 and 315 nm with a major trough at ~ 600 nm and minor troughs at ~ 450 and 350 nm (Figure 3.6.C). Thus, the visible CD data of the GlxA variants are significantly different from the WT protein. C121S and W288A display almost identical visible CD and UV-visible spectra, whereas C121G displays peak variations in the CD and in the UV-visible spectra (λ_{max} ~ 490 nm) in addition to W288F, which displays much broader and weaker transitions (Figure 3.6.C). As noted for the UV-vis spectra, visible CD of the GlxA variants is again more comparable to Gox than WT GlxA. The visible CD spectra of Gox displays distinct minima at approximately ~ 600 and 440 nm and maxima at 320 nm, with clear differences between inactive and activated enzyme states [81]. As was also the case with UV-vis spectroscopy, no notable perturbations in the visible CD were identified after treatment of the GlxA variants with the $[Fe(CN)_6]^{3-}$.

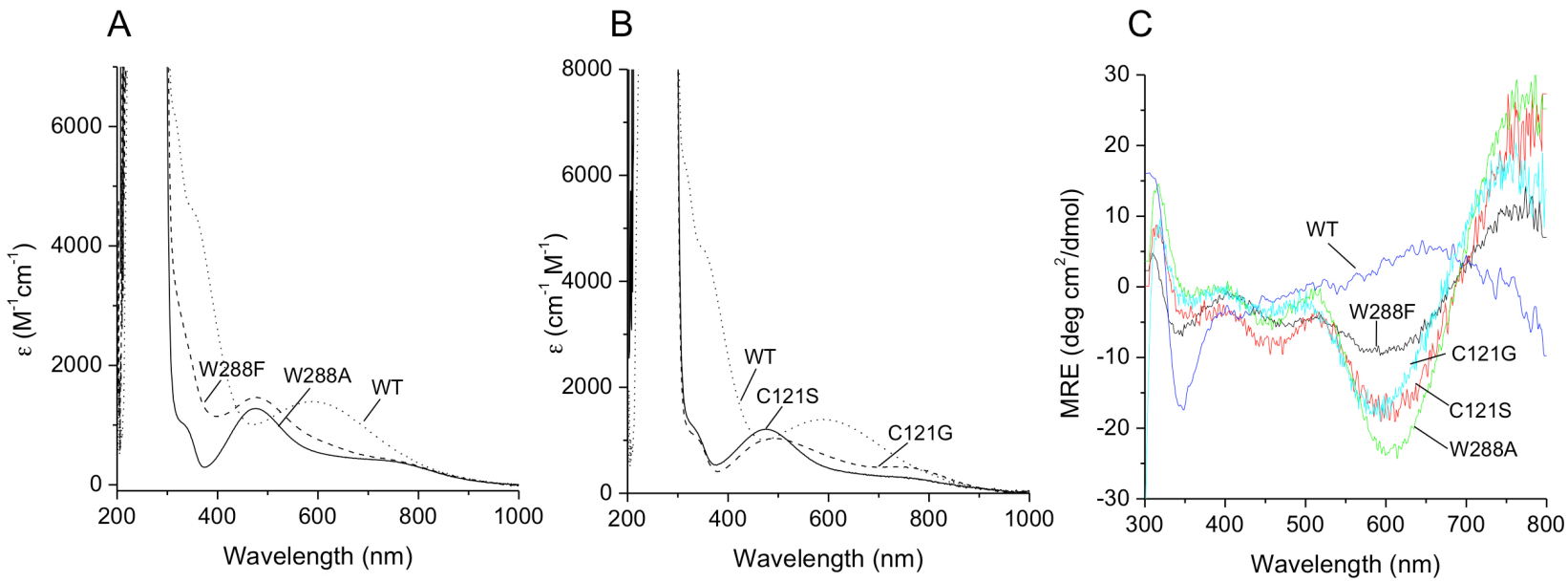


Figure 3.6: UV-visible and CD absorbance spectra of GlxA variants at 20 °C. A) UV-visible absorbance spectra of WT GlxA as isolated compared to Trp variants, W288A and W288F after the addition of 1 equivalent Cu(II)SO_4. B) UV-visible absorbance spectra of WT GlxA as isolated compared to Cys variants, C121G and C121S after the addition of 1 equivalent Cu(II)SO_4. Protein concentrations ranged from 80-200 μM in in 50 mM Tris/HCl, 150 mM NaCl, pH 8. C) Visible CD spectra of as isolated WT GlxA (blue) and four variants after loading with 1 equivalent Cu(II)SO_4, C121G (cyan), C121S (red), W288F (black) and W288A (green) measured between 800 and 300 nm. All CD spectroscopy was carried out in 10 mM KPi, 50 mM KF, pH 7 with protein concentrations ranging between 130-200 μM.

3.3.6. Stopped-flow kinetics of Cu(II) binding to the GlxA variants

Upon mixing apo-C121S and apo-W288F with excess Cu(II) solutions, a slow phase could be observed as an increase in absorbance at ~ 480 nm (Figure 3.7.A and B). Single exponential functions could be fitted to the kinetic traces and the pseudo-first order rate constants (k_{obs}) obtained plotted against [Cu(II)] to display a non-linear concentration dependence (Figure 3.7.C and D). A lag phase within the first 5-10 s was noted within these slow phases, suggesting that this slower phase is not the initial binding event supported by the non-linear [Cu(II)] dependence (Figure 3.7.A and B). Upon monitoring other wavelengths a fast phase could be detected at 280 nm seen by a small decrease in absorbance with a time course corresponding to the lag phase identified at ~ 480 nm (5-10 s) (Figure 3.7.E and F). These kinetic traces were fitted to single exponential functions and the pseudo-first order rate constants (k_{obs}) plotted against [Cu(II)] to display a linear dependence on [Cu(II)] (k_1 Table 3.3) (Figure 3.7.G and H). This data suggests a sequential mechanism for binding Cu(II) comprising a rapid weak initial interaction (280 nm) between the metal and the site followed by a conformational change (480 nm) in the protein, which alters the Cu-bound site to one with a very high affinity for the metal. A simple formulation of this mechanism is as follows:

Scheme 1:

$$\text{GlxA + Cu(II)} \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} \text{GlxA*Cu(II)} \xrightarrow{k_2} \text{GlxA-Cu(II)}$$

Here no rate constant is assigned to dissociation of Cu(II) from GlxA-Cu(II) because the titration data shown in Figure 3.4 suggest an upper-limit for a K_d of 10^{-8} M and thus the dissociation rate constant is very small. Analysis of this model, in which the first step is much faster than the second and may be considered to be in an equilibrium throughout all times of the reaction, yields a slight [Cu(II)] dependence of the slow step due to the observed rate constant (k_{obs}) given in Equation 3.3. From fitting the data of the slow phase at ~ 480 nm in Figures 3.7 for each variant to Equation 3.3 gives k_2 for the slow phase and a K_d for the initial Cu(II) binding phase reported in Table 3.4. This enabled for k_{-1} values to be determined ($K_d = k_{-1}/k_1$) and reported in Table 3.4. These values are somewhat larger than the intercepts of k_{obs} of the fast phase, but are still in reasonable agreement.

For the W288A and C121G variants only the initial Cu(II) binding phase at 280 nm could be monitored by stopped-flow spectroscopy (Figure 3.8) as the subsequent rearrangement at ~ 480 nm was too slow to measure (*vide supra*) (Figure 3.5). The kinetic traces were fitted to single exponential functions (Figure 3.8.A and B) and the pseudo-first order rate constants (k_{obs}) plotted against [Cu(II)] to reveal a linear dependence on [Cu(II)] (Figure 3.8.C and D). From the slopes of these plots k_1 values were determined and reported in Table 3.4. It is apparent that the second order rate constants (k_1) (Table 3.4) of initial Cu(II) binding to all four variants are of similar values, suggesting that Cys^{121} and Trp^{288} are not important for initial Cu(II) capture and only influence the rearrangement step *i.e.* k_2.

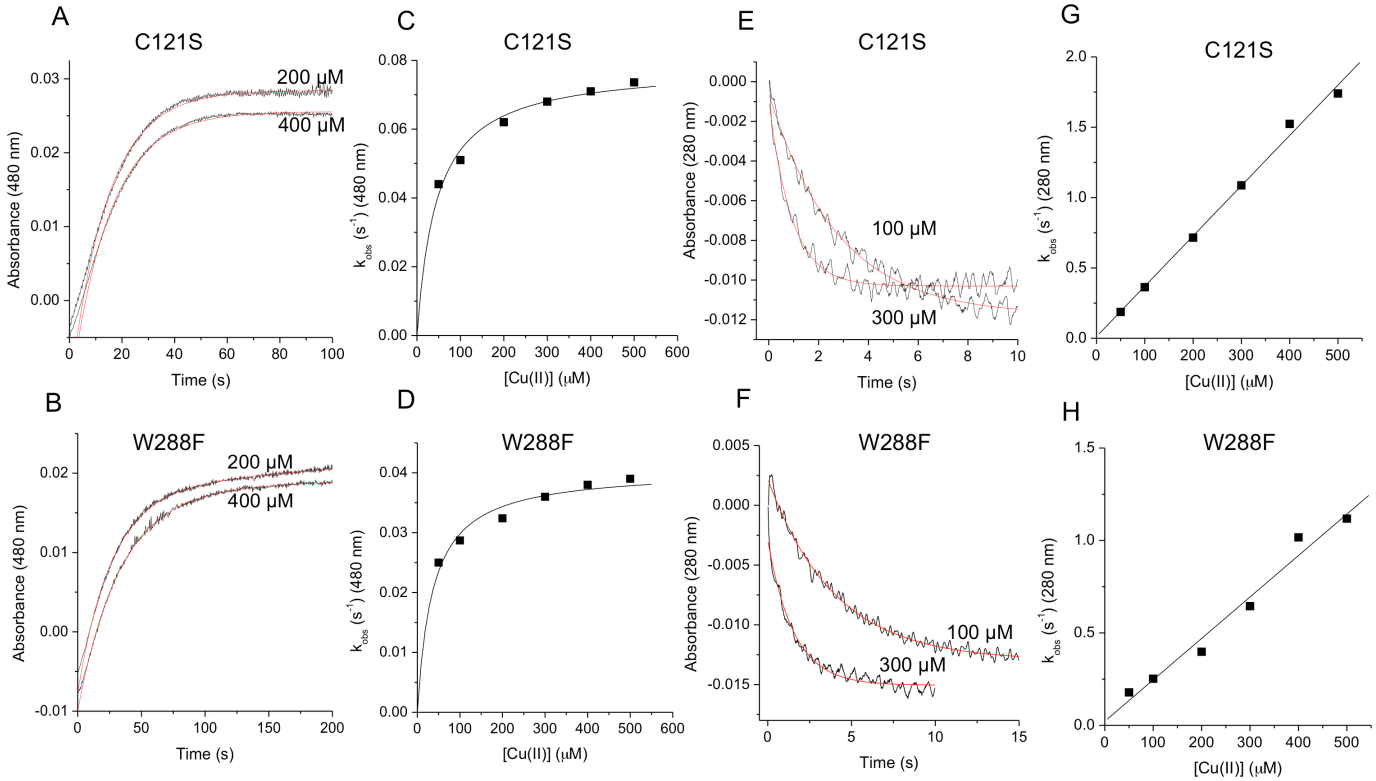


Figure 3.7: Stopped-flow kinetics of Cu(II) binding to C121S and W288F GlxA variants at 25 ^{o}C in 50 mM Tris/HCl, 150 mM NaCl, pH 8. A and B) Time courses over 100-200 s observed on reacting C121S and W288F (17 µM) with Cu(II)SO_4 (200 µM and 400 µM after mixing). C and D) Observed rate constants (k_{obs}) for Cu(II) binding to C121S and W288F over a range of Cu(II) concentrations for the slow phase at 480 nm and fitted to obtain k_2 and K_d values using Equation 3.3. E and F) Time courses over 10-15 s at 280 nm observed on reacting C121S and W288F (17 µM) with Cu(II)SO_4 (100 µM and 300 µM after mixing). G and H) Observed rate constants (k_{obs}) for Cu(II) binding to C121S and W288F over a range of Cu(II) concentrations for fast initial binding at 280 nm. Linear fits yield second order rate constants given in Table 3.4.

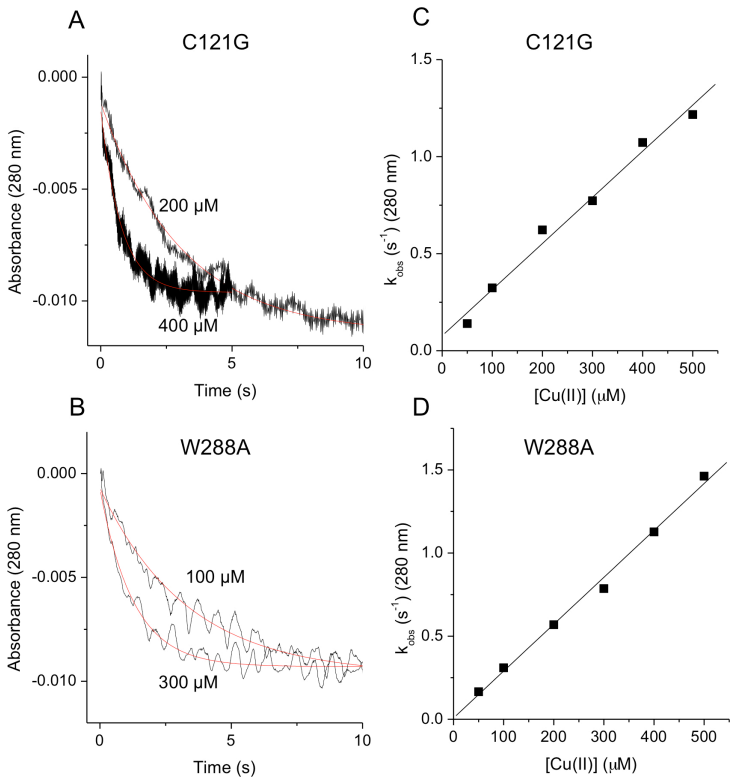


Figure 3.8: Stopped-flow kinetics of initial Cu(II) binding to C121G and W288A GlxA variants at 25 ^{o}C in 50 mM Tris/HCl, 150 mM NaCl, pH 8. A and B) Time courses over 5-10 s at 280 nm observed on reacting C121G and W288A (17 µM) with Cu(II)SO_4 concentrations as indicated. C and D) Observed rate constants (k_{obs}) for Cu(II) binding to C121G and W288A over a range of Cu(II) concentrations for fast initial binding at 280 nm.

Table 3.4: Rate constants for Cu(II) binding to GlxA variants. Second order rate constants (k_1) obtained from linear fits of the stopped-flow kinetics through monitoring the absorbance change at 280 nm and k_2 and K_d obtained from the slow phase at 480 nm by fitting to Equation 3.3 for C121S and W288F. Standard errors are reported in parenthesis.

GlxA variant	k_1 (M^{-1} s^{-1})	k_2 (s^{-1})	k_{-1} (s^{-1})	K_d (µM)
C121G	2.4 (± 0.1) x 10^3	-	-	-
C121S	3.6 (± 0.08) x 10^3	0.078 (± 0.01)	0.17 (± 0.01)	46 (± 1)
W288A	2.8 (± 0.06) x 10^3	-	-	-
W288F	2.2 (± 0.2) x 10^3	0.041 (± 0.01)	0.08 (± 0.01)	36 (± 3)

3.3.7. Cu(II) EPR spectra of the GlxA variants

Once the GlxA variants were loaded with Cu(II) the EPR spectra were recorded and could be directly compared to the WT GlxA spectrum as discussed in Chapter 2 (See Chapter 2, Figure 2.8). Similarly to WT the EPR spectra of the four variants represent typical Cu(II) EPR spectra (Figure 3.9). Previously in Chapter 2 in the absence of any Cu(II) simulations a hyperfine splitting of 1:4:8:10:8:4:1 was identified in the $g_{\perp}$ (perpendicular) region of the WT GlxA spectrum (See Chapter 2, Figure 2.8). This was assigned to two nitrogen atoms and a further hyperfine interaction not immediately apparent. Upon analysis of the GlxA variant spectra, C121S, W288A and W288F appear very similar to the WT spectrum (Figure 3.9). However, C121G displays more detailed hyperfine structures in the $g_{\perp}$ area (Figure 3.9.B). It is also apparent for all four GlxA variants that the $g_{\parallel}$ (parallel) region displays evidence that the measured Cu(II) spectrum is a combination of at least two different Cu(II)-EPR spectra. Once axial symmetry was assumed for C121G, it was possible to simulate the Cu(II) spectra using a simplistic model with only two coordinating nitrogen's present (Figure 3.9.B). All the spectra shown in Figure 3.9 can therefore be represented as a sum of two signals from two different species, one of which, Sim2 has a smaller g_z and a larger A_z than Sim1 and, as a result, exhibits a clear "overshoot" line (Table 3.5). Interestingly, the g_z values for Sim1 and Sim2 are practically conserved across the five variants - 2.27-2.28 for Sim1 and $g_z = 2.18$ for Sim2 (Table 3.5). One obvious feature that singles out the C121G mutant is the well-resolved hyperfine structure in the $g_{\parallel}$ area of the spectrum. The simulation parameters of the spectra (Table 3.5) strongly indicate that this must be the consequence of the fact that this mutant is the only variant that is characterised by axial g-factor anisotropy $g_x = g_y$ in both Sim1 and Sim2. None of the other experimental spectra can be simulated for axial g-factor anisotropy and instead display rhombic symmetry (different to that previously assumed for WT, Chapter 2). Complete sets of simulation parameters used to simulate all the spectra in Figure 3.9 are presented in Table 3.5.

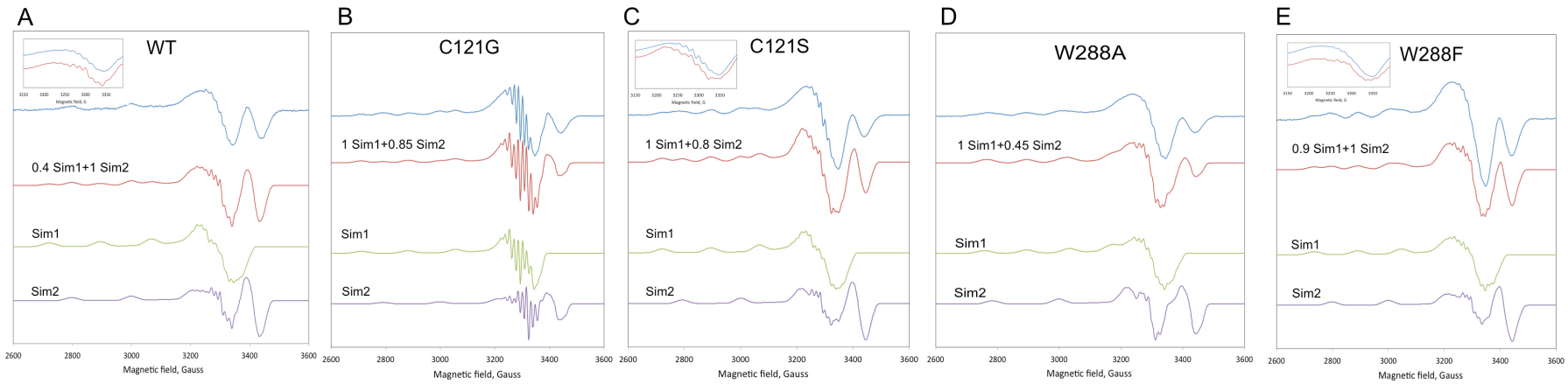


Figure 3.9: EPR spectroscopy of WT GlxA compared to four GlxA variants from *S. lividans*. (A) As purified Cu(II)-GlxA EPR spectrum, (B) C121G GlxA variant loaded with 1 equivalent of Cu(II), (C) C121S GlxA variant loaded with 1 equivalent of Cu(II), (D) W288A GlxA variant loaded with 1 equivalent of Cu(II), (E) W288F GlxA variant loaded with 1 equivalent of Cu(II). All measured spectra are shown in blue, the approximate simulations in red, which are produced by a combinations of individual simulated EPR signals of Sim1 in green and Sim2 in purple which are simulations of Cu(II) ion coordinated by two nitrogen's. WT GlxA was at 85 μM and all variants at 120 μM all at pH 7. Simulation parameters are given in Table 3.5. All EPR spectra were measured at 40 K on a Bruker EMX EPR spectrometer (X band) with a microwave power P = 3.18 mW, a modulation frequency of 100 kHz, modulation amplitude A_M = 5 G, time constant τ = 82 ms, scan rate V = 22.6 G s^{-1}, number of scans per spectrum NS = 1.

Table 3.5: The ERR simulation parameters used to simulate Sim1 and Sim2 for WT GlxA and the four GlxA variants, C121G, C121S, W288A and W288F in Figure 3.9. In bold are the g_z and hyperfine splitting constants (A_z) determined directly from the spectral parallel components.

Sample	Simulation	x, y, z	g	A_{Cu2+}, G	A_N[a)], G	A_N[b)], G	ΔH[c)], G
WT	Sim1	x	2.037	17.2	14	14	12
		y	2.083	17.2	14	14	12
		z	**2.27**	**172**	14	14	25
	Sim2	x	2.033	20.2	14	14	12
		y	2.057	20.2	14	14	10
		z	**2.182**	**202**	14	14	19
C121G	Sim1	x	2.061	15.996	14	14	11
		y	2.061	15.996	14	14	11
		z	**2.279**	**172**	14	14	28
	Sim2	x	2.046	20.8	15	15	11
		y	2.046	20.8	15	15	10
		z	**2.181**	**208**	15	15	20
C121S	Sim1	x	2.0415	17.2	14	14	12
		y	2.0855	17.2	14	14	12
		z	**2.27**	**172**	14	14	26
	Sim2	x	2.028	20.8	14	14	12
		y	2.068	20.8	14	14	10
		z	**2.179**	**208**	14	14	19
W288A	Sim1	x	2.029	13.7	14	14	12
		y	2.069	13.7	14	14	12
		z	**2.283**	**137**	14	14	26
	Sim2	x	2.054	21.5	14	14	12
		y	2.064	21.5	14	14	12
		z	**2.178**	**215**	14	14	19
W288F	Sim1	x	2.027	15.5	14	14	12
		y	2.073	15.5	14	14	12
		z	**2.28**	**155**	14	14	30
	Sim2	x	2.02	20.2	14	14	12
		y	2.06	20.2	14	14	11
		z	**2.182**	**202**	14	14	21

[a)] On one coordinating nitrogen. [b)] On the other coordinating nitrogen. [c)] Increased linewidth.

3.3.8. Detection of a radical species in the C121G and C121S variants

Upon addition of strong oxidising agents to W288A and W288F, no radical species could be detected by EPR spectroscopy as has been reported for the WT GlxA. Interestingly, following addition of $[Fe(CN)_6]^{3-}$ to the C121G and the C121S variants a radical species was observed that is very different to that of the triplet observed for the WT (Figure 3.10).

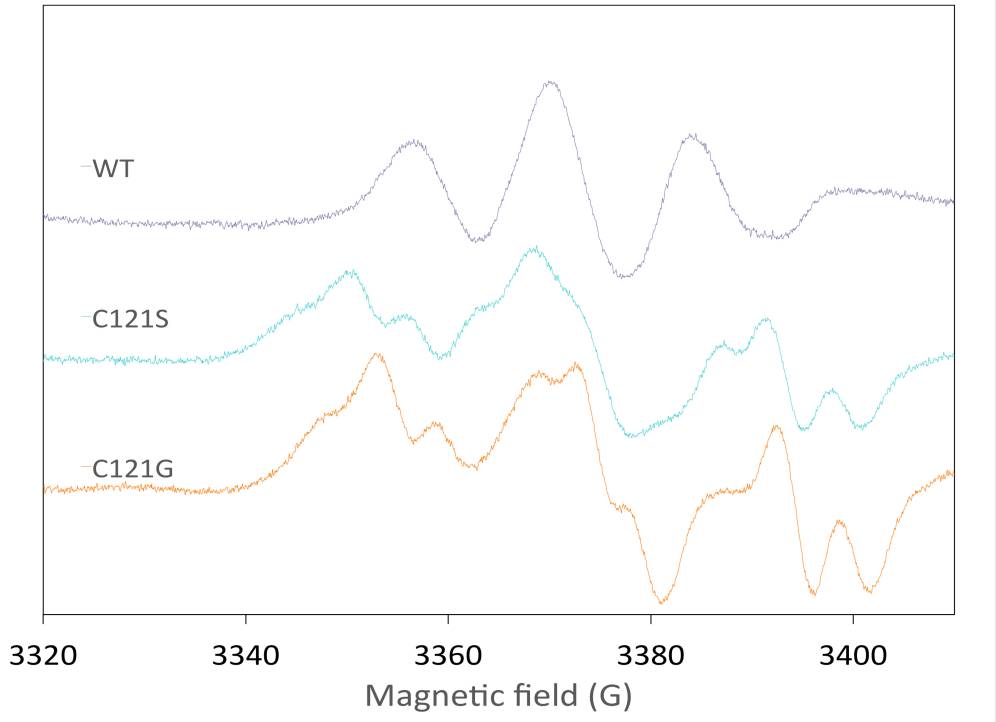


Figure 3.10: EPR spectra of protein radicals present in WT *S. lividans* GlxA, C121S and C121G GlxA variants. Cys-Tyr radical species of WT (grey), protein radical of C121S (blue) and C121G (orange) after the addition 50-fold excess $[Fe(CN)_6]^{3-}$ to as isolated WT GlxA, C121S and C121G following addition of 1 equivalent $Cu(II)SO_4$. All spectra were recorded at 40 K using 120 μM of GlxA protein in a mixed buffer system consisting of 10 mM each of Tris, potassium acetate, MES, MOPS and 200 mM KCl, pH 7. Spectra were recorded with experimental conditions of, microwave frequency ν_{MW} = 9.47 GHz, microwave power P = 3.18 mW, modulation frequency ν_M = 100 kHz, modulation amplitude A_M = 3 G, time constant τ = 82 ms, scan rate V = 0.596 G s^{-1}.

3.3.9. C121G is active with glycolaldehyde

As previously discussed in Chapter 2 enzymatic activity of WT GlxA was measured with a range of putative substrates. However only glycolaldehyde displays any appreciable activity with a K_m of 115 mM, and maximal aerobic turnover number (k_{cat}) of 0.14 s^{-1} giving a V_{max}/K_m of 1.22 M^{-1} s^{-1} (Figure 3.11). All four GlxA variants were assessed for their activity with glycolaldehyde, with only C121G displaying any turnover activity. With increasing glycolaldehyde concentrations C121G displayed Michaelis-Menten kinetics similar to WT GlxA enabling for a K_m of 67 mM to be determined, and a maximal aerobic turnover number (k_{cat}) of 0.08 s^{-1} to give a V_{max}/K_m of 0.96 M^{-1} s^{-1} (Figure 3.11). Therefore, C121G displays a slightly lower catalytic turnover for glycolaldehyde than WT GlxA.

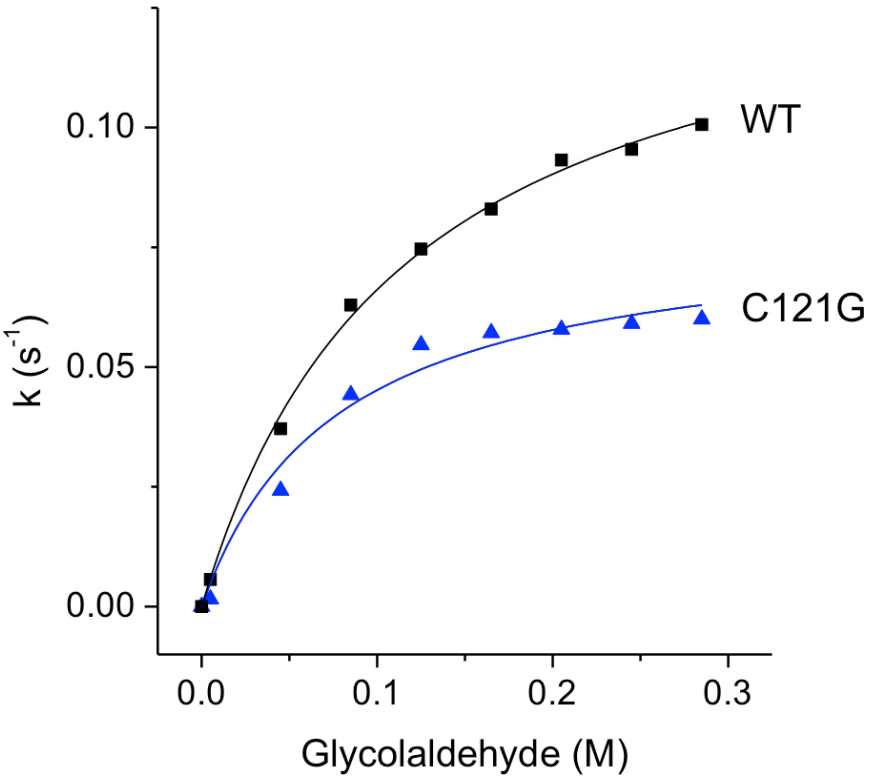


Figure 3.11: Enzyme activity of WT GlxA compared to the C121G variant at 25 °C. Data points were fitted to the Michaelis-Menten equation (Chapter 2, Equation 2.1) to yield a K_m value and turnover rate (k_{cat}) reported in the text. Assays were all carried out in 0.1 M NaPi, pH 7.4.

3.3.10. X-ray crystal structure of W288A, Cu(II)-soaked W288A and W288F

Single needle shaped crystals of W288A and W288F GlxA variants grew in a crystallisation condition identical to the WT protein (Figure 3.12). The W288A crystals were also used for Cu(II) soaking (see experimental procedures section 3.2.10). W288A contains two protein molecules in the crystallographic asymmetric unit, molecule A (residues 38-645) and molecule B (residues 37-645). Molecule A contains a disordered region between residues 198-206 of domain 2 (not visible in the electron density), which in molecule B creates a long β-hairpin loop as seen in the WT structure. Molecule A however contains a higher quality electron density with significantly lower B–factors. This is identical to the W288F variant structure. In contrast Cu(II)-soaked W288A contains only one molecule in the asymmetric unit (residues 43-645) comprising of a disordered region between 200-205, again in the β-hairpin loop region of domain 2. A Cα superposition of the GlxA variants with the WT structure is illustrated in Figure 3.13 with RMSD values for all Cα atoms of 0.57 Å for Cu(II)-soaked W288A, 0.31 Å for W288A and 0.27 Å for W288F [168, 169].

The W288A structure reveals no electron density that could be assigned to a metal ion in the active site region (Figure 3.14.A). Furthermore, the electron density in the WT structure attributed to the Cys-Tyr crosslink is also absent indicating that the cross-linked cofactor is not formed in this variant. The absence of the stacking Trp^{288} creates space to enable the Tyr^{289} ring to rotate ~ 90^{o} away from its Cu coordinating and cross-linked orientation compared to the WT structure (Figure 3.14.B and D). The other active site ligands remain in positions posed to coordinate Cu (Figure 3.14.A and B). Upon soaking the W288A crystals with Cu(II) a structure was determined in which a clear 23 σ peak in the σ-weighted F_o-F_c difference map was observed and a Cu ion was modelled. However, based upon the two orientations identified for the side chains of Tyr^{289} and Cys^{121} the crystal appeared to contain a mixture of apo (30 %) and Cu(II) bound (70 %) forms (Figure 3.14). However, it is clear that in the presence of Cu(II) the cross-link is now formed. The Cu site geometry is maintained in the Cu(II)-soaked crystal with bond lengths compared with those of WT in Table 3.6.

The W288F structure displayed electron density within the active site consistent with the presence of a metal ion; therefore a Cu ion could be modelled (Figure 3.14.A and B). Similar to the Cu(II)-soaked W288A and WT GlxA the first coordination sphere residues in W288F are superimposable, however clear variations in bond lengths between the Cu ion and the coordinating residues exist with prominent changes between molecule A and molecule B (Table 3.6). In particular significantly longer bond lengths can be observed for His^{502} and His^{589} to the Cu ion compared to W288A and WT (Table 3.6). Notably, Tyr^{289} appears to be flipped towards Phe^{288} by ~ 20^{o} but not as drastically as is seen in W288A (~ 90^{o}) (Figure 3.14). When compared to WT GlxA it is clear that Phe^{288} does not stack directly over the Cys-Tyr cross-link therefore minimising any possible stacking interactions but allowing some rigidity to the active site to remain (Figure 3.14.C and D).

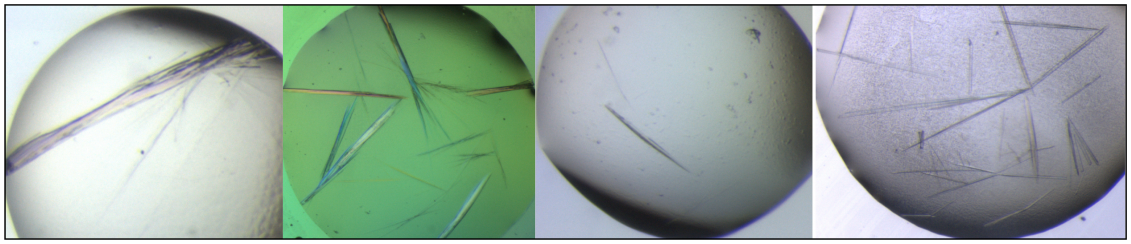


Figure 3.12: Crystallisation of Trp^{288} GlxA variants. A) W288A GlxA crystal (15 mg ml^{-1}) in a 96 well drop containing 20 % PEG 20 K, 0.1 M sodium acetate, pH 4.6. B) W288A GlxA crystals (15 mg ml^{-1}) in a 24 well drop containing 20 % PEG 20 K, 0.1 M sodium acetate, pH 4. C) W288F GlxA crystal (15 mg ml^{-1}) in a 96 well drop containing 20 % PEG 20 K, 0.1 M sodium acetate, pH 4.6. D) W288F GlxA crystals (15 mg ml^{-1}) in a 24 well drop containing 20 % PEG 20 K, 0.1 M sodium acetate, pH 4.

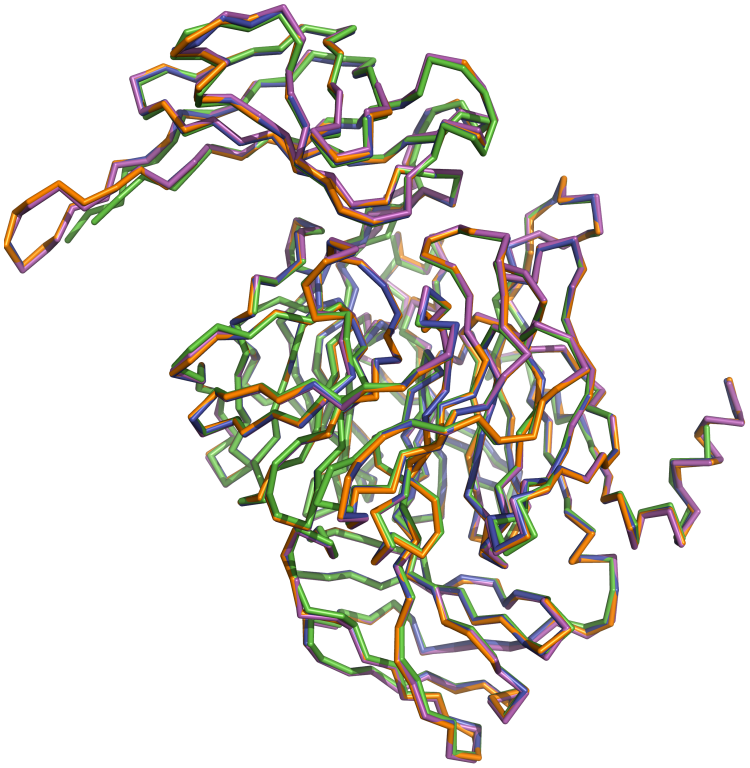


Figure 3.13: Superposition of X-ray crystal structures of WT GlxA (blue), W288A (purple), Cu(II)-soaked W288A (green) and W288F (orange) displayed as ribbons. Image created using PyMol.

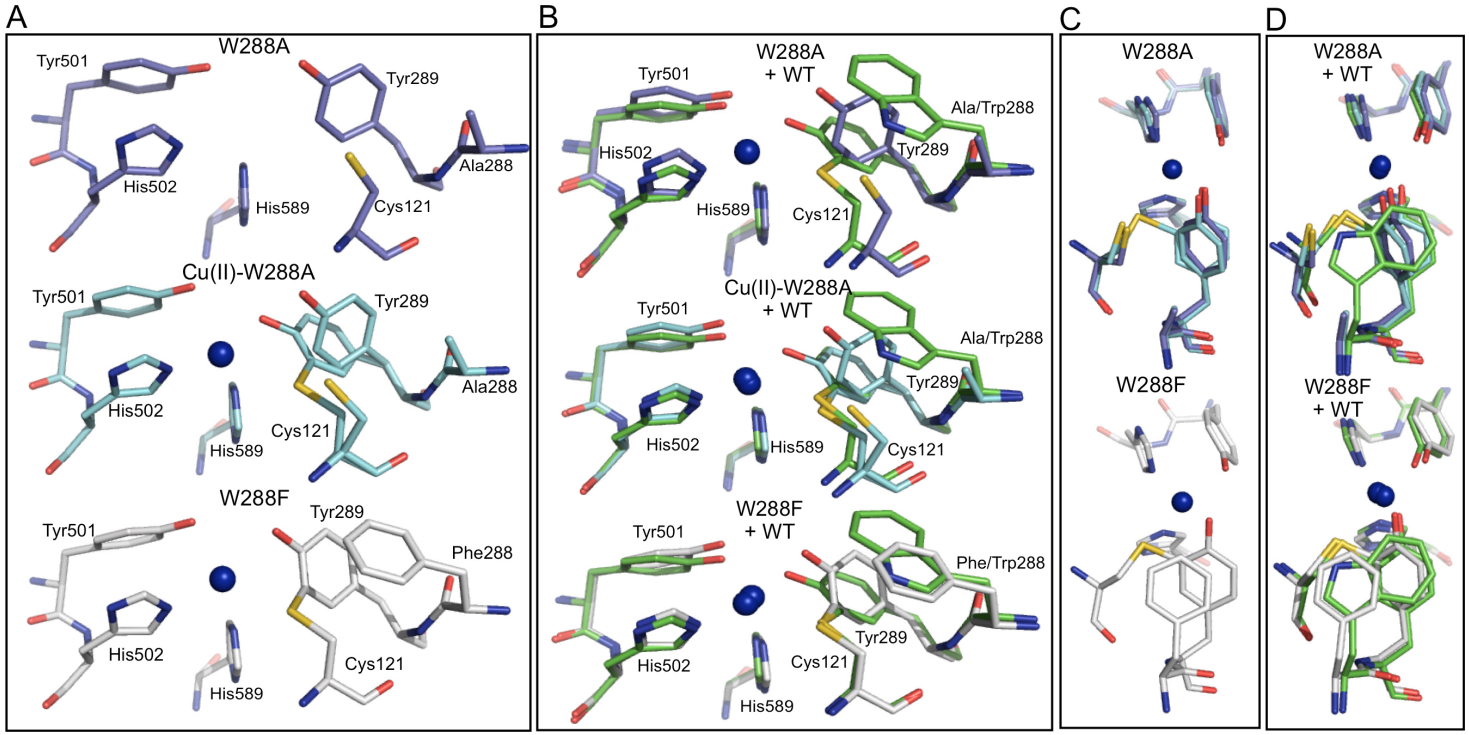


Figure 3.14: Comparison of the Cu sites of W288A and W288F GlxA variants. A) Cu sites of W288A (purple), Cu(II)-soaked W288A (blue) and W288F (silver) showing the orientation of the Cu active site residues as depicted. B) Cu sites of W288A (purple), Cu(II)-soaked W288A (blue) and W288F (silver) compared to WT GlxA (green) (PDB entry 4UNM). C) An above view of the Cu site for apo/Cu(II)-soaked W288A (purple and blue) and W288F (silver). D) An above view of the Cu site for apo/Cu(II)-soaked W288A (purple and blue) and W288F (silver) compared to WT GlxA (green) (PDB entry 4UNM). The Cu ion is shown as a blue sphere in all images. Waters are omitted from the images for clarity. Images were made using PyMol.

Table 3.6: A comparison of bond lengths in the Cu active site of WT GlxA, Cu(II)-soaked W288A and W288F variants. (A/B) indicate monomers A and B.

	WT-GlxA (A/B) Distance (Å)	Cu(II)-soaked W288A (A) Distance (Å)	W288F (A/B) Distance (Å)
Cu-Tyr^{289} OH	1.84/1.97	2.06 (3.99)*	1.82/2.69
Cu-Tyr^{501} OH	2.24/2.14	2.83	2.66/2.39
Cu-His^{502} Nε2	2.13/2.18	2.03	2.41/2.45
Cu-His^{589} Nε2	2.18/2.23	2.07	2.75/2.49
Cu-H_2O	2.44/2.50	2.14	2.76
Cys^{121}-Tyr^{289}	1.92/1.97	2.15 (3.28)*	2.18/ 2.28

* Brackets indicate distances from the Cu ion to the uncoordinated Tyr and Cys residues in the 30 % apo-structure of Cu(II)-soaked W288A.

3.3.11. Western blotting to assess GlxA maturation

GlxA functionality requires the formation of a Cys-Tyr cross-link and the incorporation of a Cu ion [166]. To assess whether this maturation affects GlxA mobility, purified apo, Cu(II)(holo)-GlxA and the C121G variant were run on an SDS-PAGE gel and migration patterns detected by GlxA polyclonal antibodies (Figure 3.15). It is apparent that apo-GlxA and the C121G variant, in which a Gly replaces the cross-linking Cys residue, migrate slower than the holo-GlxA form (Figure 3.15). This indicates that the Cys-Tyr cross-link is absent in both the apo-GlxA and the C121G variant accounting for this retardation of electrophoretic mobility (Figure 3.15).

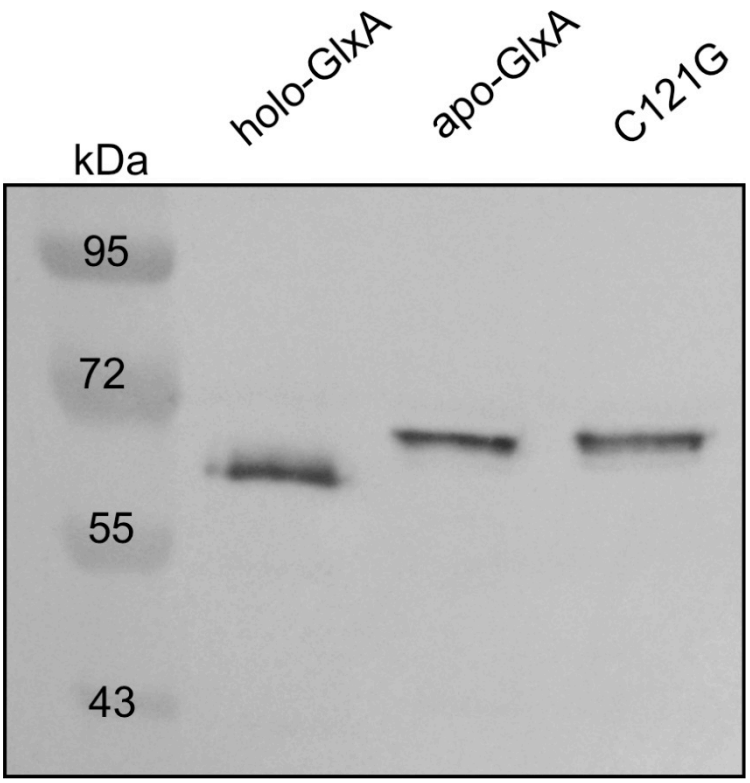


Figure 3.15: Western blot analysis of GlxA maturation. GlxA is over-expressed and purified in either Cu-replete conditions (Cu(II)-holo-GlxA) or Cu-starved conditions (apo-GlxA). The C121G variant is unable to form the Cys-Tyr cross-link and migrates together with apo-GlxA. Molecular weight markers are indicated in kDa in the far left lane.

3.3.12. Western blotting of W288A and W288F GlxA maturation

Maturation of the Cys-Tyr cross-link in GlxA was monitored over time for the W288A and W288F variants. Both proteins were expressed and purified in essentially Cu-starved conditions and the uptake of Cu(II) could be monitored using UV-visible and stopped-flow kinetic spectroscopy (*vide supra*). Thus, samples of purified apo-Trp^{288} variants were taken and the maturation process monitored over a period of time after 1-equivalent of Cu(II) was loaded. Samples were run on an SDS-PAGE gel and migration patterns detected by GlxA polyclonal antibodies. The Trp^{288} variants were compared to controls of holo-GlxA and the C121G variant, known to form and to not form the Cys-Tyr cross-link, respectively (Figure 3.16). W288A displays a single band, which migrates slower than WT-holo-GlxA and is in line with the C121G variant confirming the absence of the Cys-Tyr cross-link (Figure 3.16.A). This is also seen in apo-W288F, with the addition of a faint band migrating in line with holo-GlxA showing a small amount of cross-linked W288F protein (Figure 3.16.B). Over time without any addition of Cu(II) the apo band for W288A and W288F remains relatively unchanged. However, upon addition of Cu(II) to W288A and W288F two clear bands become apparent, one consistent with the formation of the Cys-Tyr cross-link and the other the pre-mature form (Figure 3.16). Over time the lower band corresponding to the mature protein appears to become thicker than the slower migrating band. However, after ~ 72 h the higher band still remains illustrating that a mixture of the mature and pre-mature forms persist (Figure 3.16). This may be correlated with the Cu(II)-soaked W288A X-ray crystal structure, which reveals a mixture of formed and unformed Cys-Tyr cross-link (Figure 3.14).

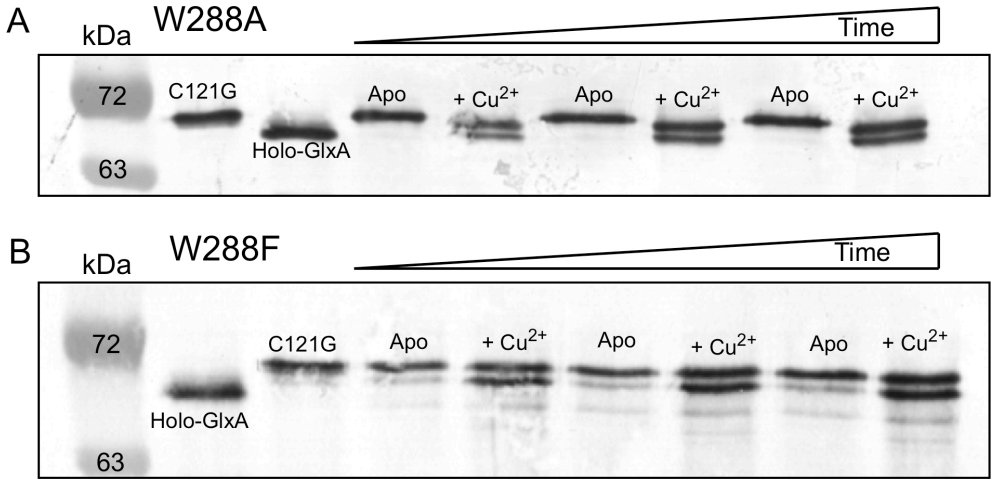


Figure 3.16: Western blot analysis of W288A and W288F GlxA maturation. Trp^{288} variants W288A (A) and W288F (B) were over-expressed and purified in Cu-starved conditions (apo) then incubated with 1 equivalent of $Cu(II)SO_4$ and samples taken after ~ 1 h, ~ 10 h and 72 h. They were compared to WT GlxA over-expressed and purified in Cu-replete conditions (holo-GlxA) and to the C121G variant, which is unable to form the Cys-Tyr cross-link. Molecular weight markers are indicated in kDa.

3.4. Discussion

3.4.1. Spectroscopic properties of the GlxA variants are distinct from WT

The large variation in spectroscopic properties of the four GlxA variants compared to the WT GlxA signifies the importance of Cys^{121} and Trp^{288} in GlxA. Replacement of these residues clearly alters the electronic transitions of the Cu in the active site and results in distinct differences in their UV-visible absorbance spectrum (Figure 3.6). Interestingly, unlike WT GlxA the variants express and purify largely free from Cu, and are all capable of binding Cu(II) readily and once bound display absorption spectra distinct from WT GlxA (Figure 3.6). On the contrary the GlxA variants all display spectra more comparable to the Cu(II) Gox spectrum [81, 83]. The UV-visible spectra of the Cu(II)-Trp^{290} Gox variants W290H, W290F and W290G share some similarities to WT Gox in their fully oxidised and semi-reduced states, especially the W290F variant, however larger shifts in the absorption peaks are noted for W290G and W290H [123]. The Cu(II)-C228G Gox variant also displays similarities to the WT but with a shift in the peak maxima to 476 nm and no recognisable 800 nm band [122]. Loss of the 800 nm band in the C228G variant of Gox may reflect protonation (and dissociation from Cu) of Tyr^{495} [83]. Therefore, the UV-visible spectra of the GlxA variants display more significant changes compared to WT GlxA than observed with similar variants in Gox.

The absorption band between 400-500 nm is assigned to LMCT of the axial Tyr^{495} to Cu(II) metal ion and π - π^* transitions of Tyr^{272} in Gox [122]. Therefore, the shift from λ_{max} 577 nm in WT GlxA to ~ 480 nm in the GlxA variants suggests a change in the transitions of the Tyr^{501} residue to Cu(II) and π - π^* transitions of Tyr^{289} in GlxA (Figure 3.6). This transitional change shown in WT GlxA compared to Gox is therefore influenced by the stacking Trp^{288} residue and the thioether bond Cys^{121}-Tyr^{289} [122]. Table 3.7 displays a comparison of the bond lengths from the Cu ion to the coordinating Tyr residues in WT GlxA, Cu(II)-soaked W288A, W288F and Gox, clearly showing that the largest difference noted is from the Cu centre to the axial Tyr (Tyr^{501} in GlxA and Tyr^{495} in Gox). Strikingly, the bond length for Cu-Tyr^{501} in WT-GlxA is ~ 2.2 Å, which is significantly shorter than observed in W288A (~ 2.8 Å), W288F (~ 2.5 Å) and Gox (~ 2.7 Å) (Table 3.7). This decreased bond length in WT GlxA could account for the differences observed in the UV-visible absorption spectra and it could be postulated to be the reason why the spectra of the GlxA variants appear more comparable to Gox. Studies using azide binding on the Trp^{290} Gox variants illustrates that the stacking Trp modulates the function of the catalytic base (Tyr^{495} in Gox) and this could therefore also be the case for Trp^{288} in GlxA [123]. Removal of Trp^{288} or Cys^{121} disrupts the electronic transitions unique to GlxA, thus these two residues must contribute significantly to the differences observed in the WT protein.

Upon addition of an oxidising agent to Gox an intense green colour has been reported with distinct spectral transitions [81, 83]. The optical spectra of oxidised Gox variants W290F, W290G and W290H are comparable to WT-Gox but W290G and W290H display much smaller extinction coefficients and are highly unstable [123]. Even though the Cu(II) spectrum of the GlxA variants share similarities to Gox no observed change in the spectra or colour could be observed after the addition of an oxidising agent. This is also seen in the WT protein and is likely to be a consequence of the unique tertiary structural fold which gives rise to a more insulated Cu site.

Table 3.7: Comparison of bond lengths of the Cu sites (monomer A and B) of WT, Cu(II)-soaked W288A, W288A GlxA from *S. livdians* and Gox from *F. graminearum* (PDB 1GOF) [76].

	WT (A/B) (Å)	W288A (Å)	W288F (A/B) (Å)		Gox (Å)
Cu-Tyr^{289}	1.84/1.97	2.06	1.82/2.69	Cu-Tyr^{272}	1.93
Cu-Tyr^{501}	2.24/2.14	2.83	2.66/2.39	Cu-Tyr^{495}	2.69

Interestingly, the visible CD spectra of the GlxA variants are also more comparable to Gox and again display distinct differences to WT GlxA (Figure 3.6.C). WT GlxA displays CD spectra with minima at 350 nm and maxima at ~ 320 and 650 nm, which compares to WT Gox and its Trp^{290} variants, W290G, W290F and W290H which all produce a prominent CD band at 625 nm and positive peaks at ~314 - 325 nm [83, 123]. Similarly, the GlxA variants (W288A, W288F, C121S, C121G) show a prominent CD bands at ~ 600 nm and ~315 nm (Figure 3.6.C). Therefore, the CD spectra of the GlxA variants again share features similar to Gox and clear differences to WT GlxA. The band at 600 nm has been assigned in Gox to the $^2B_1$ to $^2B_2$ ($d_{x^2-y^2}$ to d_{xy}) transition assuming square pyramidal C_{4V} Cu(II) geometry [83, 123]. This band is present in the GlxA variants but not in the WT, thus the active site geometry in the WT must be altered in order to perturb this electronic transition. The positive peak between ~ 314 - 325 nm in Cu(II)-Gox has been assigned as an N(π) imidazole to Cu(II) ligand-to-metal charge transfer (LMCT) arising from the His ligands. This CD feature (~ 314 - 325 nm) can also be seen in WT GlxA and the four GlxA variants, hence the His Cu(II) ligations are not affected. In contrast to GlxA, the Trp^{290} variants in Gox display very similar CD spectra to WT Gox, illustrating that the second coordination shell Trp^{290} residue does not perturb the Cu(II) centre in Gox as would appear to be the case in GlxA [123].

3.4.2. Cu(II) binding rates differ in the four GlxA variants

The kinetics of Cu(II) binding to the GlxA variants can be described by Scheme 1. In all cases an initial binding event occurs with a [Cu(II)] dependence to give second order rate constants (k_1) all of similar magnitude (Table 3.4). The second step is a first order process and is assigned to a rearrangement step to yield the final Cu(II)-complex, which is distinctly different in rate (k_2) depending upon the type of residue replacing the Cys^{121} and Trp^{288}. For the smaller amino acid substitutions *i.e.* C121G and W288A, k_2 values could not be determined using stopped-flow kinetics. Instead the slow formation of the 480 nm band were monitored following a stoichiometric Cu(II) addition. In contrast, C121S and W288F, contain larger more comparable amino acid substitutions and display much faster rate constants for the formation of the final Cu(II)-complex, enabling k_2 values to be determined from stopped-flow kinetics (Table 3.4). The most feasible explanation for these differences is that C121S and W288F hold the Cu active site of GlxA in a more rigid conformation allowing the final Cu site geometry to be reached more easily. While the GlxA variants, C121G and W228A create more flexibility in the Cu site and thus yield a slower, less well defined, rearrangement of the site into its final conformation.

3.4.3. C121G displays a distinct Cu(II) EPR spectrum

EPR spectra of the WT as isolated GlxA was measured and assumed to display axial symmetry with a $g_{\perp}$ of 2.050 and $g_{\parallel}$ of 2.182 as discussed in Chapter 2 (Figure 2.8). To simulate the EPR spectra a combination of two simulations (Sim1 and Sim2, Table 3.5) was required. The simulations shown in Figure 3.9 clearly show that WT, C121S, W288A and W288F all display similar Cu(II) EPR spectra with simulation parameters reported in Table 3.5. It is clearly apparent that they display rhombic g-factor anisotropy ($g_x \neq g_y \neq g_z$) as opposed to the previously presumed axial symmetry ($g_x = g_y < g_z$) for WT GlxA (see Chapter 2). C121G is the only variant which displays a significantly different EPR spectrum with a well-resolved hyperfine structure in the $g_{\parallel}$ area. This variant was simulated by axial g-factor anisotropy $g_x = g_y$ in both Sim1 and Sim2 (Figure 3.9 and Table 3.5). Therefore, C121G must possess a different Cu(II) active site environment compared to WT, C121S, W288A and W288F. The most plausible explanation for the axial symmetry observed in the EPR spectrum of C121G must be due to a lack of perturbation in the location of the unpaired electron. In C121G, a Gly residue, the smallest amino acid residue consisting of a hydrogen substituent as its side chain, replaces Cys^{121}. Therefore, in this variant lack of the side chain allows for minimal disturbance in the surroundings of the unpaired electron. Once a larger side group is substituted, such as Ser in C121S, the symmetry of the environment of the unpaired electron becomes shifted to display rhombic g-factor anisotropy (Figure 3.9 and Table 3.5). Similarly, this is observed in WT, W288A and W288F illustrating that substitution of the second coordination sphere stacking Trp residue does not effect the

symmetry of the Cu(II) EPR spectra. Consequently, C121G displays two equal axes in the XYZ reference frame as opposed to all axes being different in WT and the other GlxA variants.

3.4.4. Trp^{288} and Cys^{121} control the unique radical behaviour of GlxA

Substitution of Trp^{288} or Cys^{121} clearly effects radical formation. The W288A and W288F variants abolish the radical species identified in the WT protein observed through EPR spectroscopy and consistent with this generate no catalytic turnover with glycolaldehyde. Therefore, the stacking Trp in GlxA clearly modulates the radical identified in GlxA. Similar observations have been reported for Trp^{290} variants in Gox, with Trp^{290} shown to control the properties and reactivity of the Cys-Tyr radical. W290G and W290H display elevated redox potentials and are highly unstable compared to the radical in W290F, which displays similar properties to the WT Gox radical [123]. W290G and W290H required a 12-fold excess of cesium octacyanomolybdate ($E^{0\prime}$ 892 mV) in order to form a radical [123], whereas even after the addition of up to > 50-fold excess $[Ir(Cl)_6]^{3-}$ ($E^{0\prime}$ 870 mV) no radical was observed in W288A and W288F GlxA variants. In Gox W290F the radical behavior is similar to WT, differing to W288F in GlxA. Strikingly, both C121S and C121G after the addition of $[Fe(CN)_6]^{3-}$ display a unique radical species (Figure 3.10). The radical species identified is unusual because removal of Cys^{121} eradicates the ability of the Cys-Tyr cross-link to form, which is thought to be essential for lowering the redox potential of the co-factor and establishing activity of the enzyme. Therefore, removal of Cys^{121} allows migration of the radical species. The most probable location of this radical species is on the stacking Trp^{288} residue, although simulations of the spectra need to be determined in order to elucidate the residue housing this unusual radical species.

3.4.5. C121G is active on glycolaldehyde

WT GlxA has the highest activity for the substrate, glycolaldehyde that is the smallest molecule to contain an aldehyde and alcohol group. Strikingly, among the four GlxA variants studied here, C121G is the only one that produced catalytic turnover of glycolaldehyde (Figure 3.11). V_{max}/K_m rates of 1.22 M^{-1} s^{-1} and 0.96 M^{-1} s^{-1} were determined for WT and C121G, respectively, presenting clear similarities. However, C121G displays a smaller K_m value of 67 mM compared to 115 mM of the WT protein, illustrating that C121G possess a higher binding affinity for glycolaldehyde. Interestingly, the k_{cat} of C121G is lower than for the WT with a value of 0.08 s^{-1} compared to 0.14 s^{-1}. This therefore demonstrates that although the affinity of C121G for glycolaldehyde is higher the turnover efficiency is considerably lower than for the WT enzyme (Figure 3.11). It is not implausible to observe activity of C121G with glycolaldehyde, because although Cys^{121} is no longer available to

form the Cys-Tyr cross-link, a different radical species is detected through EPR spectroscopy (Figure 3.10). Interestingly, C121S also displays a similar radical species identified through EPR spectroscopy after treatment with $[Fe(CN)_6]^{3-}$ (Figure 3.10). Nevertheless, no activity could be observed for C121S with any of the substrates verified with WT GlxA. Thus, substitution of Cys^{121} to a Gly residue rather than a Ser allows GlxA activity to remain. The reason for this difference in activity between the two Cys variants may be due to C121G allowing more flexibility in the active site enabling accommodation of the substrate molecule. This change in Cu site geometry has been illustrated by the EPR spectroscopy and Cu(II) binding studies. Cu(II)-C121G EPR produces significantly different hyperfine structures with axial g-factor anisotropy compared to rhombic symmetry in the WT and other GlxA variants and Cu(II) binding to C121G is significantly slower than for C121S (Figure 3.9 and Table 3.4). A radical species for the C228G Gox variant has not been reported and a 10,000 times lower activity observed with D-galactose compared to WT Gox [122]. Therefore, the ability of C121G to form a radical and turnover glycolaldehyde is a unique property. No activity could be identified for either Trp^{288} variants W288A and W288F, both lacking a radical species upon addition of $[Fe(CN)_6]^{3-}$ ($E^{0\prime}$ 424 mV) or $[Ir(Cl)_6]^{3-}$ ($E^{0\prime}$ 870 mV). In contrast, radicals were observed for the stacking Trp^{290} variants in Gox with W290F oxidized by $[Fe(CN)_6]^{3-}$ ($E^{0\prime}$ 424 mV) and W290H and W290G oxidized by octacyanomolybdate ($E^{0\prime}$ 892 mV) [123]. Therefore, these Gox variants still displayed activity with D-galactose but with significantly lower activity, with W290H ~ 1000-fold lower than WT Gox and W290F and W290G linearly dependent on substrate concentration with ~ 50-fold and ~ 1000-fold decrease in activities, respectively [123].

3.4.6. Cu sites of W288A and W288F differ to Gox variants

The X-ray crystal structures of W288A, Cu(II)-soaked W288A and W288F illustrate further the properties of the active site in GlxA. The active site of W288A is free from any Cu metal ion and can therefore be directly compared to pre-mature Gox (Figure 3.17.A) (PDB entry 2VZ1) [87]. In pre-mature Gox, Tyr^{272} is not well ordered with two conformations observed, one similar to the Cu(II) loaded Gox and the other displaying the apo pre-mature form. Similarly, Trp^{290} in pre-mature Gox is not present in a mature stacking position with residues 290-293 not modelled into the final structure (Figure 3.17.A) [87]. Regardless of these unclear orientations, W288A GlxA and pre-mature Gox clearly show similarities in their pre-formed active sites (Figure 3.17.A). The largest change occurs in the Tyr and Cys residues forming the thioether cross-link. The Tyr (272 in Gox and 289 in GlxA) is flipped 90^{o} in relation to the Tyr position when bound to Cu(II), the Cys (228 in Gox and 121 in GlxA) is also located away from the Tyr preventing cross-link formation (Figure 3.17.A) [87]. The poorly resolved Trp^{290} in pre-mature Gox could be due to a large movement in this second coordination sphere residue upon Cu(II) binding, locking the Cu site

in place, thus movement of this residue must allow for the 90^{o} flip in the coordinating Tyr (272 Gox and 289 in GlxA) as also seen in the W288A GlxA variant.

Upon soaking W288A crystals with Cu(II) a mixture of apo and Cu(II) loaded W288A could be observed (Figure 3.17.B). This mixture of Cu(II) and apo forms could be due to the slow uptake of Cu(II) in this variant or that some of this enzyme is no longer capable of Cu(II) uptake as suggested by Western blotting (Figure 3.15). All bond lengths are shorter in the Cu(II) bound structure, and the most noticeable changes are again in the Tyr and Cys cross-linking residues (Table 3.6). Two conformations of Tyr^{289} and Cys^{121} could be modelled into the electron density effectively illustrating the flipping motion of these residues from the apo structure to the mature protein conformation. Cu(II)-soaked W288A GlxA can be directly compared to the W290G Gox variant showing noticeable similarities in the active sites but also distinct differences predominately in the location of the Cys-Tyr cross-link (Figure 3.17.B) [123]. The X-ray crystal structure of the W288F GlxA variant displays full Cu occupancy with complete formation of the Cys-Tyr cross-link (Figure 3.17.C). Full Cu(II) occupancy of W288F compared to only partial Cu(II) occupancy in W288A, must be explained by the incorporation of Cu(II) during growth of W288F locking the final conformation into place. W288F could be compared to the structures of W290H and W290F variants of Gox (Figure 3.17.C). Noticeable differences can be observed in the two His residues with them appearing shifted compared to the Gox variants. These changes in bond lengths must be due to replacement of Trp^{288} for a Phe in GlxA affecting the Cu site geometry allowing movement in these coordinating residues. Substitution of Trp^{288} in GlxA to a Phe residue does not have the same affect as in Gox, and is largely due to residue 288 being adjacent to the Tyr^{289} in GlxA, unlike in Gox where 290 is not neighbouring Tyr^{272}. Phe^{288} therefore does not stack directly on top of the Cys-Tyr cross-link as clearly seen in W290F (Figure 3.17.D). As a consequence of this, Tyr^{289} in GlxA is rotated vertically by ~ 20^{o}. Stacking of these residues can be observed in Figure 3.17.C and D showing that in W290F of Gox the Phe residue is clearly stacking over the Cys-Tyr cross-link and in W290H the His is directly stacking over the cross-linking Tyr residue (Figure 3.17.E). This change in the stacking properties observed in the W288F GlxA variant must be the reason why no radical could be detected as opposed to W290F in Gox, which displays a similar radical to WT Gox.

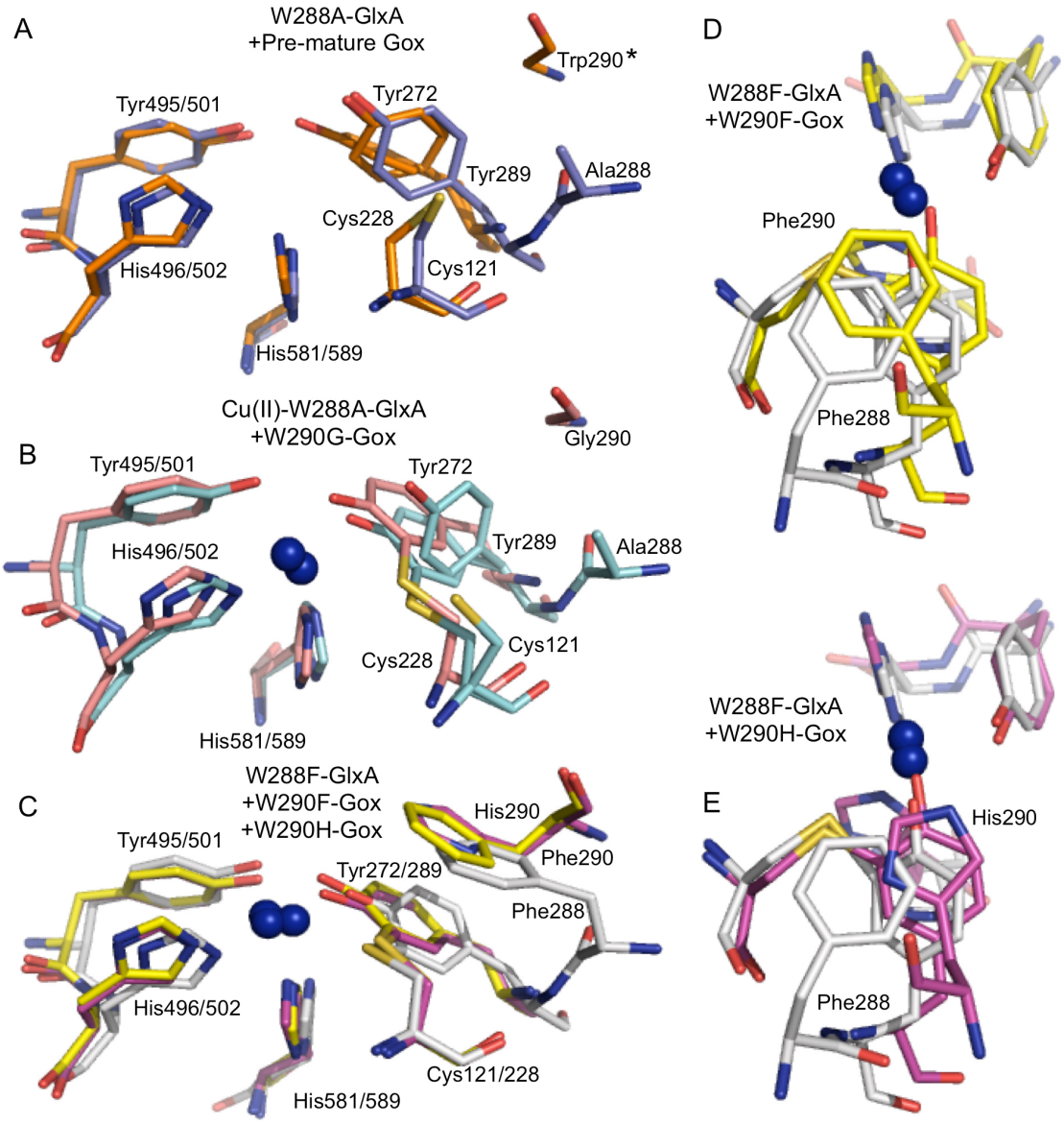


Figure 3.17: Comparison of the Cu sites of the GlxA variants W288A and W288F with Gox. A) Cu sites of W288A-GlxA (purple) overlaid with pre-mature Gox (orange) (PDB entry 2VZ1). The * symbol for Trp^{290} is illustrating that this residue could not be modelled into the final structure of pre-mature Gox. B) Cu sites of Cu(II)-soaked W288A-GlxA (blue) compared to W290G-Gox (salmon) (PDB entry 2EID). C) Cu site of W288F-GlxA (silver) compared to W290F-Gox (Yellow) (PDB entry 2EIC) and W290H-Gox (pink) (PDB entry 2EIB). D) Cu sites of W288F GlxA (silver) and W290F-Gox (Yellow) (PDB entry 2EIC) from an above view. E) Cu sites of W288F GlxA (silver) and W290H-Gox (pink) (PDB entry 2EIB) from an above view. The Cu ion is shown as a blue sphere in all images. Waters were omitted from the images for clarity. Images were made using PyMol.

3.4.7. GlxA maturation displays similarities to Gox

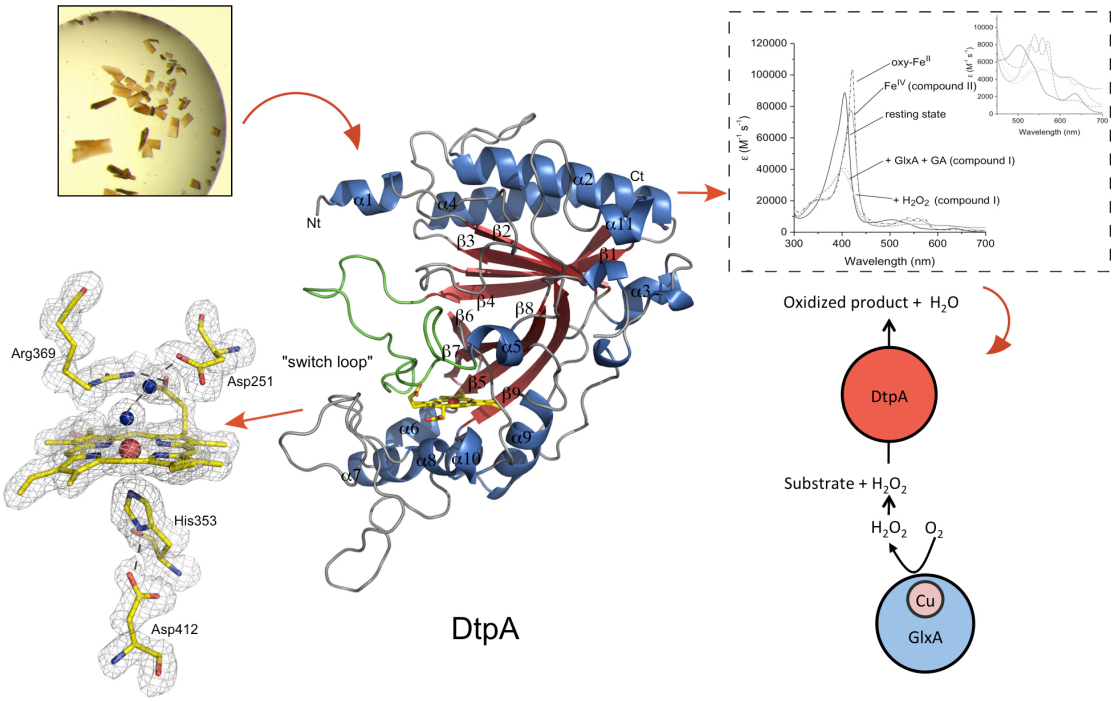
Complete processing of Gox requires at least four post-translational modification events; cleavage of the secretion signal, removal of an N-terminal pro-sequence, formation of the Cys^{228}-Tyr^{272} thioether bond and one electron oxidation resulting in a Cys-Tyr radical [150, 170]. Formation of the Cys-Tyr cross-link can be monitored by SDS-PAGE, with the pre-mature form lacking the Cys-Tyr cross-link migrating slower than the fully mature protein [87, 150, 170]. Upon formation of the Cys-Tyr bond a stable loop is produced, which prevents the protein from fully unfolding upon treatment with SDS and thus allowing faster migration [122]. Maturation of GlxA from *S. lividans* was also monitored through different migration patterns on SDS-PAGE and through detection with GlxA specific antibodies. Both apo-GlxA and the C121G variant, migrate slower than holo-GlxA, indicating that the Cys-Tyr cross-link is absent in both apo-GlxA and in the variant (Figure 3.15). An identical electrophoretic mobility pattern has been observed for pre-mature Gox and the C228G variant [122, 150]. Maturation of GlxA was monitored over time using the W288A and W288F variants (Figure 3.16). The apo variants migrate at the same height as C121G indicating the absence of the Cys-Tyr cross-link, whereas after incubation with Cu(II) a lower band assigned to cross-link formation becomes more prominent over time. However, unlike Gox maturation, the Trp^{288} GlxA variants never display a single band for the fully mature form following incubation with Cu(II). In relation to this the crystal structure of the Cu(II)-soaked W288A reveals incomplete processing to the Cys-Tyr bond (Figure 3.14). An explanation for this is not readily forthcoming, but it is notable that this is also the case for the apo-WT GlxA following incubation with Cu(II) and therefore may not be variant specific (not shown here). From *in vivo* studies using GlxA antibodies to detect maturation from mycelial extracts it is also noted that the complete processing of the Cys-Tyr cross-link to give the fully mature enzyme is not observed [171]. It may be that during expression of the apo-forms, some of the GlxA may mis-fold into a conformation that does not allow access to Cu(II) and therefore does not allow for Cys-Tyr thioether bond formation to occur

3.4.8. Conclusions

This chapter brings to attention two important findings. The first is that the second coordination sphere residue, Trp^{288}, plays a major role in tuning the electronic properties of the buried Cu site in GlxA. Remarkably, its removal abolishes the detection of the protein radical and perturbs the WT spectroscopic properties to resemble the more solvent accessible Cu site in Gox. However, perhaps most surprisingly is the finding that the Cys-Tyr crosslink is not required to form a protein based radical in GlxA. This is a highly unexpected discovery and although the identification of the new radical in the Cys variants remains elusive, a prominent role for Trp^{288} seems likely. Further work into this is required.

Chapter Four

Structural and kinetic characterisation of the *Streptomyces lividans* dye-decolourising peroxidase, DtpA, and its synergy with GlxA



Synopsis: DtpA is a dye-decolourising-type peroxidase from *Streptomyces lividans* that functions as a peroxidase and can do so in the presence of GlxA.

Some of the results from this Chapter have been published in: M.L.C. Petrus, E. Vijgenboom, A.K. Chaplin, J.A.R. Worrall, G.P. van Wezel, D. Claessen 2016 *Open Biology*, **6**

4.1. Introduction

The gram-positive soil dwelling bacterium *Streptomyces lividans* is distinctly dependent upon copper (Cu) to initiate a morphological switch between vegetative to aerial growth [24, 172]. Previous work has led to the discovery of a Cu-trafficking pathway involved in this Cu-dependent development. One of the proteins in this pathway, the Cu chaperone Sco, is required for morphogenesis under low Cu availability, *i.e.* homeostasis conditions [37, 38]. However, morphogenesis of the *sco* null-mutant is restored upon the addition of exogenous Cu [37]. Sco receives its Cu ion from the extracytoplasmic Cu chaperone ECuC [38] and in turn delivers Cu to the Cu_A site of an aa_3-type Cytochrome *c* oxidase (C*c*O) and to a second target, possibly the cuproenzyme GlxA. GlxA has been investigated and discussed in Chapter 2 where it was shown that the *glxA* null-mutant stalls aerial hyphae development on solid media and in contrast to the *sco* null-mutant cannot be rescued upon addition of exogenous Cu [166]. The *ΔglxA* also lacks the hyphal tip glycan and dramatically changes the morphology of *S. lividans* in liquid cultures [166].

GlxA is a new structural member of the radical Cu oxidase family and is capable of carrying out two-electron oxidation of primary alcohols to aldehydes with the reduction of dioxygen to H_2O_2 [166]. GlxA displays low activity with classical Gox substrates but is able to turnover glycolaldehyde, the smallest molecule to contain both an aldehyde and a hydroxyl group and produce H_2O_2, [166]. The gene for *sco* (*SLI4214*) and *ecuc* (*SLI4213*) are the first two genes of an operon that also harbours genes for a putative Cu-transport protein (*SLI4212*) and a secreted protein (*SLI4211*) with a putative Tat (twin arginine translocation) signal sequence [39] and a DyP (dye-decolourising peroxidase)-type domain (Figure 4.1) [25, 37, 40]. *SLI4211* encodes a protein hereinafter denoted DtpA for DyP-type peroxidase A, belonging to the A subclass of this haem peroxidase family due to sequence homology and the presence of a Tat signal sequence. It is therefore possible due to the genome proximity of DtpA, that a role in the Cu-dependent morphogenesis could be inferred. One possibility that is advanced in this Chapter is that DtpA may utilize the H_2O_2 produced by GlxA during the glycan substrate oxidation and whether this is required for executing the final enzymatic steps in the cascade of the GlxA-dependent morphological development of *S. lividans*.

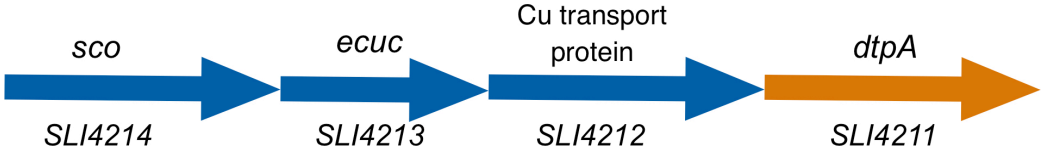


Figure 4.1: Genome environment of the *dtpA* locus in *S. lividans*. Gene annotations are as follows: *SLI4214* Sco a secreted Cu chaperone, *SLI4213* ECuC a secreted Cu chaperone, *SLI4212* a putative Cu transport protein and *SLI4211* (orange) DtpA a dye-decolourising peroxidase A type.

Peroxidases (Enzyme commission number (EC) 1.11.1.x)) encompass a large group of oxidoreductases that catalyse the oxidation of a substrate using hydrogen peroxide (H_2O_2) as the electron acceptor. The majority of peroxidases contain a b-type haem cofactor consisting of an iron (Fe) coordinated in the centre of a large heterocyclic organic ring called a protoprophyrin IX [173]. There are two principal peroxidase super-families: the plant peroxidases and the animal peroxidases. The plant peroxidase superfamily also contains evolutionary related haem containing peroxidases from fungi and bacteria, subsequently leading to a subdivision into three classes depending upon the enzymes function and cellular location [174, 175]. Class I plant peroxidases are intracellular enzymes and include yeast cytochrome *c* peroxidase (C*c*P) and the chloroplast ascorbate peroxidase (APX). Class II plant peroxidases are extracytoplasmic fungal peroxidases such as lignin peroxidase (LiP) and manganese peroxidase (MnP). Class III plant peroxidases includes extracytoplasmic peroxidases such as horseradish peroxidase (HRP). The plant peroxidases are not related through sequence to the animal peroxidases [176]. Recently, a unique superfamily of haem peroxidases has emerged not related by primary sequence, structure or substrate turnover to the plant or animal peroxidases. This new class of haem peroxidase superfamily is identified as a separate family within the PeroxiBase and Pfam databases [177, 178] and has been named as the dye-decolourising peroxidase (DyP-type) superfamily (EC 1.11.1.19) [41, 179-182].

DyPs were initially discovered in 1995 from the basidiomycetous fungus *Bjerkandera adusta* (*bad*DyP), which was misidentified as originating from *Geotrichum candidum* then from *Thanatephorus cucumeris* Dec1 (DyP_{DEC1}) (Table 4.1) [179]. *Bad*DyP was found to degrade a wide range of anthraquinone (AQ) dyes, hence the name of this new peroxidase superfamily [180]. EfeB (formerly YcdB) from *Escherichia coli* was the first recognized bacterial DyP-type peroxidase [183], with many more DyP-type peroxidases subsequently identified in bacterial genomes and characterised (Table 4.1). Phylogenetic analysis has led to DyPs being classified into four subfamilies (A-D) based on their primary sequence, with classes A, B and C including DyPs from bacteria belonging to the phyla *Actinobacteria*, and α-, β-, γ-, and δ-*proteobacteria*, while class D includes fungal DyPs [40, 184-187]. The percentage identity between subclasses is ~15 - 20 % with each having unique characteristic sequences. The A-type DyPs contain a Tat-dependent signal sequences and are therefore extracytoplasmic as opposed to classes B and C, which includes cytoplasmic enzymes suggesting these enzymes play a role in intracellular metabolic pathways [188] (Table 4.1).

Table 4.1: DyP-type peroxidases. DyPs can be subdivided into classes A-D. Individual protein name, the organism they originate from and the references associated are listed.

Subclass	Protein	Organism	Reference
A	*Tc*DyP	*Thermomonospora curvata*	[189]
	*Tfu*DYP	*Thermobifida fusca*	[174]
	DyPA	*Rhodococcus jostii RHA1*	[190]
	*Bs*DyP	*Bacillus subtilis*	[191]
	EfeB	*Escherichia coli*	[183]
B	DyPB	*Rhodococcus jostii RHA1*	[190]
	*Vc*DyP	*Vibrio cholerae*	[192]
	$DyPP_a$	*Pseudomonas aeruginosa*	[193]
	TyrA	*Shewanella oneidensis*	[194]
	*Mt*DyP	*Mycobacterium tuberculosis*	[195]
	*Pp*Dyp	*Pseudomonas putida*	[191]
C	DyP2	*Amycolatopsis sp.*	[196]
	AnaPX	*Anabaena sp.*	[184]
D	*Bad*DyP	*Bjerkandera adusta*	[180]
	AauDyP	*Auricularia auricula-judae*	[197]
	*Il*DyP	*Irpex lacteus*	[198]
	*Mep*DyP	*Mycena epipterygia*	[199]
	*Egl*DyP	*Exidia glandulosa*	[199]
	TAP	*Termitomyces albuminosus*	[200]

X-ray crystal structures of several DyP-type peroxidases have been reported including *Bad*DyP [186, 201, 202] and several from bacteria [194, 196, 203-205]. The structures of classical plant peroxidases such as HRP [206] contain primarily α-helices (Figure 4.2.A), which differ greatly to the overall fold of DyPs, rich in both α-helices and β-sheets. DyPs share structural homology with chlorite dismutases (Clds), which catalyse the disproportionation of chlorite into chloride and oxygen, with Clds, DyPs and EfeBs all belonging to a CDE structural superfamily [207]. DyPs include two domains from ancestral duplication, both with ferredoxin-like folds formed by a four-stranded anti-parallel β-sheet covered at one side by 2-3 α-helices and a haem molecule present in the C-terminal domain (Figure 4.2.B) [208]. The core CDE fold is the same across different species; however, *Bad*DyP contains a higher proportion of loops compared to EfeB [194]. The haem pocket architecture in DyPs is similar to other haem-peroxidases, with the equivalent proximal His seen in plants (His^{170} in HRP) also present in DyPs (His^{308} in *Bad*DyP) (Figure 4.2). This residue coordinates to the haem Fe in the fifth coordination position (Figure 4.2). It also forms a H-bond to a conserved acidic residue (Glu^{391} in *Bad*DyP) (not shown in Figure). The distal face of the haem differs in DyPs compared to HRP, with the catalytic His residue present in plant peroxidases (His^{42} in HRP) replaced by an Asp residue (Asp^{171} in *Bad*DyP) (Figure 4.2) [186, 194]. A D171N mutant of the *Bad*DyP enzyme identified the role of this Asp residue as the catalytic proton acceptor from H_2O_2 upon Compound I formation [186]. Similarly, in *Tfu*DyP from *T. fusca*, substitution of Asp^{242} to Ala inactivated the enzyme illustrating the role of this residue in the catalytic mechanism [174]. However, this Asp residue is not so important for B-type DyPs and some A-type DyPs [40, 186, 209]. Notably, DyPs are reported to contain higher activity in a more acidic environments which is therefore consistent with the acidic Asp residue ($pK_a = 3.9$) acting as the acid-base catalyst as opposed to a His residue ($pK_a = 6.0$) [186]. In plant peroxidases, an Arg residue is present within the haem distal area (Arg^{38} in HRP) and is reported to be essential for peroxidase activity. Poulos and Kraut [210] described that Arg^{48} present in C*c*P stabilises the negative charge on the oxygen leaving group through H-bonding upon heterolytic cleavage of the O-O bond [210]. Similarly, an Arg residue is present in DyPs (Arg^{329} in *Bad*DyP) but the role of this residue is still debatable (Figure 4.2). The Arg residue present in DyPB from *R. jostii* and *Pp*DyP from *P. putida* is crucial for enzymatic activity and postulated to play the role as the acid-base catalyst as opposed to the Asp residue [209]. However, in contrast when the Arg residue is replaced with a Lys in *Bad*DyP partial peroxidase activity is preserved [186, 191, 209-211]. Recently, Asp and Arg were found to be essential for activity of *Vc*DyP [192], however, they have been found to be important for H_2O_2 binding in *Bs*DyP but not essential for H_2O_2 (de)protonation or O-O bond cleavage [212]. Therefore, the roles of these residues are different in distinct DyP families.

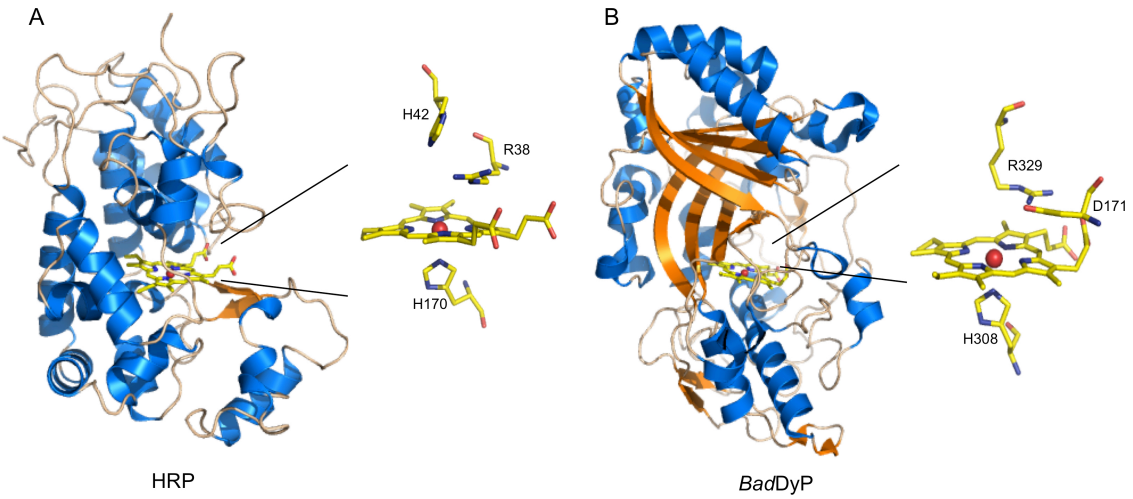


Figure 4.2: Comparison of the structures of HRP and DyP from *Bjerkandera adusta*. A) Overall cartoon structure of HRP with β-sheets shown in orange, α-helices in blue, loops in beige and the haem illustrated as yellow sticks with the Fe shown as a red sphere. Inset, haem environment of HRP, residues and the haem shown as yellow sticks with nitrogen atoms in blue, oxygen in red and the Fe shown as a red sphere (PDB accession number: 1ATJ). B) Structure of DyP from *B. adusta* with β-sheets shown in orange, α-helices in blue, loops in beige and the haem depicted as yellow sticks with the Fe shown as a red sphere. Inset, haem environment of *Bad*DyP, residues and the haem shown as yellow sticks with nitrogen atoms in blue, oxygen in red and the Fe shown as a red sphere (PDB accession number: 3AFV).

Azo, AQ and phthalocyanine-based synthetic dyes are used throughout the world with ~10^6 tons produced annually by textile, food and dye industries, of which 1-1.5 x 10^5 tons of this is released into the environment as waste [213]. These large dye molecules are extremely stable and consequently can remain within the environment for a prolonged period of time converting into harmful compounds creating significant problems to nature [214, 215]. With increased public concern and ecological awareness, interest has increased into various methods to degrade these recalcitrant dyes by decolourisation. Physiochemical processes such as coagulation, adsorption and oxidation using ozone are effective but require hazardous chemical additives, are expensive and create large volumes of sludge [215]. Microorganisms are capable of decolourising these synthetic dyes largely attributed to expression of enzymes including; laccases (Lac) [216-218], azoreductases [219] and

peroxidases such as LiP [220-222], MnP [223, 224] and versatile peroxidases (hybrid of MnP and LiP) [225]. LiP acts by oxidizing the non-phenolic unit of lignin to produce a cation radical on the aromatic ring, whereas MnP and Lac withdraw one electron from the phenol moiety generating a phenoxy radical and are therefore able to breakdown large dye molecules [226, 227]. Despite their activity, microorganisms still degrade the dyes with poor efficiency and generate toxic by-products. Therefore, the use of pure enzymes instead creates several advantages including; enhanced stability, activity, and lower costs [184]. In addition to AQ dyes such as reactive blue 19 (RB19), DyPs have also shown activity with carotenoids, methoxylated aromatics such as veratryl alcohol (VA), lignin model compounds and more typical peroxidase substrates such as 2,2-azinobis (3-ethylbenthiazoline-6-sulfonic acid) (ABTS) [41, 174, 228, 229]. The different DyP subfamilies have remarkably different specificities for the vast range of reductive substrates. This therefore illustrates the potential use of DyPs in the bioremediation of synthetic dyes.

The peroxidative cycle of DyPs are suggested to possess a catalytic mechanism similar to plant peroxidases [186, 211]. The resting state ferric (Fe^{3+}) enzyme reacts with H_2O_2 and produces Compound I [Fe^{IV}=O Por•]$^{+}$, which is two reducing equivalents more oxidized than the ferric enzyme (Figure 4.3). Compound I then reacts with one equivalent of a reducing substrate, to form Compound II (Fe^{IV}=O) that upon addition of a second reducing substrate results in ferric resting state enzyme (Figure 4.3) [204]. Addition of another electron can reduce the enzyme to the ferrous (Fe^{2+}) state (Figure 4.3). Notably, some differences between DyPs and plant peroxidases are apparent, including the incredibly stable Compound I state of DyP with a half-life of ~9 min reported for *Bad*DyP, which then slowly decays to the ferric enzyme with no detectable Compound II intermediate [211]. The physiological function of all DyP-type peroxidases is still unclear, although many functions have been suggested. Due to their peroxidase function, DyPs have been proposed to play a role in the bacterial oxidative stress response due to their co-induction with gas vesicles (scavenge O_2) under stress conditions [230]. They have also been shown to cause seedlings to wilt and are toxic to plant cells, so they have also been proposed to act as virulence factors in plant pathogens [231]. Recently, the *E. coli* DyPs, YfeX and EfeB, have been shown to capture Fe from haem, while still preserving the tetrapyrrole ring [232]. Interestingly, some DyPs are also located within shells formed by the shell-forming protein encapsulin, which is thought to compartmentalize some bacterial ferritin-like proteins [233]. In light of these discoveries it is likely that different subfamilies of DyPs possess different physiological roles, with more extensive research needed to elucidate the specific role of this new peroxidase family.

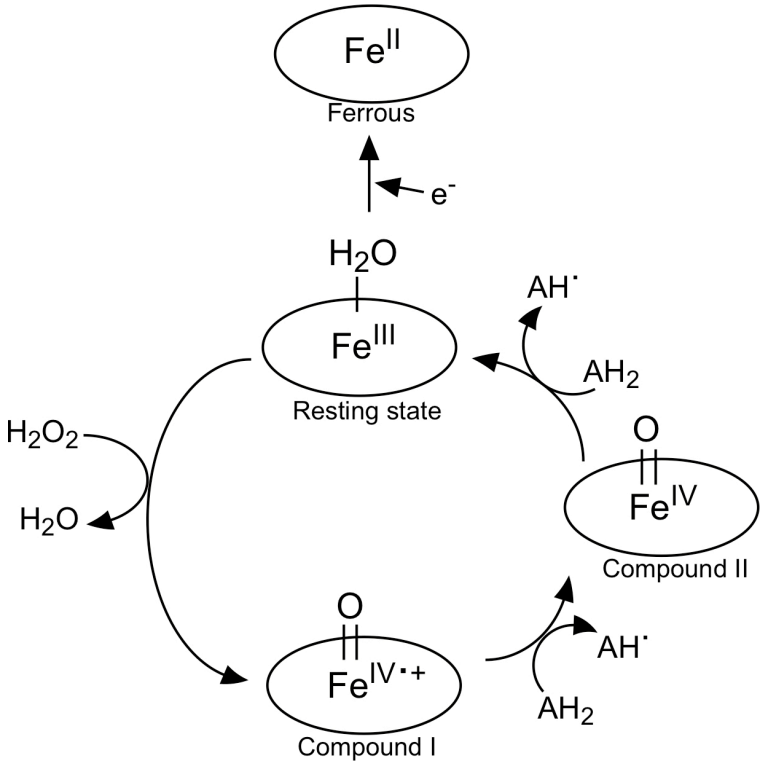


Figure 4.3: Peroxidase cycle of DyP-type peroxidases. AH_2 and AH represent the substrate and its oxidized product, respectively. The oval shape surrounding the Fe represents protoporphyrin IX.

In this Chapter the putative DyP-type peroxidase, DtpA from *S. lividans* is charactersied through structural and enzymatic studies. Using a diverse range of substrates and numerous spectroscopic techniques the peroxidase activity of DtpA has been investigated. Crucially, the peroxidase function of DtpA was assessed in the presence of GlxA. The findings highlight that DtpA does function as a peroxidase and can utilize the H_2O_2 produced by GlxA during turnover of glycolaldehyde, thus signifying an *in vitro* relationship between these two novel enzymes. Expanding on recent *in vivo* studies [171] a model for the Cu-dependent morphogenesis pathway in hyphal tips of *S. lividans* is presented.

4.2. Experimental procedures

4.2.1. Cloning of DtpA

SLI4211 encoding DtpA was amplified from the genomic DNA of *S. lividans* strain 1326 by PCR using a forward primer with a flanking 5'-NdeI restriction site and a reverse primer with a flanking HindIII restriction site (see below). The N-terminal predicted Tat signal sequence (residues 1-68) was removed during cloning resulting in a construct consisting of residues 69-445. The PCR volume was 50 µl, consisting of 2.5 µl genomic DNA, 1 µl of each primer, 2.5 µl dNTP (10 mM), 5 µl 10 x pfu + $MgSO_4$, 5 µl DMSO (100 %), 32.5 µl deionised water and 0.5 µl pfu polymerase. The following program was used to carry out the PCR: (i) 95 ºC for 3 min; (ii) 95 ºC for 1 min; (iii) 62 ºC for 1 min; (iv) 72 ºC for 2 min; (v) repeat steps (ii-iv) 35 times; 72 ºC for 8 min and a 4 ºC hold. The resulting PCR product (1,134 bp) was ligated into the NdeI and HindIII sites of a pET28a (Novagen) vector to create an N-terminal His_6-tag *pET4211* and transformed into *E. coli* JM109 cells. Colonies were selected and sequenced to corroborate the presence of the *SLI4211* gene (Figure 4.4).

Forward and reverse primers used for amplifying the *4211* gene from *S. lividans* 1326. The restriction sites for NdeI and HindIII are underlined.

SLI4211-F

5′-CTAACATATGGCGACTCCCCTCACCTCGCTC-3′

SLI4211-R

5′-CATAAGCTTTCACCCCTCCAGCAGCCGCTGA-3′

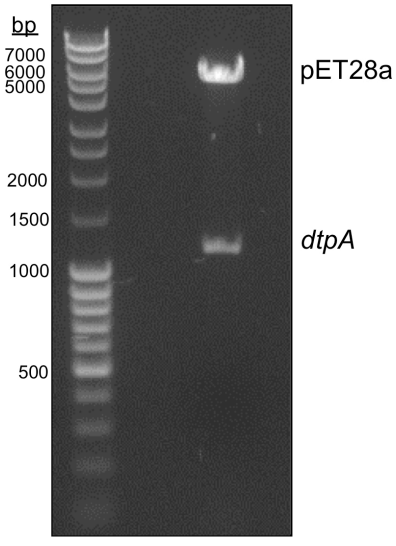


Figure 4.4: DNA 1 % agarose gel of the *dtpA (SLI4211)* pET28a construct. DNA consisting of the *pET4211* construct was digested using NdeI and HindIII restriction enzymes and corroborated against the DNA ladder (far left lane) (pET28a – 5,369 bp and *dtpA* – 1,134 bp).

4.2.2. Over-expression and purification of DtpA

pET4211 (Kan^r) was transformed into chemically competent *E. coli* BL21 (DE3) cells using the heat-shock method. Cells containing the plasmid were selected by growing on 50 μg ml^{-1} Kan agar plates and single colonies transferred to 10 ml of LB medium (Melford) containing 50 μg ml^{-1} Kan at 37 °C with shaking at 120 rpm overnight. Overnight pre-cultures were successively used to inoculate 1.4 L of high salt LB medium (10 g tryptone, 10 g NaCl, 5 g yeast extract per litre) also containing 50 μg ml^{-1} Kan and grown at 37 °C, 180 rpm. At an OD_{600} ~1.2 5-aminolaevulinic acid (ALA) (0.25 mM final concentration) and iron citrate (100 μM final concentration) were added consecutively for their use as a haem-precursor and iron supplement. Cultures were then induced by adding isopropyl β-D-thiogalactopyranoside (IPTG; Melford) to a final concentration of 0.5 mM and pure carbon monoxide (CO) gas bubbled through the culture for 20-30 s. Flasks were then sealed and incubated for a further 18 h at 30 °C at 100 rpm. Cells were harvested via centrifugation (10,000 g**,** 10 min, 4 °C) and the cell pellet resuspended in 50 mM Tris/HCl, 500 mM NaCl (Fisher) and 20 mM imidazole (Sigma) at pH 8 (buffer A). The resuspended cell suspension was lysed using an EmulsiFlex-C5 cell disrupter (Avestin) followed by centrifugation (22,000 g, 30 min, 4 °C). The clarified supernatant was loaded onto a 5-ml nickel-nitrilotriacetic acid-sepharose column (GE Healthcare) equilibrated with buffer A and eluted by a linear imidazole gradient using buffer B (buffer A with 500 mM imidazole). The protein peak eluted from the column at ~ 40 % buffer B and was collected and verified to be DtpA by SDS-PAGE gel analysis. The protein was then pooled and concentrated using a Centricon (VivaSpin) with a 10 kDa cut-off at 4 °C and loaded onto a PD-10 desalting column to remove any remaining CO and imidazole. It was then applied to a S200 Sephadex column (GE Healthcare) equilibrated with 20 mM NaPi, 100 mM NaCl, pH 7. Fractions eluting were assessed for purity by SDS-PAGE gels (see Chapter 2 section 2.2.4) then concentrated and stored at - 20 °C.

4.2.3. UV-visible spectroscopy of DtpA

All UV-visible spectra were recorded at 20 ^oC in 20 mM NaPi, pH 7, 100 mM NaCl. Concentrations of DtpA were determined using a Cary 60 UV-visible spectrophotometer (Agilent), with a 1 cm path-length Quartz cuvette (Hellma). The absorbance at 280 nm was measured and concentrations calculated using the beer-lambert law with an extinction coefficient (ε) of 37,470 M^{-1} cm^{-1} for DtpA determined using ProtParam ExPASy [93]. Hydrogen peroxide (H_2O_2) solutions (Sigma) were prepared by diluting stocks using deionised water and concentrations determined using ε of 43.6 M^{-1} cm^{-1} at 240 nm [234] Concentrations of H_2O_2 were used to form Compound I DtpA and in H_2O_2 binding and

substrate oxidation experiments. Compound II was created by the addition of 1 equivalent of ferrocyanide ($[Fe(CN)_6]^{4-}$) (Sigma) to DtpA following formation of Compound I. Reduction of DtpA was carried out by the addition of excess sodium dithionite ($Na_2S_2O_4$) (Sigma), which was then subsequently removed using a PD-10 desalting column (GE Healthcare) in an anaerobic chamber (DW Scientific $[O_2]$ < 2 ppm). The CO spectrum of DtpA was obtained by bubbling pure CO gas into the reduced enzyme following $Na_2S_2O_4$ addition within in a sealed quartz cuvette. Compound III was created by the addition of 50-fold excess H_2O_2 to the resting state enzyme. Decay of DtpA Compound I was determined by adding 1 equivalent H_2O_2 to resting state DtpA (10 μM) to generate Compound I, followed by scanning every 1.2 min between 200-800 nm to monitor decay. The absorbance at 406 nm (resting state DtpA) was measured and fitted to a single exponential function (Equation 3.2, Chapter 3) to obtain a rate of decay for DtpA Compound I.

4.2.4. Circular Dichroism spectroscopy

DtpA samples used for circular dichroism (CD) analysis were exchanged into 10 mM KPi, 50 mM KF, pH 7 using a PD-10 desalting column (GE Healthcare). Far UV-CD spectra were recorded between 260 and 175 nm at 20 ^{o}C measured on an Applied Photophysics Chirascan CD spectrophotometer equipped with a thermostatic cell holder controlled with a Peltier system (Leatherhead, UK). DtpA CD spectra was analysed using Dichroweb [235, 236] with the programs CDSSTR [237-239] and Contin-LL [240, 241] with the databases 1-7 [239] and SP175 [242]. Good fits and secondary structure estimates were obtained if the NRMSD (normalized root mean square deviation) was < 0.1 and these could then be used to calculate the α-helical and β-sheet content of the protein [243].

4.2.5. Hydrogen peroxide binding to DtpA

Binding of H_2O_2 to DtpA was initially determined by titration experiments monitored using a Cary 60 UV-visible spectrophotometer (Agilent) at 20 ^{o}C in 20 mM NaPi, pH 7, 100 mM NaCl. H_2O_2 was prepared in section 4.2.3 then titrated into the resting state enzyme (12 μM) until no further change was detected. The absorption change at 406 nm was plotted against a ratio of $[H_2O_2]$/[DtpA] to determine the stoichiometry of Compound I formation.

Transient-state kinetics of H_2O_2 binding were performed using a SX20 stopped-flow spectrophotometer (Applied photophysics) equipped with a diode array detector and monochromator, thermostatted at 25 °C with a Peltier system. A solution of DtpA enzyme (5 μM after mixing) was prepared in 20 mM NaPi, pH 7, 100 mM NaCl and rapidly mixed with a series of H_2O_2 concentrations (12.5 - 75 μM after mixing). Time courses were taken at 406 nm and the data fitted best to single exponential functions (Equation 3.1 Chapter 3). All

kinetic data were analysed using the Applied Photophysics ProKinetist software. The obtained rates were an average of three shots and were plotted against H_2O_2 concentration and analysed using a linear fit. Second-order rate constants for the formation of Compound I could be calculated from the slope of the linear fit of k_{obs} versus $[H_2O_2]$ and the standard error determined from the three shots.

4.2.6. DtpA crystallisation

DtpA (20 mg ml^{-1}) was used to screen against a range of crystallisation conditions using an ARI-Gryphon 96-well crystallisation robot. Initial crystal hits were discovered in a diverse range of crystallisation screens, but were predominately found in the JCSG$^+$ screen (Molecular Dimensions). Scaling-up and optimisation of DtpA crystals from the initial hits were carried out in 24-well plates using the hanging drop vapour diffusion method at 18 ^oC. 2 µl of protein solution at 15 mg ml^{-1} was mixed with an equal volume of reservoir solution containing 22 % 3 K polyethylene glycol (PEG) (Sigma) and 0.2 M sodium citrate, pH 5.5. Crystals grew within a week in various shapes and sizes but single flat irregular rhombohedron crystals were used for diffraction. Crystals were transferred into a cryoprotectant solution containing 22 % PEG 3 K, 0.2 M sodium citrate, pH 5.5 and 30 % sucrose for 30 s, prior to flash cooling to 100 K by plunging into liquid nitrogen. Crystallographic data were measured to 1.53 Å resolution at the I02 beamline, Diamond Light Source using an X-ray wavelength of 0.979 Å and a Pilatus 6-M-F detector (Dectris).

4.2.7. DtpA structure determination

Diffraction data obtained from DtpA crystals were processed automatically in XIA2 and scaled and merged using aimless [95] in the CCP4i suite. The structure was initially solved by molecular replacement using Molrep with a DyP structure from *S. coelicolor* as input (PDB entry 4GRC). This produced a solution for two molecules in the asymmetric unit that was predicted from the Matthews coefficient. The final structure was refined using Refmac5 [98], with model building between refinement cycles in Coot [100]. Riding hydrogen atoms were added when refinement of the protein atoms had converged. The DtpA structure was then validated using the Molprobity server [101], the JCSG Quality Control Server and tools within Coot [100]. A summary of data, refinement statistics and the quality indicators for the structure are given in Table 4.2.

Table 4.2: X-ray data processing and refinement parameters of DtpA. The DtpA crystal structure was in space group $P2_1$ with unit cell parameters 59.76, 70.18, 77.62 Å, 90, 93.21, 90°. Values in parentheses refer to the outermost resolution shell (1.53 Å)

Resolution (Å)	77.5 - 1.53
Unique reflections	96347 (4773)
Completeness (%)	99.8 (99.5)
Redundancy	4.1 (4.0)
R_{merge} (%)	0.082 (0.555)
Mn (I/σ)	8.4 (2.0)
Wilson B factor (Å^2)	18.7
R_{cryst}	0.1469
R_{free}	0.1822
RMSD Bond lengths (Å)	0.0214
RMSD Bond angles (°)	2.0383
ESU based on ML (Å)	0.053
Ramachandran favoured (%)	97.4

4.2.8. Computation of tunnels in DtpA

CAVER analyst 1.0 was used for the identification of tunnels in DtpA using molecule A [106]. The tunnels were calculated using the tunnel computation tool within CAVER defined with the haem as the starting point. Default settings for tunnel analysis were used, including the minimum probe radius set to 0.9 Å [106].

4.2.9. Steady state kinetics

Steady-state kinetics were carried out with the following substrates; ABTS, VA, DMP (2,6-dimethoxyphenol), guaiacol, RB19 and $[Fe(CN)_6]^{4-}$ (all purchased from Sigma). Samples were prepared in 1 ml cuvettes in 10 mM sodium actetate, pH 5, 150 mM NaCl or 20 mM NaPi, pH 7, 100 mM NaCl or 20 mM Boric acid, 100 mM NaCl, pH 10 or pH 10.5. They were then measured using a Hewlett-packard 8453 diode-array spectrophotometer scanning between 190 and 1100 nm thermostatted at 25 ºC. 30-1000 nM DtpA enzyme was

used with 0-60 mM ABTS, 0-90 mM VA, 0-40 mM DMP, 0-50 mM Guaiacol, 0-40 mM RB19 and 0-80 mM $[Fe(CN)_6]^{4-}$, which were baselined before the reactions were started by the addition of 0.2 mM H_2O_2. Plots of turnover rate constant (k, s^{-1}) versus substrate concentration were then constructed, whereby k which is the initial rate normalised to the enzyme concentration calculated from $((\Delta A/\varepsilon/t)/[DtpA])$ where ΔA is the absorbance change at the specific substrate oxidation wavelength (ABTS = 436 nm, VA= 310 nm, DMP = 469 nm, guaiacol = 470 nm, RB19 = 595 nm and $[Fe(CN)_6]^{4-}$ = 420 nm), ε is the extinction coefficient of the substrate oxidation product (ABTS = 29.3 $mM^{-1}cm^{-1}$, VA = 9.3 mM^{-1} cm^{-1}, DMP = 27.3 $mM^{-1}cm^{-1}$, guaiacol = 5.57 $mM^{-1}cm^{-1}$, RB19 = 10 mM^{-1} cm^{-1} and $[Fe(CN)_6]^{4-}$ = 1.04 $mM^{-1}cm^{-1}$), t is the time in seconds and [DtpA] is the total millimolar concentration of DtpA in the assay. The Michaelis constant, turnover number and catalytic efficiency were estimated by non-linear least squares fitting to the Michaelis–Menten model where possible (see Chapter 2 Equation 2.1).

Activity of DtpA with the various substrates was tested over a range of pH values using a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 and 20 mM boric acid, 100 mM NaCl pH 10-10.5. Concentrations of DtpA were 30-1000 nM depending upon the substrate being used, the substrate concentrations used were at their specific V_{max} values (30 mM ABTS, 80 mM $[Fe(CN)_6]^{4-}$,100 mM VA, 35 mM RB19, 40 mM DMP and guaiacol) with all reactions started by the addition of 0.2 mM H_2O_2 and measured using a Hewlett-packard 8453 diode-array spectrophotometer nm thermostatted at 25 ºC. The initial rates (Au/s) obtained were converted into the relative activity (%) with the highest obtained activity taken as 100 %.

4.2.10. Substrate titrations

DtpA enzyme solutions (5-8 μM in 20 mM NaPi, pH 7, 100 mM NaCl) were pre-loaded with 1 equivalent H_2O_2 creating Compound I and then titrated with either ABTS or $[Fe(CN)_6]^{4-}$ and the absorbance change monitored at 420 nm (Compound II) then plotted against a ratio of substrate to DtpA concentration. This enabled the stoichiometry of substrate necessary to reduce Compound I to Compound II and Compound II to resting state enzyme to be elucidated.

4.2.11. Stopped-flow kinetics of substrate oxidation by DtpA

DtpA protein solutions (5 μM after mixing) were prepared in either 10 mM sodium acetate, pH 5, 150 mM NaCl for ABTS and 20 mM NaPi, pH 7, 100 mM NaCl for $[Fe(CN)_6]^{4-}$ and were pre-loaded with 1 equivalent H_2O_2 to form Compound I before rapidly mixing with excess reducing substrate (0.125-1 mM ABTS and 25-150 μM $[Fe(CN)_6]^{4-}$) at 25 ^oC. Spectra were recorded upon mixing using the photo-diode array multi-wavelength unit to observed overall spectral transitions. Reactions were then monitored using a single

wavelength at 420 nm following formation and decay of Compound II. All kinetic data were analysed using the Applied Photophysics ProKinetist software. Double exponential equations were fitted to the data (Equation 4.1) to obtain pseudo-first-order rate constants (k_{obs}). The obtained rate constants were an average of three shots and were plotted against substrate concentration and analysed using a linear fit analysis. Second-order rate constants for Compound I to Compound II and Compound II to ferric (resting state) were calculated from the linear fits and the standard error determined from the three shots.

$$y = A\,(1\text{-}\exp^{-kt}) + A_2\,(\exp^{-k_2 t_2}) \qquad \text{Equation 4.1}$$

4.2.12. UV-visible spectroscopy and coupled peroxidase assay with GlxA

A Cary 60 UV-visible spectrophotometer (Agilent) at 20 ^{o}C was used to monitor various haem oxidation states of DtpA in the presence of GlxA and its substrate glycolaldehyde (Sigma). Resting state DtpA (10 μM) was prepared in 20 mM NaPi, pH 7, 100 mM NaCl in a 1 ml quartz cuvette (Hellma). Samples for UV-visible spectral analysis were prepared upon additions to the resting state DtpA enzyme of: 1 equivalent of H_2O_2; 1 equivalent of H_2O_2 followed by addition of 0.2 M glycolaldehyde; 20 μM GlxA and 0.2 M glycolaldehyde; 0.2 M glycolaldehyde; 1 equivalent of H_2O_2 followed by 1 equivalent of $[Fe(CN)_6]^{4-}$.

To assess the potential of DtpA to act as a peroxidase in conjunction with GlxA, enzymatic assays were conducted in the presence of both enzymes. Samples were prepared in 1 ml cuvettes with the following substrates assayed; glycolaldehyde (0.2 M), galactose (0.6 M), glucose (0.6 M) and N-acetyl-D-glucosamine (0.1 M) (all from Sigma) in a coupled assay containing 20 mM NaPi, 100 mM NaCl, pH 7, 30 mM ABTS, 20 μM GlxA and 5 μM DtpA or 1 μl HRP (10 mg ml^{-1}), with reactions started by the addition of GlxA. Oxidation of ABTS was monitored at 436 nm using a Hewlett-packard 8453 diode-array spectrophotometer scanning between 190 and 1100 nm thermostatted at 30 ºC. Turnover rates (k, s^{-1}) were calculated from $((\Delta A_{436}/\varepsilon_{ABTS})/t/[GlxA])$ where ΔA_{436} is the absorbance change at 436 nm upon ABTS oxidation, ε_{ABTS} is the extinction coefficient of the ABTS oxidation product taken as 29.3 $mM^{-1}cm^{-1}$, t is the time in seconds and [GlxA] is the total millimolar concentration of GlxA in the assay.

4.3. Results

4.3.1. Sequence alignment of S. lividans DtpA

DtpA amino acid sequence was aligned with various other DyP-type peroxidase subclasses including, EfeB from *E. coli* (class A), TyrA from *S. oneidensis* (class B), DyP2 from *Amycolatopsis sp* (class C) and *Bad*DyP from *B. adusta* (class D) using Clustal Omega (Figure 4.5) [180, 183, 194, 196]. The secondary structure of DtpA was predicted using Jpred4 and consists of 12 β-sheets (orange) and 8 α-helices (Teal) (Figure 4.5 and Table 4.3) [108, 113]. DtpA shares ~ 20 % identity with subclasses B, C and D (TyrA, DyP2 and *Bad*DyP) and ~ 33 % identity with other class A enzymes such as EfeB. Although the sequence identity is low, the main haem pocket residues are conserved across all subclasses and species, including the proximal His residue (His^{353} in DtpA shown by the ^ symbol in Figure 4.5), and the distal site residues, Arg and Asp (Arg^{369} and Asp^{251} in DtpA, shown by * and + symbols, respectively, in Figure 4.5). This Asp residue is present within the GXXDG motif (where X is any amino acid) (shown by a red dashed box in Figure 4.5) that has been assigned as the fingerprint of all DyP-type peroxidases [186, 189, 211]. The conserved Tyr residue postulated to have a role in LRET (long range electron transfer) is also indicated in Figure 4.5 (: symbol for Tyr^{374} in DtpA) [244].

Figure 4.5: Sequence alignment of DtpA from *S. lividans*. Clustal Omega sequence alignment of DtpA from *S. lividans* with DyP2 from *Amycolatopsis sp, EfeB* from *E. coli, TyrA* from *S. oneidensis* and *BadDyP* from *B. adusta.* Completely and partially conserved residues are coloured in dark to light grey, respectively. The secondary structure prediction for DtpA from Jpred4 [113] is indicated, with α-helices in teal and β-sheets in orange. Symbols identify the position of important haem pocket residues, with the * showing the position of Arg^{369}, the + the position of Asp^{251} and ^ the position of His^{353}. The conserved DyP motif GXXDG is shown by a dashed red box. The conserved Tyr residue with possible importance in LRET is shown by the : symbol (Tyr^{374} in DtpA). The "switch loop" is shown by a green line between residues Asn^{231}-Asp^{262} of DtpA. The Tat signal sequences for DtpA and EfeB are not included in the alignment.

Table 4.3: Jpred4 and Dichroweb analysis of DtpA. Jpred4 was also used to predict the secondary structure of DtpA [113]. Dichroweb was used to analyse the CD spectrum of DtpA with significant fits (NRMSD < 0.1) selected and averaged to calculate the secondary structural content [235].

Name	α-helix	β-sheet	Turns	Unordered
Jpred4	18%	17%	65%	
Dichroweb	16%	41%	16%	27%

4.3.2. Cloning, over-expression and purification of DtpA

dtpA was cloned from *S. lividans* to create a pET28a construct that encoded residues 69-445 (*i.e.* the N-terminal Tat signal sequence was absent) (Figure 4.6). DtpA (378 residues including the Met start codon) was over-expressed in *E. coli* to produce a pink cell pellet (Figure 4.6.A) and protein detected using 15 % SDS-PAGE gels with samples taken before and after induction (Figure 4.6.A). The overall yield of DtpA expressed in *E. coli* is ~ 7 mg L^{-1}. DtpA was purified by first using a Ni^{2+}-NTA sepharose column producing a single peak at 280 nm at ~ 30 % buffer B (Figure 4.6.B), followed by further purification on application to a S200 size-exclusion column. This produced a single peak with an elution volume of ~ 77 ml consistent with a monomeric DtpA protein species (Figure 4.6.C). Fractions eluting from the Ni^{2+}-NTA and S200 size-exclusion were analysed on 15 % SDS-PAGE gels and both displayed single pure bands running between 40-50 kDa consistent with the approximate molecular weight of DtpA (~ 40.5 kDa) (Figure 4.6.B and C).

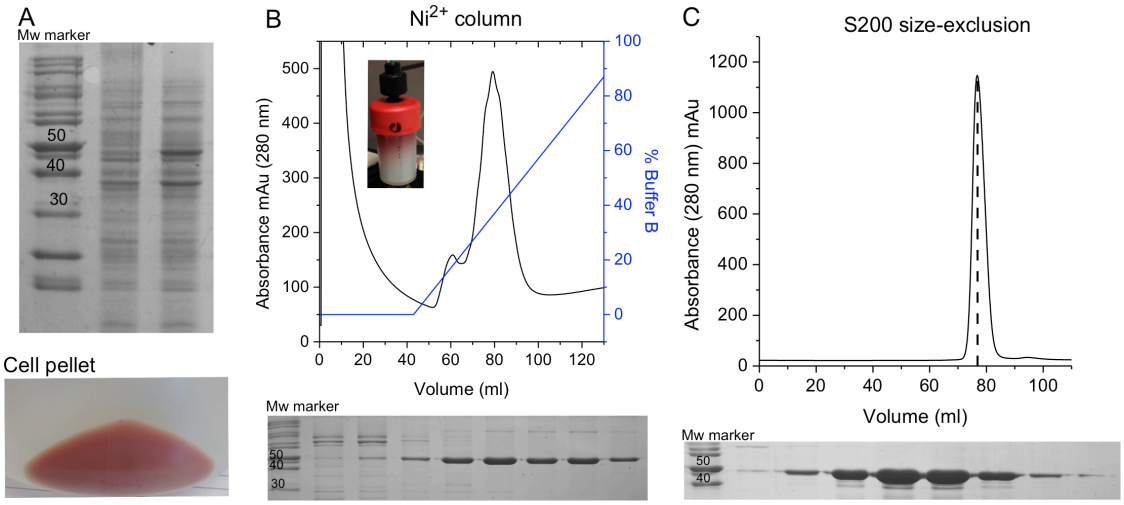


Figure 4.6: Over-expression and purification of DtpA. A) Over-expression of DtpA. A coomassie stained 15 % SDS-PAGE gel of DtpA before (lane 1) and after IPTG (lane 2). Below shows the pink pellet after over-expression. B) Ni^{2+}-NTA profile of DtpA showing an absorbance peak (mAu) at 280 nm in black at ~30 % buffer B shown in blue. Inset, a Ni^{2+}-NTA column with red DtpA bound. Below shows a coomassie stained 15 % SDS-PAGE gel of DtpA fractions from the Ni^{2+}-NTA column. A single protein band between 40-50 kDa can be identified as DtpA (containing the His_6-tag). C) S200 size-exclusion sephadex column absorbance 280 nm (mAu) profile of DtpA eluting at ~ 77 ml in 20 mM NaPi, 100 mM NaCl, pH 7. Below, Coomassie stained 15 % SDS-PAGE gel analysis of DtpA fractions from an S200 size-exclusion sephadex column. A single protein band between 40-50 kDa can be identified as DtpA.

4.3.4. Secondary structure examination of DtpA

Pro-origami was used to create a 2D protein topology diagram of DtpA [112]. A mixture of α-helices and β-sheets is predicted with 9 β-sheets and 18 small α-helices. This differs to Jpred4 which predicted 12 β-sheets and only 8 α-helices (Figure 4.6/4.7.A and Table 4.3) [113]. The topology map also predicts the orientation of the secondary structural elements with a central core of anti-parallel β-sheets surrounded at either end by several α-helices (Figure 4.7.A). Far UV-CD of DtpA confirms that the protein is folded and also displays features consistent with predictions that the secondary structure consists of a mixture of α and β content revealed by the minima at 208 nm and 230 nm and maxima at ~196 nm (Figure 4.7.B). Dichroweb was used to analyse the DtpA CD spectrum with significant fits (NRMSD < 0.1) selected and averaged to calculate the secondary structural content [235, 236]. Examples of fits from Dichroweb analysis for DtpA are given in Appendix 2.1. Dichroweb gave the secondary structure % in Table 4.3 predicting a much higher proportion of β-sheet content compared to that predicted from Jpred4 [245]. However, these data provide strong evidence prior to tertiary structural information that DtpA contains a mixed α/β structure.

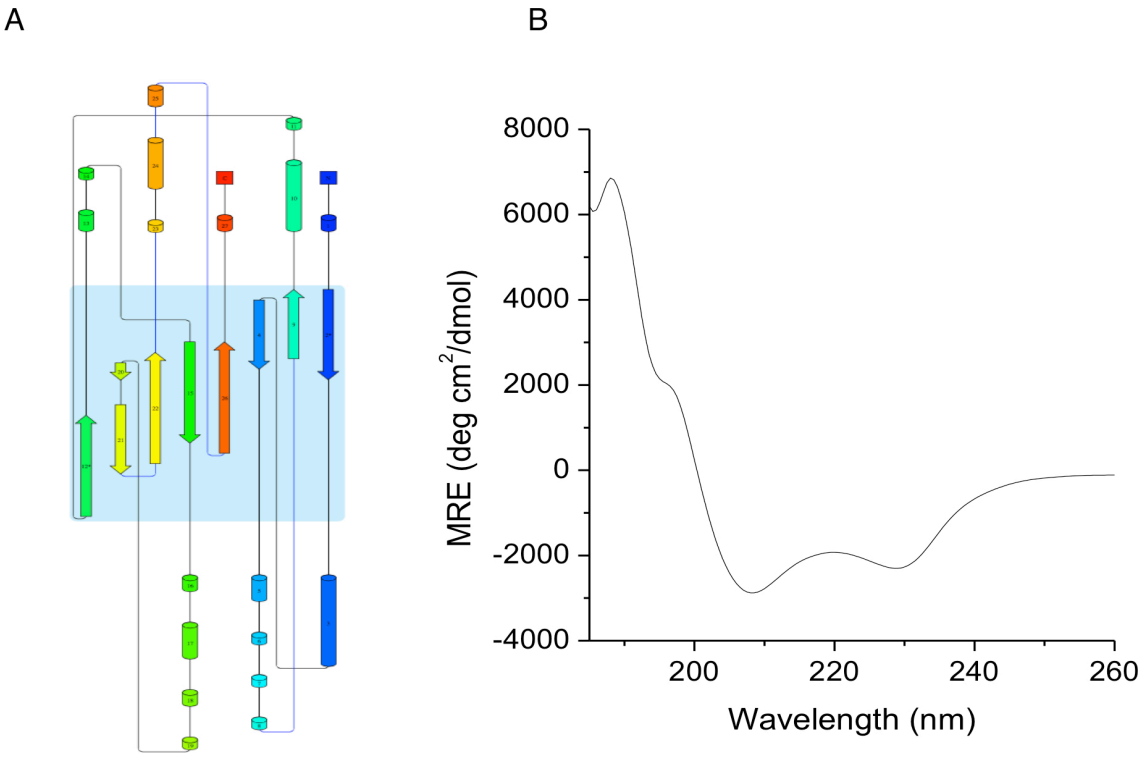


Figure 4.7: Secondary structural analysis of *S. lividans* DtpA. A) The secondary structure topology map of DtpA showing 9 β-sheets and 18 α-helices with the predicted 2D assembly (Pro-origami) (Stivala *et al.*, 2011). B) Far UV-CD of isolated DtpA (10 μM) in 10 mM KPi, 50 mM KF, pH 7 recorded between 260 and 180 nm at 20 ^{o}C.

4.3.5. Spectral characteristics of DtpA

As isolated, DtpA is brown in colour and gives rise to spectral transitions consistent with a resting state haem peroxidase, comprising a soret peak at 406 nm and peaks at 502 nm and 635 nm corresponding to the β and charge transfer bands and also a shoulder at 540 nm [180, 186] (Figure 4.8, Table 4.4). The *Reinheitszahl* value which reflects the purity and spectral characteristics (R_z=A_{406}/A_{280}) was ~ 2.2 for DtpA at pH 7 consistent with other DyP-type peroxidases [180, 189, 190, 209, 246, 247]. Upon addition of 1 equivalent H_2O_2, the soret absorption band shifts to 399 nm, becoming broader and decreases in intensity by ~ 46 %, with a shoulder at ~ 350 nm (Figure 4.8 and Table 4.4). The visible region displayed broad absorption peak maxima at 530, 557, 614 and 644 nm [180, 186] (Table 4.4 and Figure 4.8). This indicates that DtpA has undergone a two-electron oxidation process to form Compound I (Por•+ Fe^{IV}=O). The spectral characteristics of resting state and Compound I are indistinguishable from other DyP-type peroxidases and are also comparable to plant peroxidase representatives such as MnP, LiP and HRP. This therefore implies that the primary reaction state of DtpA Compound I is the same Fe^{IV} oxoferryl centre and either a protein or porphyrin-based cation radical intermediate. Upon addition of 1 equivalent of $[Fe(CN)_6]^{4-}$, the DtpA Compound I species transforms into a species with a Compound II spectrum *i.e.* a ferryl species (Fe^{IV}=O) with maxima at 419, 528, 557, 621 and 728 nm (Table 4.4 and Figure 4.8), which differs to the Compound II spectrum of *Bad*DyP but is consistent with other DyPs and plant peroxidases [40, 185, 248-250]. This indicates a single electron reduction of Compound I (Por•+ Fe^{IV}=O) to Compound II (Fe^{IV}=O). Upon the addition of excess H_2O_2 to the resting state enzyme, it converts the enzyme into spectral characteristics consistent with the inactive Compound III form, either a Fe^{II}=O_2 or Fe^{III}-O_2•$^-$ species with peak maxima at 416, 541 and 577 nm (Figure 4.8 and Table 4.4). This is coherent with spectral characteristics of plant peroxidases and other DyP members [180, 251-253]. When the resting state is treated with excess $Na_2S_2O_4$ the soret band shifts and flattens creating the deoxy-ferrous (Fe^{2+}) spectra of DtpA with absorbance maxima at 422 and 559 nm (Figure 4.8 and Table 4.4). By bubbling pure CO gas into the reduced DtpA enzyme generates a CO bound form with a sharp soret absorbance peak at 420 nm and α/β peaks of 571 and 539 nm, respectively (Figure 4.8, Table 4.4).

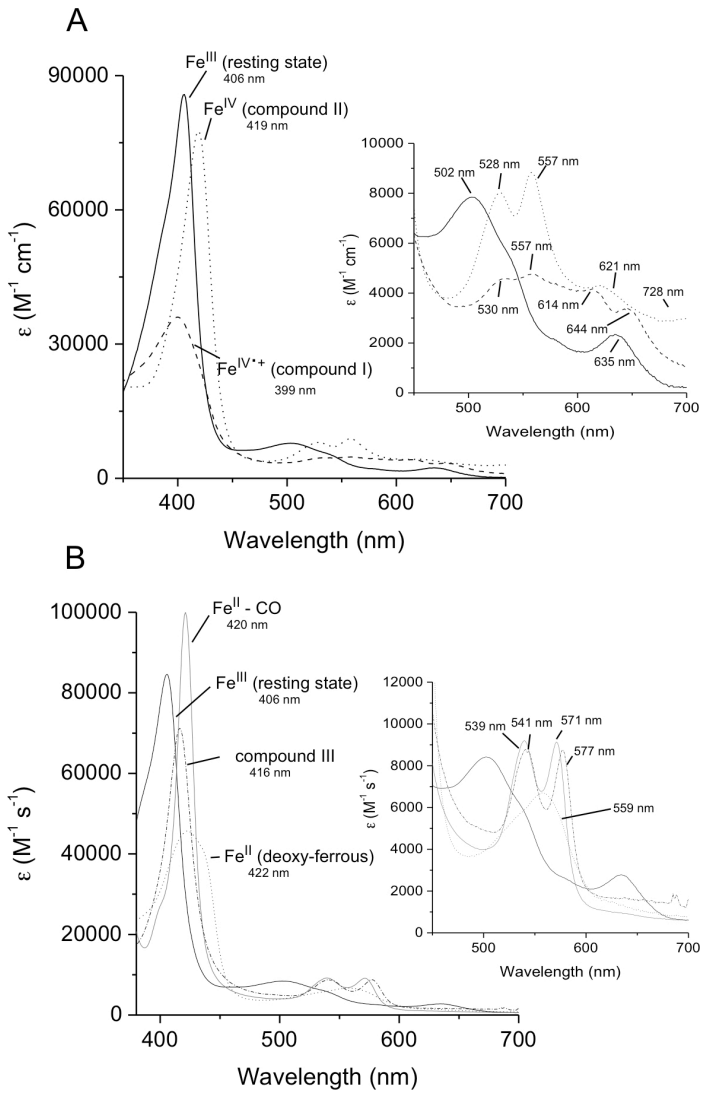


Figure 4.8: UV-visible absorbance spectra of DtpA. A) Various haem oxidation states of DtpA (8 μM) after the addition of 1 equivalent H_2O_2 to the resting state enzyme forming Compound I then after the addition of 1 equivalent $[Fe(CN)_6]^{4-}$ forming Compound II. Inset, shows an enlarged scale of the Q bands with wavelengths depicted. B) Resting state DtpA (18 μM), ferrous DtpA (9 μM) after the addition of excess $Na_2S_2O_4$, the ferrous-CO (23 μM) spectrum after pure CO gas is bubbled through the ferrous protein and Compound III after the addition of 50-fold excess H_2O_2 to resting state DtpA. All spectra recorded in 20 mM NaPi, pH 7, 100 mM NaCl at 20 ^oC.

Table 4.4: Absorption maxima of oxidized haem intermediates of DtpA. (*sh*) stands for shoulder.

	Resting	Compound I	Compound II	Compound III	Deoxy-Ferrous	Ferrous-CO
DtpA	**406**, 502, 540 (*sh*), 635	**399**, 350 (*sh*), 530, 557, 614, 644	**419**, 528, 557, 621, 728	**416,** 541, 577	**422**, 559	**420**, 539, 571

4.3.6. The rate of formation and decay of Compound I

Upon addition of H_2O_2 to DtpA Compound I is formed with an absorbance spectrum consistent with other peroxidases (Figure 4.8 and 4.9) [180, 186]. On titrating known concentrations of H_2O_2 into DtpA a linear decrease in the absorbance at 406 nm was observed until a break point was reached at a ratio of ~ 1 H_2O_2:DtpA (Figure 4.9). This indicates that it takes ~ 1 H_2O_2 molecule to form Compound I, which is expected, as H_2O_2 is capable of removing two-electrons from DtpA.

Compound I formation was also observed using a stopped-flow spectrophotometer (Figure 4.10) enabling for the rate of formation of Compound I to be determined (at pH 7) by monitoring the decay of the soret band at 406 nm upon rapid mixing of the resting state enzyme with various H_2O_2 concentrations. Within ~ 30 ms the observation of bands characteristic of Compound I could be observed (Figure 4.10.A). Single-order exponential functions were fitted to the kinetic traces and the pseudo-first order rate constants (k_{obs}) plotted against H_2O_2 concentration (Figure 4.10.B). No intermediate was resolved and a two-step model of A to B was used to fit the data. A second-order rate constant for Compound I formation of 8.9 (± 0.25) x 10^6 M^{-1} s^{-1} was determined (Figure 4.10.B). The linear fit runs through or very close to the origin, showing that the reverse chemical reaction of Compound I formation is very low (*i.e* the reduction of Compound I is likely to be negligible compared to its formation) (Figure 4.10.B). This rate of oxidation is comparable to other DyP-type peroxidases such as *Tc*DyP from *T. curvata* (3.4 x 10^6 M^{-1} s^{-1}) but is higher than some other DyPs such as the B-type from *R. jostii* RHA1 (1.79 x 10^5 M^{-1} s^{-1}) [180, 209].

To assess the stability of Compound I, 1-equivalent of H_2O_2 was added to the resting state enzyme to form Compound I, and the decay monitored by following the absorption at 406 nm (Figure 4.11). A single exponential function fitted to the data (Equation 3.2, Chapter 3) gave a $t_{1/2}$ ~ 2.5 min (Figure 4.11.B). This very stable Compound I species decayed to a spectrum consistent with a ferric enzyme without any detectable Compound II-type intermediate, similarly noted for other DyPs [183, 186, 204, 211, 254].

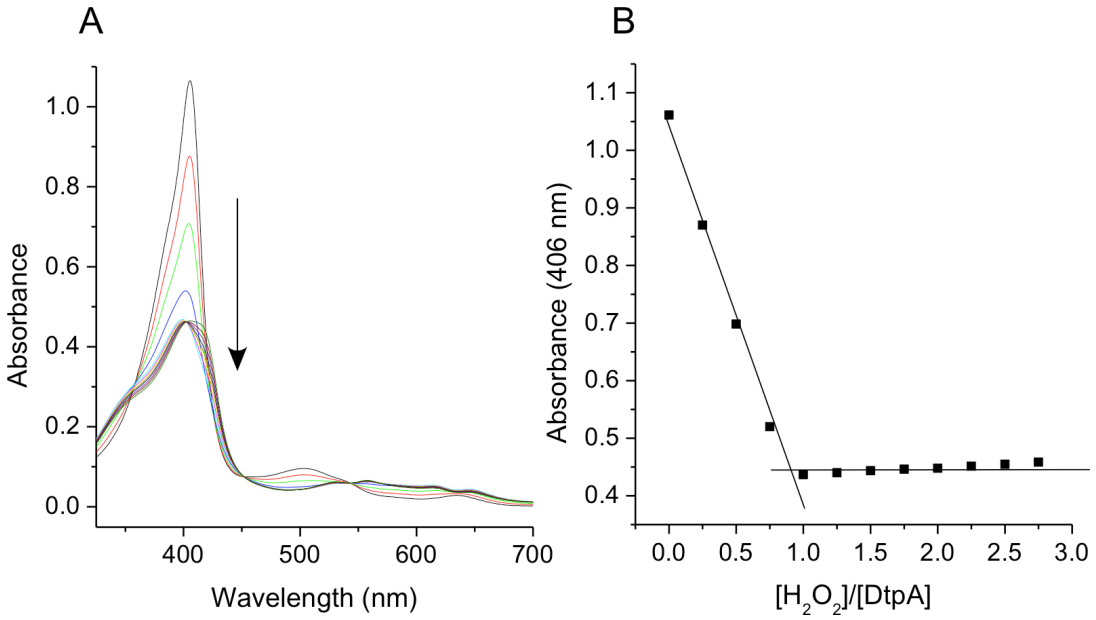


Figure 4.9: Formation of DtpA Compound I by titrating with H_2O_2. Compound I formation was monitored using UV-visible absorbance changes at 20 ^oC in 20 mM NaPi, pH 7 100 mM NaCl upon titration of known H_2O_2 concentrations to the resting state enzyme (12 μM). A) The absorbance spectrum of DtpA upon addition of H_2O_2. B) The absorbance changes at 406 nm plotted as a function of $[H_2O_2]/[DtpA]$. Lines indicate the point of intersection.

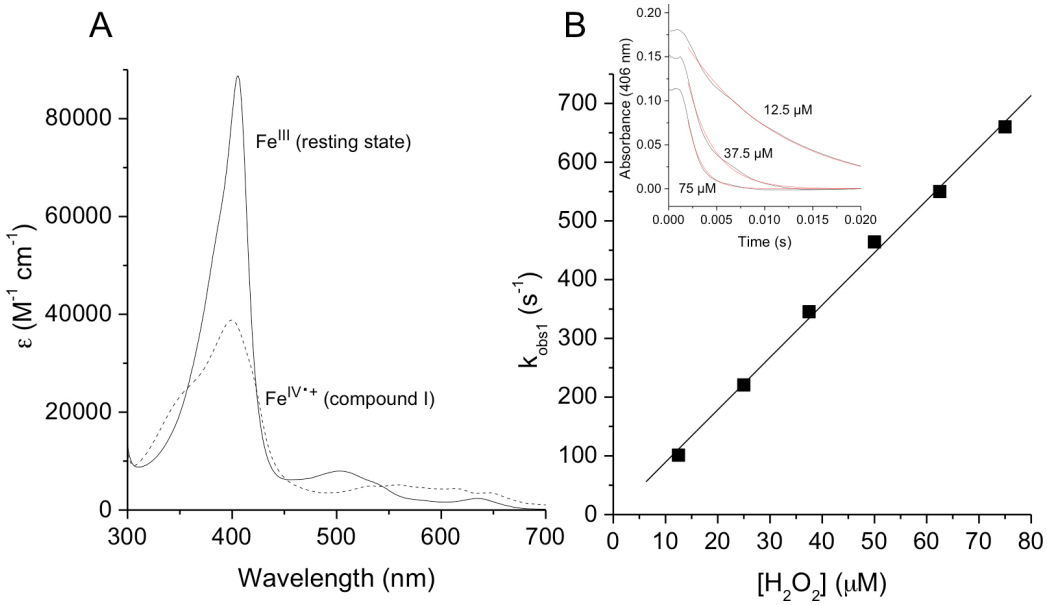


Figure 4.10: Formation of DtpA Compound I. A) UV-visible spectra of resting state (ferric) DtpA (12 μM) (solid line) and compound I (dashed line) after the addition of 1 equivalent H_2O_2. B) Stopped-flow kinetics of DtpA compound I formation with H_2O_2. Inset, example time traces (12.5, 37.5 and 75 μM H_2O_2) at 406 nm fitted to single exponential functions. The mixing period of ~ 2 ms can be observed. DtpA (5 μM after mixing) solutions were mixed with a range of excess H_2O_2 concentrations (12.5-75 μM) and the rates obtained from the single exponential fits with each plotted rate an average of three traces and thus the error is accounted for in the size of the marker. All experiments were performed at 25 ^oC in 20 mM NaPi, pH 7, 100 mM NaCl.

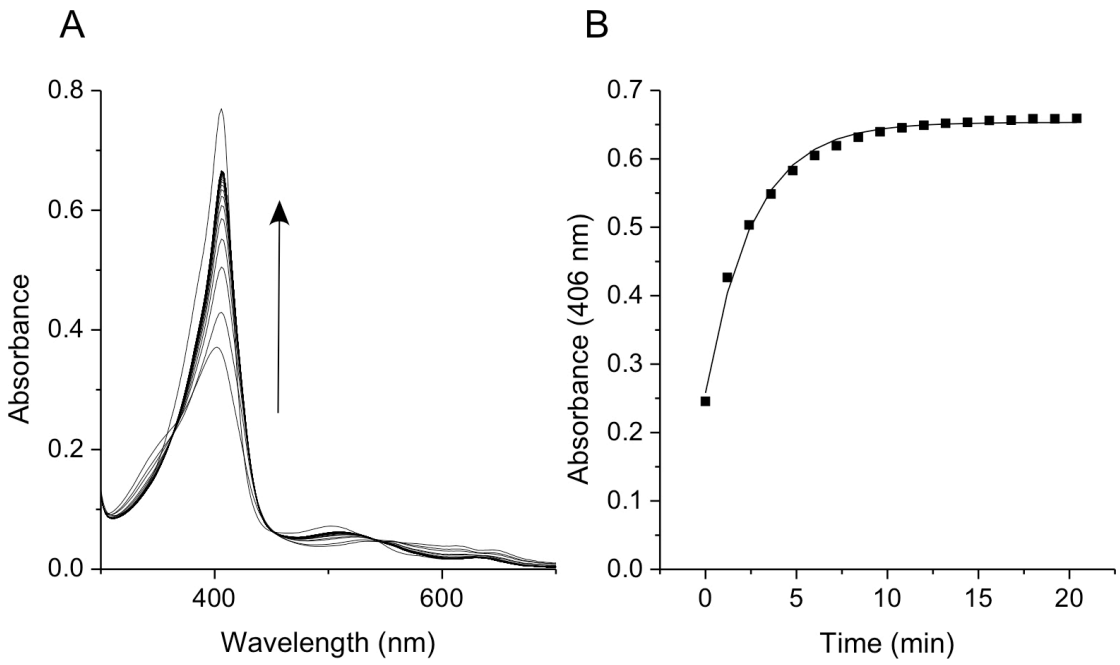


Figure 4.11: The decay of DtpA Compound I. Decay of Compound I DtpA was monitored using UV-visible absorbance changes at 20 ^{o}C in 20 mM NaPi, pH 7, 100 mM NaCl. A) The absorbance spectrum of DtpA (10 μM) showing the resting state enzyme (top most spectrum) then after the addition of 1 equivalent H_2O_2 showing Compound I formation (bottom spectrum) then scans taken every 1.2 min as Compound I decays. B) The absorbance changes at 406 nm as a function of time fitted to a single exponential function (Equation 3.2, Chapter 3).

4.3.7. Crystal structure of DtpA from S. lividans

Single flat irregular rhombohedron crystals of DtpA were obtained and used for X-ray measurements after being transferred into a cryoprotectant and flash cooled to 100 K (Figure 4.12). Crystallographic data were measured to 1.53 Å resolution, with a space group of $P2_1$ assigned and two molecules identified in the asymmetric unit (Table 4.2). The orientation of molecule A (dark grey) and molecule B (light grey) (both modelled from Thr^{70} – Gly^{445}) can be seen in Figure 4.13.A. Both monomers display similar B-factors with a disordered region (not visible in the electron density) for residues 272-277 and 272-279 for molecule A and B, respectively (Figure 4.13.A). Superposition of the two monomers gives r.m.s deviation of less than 0.4 Å for all corresponding Cα atoms. The orientation of the two-haem centres can also be observed towards the C-terminal domain of each monomer with an Fe present at the centre of each (Figure 4.13.A). Although two molecules are identified in the crystallographic asymmetric unit, DtpA exists in solution as a monomer, as confirmed by the elution profile from a S200 size-exclusion column, previously calibrated with other proteins with known molecular weights (Figure 4.6). The overall structure of the DtpA monomer (~ 40.5 kDa) has dimensions of ~ 65 x 46 x 35 $Å^3$ consistent with other peroxidase structures [186] (Figure 4.13.B). Of the 377 total residues in DtpA, 104 residues form α-helices (27.5 %) and 72 β-strands (19 %), consistent with results from Jpred4 and similar to dichroweb analysis (Table 4.3). In general DyPs contain more β-strands than other peroxidases such as LiP, MnP and HRP (Figure 4.2) [206, 255, 256]. The overall structure of DtpA is formed by two main domains, each containing an antiparallel large β-sheet with either four-strands (β-strands 1, 2, 3 and 4) or five strands (β-strands 5, 6, 7, 8 and 9) (shown in red in Figure 4.13.B). However, strands 7 and 8 are displayed as one large strand in CCP4MG. Each domain also contains multiple helices (α-1-4 and α–6-10 shown in blue in Figure 4,13.B) resulting in a ferredoxin-like fold. CCP4MG also picks up on two small β-strands not identified using PyMol between α7 and α8, although these are not always present in all DyP structures. This arrangement was largely predicted prior to X-ray structural information using Pro-origami (section 4.3.4) (Figure 4.7.A).

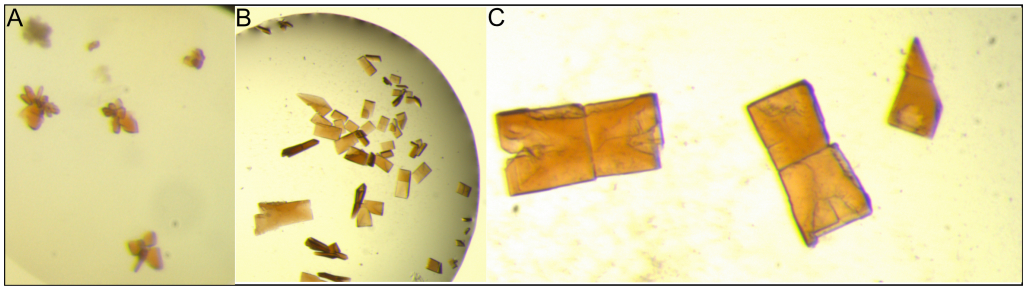


Figure 4.12: Crystallisation of DtpA. A) Small clusters of DtpA crystals (20 mg ml^{-1}) in a 96 well drop containing 20 % PEG 3 K and 0.2 M sodium citrate, pH 5.5. B and C) Larger DtpA crystals (15 mg ml^{-1}) from 24-well drops containing 22 % PEG 3 K and 0.2 M sodium citrate, pH 5.5.

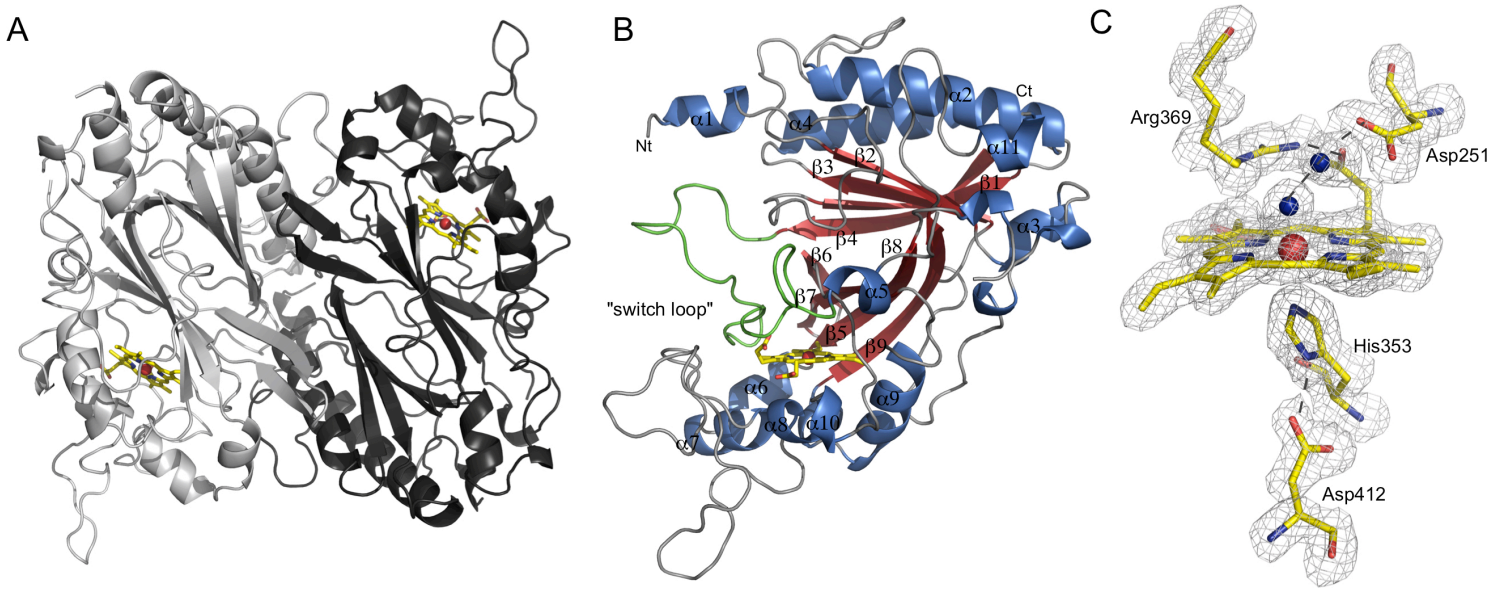


Figure 4.13: X-ray crystal structure of *S. lividans* DtpA. A) Arrangement of DtpA molecules in the asymmetric unit shown as a cartoon representation with one molecule coloured light grey (molecule B) and the other dark grey (molecule A). The haem centres are represented as yellow sticks and the Fe as red spheres. B) Cartoon representation of a single DtpA molecule (Chain B). The α-helices are coloured in blue and the β-sheets in red with their numbers depicted. The haem is shown as yellow sticks and the Fe as a red sphere. The "switch loop" is shown in green between residues Asn^{231}-Asp^{262} of DtpA. C) A 2 F_o-F_c electron-density map contoured at 1 σ of the DtpA active site. The haem and ligands are represented as yellow sticks, the Fe as a red sphere and two water molecules as blue spheres. Nitrogen atoms are coloured blue and oxygen atoms red. Dashed lines indicate H-bonds. Images were prepared in PyMol.

4.3.8. The haem environment of DtpA

Fe metal ions were modelled into two clear 26 σ peaks in the σ-weighted F_o-F_c difference map of DtpA (Figure 4.13). These were present within two protoporphyrin IX rings, found in each DtpA monomer within the asymmetric unit. The haem is present within the C-terminal domain of DtpA in a large hydrophobic cavity (Figure 4.13). The cavity is inserted between α-helices 6-10 and β-strand 8 and contains a volume of ~ 695 $Å^3$ identified using CAVER [106]. Within this pocket the porphyrin ring creates Van der Waals interactions with various hydrophobic residues including; Met^{247}, Ile^{294}, Met^{296}, Ile^{313}, Ala^{354}, Pro^{371}, Leu^{388}, Phe^{390}, Phe^{401}, Val^{404}, Leu^{408}, Leu^{414}, Ile^{418}. A flexible loop can be identified in DtpA covering and interacting with the haem as has previously been identified in EfeB from *E. coli*. This flexible loop is called the "switch loop", which connects the N- and C-terminal domains of the enzyme (residues Asn^{231}-Asp^{262} in DtpA shown in green in Figure 4.13.B) and has been proposed to be important in substrate stability [203]. The loop is formed by 32 amino acid residues in DtpA (Asn^{164} - Asp^{195}) as opposed to only 22 in EfeB and is believed to flip in and out to enable substrate and product entry and exit during peroxidase turnover [203]. This loop also participates in dimer interactions between the two monomers in the crystallographic unit, with haem binding stabilising the loop and assisting in dimerization.

Consistent with other peroxidase structures, the fifth ligand to the haem iron is a His residue (His^{353}). This His residue is within the C-terminal region of DyPs whereas equivalent His residues in other peroxidases are generally located within the middle region of the protein [175]. The Nε atom of the imidazole ring of His^{353} is 2.16 Å away from the Fe atom and the Nδ atom is H-bonding to the carboxylate of a nearby Asp^{412} residue (2.67 Å) (Figure 4.13.C). This Asp is conserved among some DyP members and plant peroxidases, however it is replaced by a Glu in EfeB (Glu^{389}) and *Bad*DyP (Glu^{391}) [186, 203, 257] (Figure 4.13.C). Two H_2O molecules are found in the assumed binding pocket for H_2O_2 on the distal face of DtpA with a distance of 1.83 Å between the two H_2O molecules (H_2O 95 and 299) and H_2O 95 is 2.17 Å away from the Fe (Figure 4.13.C). DtpA lacks the distal His ligand that is conserved among classical peroxidase families and instead the pocket contains Arg^{369} and Asp^{251} (Figure 4.13.C). The Asp residue is conserved in all DyPs (Figure 4.13.C Asp^{251} DtpA shown by + symbol in the alignment) and appears to replace the role of the distal His in other peroxidases. The catalytic importance of this residue seems to depend upon the enzyme type, with it essential for D-type DyPs but is not require for B-type enzymes and its role in A-type enzymes is still questionable [40, 189, 203]. The Arg residue identified in DtpA (Arg^{369}) H-bonds to Asp^{251} and is also present in other DyPs and plant peroxidases and acts as an essential residue in the activity of the peroxidase (Figure 4.13.C) [186, 189, 206, 209]. The two propionic acid groups of the haem are orientated such that they face outwards towards the solvent (Figure 4.13.C). The propionate group on pyrrole A bends almost 90 ^{o} and is

stabilised through H-bonds to the NH1 group of Arg^{369}, the backbone amide group of Gly^{252} and to a H_2O molecule (H_2O 13), a similar arrangement is found in other DyP structures (Figure 4.13.C) [186, 194, 203, 204]. The second propionic acid moiety present on pyrrole ring D also H-bonds to the same H_2O molecule (H_2O 13) and also to the Nε and NH1 of Arg^{254} and also to two other H_2O molecules (H_2O 35 and 104) (Figure 4.13.C). Finally, another important residue has been revealed in the distal haem pocket, which varies between DyP subclasses. In DyPB from *R. jostii* RHA1 it is an Asn residue (Asn^{246}) [209], a Gly in DyP2 from *Amycolatopsis sp* (Gly^{350}) [196] and a Ser in *Bad*DyP from *B. adusta* (Ser^{331}) [186]. The Asn residue (Asn^{246}) in *R. jostii* RHA1 was shown through a N246A variant to play a minor role in Compound I formation but is important in stabilizing this intermediate [209]. EfeB a class A Dyp-type peroxidase, displays a Gly residue at the equivalent position [203]. However, in DtpA this residue is a Pro (Pro^{304}) (identical to 4GT2 from *S. coelicolor* unpublished). Interestingly, in other *S. coelicolor* DyP structures this residue also varies, with either a Gly or Asn residue present (unpublished PDB entry 4GU7 is Asn^{245} and Gly^{344} in 4GRC).

4.3.9. Substrate access or product exit tunnels in DtpA

In DtpA putative substrate tunnels were measured from the haem to the protein surface. CAVER calculated the presence of nine tunnels, however of these only two can be identified as main tunnels with the others only creating extensions of these two (Figure 4.14). Tunnel I is the shorter of the two with a distance to the surface of ~ 9.7 Å which extends up from the two haem coordinated H_2O molecules on the distal face (Figure 4.14.B). The channel contains six H_2O molecules leading to the entrance of the tunnel (including the two coordinating the Fe) and a further two H_2O molecules past the opening to the channel as identified by CAVER (blue spheres, Figure 4.14.C). The radius of Tunnel I is significantly smaller than Tunnel II with an opening of ~ 7 x 7 Å and an average diameter of ~ 4 Å (Figure 4.14.B). Tunnel II creates a large pocket surrounding and including the haem made up of ~ sixteen H_2O molecules (red spheres, Figure 4.14.C). Tunnel II is longer than Tunnel I with a length of ~ 20 Å and also consists of a wider opening of ~10 x 10 Å and an average diameter of ~ 5.5 Å (Figure 4.14.B). Many of the substrates used by DtpA such as ABTS and RB19 are large and bulky molecules and therefore may not be able to easily access the buried haem cofactor through either of these tunnels. However, Tunnel I includes 8 and Tunnel II 10 residues which are part of the "switch loop", and therefore these residues may change their position and orientations through substrate binding or on Compound I formation [203].

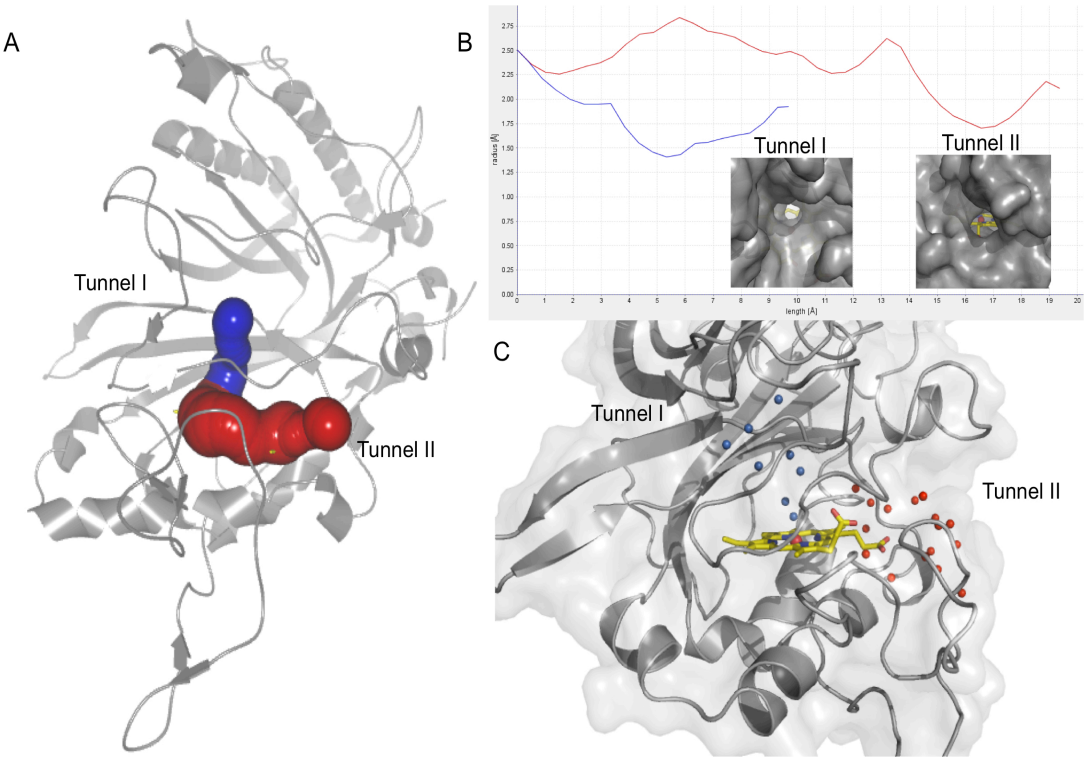


Figure 4.14: Tunnels identified in DtpA from *S. lividans*. A) The position of two surface-to-haem tunnels in DtpA, tunnel I in blue and tunnel II in red identified using the programme CAVER [106]. B) A graph of the length (Å) and radius (Å) of both tunnel I (blue) and tunnel II (red). Inset, shows a partial transparent surface view of the entrances to tunnel I and tunnel II. C) A partial transparent view of the water molecules present in the two identified tunnels, blue spheres represent water molecules in tunnel I and red spheres illustrates the waters present in tunnel II. The haem is represented as yellow sticks and the Fe as a red sphere.

4.3.10. Electrostatics of DtpA

DtpA has a pI of 6.1, which is significantly higher than many of the other reported DyPs such as *Bad*DyP with a pI of 4.48 and DyPB a pI of 4.39 [204, 258]. Electrostatic surface representations of DtpA show a somewhat disperse distribution of charge across the surface of the protein (Figure 4.15). However, the entrance to Tunnel II contains a large area of positive charge (blue, entrance shown by a dashed circle Figure 4.15), inferring the possibility of entry or exit of negatively charged substrate or product. In contrast, proposed entrances to Tunnel I display largely neutral charges (Figure 4.15).

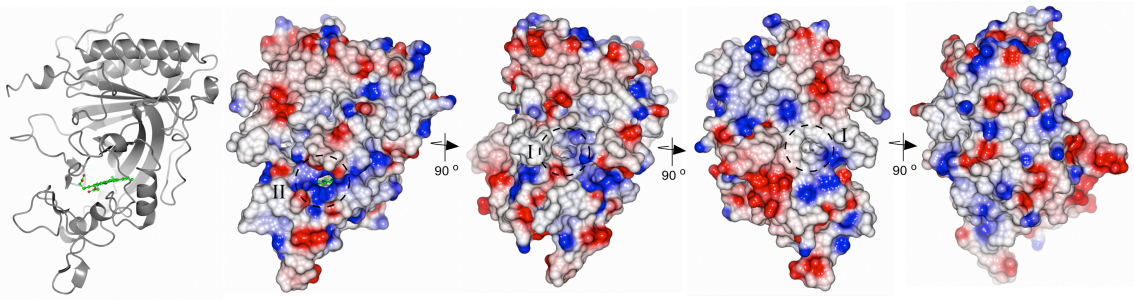


Figure 4.15: Electrostatic surface representations of DtpA. A cartoon representation and four electrostatic orientations of DtpA are shown, with 90^{o} rotations displayed. Positive charges are indicated in blue and negative charges in red, with entrances to the tunnels leading to the haem shown by dashed circles and the tunnel numbers depicted (either I or II). Images made in CCP4MG [121].

4.3.11. Steady state enzyme kinetics of DtpA with various substrates

In order to explore the substrate specificity of DtpA a panel of well-known DyP-type peroxidase substrates were evaluated. These included the classical peroxidase substrate ABTS, the AQ dye RB19, phenolic substrates DMP and guaiacol, the non-phenolic substrate VA and the small inorganic reductant $[Fe(CN)_6]^{4-}$. The chemical structures of these substrates are shown in Figure 4.16. Steady-state kinetics of DtpA were performed over a range of substrate concentrations and pH values. pH profiles were determined for each substrate and steady state kinetics were carried out at optimum pH values. Michaelis-Menten kinetics were observed for all substrates with the exception of RB19. The K_m, k_{cat} and k_{cat}/K_m values are reported in Table 4.5. For the substrates displaying Michaelis-Menten kinetics, $[Fe(CN)_6]^{4-}$ displayed the highest activity (k_{cat}/K_m) at pH 5 with a second peak of activity at pH 7 (Figure 4.17 and Table 4.5). The second highest activity was observed for ABTS also at pH 5 but a much broader pH profile was observed with a lower activity optimum at pH 7 (Figure 4.18.A and Table 4.5). DtpA displayed much lower activity for VA, DMP and guaiacol over the pH range explored (Table 4.5). It is noted that for both the phenolic substrates (DMP and guaiacol) alkaline pH proved optimal for highest activity (Figure 4.20 and 4.21). Overall however, no significant differences were observed for the k_{cat}/K_m values amongst these substrates (Figure 4.19/4.20/4.21 and Table 4.5). For RB19 a clear pH optimum is observed at pH 4 (Figure 4.22). However, unlike the other substrates tested, activity is now observed with μM substrate concentrations but Michaelis-Menten kinetics is not followed. The data obtained is therefore not easily accountable and suggests a more complicated substrate oxidation model. Examples of wavelength changes upon oxidation of the substrates and time courses observed are given in Appendix 2.2 and 2.3.

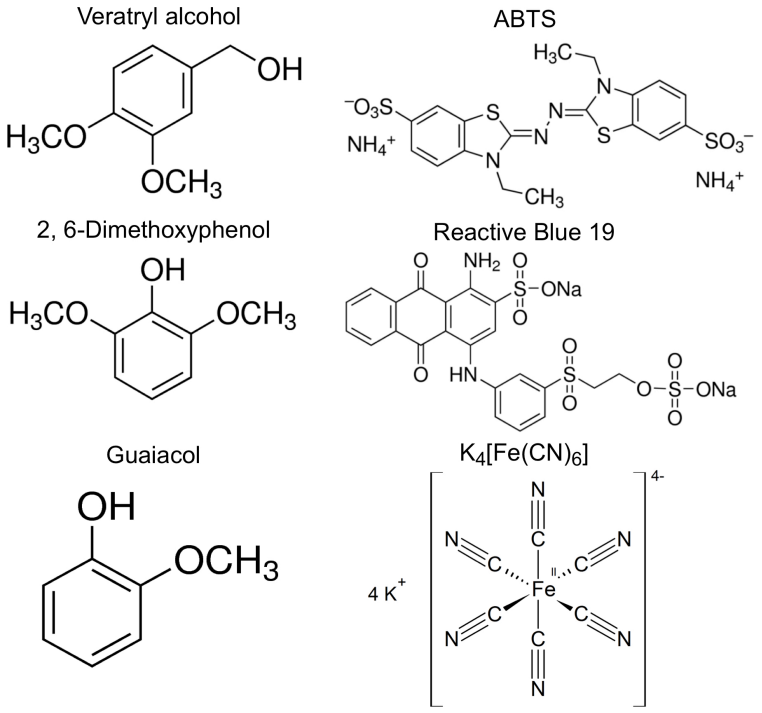


Figure 4.16: Structures of reducing substrates used to determine substrate activity of DtpA.

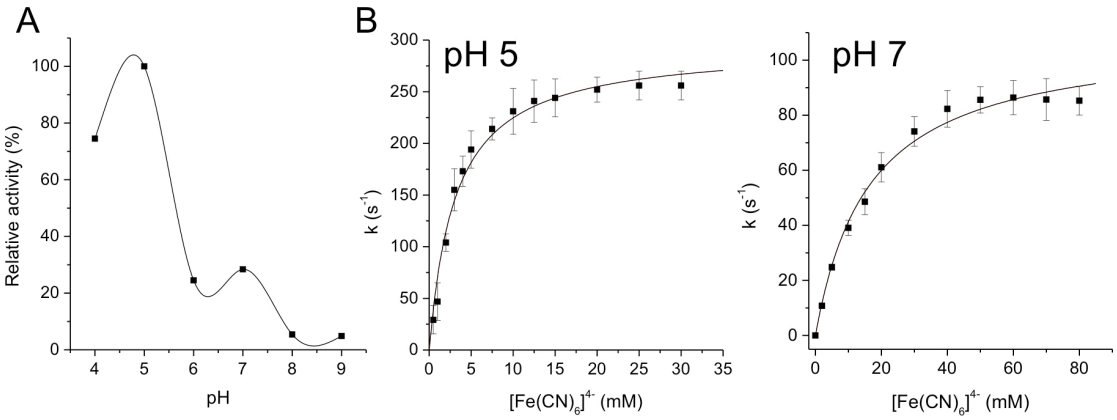


Figure 4.17: Steady-state kinetics of $[Fe(CN)_6]^{4-}$ oxidation by DtpA. A) pH profile of DtpA with $[Fe(CN)_6]^{4-}$. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 with 80 mM $[Fe(CN)_6]^{4-}$ and 40 nM DtpA. B) Steady-state kinetics of DtpA with $[Fe(CN)_6]^{4-}$ at pH 5 and pH 7. Reactions were performed using 40 nM enzyme, 0.2 mM H_2O_2 and different $[Fe(CN)_6]^{4-}$ concentrations at pH 5 (10 mM Na acetate, 150 mM NaCl) and pH 7 (20 mM NaPi,100 mM NaCl). The data were fitted by non-linear least squares Michaelis–Menten model (see Chapter 2 Equation 2.1).

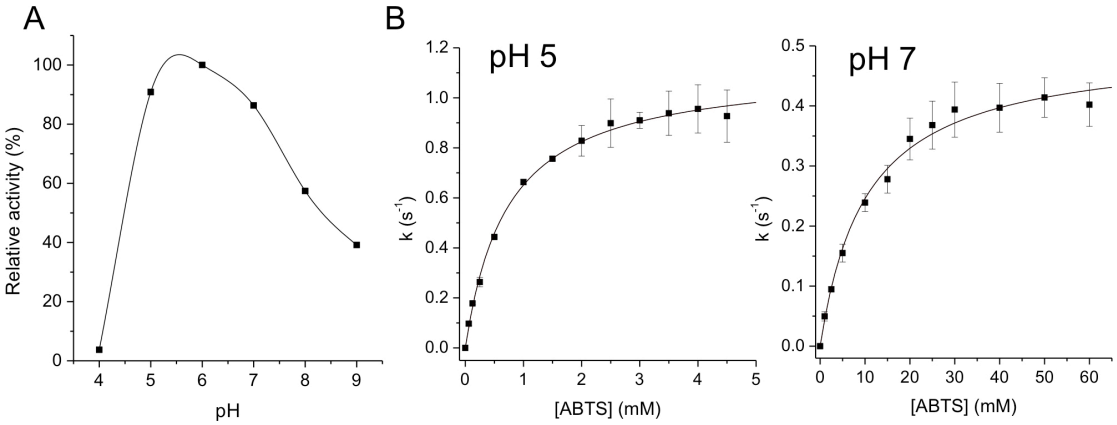


Figure 4.18: Steady-state kinetics of ABTS oxidation by DtpA. A) pH profile of DtpA with ABTS. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 with 30 mM ABTS and 200 nM DtpA. B) Steady-state kinetics of DtpA with ABTS at pH 5 and pH 7. Reactions were performed using 200-300 nM enzyme, 0.2 mM H_2O_2 and different ABTS concentrations at pH 5 (10 mM Na acetate, 150 mM NaCl) and pH 7.0 (20 mM NaPi, 100 mM NaCl). The data were fitted by non-linear least squares Michaelis–Menten model (see Chapter 2 Equation 2.1).

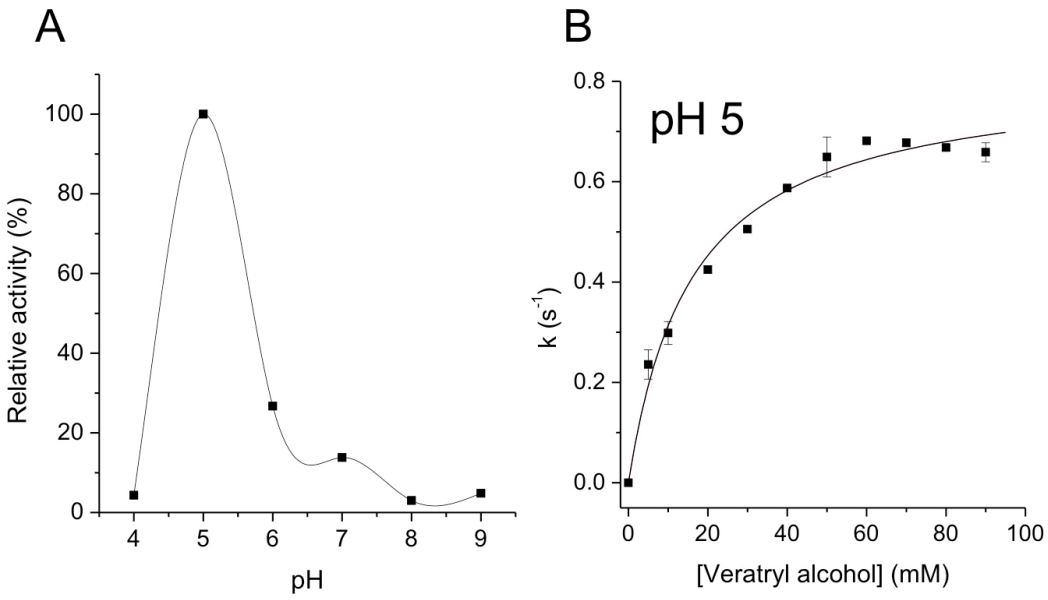


Figure 4.19: Steady-state kinetics of VA oxidation by DtpA. A) pH profile of DtpA with VA. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 with 100 mM VA and 1 μM DtpA. B) Steady-state kinetics of DtpA with VA at pH 5. Reactions were performed using 1 μM enzyme, 0.2 mM H_2O_2 and different VA concentrations at pH 5 (10 mM Na acetate, 150 mM NaCl). The data were fitted by non-linear least squares Michaelis–Menten model (see Chapter 2 Equation 2.1).

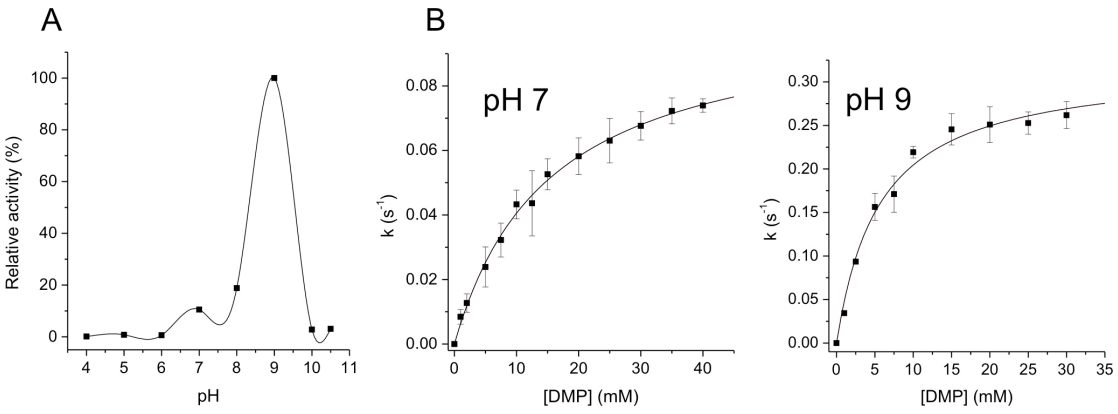


Figure 4.20: Steady-state kinetics of DMP oxidation by DtpA. A) pH profile of DtpA with DMP. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 and 20 mM boric acid, 100 mM NaCl pH 10-10.5 with 40 mM DMP and 1 μM DtpA. B) Steady-state kinetics of DtpA with DMP at pH 7 and pH 9. Reactions were performed using 1 μM enzyme, 0.2 mM H_2O_2 and different DMP concentrations at pH 7 (20 mM NaPi, 100 mM NaCl) and pH 9 (10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl). The data were fitted by non-linear least squares Michaelis–Menten model (see Chapter 2 Equation 2.1).

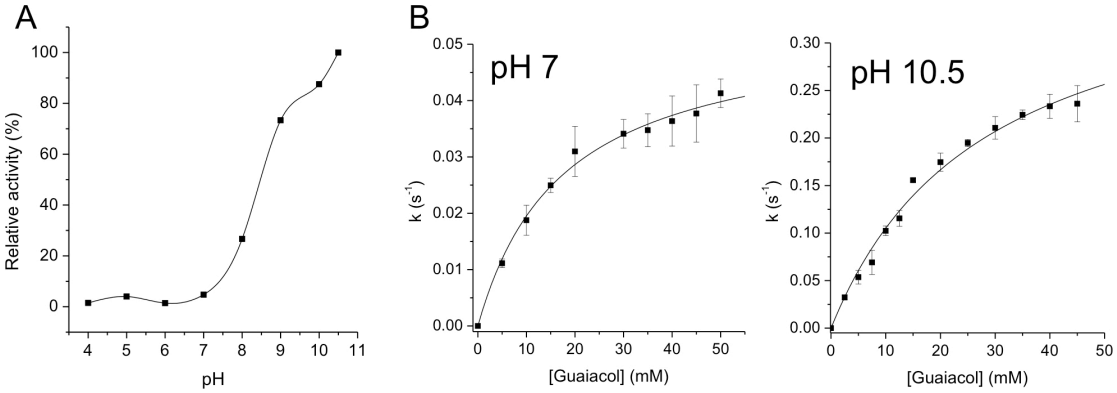


Figure 4.21: Steady-state kinetics of guaiacol oxidation by DtpA. A) pH profile of DtpA with guaiacol. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 and 20 mM boric acid, 100 mM NaCl pH 10-10.5 with 40 mM guaiacol and 1 μM DtpA. B) Steady-state kinetics of DtpA with guaiacol at pH 7 and pH 10.5. Reactions were performed using 1 μM enzyme, 0.2 mM H_2O_2 and different guaiacol concentrations at pH 7 (20 mM NaPi, 100 mM NaCl) and pH 10.5 (20 mM boric acid, 100 mM NaCl). The data were fitted by non-linear least squares Michaelis–Menten model (see Chapter 2 Equation 2.1).

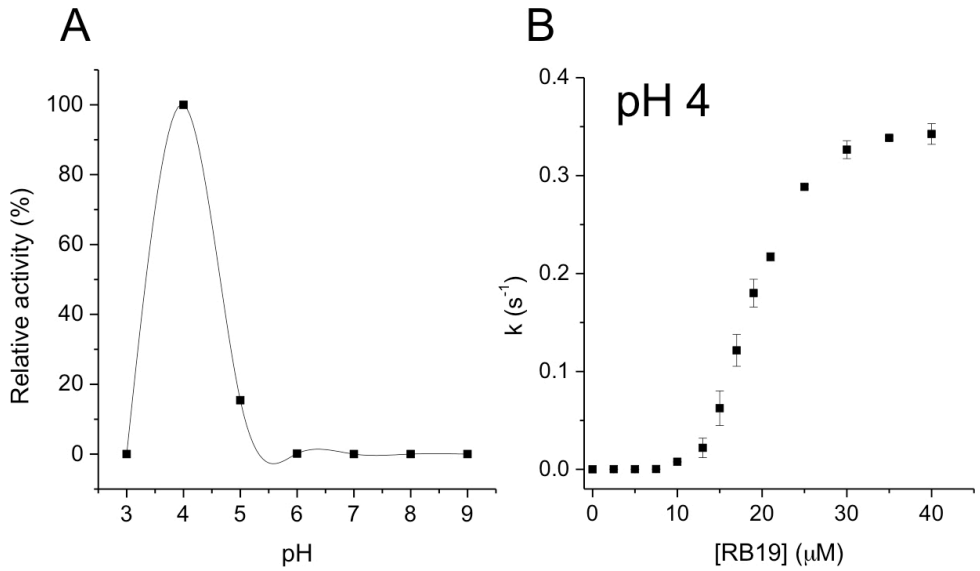


Figure 4.22: Steady-state kinetics of RB19 oxidation by DtpA. A) pH profile of DtpA with RB19. Reactions were performed in a mixed buffer system comprising, 10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl with pH values between 3-9 with 35 μM RB19 and 1 μM DtpA. B) Steady-state kinetics of DtpA with RB19 at pH 4. Reactions were performed using 1 μM enzyme, 0.2 mM H_2O_2 and different RB19 concentrations at pH 4 (10 mM Tris, MES, MOP, sodium acetate and 200 mM KCl).

Table 4.5: Steady-state kinetic parameters for DtpA. Several substrates and pH values were determined with K_m (mM), k_{cat} (s^{-1}) and k_{cat}/K_m ($M^{-1}s^{-1}$). k_{cat}/K_m value for ABTS at pH 5 is shown in orange and for $[Fe(CN)_6]^{4-}$ at pH 7 in blue corresponding to Table 4.6.

Substrate	pH	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}s^{-1}$)
ABTS	5	0.73 (0.2)	1.12 (0.15)	1.55×10^3
	7	10.4 (0.8)	0.5 (0.05)	48
$[Fe(CN)_6]^{4-}$	5	3.1 (0.6)	295 (8.8)	9.5×10^4
	7	16.3 (1.8)	109 (3.7)	6.7×10^3
RB19	4	-	$\geq$ 0.35 (0.01)	-
VA	5	16 (1.2)	0.82 (0.03)	50.9
DMP	5	-	0.0026 (0.0005)	-
	7	15.3 (1.6)	0.103 (0.02)	6.7
	9	5.6 (0.07)	0.32 (0.02)	57
Guaiacol	7	17.5 (0.3)	0.054 (0.005)	3.6
	10.5	28.1 (4.9)	0.4 (0.06)	14.3

4.3.12. Stopped-flow kinetics of DtpA with ABTS and $[Fe(CN)_6]^{4-}$

Stopped-flow kinetics were employed to assess the mechanism of substrate oxidation for ABTS and $[Fe(CN)_6]^{4-}$. Prior to this, static UV-vis titration experiments with either ABTS or $[Fe(CN)_6]^{4-}$ were carried out. These produced spectral absorption properties consistent with one reducing equivalent of the substrate required to reduce Compound I to Compound II and a second equivalent to reduce Compound II to the resting state enzyme (Figure 4.23). To initiate stopped-flow experiments one equivalent of H_2O_2 was added to DtpA then mixed rapidly with excess ABTS or $[Fe(CN)_6]^{4-}$ (Figure 4.24 and 4.25). A spectrum consistent with DtpA Compound I (soret band at ~399 nm) was initially observed which transforms rapidly into a spectrum resembling Compound II (soret band at 420 nm), followed by a slower transformation to the resting state enzyme (soret at 406 nm). Rates were obtained by monitoring the absorbance change at 420 nm. Data were fitted using double exponential equations (Equation 4.1) to obtain pseudo-first-order rate constants (k_{obs}). The k_{obs} determined for both phases were plotted against substrate concentration (Figure 4.24 and 4.25) and are linearly dependent on the concentration of substrate. The slopes yield second-order rate constants reported in Table 4.6 (Figure 4.24 and 4.25). The second-order rate constants derived from k_{obs2} for Compound II reducing to the ferric enzyme is the rate limiting step and therefore correlates to the k_{cat}/K_m values obtained from the steady-state kinetics with ABTS and $[Fe(CN)_6]^{4-}$ (orange and blue, Table 4.5 and 4.6).

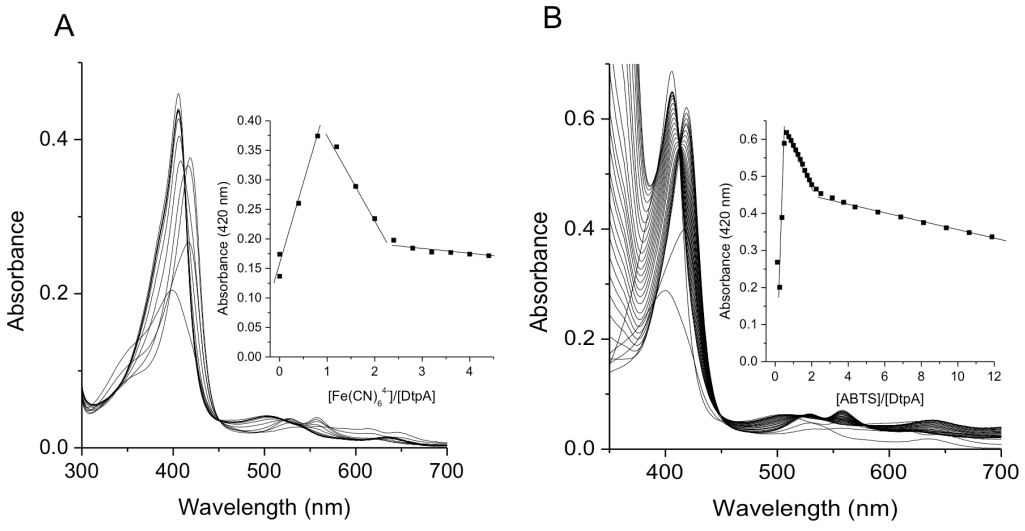


Figure 4.23: DtpA Compound I titrations with $[Fe(CN)_6]^{4-}$ and ABTS. A) UV-visible absorbance changes upon titrating $[Fe(CN)_6]^{4-}$ into DtpA (5.4 μM). Inset, absorbance changes at 420 nm plotted as a function of the ratio $[Fe(CN)_6]^{4-}$/[DtpA] concentration. B) UV-visible absorbance changes upon titrating ABTS into DtpA (8 μM). Inset, absorbance changes at 420 nm plotted as a function of the ratio of [ABTS]/[DtpA] concentrations. All spectra were measured at 20 ^{o}C in 20 mM NaPi, pH 7, 100 mM NaCl.

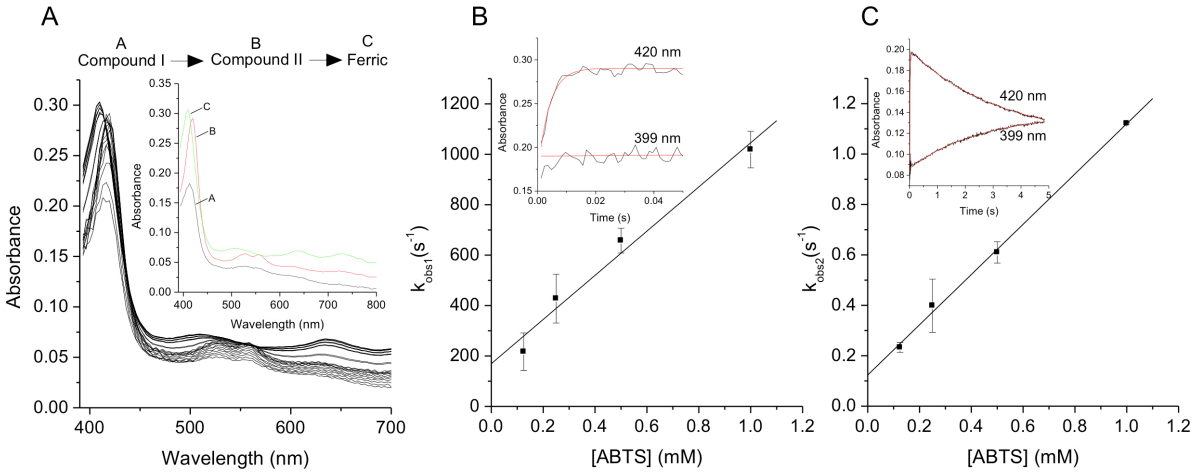


Figure 4.24: Stopped-flow kinetic analysis of the reaction of DtpA with ABTS (10 mM sodium acetate, pH 5, 150 mM NaCl (25 ^{o}C)). A) UV-visible spectral changes upon mixing DtpA (5 μM after mixing) pre-loaded with 1 equivalent H_2O_2 creating Compound I with ABTS (0.125 mM after mixing). Inset, global fitting model using a sequential mechanism of a > b > c with A, B and C, identified as Compound I, Compound II and Fe^{3+}, respectively. B) Observed rate constants (k_{obs1}) over a range of ABTS concentrations for the fast phase measured at 420 nm. Inset, time courses over the first 0.05 s at 420 and 399 nm observed on reacting DtpA (5 μM) with ABTS (0.125 mM after mixing). C) Observed rate constants (k_{obs2}) over a range of ABTS concentrations for the slow phase measured at 420 nm. Inset, time courses over 5 s at 420 and 399 nm observed on reacting DtpA (5 μM) with ABTS (0.125 mM after mixing).

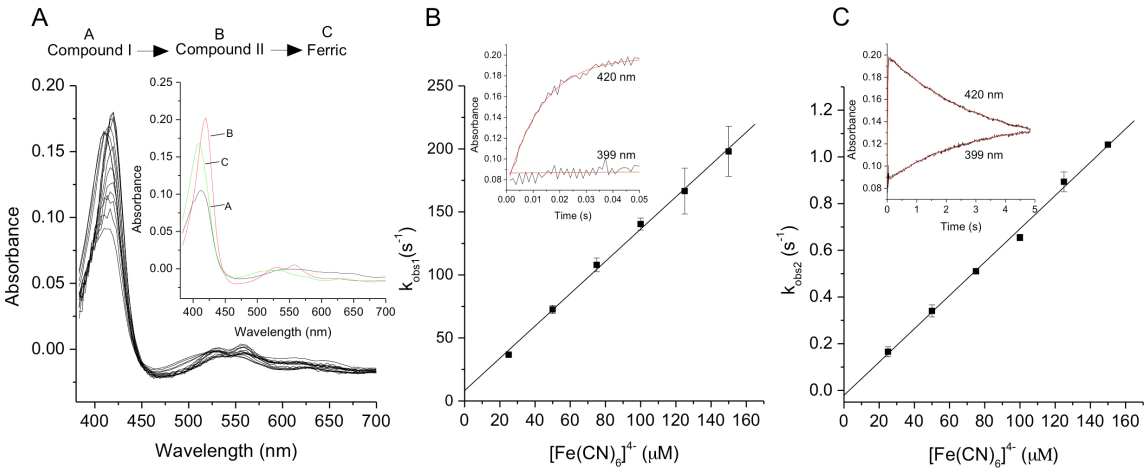


Figure 4.25: Stopped-flow kinetic analysis of the reaction of DtpA with $[Fe(CN)_6]^{4-}$ (10 mM NaPi, pH 7, 100 mM NaCl (25 ^{o}C)). A) UV-visible spectral changes upon mixing DtpA (5 μM after mixing) pre-loaded with 1 equivalent H_2O_2 creating Compound I with $[Fe(CN)_6]^{4-}$ (125 μM after mixing). Inset, global fitting model using a sequential mechanism of a > b > c with A, B and C identified as Compound I, Compound II and Fe^{3+}, respectively. B) Observed rate constants (k_{obs1}) over a range of $[Fe(CN)_6]^{4-}$ concentrations for the fast phase measured at 420 nm. Inset, time courses over the first 0.05 s at 420 and 399 nm observed on reacting DtpA (5 μM) with $[Fe(CN)_6]^{4-}$ (125 μM after mixing). C) Observed rate constants (k_{obs2}) over a range of $[Fe(CN)_6]^{4-}$ concentrations for the slow phase measured at 420 nm. Inset, time courses over 5 s at 420 and 399 nm observed on reacting DtpA (5 μM) with $[Fe(CN)_6]^{4-}$ (125 μM after mixing).

Table 4.6: Transient kinetic parameters for DtpA. Second-order rate constants determined using stopped-flow kinetics, orange shows the second order rate constant for Compound II to resting state DtpA using ABTS and blue using $[Fe(CN)_6]^{4-}$ corresponding to those determined using steady-state kinetics in Table 4.5.

	k_{obs1} (M^{-1} s^{-1})	k_{obs2} (M^{-1} s^{-1})
ABTS[a]	8.8 (± 0.44) x 10^5	1.02 (± 0.07) x10^3
$[Fe(CN)_6]^{4-}$ [b]	1.28 (± 0.07) x 10^6	7.1 (± 0.15) x10^3

[(a)] Experiments performed at pH 5

[(b)] Experiments performed at pH 7

4.3.13. DtpA acts as a peroxidase in the presence of GlxA

Recent *in vivo* studies in *S. lividans* have provided evidence that DtpA is required for GlxA maturation and morphogenesis [171]. To test whether DtpA and GlxA have a synergistic role, *in vitro* studies were performed. No evidence for complex formation between DtpA and GlxA was observed by size-exclusion chromatography (Appendix 2.4). This method favours the detection of tight interactions (K_d < μM) so a weaker transient interaction cannot be ruled out. Upon mixing equimolar amounts of resting state DtpA and GlxA, no change in the DtpA absorbance spectrum could be observed. As discussed in Chapter 2, GlxA is relatively inactive with substrates readily turned over by fungal Gox [166], with the highest activity observed for glycoladehyde. After addition of glycolaldehyde to GlxA:DtpA, the spectrum of DtpA transformed immediately (< 1 min) into a spectrum consistent with Compound I formation, although slight perturbations are observed (soret band at 403 nm and peaks at 536, 576, 614 and 645 nm as compared to 399, 530, 557, 614 and 644 nm) (Figure 4.26.A), providing direct proof that DtpA is a true peroxidase and GlxA is producing H_2O_2, being used by DtpA. Over a period of time (> 20 min) Compound I is transformed to a new species with maxima at 420, 539, 571, 635 and 696 nm (Figure 4.26.A). These spectral properties are not consistent with the formation of the one electron reduced Compound II peroxidase form *i.e.* a ferryl species (Fe^{IV}=O), which in the case of DtpA has maxima of 419, 528, 557, 621 and 728 nm (Figure 4.26.A and Table 4.4). Instead, the new species has a spectrum that resembles an oxy-ferrous form suggesting that under the conditions employed DtpA eventually becomes fully reduced (Figure 4.26.A). The same spectral species is formed upon addition of excess glycolaldehyde to DtpA Compound I generated through addition of 1 equivalent of H_2O_2 and on addition of excess glycolaldehyde to the resting state ferric enzyme (See Appendix Figure 2.5). This illustrates that glycolaldehyde can act as a reductant with DtpA, it does not form Compound I, and therefore when GlxA is present, glycolaldehyde is its substrate leading to the production of H_2O_2 (Figure 4.26) [186].

One possibility that could arise is that GlxA activity is dependent on the peroxidase used. Therefore a coupled peroxidase assay was performed comparing DtpA to HRP. Figure 4.26.B compares the turnover rates (k s^{-1}) for various substrates used in Chapter 2 for GlxA in the presence of HRP or DtpA as the peroxidase. The rates observed were similar to those previously reported for GlxA in the presence of guaiacol and HRP. It is also clearly apparent that for all the substrates tested no boost in activity occurs when DtpA replaces HRP (Figure 4.26.B). It does however further substantiate that DtpA can in fact utilise the H_2O_2 produced by GlxA acting as an effective peroxidase.

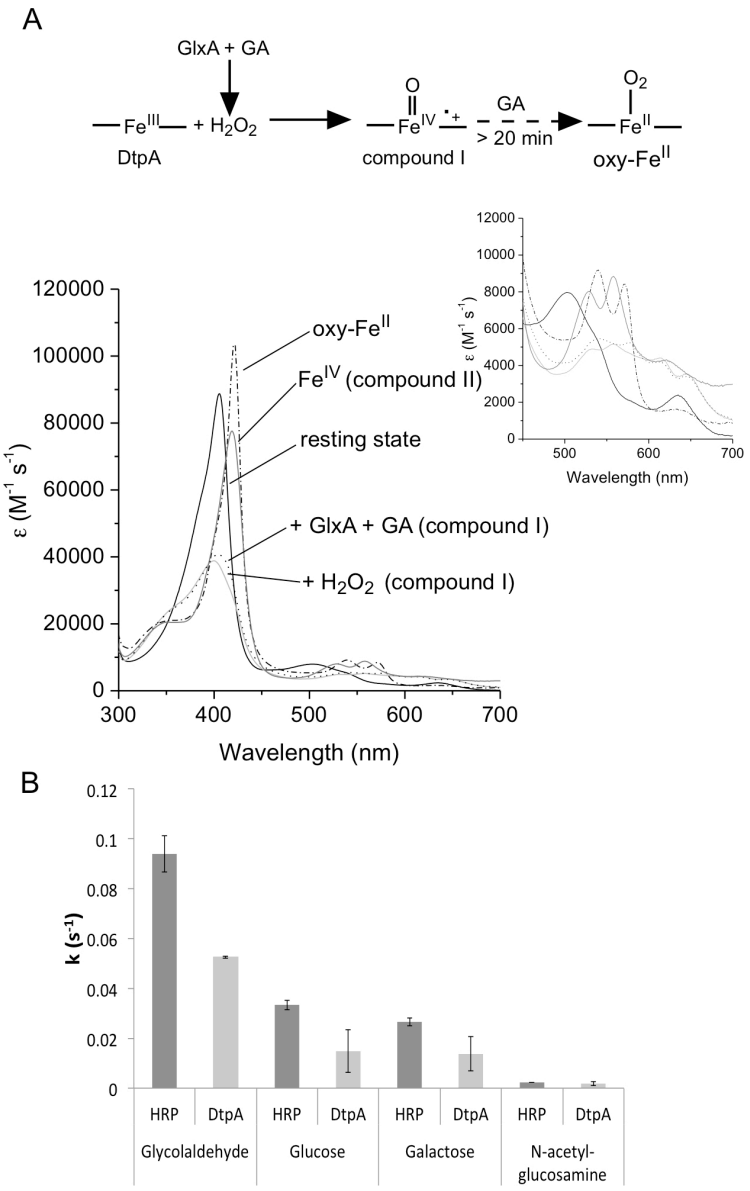


Figure 4.26: Peroxidase activity of DtpA. (A) Static UV-visible spectra of various haem oxidation states of DtpA (20 mM sodium phosphate pH 7, 100 mM NaCl) as illustrated by the reaction scheme. Addition of H_2O_2 (light grey solid line), or addition of GlxA and glycolaldehyde (GA) (dotted line) to resting state (Fe^{III}) DtpA leads to a Compound I spectrum (formed within ~ 1 min). Over time (> 20 min) the compound I species is converted to a species with an oxy- ferrous like spectrum (dashed-dot line) in the presence of GA. Note that no Compound II species is observed in this process. The Compound II spectrum shown was generated by formation of DtpA Compound I followed by addition of $[Fe(CN)_6]^{4-}$ (dark grey solid line). The inset shows a zoomed-in region of the weaker intensity absorbance bands. (B) Turnover rates (k) for GlxA with four different substrates (30 ^oC) in the presence of HRP or DtpA determined through the subsequent oxidation of ABTS. Error bars indicate the standard deviation from triplicate experiments.

4.4. Discussion

4.4.1. DtpA is a DyP-type peroxidase

Initial over-expression of DtpA in *E. coli* yielded small quantities, as previously reported for other DyPs [259]. Linde *et al* describe a refolding protocol for DyP from *A. auricula-judae,* extracting higher quantities of soluble protein extracting it from insoluble inclusion bodies [259]. This refolding technique was not necessary for DtpA but instead the use of pure CO gas during over-expression was used to significantly boost the expression of soluble DtpA (Figure 4.6). Addition of CO directly into the cultures helps to stabilise the enzyme and increases the production of soluble protein expression, which has been described for the over-expression of globins [260, 261]. The ability of CO to bind to DtpA was observed through UV-visible spectroscopy displaying signature CO haem absorption peaks (Figure 4.8 and Table 4.4) and also detected by the presence of a red pellet following growth and a bright red supernatant after cell-lysis (Figure 4.6). Upon removal of CO a transformation in colour from red to brown is observed, typical for a b-type haem peroxidase.

Sequence alignments of DtpA with other well-characterised DyPs clearly point towards DtpA being a member of the DyP-type peroxidase family (Figure 4.5 and 4.6). Spectral characteristics of resting state, Compound I and Compound II DtpA are compared to other DyPs and plant peroxidases listed in Table 4.7. In agreement with spectral properties previously reported for other DyPs, DtpA retains characteristic spectral transitions [180, 186, 248-250]. The soret band is the representative absorption peak of peroxidases and contains a peak maximum of 406 nm for resting state DtpA [180, 262, 263]. Upon addition of H_2O_2 the soret band shifts to 399 nm, a signature typical of Compound I formation, as seen for *Bad*DyP and is comparable to HRP [180, 248]. One electron reduction leads to Compound II formation with a soret band with peak maxima at 420 nm also comparable to many DyPs and plant peroxidases, only differing to *Bad*DyP which surprisingly reports a peak maxima at 399 nm (Table 4.7) [186]. Fundamentally, the catalytic cycle of DyPs is the same as seen in other peroxidases. Interestingly, when DtpA Compound I is formed, in the absence of a reducing substrate it returns to resting state enzyme within 10 min ($t_{1/2}$ ~ 2.5 min) (Figure 4.11), whereas Compound I of other peroxidases such as LiP converts first to a Compound II spectrum within 1 min before reverting to the resting state enzyme [183, 186, 204, 211, 250, 254]. It is also apparent that in the presence of excess H_2O_2, Compound II is converted to Compound III, an inactivated state, which has been reported to decrease peroxidase activity (Figure 4.8 and Table 4.4) [251, 252]. Anaerobic reduction of DtpA with $Na_2S_2O_4$ altered the spectral shape significantly, with the soret band decreasing and shifting towards 422 nm and a broad peak at 559 nm, similarly observed with other haem peroxidases (Figure 4.8 and Table 4.4) [174, 180].

Table 4.7: Absorption maxima of resting state and catalytic intermediates of DtpA and other haem peroxidases. (*sh)* stands for shoulder.

	pH	Resting state	Compound I	Compound II	Reference
DtpA	7.0	406, 502, 635	399, 350 (*sh*), 530, 557, 614, 644	419, 528, 557, 621, 728	This work
*Bad*DyP	3.2	406, 506, 636	401, 530, 556, 615, 644(*sh*)	399, 529, 555, 615, 644(*sh*)	[186]
*Tc*DyP	7.8	406, 506, 544(*sh*), 567(*sh*), 624	407, 512, 546, 623, 648(*sh*)	416, 524, 555, 623	[189]
DypA	7.5	408, 502, 632	-	419, 528, 557, 619	[204]
DypB	7.5	404, 503, 634	397, 340 (*sh*), 580, 613, 648	-	[204]
HRP	6.6	403, 498, 640	400, 525 (*sh*), 577, 622 (*sh*), 651	420, 527, 555	[248]
LiP	6.0	408, 496, 630	408, 550, 608, 650	420, 525, 556	[250]
MnP	4.5	406, 502, 632	407, 558, 617, 650	420, 528, 555	[249]

A second-order rate constant of 8.9 (± 0.25) x 10^6 $M^{-1}s^{-1}$ was determined for Compound I formation (Figure 4.10). Recently, Compound I was observed for the first time in an A-type DyP, *Tc*DyP, displaying a second order rate constant of 5.92 x 10^6 $M^{-1}s^{-1}$ [189]. This is comparable to DtpA and is ~30-fold higher than found in the class B DypB [189, 209]. Furthermore, a large cavity can be identified in the distal face of the haem centre for DtpA a feature previously been reported to be essential for H_2O_2 binding (Figure 4.27) [258]. Four residues are clearly involved in facilitating the binding of H_2O_2 in the distal pocket of DyPs including, Asp, Arg, Leu and Phe. Among all DyP family members (classes A, B, C and D) these four residues are conserved. Figure 4.27 shows the structural arrangement of these residues in DtpA (class A), EfeB (class A), TyrA (class B), DyP2 (class C) and *Bad*DyP (class D). It is clearly apparent that these residues superimpose in all four DyP-type subclasses. The only slight variation is the position of the Asp residue in DyP2 from *Amycolatopsis sp* in which the side chain is flipped ~ 90° (Figure 4.27) [196]. Recently, this Asp residue was reported to be involved in a 'swinging' mechanism for Compound I formation, observed using cyanide as a structural analogue of H_2O_2 [258]. H_2O_2 binds to the DyP and the Asp residue 'swings' into an optimal position so that it can accept a proton from H_2O_2. Once Compound I is formed the Asp 'swings' back into its original position [258]. This movement in Asp has also been observed in the crystal structure of *Aau*DyP postulating a role not only in Compound I formation but as a 'gatekeeper' that allows the passage of substrate molecules larger than H_2O_2 (*i.e.* small organic molecules), allowing direct oxidation of substrates [197].

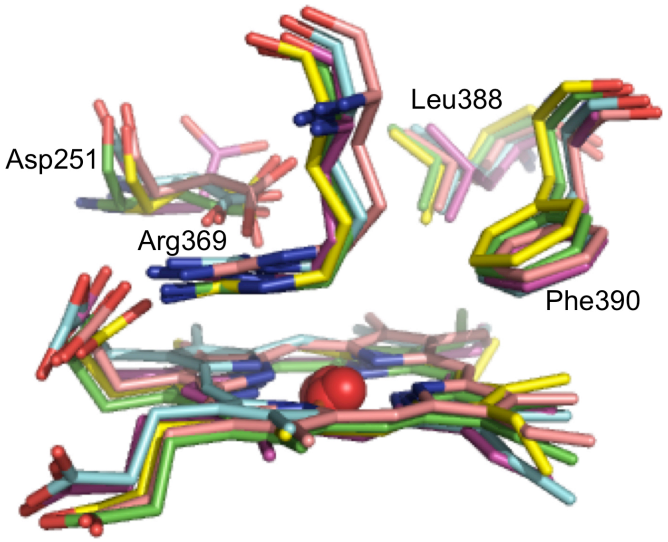


Figure 4.27: Comparison of the H_2O_2 binding site in DyP-type peroxidases. DtpA from *S. lividans* in green, EfeB from *E. coli* an A-type DyP in yellow (PDB entry 2WX7), TyrA from *S. oneidensis* a B-type DyP in pink (PDB entry 2IIZ), DyP2 from *Amycolatopsis sp* a class C DyP in purple (PDB entry 2G2C) and *Bad*DyP from *B. adusta* a D-type DyP shown in blue (PDB entry 2D3Q) [196, 258].

4.4.2. Unique loops differentiate the three DyPs in Streptomyces

Analysis of the *S. lividans* genome [4] led to the identification of three DyP-type peroxidases. Two class A-type peroxidases, SLI4211 (DtpA) and SLI2602 an EfeB homologue and a class B-type peroxidase SLI7409. SLI4211 and SLI2602 contain Tat signal sequences with SLI2602 containing a clear signal peptidase cleavage site, whereas, SLI7409 contains no signal sequence and is thus predicted to reside in the cytoplasm. This leads to the conclusion that the three DyP-type peroxidases in *S. lividans* are strategically positioned within the cell, *i.e.* cytoplasmic (SLI7409), membrane associated (SLI4211) and secreted in the extracytoplasmic environment (SLI2602). The X-ray structural coordinates of the homologs of these three *S. lividans* DyPs from *S. coelicolor* (SCO3963 (PDB entry 4GT2) (DtpA), SCO7193 (PDB entry 4GU7) (SLI7409) and SCO2276 (PDB entry 4GRC) (SLI2602)) have been deposited to the PDB without any associated publication. It is clear from sequence alignment (Figure 4.28) that these DyPs share a high level of sequence identity. DtpA shares 99.5 % sequence identity to SCO3963 with only two amino acid changes as indicated in Figure 4.28. The first amino acid change is an Arg in SCO3963 (Arg^{79}) to a Gln residue in DtpA (Gln^{79}) with this residue located on the periphery of the enzyme so is unlikely to cause any affects to the enzyme. The second change is a Glu in SCO3963 (negative side chain, Glu^{168}) to a Lys in DtpA (positive side chain, Lys^{168}) but again this residue is present on the exterior of the protein and is not likely to cause any significant effect. Therefore, for structural comparisons DtpA was compared with SCO2276 and SCO7193. DtpA shares 25.5 % sequence identity to SCO7193 and 45.5 % to SCO2276 (SCO2276 shares 23 % identity to SCO7193). This is as expected because DtpA and SCO2276 are class A-type DyPs and SCO7193 is a class B-type peroxidase. This identity is comparable to the RMSD values for all Cα atoms when the structures are overlaid using superpose, with DtpA and SCO7193 giving a value of 1.97 Å and DtpA with SCO2276 a value of 1.18 Å (SCO2276 and SCO7193 give a value of 1.18 Å) [168, 169].

Figure 4.28: Sequence alignment of DtpA from *S. lividans* with SCO3963, SCO7193 and SCO2276 from *S. coelicolor*. Clustal Omega sequence alignment with completely and partially conserved residues coloured in dark to light blue, respectively. Symbols identify the position of important haem pocket residues, with the * showing the position of Arg^{369}, the ^ the position of Asp^{251} and + the position of His^{353} (DtpA numbering). The conserved DyP motif GXXDG is shown by a dashed box. The conserved Trp residue with possible importance in LRET is shown by the " symbol (Trp^{301} in DtpA) and the Tyr residue (Tyr^{374} in DtpA) also possible to play a role in LRET is shown by a " symbol. The changes in amino acids between DtpA and SCO3963 are shown by the : symbol. The Tat signal sequences for DtpA and EfeB are not included in the alignment.

Notably, the most significant difference between all three DyPs is the length and structure of the loop regions (Figure 4.29). SCO7193 displays smaller loops compared to SCO2276 and DtpA (Figure 4.29). These differences can be observed through surface representations showing the flat surface of SCO7193 compared to DtpA and SCO2276 (Figure 4.30). Interestingly, the switch loop also differs between the DyPs; this loop is formed by 32 amino acids in DtpA (Asn^{164} - Asp^{195}) as opposed to 29 in SCO2276 (Lys^{220}-Glu^{248}) and only 20 in SCO7193 (Phe^{139}- Thr^{158}) (Figure 4.29.E). A large tunnel is formed by this switch loop in EfeB and has been proposed to provide a potential cofactor/substrate binding site [203]. This loop displayed major conformational changes between the structures of apo-EfeB (PDB entry 2Y4D, unpublished) and EfeB in complex with protoporyhrin IX (PDB entry 2Y4E, unpublished) suggesting its role in substrate/cofactor binding [203]. Many of the residues identified in the Tunnels in the three DyPs (*vide infra*) are part of this flexible switch loop and therefore this variation in loop length maybe important in substrate access, specificity and turnover. Although differences in sequence and tertiary structure are apparent, the main haem pocket residues of the three DyPs align precisely within sequence alignments and in structural comparisons (Figure 4.28 and 4.29.D/E).

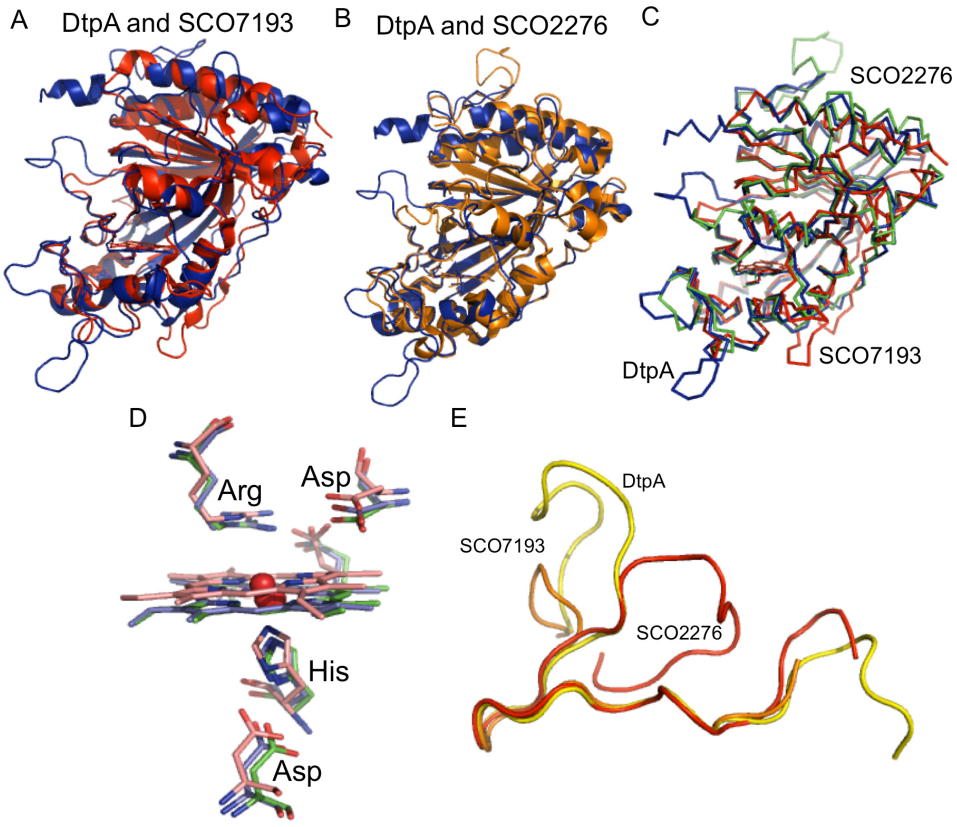


Figure 4.29: Structural comparison of DtpA with SCO7193 (PDB entry 4GU7) and SCO2276 (PDB entry 4GRC) from *S. coelicolor*. A) Tertiary structural alignment of DtpA (blue) with SCO7193 (red). B) Tertiary structural alignment of DtpA (blue) with SCO2276 (orange). C) Ribbon representation of DtpA (blue) with SCO7193 (red) and SCO2276 (green). D) Haem environment of DtpA (purple) with SCO7193 (pink) and SCO2276 (green). E) Switch loop of DtpA (yellow) residues Asn^{164} - Asp^{195} with SCO2276 (red) residues Lys^{220}-Glu^{248} and SCO7193 (orange) residues Phe^{139}-Thr^{158}.

Surface electrostatics of the three DyP enzymes are shown in Figure 4.30. CAVER was used to analyse the tunnels of the *S. coelicolor* enzymes as previously described for DtpA (section 4.3.9, Figure 4.14). SCO7193 does not display any major tunnels, and instead a large cavity can be identified ~1.8 Å from the surface (Figure 4.30.D). The entrance to this large cavity can be seen by the surface electrostatic representation, in which the haem cofactor can be clearly observed with a large opening of ~10 x 4 Å. SCO7193 exhibits a significantly high proportion of negative charge (pI 4.6), which form distinct surface patches (Figure 4.30A). Notably, one of these patches surrounds the entrance to this large cavity leading to the haem pocket (Figure Figure 4.30.A). SCO2276 and DtpA possess two clear tunnels leading from the solvent to the haem environment (Figure 4.30.E and F). Tunnel I and II identified in SCO2276 are similar in length of ~ 17 Å, with Tunnel II in a similar position to Tunnel II of DtpA (Figure 4.30.E). The entrance to Tunnel II of SCO2276 can be seen in the electrostatic map with a small opening to the haem pocket but with a largely positive area as also seen in DtpA (Figure 4.30.B and C). Tunnel I of SCO2276 is different to Tunnel I of DtpA and exits in an opposite direction to Tunnel II (Figure 4.30.E). Tunnel II in all *Streptomyces* structures analysed bares striking resemblance to the Tunnel II in *Bad*DyP, whereas Tunnel I appears to be more variable [202]. Furthermore, analysis of pI values from various subclasses of DyP enzymes would suggest that classes B, C and D contain more acidic values than the A subclass. The variation in tunnels and pI values within this class of enzymes hint at varied substrate specificity.

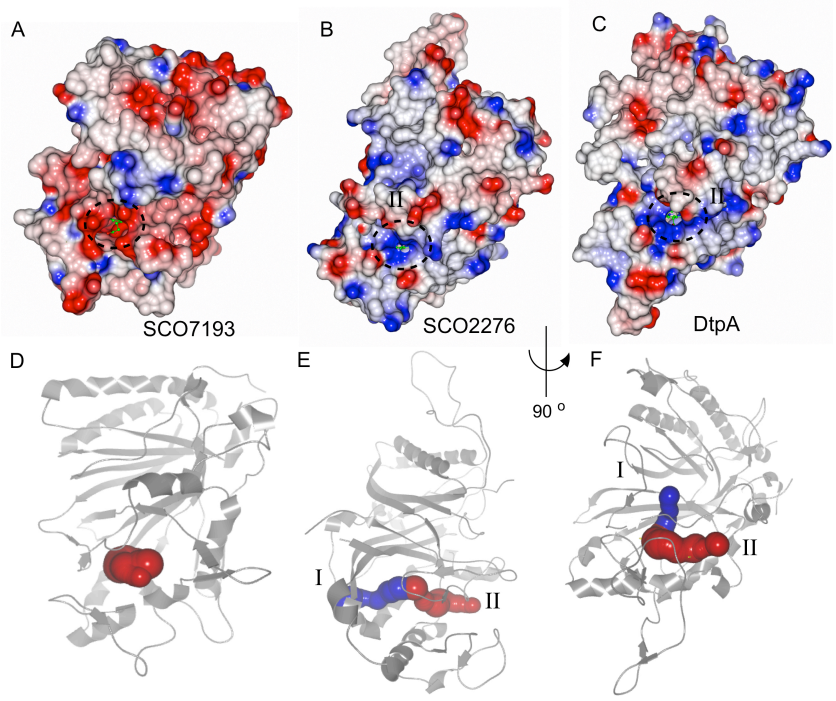


Figure 4.30: Surface electrostatic and tunnel comparisons of DtpA with SCO7193 (PDB entry 4GU7) and SCO2276 (PDB entry 4GRC) from *S. coelicolor*. Electrostatic surface representation of A) SCO7193, B) SCO2276 and C) DtpA. Surface-to-haem Tunnels in D) SCO7193, E) SCO2276 and F) DtpA identified using the programme CAVER [106].

4.4.3. The biologically relevant substrate of DtpA is unknown

The physiological substrate of all DyP-type peroxidases is as yet unknown. Using a range of reducing substrates such as those illustrated in Figure 4.16 it is apparent that A and B-type DyPs show lower activity compared to C and D classes [40, 190, 204]. However, the A-class *Tc*DyP has proved to be an exception, with high catalytic efficiency determined with ABTS and AQ dyes comparable to the more active D-type fungal DyPs [189]. From the substrates tested at varying pH it is apparent that DtpA is not operable at a specific pH value and instead seems to vary depending upon the substrate used, an observation also made for other related enzymes (Figures 4.17-4.22) [264]. However, this is unusual as all other reported DyPs characterised exhibit significantly higher activity under very low acidic pH conditions, with a considerably lower pH optimum (pH ~ 2) for the substrate VA reported [198, 259]. DtpA appears to be optimal at pH 5 with ABTS, $[Fe(CN)_6]^{4-}$, VA and pH 4 for RB19 (Figures 4.17 - 4.19 and 4.22) but is also clearly active at alkaline pH values (9 - 10.5) with the phenolic substrates DMP and guaiacol (Figures 4.20 and 4.21). The highest activity is observed with $[Fe(CN)_6]^{4-}$ (Figure 4.17). This small inorganic molecule has not previously been characterised as a substrate with DyPs. However with the exception of *Tc*DyP, the catalytic efficiency shown with $[Fe(CN)_6]^{4-}$ is in line with other A-type DyPs [189, 191, 204]. ABTS is a more commonly used peroxidase substrate and for DtpA the activity is an order of magnitude lower than $[Fe(CN)_6]^{4-}$ but still in line with other A-class DyPs (Figure 4.18) [191, 204]. The small molecules VA, DMP and guaiacol have also been used as substrates with other DyPs and tend to give lower activities than for example ABTS [174, 189, 191, 198, 259]. In fact for *Tc*DyP no activity was found for DMP and VA [189]. However, for DtpA activity was observed for VA, DMP and guaiacol (Figures 4.20 – 4.21) but as was the case with guaiacol for *Tc*DyP, significantly lower k_{cat}/K_m values compared to ABTS are observed (Table 4.5). An explanation for this may be due to the neutrality of these substrates rather than size.

As the name suggests DyPs have the ability to oxidise and decolourise some industrially used AQ dyes using a two-electron oxidation process [208, 229]. RB19 is a common dye substrate, which has received much attention and was therefore tested. DtpA displayed the ability to decolourise RB19 at an optimum pH of 4, similar to other DyPs (Figure 4.22) [192, 198, 246, 247]. However, the steady state kinetics of DtpA suggested a more complex pattern of oxidation compared to the substrates discussed above (Figure 4.22.B). This precluded the ability to obtain a K_m value for the reaction and thus make a true comparison of the catalytic efficiency of RB19 with other DyPs for which activity has been reported (Table 4.5). Interestingly, the behaviour observed here with DtpA has similarities to the D-type *Aau*DyP enzyme [197, 244, 259]. Here the authors observe biphasic (sigmoidal) kinetics and interpret this to mean the existence of more than one substrate oxidation site with different turnover numbers and substrate affinities [244, 259]. The high-turnover site for oxidation ($k_{cat} > 200\ s^{-1}$) was assigned to Trp^{377}, a conserved surface exposed residue, since

activity was absent in a W377S variant. Whereas, the low-turnover site/s (RB19 k_{cat} ~ 20 s^{-1}) was proposed to correspond to the haem access-channel, since activity was decreased when the haem channel was occluded by a G169L variant [244]. A similar behaviour has been reported for *P. eryngii* VP that can oxidise some phenols and dyes (such as DMP and ABTS) at a high and low affinity site [265], also found to localize at an exposed catalytic Trp residue [155] and at the main haem access channel [266], respectively.

An alternative suggestion to the steady-state kinetics observed in this Chapter is that at low RB19 concentrations the two electron oxidation required to form the decolourised product is not occurring. The RB19 may still bind to the DtpA but only enabling the first electron oxidation to occur giving a product with an extinction coefficient that is no different to the unreacted substrate, therefore giving rise to the first part of the kinetic profile in Figure 4.22.B. As the dye concentration increases the already bound RB19 either modifies the binding site to enable for the two-electron oxidation to occur or the RB19 stacks on top of the already bound dye. Further investigations using stopped-flow kinetics may help to elucidate this complicated scenario.

4.4.4. Long range electron transfer (LRET)

The majority of dyes, which are typically oxidised by DyPs, such as RB19, cannot directly interact with the buried haem and thus a LRET scenario has been proposed. DyPs possess several aromatic residues, which are potentially able to house reactive radical species and facilitate LRET [197, 267]. Aromatic surface residues involved in a LRET route were first reported in *Phanerochaete chrysosporium* LiP [268] and *Pleurotus eryngii* VP, [155] then later identified in many LiPs and VPs through sequence analysis [269]. Compound I is formed by Fe^{IV}=O and a radical cation, that can migrate, resulting in a protein-based radical cation [270]. Thus, in such a circumstance, a protein-based substrate interaction site occurs [197, 270, 271]. DyPs originating from the fungus *basidiomycete* (class D DyPs) contain on average 9 Tyr and 6 Trp residues per protein molecule [197, 208, 267]. DtpA (class A) contains only 3 Tyr and 6 Trp (Figure 4.31.A.), compared to EfeB (class A; 10 Tyr, and 4 Trp), TyrA (class B; 12 Tyr and 2 Trp) and DyP2 (class C DyP; 10 Tyr and 5 Trp), which all contain a significantly higher number of Tyr residues. High numbers of Tyr residues in the fungal *basidiomycete* class D DyPs contrasts to a low number in the ligninolytic peroxidases such as LiPs and VPs from the same group of fungi. The absence of Tyr residues in these ligninolytic peroxidases is thought to protect these enzymes from oxidative inactivation by coupling mechanisms or other reactions between Tyr phenoxy radicals [272]. It is therefore believed that DyPs with high numbers of Tyr residues are present in less oxidative environments [208, 272]. However, DtpA contains a significantly lower number of Tyr residues and will therefore be less prone to oxidation reactions in a more oxidative environment. Similarly, the three DyPs identified in *S. coelicolor* all contain low numbers of Tyr residues with SCO7193 and SCO2276 containing 5 and SCO3963 (homologue to DtpA)

containing 3. Recently the aromatic surface residues of *Aau*DyP have been extensively studied for their ability to house a reactive radical [244]. A mixed Tyr^{337} and Trp^{377} radical was detected by EPR with QM/MM (quantum mechanics/molecular mechanics) simulations describing a LRET pathway from Trp^{377} to the haem [197, 267, 273]. Trp^{377} is conserved amongst *basidiomycete* DyPs and a W377S variant, abolished the EPR radical and decreased enzymatic activity compared to native protein, revealing the central role of this residue in catalysis [244]. When *Aau*DyP (PDB 4W7J) was superposed with DtpA, differences in the structures are apparent with an RMSD of 2.1 Å for all Cα atoms [168, 169] (Figure 4.28.B). Interestingly, Trp^{377} shown to house a radical species and assist in LRET in *Aau*DyP is not in close proximity to any of the 6 Trp residues in DtpA. The closest Trp residue to the haem centre in DtpA is Trp^{301}, which is 11.2 Å away, which is an identical to the distance of Trp^{377} from the haem in *Aau*DyP. Trp^{301} also aligns with Trp residues from EfeB and TyrA (Figure 4.28 " symbol), and is the only Trp residue, which aligns in DtpA with SCO2276, SCO7193 and SCO3963 (Figure 4.28 " symbol and Figure 3.1.C). Intriguingly, Tyr^{337} in *Aau*DyP overlays exactly with Tyr^{374} in DtpA however, its role within *basidiomycetes* to form a catalytic radical has been disregarded (: symbol in Figure 4.5, " in Figure 4.28 and Figure 4.31.B) [244].

DtpA displays a classical Compound I UV-visible spectrum after the addition of H_2O_2. This illustrates that the radical likely resides on or very close to the prophyrin ring. This is also the case in reported Compound I absorption spectra of other DyPs (Table 4.7) [180, 186, 248]. However, a protein radical species has been identified through EPR spectroscopy in other DyP-type peroxidases, albeit at cryogenic temperatures, suggesting that once Compound I is formed the radical migrates (as discussed above). In *Aau*DyP this location has been assigned to Trp^{377} not present in DtpA. Thus, future work should be focused on using EPR spectroscopy to resolve the radical migration properties of DtpA.

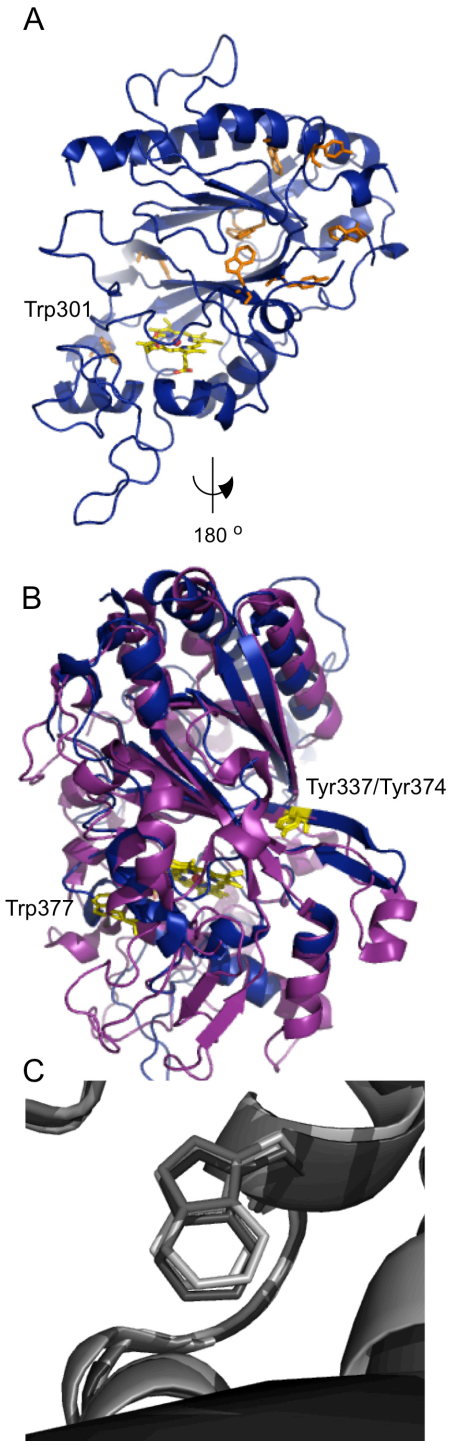


Figure 4.31: Surface aromatic amino acid comparisons. A) DtpA shown by a cartoon representation in blue with all the Tyr and Trp residues shown as orange sticks and the haem in yellow with Fe as a red sphere. B) Superposition of DtpA (blue) and *Aau*DyP (PDB 4W7J) from *Auricularia auricula-judae* with important aromatic residues depicted as yellow sticks and the haem in yellow with Fe as a red sphere. C) Superposition of DtpA Trp^{301} with SCO2276 (PDB entry 4GRC) and SCO7193 (PDB entry 4GU7).

4.4.5. DtpA is required for GlxA maturation and morphogenesis in S. lividans

The extracytoplasmic Cu chaperone proteins Sco and ECuC facilitate a Cu-trafficking pathway involving Sco receiving a Cu(I) ion from ECuC [38]. Recently, it has been reported that these chaperones also play a role in the maturation of GlxA, linking this to the Cu-dependency of *Streptomyces* development [171].In Chapter 3 Western blotting using GlxA variants identified that Cu and the Cys-Tyr cross-link are required to form the fully mature protein (see Figure 3.15 in Chapter 3). A recent collaborative study has shown from Western blotting of mycelium extracts that immature GlxA (absence of the Cys-Tyr cross-link) is predominately present in the deletion mutants Δ*sco* and Δ*ecuc* [171]. This corroborates the initial suggestion in Chapter 2 that Sco is involved in the maturation of GlxA (section 2.3.13, Chapter 2). Surprisingly, a Δ*dtpA* null-mutant also blocks aerial hyphae and pellet formation under Cu limiting conditions as also observed for the Δ*sco* mutant, with both these phenotypes becoming restored upon addition of exogenous Cu (section 2.3.13 Chapter 2). Western blotting has also shown that GlxA maturation is perturbed in the Δ*dtpA* strain with no mature GlxA identified [171]. This implies that DtpA has a role in the maturation of GlxA. Its exact role remains open but it may be that DtpA oxidises Sco-bound Cu(I) to Cu(II). This would not only explain why GlxA maturation is impaired in both *dtpA* and *sco* mutants, but also why the C*c*O activity is reduced in the *dtpA* mutant (Fujimoto *et al.*, 2012 and our unpublished results). Sco proteins bind both Cu(I) and Cu(II) and Cu transfer to acceptor proteins may depend on the oxidation state of the metal [38]. A role for DtpA in oxidizing metal ions would thus be very similar to the function of the *B. subtilis* DyP-type peroxidase EfeB, which oxidizes Fe(II) to Fe(III) before cellular uptake [274].

Based on results presented here, it has been shown that DtpA functions as a peroxidase in the presence of GlxA and glycolaldehyde, utilising the H_2O_2 produced by GlxA. Therefore it may be speculated that DtpA acts in tandem with GlxA removing this potentially toxic reactive oxygen species from the extracellular environment. In addition, results in Chapter 2 highlight the roles of CslA and GlxA in the synthesis, modification and deposition of a β-(1-4) glycan at the hyphal tips of *S. lividans*. Taken together these results have led to the proposed model for the Cu-dependent morphogenesis pathway and glycan processing in the hyphal tips as shown in Figure 4.32.

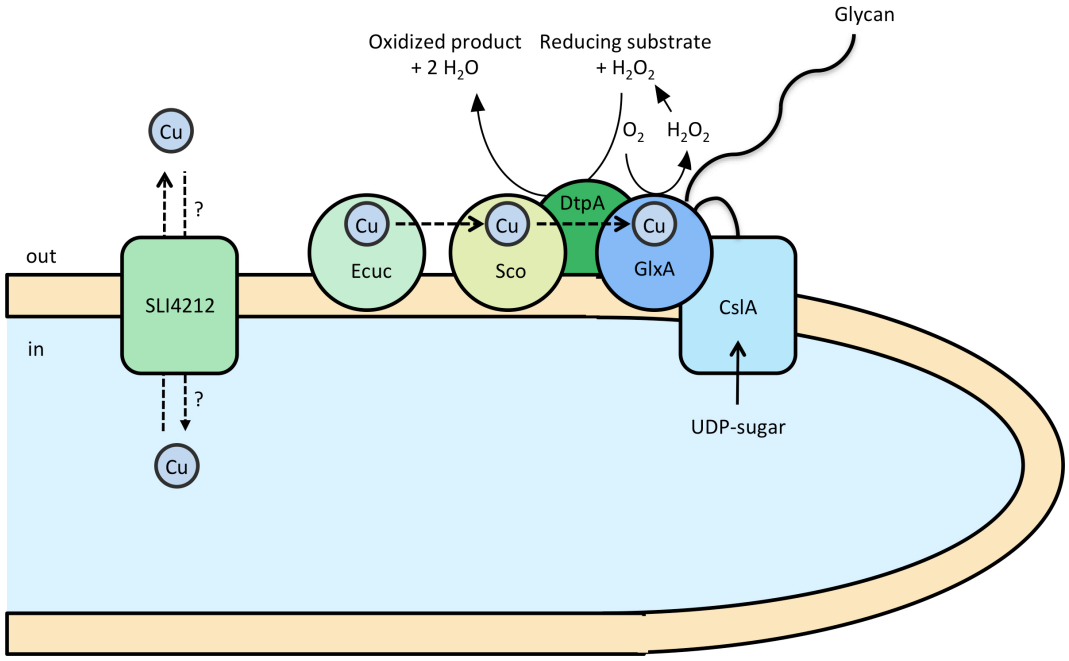
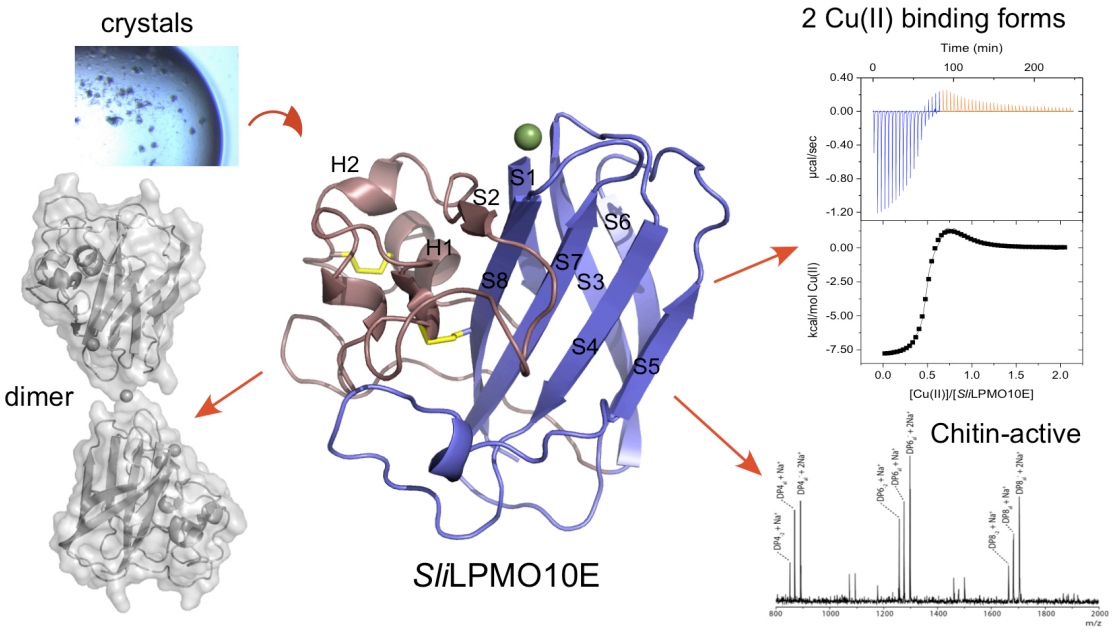


Figure 4.32: Proposed model for the Cu-dependent morphogenesis pathway in hyphal tips of *S. lividans*. Mature GlxA requires the incorporation of a Cu ion and the formation of a Cys-Tyr covalent bond. GlxA receives its Cu from the extracellular chaperone Sco, which in turn receives Cu from the chaperone ECuC. The putative Cu transporter SLI_4212 may be involved in shuttling Cu ions over the membrane but has not yet been determined. DtpA is required for maturation of GlxA and removes the reactive H_2O_2 generated by GlxA while oxidizing its substrate. Mature GlxA acts cooperatively with the cellulose synthase-like protein CslA in formation of an extracellular glycan.

Chapter Five

Structural and kinetic characterisation of a lytic polysaccharide monooxygenase from *Streptomyces lividans*



Synopsis: *Sli*LPMO10E is one of seven lytic polysaccharide monooxygenases identified in *Streptomyces lividans*. It is a member of the AA10 family displaying activity on chitin and presents the existence of two Cu-binding conformers within the active site.

5.1. Introduction

Complex sugars, termed "polysaccharides" from plants, insects and marine organisms embody some of nature's largest reservoirs of organic carbon on Earth. Chitin, cellulose and starch make up a large proportion of this in the form of hexose sugars linked linearly by glycosidic bonds. Chitin is formed by β-(1-4) linked N-acetyl glucosamine sugar units and is predominately found within crustaceans and insects, thus is readily available from seafood industry waste [275]. Lignocellulose makes up the woody part of plants and is composed of cellulose, hemicelluloses and lignin. Long chains of β-(1-4) linked glucose subunits, which can hydrogen bond to each other, form the crystalline solid cellulose. Starch is the main energy storage polysaccharide in plants [276]. It is formed of two types of glucose polymers, amylose, which is produced by long chains of α-(1-4) glucose units and amylopectin, consisting of shorter α-1(-4) linked glucose units in addition to α-(1-6) linked side chains (structures are shown in Figure 5.1) [277]. The increasing demand for fuel, coupled with ever depleting stocks of crude oil, has driven a great deal of attention towards the use of natural recoverable sources for biofuel production. Recalcitrant polysaccharides have enormous potential as renewable energy rich sources for second-generation biofuel production [278, 279]. The use of these polysaccharides for biofuel production involves two processes: degradation into fermentable sugars catalyzed by hydrolytic/oxidative enzymes and fermentation into ethanol by yeast or bacteria [280]. Despite their potential, the crystalline arrangement of these polysaccharides, crucial for biological function, ensues a major challenge of breakdown into fermentable sugar due to their extraordinary stability. Therefore, high costs are associated with the cocktail of enzymes necessary to breakdown these insoluble substrates, largely due to the enzymes inefficiency [281]. To solve the problem of recalcitrance, research has progressed towards the use of a cost-effective and robust biocatalyst [282, 283]. Degradation of biomass was principally believed to solely involve a consortium of different endoglucanases and exo-acting cellobiohydrolases (collectively termed cellulases) with their diversity recently becoming more apparent [269, 284-289]. These enzymes all belong to different glycoside hydrolase families (GH) classified in the CAZy database (Carbohydrate Active enZyme, www.cazy.org) and once coupled with hemicellulases, form the basis of enzyme cocktails that can be used in biorefinery [67, 290]. Nevertheless, an initial enzymatic reaction must exist in order to overcome the inert nature of the crystalline substrate enabling subsequent attack by these classical hydrolytic enzymes.

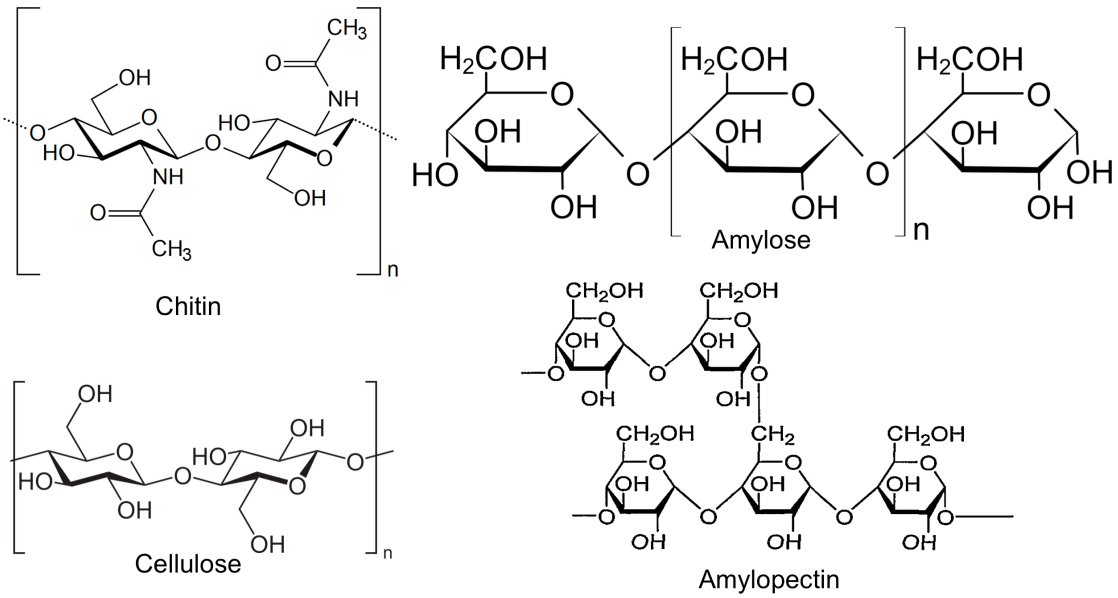


Figure 5.1: Structures of chitin, cellulose and starch (built from amylose and amylopectin).

Lytic polysaccharide monooxygenases (LPMOs termed PMOs by some) are a newly discovered group of copper (Cu) dependent carbohydrate-active enzymes that act by initial action on these crystalline substrates enhancing depolymerisation of recalcitrant polysaccharides in nature [287, 290-293]. Identification and characterisation of members of this cuproenzyme family from fungal and bacterial species has gathered pace in recent times due to their promising uses in deriving second generation biofuels and advancing the understanding of biological mechanisms of biomass breakdown [294-300]. LPMOs are classed into auxiliary activity (AA) families (as of 2013 [301]) in the CAZy database with four AA families, AA9, AA10, AA11 and AA13 so far identified [290, 301]. AA9 (formerly GH61 (glycoside hydrolase family 61)) and AA10 (formerly CBM33 (carbohydrate binding module family 33)) families have been the most extensively studied, with AA9 members having activity for cellulose and hemicelluose substrates [287, 302, 303] and AA10 members displaying activity with either cellulose or chitin substrates [291, 294, 304]. This class of enzymes was later expanded to include AA11 and AA13 families that have so far been determined to contain activity with chitin and starch substrates, respectively [305-307].

These LPMO families use an oxidative mechanism to cleave the β-1,4 glycosidic bonds within the chitin or cellulose substrate or the α-1,4 and α-1,6 linkages in starch [287, 305, 307]. Families AA9, AA11 and AA13 so far include only fungal enzymes, whereas the AA10 family encompasses enzymes from all domains of life. Initial investigations on LPMOs were reported in 2005 by Vaaje-Kolstad *et al* describing the harmonious relationship between a chitin-active AA10 (Chitin-binding protein 21, CBP21) and hydrolases in *Serratia marcescens* [308]. Later it was reported that an AA9 LPMO contained the ability to boost

enzyme activity on lignocellulosic biomass [294]. Subsequently, oxidative cleavage of glycosidic linkages by CBP21 was described by Vaaje-Kolstad *et al* [291]. The major breakthrough occurred when an AA9 from *Thermoascus aurantiacus* (*Ta*AA9) was shown to act as a monooxygenase, with its activity depending upon Cu [287]. Consequently, work with two AA9s from *Neurospora crassa* [309] and several AA10 members, established LPMOs as Cu-dependent monooxgenases [282, 310, 311]. Following on from these findings, knowledge on this new class of enzyme has increased significantly with expansion of the substrates upon which LPMOs can function.

All LPMOs share a common core immunoglobulin-like β-sandwich structure [292, 293]. Unlike canonical GH enzymes, which have substrate-binding grooves or tunnels, LPMOs consist of their active site at the center of a flat surface (structural comparisons shown in Figure 5.2). Aromatic side chains present on this flat surface and loop regions within the protein are implicated to be important in substrate binding and regioselectivity [193, 305, 306, 312]. All LPMOs coordinate a single Cu ion through three nitrogen ligands. Two of these are provided from the amino group and side-chain of the N-terminal His together with another nitrogen from a distal His side chain to give T-shaped coordination geometry referred to as the "His-brace" [292, 293]. The three nitrogen ligands and in some cases additional water molecules are located in the equatorial plane, with a Tyr or Phe and Ala (Gly in AA13) residue present in the axial positions [282, 287, 305-307, 313, 314]. Examples of the different Cu active site architectures present in the LPMO families are shown in Figure 5.3. Interestingly, when LPMOs are expressed in filamentous fungi the N-terminal His side chain that coordinates the Cu becomes Nε-methylated, however in systems lacking the machinery to carry out this modification, LPMOs still remain active, thus a role of this modification, if any, remains to be elucidated (Figure 5.3.A and E) [193, 287, 305-307, 315].

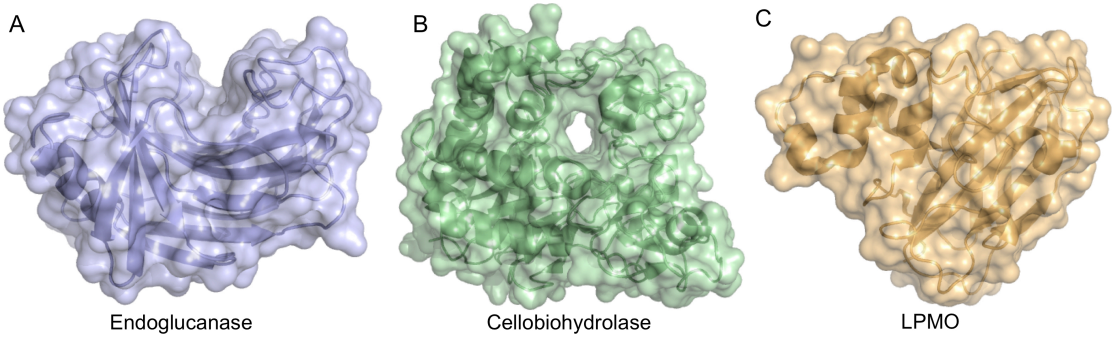


Figure 5.2: Comparison between the active sites of glycoside hydrolases and LPMOs. A) Endoglucanase shown in purple from *S. lividans,* PDB entry 1NLR, displaying a cleft for substrate binding. B) Cellobiohydrolase shown in green from *Thermobifida fusca,* PDB entry 4B4H, showing a substrate-binding tunnel. C) *Sc*LPMO10B shown in orange from *S. coelicolor,* PDB entry 4OY6, displaying the flat substrate-binding surface. Images were made using PyMOL.

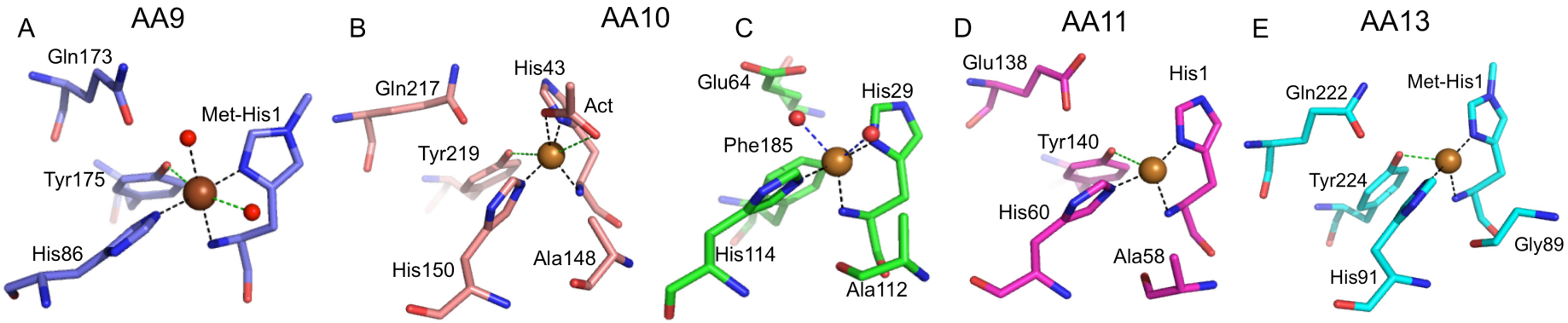


Figure 5.3: Comparisons of the active site architecture in LPMOs. A) AA9 from *T. aurantiacus* (PDB: 2YET) [287]. B) AA10 *Sc*LPMO10B from *S. coelicolor* (PDB: 4OY6) [313]. C) AA10 from *E. faecalis* (PDB: 4ALC) [316]. D) AA11 from *A. oryzae* (PDB: 4MAI) [305]. E) AA13 from *A. oryzae* (PDB: 4OPB) [306]. Cu is shown as a brown sphere. Cu equatorial coordinating groups are indicated by black broken lines and axial coordinating groups by green broken lines. Two waters, which are distorted from typical, type 2 geometry in (C) in the chitin-active AA10 are shown with blue broken lines. Second sphere residues, Glu/Gln and Ala/Gly are also shown.

To initiate enzymatic activity an electron source (either protein or small molecule) is utilised to reduce Cu(II) to the Cu(I) oxidation state, with the latter required to activate dioxygen [193, 287, 303, 309, 315, 317]. Chain breaks can result from oxidation of either the C1 or C4 of a sugar ring, with C1 oxidation generating soluble oligosaccharides with an aldonic acid at their reducing end and C4 oxidation yielding ketoaldose at the non-reducing end [193, 287, 302, 303, 309, 315, 317]. Two principle mechanisms of glycosidic bond cleavage have been proposed [309, 318]. The mechanisms proceed via the formation of either a Cu(II)-superoxo complex or by a Cu(II)-oxyl mechanism [309, 318]. The resulting C1 or C4 oxidation of a sugar ring instigate chain breakage (hence 'lytic') in the polysaccharide that are then susceptible to hydrolases for further degradation and glycan processing [193, 287, 303, 309, 315, 317].

Streptomycetes are the dominant bacterial genus responsible for aerobic biomass decomposition in soil environments and are considered vital players in the decomposition of cellulose and other biomass polymers [7, 319, 320]. Nearly all publically available streptomycetes genomes encode a relatively high percentage of genes for putative cellulolytic and chitinolytic enzymes. Recently, *Streptomyces* sp. SirexAA-E (ActE) has been identified and is associated with a symbiotic relationship with the pine-boring woodwasp *Sirex noctilio* [321]. The secretome of *Streptomyces* sp. ActE has been analysed through growth on a variety of carbon sources including cellulose and chitin [322]. In each case an LPMO was identified in the five most abundant proteins secreted, which also included exo- and endoglucanases, chitinase and N-acetylglucosaminidases [322]. Comparison of growth of ActE on cellulose to that of *Trichoderma reesei* Rut-C30, a model cellulolytic microbe, revealed that both completely deconstruct filter paper in 5-7 days [322]. Likewise, enzymatic activities of the ActE secretome showed comparable specific activity to the commercial enzyme cocktail prepared from the secretrome of *T. reesei* Rut-C30, the major progenitor organism of biofuels enzymology [322]. *S. coelicolor* and *S. griseus* have comparable putative genomic compositions of CAZy proteins to ActE, however these two streptomycetes are less capable of growing on cellulose [322]. This suggests that not all streptomycetes, despite their abilities to grow on plant polysaccharides [320], are efficient biomass decomposers.

The growth of many streptomycetes in liquid cultures is characterized by the formation of large, biofilm-like aggregates, called pellets [50, 51]. Streptomycetes possess the ability to directly secrete enzymes into the culture broth making them attractive hosts for the heterologous production of commercially valuable biomolecules [43, 44]. However, the pellets cannot efficiently take up oxygen and nutrients, thus they compromise the overall efficiency of the fermentation and subsequent secretion. Recently, synthesis of extracytoplasmic glycans has shown to be important for pellet formation [57, 64, 65, 166]. In

combination with results discussed in Chapter 2, enzymes encoded by the conserved *cslA-glxA* locus were found to produce a β-(1,4)-glycan at hyphal tips, providing protection during the ongoing cell wall remodeling at these sites [50, 64, 65, 166]. CslA belongs to family 2 of the glycosyl transferases, which contains cellulose and chitin synthases, amongst others [290]. Mutation of *cslA* abolishes pellet formation in liquid-grown cultures and also blocks aerial growth on solid media [64, 65]. As presented in Chapter 2 GlxA is a mono-copper oxidase containing a Cys-Tyr cross-linked redox cofactor and a Cu active site similar to fungal galactose oxidases (Gox) [166]. Deletion of *glxA* in *S. lividans* blocks development and abolishes pellet formation coinciding with the loss of glycan deposition at hyphal tips – the same phenotype as observed with the *cslA* mutation (Chapter 2) [72, 166]. This is consistent with a model whereby CslA and GlxA cooperatively function in glycan deposition and are responsible for the formation of reproductive aerial structures and for the formation of pellets in liquid cultures (Chapter 2) [166]. The chemical composition of the extracytoplasmic glycan is not presently known, however GlxA has a 1000-fold lower activity with canonical Gox substrates (Chapter 2) [166].

Recent *in vitro* and *in vivo* studies in *S. lividans* have identified an extracellular Cu-trafficking pathway involving Cu chaperone proteins (Sco and ECuC) [37, 38] that are required for GlxA maturation (as discussed in Chapter 4) (*i.e*. Cu-loading and subsequent formation of the Cys-Tyr cross-link [166] together with a role for a DyP-type haem peroxidase (DtpA) (Chapter 4) [171]. Based on these studies it has been suggested that a large protein complex is formed at the hyphal tips for synthesis and deposition of the glycan required for morphogenesis . Downstream of *glxA* is the *cslZ* gene, encoding for an endoglucanase, whilst upstream of *cslA* is a gene (*SLI3182*), predicted to encode for an AA10 LPMO (Chapter 2, genome environment Figure 2.1). The *S. lividans* 1326 genome [4] has seven genes that encode for putative AA10s (*Sli*LPMO10A-G) (Table 5.1). RNA-seq data of *S. lividans* 1326 grown in a mannitol/glucose media reveals two AA10 transcripts to be constitutively expressed (*Slilpmo10B* and *Slilpmo10E*), whilst the remaining five are likely under transcriptional control of a repressor or activator and depend upon induction by a cellulose or chitin substrate (Table 5.1). The proximity of *Slilpmo10E* and *cslZ* to the *cslA-glxA* operon could imply combined oxidative and hydrolytic activity is required in the processing and degrading of the *cslA-glxA* glycan produced during the development programme.

The intention of this investigation was to extend the glycan processing model proposed in Chapter 4 by characterizing the putative LPMO, *Sli*LPMO10E, within the *cslA-glxA* operon in *S. lividans,* and elucidate whether it is specific for cellulose or chitin-like substrates. *Sli*LPMO10E was found to be exclusively active on a chitin substrate and have a tertiary structure and Cu coordination sphere consistent with activity confined to oxidising the C1 ring position of the substrate. In addition, thermodynamics and kinetics of Cu(II) binding to *Sli*LPMO10E have elucidated the existence of two apo-protein forms that can bind

Cu(II) to a single site with three or two nitrogen coordination geometry, each with distinct kinetic and thermodynamic parameters.

Table 5.1: Transcription analysis of the seven *lpmo* genes in *S. lividans*. RNA-seq data annotations and expression levels in liquid MM (minimal media) + glucose/mannitol were analyzed with the *S. lividans* 1326 genome sequence as input [4].

Gene	Protein	*S. lividans* 1326 (RPKM)*	*S. lividans* 1326 + Cu (RPKM)*
SLI0440	*Sli*LPMO10A	0	0
SLI0614	*Sli*LPMO10B	10	9
SLI1466	*Sli*LPMO10C	0	6
SLI2039	*Sli*LPMO10D	0	2
SLI3182	*Sli*LPMO10E	22	27
SLI6742	*Sli*LPMO10F	0	0
SLI7441	*Sli*LPMO10G	0	3

*The transcriptomes of wild-type *S. lividans* 1326 was obtained without addition of Cu to the growth medium. The Cu induced transcriptome was obtained after a 2 hour exposure to 400 μM Cu(II). RPKM represents reads/kb of exon/million mapped reads.

5.2. Experimental procedures

5.2.1. Cloning of LPMOs and site-directed mutagenesis

To ensure that the first N-terminal residue in all mature LPMO proteins is the Cu coordinating His residue, the pelB leader sequence within the pET26b plasmid (Novagen) was mutated. The original NcoI site (CCATGG in bold Figure 5.4) was deleted and reintroduced 6 nucleotides downstream ensuring the pelB leader sequence was intact and removing the starting methionine residue (ATG). Mutagenesis primers used to mutate the pelB leader sequence are reported in Table 5.2 with the two changed nucleotides underlined. The Quikchange (Stratagene) method was used using the original pET26b plasmid as the template and all other reagents shown in Table 5.3 and the PCR protocol in Table 5.4. DpnI was added to the PCR mix to remove any template DNA. The plasmid DNA was then sequenced to corroborate that the correct change in the pelB region was incorporated. Wild-type (WT)- *Sli*LPMO10E was prepared by amplifying the *SLI3182* (519 base pairs – nt 88 to 606) gene from genomic DNA using the primers shown in Table 5.2 containing the altered NcoI site. The reagents used in the PCR mixture for amplifying *SLI3182* are shown in Table 5.5 and the PCR protocol in Table 5.6. The *SLI3182* gene product was cloned into the modified pET26b plasmid using the NcoI and HindIII restriction sites. This creates a *pET3182* construct containing His^{30} as the first residue of the mature LPMO protein following cleavage of the pelB leader sequence. A similar cloning strategy was used for *Sli*LPMO10B (*SLI0614*, 561 bp - nt 127 to 687) using the primers shown in Table 5.2 (carried out by collaborators at Leiden University). The *pET0614* construct was created containing His^{43} as the N-terminal start residue. All constructs were sequenced to corroborate the presence of the modified pET26b plasmid and the correct gene insert (Figure 5.6.C - *pET3182*)

Site-directed mutagenesis of His^{187} to an Ala in *Sli*LPMO10E was carried out using the mutagenic primers reported in Table 5.2. The WT construct was used as a template with the reagents shown in Table 5.3 and the protocol in Table 5.4 to produce the H187A mutant (identical to the mutagenesis of pET26b).

The N-terminal *Sli*LPMO10E extension variant (*Sli*LPMO10E-Ext) whereby His^{30} is no longer the N-terminal residue was created by introducing a Gly residue prior to His^{30} and cloning into an unmodified pET26b plasmid to give a mature protein with an N-terminal sequence of Met-Gly-His^{30}. Primers used to amplify the gene are shown in Table 5.2 with the reagents used given in Table 5.5 and the protocol in Table 5.6 (identical to the WT-*Sli*LPMO10E PCR). The gene product was subsequently cloned into an unmodified pET26b plasmid creating the *pET3182-Ext* construct (Figure 5.6.F). All constructs were DNA sequenced to corroborate the correct mutations and insertions.

5’-
AAATACCTGCTGCCGACCGCTGCTGCTGGTCTGCTGCTCCTCGCTGCCCAGCCGG
CGATGG**CC|ATGG**……-3’

Figure 5.4: pET26b pelB nucleotide sequence. The line | indicates the signal peptidase cleavage site. Bold is the original NcoI site and the underlined nucleotides are the ones that were mutated to create the new NcoI site (also shown below in Table 5.2).

Table 5.2: Primer sequences for the pelB mutation of the pET26b plasmid, *Sli*LPMO10E, *Sli*LPMO10B, H187A and *Sli*LPMO10E-Ext.

Primer Name	Primer sequence	Tm (^{o}C)
Pelb Forward	5′-CCAGCCGGCCATGGCGATGGATATCGG-3′	62
Pelb Reverse	5′-CCGATATCCATCGCCATGGCCGGCTGG- 3′	
*Sli*LPMO10E Forward	5′-ATAATTACCATGGCGCACGGCTACACCGACCTG-3′	62
*Sli*LPMO10E Reverse	5′-CTTAAAGCTTTCAGAAGGTGACGTCCGAGCAGGC-3′	62
*Sli*LPMO10B Forward	5′-GCGGAATTCCATGGCGCACGGCTCGGTCGTCGACC-3′	65
*Sli*LPMO10B Reverse	5′-CGCAAGCTTTCAGCCGAAGTCGACGTCGCTGC-3′	68
H187A Forward	5′-GTGTGGACGGTCGCCGACACGGGCAA-3′	62
H187A Reverse	5′-TTGCCCGTGTCGGCGACCGTCCACAC-3′	62
*Sli*LPMO10E-Ext Forward	5′-TAATTATCCATGGGTCACGGCTACACCGACCTGC-3′	62
*Sli*LPMO10E-Ext Reverse	5′-CTTAAAGCTTTCAGAAGGTGACGTCCGAGCAGGC-3′	62

Table 5.3: Reagents in the PCR protocol used to amplify the mutated pelB sequence within the pET26b plasmid and the H187A mutant.

Reagent	Concentration	Volume (µl)
Plasmid DNA	N/A	1
Forward Primer	75 ng/µl	1
Reverse Primer	75 ng/µl	1
dNTP's	10 mM	0.6
Pfu Turbo buffer	10 x	3
DMSO	100%	1.8
Sterile, deionized water	-	21
Pfu Turbo	-	0.6
Total		30

Table 5.4: The PCR protocol used for the amplification of the mutated pET26b plasmid and H187A mutant. Numbers highlighted in bold were cycled 15 times.

Temperature (°C)	Time (min)
95	3
95	**0.5**
58	**1**
68	**13**
72	8
10	Final Extension

Table 5.5: Reagents used in the PCR protocol to amplify *Sli*LPMO10E and *Sli*LPMO10E-Ext from the genomic DNA of *S. lividans*.

Reagent	Concentration	Volume (µl)
Genomic *S. lividans* DNA	N/A	2.5
Forward Primer	N/A	1
Reverse Primer	N/A	1
dNTP's	10 mM	2.5
10X Buffer Pfu + $MgSO_4$	-	5
DMSO	100 %	5
Sterile, deionized water	-	32.5
Pfu DNA polymerase	-	0.5
Total		50

Table 5.6: The PCR protocol used for the amplification of *Sli*LPMO10E and *Sli*LPMO10E-Ext from the genomic DNA of *S. lividans*. Numbers highlighted in bold were cycled 35 times.

Temperature (°C)	Time (min)
95	3
95	**1**
62	**1**
72	**2**
72	7.5
4	Final Extension

5.2.2. Over-expression and purification of LPMOs

pET3182, pET0614, H187A and *pET3182-Ext* (encoding *Sli*LPMO10E, *Sli*LPMO10B, H187A and *Sli*LPMO10E-Ext) were transformed into chemically competent *Escherichia coli* BL21 (RIL) cells using the heat-shock method. Cells containing the desired plasmid were selected by growing on 50 μg ml^{-1} kanamycin (Kan) and 35 μg ml^{-1} chloramphenicol (Cm) agar plates. Single colonies were transferred into 3 ml of LB (Luria broth) medium (Melford) containing 50 μg ml^{-1} Kan and 35 μg ml^{-1} Cm. These were incubated overnight at 37 °C with shaking at 180 rpm. Overnight pre-cultures were successively used to inoculate 500 ml LB cultures in 2 L flasks incubated at 37 °C 180 rpm. At an OD_{600} of 0.4, the temperature was lowered to 16 °C and upon reaching an OD_{600} of 0.6-0.8 isopropyl β-D-1-thiogalactopyranoside (IPTG; Melford) was added to a final concentration of 1 mM. Cells were harvested after 16 h at 3,501 g for 20 min at 4 °C. Cell pellets were combined, weighed and resuspended in $1/30^{th}$ of the total culture volume in ice-cold 50 mM Tris/HCl 1 mM EDTA 20 % w/v sucrose, pH 8. The cell paste was stirred at 4 °C for 1 h, then 60 μl of 1 M $MgSO_4$ was added for every gram of cell pellet and left for a further 30 min stirring at 4 °C. The smooth cell suspension was then centrifuged at 38,724 g 20 min at 4 °C and the supernatant removed and stored at 4 °C. The pellet was then resuspended in ice-cold water ($1/30^{th}$ the volume of the culture), stirred at 4 °C for 1 h followed by centrifugation at 38,724 g 20 min at 4 °C with the supernatant collected and combined with that from the previous sucrose fractionation. The combined supernatants were then dialysed overnight against 5 mM Tris/HCl 1 mM EDTA, pH 7 at 4 °C.

For *Sli*LPMO10E (pI 8.42), *Sli*LPMO10E-Ext (pI 8.39) and H187A (pI 8.43) the dialysate was diluted with buffer A (5 mM Tris/HCl, pH 7) and loaded to a 5 ml HiTrap SP-column (GE-Healthcare) equilibrated with buffer A and eluted with a linear salt gradient using buffer A with 1 M NaCl. For *Sli*LPMO10B (pI 4.24) the dialysate was diluted with buffer B (10 mM Tris/HCl, pH 5.5) and loaded to a pre-equilibrated DEAE column (GE-Healthcare) and eluted with a linear salt gradient using buffer B with 500 mM NaCl. Fractions containing LPMO were pooled and concentrated at 4 °C using a 5 kDa cut-off centricon (Vivaspin). Concentrated protein was then loaded to a HiLoad 26/60 Superdex-75 size-exclusion column (GE Healthcare) equilibrated with buffer C (10 mM sodium acetate pH 5, 150 mM NaCl). Fractions eluting from the major peak (retention volume ~ 80 ml consistent with a monomer species) were concentrated and stored at - 20 °C until required.

5.2.3. UV-visible spectroscopy

Concentrations of the LPMO proteins were determined using a Varian Cary 50 UV-visible spectrophotometer with a 1 cm pathlength Quartz cuvette (Hellma) at 20 ^{o}C. The absorbance at 280 nm was measured and concentrations calculated using the beer-lambert

law with an extinction coefficient (ε) (280 nm) of 36,690 M^{-1} cm^{-1} for WT *Sli*LPMO10E, H187A and *Sli*LPMO10E-Ext and 63,160 M^{-1} cm^{-1} for *Sli*LPMO10B [93].

5.2.4. Protein preparations

Prior to experiments requiring the addition of Cu(II) salts, proteins were incubated for prolonged periods (~ 1-2 h) with excess EDTA (Ethylenediaminetetraacetic acid). The excess was removed by passing the protein/EDTA mix down a HiLoad 26/60 Superdex-75 size-exclusion column (GE Healthcare). For Cu(II) titrations and preparation of Cu(II)-loaded enzyme a stock solution of 100 mM $Cu(II)SO_4$ (Sigma) was prepared and diluted as required. Cu(II)-loaded enzymes were prepared by stoichiometric addition of $Cu(II)SO_4$ followed by desalting to ensure no unbound metal remained.

5.2.5. Fluorescence spectroscopy

Fluorescence spectroscopy of *Sli*LPMO10E, *Sli*LPMO10E-Ext and *Sli*LPMO10B were measured on a LS 50B fluorimeter (Perkin Elmer) thermostatted at 20 ^{o}C. Trp fluorescence emission spectra were collected between 300 - 400 nm following excitation at 295 nm with excitation and emission slits set at 5 nm. LPMO proteins (2 μM) were prepared in 3 ml Quartz cuvettes (Hellma) in 10 mM sodium acetate, pH 5, 150 mM NaCl. The decrease in the observed emission was monitored upon microliter aliquots of $Cu(II)SO_4$ solution of known concentration.

5.2.6. Circular Dichroism spectroscopy

LPMO samples used for circular dichroism (CD) analysis were exchanged into 10 mM KPi, 50 mM KF, pH 7 using a PD-10 desalting column (GE-Healthcare). Far UV-CD spectra were recorded between 260 and 190 nm at 20 ^{o}C measured on an Applied Photophysics Chirascan CD spectrophotometer equipped with a thermostatic cell holder controlled with a Peltier system (Leatherhead, UK).

5.2.7. Analytical gel filtration

Speciation of WT-*Sli*LPMO10E was examined using a 10/300 GL G-75 Superdex column (GE-Healthcare) equilibrated in 10 mM Na acetate, 150 mM NaCl, pH 5. Samples of apo-*Sli*LPMO10E and following addition of 5-fold excess $Cu(II)SO_4$ (300 μM) were injected onto a pre-equilibrated column and elution volumes monitored at 280 nm.

5.2.8. Redox potentials

The cell potentials for WT-*Sli*LPMO10E-Cu(II)/WT-*Sli*LPMO10E-Cu(I) and *Sli*LPMO10E-Ext-Cu(II)/*Sli*LPMO10E-Ext-Cu(I) were carried out by reacting reduced oxygen-free N,N,N',N'-tetramethyl-1, 4-phenylenediamine (TMP) (50 µl at either 150 µM or 300 µM) with LPMO-Cu(II) (50 µl at 70 µM) (pH 5, 25 °C) and letting the reaction reach equilibrium under anaerobic conditions. The equilibrium constant was determined by calculating the concentration of TMP_{ox} by measuring the absorbance from the TMP radical cation (TMP_{ox}) at 610 nm and using an extinction coefficient of 14 mM^{-1} cm^{-1}, which is thus equal to the concentration of LPMO-Cu(I) (Equation 5.1) and hence the equilibrium constant can be calculated (Equation 5.2).

$$TMP_{red} + LPMO\text{-}Cu(II) \leftrightharpoons TMP_{ox} + LPMO\text{-}Cu(I) \qquad \text{Equation 5.1}$$

$$K = \frac{[TMP_{ox}][LPMO\text{-}Cu(I)]}{[TMP_{red}][LPMO\text{-}Cu(II)]} \qquad \text{Equation 5.2}$$

The relationship between the free energy change (ΔG_r°), the equilibrium constant (K) and the cell potential (E°) is shown in Equation 5.3, where R is the gas constant, T is the temperature in Kelvin, n is the number of electrons and F is the Faraday constant. Subtraction of the E° for the TMP_{ox}/TMP_{red} (273 mV) redox couple yields a mid-point redox potential determined at pH 5 ($E_{m,5}$).

$$\Delta G_r^\circ = -RT \ln K = -nFE^\circ \qquad \text{Equation 5.3}$$

5.2.9. Electron Paramagnetic Resonance (EPR) spectroscopy

Continuous-wave X-band frozen solution EPR spectra of *Sli*LPMO10E (80 µM) were prepared in duplicate in 5 mM Tris, 150 mM NaCl, pH 7. Samples of *Sli*LPMO10E were prepared as isolated, with addition of 0.9 equivalent $Cu(II)SO_4$. Wilmad SQ EPR tubes (Wilmad Glass, Buena, NJ) were filled with the LPMO solutions and frozen in methanol kept on dry ice, then transferred to liquid nitrogen. All EPR spectra were measured at 10 K on a Bruker EMX EPR spectrometer at a modulation frequency of 100 kHz. A spherical high-quality Bruker resonator ER 4122 SP 9703 and an Oxford Instruments liquid helium system were used to measure the low-temperature EPR spectra. Blank frozen water samples were subtracted from the EPR spectra of the protein samples to eliminate the baseline caused by

the resonator's walls, quartz insert or quartz EPR tube. The baseline was corrected by subtracting a polynomial line drawn through a set of points randomly chosen on the baseline using WinEPR (v2.22, Bruker Analytik, GmbH).

5.2.10. EPR Cu(II) simulations

Spectral simulations were performed in SimFonia v. 1.26 (Bruker Analytik, GmbH). The g_z value and its hyperfine splitting constant $|A_z^{Cu}|$ were determined directly from the spectra. Rather than varying g_x and g_y independently, two other values, g_{av} and Δg, were varied, where $g_{av} = (g_x + g_y)/2$ and $\Delta g = g_y - g_{av} = g_{av} - g_x$. Δg was always assumed zero at the beginning of simulation process, then, if a good fit was impossible to achieve it was allowed to be different from zero. To maintain consistency in the simulation procedure, a number of constraints were imposed on the choice of simulation parameters thus minimising the number of parameters allowed to be varied. Thus, the tensor for the hyperfine interaction of the spin density with Cu nucleus A^{Cu} ($I^{Cu} = 3/2$) was assumed axial, with the $A_x^{Cu} = A_y^{Cu} = 0.1\ A_z^{Cu}$ (small variation of the factor around the value of 0.1 were found to produce minimal effect on the simulated spectra, therefore the value has been locked at 0.1 for all simulations). Further, all A^{Cu}-tensors were assumed collinear with the g-tensor (all Euler angles were assumed zero). Cu(II) coordinating nitrogen atoms ($I_N = 1$) were assumed to produce isotropic hyperfine interactions ($A_x^N = A_y^N = A_z^N = 14$ G). In accordance with earlier reported relationship between principal g-factor components and individual line width values along those directions [167], the x-, y- and z- components of the spectra line width ΔH were assumed to be in a close to linear dependence on the g-values, the higher the g_i, the higher the ΔH_i value (where i is x, y or z). This dependence was maintained for all simulations.

5.2.11. Substrate binding assay

WT- *Sli*LPMO10E (~ 20 μM) was prepared in 50 μl of 50 mM Bis-Tris/HCl pH 6.2 and incubated with ~ 5-10 mg of excess substrate; cellulose (Whatman #1, 50 mm), or squid pen β-chitin at room temperature for ~ 3-4 h with occasional agitation. The supernatant was then removed by centrifugation and kept as the unbound fraction. Substrate was then washed three times with 500 ml of buffer (50 mM Bis-Tris/HCl pH 6.2) to remove any unspecific enzyme. Bound enzyme was removed by the addition of 50 μl of SDS-PAGE loading buffer containing DTT (see Chapter 2 section 2.2.4) and heated to 95 ^{o}C for 10 min. All samples were then analysed for bound and unbound enzyme content by loading and visualising on a 15 % SDS-PAGE gel (BioRad) (see Chapter 2 section 2.2.4).

5.2.12. Mass spectrometry

Oxidative activity was assayed by mass spectrometry (MS) using a range of substrates (squid pen β-chitin, Avicel (cellulose), Konjac glucomannan, Guar gum, Ivory nut mannan and Locust bean gum). Reaction mixtures (1 ml total volume) contained 0.2 % w/v solid substrate in 10 mM ammonium acetate, pH 5, 1 mM ascorbic acid and 10 μM Cu-loaded LPMO. Control samples were also prepared in the absence of ascorbic acid and protein. Reactions were incubated at 30 °C rotating for 5-12 h before the remaining solid substrate was removed by centrifugation at 16,000 g, 4 °C for 5 min and the supernatant used for analysis. 1 μl of sample was mixed with an equal volume of 10 mg/ml 2,5-dihydroxybenzoic acid in 50 % acetonitrile, 0.1 % trifluoroacetic acid on a SCOUT-MTP 384 target plate (Bruker). The spotted samples were then dried in a vacuum desiccator before being analyzed on an Ultraflex III matrix-assisted laser desorption ionization–time of flight/time of flight (MALDI-TOF/TOF) instrument (Bruker), as described previously [291].

5.2.13. SliLPMO10E crystallisation and structure determination

An ARI-Gryphon 96-well crystallisation robot was used to screen crystallisation conditions for Cu-loaded *Sli*LPMO10E. A crystal hit was found in 0.1 M sodium acetate pH 4.6, 25 % PEG 4,000 (PEG suite, Qiagen). Scaling-up and optimisation of *Sli*LPMO10E crystals from the initial hit was carried out in 24-well VDX plates (Molecular Dimensions) using the hanging drop vapour diffusion method at 18 ^{o}C. Equal volumes of protein (15 mg ml^{-1}) and reservoir solution were mixed and crystals suitable for diffraction studies grew within 2-3 weeks. A single crystal grown from 0.05 M sodium acetate pH 4.6, 25 % PEG 4,000 was transferred to a cryoprotectant solution containing the respective reservoir solution and 20 % glycerol prior to flash-cooling by plunging into liquid nitrogen. Crystallographic data were measured at the beamline I03, Diamond Light Source, using an X-ray wavelength of 1.00 Å and a Pilatus 6 M-F detector (Dectris). Data were processed automatically in XIA2 and scaled and merged using Pointless/Scala [95] in the CCP4i suite. The structure was solved using automated molecular replacement in BALBES with an initial model built in ARP-wARP. The structure was refined using Refmac5 [98], with model building between refinement cycles in Coot [100]. Riding hydrogen atoms were added when refinement of the protein atoms had converged. Structures were validated using the Molprobity server [101], the JCSG Quality Control Server and tools within Coot [100]. A summary of data, refinement statistics and the quality indicators for the structure are given in Table 5.7. Additional data were collected at wavelengths of 1.20 Å and 1.33 Å to generate anomalous maps for validation of modelled Cu atoms in the structure.

Table 5.7: Crystallographic data processing and refinement statistics of *Sli*LPMO10E. Values in parentheses refer to the outermost resolution shells (1.40-1.38, 2.07-2.02, 2.07-2.01 Å, respectively). The unit cell was 69.6, 32.4, 61.3 Å, 90, 97.8, 90 ° in space group C2.

Wavelength (Å)	1.0	1.33	1.2
Resolution (Å)	1.38	2.02	2.01
Unique reflections	27360	8755	8844
Completeness (%)	97.3 (92.2)	97.6 (90.6)	96.7 (87.2)
R_{merge}	0.055 (0.680)	0.060 (0.087)	0.049 (0.084
Mn (I/Sd)	14.8 (1.8)	20.3 (13.7)	24.1 (15.9)
Redundancy	4.9 (3.5)	5.5 (4.7)	5.5 (4.7)
Wilson B-factor ($Å^2$)	9.3	8.3	9.8
R_{cryst}	0.135		
R_{free}	0.161		
ESU base on ML (Å)	0.038		
RMS dev. Bond lengths (Å)	0.019		
RMS dev. Bond angles (^{o})	1.9		
Ramachandran favoured (%)	98.8		

5.2.14. Fluorescence stopped-flow kinetics

Kinetic experiments to measure fluorescence emission changes on Cu(II) binding were carried out on a SX20 stopped-flow spectrophotometer (Applied Photophysics) thermostatted at 25 °C with a Peltier system. Trp residues were excited at 295 nm with a filter absorbing light at ~ 295 nm (excitation band) and transmitting light above 350 nm was used. Apo proteins *Sli*LPMO10E, *Sli*LPMO10B, *Sli*LPMO10E-Ext and H187A (2.5 or 5 μM) were prepared in 10 mM sodium acetate, pH 5, 150 mM NaCl and rapidly mixed with a series of Cu(II)SO_4 concentrations (25-500 μM after mixing). Following excitation at 295 nm time courses were taken collecting the total fluorescence emission. The data were fitted best to double exponential functions for *Sli*LPMO10E, *Sli*LPMO10B and H187A (Equation 5.4) and triple exponential functions for *Sli*LPMO10E-Ext (Equation 5.5). All kinetic data were analysed using the Applied Photophysics ProKinetist software. The obtained rates were plotted against Cu(II) concentration and where appropriate were analysed using a linear fit analysis. The second order rate constant (k_{on}) is the gradient of the linear analysis and could be calculated for the fast phase of all LPMO proteins (and slow phases of *Sli*LPMO10E, *Sli*LPMO10B and H187A), whereas, k_{off} is determined as the y-intercept and could only be calculated for the fast phase of *Sli*LPMO10E-Ext. By combining the k_{off}/k_{on} a K_d could be determined for initial Cu(II) binding to *Sli*LPMO10E-Ext.

Stoichiometry of Cu(II) binding to *Sli*LPMO10E and *Sli*LPMO10E-Ext were explored by monitoring fluorescence amplitude changes. Proteins (5 μM after mixing, in 10 mM sodium acetate, 150 mM NaCl, pH 5) were pre-loaded with *sub* to *super* stoichiometric amounts of Cu(II), which were then rapidly mixed with excess Cu(II). The change in the fluorescence emission amplitude was monitored for the fast and slow phases of *Sli*LPMO10E and a combination of all three phases for *Sli*LPMO10E-Ext and plotted against [Cu(II)/LPMO]. All kinetic data were analysed using the Applied Photophysics ProKinetist software.

$$y = A\ (\exp^{-kt}) + A_2\ (\exp^{-k_2 t_2}) \qquad \text{Equation 5.4}$$

$$y = A\ (\exp^{-kt}) + A_2\ (\exp^{-k_2 t_2}) + A_3(\exp^{-k_3 t_3}) \qquad \text{Equation 5.5}$$

5.2.15. Fractional saturation and K_d calculations

Trp fluorescence quenching experiments could be calculated in terms of fractional saturation defined in Equation 5.6 where ΔF is the fluorescence change and ΔF_T is the fluorescent change from zero to fully saturated enzyme with Cu(II). The K_d could then be calculated using the hyperbolic fit in Equation 5.7 (P2 is the K_d) or the hyperbolic equation in

Equation 3.3 in Chapter 3 (when [Cu(II)]>>[LPMO]), or by using the quadratic equation shown in Equation 5.8 (when [Cu(II)], [LPMO] and [LPMO-Cu(II)] are known).

$$Y=1-\frac{\Delta F}{\Delta F_T} \qquad \text{Equation 5.6}$$

$$y = \frac{(P1*[Cu(II)])}{(P2+[Cu(II)])} \qquad \text{Equation 5.7}$$

$$y = \frac{([P_T]+[S]+K_d)-\sqrt{([P_T]+[S]+K_d)^2-4([P_T][S])}}{2[P_T]} \qquad \text{Equation 5.8}$$

5.2.16. Isothermal titration calorimetry (ITC)

Calorimetric titration experiments were carried out at 25 ± 0.1 ^{o}C on a MicroCal VP-ITC calorimeter in (10 mM sodium acetate pH 5, 150 mM NaCl). Before each run, samples were degassed for 15 min at 23 ± 0.1 ^{o}C using the ThermoVac accessory. A solution of Cu(II)SO_4 (1 mM) was loaded into the injection syringe and titrated into 100 μM of WT or variant *Sli*LPMO10E in the sample cell with stirring at 307 rpm for the duration of the experiment. A reference power of 5 mcal/s was used with either an initial 3 or 5 μl injection of Cu(II)SO_4 followed by either 5 or 10 μl injection for all subsequent titration points. A 60 s initial equilibrium delay with 270 s pauses between subsequent injections were applied throughout. Raw data were analysed using Origin 7.0 software. The integrated data were corrected for the heat of dilution of buffer into buffer and Cu(II)SO_4 into buffer and the binding isotherms were fitted using binding models provided in the software package. All experiments were performed in duplicate and the standard error reported.

5.3. Results

5.3.1. Amino acid sequence alignment

Amino acid sequences of *Sli*LPMO10E and *Sli*LPMO10B were aligned with CBP21 from *S. marcescens*, CelS2 (*Sc*LPMO10C) and *Sc*LPMO10B from *S. coelicolor* using Clustal Omega (Figure 5.5) [108]. *Sli*LPMO10B is 99.5 % identical to *Sc*LPMO10B with only one amino acid difference, Glu[175] in *Sc*LPMO10B and Asp[175] in *Sli*LPMO10B. The other sequences share ~ 20-30 % sequence identity, except *Sli*LPMO10E and CBP21, which share ~ 45 % identity. The secondary structure prediction for *Sli*LPMO10E from Jpred4 [113] is indicated above the alignment (Figure 5.5) with a mixed α and β secondary structure, calculated as 5.2 % α-helix and 23 % β-sheet content.

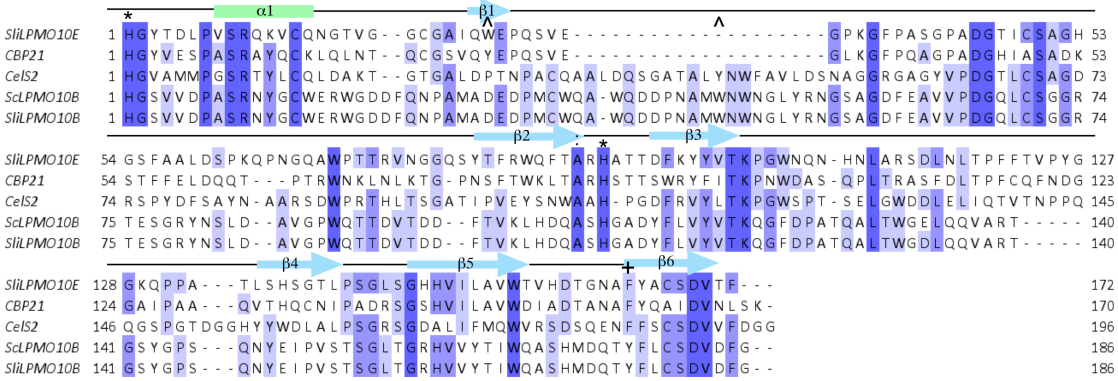


Figure 5.5: Sequence alignment of *Sli*LPMO10E and *Sli*LPMO10B. Clustal Omega sequence alignment of two out of the seven *S. lividans* LPMO proteins (*Sli*LPMO10E and *Sli*LPMO10B) with CBP21 from *S. marcescens*, CelS2 (*Sc*LPMO10C) and *Sc*LPMO10B from *S. coelicolor*. Completely and partially conserved residues are coloured in dark to light blue, respectively. The secondary structure prediction for *Sli*LPMO10E from Jpred4 [113] is indicated above the alignment, with an α-helix shown in green and β-sheets in blue. The * symbol indicates the position of the coordinating His residues, the ^ shows the presence of the surface Trp[56] in *Sli*LPMO10E and Trp[88] in *Sli*LPMO10B, the : shows the position of Ala LPMO residues (Ala[118] in *Sli*LPMO10E) and the + symbol shows the position of Tyr and Phe residues in the LPMOs.

5.3.2. Cloning, over-expression and purification of LPMOs

In LPMOs the N-terminal His residue of the mature protein coordinates to the Cu ion. To ensure this is the case for over-expression of LPMO constructs in *E. coli* the NcoI restriction site in the original pET26b plasmid was deleted and a new NcoI site created downstream (Figure 5.6.A). This facilitates overproduction of any LPMO protein, that following cleavage of the pelB leader sequence results in a His as the N-terminal residue enabling formation of the 'His brace' Cu site. Genes encoding *Sli*LPMO10E and *Sli*LPMO10B were successfully amplified from the genomic DNA of *S. lividans* (Figure 5.6.B), and subsequently sub-cloned into the modified pET26b *E. coli* expression vector containing a pelB leader sequence for periplasmic expression (Figure 5.6.C). *Sli*LPMO10E is the main focus of this work with the genomic environment shown in Chapter 2 Figure 2.1 (and introduction Figure 1.3). *Sli*LPMO10B is a homologue of the previously characterised AA10 *Sc*LPMO10B from *S. coelicolor* sharing 99.5 % sequence identity and has been used here to compare to the Cu(II) binding kinetics of *Sli*LPMO10E. The H187A variant of *Sli*LPMO10E was successfully created using a Quikchange site-directed mutagenesis method (Figure 5.6.D). Furthermore, the gene encoding *Sli*LPMO10E-Ext was also successfully cloned from the genomic DNA of *S. lividans* (Figure 5.6.E) and sub-cloned into the original pET26b plasmid encoding an extension protein of *Sli*LPMO10E with the N-terminal sequence Met-Gly-His[30] (Figure 5.6.F).

All LPMO proteins produced clear bands on SDS-PAGE gels during over-expression and clean protein bands from the supernatant fractions after the sucrose and water shocks (Figure 5.7.A). *Sli*LPMO10E, *Sli*LPMO10E-Ext and H187A were first purified using a HiTrap SP-column with all proteins producing single sharp peaks at 280 nm ~ 30 ml elution volume at ~ 20 % buffer A with 1 M NaCl (Figure 5.7.B). Whereas, *Sli*LPMO10B (pI 4.24) was first purified using a DEAE column giving multiple peaks at 280 nm with a clear peak corresponding to the protein at ~ 90 % buffer B with 500 mM NaCl (Appendix 3.1). All LPMO proteins were purified further by application to a G75 size exclusion column, all producing elution profiles with retention volumes consistent with monomeric species (Figure 5.7.C and Appendix 3.1.C). Fractions of the proteins from the G75 column migrated as single bands on denaturing PAGE gels to a mass of ~ 18 kDa for *Sli*LPMO10E (Figure 5.7.D *Sli*LPMO10E predicted mass - 18,422 Da for residues 30-201, *Sli*LPMO10E-Ext mass - 18,610 Da and H187A - 18,356 Da) and ~20 kDa for *Sli*LPMO10B (Appendix 3.1.D predicted mass 20,723.4 Da for residues 43-229).

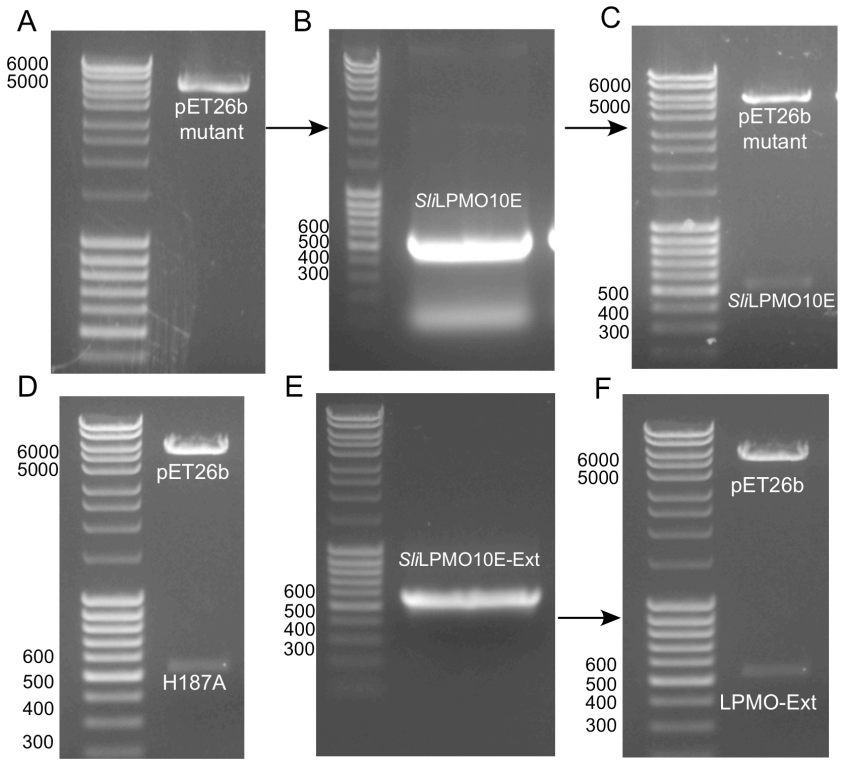


Figure 5.6: Cloning of *Sli*LPMO10E, H187A and *Sli*LPMO10E-Ext. A) DNA agarose gel (1 % w/v) of the mutated pET26b plasmid (5,360 bp). B) DNA agarose gel (1 % w/v) of the *Sli*LPMO10E PCR product amplified from the genomic DNA of *S. lividans*. The correct band of 519 bp was corroborated with the DNA ladder (Fermentas DNA mass ruler). C) *pET3182* construct digested using NcoI and HindIII restriction enzymes and corroborated against the DNA ladder (far left lane) (pET26b – 5,360 bp and *Sli*LPMO10E – 519 bp). D) H187A *Sli*LPMO10E variant construct digested using NcoI and HindIII restriction enzymes (pET26b – 5,360 bp and H187A – 519 bp). E) DNA agarose gel (1 % w/v) of the PCR product of *Sli*LPMO10E-Ext amplified from the genomic DNA of *S. lividans*. F) *pET3182-Ext* construct digested using NcoI and HindIII restriction enzymes and corroborated against the DNA ladder (pET26b – 5,360 bp and *Sli*LPMO10E-Ext – 525 bp).

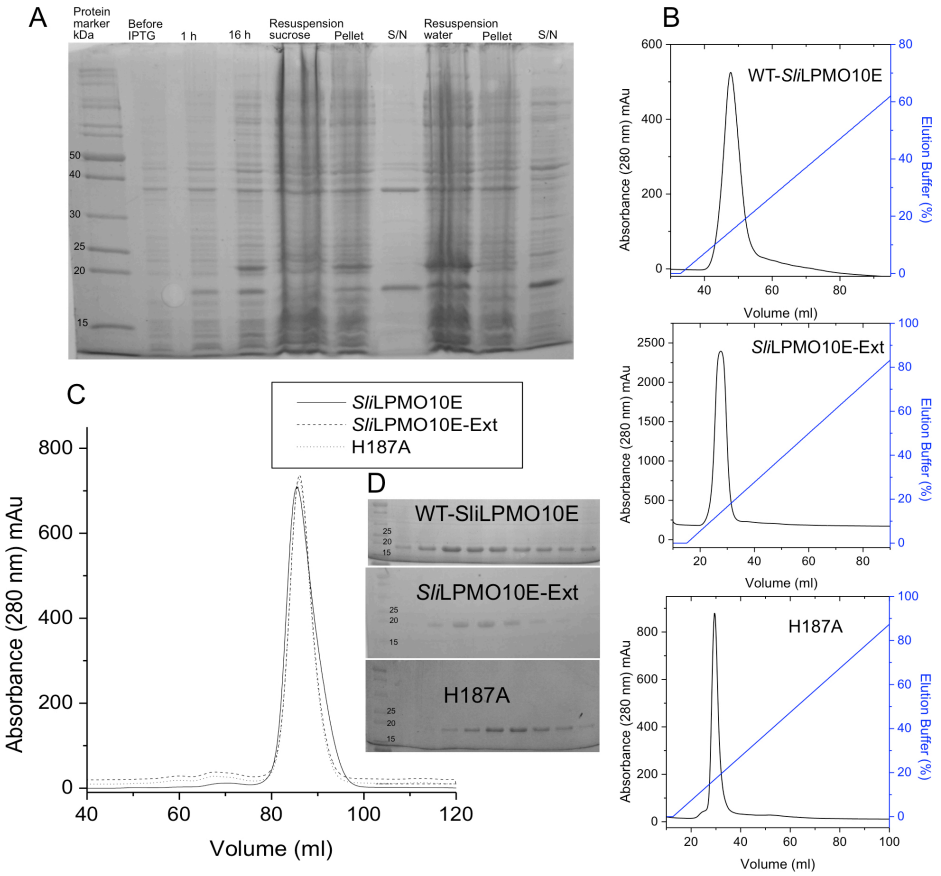


Figure 5.7: Over-expression and purification of *Sli*LPMO10E, *Sli*LPMO10E-Ext and H187A. A) Coomassie stained 15 % SDS-PAGE gel analysis of *Sli*LPMO10E during over-expression and osmotic shock treatment, with the lane contents indicated. B) SP column profiles of *Sli*LPMO10E, *Sli*LPMO10E-Ext and H187A showing an absorbance peak (mAu) at 280 nm in black at ~20 % buffer B shown in blue. C) G75 size-exclusion Sephadex column absorbance 280 nm (mAu) profile of *Sli*LPMO10E, *Sli*LPMO10E-Ext and H187A all eluting at ~ 83 ml, consistent with monomeric forms of the proteins, in 10 mM sodium acetate pH 5, 150 mM NaCl. D) Coomassie stained 15 % SDS-PAGE gel analysis of purified *Sli*LPMO10E, *Sli*LPMO10E-Ext and H187A displaying single protein bands between 15-20 kDa consistent with the mass of the proteins (*Sli*LPMO10E– 18,422 Da, *Sli*LPMO10E-Ext -18,610 Da and H187A – 18,356 Da).

5.3.3. CD spectroscopy of LPMOs

Far UV-CD of *Sli*LPMO10E and the variants (-Ext and H187A) confirms that the proteins are folded and that the proteins consist of mainly β-sheets with some contribution from α-helices revealed by the minima at ~ 206 nm and maxima at ~ 197 nm (and at ~234 nm) (Figure 5.8). This coincides with predictions from the amino acid sequence (Figure 5.8) (Jpred4) [113] yielding 5.2 % α-helices and 23 % β-sheet content. When the far UV-CD of *Sli*LPMO10E was compared to *Sli*LPMO10B (Appendix 3.2), differences were clearly apparent with *Sli*LPMO10B displaying a broad minima at ~ 210-220 nm and maxima at ~ 196 nm. These differences demonstrate the more α-helical content of *Sli*LPMO10B compared to *Sli*LPMO10E as predicted from the sequence of *Sli*LPMO10B, producing 10 % α-helices and 21.5 % β-sheet from Jpred4 predictions (Appendix 3.2) [113].

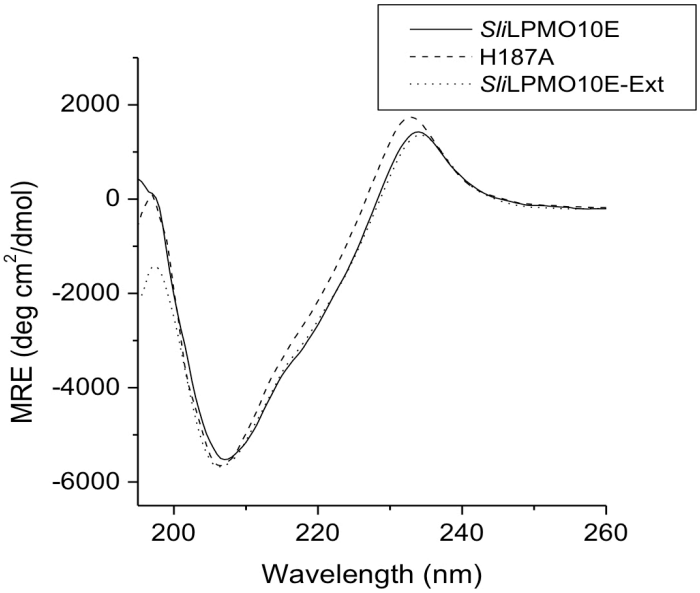


Figure 5.8: Far UV-CD spectra of *Sli*LPMO10E, H187A and *Sli*LPMO10E-Ext in 10 mM KPi, 50 mM KF, pH 7, 20 °C. Protein concentrations are all 20 μM with *Sli*LPMO10E – Solid line, H187A - dashed line and *Sli*LPMO10E-Ext – dotted line.

5.3.4. Cu(II) binding monitored by Fluorescence spectroscopy

Weak spectral transitions are observed upon Cu binding to LPMO proteins due to the nature of the Cu binding site, it is therefore necessary to monitor Cu binding through another avenue. Fluorescence spectroscopy was used to monitor Cu(II) binding. Quenching of the Trp fluorescence at 340 nm and 350 nm for *Sli*LPMO10E and *Sli*LPMO10B, respectively, was observed upon addition of a Cu(II) solution (Figure 5.9). A break point was reached at ~ 0.8 to 0.9 equivalents of Cu(II) to each protein, suggesting a binding stoichiometry close to 1:1 (Figure 5.9). Incubation of purified *Sli*LPMO10E or *Sli*LPMO10B for prolonged periods with excess EDTA and subsequent removal did not change the stoichiometry of the titration. Notably, there is no difference in Cu(I) or Cu(II) binding, thus all further studies were carried out with Cu(II). Purified samples of *Sli*LPMO10E were subjected to ICP-MS analysis to determine Cu content (performed by Robert Hider at Kings College London). The results from numerous preparations were always consistent with the presence of Cu but concentrations determined were always only ~ 3 % (0.03 mg atom/L) of the protein as suggested from the fluorescence titrations. Consequently, preparations of Cu-loaded LPMOs were made by the addition of stoichiometric amounts of Cu(II)SO_4, followed by desalting using a PD-10 (GE-Healthcare) column to remove any unbound Cu ions.

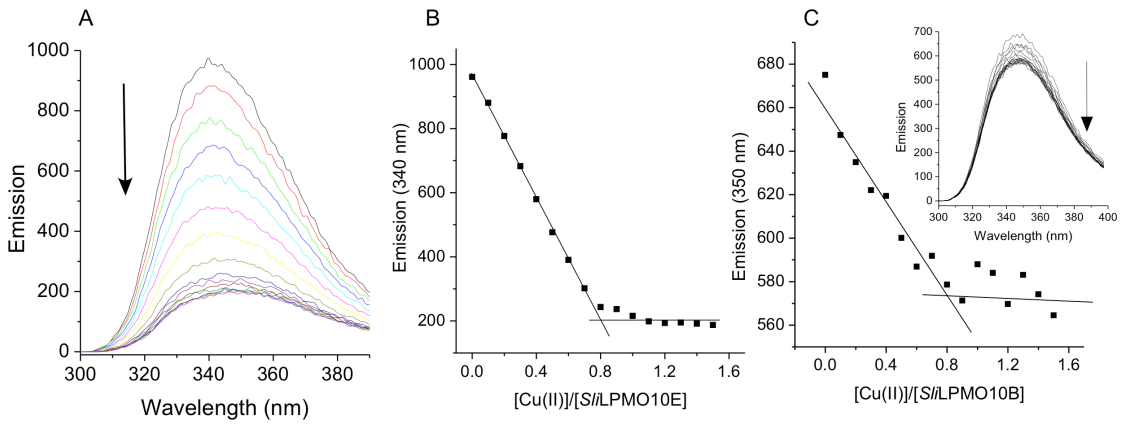


Figure 5.9: Fluorescence change upon Cu(II) binding to *Sli*LPMO10s from *S. lividans*. A) Changes in the Trp emission spectrum (Excitation 295 nm) (10 mM sodium acetate pH 5, 150 mM NaCl, 20 ^{o}C) of *Sli*LPMO10E (2 μM) upon titration with 1 mM Cu(II)SO_4. The arrow indicates the direction of the emission change. B) The emission at 340 nm plotted as a function of [Cu(II)]/[*Sli*LPMO10E]. The stoichiometry of the reaction is indicated by the intersection of the two lines. C) Inset, *Sli*LPMO10B (2 μM) upon titration with 1 mM Cu(II)SO_4. The arrow indicates the direction of the emission change. The emission at 350 nm plotted as a function of [Cu(II)]/[*Sli*LPMO10B]. The stoichiometry of the reaction is indicated by the intersection of the two lines.

5.3.5. EPR spectroscopy of SliLPMO10E

Apo-*Sli*LPMO10E gave a very weak EPR signal consistent with the presence in the sample of a small amount of Cu(II), in keeping with ICP-MS analysis (Figure 5.10.A). On preparation of holo-*Sli*LPMO10E, a much stronger EPR signal clearly originating from Cu(II) centres was obtained (Figure 5.10.B). Attempts to simulate this spectrum as a single Cu(II) species proved unsuccessful. A simulation was possible, however, if a linear combination of two different line-shapes was assumed. The two simulated spectra are shown in Figure 5.10.C, and the simulation parameters are reported in Table 5.8. Thus the experimental EPR spectrum recorded at 10 K for *Sli*LPMO10E is best approximated by a superposition of two different EPR lines, further referred to as Sim1 and Sim2, from two different Cu(II) complexes. Sim1 was simulated for three nitrogen atoms coordinating Cu and Sim2 for two. One noticeable difference between the line shapes of Sim1 and Sim2 is the presence of a strong 'overshoot' line in Sim2. This is a consequence of a combination of two factors: a smaller g-factor anisotropy (resulting in a smaller g_z value) and a larger A_z^{Cu} value in Sim2, than in Sim1. Sim1 and Sim2 were combined in equal proportions to produce a simulated spectrum equivalent to that of the Cu loaded sample thus assuming a roughly equal concentration of the two species (Figure 5.10.B).

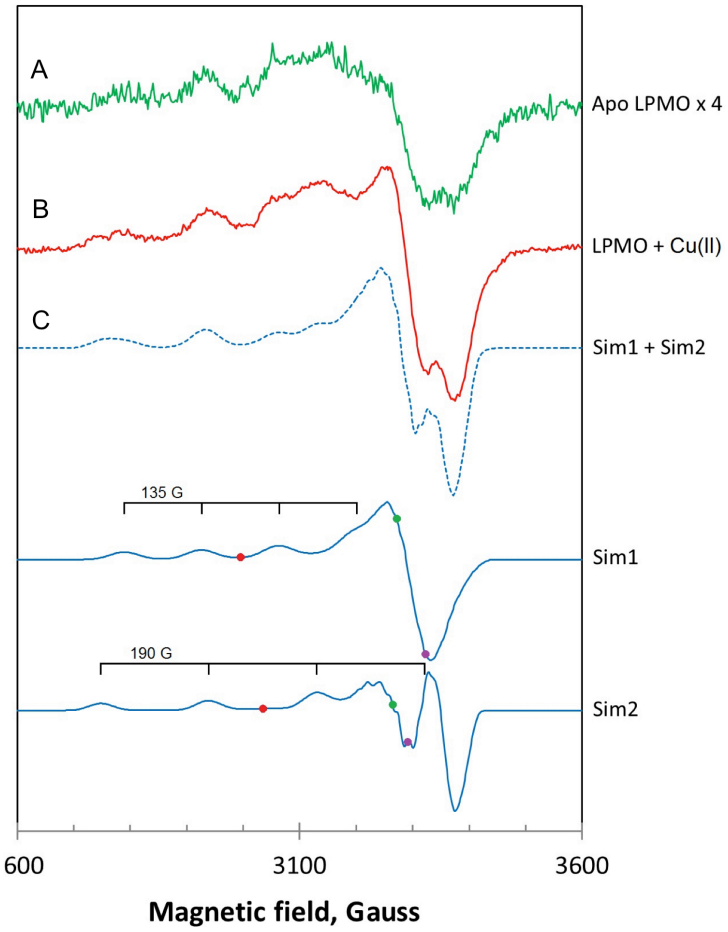


Figure 5.10: EPR spectroscopy of *Sli*LPMO10E at pH 7 and 10 K. A) As purified magnified x 4, B) loaded with 0.9 equivalents of Cu(II). The experimental spectrum in (B) (solid lines) has been approximated by linear combinations (dotted line) of the simulated EPR signals Sim1 and Sim2. C) Simulated EPR spectra of the Cu(II) ion coordinated by three N ligands (Sim1) and by two N ligands (Sim2). The coloured dots on the two simulated spectra indicate the positions of the principal g-values used in the simulations, magenta – g_x, green – g_y, red – g_z. Each simulated spectrum is labelled with the equidistant quartet tick marks, cantered at g_z and showing the positions of lines originated from the hyperfine interaction of the of Cu(II) electron spin (S = 1/2) with the Cu nuclear spin (I = 3/2). Other simulation parameters are given in Table 5.8.

Table 5.8: The EPR simulation parameters of *Sli*LPMO10E used to simulate spectra Sim1 and Sim2 in Figure 5.10. In bold are the g_z and hyperfine splitting constant (A_z^{Cu}) determined directly from the spectral parallel component.

		g (G)	A^{Cu} (G)	A^{N} (G)	ΔH (G)
Sim1[a]	x	2.033	13.5	14	12
	y	2.067	13.5	14	13
	z	**2.260**	**135**	14	26
Sim2[b]	x	2.055	19	14	12
	y	2.071	19	14	13
	z	**2.23**	**190**	14	23

[a]Simulated on 3-coordinating N atoms. [b]Simulated on 2-coordinating N atoms

5.3.6. Enzymatic activity of SliLPMO10E

SDS-PAGE gels were used to monitor initial substrate binding of *Sli*LPMO10E using cellulose and β-chitin. *Sli*LPMO10E clearly binds preferentially to β-chitin over cellulose based on the proportion of protein unbound relative to the bound fraction observed on the gel (Figure 5.11, inset). To assess substrate specificity and activity of *Sli*LPMO10E, Cu-loaded protein was incubated overnight in the presence of ascorbate as the reducing agent with a diverse range of glycan substrates. Analysis of the reaction products by MALDI-TOF MS revealed *Sli*LPMO10E activity only for β-chitin, in the presence of ascorbate (Figure 5.11). These observed results are mostly consistent with previous reports on chitin active AA10 LPMOs [291]. The breakdown products of β-chitin can be identified as predominantly reducing end aldonic acid oligosaccharides (C1 oxidation) with even numbered degrees of polymerization (DPn_{al}). However, not consistent was the observation of -2 Da (DPn_{-2}) species, which can either be due to reducing end lactones prior to ring opening or non-reducing end ketoaldoses resulting from C4 oxidation (Figure 5.11). Distinguishing these species by MS is challenging and so from the present data it is unclear whether *Sli*LPMO10E is capable of oxidizing at the C4 as well as the C1 position.

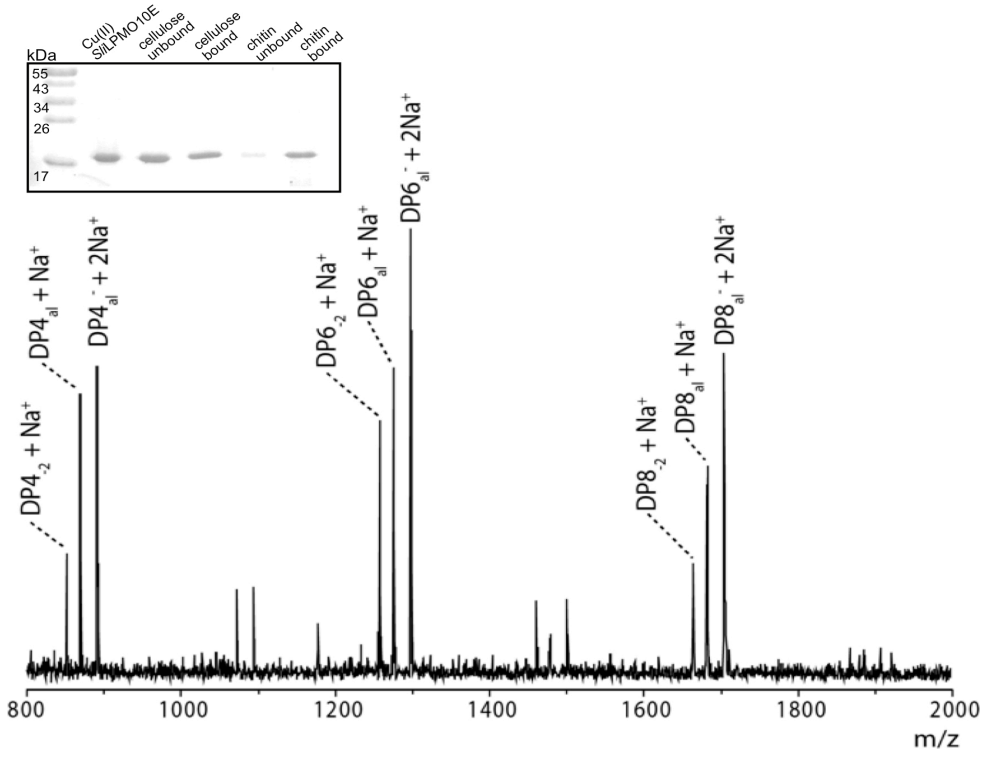


Figure 5.11: MALDI-TOF analysis of *Sli*LPMO10E products from squid-pen β chitin. DPn_{al} = aldonic acid, DPn_{-2} = lactone or C4 ketoaldose, (measured MW). DP4$_{-2}$ +Na^+ (851.3), DP4$_{al}$ +Na^+ (869.3), DP4$_{al^-}$ +2Na^+ (891.3), DP6$_{-2}$ +Na^+ (1257.5), DP6$_{al}$ +Na^+ (1275.5), DP6$_{al^-}$ +2Na^+ (1297.5), DP8$_{-2}$ +Na^+ (1663.6), DP8$_{al}$ +Na^+ (1681.7), DP8$_{al^-}$ +2Na^+ (1703.6). Inset, shows a 15 % SDS-PAGE gel of bound and unbound *Sli*LPMO10E fractions to cellulose and β-squid pen chitin, protein marker is indicated (*Sli*LPMO10E- 18,422 Da).

5.3.7. X-ray crystal structure of SliLPMO10E

*Sli*LPMO10E was loaded with 1 molar equivalent of Cu(II)SO_4 and used to screen a range of crystallisation conditions. Initial crystal clusters were identified in 0.1 M sodium acetate pH 4.6, 25 % PEG 4,000 (PEG suite, Qiagen) (Figure 5.12.A). These were then optimised in 24-well plates and small square crystals of Cu(II)- *Sli*LPMO10E grew within 2-3 weeks (Figure 5.12.B and C) in 0.05 M sodium acetate pH 4.6, 25 % PEG 4,000. Single LPMO crystals were used for X-ray measurements. Crystallographic data were collected at the Diamond Light Source and the X-ray structure of Cu(II) loaded *Sli*LPMO10E was determined to 1.38 Å resolution. One molecule was identified in the crystallographic asymmetric unit and a space group of C2 assigned. All residues (30 to 201) of *Sli*LPMO10E were modelled into a continuous and well-defined electron density with the overall structure of Cu(II)-*Sli*LPMO10E shown in Figure 5.13.A. The structure consists of a distorted β-sandwich fold containing seven anti-parallel β-strands also observed in other AA10 LPMO family members whose structures have been reported [282, 310, 313, 316]. This fold consists of two β-sheets, one containing three anti-parallel strands (S1, S3, S6) and the other, four antiparallel strands (S4, S5, S7, S8) (Figure 5.13.A). Between strands S1 and S3 lies a region designated as Loop2 (Figure 5.13.A,), which shows the most variability in size between LPMO families. In *Sli*LPMO10E this region consists of two α-helices (H1 and H2) and a small β-strand (S2). Two disulphide bonds are also present within the Loop2 region, Cys^{43}/Cys^{51} and Cys^{78}/Cys^{196}, with the latter connecting the Loop2 to the C-terminal β-strand (S8) (Figure 5.13). Disulphide bonds have been identified in other LPMO members and are suspected to confer stability and have been reported to be thermo-tolerant [282, 305]. In the vicinity of His^{30} and His^{120} a well-defined 16 σ peak is observed in the σ weighted F_o-F_c difference map into which a Cu ion was modelled. Three equatorial coordinate bonds to the Cu are formed from the Nδ1 and Nε2 of His^{30} and His^{120}, respectively, and the N-terminal amino-group of His^{30} to create the His-brace coordination signature (Figure 5.13.B). The bond lengths and angles are reported in Table 5.9, and are comparable to other AA10 LPMOs. No water molecule(s) were found to directly coordinate the Cu ion as seen in other AA10 members (Figure 5.13). A number of other residues in the vicinity of the Cu site that are commonly found in other characterised chitin-active AA10s are noted and include Ala^{118}, suggested to restrict access to the solvent facing axial position of the Cu site, also present in *Sc*LPMO10B from *S. coelicolor* (homologue to *Sli*LPMO10B) and is largely conserved with other AA10 members (Figure 5.5 : symbol) [308]. Another residue within the Cu site, Phe^{193}, was consistently found in AA10 members until the identification of a Tyr residue in the cellulose active *Sc*LPMO10B (homologue to *Sli*LPMO10B) resembling AA9 members (Figure 5.5 + symbol) (Figure 5.13.C) [313]. Also observed are the side chain orientations of Glu^{62} and Val^{186} in the vicinity of the Cu site acting to create a cavity, which is a feature found in other AA10s (Figure 5.13.C).

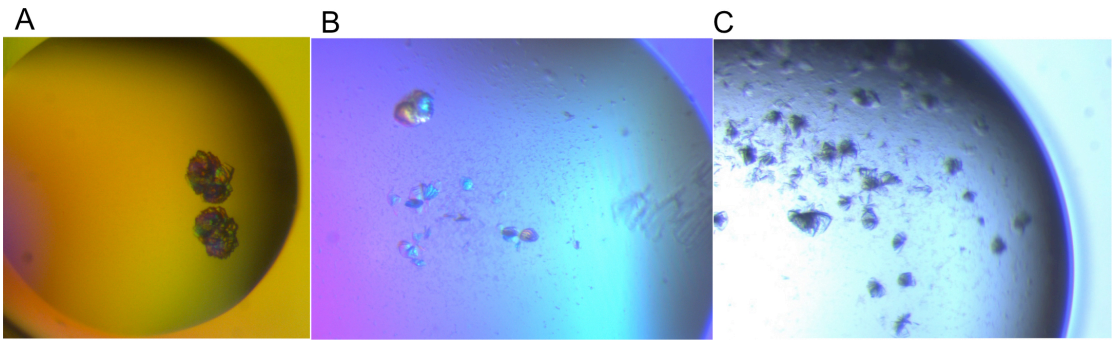


Figure 5.12: *Sli*LPMO10E crystals. A) *Sli*LPMO10E (15 mg ml^{-1}) crystals in a 96-well plate in 0.1 M sodium acetate pH 4.6, 25 % PEG 4,000 (PEG suite, Qiagen). B and C) Optimised crystals in 24-well plates in 0.05 M sodium acetate pH 4.6, 25 % PEG 4,000. A polarised filter was used in A and B to verify the presence of the small crystals.

An electrostatic surface representation of the putative substrate-binding surface of *Sli*LPMO10E reveals a flat surface with two small stretches of opposite surface charges, running either side of the positively charged Cu ion (Figure 5.13.D). Aromatic residues are often involved in enzyme-carbohydrate interactions and have been identified on the surface of AA9 LPMOs with a role suggested in substrate binding [193, 312]. A distinct surface aromatic residue, Trp^{56} located in the Loop2 region can be identified in *Sli*LPMO10E and is orientated such that its rings are parallel to the binding surface (Figure 5.13.D). This Trp aligns through sequence analysis (~ 45 % sequence identity) (Figure 5.5 ^ symbol) and structurally with Tyr^{54} in CBP21 (PDB 2BEM) [323]. Trp^{56} in *Sli*LPMO10E is also in a structurally equivalent position to Tyr^{79} of CelS2 and Trp^{88} in *Sc*LPMO10B (homologue to *Sli*LPMO10B) from *S. coelicolor* but align differently in sequence alignments (Figure 5.5 ^ symbol) [313]. This single aromatic residue is also conserved in a similar structural position in other AA10 LPMOs such as Trp^{58} in *Ef*LPMO10A (PDB 4A02) from *E. faecalis* [310] and Trp^{50} in *Ba*LPMO10A (PDB 2YOW) from *B. amyloliquefaciens* [282]. Therefore this residue is considered important for substrate binding and specificity.

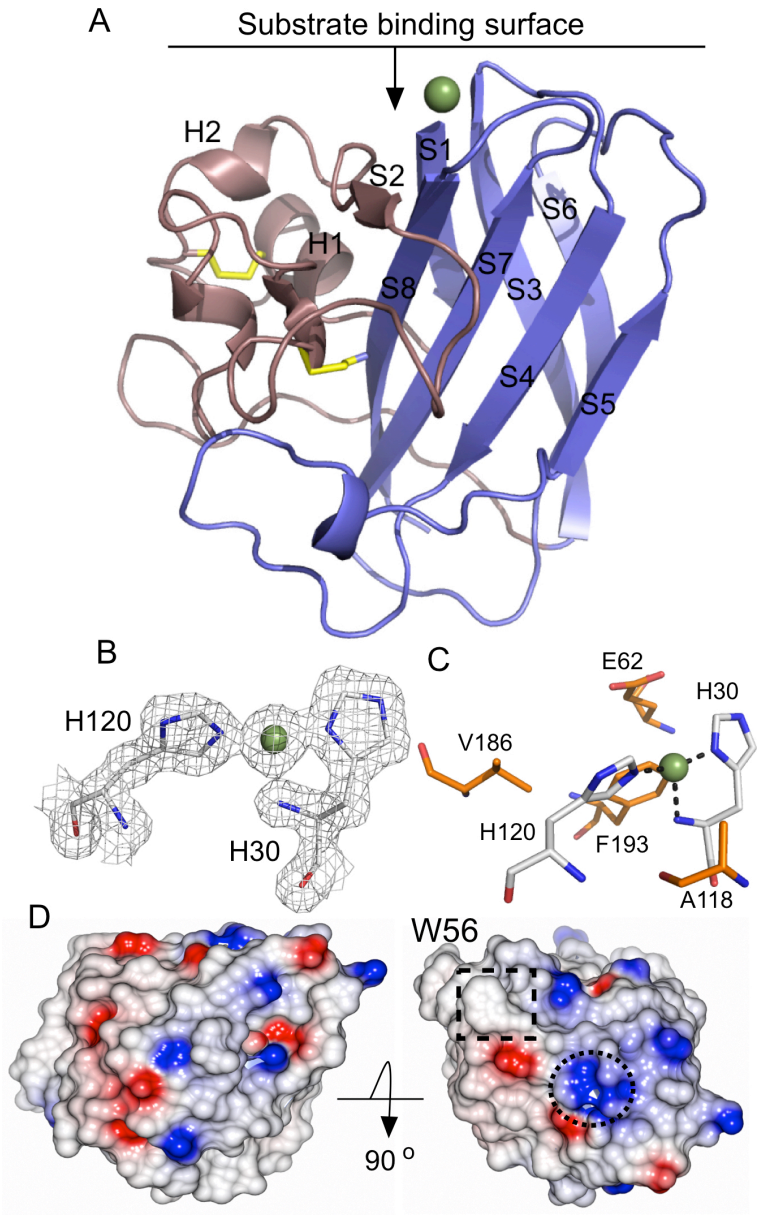


Figure 5.13: X-ray crystal structure of *Sli*LPMO10E. A) Tertiary structure shown in a cartoon representation with the core β-sheet fold shown in purple and the Loop2 region displayed in dirty violet. The two-disulphide bonds are shown in yellow and the Cu atom is represented as a sphere and coloured smudge. B) A 2 F_o-F_c electron-density map contoured at 1 σ of the Cu site of *Sli*LPMO10E, the Cu ion shown as a sphere, coordinated in a 'His-brace' structure by His^{120} and His^{30} shown as sticks. C) A close-up stick representation of the active site and surrounding residues. Coordinating His residues are coloured silver and residues surrounding the first Cu coordination sphere are shown in orange. D) Two electrostatic views of *Sli*LPMO10E, showing the proposed substrate-binding surface with W^{56} indicated by the dashed square and the Cu ion by dashed line circle. Images were prepared using CCP4MG and PyMol [121].

Table 5.9: Bond lengths and angles of active site and dimer site of *Sli*LPMO10E.

Bond	Length/Angle (Å, ^{o})
Active site	
His^{30} Nδ1	1.93
His^{30} N	2.24
His^{120} Nε2	1.91
His^{30} Nδ1-Cu-His^{30} N	95
His^{30} Nδ1-Cu-His^{120} Nε2	160
His^{30} N-Cu-His^{120} Nε2	103
MMD	
His^{187} Nε2	1.98
Asp^{188} Oδ2	2.25
His^{187} Nε2*	1.94
Asp^{188} Oδ2*	2.39

* Symmetry related molecule to form the metal mediated dimer (MMD)

During catalysis the Cu-site will be orientated towards the substrate and thus occluded from direct interaction with a redox partner or electron source. It has been proposed in LPMOs that electrons tunnel towards the Cu via an entry point on the distal surface that can be up to 21 Å away from the Cu site [193, 306]. Consistent with this, *Sli*LPMO10E possesses a chain of close-packed aromatic residues running through the β-sheet core to the Cu site (Figure 5.14.A). Putative electron entry points into this chain could occur at either Tyr^{110} or Phe^{151} (Figure 5.14.B), with a distance to the Cu site of 23 and 21 Å, respectively and are therefore favourable for long range biological electron-transfer.

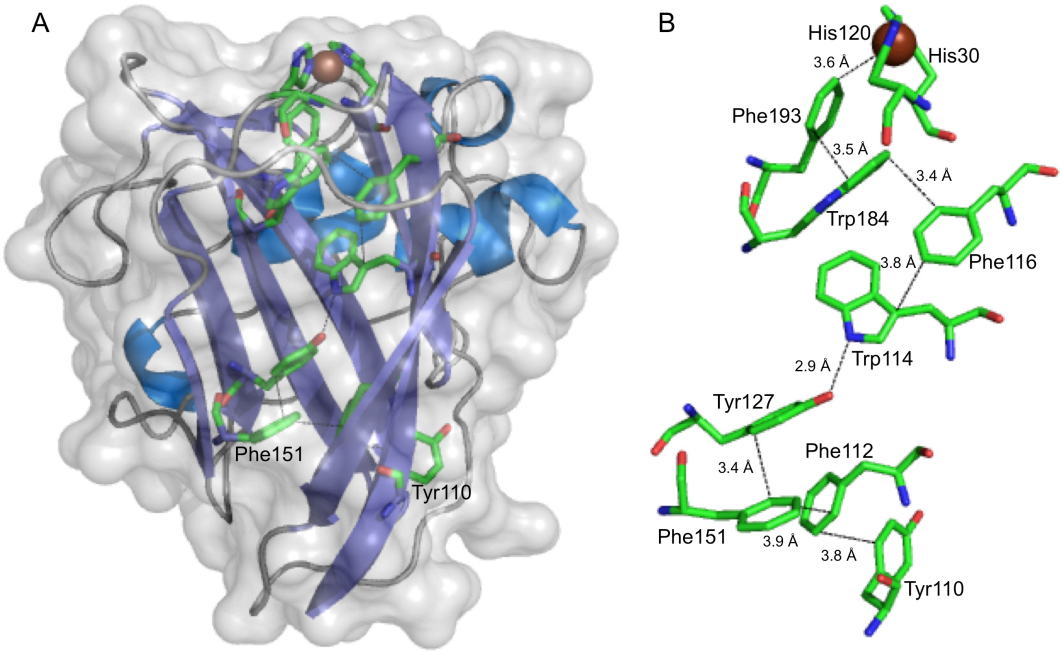


Figure 5.14: Putative electron transfer pathway from the Cu site to Tyr^{110}. A) Overall structure of *Sli*LPMO10E showing placement of the putative electron transfer pathway. B) The putative electron transfer pathway with residues involved shown as sticks. The Cu ion is shown as a brown sphere. Images were made in PyMOL.

In addition to the active site Cu ion identified by the 16 σ peak a second peak of 26 σ could be identified in the σ weighted F_o-F_c difference map of *Sli*LPMO10E assigned to a second metal ion. Generation of a symmetry related molecule of *Sli*LPMO10E revealed the presence of a metal mediated dimer, with the metal ion coordinated by His[187] and Asp[188] deriving from both monomers of the protein (Figure 5.16.A). Crystallisation conditions contained no metal-salts and the excess Cu(II) used to form the Cu(II)-loaded *Sli*LPMO10E was removed prior to crystallisation. Although every care was taken to ensure no serendipitous binding of rogue metals to the protein was taken there is the possibility that a divalent metal such as Zn(II) could out compete Cu(II) for this site. A fluorescence Cu edge scan of a Cu-*Sli*LPMO10E crystal was taken between 8.875-9.075 keV and produced a strong K-edge at ~8.993 keV consistent with Cu present in the crystal (Figure 5.15). However, to confirm the identity of the metals in both sites (Cu or Zn) anomalous difference maps were generated by measuring a single crystal at two different X-ray wavelengths (1.2 Å and 1.33 Å). At 1.2 Å the energy is above that of both Cu and Zn K-edges, whereas at 1.33 Å only the Cu K-edge contributes to the anomalous scattering. The anomalous difference maps generated showed similar σ weighted electron density peaks for the metal ion sites (active and dimer) at both wavelengths, therefore proving that Cu must be the metal ion present in both the active and the dimer site (if Zn was present there would be a decrease in the peak at 1.33 Å) (Figure 5.16.B). In the dimer site the Cu ion is coordinated with a distorted square planar geometry with the Nε2 atoms of the His residues linearly coordinated and the Oδ2 atoms of the aspartates are bent. Bond distances and angles are reported in Table 5.9. The two monomers of *Sli*LPMO10E are orientated such that the substrate binding surfaces remain exposed and fully accessible, thus suggesting a possible functional role. However, the stoichiometry derived from the fluorescence data of Cu ion binding is not consistent with such an assembly (*vide supra*) (Figure 5.9). Gel filtration chromatography also confirms that upon addition of Cu(II), *Sli*LPMO10E does not form a dimer assembly (Figure 5.17). No other known LPMO structures report the presence of a metal mediated dimer.

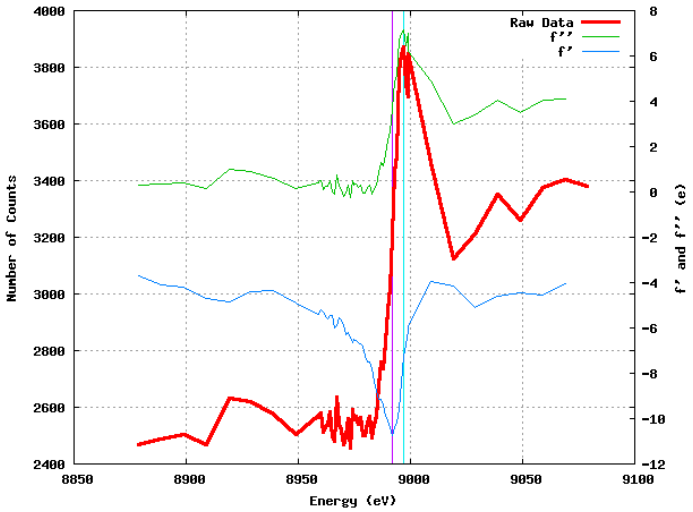


Figure 5.15: Fluorescence Cu-edge scan of a Cu-*Sli*LPMO10E crystal between 8875-9075 eV.

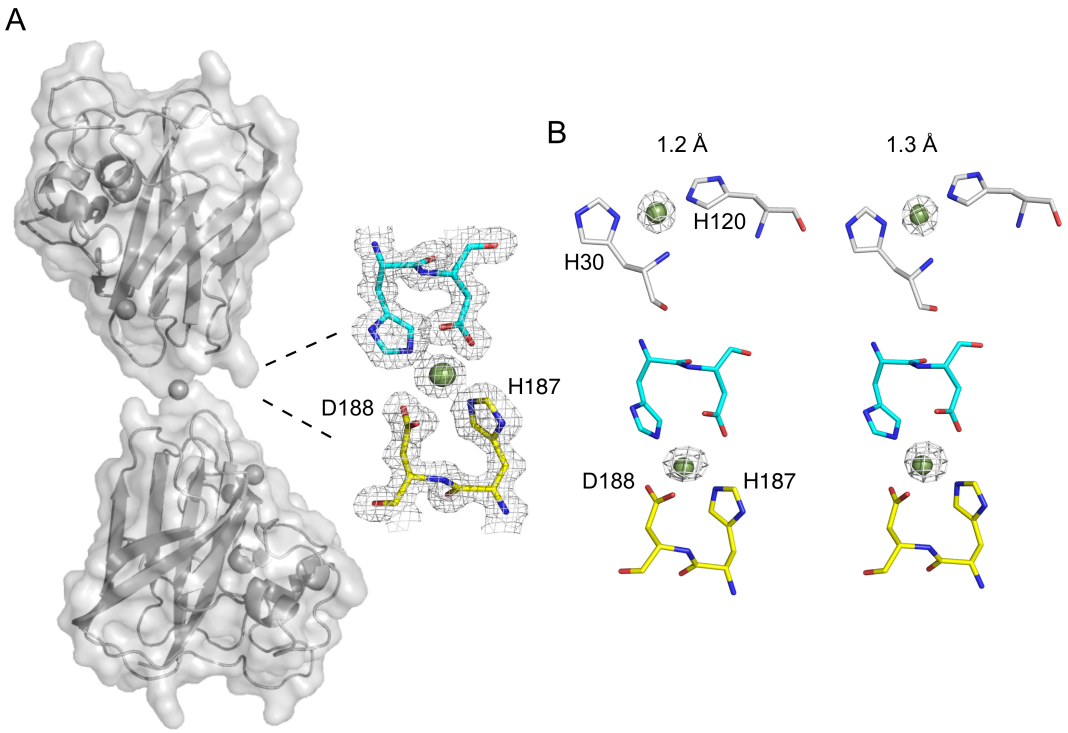


Figure 5.16: Properties of the metal mediated dimer of *Sli*LPMO10E. A) The dimer of *Sli*LPMO10E shown in grey with Cu ions represented as spheres. A 2 F_o – F_c electron density map contoured at 1 σ of the Cu coordination site at the dimer interface. Residues are labelled and shown as sticks with yellow from one monomer and cyan from the other, Cu is shown as a sphere. B) Anomalous difference maps contoured at 5 σ at 1.2 Å (above the Cu and Zn K-edges) and 1.33 Å (below the Zn but above the Cu K-edges) for the active and dimer sites. This is consistent with Cu being present in both sites.

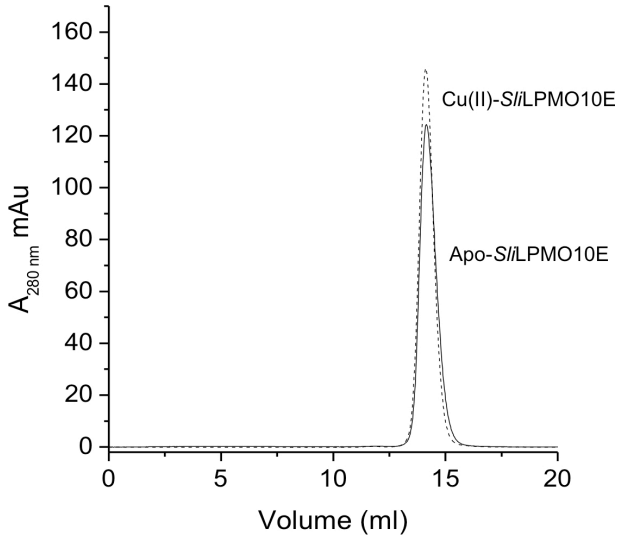


Figure 5.17: Analytical gel filtration of *Sli*LPMO10E. Apo-*Sli*LPMO10E (solid line) and *Sli*LPMO10E with 5-fold excess Cu(II) (dashed line) (300 µM *Sli*LPMO10E) in 10 mM sodium acetate pH 5, 150 mM NaCl.

5.3.8. Kinetics of Cu(II) binding to SliLPMO10E and H187A

To explore further the metal binding properties of *Sli*LPMO10E, fluorescence stopped-flow kinetic experiments were employed. Apo-*Sli*LPMO10E was mixed with various known concentrations of Cu(II)SO_4 and the total quenching of Trp fluorescence monitored as a function of time in a stopped-flow spectrophotometer. Upon mixing, two phases of fluorescence quenching were observed (Figure 5.18.A), with the observed pseudo first-order rate constants (k_{obs}) linearly dependent on [Cu(II)], with the slopes yielding second-order rate constants, k_1 and k_2 (Figure 5.18.B and Table 5.10). These linear fits run through or very close to the origin, showing that the affinity of Cu(II) binding is very high (*i.e* the k_{off} rates are very small meaning the dissociation rate constants for Cu are small compared to the association rates) and the relative amplitudes of the two phases are independent of the [Cu(II)] (Figure 5.18.B).

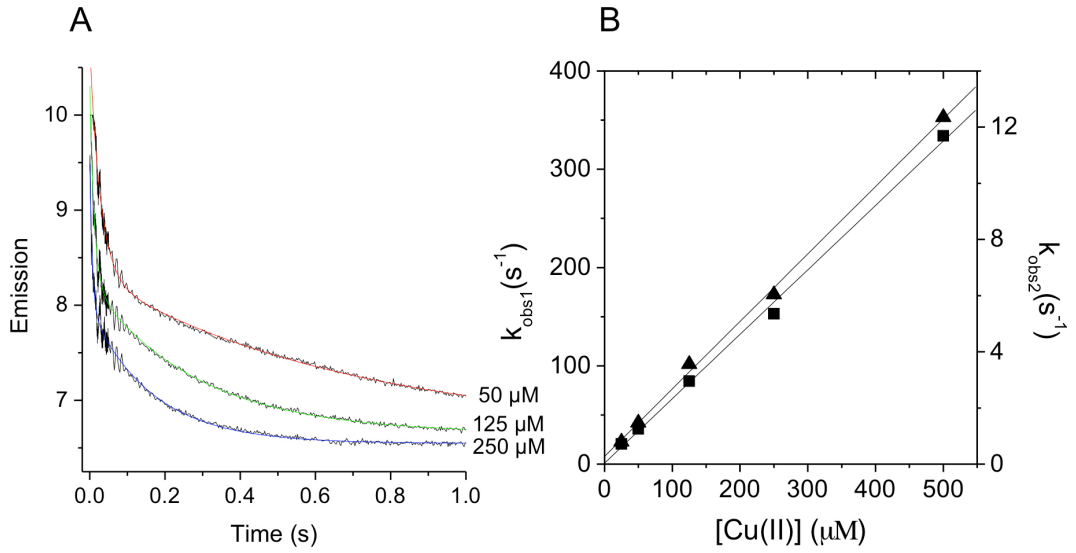


Figure 5.18: Fluorescence stopped-flow kinetics of Cu(II) binding to WT-*Sli*LPMO10E. A) Fluorescence time courses of WT-*Sli*LPMO10E (2.5 µM after mixing) reacting with various [Cu(II)], with example spectra and concentrations indicated. The colours (red, green and blue) indicate biphasic fits to the different [Cu(II)]. B) Observed rate constants of Cu(II) binding to *Sli*LPMO10E, for the fast (k_{obs1} in squares) and the slow (k_{obs2} in triangles) phase as a function of [Cu(II)] (10 mM sodium acetate, pH 5, 150 mM NaCl, 25 ^{o}C).

The stoichiometry of Cu(II) binding to *Sli*LPMO10E could be identified by loading the protein with *sub* to *super* stoichiometric amounts of Cu(II) followed by mixing with excess Cu(II) and monitoring the changes in the amplitudes of fluorescence emission. Upon mixing, two phases were retained, but as expected the total amplitude of the fluorescence quench decreased as the concentration of the Cu(II) preloaded into the protein increased (*i.e.* Cu(II) can now only bind to the empty sites). The relative amplitudes of the two phases changed as a result of pre-mixing with Cu(II), the faster phase being abolished at lower preloaded Cu(II) additions and the slow phase at higher preloaded [Cu(II)]. This allows the stoichiometry of each phase to be determined as illustrated in Figure 5.19. The faster phase titrated with a stoichiometry of ~ 0.5, and the slower phase, which titrates once the site responsible for the faster binding is filled, also titrates with a stoichiometry of ~ 0.5. Thus, the overall stoichiometry is 1:1, consistent with the interpretation of the static fluorescence titrations (Figure 5.19).

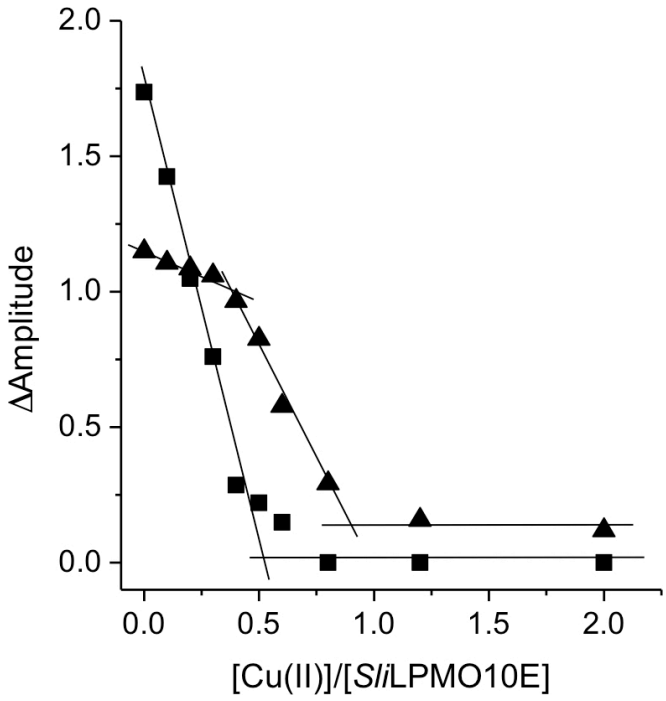


Figure 5.19: Cu(II) binding titration to *Sli*LPMO10E monitored by fluorescence stopped flow. Stoichiometry of Cu(II) binding to *Sli*LPMO10E monitored by fluorescence amplitude changes when *Sli*LPMO10E was pre-loaded with *sub* to *super* stoichiometric amounts of Cu(II) then rapidly mixed with excess Cu(II). The fast phase (k_{obs1}; squares) titrates with a ~ 0.5 stoichiometry with the slow phase (k_{obs2}; triangles) being observed following completion of the fast phase with a ~ 0.5 stoichiometry.

Although it was previously considered unlikely that the metal mediated dimer observed in the crystal (*vide supra*) existed in solution, the two [Cu(II)] dependent phases observed in the kinetic measurements promoted consideration as to whether the slow phase of Cu(II) binding was due to the presence of the second Cu binding site (dimer site). Kinetic experiments, with the H187A variant in which one of the dimer site ligands (His^{187}) was removed (H187A), were repeated. Essentially, identical results to the WT were obtained *i.e.* two clear phases both [Cu(II)] dependent (Figure 5.20) with k_1 and k_2 values reported in Table 5.10 and comparable to WT. Thus, the contribution of this dimer site observed in the crystal structure to the quenching of fluorescence in the WT protein is discounted and the kinetic data are most consistent with Cu(II) binding to two forms of a single site.

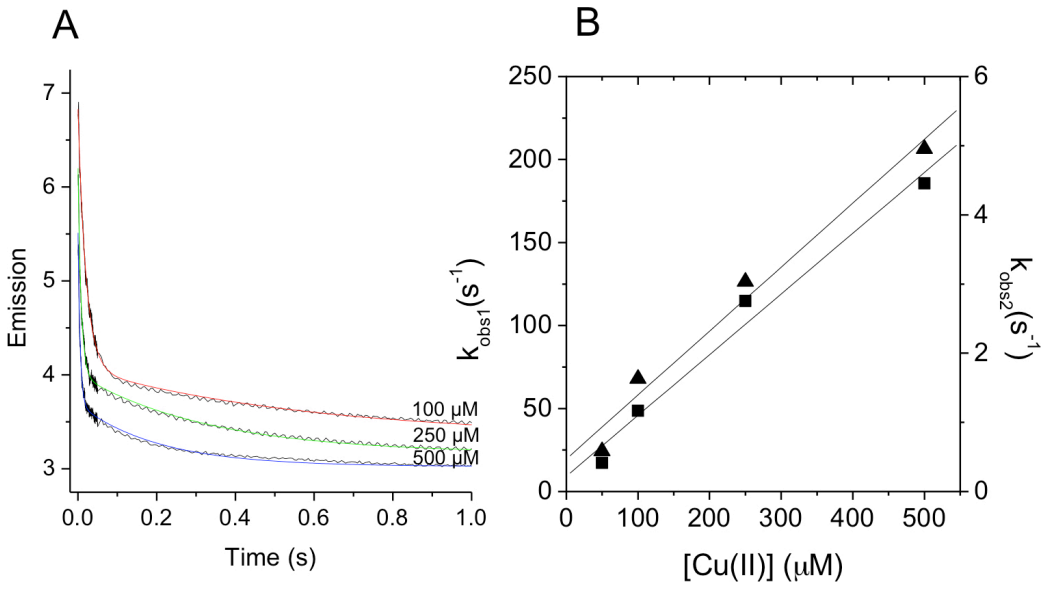


Figure 5.20: Fluorescence stopped-flow kinetics of Cu(II) binding to *Sli*LPMO10E H187A variant. A) Fluorescence time courses of H187A (5 µM after mixing) reacting with various concentrations of Cu(II), with example spectra and concentrations indicated. The colours (red, green and blue) indicate biphasic fits to the different [Cu(II)]. B) Observed rate constants of Cu(II) binding to H187A, for the fast (k_{obs1} in squares) and the slow (k_{obs2} in triangles) phase as a function of [Cu(II)] (10 mM sodium acetate, pH 5, 150 mM NaCl, 25 ^{o}C).

Table 5.10: Kinetic and thermodynamic parameters determined from stopped-flow fluorescence, fluorescence titrations and ITC at 25 ^{o}C for Cu(II) binding to the various forms of *Sli*LPMO10E.

[a]Fluorescence					
Protein	k_1 $M^{-1}s^{-1}$	k_2 $M^{-1}s^{-1}$	k_2 s^{-1}	k_3 s^{-1}	K_d M
*Sli*LPMO10E	6.6 (0.3) x 10^5	2.2 (0.04) x 10^4	-	-	-
H187A	3.6 (0.3) x 10^5	9.2 (0.3) x 10^3	-	-	-
*Sli*LPMO10E-Ext	2.0 (0.2) x 10^6	-	60 (1.7)	0.2 (0.01)	[b]2.4 x 10^{-6}
					[c]0.5 x 10^{-6}
*Sli*LPMO10B	1.0 (0.1) x 10^6	5.0 (0.2) x 10^4	-	-	-
[d]ITC					
Protein	*N*	K_d (M)	ΔG_b (kcal mol^{-1})	ΔH_b (kcal mol^{-1})	$-T\Delta S$ (kcal mol^{-1})
[e]*Sli*LPMO10E	0.48 (0.02)	62.5 (0.2) x 10^{-9}	-9.8 (0.3)	-7.9 (0.4)	-1.9 (0.8)
	0.43 (0.05)	4.2 (0.6) x 10^{-6}	-7.3 (0.2)	2.3 (0.5)	-9.6 (0.3)
[e]H187A	0.40 (0.05)	55.8 (2.0) x 10^{-9}	-9.9 (0.8)	-7.9 (0.2)	-2.0 (0.1)
	0.42 (0.1)	4.1 (0.2) x 10^{-6}	-7.4 (0.9)	2.5 (0.4)	-9.9 (0.3)
[f]*Sli*LPMO10E-Ext	0.94	2.6 x 10^{-6}	-7.6	20.0	-27.6

[a]Experimental uncertainties are reported in parenthesis and are the standard deviations from triplicate experiments. [b]Determined from static fluorescence titration. [c]Determined from stopped-flow fluorescence. [d]Experimental uncertainties are reported in parenthesis and are the standard deviations from triplicate experiments for *Sli*LPMO10E and duplicates for H187A and *Sli*LPMO10E-EXT. [e]Parameters derived from a sequential binding model. [f]Parameters derived from a one set of sites binding model.

5.3.9. Thermodynamics of Cu(II) binding to SliLPMO10E and H187A

To assess the thermodynamics of Cu(II) binding ITC studies were undertaken. Cu(II) was titrated into apo-*Sli*LPMO10E with a typical ITC profile observed in Figure 5.21, where it is apparent that two binding phases are detected. The first of these two phases is an exothermic binding process and the second an endothermic process. The profile could be fitted clearly to a two-site binding model corroborating the presence of the two distinct binding events with the parameters reported in Table 5.10. Notably, the stoichiometry for each binding phase is ~ 0.5 giving a total of 1 Cu(II) per *Sli*LPMO10E protein in keeping with the kinetic data. The first binding phase has a K_d for Cu(II) in the nM region and exhibits a favourable enthalpic binding (ΔH_b) contribution with a small but favourable entropic ($-T\Delta S$) contribution leading to an overall negative favourable free energy of binding (ΔG_b) (Table 5.10). The second phase displays a weaker affinity for Cu(II) with a K_d now in the μM range (Table 5.10). Moreover, this binding event is characterised by a positive ΔH_b (unfavourable) that is compensated for by an increased $-T\Delta S$, thus maintaining a favourable negative ΔG_b (Table 5.10).

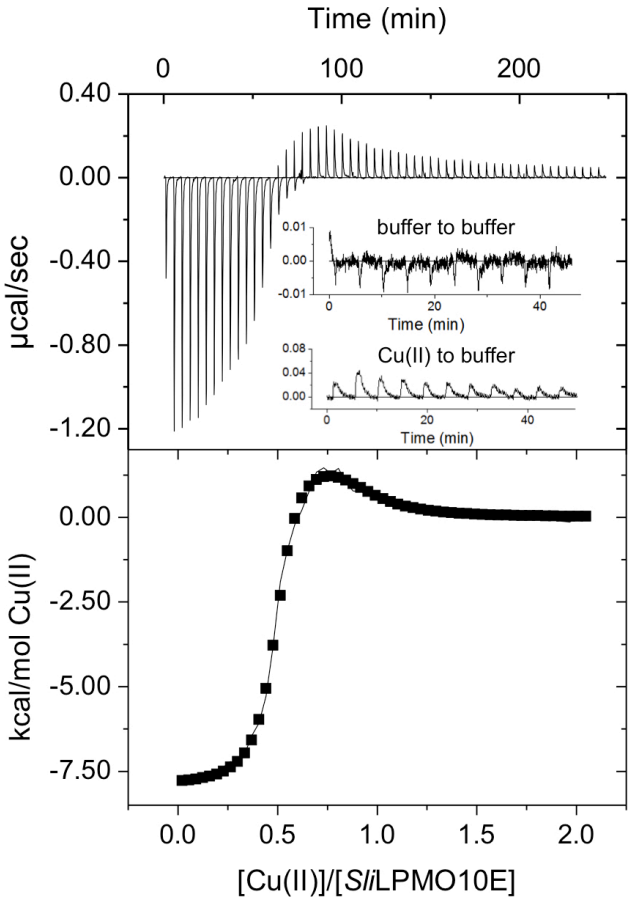


Figure 5.21: Thermodynamics of Cu(II) binding to WT-*Sli*LPMO10E. Cu(II) titration into *Sli*LPMO10E, shown as an ITC binding profile and as a fit to a sequential binding model (solid line). Inset shows the heat of dilutions from buffer to buffer and Cu(II) to buffer titrations The thermodynamic fit parameters are given in Table 5.10 (25 ^{o}C in 10 mM sodium acetate pH 5, 150 mM NaCl).

ITC experiments were repeated using the H187A variant of *Sli*LPMO10E and produced an essentially identical profile to the WT (Figure 5.21). The profile could be fitted to a two-site binding model and produced thermodynamic parameters reported in Table 5.10, which are indistinguishable from the WT protein. This reconfirms that the second metal site observed in the crystal is not that detected by either kinetics or ITC.

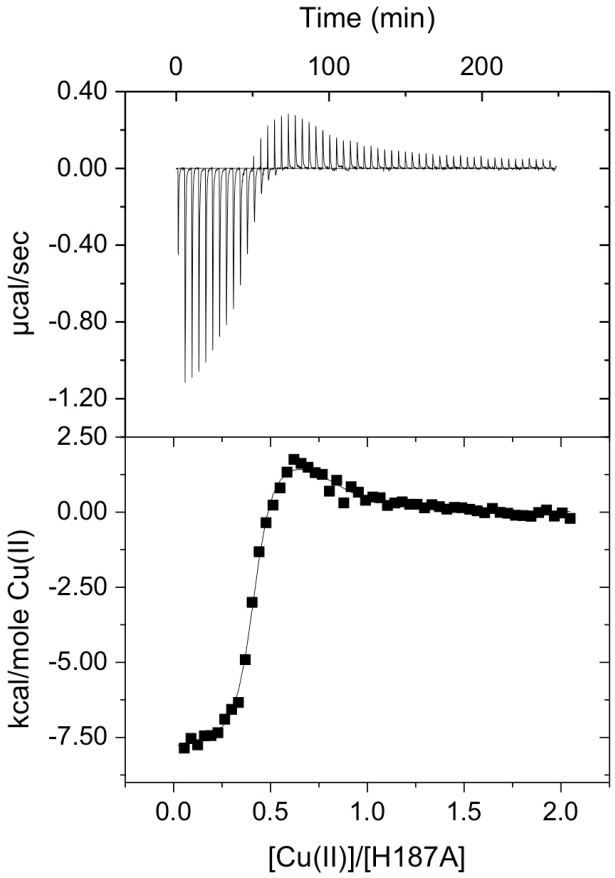


Figure 5.22: Thermodynamics of Cu(II) binding to the *Sli*LPMO10E H187A variant. ITC binding profile and fit to a sequential binding model (solid line) upon titrating H187A with Cu(II). Thermodynamic parameters obtained from the fit are reported in Table 5.10 (25 ^{o}C in 10 mM sodium acetate pH 5, 150 mM NaCl).

5.3.10. Kinetics of Cu(II) binding to SliLPMO10B

In order to elucidate whether the intricate Cu(II) binding observed in *Sli*LPMO10E is unique to this protein or is common to other AA10 LPMOs, *Sli*LPMO10B (homologue of *Sc*LPMO10B from *S. coelicolor*, 99 % genome sequence identity to *S. lividans*) was cloned, over-expressed and the Cu(II) binding determined. Size-exclusion chromatography was again consistent with a monomer species which migrated as a single band on a denaturing PAGE gel to a mass of ~ 20 kDa (predicted mass 20,723.4 Da for residues 43-229) (*vide supra*) (Appendix 3.1). Addition of Cu(II) aliquots quenched the Trp fluorescence at 350 nm, until a break point was reached at ~ 0.8 to 0.9 equivalents of Cu(II), again suggesting a binding stoichiometry close to 1:1 (*vide supra*) (Figure 5.9). On rapid mixing in the stopped-flow spectrophotometer with excess Cu(II) the fluorescence was quenched as seen for *Sli*LPMO10E but the signal was smaller as also noted from the static titration (Figure 5.23). However, time courses were again found to comprise two phases the rate constants of which were linearly dependent on [Cu(II)] (Figure 5.23 and Table 5.10). Therefore, these data demonstrate the presence of a single Cu(II) binding site in two forms generating distinct kinetic parameters comparable to those observed for *Sli*LPMO10E. It is thus clear that LPMOs other than *Sli*LPMO10E are capable of this unique binding nature.

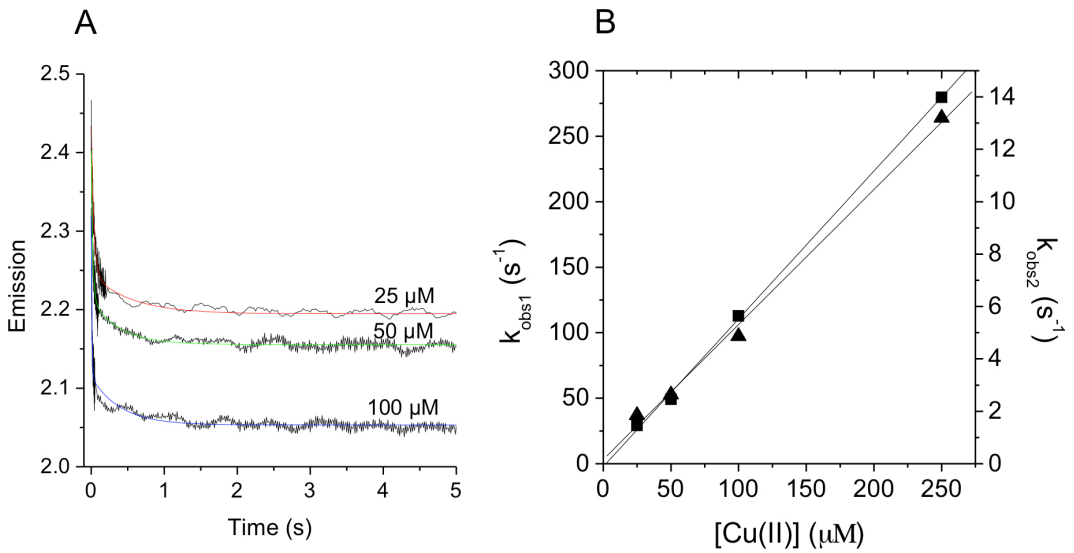


Figure 5.23: Fluorescence stopped-flow kinetics of Cu(II) binding to *Sli*LPMO10B. A) Fluorescence time courses of *Sli*LPMO10B (5 µM after mixing) reacting with various [Cu(II)], with example spectra and concentrations indicated. The colours (red, green and blue) indicate biphasic fits to the different [Cu(II)]. B) Observed rate constants of Cu(II) binding to *Sli*LPMO10B, for the fast (k_{obs1} in squares) and the slow (k_{obs2} in triangles) phase as a function of [Cu(II)] (10 mM sodium acetate, pH 5, 150 mM NaCl , 25 ^{o}C).

5.3.11. Kinetics and thermodynamics of Cu(II) binding to SliLPMO10E-Ext

To assess the importance of the 'His-brace' Cu active site in LPMOs an N-terminal extension protein (*Sli*LPMO10E-Ext) was created whereby His^{30} is no longer the first residue in the mature protein. Instead the mature protein has a Met and Gly before His^{30} (Met-Gly-His^{30}-) with the aim being to abolish/perturb the amino group coordination of the His-brace. *Sli*LPMO10E-Ext remains capable of binding Cu(II) as determined from the quenching of the Trp fluorescence in the static titration (Figure 5.24.A). However, rather than obtaining a stoichiometry as seen for WT- *Sli*LPMO10E a binding event could be observed allowing the fractional saturation to be calculated using Equation 5.6 and fitted using Equation 5.8 (Figure 5.24.B). *Sli*LPMO10E-Ext clearly possessed a weaker affinity for Cu(II) enabling a K_d to be determined (Table 5.10).

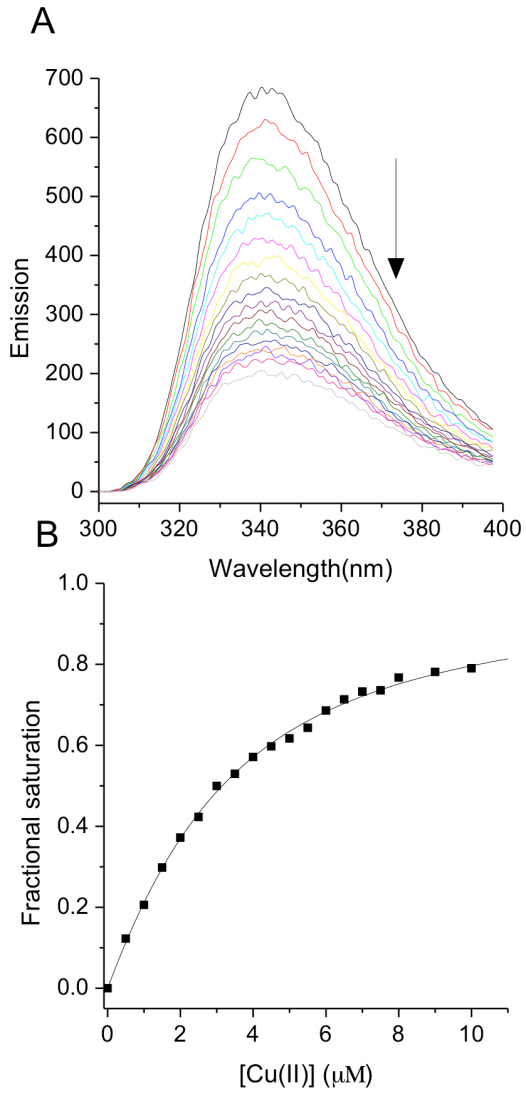


Figure 5.24: Fluorescence change upon Cu(II) binding to *Sli*LPMO10E-Ext. A) Changes in the Trp emission spectrum (Excitation 295 nm) (10 mM sodium acetate pH 5, 150 mM NaCl, 20 ^{o}C) of *Sli*LPMO10E-Ext (2 μM) upon titration with 1 mM Cu(II)SO_4. The arrow indicates the direction of the emission change. B) *Sli*LPMO10E-Ext plotted as a function of fractional saturation defined as $Y = 1 - \Delta F/\Delta F_T$ (Equation 5.6) where ΔF is the fluorescence change and ΔF_T is the fluorescent change from zero to fully saturated enzyme giving a K_d of 2.4 μM when fitted to Equation 5.8.

To evaluate the mechanism of Cu(II) binding to *Sli*LPMO10E-Ext, fluorescence stopped-flow kinetic measurements on rapidly mixing with excess Cu(II) yielded time courses now comprising of three distinct phases (Figure 5.25.A). The observed pseudo first-order rate constants (k_{obs}) for all three phases were plotted against [Cu(II)] (Figure 5.25). Only the fastest phase was linearly dependent on the [Cu(II)] (Figure 5.25.B) with a second-order rate constant (k_1) reported in Table 5.10. Although binding to this site is rapid it is with low affinity as seen by the intercept on the k_{obs1} axis, which is interpreted as the dissociation rate constant (k_{off}) of Cu(II) from this site (k_{off} = 230 s^{-1}) (Figure 5.25.B). Combining k_{off} with the k_1 gives a K_d of 1.2 x 10^{-4} M. Similarly, when the amplitudes of this fast phase were calculated in terms of fractional saturation and fitted to a hyperbola, this produced a K_d of 0.96 x 10^{-4} M (Figure 5.25.B, inset). This confirms that the initial Cu(II) binding of *Sli*LPMO10E-Ext is ~10^{-4} M, thus significantly lower than for the WT protein. The two slower kinetic phases display little [Cu(II)] dependence and are assigned to two first order reorganisation steps (k_2 and k_3 Table 5.10) in the Cu(II) bound protein that each lead to enhanced quenching of the Trp fluorescence. The observed rate of the third phase (k_{obs3}) could be fitted to a hyperbola Equation 3.3 as discussed in Chapter 3, in which when the first step is much faster than the second and third, a slight [Cu(II)] dependence can be observed (Figure 5.25.D).

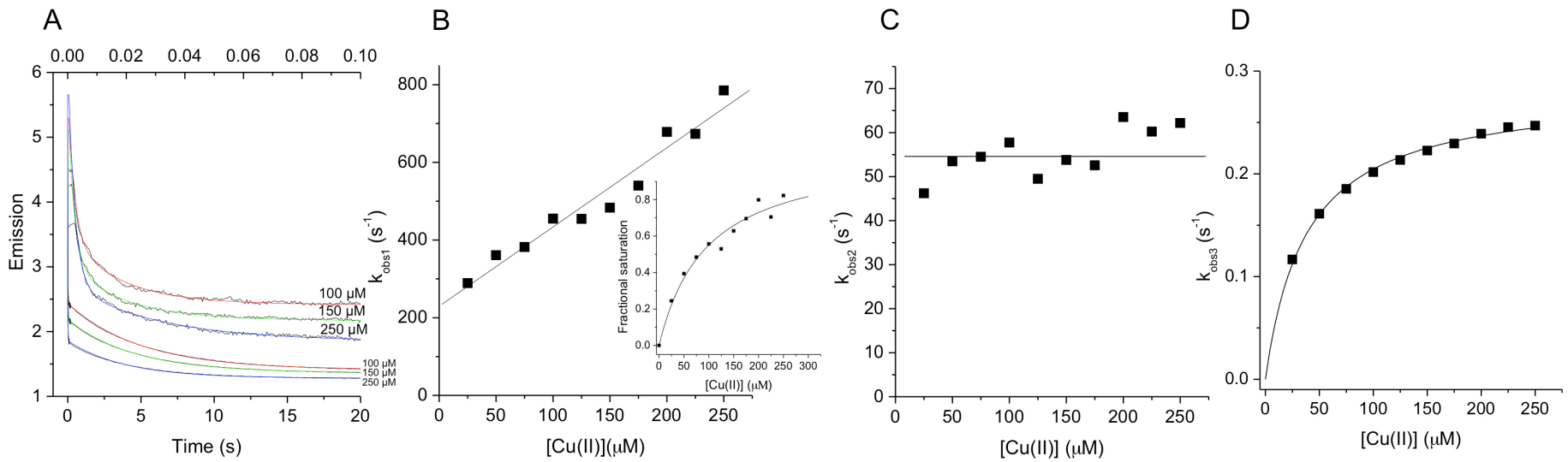


Figure 5.25: Fluorescence stopped-flow kinetics of Cu(II) binding to *Sli*LPMO10E-Ext (25 ^{o}C). A) Fluorescence time courses on reacting 5 µM (after mixing) *Sli*LPMO10E-Ext with various [Cu(II)] as indicated. The colours (red – 100 µM, green – 150 µM and blue – 250 µM) indicate triphasic fits to the different [Cu(II)]. Two x-axis are used to display the three phases in more detail. B) Observed rate constant for the fast phase (k_{obs1}) of Cu(II) binding to *Sli*LPMO10E-Ext. Inset, amplitudes calculated as fractional saturation defined as $Y = 1 - \Delta F/\Delta F_T$ (Equation 5.6) where ΔF is the fluorescence change and ΔF_T is the fluorescent change from zero to fully saturated enzyme, giving a K_d of 96 µM when fitted to Equation 5.7. C) Observed rate of the second phase (k_{obs2}). D) Observed rate of the third phase (k_{obs3}) fitted to Equation 3.3 (Chapter 3).

Titration experiments identical to WT were performed by mixing Cu(II) solutions with *Sli*LPMO10E-Ext preloaded with *sub* to *super* stoichiometric [Cu(II)]. Unlike for WT the three distinct phases all titrate together indicating binding to a single site that undergoes rearrangement. A striking result is that the K_d determined following preloading with Cu(II) (Figure 5.26 and Table 5.10), is two orders of magnitude lower (tighter) than the initial (fast) binding site (K_d 1.2 x 10^{-4} M) and in line with the K_d determined from static fluorescence titration experiments (Table 5.10). Together this indicates that following initial Cu(II) binding subsequent rearrangements detected in the slower kinetic phases lock Cu(II) into the site (< K_d). This is supported by the observed decrease in the amplitude of the slower phase as the [Cu(II)] with which it is mixed increases (Figure 5.26). This is expected when the [Cu(II)] span the value of the K_d for the initial binding site. Thus, at 100 μM Cu(II) the amplitude of the fast phase reflects half occupancy of this site, full binding then occurs at the rate of the slower phase. At 250 μM Cu(II), approximately 70 % of the initial site is rapidly filled and reflected in the amplitude of the initial binding process and the remaining sites are filled at the rate of the reorganisation (as explained by the hyperbolic equation).

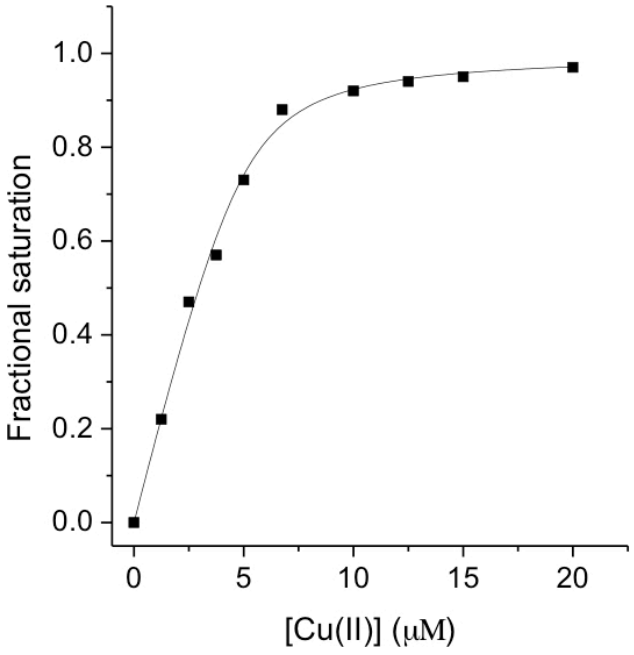


Figure 5.26: Titration of Cu(II) binding to *Sli*LPMO10E-Ext monitored by fluorescence stopped-flow kinetics. Stoichiometry of Cu(II) binding to *Sli*LPMO10E-Ext monitored by fluorescence amplitude change for all three phases when *Sli*LPMO10E-Ext was pre-loaded with *sub* to *super* stoichiometric amounts of Cu(II) then rapidly mixed with excess Cu(II). Plotted as a function of fractional saturation defined as $Y = 1 - \Delta F/\Delta F_T$ (Equation 5.6) where ΔF is the fluorescence change and ΔF_T is the fluorescent change from zero to fully saturated enzyme with Cu(II) versus Cu(II) concentration, give a K_d of 0.5 μM when fitted to the quadratic Equation 5.8.

The thermodynamics of Cu(II) binding to *Sli*LPMO10E-Ext were investigated by ITC. The profile in Figure 5.27 differs to WT protein showing a single endothermic binding event, which may be fitted to a single-site binding model to give the thermodynamic parameters reported in Table 5.10. Unlike the WT a single stoichiometry of ~ 1.0 Cu(II) per protein was determined with a K_d in the micromolar range correlating well to that determined from the static fluorescence titration and stopped-flow kinetic experiments (Table 5.10). An unfavourable ΔH_b is compensated for by a large favourable -TΔS contribution, maintaining a favourable ΔG_b (Table 5.10). These thermodynamic parameters are reminiscent to those determined for the second binding phase in the WT protein (Table 5.10) suggesting that this phase may stem from the binding of Cu(II) only to the side chain nitrogen's of the coordinating His residues.

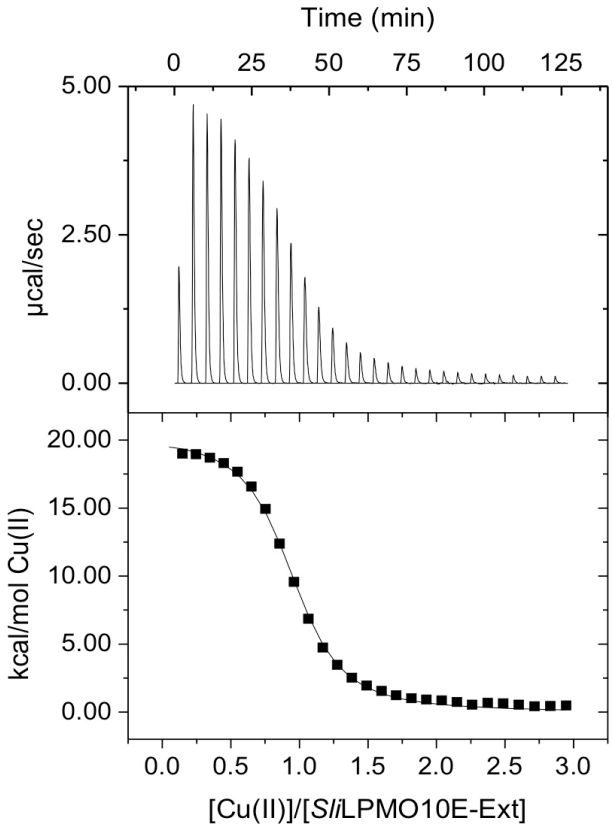


Figure 5.27: Thermodynamics of Cu(II) binding to *Sli*LPMO10E-Ext. ITC binding profile and fit to a single-site binding model (solid line) upon titrating *Sli*LPMO10E-Ext with Cu(II). Thermodynamic parameters obtained from the fit are reported in Table 5.10 (25 ^{o}C in 10 mM sodium acetate pH 5, 150 mM NaCl).

5.3.12. Redox potentials of SliLPMO10E and SliLPMO10E-Ext

The cell potential of the redox couples *Sli*LPMO10E-Cu(II)/*Sli*LPMO10E-Cu(I) and *Sli*LPMO10E-Ext-Cu(II)/*Sli*LPMO10E-Ext-Cu(I) were determined at pH 5 ($E_{m,5}$) using the redox indicator TMPD (Experimental procedures section 5.2.6). An $E_{m,5}$ of + 244 (± 7) mV was determined for WT-*Sli*LPMO10E (Figure 5.28.A), which aligns well with the homologue of *Sli*LPMO10B in *S. coelicolor*, *Sc*LPMO10B which has an $E_{m,5.5}$ of 251 (± 15) mV at pH 5.5 [313]. These E_m values also correlate well with those determined previously for other LPMOs under similar conditions. An $E_{m,5}$ of + 202 (± 4) mV was determined for *Sli*LPMO10E-Ext (Figure 5.28.B). This decrease of ~ 40 mV implies a destabilisation of the Cu(I) state, potentially due to loss of a coordinating Cu ligand.

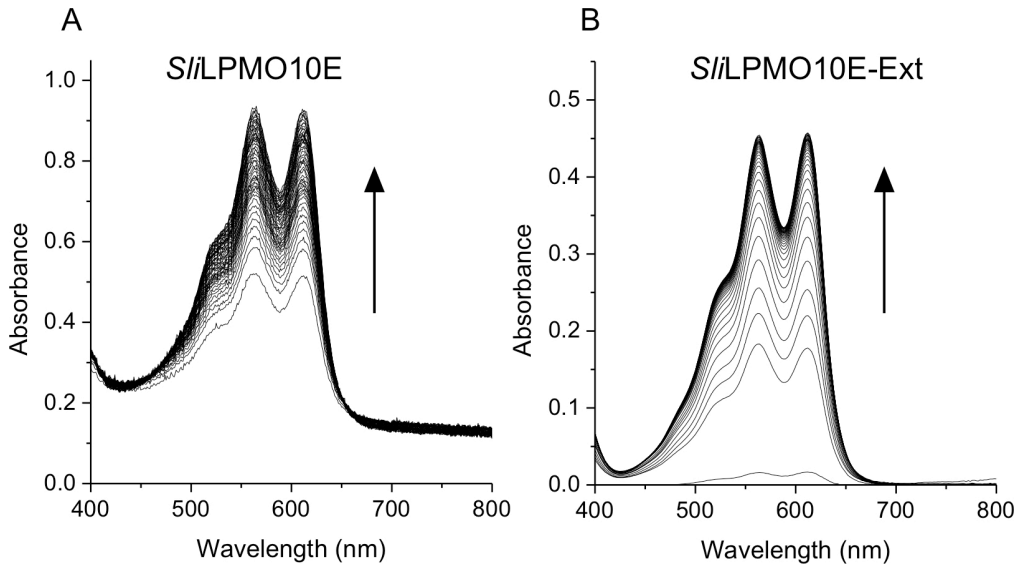


Figure 5.28: Redox potential measurements of *Sli*LPMO10E and *Sli*LPMO10E-Ext. Absorbance from the TMP radical cation (TMP_{ox}) is shown over time when mixed with Cu(II)-*Sli*LPMO10E (A) and Cu(II)-*Sli*LPMO10E-Ext (B) until no further absorbance change is observed. The absorbance at 610 nm was then used as the concentration of TMP_{ox} and LPMO-Cu(I) to calculate their redox potentials. Measurements were performed in 10 mM sodium acetate pH 5, 150 mM NaCl at 25 ^{o}C.

5.4. Discussion

5.4.1. LPMOs in Streptomycetes

Streptomycetes are economically and ecologically important microbes capable of secreting a host of CAZys, including LPMOs, in the presence of various plant polymers and polysaccharide substrates, which have the potential for the commercial construction of potent enzyme cocktails for industrial biomass decomposition [3, 322]. LPMOs may also act on internal substrates and be involved in assisting the degradation of glycans in response to cellular signals. By carrying out oxidative cleavage of otherwise inaccessible polysaccharide chains, LPMOs create access points for classical hydrolytic enzymes such as cellulases [302]. The interest in the context of this thesis stems from an extracytoplasmic glycan produced and modified by the CslA-GlxA combination for hyphal tip development in *S. lividans* (as discussed in Chapter 2) [166], which during the life cycle of the microbe will need to be degraded. Therefore, a combination of oxidative (*Sli*LPMO10E) and hydrolytic activity (CslZ) may be an efficient means to achieve this glycan breakdown.

5.4.2. The substrate of SliLPMO10E is chitin-like

Two cellulose degrading AA10s from *S. coelicolor* (*Sc*LPMO10B and *Sc*LPMO10C) and a chitin degrading AA10 from *S. gresius* (*Sg*LPMO10F), have been characterised [313, 324]. *Sc*LPMO10C and *Sg*LPMO10F have been shown to oxidise strictly at the C1 carbon of the sugar [313, 324], whereas *Sc*LPMO10B along with *Thermobifida fusca Tf*LPMO10A are examples of AA10s that can oxidise at the C1 and C4 positions. *Sli*LPMO10E is only active on chitin, with the predominant oxidation product indicative of C1 cleavage (Figure 5.11). Notably in order to measure the activity of *Sli*LPMO10E it was necessary to add 10-fold more enzyme than would typically be used in these assays. This may imply that chitin is not the direct target of *Sli*LPMO10E. Transcriptome and secretome analysis of *Listeria monocytogenes* recently demonstrated that the expression of its AA10 was not induced during growth on chitin [325]. Though the enzyme displayed chitinase-boosting activity, given its lack of expression when chitin was the sole carbon source it was suggested that it may play some other role during virulence, presumably by targeting host N-acetyl glucosamine containing polysaccharides [325]. The chemical composition of the glycan synthesized by CslA has not yet been characterized. Typically, genes involved in bacterial cellulose synthesis are organized in an operon, containing the *bcsABCD* genes [68]. However, *Streptomyces* lacks a c-di-GMP binding protein (BcsB), conserved in cellulose-producing organisms and essential for the synthesis of cellulose [69]. In light of this and the activity only for chitin hints at the possibility that the glycan produced by CslA has N-acetyl glucosamine functionalities (*i.e* chitin-like) enabling the constitutively expressed *Sli*LPMO10E to process.

Further support for *Sli*LPMO10E being specific for a glycan substrate containing N-acetyl glucosamine groups can be found from features inherent in the X-ray structure. In the structure the L2 region is smaller than for cellulose active members, and the electrostatics of the substrate-binding surface reveals clusters of positive and negative charge straddling the Cu site, in contrast to that of cellulose active AA10s where the positive Cu site is surrounded by predominately negative charge [313]. A cavity created through the orientation of the side chains of Glu^{62} and Val^{186} (Glu^{60} and Ile^{180} in CBP21) is absent in cellulose active AA10s *Sc*LPMO10B and *Sc*LPMO10C with this cavity suggested to accommodate oxygen or the acetyl group of an N-acetyl glucosamine moiety (Figure 5.29) [313]. The cellulose-active LPMOs have a positively charged side chain (Figure 5.29) (Arg^{212} in CelS2 and His^{214} in *Sc*LPMO10B) that fills the cavity present in the chitin-active LPMO10s (Figure 5.29) [313].

Spectral simulation (Sim1 and Sim2) of the EPR spectrum of *Sli*LPMO10E indicates the presence of two Cu(II) species. The g_z and $|A_z^{Cu}|$ values for these are 2.26/2.23 and 135/190 G respectively, with the 2.26 and 135 G pair representing the 'typical' distorted AA10 coordination where the three N atoms of the His brace are all coordinating to the Cu(II). The g_z and $|A_z^{Cu}|$ values determined directly from the spectral parallel region are the most reliable parameters for comparison with EPR data obtained from other LPMOs. In contrast the perpendicular g values (g_x and g_y) are estimated through simulation. From the present study it is clear that the simulation of spectra for a mixed species gives a significantly lower degree of rhombicity (difference between g_x and g_y) compared to other AA10s [314, 326]. Thus previous AA10 values may reflect a mixture of species rather than true rhombicity as originally reported [326]. For chitin active AA10s $|A_z|$ values of between 109 and 118 G have been reported [314], placing them between the Peisach-Blumberg classification of type 1 and type 2 Cu centres (Figure 5.30) [115]. In contrast cellulose active AA10s and AA9s have higher $|A_z|$ values of *c.a.* 147 G and are firmly within the type 2 classification (Figure 5.30) [314]. Thus *Sli*LPMO10E lies in between cellulose active AA9s/AA10s and chitin active AA10s, suggesting a more axial geometry compared to other chitin active AA10s. The second species has a lower g_z value of 2.23 and is in a clearly axial geometry ($|A_z|$ = 190 G). This could arise from a Cl^- ion binding to the Cu(II), as a consequence of 150 mM NaCl being present in the buffer. The axial geometry may arise from two N atoms coordinating to the Cu(II), with the decomplexation of one of the His residues or the N-terminal amino group a possibility, with the coordinated mono-atomic Cl^- ion able to move around until it overlaps well with the $d(x_2\text{-}y_2)$ orbital giving rise to the lower g_z of 2.23 value owing to enhanced covalency of the SOMO across the Cl ligand.

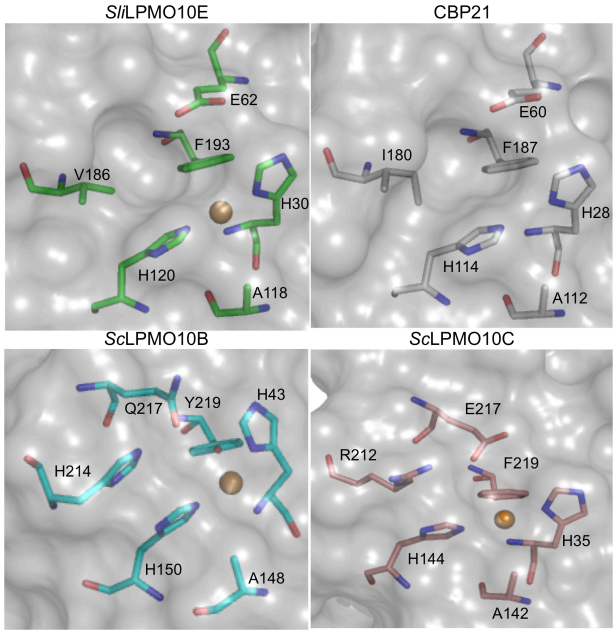


Figure 5.29: Comparison of cavities on the substrate-binding surface of AA10 LPMOs. The picture shows the surfaces of two chitin-active LPMOs, *Sli*LPMO10E from *S. lividans* and CBP21 from *S. marcescens* (PDB entry 2BEM) [323] and two cellulose-active LPMOs *Sc*LPMO10B (4OY6) and *Sc*LPMO10C (CelS2) from *S. coelicolor* (4OY7) [313]. The cellulose-active LPMOs have a positively charged side chain (Arg^{212} in *Sc*LPMO10C and His^{214} in *Sc*LPMO10B) that fills the cavity present in the chitin-active AA10 LPMOs. *Sli*LPMO10E and CBP21 have a hydrophobic side chain in the bottom of the cavity (Val^{186} in *Sli*LPMO10E and Ile^{180} in CBP21). Protein surfaces are shown in a transparent representation of the molecular surface. Amino acid side chains are shown in stick representation and the Cu ions are shown as brown coloured spheres.

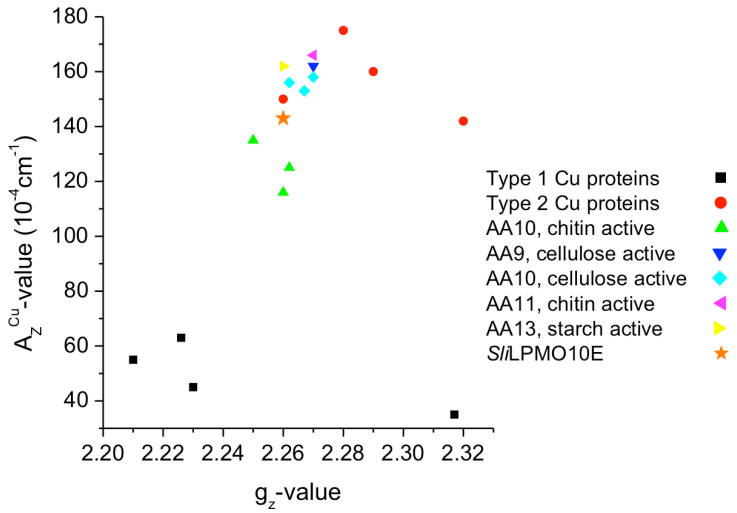


Figure 5.30: A Peisach-Blumberg plot of different LPMOs compared to some typical type 1 and type 2 Cu proteins. The type 1 Cu proteins are: horseradish umecyanin [327], plastocyanin [328], cucumber basic protein [329] and rusticyanin [330]. Type 2 proteins are: galactose oxidase [331], pig plasma benzylamine oxidase [332], pig kidney diamine oxidase [333] and spinach diphosphate carboxylase [334] (●); chitin active AA10s *Sm*LPMO10A (CBP21), *Bl*LPMO10A [314], and presumably chitin-active *Ba*LPMO10A [282](▲); cellulose-active AA9, *Ta*AA9A [287] (▼); cellulose-active AA10s, CelS2 (*Sc*LPMO10C), *Sc*LPMO10B and *Tf*LPMO10B (E8-N) [314] (◆); chitin-active AA11, *Ao*LPMO11 [305] (◀); starch-active AA13, *An*AA13 (▲) [306] and *Sli*LPMO10E (★).

5.4.3. SliLPMO10E performs C1 oxidation

The characteristic His-brace Cu coordination associated with all LPMOs is also observed for *Sli*LPMO10E, with each nitrogen ligand coordinating the Cu in an equatorial position. The Cu is likely to be in the Cu(I) state, favoured by this T-shaped geometry as a result of reduction from the high intensity X-ray beam. As described for other AA10s such as CBP21 a Phe and a conserved Ala residue are located in the vicinity of the Cu site with their hydrophobic side chains orientated so as to impede access of exogenous ligands to the axial positions of the Cu ion [316, 323]. This has led to the suggestion that dioxygen activation is most likely at the fourth equatorial position in AA10s resulting in strict C1 substrate oxidation. Conversely, access to the solvent-facing axial position has been suggested to be responsible for C1/C4 oxidation in AA9 members lacking an Ala as in *Ta*AA9 and in *Sc*LPMO10B where the Ala is present but its side chain is positioned such as not to impede ligand access. In *Sli*LPMO10E the Ala^{118} side chain superimposes with other strict C1 oxidising AA10s. This structural feature is consistent with the MS interpretation of C1 oxidised products of the chitin substrate and predicts that the -2 Da species identified in the MS likely arise from reducing end lactones rather than C4 oxidation products.

5.4.4. Metal-mediated dimer in SliLPMO10E crystal structure

The X-ray structure of *Sli*LPMO10E reveals a site for a second coordinating Cu ion. Although other LPMO structures have been shown to form dimer packing in the crystal structure of CBP21 and higher assembly packing with eight molecules identified in the structure of CelS2 none have identified a coordinating metal ion as seen in *Sli*LPMO10E [313, 323]. However, solution studies provide no evidence to indicate that this site is occupied or that the metal mediated dimer that can be generated in the crystal is formed. Despite this the orientation of the two monomers are not occluded from interaction with the substrate (Figure 5.16). This suggests a functional assembly with a distance between the two active site Cu ions of 29 Å. If such a dimer were to form then its organisation would allow for chain breaks in the substrate at defined distances that for chitin would correspond to six subunits.

5.4.5. Two Cu(II) binding modes

With the exception of *Sli*LPMO10D all other LPMOs identified in *S. lividans* 1326 are Sec substrates and are thus predicted to fold in the extracellular environment upon secretion. This suggests that Cu acquisition occurs during or post folding. Whilst buffering of metal ions in the cytoplasm, particularly Cu(I), is tightly regulated, metal concentrations in the oxidising extracytoplasmic environment are less well regulated. Thus designated trafficking systems may not always be operable and metallation can occur directly through partially hydrated metal ions or a polydisperse metal pool. The kinetic data of Cu(II) binding

to *Sli*LPMO10E is most consistent with a model in which Cu(II) binds to a single site in two binding forms, each with distinct kinetic parameters. This suggests that the apo-form of *Sli*LPMO10E possesses a Cu(II) binding site that exists in two energetically stable forms that do not interconvert on the timescale of the stopped-flow experiments. A possibility is that the Cu(II) bound forms have different coordination numbers as suggested from the EPR simulations, with one form having three N atoms coordinated and the second form two. Further support for this arises from the kinetics of the *Sli*LPMO10E-Ext variant. The presence of a single [Cu(II)] dependent step and a stoichiometry of one suggests that a single form, possibly with a N atom from each of the His side chains coordinating the Cu(II). Calorimetry measurements further advance the idea of Cu(II) binding at the same site in two forms each with a 0.5 stoichiometry. The first form possesses a high Cu(II) affinity with a favourable ΔH_b (three coordinate) whereas the second form (two coordinate) has a lower affinity with an unfavourable ΔH_b. As with the kinetic experiments the *Sli*LPMO10E-Ext variant behaves differently. A single binding event with a stoichiometry of one and an unfavourable ΔH_b and K_d mirroring the second binding form of the WT *Sli*LPMO10E is observed. This suggests that the two apo-forms are not present in the *Sli*LPMO10E-Ext variant and that Cu(II) binds to a single form. This leads to the suggestion that the first exothermic binding event for WT *Sli*LPMO10E is attributed to all three coordinating ligands, whereas the second, endothermic binding event, we attribute to a two coordinate form, which does not now utilise the N-terminal amino group.

Based on this data, a mechanism for Cu(II) binding in AA10 members is proposed in Figure 5.31. Two non-equilibrating apo-species exist (I and I*) with Cu(II) first binding to species I in an exothermic reaction with a nM K_d, which we attribute to Cu(II) coordinating all three N ligands of the His-brace. Once the Cu(II) binding site in species I is filled, the second apo-species (I*) binds Cu(II) in an endothermic manner with a μM K_d, and is attributed to Cu(II) coordinating two N ligands of the His brace, with water (low salt conditions) or Cl^- ions (high salt conditions) completing the coordination sphere (X in Figure 5.31). The mixed Cu(II)-bound species arising from two forms of a single site is further supported through the EPR measurements. However, the exact nature of the two coordinate N species is open to interpretation. In Figure 5.31 two possibilities have been depicted. The first whereby the second coordinating His is not bound to the Cu(II) and the second with an unbound Nt (amino-terminal). Furthermore, apo crystal structures of certain LPMOs reveal that the Nt is "flipped-out" (Hemsworth *et al.* unpublished data). This may suggest that one of the apo-species in solution is similar to that trapped in the crystal and binds Cu(II) with an unfavourable ΔH_b and a weaker affinity compared to the other apo-form, which likely has its amino group in the inward orientation to give a favourable ΔH_b and a higher binding affinity due to three N coordination to the Cu(II). Comparison of the kinetic and thermodynamic data for the *Sli*LPMO10E-Ext variant with that of WT *Sli*LPMO10E gives a strong indication that the two coordinate N species arises from the absence of Nt coordination (Figure 5.31).

Finally, the proposition that the kinetics and EPR results arise from two forms of a single site must be set against the structural data that shows a single unique Cu binding configuration. One way to account for this apparent discrepancy is by proposing that one of the forms discerned in the kinetic experiments resides at lower energy than the other form and that once Cu(II) is bound the whole of the protein slowly converts to the lower energy form with the Nt moving inwards and completing the three coordinate His-brace geometry seen in the X-ray crystal structure. This final conversion would be difficult to detect in fluorescence experiments, as it would occur in the Cu(II) bound form in which fluorescence is largely quenched. Alternatively, the three N Cu(II) coordinate species being of lower energy is the one 'selected' for crystallisation.

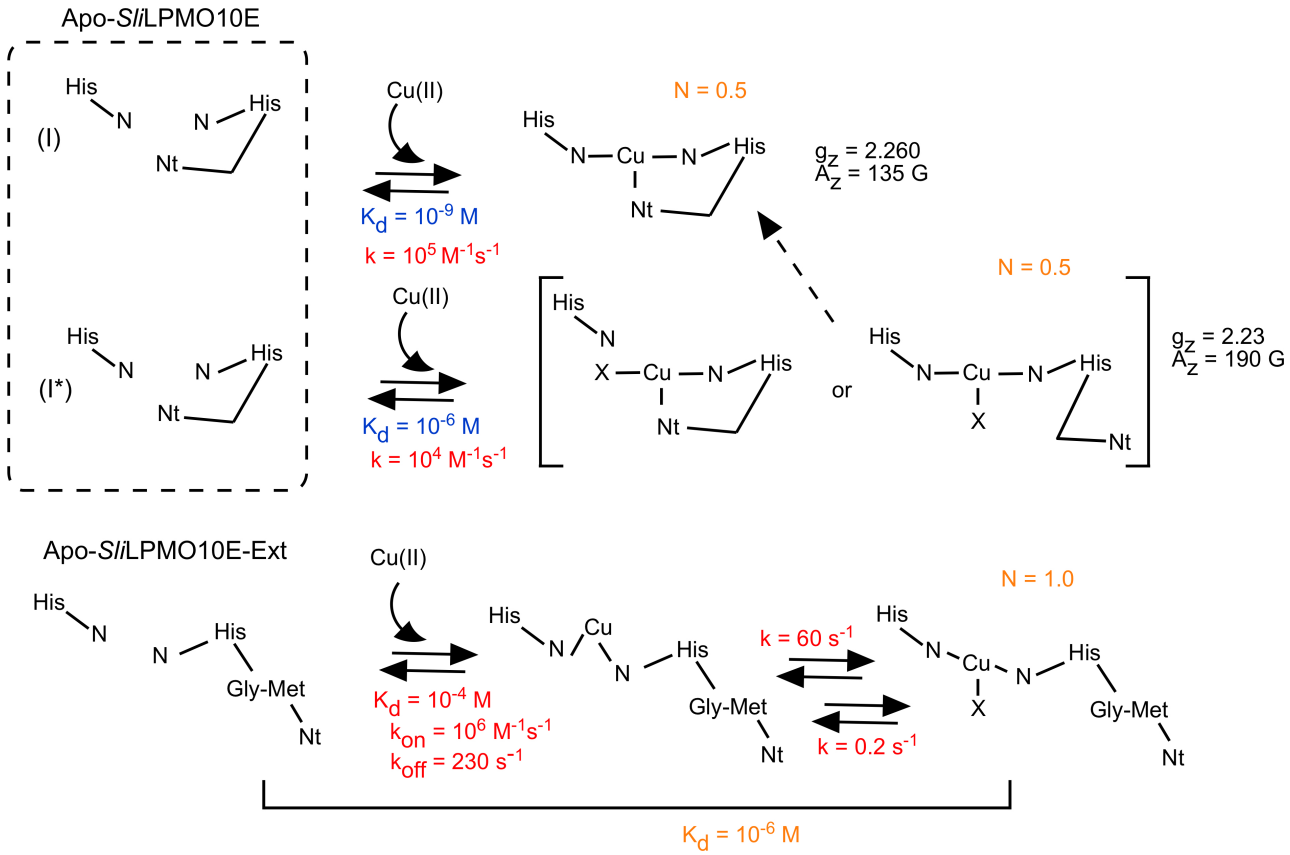
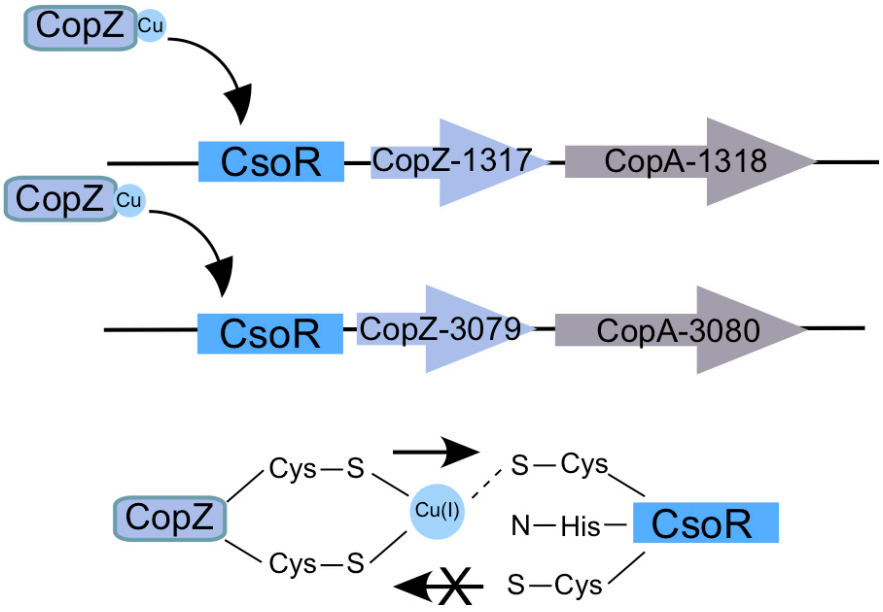


Figure 5.31: Mechanism of Cu(II) binding to *Sli*LPMO10E. Dissociation constants (K_d) and rates (k) are reported in blue (ITC measurement) and red (stopped-flow kinetics), with orange from both experimental approaches, including the stoichiometry of binding (N). X indicates bound water under low chloride concentrations and a chloride ion under higher concentrations.

Chapter Six

Copper trafficking in the CsoR regulon of *Streptomyces lividans*



Synopsis: Two out of four CopZ Cu(I) chaperones from *Streptomyces lividans* have been characterized and can traffic Cu(I) to CsoR under Cu stress conditions upregulating the CsoR regulon.

Some of the results from this Chapter have been published in:

Chaplin, A.K., Tan, B.G., Vijgenboom, E. and Worrall, J.A.R. "Copper trafficking in the CsoR regulon of *Streptomyces lividans*" 2015 *Metallomics* **7,** 145-55.

6.1. Introduction

Copper (Cu) is a vital transition metal in biology. The ability of Cu to undergo one electron redox cycling between Cu(I) and Cu(II) makes it a useful cofactor in many human enzymes including, Cu-Zn superoxide dismutase (SOD), cytochrome *c* oxidase, lysyl oxidase and ceruloplasmin [335-338]. Despite being an essential micronutrient, 'free' or uncomplexed Cu is toxic, with its ability to kill bacteria, yeast and viruses known since ancient times. The earliest medicinal use of Cu was recorded in the Edwin Smith Papyrus, a piece of medical text written by the ancient Egyptians that describes the use of Cu in sterilising chest wounds and drinking water [339, 340]. Cu was also widely used by the ancient Greeks, Romans and Aztec civilizations to treat a wide variety of infectious conditions until the modern era of antibiotics [339, 340]. Recently, the phenomenon of Cu toxicity has been exploited for its use in medical surfaces leading to it becoming the first solid antimicrobial material registered at the U.S. Environmental Protection Agency [341]. Cu manifests its toxicity by high affinity binding to adventitious sites and by the production of reactive oxygen species, responsible for devastating cellular damage [342-347]. Furthermore, recent studies have elucidated one of the primary mechanisms of Cu toxicity. It has been reported both *in vivo* and *in vitro* that under Cu stress conditions the maturation, stabilisation and activation of iron-sulphur (Fe-S) cluster enzymes become impaired [348, 349]. The Cu ion displaces the Fe atoms through coordination to the thiolate or inorganic sulphur ligands, resulting in cellular environments containing such enzymes not able to tolerate even low levels of Cu [348]. Thus, tight regulation of 'free' Cu is crucial for cell survival. In bacteria the complex relationship between proteins that sense, traffic and transport Cu is essential for maintaining the delicate balance between minimizing the toxicity of free Cu and ensuring the metabolic needs of the cell are met. Levels of unbound Cu within the cytoplasmic environment of bacteria are negligible ($< 10^{-18}$ M) with Cu metallochaperones playing a key role in maintaining cellular homeostasis [350].

Metallochaperone proteins were discovered in the 1990's and function as cytosolic cuprous ion transporters trafficking Cu(I) to specific high affinity target proteins or membrane transporters [343, 351-354]. Atx1 (anti-oxidant) from the yeast *Saccharomyces cerevisiae* was the first Cu metallochaperone to be discovered with an apparent role in protecting cells against toxicity from the superoxide anion (O_2^-) and hydrogen peroxide (H_2O_2) in strains lacking SOD [355-357]. Subsequent evidence led to the specific role of Atx1 in trafficking Cu(I) from the Cu ion transporter Ctrp1 to the P-type ATPase Ccc2, which pumps Cu into post golgi vesicles. This leads to Cu incorporation in multi-Cu oxidases such as Fet3 that, together with Ftr1 permease mediates high affinity iron up-take [351, 355, 358-360]. Following the identification of Atx1 the human homologue was identified denoted as HAH1 (human ATX1 homologue 1) or Atox1 (anti-oxidant protein 1) [361]. HAH1 transports Cu(I) to the N-terminal metal binding domain (MBD) of the ATP7A (Menkes) and

ATP7B (Wilsons) P-type ATPases [362-366]. Mutations in ATP7A or ATP7B disrupts the Cu homeostatic balance, leading to either Cu deficiency in Menkes disease or Cu overload in Wilson disease, respectively [365-367]. Subsequent Atx1 homologues have been identified in cyanobacteria (Atx1), plants (CCH), and bacteria (CopZ) [357, 368] all possessing a classic ferredoxin-like βαββαβ-fold. This consists of four β-strands creating an antiparallel β-sheet located underneath the two α-helices, thus creating an arrangement known as an 'open-faced β-sandwich' (Figure 6.1.A) [369]. A $MX_1CX_2X_3C$ metal binding motif (where X is any amino acid) is located on a solvent exposed loop connecting β-sheet 1 and α-helix 1 (Figure 6.1.A), with the two Cys residues creating a linear biscysteinate Cu(I) coordination [370, 371]. This exposed Cu(I) must play a functional role in these chaperones in their delivery and acquisition of the metal ion. The pH and ionization properties of the Cys residues have been reported to be determining kinetic factors in eukaryotic Cu(I) trafficking [370-372]. A single $MX_1CX_2X_3C$ motif and βαββαβ-fold is also present in CopA as well as two identified in Ccc2 and six in the Menkes and Wilson proteins.

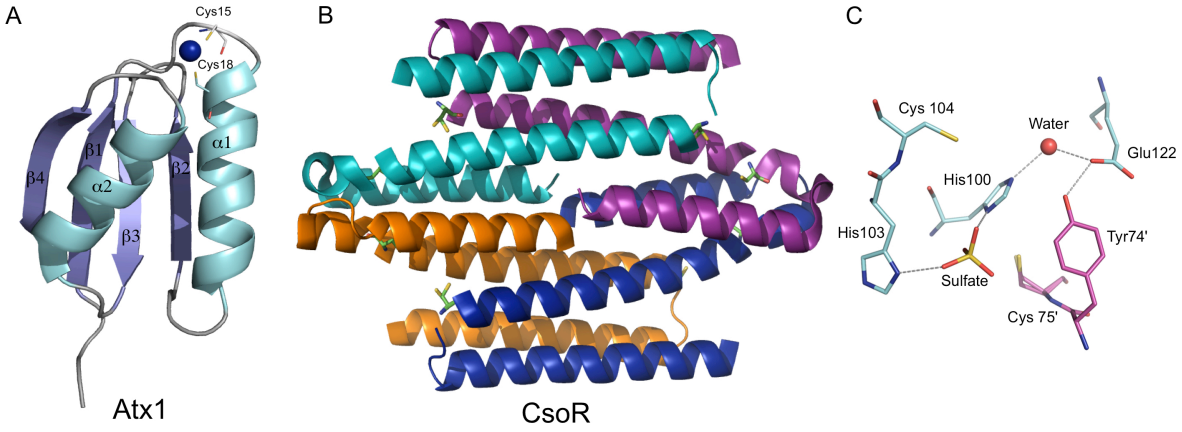


Figure 6.1: Structures of Atx1 from *S. cerevisiae* and CsoR from *S. lividans*. A) Solution structure of Cu(I)-Atx1 from *S. cerevisiae* with β-sheets shown in light blue, α-helices in cyan, loops in grey and the Cu ion shown as a blue sphere. Cu(I) is coordinated by Cys^{15} and Cys^{18} on the loop region between β1 and α1 with Sγ atoms coloured yellow (PDB accession number: 1FD8) [371]. B) Physiologically relevant tetramer assembly of apo-CsoR from *S. lividans* with protomers coloured blue, orange, teal and purple. The Cys residues predicted to be involved in Cu(I) binding are shown as sticks with the Sγ atom coloured yellow. C) The putative Cu(I) binding region of CsoR. Hydrogen bonds are shown as grey dashed lines Residues originating from each protomer are coloured either purple or teal. The proposed copper binding residues $Cys^{75'}$ and Cys^{104} lie some 7.6 Å apart and are separated by the imidazole side chain of His^{100}, which forms a hydrogen bond to a well ordered water molecule. A sulfate anion is present at the protein surface, forming bonds to His^{100} and His^{103} (PDB accession number: 4ADZ) [373].

The Atx1 homologue in bacteria is the CopZ metallochaperone, first identified in *Enterococcus hirae* as part of the *copYZAB* operon [374]. This operon consists of genes encoding two Cu P-type ATPases, CopA (initially identified as an importer but now all assumed as to be exporters) and CopB (exporter), a Cu responsive repressor CopY controlling expression of this operon and a CopZ metallochaperone [374-376]. Metalloregulatory proteins control the expression of these ATPase and metallochaperone proteins allowing them to quickly adapt to toxicity or deficiency of metal ions [377, 378]. In *E. hirae* transcription of the *copYZAB* operon is regulated by the transcriptional repressor Zn(II)-CopY [376]. During Cu stress conditions CopZ delivers two Cu(I) ions to the homodimer CopY repressor dissociating the bound Zn(II) ion leading to Cu(I)-CopY de-repressing the *cop* operon [374, 375]. Extensive work has also been carried out on the *copAZ* operon of *Bacillus subtilis*, encoding a P-type ATPase CopA and a Cu chaperone CopZ [379-381]. Delivery of a cuprous ion from the CopZ to the MBD's (two) of CopA (also possessing a βαββαβ-fold and a $MX_1CX_2X_3C$ motif) has been determined *in vitro* and *in vivo*, and is essential for regulating the activity of the ATPase [381-386] However, unlike CopZ from *E. hirae,* CopZ from *B. subtilis* can form higher order assemblies upon addition of Cu(I), with X-ray structures of tetranuclear and trinuclear assemblies reported, retaining their ability to transfer Cu(I) to CopA [387-389]. In *B. subtilis* and other bacteria, in the absence of the CopY repressor, a gene encoding for a copper sensitive operon repressor (CsoR) protein has been identified. A CsoR was first reported in *Mycobacterium tuberculosis* where it was shown to regulate the *rv0967-rv0970* operon in which *rv0969* encodes for a P_1-type ATPase involved in Cu transport and *rv0967* for CsoR [390, 391]. CsoR proteins are specific Cu(I) regulators that are widely spread across the bacterial genome, unlike the CopY repressor [390]. The structures of several CsoR members have been determined by X-ray crystallography and reveal a CsoR protomer consisting of three α-helices of varying lengths arranged in a disc-shaped D_2-symmetric homotetramer that is notable for the absence of a recognizable DNA-binding domain [373, 390, 392-394] (Figure 6.1.B). Under homeostasis conditions apo-CsoR binds to its DNA operator sequence(s) in a 2:1 CsoR tetramer:DNA stoichiometry, repressing the operon [395]. However, once cytoplasmic Cu(I) levels increase above the cells threshold, binding of Cu(I) to CsoR leads to an allosteric structural rearrangement releasing it from its target DNA [393, 395]. This in turn upregulates downstream expression of genes linked to Cu(I) efflux often including a CopZ-like chaperone and a CopA-like ATPase.

The soil dwelling antibiotic producing Gram-positive bacterium *Streptomyces lividans* displays a distinct dependence on the bioavailability of Cu in order to fully initiate morphological development [5, 23-25, 37, 38]. A gene *SLI4375* encoding a CsoR has been identified in *S. lividans* and shown to regulate a 3-loci regulon consisting of two *copZA*-like

operons and its own *csoR* gene (Figure 6.2) [373]. Under Cu stress conditions, the transcription levels of the *copZ/copA* genes increase in parallel with the eventual increase in expression of the *csoR* gene [373]. The two *copZA*-operons encode two putative CopZ-like metallochaperone proteins CopZ-1317 and CopZ-3079 and two CopA P-type ATPases CopA-1318 and CopA-3080 (Figure 6.2). CsoR from *S. lividans* regulates Cu(I) with attomolar affinity (10^{-18} M) coordinating the metal ion with His^{100}, Cys^{104} and $Cys^{75'}$ (Figure 6.1.C) [373]. The ability of CopZ proteins to safely shuttle cuprous ions to CopA-like ATPases for efflux under Cu stress conditions is largely accepted. However, CopZs may also harbour the ability to transfer Cu(I) to the DNA bound apo-CsoR regulator, similar to that observed for Cu(I) delivery to the CopY repressor in *E. hirae* [374, 375]. This possibility has not been explored.

To fully elucidate the importance of CopZ metallochaperones and to explore their Cu(I) trafficking potential, CopZ-1317 and CopZ-3079 both under the transcriptional control of CsoR in *S. lividans* were chosen for investigation [4, 373]. Spectroscopic techniques were utilised to determine the Cu(I) binding affinities and to assess the pK_a properties of the Cys residues in the $MX_1CX_2X_3C$ motif. CopZ-1317 variants were prepared (H22G and Y71F) to offer further insights into factors tuning the chaperoning properties of these CopZ proteins. Elucidation of higher order assemblies of the CopZ pair were investigated using size exclusion chromatography and together with EMSA studies allowed *in vitro* transfer of Cu(I) between the CopZ proteins and the CsoR to be studied. In combination with the re-analysis of RNA-seq data using the *S. lividans* genome as input, a more complete model to that reported previously [373] for the CsoR regulon in *S. lividans* is proposed.

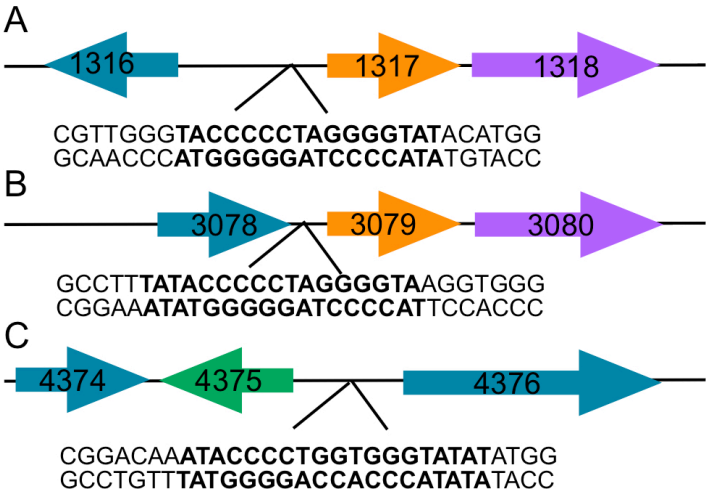


Figure 6.2: The 3-loci CsoR regulon in *S. lividans*. A and B) The gene environment for two operator DNA sequences of CsoR upstream of genes encoding for putative CopZ (1317 and 3079, orange) and CopA P-type ATPase (1318 and 3080, purple) proteins. C) The gene environment of *SLI4375* encoding CsoR (green) and its operator sequence. Sequences were predicted using PREDetector and confirmed experimentally by Dwarakanath *et al* 2012 with the consensus operator sequence shown in bold [373, 396].

6.2. Experimental procedures

6.2.1. Transcription analysis of csoR and the copZA operons in S. lividans

The following experimental procedures were carried out by collaborators at Leiden University (Dr. Erik Vijgenboom).

RNA-seq data obtained and previously reported [373] was analysed with the genome of *S. lividans* as reference, which has since then become available [4]. The total RNA was previously isolated with Kirby mix according to standard procedures from mycelium in early log phase. The transcriptome of wild-type (WT) *S. lividans* 1326 and the *ΔcsoR* mutant strain were obtained without addition of Cu to the growth medium and with addition of Cu, in which the induced transcriptome was obtained after a 2 hour exposure to 400 μM Cu(II). RNA integrity was confirmed by agarose gel electrophoresis and the absence of genomic DNA checked by PCR. Transcriptome analysis by RNA-seq was then carried out and the expression measured in RPKM defined as the reads/kb of exon/million mapped reads *i.e.* dividing the total number of exon reads (in this case one exon per reference sequence) by the number of mapped reads (in millions) times the exon length (in kb) [373] [397].

6.2.2. Cloning of SLI1317 and SLI3079

Full-length *SLI1317* and *SLI3079* genes (234 and 243 base pairs, respectively) were amplified from the genomic DNA of *S. lividans* (*S. lividans* 66, stock number 1326 John Innes Centre). The genes were then ligated into the NdeI and HindIII restriction sites of a pET28a vector (Novagen) to create N-terminal His_6-tagged constructs for over-expression in *Escherichia coli*. Primers were created containing NdeI and HindIII restriction sites (see below) to amplify the two CopZ genes from the genomic DNA of *S. lividans*. Table 6.1 displays the composition of the reagents and Table 6.2 the PCR protocol used to amplify *SLI1317* and *SLI3079*. The resulting PCR products were analysed on an agarose gel (1 % w/v) and gel extracted using a QIAquick gel extraction kit. The CopZ genes were then ligated into the NdeI and HindIII sites of a pET28a (Kan^r) vector and transformed into *E. coli* JM109 cells and plated on LB (Luria broth) kanamycin (Kan 50 μg ml^{-1}) agar plates. Colonies were selected, mini-prepped (Fermentas GeneJet Plasmid Miniprep kit) and sent for sequencing to corroborate the presence of the *SLI1317* and *SLI3079* genes. This gave *pET1317* and *pET3079* plasmids consisting of N-terminal His_6-tagged constructs with thrombin cleavage sites between the His_6-tag and the N-terminal start codons, under the control of the T7 promoter.

Forward and reverse primers used to amplify *SLI1317* and *SLI3079* genes from the genomic DNA of *S. lividans*. The restriction sites for NdeI and HindIII are underlined.

SLI1317- Forward

5′-GTTACACATATGCCCTCGAACGTCACCG-3′

SLI1317- Reverse

5′-CTTAAAGCTTTCAGACGCGGCCGGTCAG-3′

SLI3079 – Forward

5′-GTTACACATATGACCGCCCAGACCGAC-3′

SLI3079 - Reverse

5′-CTTAAAGCTTTCAGACCCGGCCGACGAGC-3′

Table 6.1: The reagents and volumes used to amplify CopZ-*1317* and *CopZ-3079* from the genomic DNA of *S. lividans* 1326.

Reagent	Concentration	Volume (μl)
Genomic *S. lividans* 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 6.2: PCR protocol used to amplify *CopZ-1317* and *CopZ-3079* from the genomic DNA of *S. lividans*. The values in bold indicate that these steps were cycled 35 times and the remaining steps only once.

Temperature (°C)	Time (min)
95	3
95	**1**
52	**0.5**
72	**2**
72	7.5
4	Final Extension

6.2.3. Site-directed mutagenesis

The Quikchange (Stratagene) site-directed mutagenesis method was used to create the H22G and Y71F variants of CopZ-1317. The nucleotide and corresponding amino acid sequence of CopZ-1317 with the respective codon changes for His^{22} and Tyr^{71} are shown in red. Forward and reverse mutagenic primers with the respective nucleotide change(s) used to amplify H22G and Y71F are shown. Table 6.3 gives the reagents and Table 6.4 the protocol used to amplify both variants from the *pET1317* construct.

CopZ-1317 nucleotide and amino acid sequence, red are the codons to be changed to create the respective mutations.

ATG	CCC	TCG	AAC	GTC	ACC	GCC	CCC	GTC	ACC	ACC
M	P	S	N	V	T	A	P	V	T	T
GCC	TAC	GCC	GTC	GCC	GGC	ATG	AGC	TGC	GGT	CAC
A	Y	A	V	A	G	M	S	C	G	H
TGC	AGC	GCG	ACC	CTC	ACC	CGG	GTG	ATC	GGC	GAG
C	S	A	T	L	T	R	V	I	G	E
CTG	GAC	GGC	GTG	ACC	GGC	GTC	GAC	GTC	CAA	CAC
L	D	G	V	T	G	V	D	V	Q	L
GAC	ACC	GGC	CGG	GTC	ACC	GTC	ACC	GCC	GAC	GCG
D	T	G	R	V	T	V	T	A	D	A
GAG	CCG	GAC	GAC	GCC	GCG	ATC	GCC	GAG	GTG	GTC
E	P	D	D	A	A	I	A	E	V	V
GAC	GAA	GCC	GGG	TAC	GAG	CTG	ACC	GGC	CGC	GTC
D	E	A	G	Y	E	L	T	G	R	V
TGA										
*										

Mutagenic primers used to create Y71F and H22G variants of CopZ-1317. The codons changed are underlined.

H22G – His to Gly, CAC to GGC

5′-CATGAGCTGCGGTGGCTGCAGCGCG-3′

5′-CGCGCTGCAGCCACCGCAGCTCATG-3′

Y71F – Tyr to Phe, TAC to TTC

5′-CGAAGCCGGGTTCGAGCTGACCGGC-3′

5′-GCCGGTCAGCTCGAACCCGGCTTCG-3′

Table 6.3: The reagents and volumes used for amplification of H22G and Y71F mutants of CopZ-1317.

Reagent	Concentration	Volume (μl)
Plasmid DNA (*pET1317*)	25 ng/μl	1.0
Forward Primer	75 ng/μl	1.0
Reverese Primer	75 ng/μl	1.0
dNTPs	10 mM	0.6
PFU Buffer	10 x	3.0
PFU Turbo polymerase	-	0.6
Sterile deionized water	-	22.8
Total	-	30.0

Table 6.4: Quikchange site-directed mutagenesis protocol used to create H22G and Y71F mutants of CopZ-1317.

Temperature (ºC)	Time (min)
95	3
95	**0.5**
58	**1**
68	**13**
72	8
10	hold

6.2.4. Over-expression and purification of CopZ-1317 and CopZ-3079

pET1317 and *pET3079* encoding for CopZ-1317 (77 amino acids) and CopZ-3079 (80 amino acids) were transformed into chemically competent *E. coli* BL21 (DE3) cells. Cells containing the plasmids were selected by growing on 50 μg ml^{-1} Kan agar plates and single colonies transferred into 3 ml of LB medium (Melford) containing 50 μg ml^{-1} Kan. These were incubated overnight at 37 °C shaking at 180 rpm. Overnight pre-cultures were successively used to inoculate 750 ml of 2xYT medium in 2 L flasks and incubated at 37 °C, 220 rpm. At an OD_{600} of 0.6 they were inoculated with 1 M isopropyl β-D-1-thiogalactopyranoside (IPTG) solution to give a final concentration of 1 mM and the temperature decreased to 25 ^{o}C. Cells were then harvested after 16 h at 3,501 g and lysed using an EmulsiFlex-C5 cell disrupter (Avestin). Cell lysate was then centrifuged at 38,724 g for 20 min at 4 °C. The clarified supernatant was loaded onto a 5 ml Ni^{2+}-NTA Sepharose column (GE Healthcare) equilibrated with buffer A (50 mM Tris/HCl, 500 mM NaCl, 20 mM imidazole, pH 8) and eluted with a linear imidazole gradient using buffer B (buffer A with 500 mM imidazole). A single peak for both CopZ-1317 and CopZ-3079 at ~ 30 % buffer B was eluted and fractions pooled and dialysed overnight at 4 °C against buffer C (50 mM Tris/HCl, 150 mM NaCl, 2 mM EDTA, 2mM DTT, pH 8). Following dialysis, the N-terminal His_6-tag was removed by incubating the CopZ proteins at room temperature overnight in the presence of 125 U of thrombin (Sigma). The protein/thrombin mixture was then reapplied to a Ni^{2+}-NTA sepharose column and the flow-through collected and concentrated at 4 ^{o}C using a centricon (vivaspin) with a 5 kDa cut-off. Concentrated proteins were loaded to a G75 HiLoad 16/60 Sephadex column (GE Healthcare) equilibrated with buffer C. Major peaks eluted with retention volumes of ~ 78 ml and ~ 82 ml for CopZ-1317 and CopZ-3079 respectively. Samples from these peaks were assessed for purity by 15 % SDS-PAGE gels (see Chapter 2 section 2.2.4) and those deemed of good purity were concentrated and stored at -20 ^{o}C until required.

Electrospray ionisation mass spectrometry (ESI-MS) using a Micromass Quattro Ultima triple quadrupole instrument operating in the positive ion detection mode was used to determine the mass of the purified CopZ proteins. CopZ-1317 and CopZ-3079 were desalted and exchanged into 1 M ammonium acetate followed by a 1:20 dilution with a 50 % methanol and 1 % formic acid solution.

6.2.5. Over-expression and purification of the H22G and Y71F variants

The H22G and Y71F variants were expressed and purified following identical procedures adopted for the wild-type (WT) CopZ-1317 protein (section 6.2.4).

6.2.6. Over-expression and purification of S. lividans CsoR

WT CsoR (SLI4375) was over-expressed in *E. coli* and purified as previously reported [373].

6.2.7. Circular dichroism and UV-visible spectroscopy

All absorbance spectral measurements were carried out using a Varian Cary 50 UV-visible spectrophotometer thermostatted at 20 ^{o}C. Apo-CopZ concentrations were determined at 280 nm using extinction coefficients (ε) of 3105 M^{-1} cm^{-1} for CopZ-1317 and H22G variant and 1615 M^{-1} cm^{-1} for Y71F variant and CopZ-3079 [93]. Apo-CsoR concentrations were also determined at 280 nm using the extinction coefficient (ε) of 3105 M^{-1} cm^{-1} for the CsoR monomer. All extinction coefficients were determined using ProtParam ExPASy and concentrations calculated using the beer lambert law [93]. CopZ samples (20 μM) for circular dichroism (CD) analysis were exchanged into 10 mM KPi, 50 mM KF, pH 7 using a PD-10 desalting column (GE healthcare). Far UV-CD spectra were recorded between 260 and 180 nm at 20 ^{o}C on an Applied Photophysics Chirascan CD spectrophotometer (Leatherhead, UK) equipped with a thermostatic cell holder controlled with a Peltier system. All CopZ CD spectra were analysed using Dichroweb [235, 236] with the programs CDSSTR [237-239] and Contin-LL [240, 241] with the databases 1-7 [239] and SP175 [242]. Good fits and secondary structure estimates were obtained if the NRMSD (normalized root mean square deviation) was < 0.1 and these could then be used to calculate the α-helical and β-sheet content of the corresponding protein [243].

6.2.8. Preparation of reduced proteins and Cu(I) solutions

Proteins were reduced in an anaerobic chamber (DW Scientific [O_2] < 2 ppm) with 5 mM DTT and desalted (twice) using a PD-10 column into the chosen buffer. Free thiol content was ensured by determining the reduction of 5,5'-dithiobis(2-nitrobenzoic acid) (DTNB) monitored at 412 nm (TNB^{2-} ion has a yellow colour and absorbance at 412 nm) (ε = 13,500 $M^{-1}cm^{-1}$) [398]. To determine Cu(I) concentrations a stock of Cu(I) chloride (Sigma) was prepared by weighing out the desired amount and dissolving it in 500 mM NaCl and 10 mM HCl under anaerobic conditions [375, 388]. Dilutions of the stock were made using 10 mM MOPS pH 7.5, 150 mM NaCl. The Cu(I) concentration was then quantified by step-wise addition to a known concentration of the Cu(I) specific bidentate chelator bicinchoninic acid (BCA) and monitoring the absorbance spectrophotometrically. A linear increase in absorbance at 562 nm was observed for the $[Cu(I)(BCA)_2]^{3-}$ complex until no further change was observed, allowing the concentration to be calculated from the absorbance at 562 nm using an extinction coefficient of ε = 7,900 M^{-1} cm^{-1} (for $[Cu(I)(BCA)_2]^{3-}$)[399].

6.2.9. Cu(I) titrations

Samples of CopZ-1317 and CopZ-3079 and all Cu(I) solutions were prepared anaerobically (section 6.2.8). Cu(I) titrations were carried out in a sealed anaerobic quartz cuvette with micro-litre aliquots of Cu(I) solution titrated using a gas-tight syringe (Hamilton). Absorbance changes were monitored using a Varian Cary 50 UV-visible spectrophotometer thermostatted at 20 ^{o}C in 10 mM MOPS pH 7.5, 150 mM NaCl.

6.2.10. Competition assays using BCA and BCS

Competition assays were set up anaerobically with either BCA or bathocuprione disulfonate (BCS) (Sigma). Increasing protein concentrations (0-80 μM) were added to solutions of $[Cu(I)L_2]^{3-}$ (L = ligand) of defined molar ratio L:Cu(I) ≥ 3 creating a series of individual solutions with constant [Cu(I)] and [L] and varying [protein]. Samples were left for ~ 1 h and the transfer of Cu(I) from the $[Cu(I)L_2]^{3-}$ complex to the protein was determined by measuring the absorbance of the $[Cu(I)L_2]^{3-}$ complex spectrophotometrically, for L = BCA at 562 nm (ε = 7,900 M^{-1} cm^{-1}) [399] and L = BCS at 483 nm (ε = 13,000 M^{-1} cm^{-1}) [400].

6.2.11. Calculation of the Cu(I) dissociation constant (K_d)

By interchanging [L], assays favouring competitive or non-competitive Cu(I) binding could be set-up. Non-competitive Cu(I) binding allowed an estimate of the binding stoichiometry and competitive Cu(I) binding allowed the dissociation constant for Cu(I) (K_dCu(I)) to be determined. The K_d can be defined as the dissociation constant of a metal from a protein as shown in Equation 6.1 whereby [M] is the metal concentration and [P] is protein concentration. The affinity of an apo-protein for a metal-ion can be determined by studying the exchange of a metal-ion between a chelator and a protein. Equation 6.2 gives the general reaction where, MP, L, ML_2 and P are the metal-protein complex, the chelator used, the metal-chelator complex and the apo-protein, respectively, whereby two molecules of chelator (L) bind to a single metal ion (M). Equation 6.3 shows this reaction specific for the CopZ proteins with BCA/BCS chelators.

$$K_d = \frac{[M][P]}{[MP]}\ (M) \qquad \text{Equation 6.1}$$

$$MP + 2L \leftrightarrow ML_2 + P \qquad \text{Equation 6.2}$$

$$\text{apo-CopZ} + BCA_2/BCS_2Cu(I) \leftrightarrow Cu(I)\text{-CopZ} + BCA_2/BCS_2 \qquad \text{Equation 6.3}$$

With a binding stoichiometry of 1:2 (Cu(I):chelator) as it is for the chelators BCA and BCS, as shown in Equation 6.2 and 6.3, a K_dCu(I) can be calculated using Equation 6.4.

$$K_{eq}\ \left(M^{-1}\right) = K_d\ \beta'_2 = \frac{([\text{apo-P}]_{tot} / [MP])-1}{\{([L]_{tot}/[ML_2])-2\}^2[ML_2]}$$ Equation 6.4

Whereby, [L] is the ligand concentration (BCA or BCS), $[L]_{tot}$ the total ligand concentration, [P] is the protein concentration and $[P]_{tot}$ is the total protein concentration, [MP] is the metal protein complex and [ML] the metal ligand complex. From carrying out a competition assay the only terms that are known in Equation 6.4 are the concentrations of P_{tot} and L_{tot}. The concentration of ML_2 in all the samples is determined from the measured absorbance values (of Cu loaded BCA and/or BCS at 562 nm and 483 nm, respectively) and the total concentration of metal *i.e.* M_{tot} is obtained from the absorbance of the sample to which no protein is added. With the known concentration of M_{tot} and ML_2 the concentration of MP can be calculated using Equation 6.5.

$$[M]_{tot} = [ML_2] + [MP]$$ Equation 6.5

Thus, values of P_{tot}, L_{tot} and MP can be substituted into Equation 6.4 to calculate K_{eq} for each sample in the assay. The K_{eq} can then be divided by the overall formation constant (β_2'), $10^{17.2}$ M^{-2} for $[Cu(I)(BCA)_2]^{3-}$ and $10^{19.8}$ M^{-2} for $[Cu(I)(BCS)_2]^{3-}$, to calculate the K_d [399, 401], where K_d and β_2 are defined as the dissociation constant for the protein and the formation constant for the chelator used. The β_2 value is the product of two different binding affinities K_1 and K_2. K_1 corresponds to the binding of the first chelator molecule to the metal-ion and K_2 corresponds to the binding of the second chelator molecule. Assays were performed in triplicate and the K_dCu(I) value for a series was initially calculated for each individual solution, then averaged. Standard deviation was used to calculate the error. By using the average K_dCu(I) value, a simulated curve was plotted using Equation 6.4. Simulation calculations are set out in Appendix 4.1.

Conditions for the plots shown in Figure 6.9 for CopZ-1317 with BCA were: 0 - 80 μM [P], 20 μM [Cu(I)] and 80 μM [BCA]; with BCS, 0 – 50 μM [P], 31 μM [Cu(I)], and 80 μM [BCS]. Conditions for CopZ-3079 with BCA were; 0 - 60 μM [P], 21 μM [Cu(I)] and 80 μM [BCS]; with BCS 0-60 μM [P], 32 μM [Cu(I)] and 120 μM [BCS]. Conditions for the plots shown in Figure 6.10 for H22G were: 0 – 80 μM [P], 23 μM [Cu(I)] and 80 μM BCA: with BCS, 0 – 80 μM [P], 25 μM [Cu(I)] and 80 μM BCS. Conditions for Y71F with BCA were: 0 – 80 μM [P], 22 μM [Cu(I)] and 80 μM BCA: with BCS, 0 – 80 μM [P], 19 μM [Cu(I)] and 80 μM BCS.

6.2.12. Analytical gel filtration

A 10/300 GL G-75 Superdex column (GE-Healthcare) was equilibrated in degassed buffer (10 mM MOPS pH 7.5, 150 mM NaCl) with or without 2 mM DTT. Speciation of CopZ proteins (1317/3079) were analysed by preparing apo, 0.5, 1.0, 1.5 and 2 molar equivalents of Cu(I) to pre-reduced CopZ anaerobically (section 6.2.8). Samples were then injected onto a pre-equilibrated column with or without the presence of DTT and the elution profile monitored by the absorbance at 280 nm. For Cu(I) transfer experiments reduced apo-CopZ-3079 and apo-CsoR were loaded as required with 10-fold excess Cu(I), incubated, then any excess Cu(I) removed using a PD-10 desalting column (10 mM MOPS, pH 7.5, 150 mM NaCl). Protein samples were then mixed (Apo-CopZ-3079 and Cu(I)-CsoR or Cu(I)-CopZ-3079 and apo-CsoR) in a 1:1 stoichiometry, incubated for ~ 1 h then injected onto a pre-equilibrated column (10 mM MOPS pH 7.5, 150 mM NaCl) and the absorbance at 280 nm monitored.

6.2.13. Determination of Cys pK_a values by absorbance at 240 nm

Determination of Cys pK_a values were carried out by monitoring the absorbance at 240 nm using two methods with both using reduced apo proteins in 5 mM MOPS pH 7, 25 mM KCl prepared in the anaerobic chamber (section 6.2.8). The first method used for the CopZ proteins, used a series of mixed buffer systems containing 10 mM each of potassium acetate, MES, MOPS, Tris/HCl and 200 mM KCl with pH values of 4 to 10 with increments of 0.5. Reduced apo-proteins (final concentration of ~ 40 μM) were added to each pH buffer and incubated for ~ 2 h. The absorbance at 240 nm was then measured spectrophotometrically in a sealed quartz cuvette (Hellma) and plotted as a function of pH. The second method (used for CsoR) was to add ~ 40 μM of pre-reduced protein into a low pH mixed buffer and gradually add NaOH while monitoring the pH and measuring the absorbance change at 240 nm spectrophotometrically. The $\varepsilon_{240\ nm}$ was then calculated and plotted by rearranging the beer lambert law, with the absorbance taken at 240 nm and the concentration of protein calculated using the extinction coefficient and absorbance at 280 nm. Models describing one (Equation 6.6) or two (Equation 6.7) non-interacting macroscopic ionisations were used to determine pK_a values of the Cys residues, where ε_o is the extinction of the thiol form and $\Delta\varepsilon$ is the difference between the extinction coefficient of a thiol and thiolate form. The pK_a values reported are an average of multiple data sets and the error reported is the standard deviation.

$$\varepsilon_{240} = \varepsilon_o + \frac{\Delta\varepsilon\ 10^{\text{pH-pKa}}}{1 + 10^{\text{pH-pKa}}} \qquad \text{Equation 6.6}$$

$$\varepsilon_{240} = \varepsilon_o + \frac{\Delta\varepsilon_1 10^{pH-pK_{a1}} + \Delta\varepsilon_2 10^{2pH-pK_{a1}-pK_{a2}}}{1 + 10^{pH-pK_{a1}} + 10^{2pH-pK_{a1}-pK_{a2}}} \qquad \text{Equation 6.7}$$

6.2.14. Determination of Cys pK_a values using Badan

Reduced apo-CopZ proteins were prepared in an anaerobic chamber in 5 mM MOPS pH 7.5, 25 mM KCl, together with a series of mixed buffer systems (*vide supra*) with the pH of each solution individually adjusted in increments of 0.5 from pH values of 3 to 10. A 10 mM stock solution of Badan (6-bromoacetyl-2-dimethylaminonaphthalene, Molecular Probes) was prepared in DMSO. Under anaerobic conditions reduced apo-proteins were added to a final concentration of ~ 1 μM to each buffered solution followed by addition of Badan to a final concentration of 13 μM in a 3 ml quartz cuvette (maintaining pseudo-first order conditions). The change in fluorescence emission of Badan at 540 nm (excitation wavelength 391 nm) was monitored every 10 s at 20 ^oC for 25 min on a Perkin-Elmer LS55 fluorescence spectrophotometer. Time courses measured at the different pH values were fitted to either a single or double exponential function to obtain an initial pseudo-first order rate constant (k_0), which was plotted as a function of pH. pK_a values were obtained by fitting the data to Equation 6.8, describing a single protonation/deprotonation process, or Equation 6.9 describing two ionization events, where k_0 is the initial pseudo-first order rate constant and k_{SH}, k_{S^-} are the rate constants for the various protonated and deprotonated forms of the protein. The pK_a values reported are an average of multiple data sets and the error is the standard deviation.

$$k_0 = k_{SH} + \frac{k_{S^-} 10^{pH-pK_a}}{1+10^{pH-pK_a}} \qquad \text{Equation 6.8}$$

$$k_0 = \frac{k_{SHSH} + k_{S^-SH} 10^{pH-pK_{a1}} + k_{S^-S^-} 10^{2pH-pK_{a1}-pK_{a2}}}{1+10^{pH-pK_{a1}} + 10^{2pH-pK_{a1}-pK_{a2}}} \qquad \text{Equation 6.9}$$

6.2.15. Electrophoretic mobility shift assays (EMSA)

DNA oligomers (Sigma) were prepared in 10 mM HEPES pH 7.5, 150 mM NaCl. Concentrations of individual oligonucleotides were determined using appropriate extinction coefficients at 260 nm on a Nanodrop 2000 (Thermo Scientific). Complementary pairs of DNA oligomers were annealed by mixing equal concentrations and heating at 95 ^oC in a water bath for 5 min and then left to cool over-night at room temperature. Reduced proteins were prepared in an anaerobic chamber in 10 mM HEPES, pH 7.5, 150 mM NaCl. For Cu(I) loaded samples, CuCl solution was added to the respective proteins and allowed to incubate

for 30 min followed by removal of any excess Cu(I) by application to a PD-10 desalting column (GE-Healthcare). 0.5 μM of a DNA duplex was incubated with the desired concentration of protein (4 - 6 μM CsoR monomer (equivalent to 0.5 μM 2 x CsoR tetramer) and 4 - 40 μM CopZ (equivalent to a 1:1 and a 1:10 CsoR-monomer:CopZ)) in 10 mM HEPES pH 7.5, 150 mM NaCl, 1 mM DTT. 1 or 10 molar equivalent of Cu(I)-loaded CopZ proteins were incubated with apo-CsoR or, 1 or 10 molar equivalent of apo-CopZ (3079/1317) proteins were mixed with Cu(I)-loaded CsoR. All samples were incubated at room temperature under anaerobic conditions for ~ 1 h and then loaded (20 μl) to a pre-run 6 % Tris-Borate EDTA (TBE) polyacrylamide gel (Reagents for a 6 % TBE gel in Table 6.5) (120 V for 40 min in 1 x TBE buffer). Gels were stained for 30 min in an ethidium bromide solution followed by imaging.

Table 6.5: The components required to construct two 6 % TBE gels for EMSA.

Chemical	Volume
ddH_2O	7.2 ml
5 x TBE	2.4 ml
30 % Acrylamide/Bisacrylamide 29:1	2.4 ml
10 % APS (Ammonium persulphate)	200 μl
TEMED (Tetramethylethylenediamine)	10 μl
80 % Glycerol	240 μl
Total	12.45 ml

6.3. Results

6.3.1. Transcription analysis of CopZ genes in S. lividans

A model for the regulon under the control of CsoR in *S. lividans* and its response to Cu stress has previously been reported [373]. This model was constructed by RNA sequencing and promoter probing experiments using the *S. coelicolor* genome as a reference [373]. However, since then the genome sequence of *S. lividans* has subsequently become available leading to the identification of a larger cytosolic Cu(I) handling capacity than was previously foreseen [4]. From transcription re-analysis of the *csoR* and the *copZA* operons using the *S. lividans* genome sequence as input, four *copZ* and five P-type ATPase (*copA*) encoding genes have been identified (Table 6.6). The four *copZ* genes, 0895, 1063, 1317 and 3079 are immediately upstream of four out of the five P-type ATPases. The five P-type ATPases, 0845, 0896, 1064, 1318 and 3080 are all part of the IB1_Cu class with only 0845 not part of an operon containing a *copZ* gene. A CsoR operator binding site has been identified (using PREDetector [396]) upstream of all *copZ* genes with *SLI1317* containing two CsoR binding sites. Analysis of the transcription before and after response to Cu reveals that *copZ*-0895, 1063 and 3079 with their cognate P-type ATPases 0896, 1064 and 3080 contribute to the majority of the transcription in response to Cu (Table 6.6). *SLI1317*/*SLI1318* only contributes a maximum of 7 % of the total transcription and *SLI0895*; the single P-type ATPase contributes even less (Table 6.6). The *csoR* gene shows high levels of transcription, with low induction upon addition of Cu as previously reported [373]. When the *csoR* gene is deleted (*ΔcsoR*) confirmed by zero CsoR transcription, only three *copZ/copA* couples are induced even though all four contain a CsoR operator-binding site (Table 6.6). Therefore, the 0895/0896 operon appears not to be within the CsoR (*SLI4375)* regulon.

Table 6.6: Transcription analysis of the *csoR* and the *copZA* operons in *S. lividans*. RNA-seq data obtained and reported previously [373] was analysed with the genome of *S. lividans* 1326 as input [4]. CsoR binding sites were determined using PREDetector [396].

Name	Gene	CsoR binding site	[A] *S. lividans* 1326 (RPKM)[b]	[a] +Cu (RPKM)	[a] *ΔcsoR* (RPKM)	[c] Fold increase + Cu	[c] Fold increase + *ΔcsoR*
P-type ATPase-IB1-Cu	*SLI0845*	No	2	0	2	0	1
CopZ	*SLI0895*		39	253	33	6	1
P-type ATPase-IB1-Cu	*SLI0896*	Yes	7	24	7	3	1
CopZ	*SLI1063*		33	128	67	4	2
P-type ATPase-IB1-Cu	*SLI1064*	Yes	18	65	40	4	2
CopZ	*SLI1317*		0	9	5	>>	>>
P-type ATPase-IB1-Cu	*SLI1318*	Yes, two	3	3	3	1	1
CopZ	*SLI3079*		16	142	80	9	5
P-type ATPase-IB1-Cu	*SLI3080*	Yes	29	54	23	2	1
CsoR	*SLI4375*	Yes	170	212	0	1	0

[a] The transcriptomes of WT *S. lividans* 1326 and the *ΔcsoR* mutant strain were obtained without addition of Cu to the growth medium. The Cu induced transcriptome was obtained after a 2 h exposure to 400 μM Cu(II).

[b] The measure of expression is RPKM defined as the reads/kb of exon/million mapped reads *i.e*. dividing the total number of exon reads (in this case one exon per reference sequence) by the number of mapped reads (in millions) times the exon length (in kb) [397].

[c] Compared to the WT strain in the absence of Cu

6.3.2. CopZ-1317 and CopZ-3079 amino acid sequences

To understand whether there is an interplay between the CopZ Cu(I) chaperones and CsoR in *S. lividans* two of the three CopZ proteins under the control of CsoR were examined further. These two proteins differ significantly in their transcription expression levels with *SLI3079* being a major player in Cu resistance, whereas *SLI1317* displays only minimal expression levels (Table 6.6). The genomic environment of these two CopZ proteins and CsoR (*SLI4375*) are shown in Figure 6.2 with the identification of the CsoR operator binding sites, predicted using PREDetector and confirmed by Dwarakanath *et al* [373, 396]. The *SLI1317* and *SLI3079* genes encoding CopZ-1317 and CopZ-3079 were aligned with the two other CopZ sequences from *S. lividans* (*SLI1063* and *SLI0895*) and with CopZ proteins from *E. hirae* and *B. subtilis*, and the *H. sapiens* HAH1 and *S. cerevisiae* Atx1 proteins using Clustal Omega (Figure 6.3) [108]. The CopZ proteins within *S. lividans* share ~ 46 - 66 % identity and ~ 20 % identity to other homologous proteins (Figure 6.3). *SLI0895* and *SLI1063* contain Cys rich N-terminal extensions, containing CCXXXXXC and CCXXC motifs, respectively. The secondary structure of CopZ-1317 and CopZ-3079 was predicted using Jpred3 and is shown in Figure 6.3 above the sequence alignment displaying a βαββαβ-fold common to Atx1 family members. The secondary structure content could be calculated from this prediction and identified as containing 27 % for both α-helix and β-sheet content. These metallochaperones bind Cu(I) through thiolate ligands originating from the $MX_1CX_2X_3C$ motif, conserved in all sequences (X any amino acid) (Figure 6.3). There are however differences at the X positions between eukaryotic and prokaryotic species, and an apparent variation at the X_2 position between the prokaryotic proteins (Figure 6.3). The alignment illustrates the presence of a His residue at the X_3 position for both CopZ-1317 and CopZ-3079 and a Tyr residue at position 71 in CopZ-1317 compared to a Phe in CopZ-3079 (residues identified with a * in Figure 6.3). These residues are of interest due to their influence on the acid/base properties of the coordinating Cys residues and will be explored in this work.

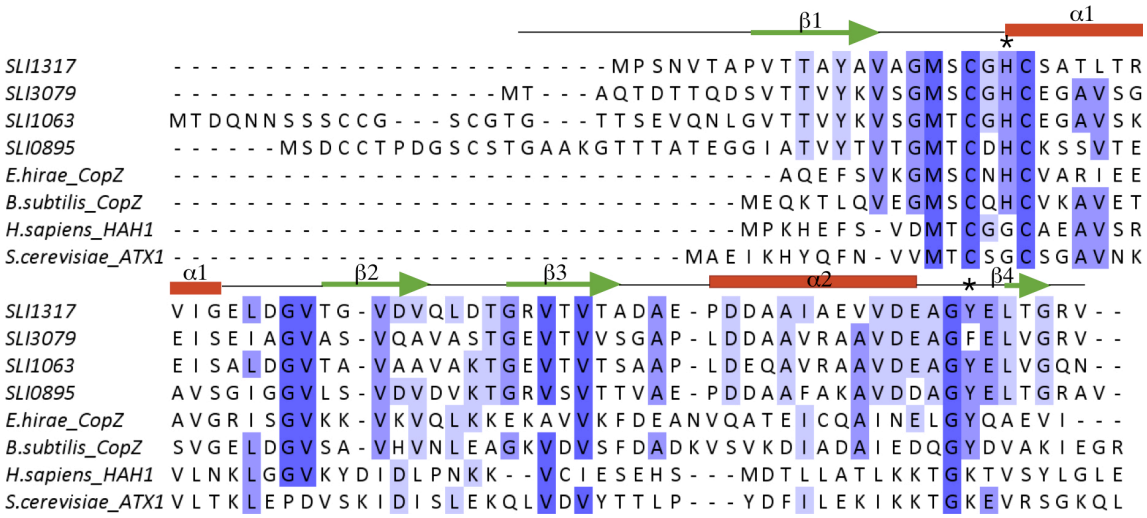


Figure 6.3: Sequence alignment of CopZ-1317 and CopZ-3079. Clustal Omega sequence alignment of the four *S. lividans* CopZ proteins (*SLI*), *E. hirae* and *B. subtilis* CopZ, and the *H. sapiens* HAH1 and *S. cerevisiae* Atx1 proteins. Complete and partially conserved residues are coloured in dark to light blue respectively. The secondary structure prediction for CopZ-1317 and CopZ-3079 from Jpred3 [245] is indicated above the sequence alignment, with α-helices shown in red and β-sheets in green. The * symbol indicates the position of His^{22} and Tyr^{71} in the CopZ-1317 numbering.

6.3.3. Cloning, over-expression and purification of CopZ-3079, CopZ-1317, and the H22G and Y71F variants

Genes encoding CopZ-1317 and CopZ-3079 were successfully amplified from the genomic DNA of *S. lividans* (Figure 6.4.A), and subsequently sub-cloned into the pET28a *E. coli* expression vector containing an N-terminal His_6-tag (Figure 6.4.B). CopZ-1317 variants, H22G and Y71F were also successfully created using a Quikchange site-directed mutagenesis approach (Figure 6.4.C).

CopZ-1317 and CopZ-3079 contain 77 and 80 amino acids respectively, and were expressed in *E. coli* using 2xYT media following induction with IPTG. CopZ proteins were purified first using a Ni^{2+}-NTA sepharose column with both proteins producing single broad peaks at 280 nm ~ 40 ml elution volume at ~ 30 % buffer B (Figure 6.5.A and 6.5.B). Following His_6-tag removal CopZ proteins were purified further by application to a G75 size exclusion column. Clear single peaks were observed for both CopZ-1317 and CopZ-3079 eluting at ~ 78 ml and ~ 82 ml respectively, consistent with monomeric forms of the proteins based on the calibration profile of the column (Figure 6.5.C). Fractions eluting from the major peaks were analysed by 15 % SDS-PAGE gel with both proteins displaying clear single bands running between 10 - 15 kDa, somewhat higher than the predicted masses (8,063 Da for CopZ-1317 and 8,200 Da for CopZ-3079). H22G and Y71F were over-expressed and purified in a similar manner with Ni^{2+}-NTA and G75 profiles and SDS-PAGE results essentially identical to the WT CopZ-1317 protein (Appendix 4.2).

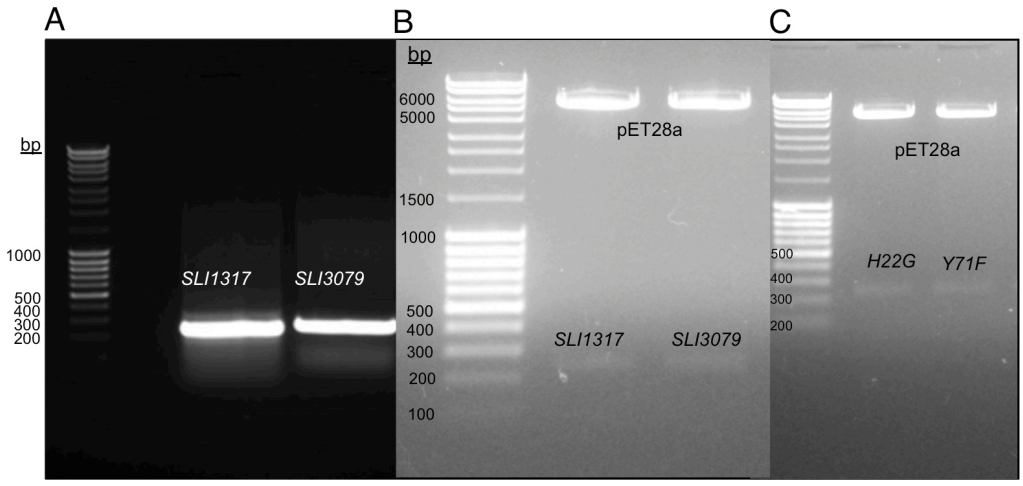


Figure 6.4: Cloning of CopZ-1317, CopZ-3079 and CopZ-1317 mutants H22G and Y71F. A) DNA agarose gel (1 % w/v) of the PCR products of *SLI1317* and *SLI3079*. The correct bands (234 bp *SLI1317* and 243 bp *SLI3079*) were corroborated with the DNA ladder (Fermentas DNA mass ruler). B) DNA agarose (1 % w/v) gel of the *pET1317* and *pET3079* constructs digested using NdeI and HindIII restriction enzymes and corroborated against the DNA ladder (far left lane) (pET28a – 5,369 bp, *SLI1317* - 234 bp and *SLI3079* - 243 bp). C) H22G and Y71F pET28a CopZ-1317 mutant constructs digested using NdeI and HindIII restriction enzymes (pET28a – 5,369 bp and *SLI1317* mutants 234 bp).

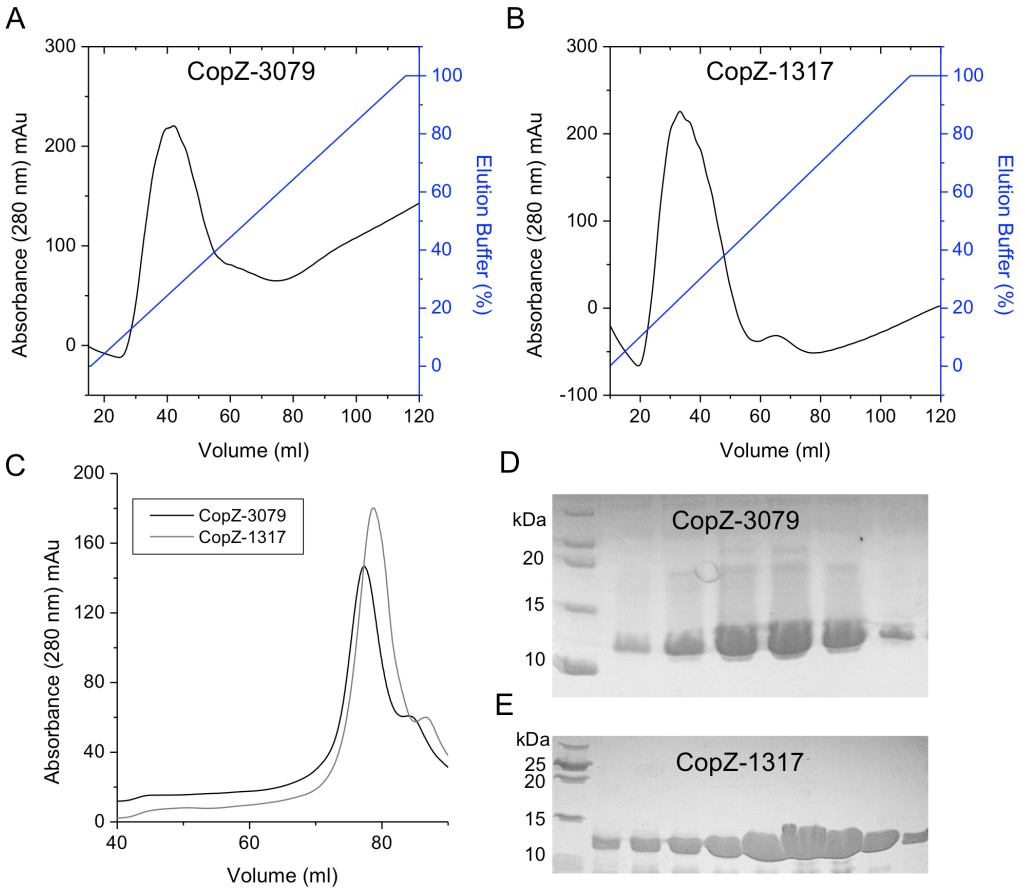


Figure 6.5: Purification of CopZ-3079 and CopZ-1317. A) Ni^{2+}-NTA profile of CopZ-3079 showing an absorbance peak (mAu) at 280 nm in black at ~ 30 % buffer B shown in blue. B) Ni^{2+}-NTA profile of CopZ-1317 showing an absorbance peak (mAu) at 280 nm in black at ~ 30 % buffer B shown in blue. C) G75 size-exclusion Sephadex column absorbance 280 nm (mAu) profile of CopZ-1317 and CopZ-3079 eluting at ~ 78 ml and ~ 82 ml respectively, consistent with monomeric forms of the proteins in 50 mM Tris/HCl, 150 mM NaCl, 2 mM EDTA, 2 mM DTT, pH 8. A shoulder was also identified for both CopZs but assigned to impurities. D) Coomassie stained 15 % SDS-PAGE gel analysis of purified CopZ-3079 displaying a single protein band between 10-15 kDa (8,200 Da CopZ-3079). E) Coomassie stained 15 % SDS-PAGE gel analysis of purified CopZ-1317 displaying a single protein band between 10-15 kDa (8,063 Da CopZ-1317).

6.3.4. Preliminary characterization of CopZ proteins

Analytical gel filtration chromatography gave elution volumes of 12.3 ml and 12.5 ml for CopZ-1317 and CopZ-3079 respectively (Figure 6.5.A). These retention volumes are consistent with a mass of ~ 10 kDa from a column calibration graph, further confirming that these CopZ proteins are monomeric in solution (also shown by G75 HiLoad 16/60 Sephadex column Figure 6.5). Purified CopZ proteins gave masses determined from denaturing mass spectrometry of 8,062 Da for CopZ-1317 (expected 8,063 Da) and 8,199 Da for CopZ-3079 (expected 8,200 Da) consistent with their expected mass. UV-visible spectroscopy reveals that both CopZs display no spectral transitions in the visible region, however do contain peaks at ~ 230 nm (peptide backbone) and at ~ 280 nm (aromatic residues) (Figure 6.6.B). Extinction coefficients for the two CopZ proteins at 280 nm differ due to variations in aromatic residues, with CopZ-1317 displaying maxima at 277 nm due to two Tyr residues and CopZ-3079 displaying a peak maxima at 279 nm due to one Tyr and one Phe residue (Figure 6.6.B). Far UV-CD spectra of both apo-CopZ proteins after His_6-tag removal confirmed that both proteins are clearly folded (Figure 6.6.C). Dichroweb was used to analyse the CD spectrum for each CopZ protein with significant fits (NRMSD < 0.1) selected and averaged to calculate the secondary structural content [235, 236]. This gave percentages of secondary structure outlined in Table 6.7 with 20 % α-helix and 22 % β-sheet for CopZ-1317 and 23 % α-helix and 18 % β-sheet for CopZ-3079 which is in good agreement with that predicted from Jpred3 (Table 6.7) (27 % for both α-helix and β-sheet) [245]. This provides strong evidence in the absence of tertiary structural information that these *S. lividans* CopZ proteins belong to the Atx1 family with βαββαβ-folds. Examples of fits from Dichroweb analysis are given in Appendix 4.3.

H22G and Y71F CopZ-1317 variants also display retention volumes from analytical gel filtration consistent with monomeric species (Appendix 4.4.A). The UV-visible spectra of both CopZ-1317 variants is similar to that observed for WT protein (Appendix 4.4.B). Importantly, far UV-CD spectra of H22G and Y71F are comparable to WT CopZ-1317 illustrating that these mutants do not affect folding of the protein (Appendix 4.4.C). Dichroweb analysis predicts 10 % α-helix and 23 % β-sheet for H22G and 13 % α-helix and 25 % β-sheet for Y71F, which although displays a decrease in the α-helical percentage compared to WT still predicts βαββαβ-folds for these mutants (Table 6.7).

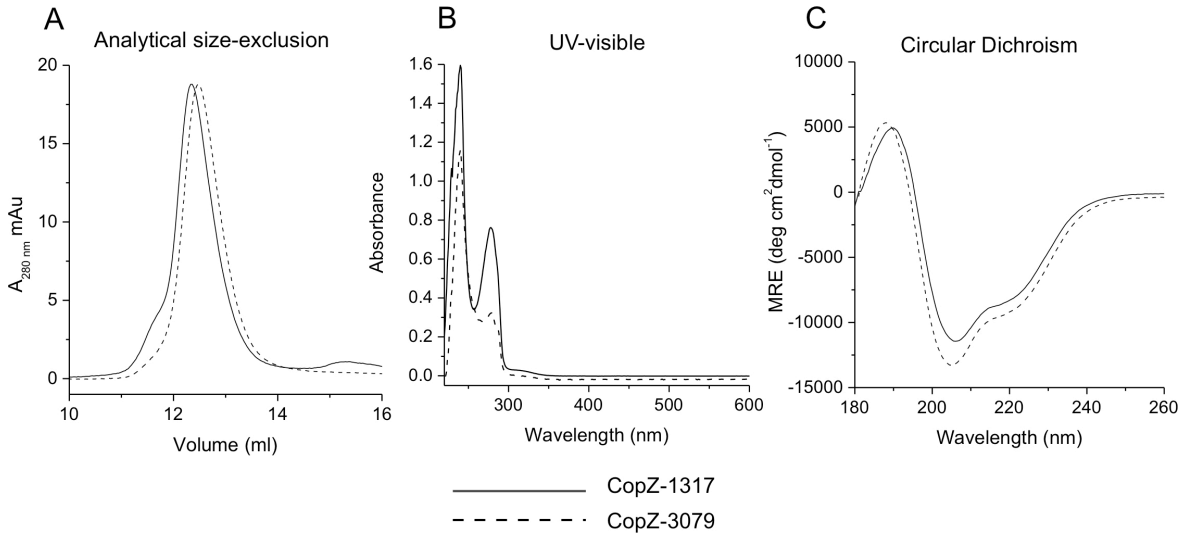


Figure 6.6: Preliminary characterization of CopZ-1317 and CopZ-3079. A) Analytical gel filtration chromatography profile for apo-CopZ-1317 and apo-CopZ-3079 in 10 mM MOPS, pH 7.5, 150 mM NaCl, 2 mM DTT. B) UV-visible spectroscopy of CopZ-1317 and CopZ-3079 in 50 mM Tris/HCl, 150 mM NaCl, 2 mM EDTA, 2 mM DTT, pH 8, 20 ^{o}C. C) Far UV-CD spectra of CopZ-1317 and CopZ-3079 at 20 °C, 10 mM KPi, 50 mM KF, pH 7 with protein concentrations of 20 μM. CopZ-1317 - solid line and CopZ-3079 - dashed line.

Table 6.7: Dichroweb analysis of CopZ-1317, CopZ-3079, H22G and Y71F CopZ-1317 variants [235]. Dichroweb was used to analyse the CD spectrum of each protein with significant fits (NRMSD < 0.1) selected and averaged to calculate the secondary structural content.

Name	α-helix (%)	β-sheet (%)	Turns (%)	Unordered (%)
CopZ-1317	20	22	20	38
CopZ-3079	23	18	20	39
CopZ-1317-H22G	10	23	20	47
CopZ-1317-Y71F	13	25	20	42

6.3.5. Multiphasic Cu(I) binding to CopZ-1317 and CopZ-3079

Unlike 0895 and 1063, CopZ-1317 and CopZ-3079 have no other Cys residues other than those in the $MX_1CX_2X_3C$ motif. Both proteins readily reduce DTNB to consistently give an average protein:thiol ratio of 1:2, indicative of a high level of Cys solvent exposure, ideal for capturing and binding Cu(I). The UV-visible absorbance spectra of both CopZ proteins display strong absorbance maxima ~ 277-279 nm due to Tyr aromatic residues but are otherwise featureless. When Cu(I) is titrated into the CopZ proteins specific spectral transitions can be observed (Figure 6.7.A). Changes in the absorbance spectrum can be detected between 240-260 nm arising from ligand to metal charge transfer from the formation of Cu(I)-S(Cys) bond and between 260-335 nm from cluster-localised Cu(I) (d-s) transitions [402]. These changes were observed in both CopZs and were plotted as a function of Cu(I)/protein displayed in Figure 6.7.B and C both illustrating multiphasic Cu(I) binding. CopZ-1317 (Figure 6.7.B) displays a steep absorbance increase at 240 nm between 0-1.0 Cu(I)/protein followed by a second increase at 1.5-2.0, differing to the absorbance at 335 nm which shows multiple multiphasic absorbance changes. In contrast, CopZ-3079 (Figure 6.7.C) shows an absorbance increase at both 240 nm and 300 nm between Cu(I)/protein of 0-0.5 and 0.5-1.0.

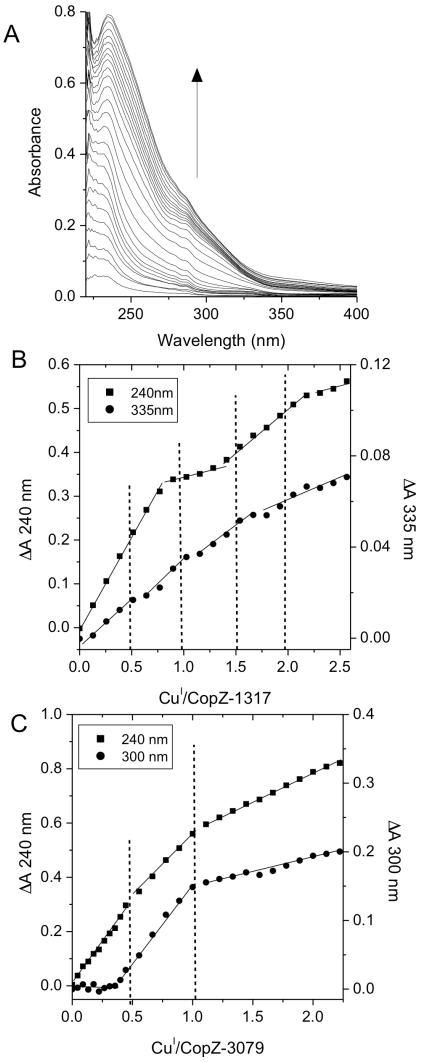


Figure 6.7: Cu(I) binding to CopZ-1317 and CopZ-3079. Cu(I) binding was monitored using UV-visible absorbance changes at 20 ^{o}C, pH 7.5. A) The absorbance spectrum of apo-CopZ-3079 (45 μM) upon titration of 5 μM Cu(I) aliquots under anaerobic conditions, baselined with apo-CopZ-3079. B) The absorbance changes observed for CopZ-1317 at 240 nm and 335 nm plotted as a function of [Cu(I)/CopZ-1317] (CopZ-1317, 39 μM). C) The absorbance changes observed for CopZ-3079 at 240 nm and 300 nm from (A), plotted as a function of [Cu(I)-CopZ-3079].

6.3.6. CopZ-1317 and CopZ-3079 speciation

It has been reported that CopZ proteins can form higher order assemblies upon addition of Cu(I), whereby each CopZ monomer binds Cu(I) with higher order assembly forming through sharing of Cys ligands to form metal clusters [387-389]. The titration results presented in Figures 6.7.B and C hint that higher order assemblies for this *S. lividans* CopZ pair is a possibility. A good way to study the formation of higher order species is to monitor elution profiles using an analytical size exclusion column upon addition of Cu(I) to the protein. CopZ-1317 and CopZ-3079 were prepared in apo-form, and with 0.5, 1.0, 1.5 and 2 molar equivalents of Cu(I) added in the presence or absence of DTT. Apo-CopZ proteins display major peaks eluting at ~ 12.3 ml and ~ 12.5 ml for CopZ-1317 and CopZ-3079, respectively, consistent with monomeric species. In the presence of 2 mM DTT these monomeric species remained, even with the addition of 2 molar equivalents of Cu(I) (Figure 6.8.A). However, in the absence of DTT additions of Cu(I) to the CopZ proteins clearly display the formation of higher order species. When 1 Cu(I) equivalent is added to CopZ-3079 the major species has an elution volume of ~ 10.9 ml with a small amount of monomeric protein at ~ 12.5 ml. Once 1.5 and 2 molar equivalents of Cu(I) are added, the species at ~ 12.5 ml disappears and the one at ~ 10.9 ml increases in intensity (Figure 6.8.B). Based on a calibration curve of the column this elution volume corresponds to a molecular weight of ~ 20 kDa, therefore correlating to a dimer species of CopZ-3079 (~ 16.5 kDa). In comparison, CopZ-1317 remains predominantly as a monomer at 1 molar equivalent of Cu(I) at ~ 12.3 ml, however, after the addition of 1.5 molar equivalents of Cu(I) two species of equal intensities are observed with a higher order assembly at an elution volume of ~ 9.4 ml (Figure 6.8.C). At 2 molar equivalents of Cu(I) the higher order species becomes the major form corresponding to a molecular weight based on the column calibration curve of ~ 43 kDa suggesting CopZ-1317 is a tetramer assembly upon addition of 2 molar equivalents of Cu(I).

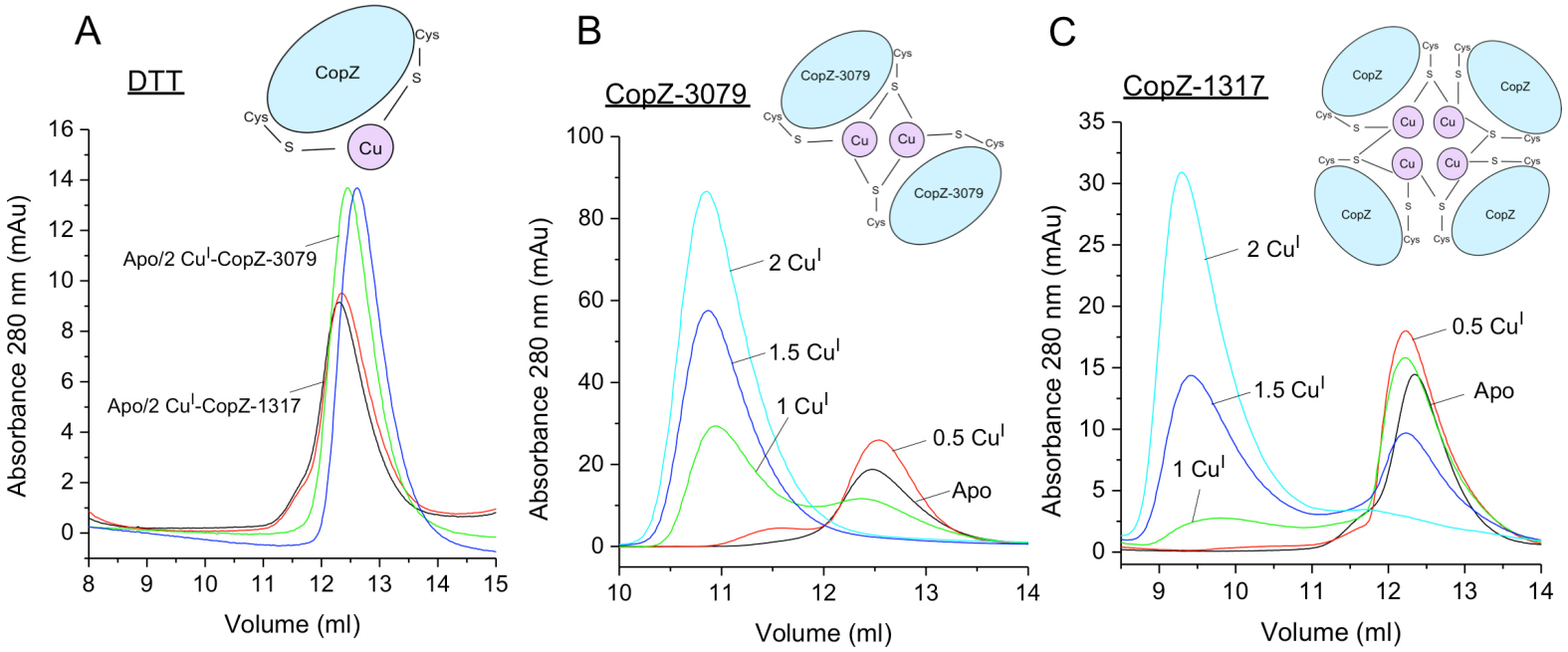


Figure 6.8: Speciation of CopZ-3079 and CopZ-1317 monitored by analytical gel filtration chromatography. A) Apo-CopZ-3079, apo-CopZ-1317 and 2 molar equivalents of Cu(I) added to both CopZ proteins prepared in 10 mM MOPS, 150 mM NaCl, pH 7 buffer containing 2 mM DTT (see experimental section 6.2.12). B) The elution profile of apo-CopZ-3079 and once loaded with, 0.5, 1, 1.5 and 2 molar equivalents of Cu(I). C) The elution profile of apo-CopZ-1317 and once loaded with, 0.5, 1, 1.5 and 2 molar equivalents of Cu(I). B and C were run in 10 mM MOPS, 150 mM NaCl, pH 7.5 buffer without any DTT present (see experimental 6.2.12). CopZ samples were prepared anaerobically ~ 100-200 μM. Insets, cartoon style diagrams of possible assemblies formed by the CopZ proteins under the conditions imposed, the Cu ion is shown in purple and the CopZ proteins in blue.

6.3.7. BCA and BCS competition assays

To assess the Cu(I) binding affinities of the CopZ proteins the Cu(I) bidentate chelators BCA and BCS were used in competition assays (Figure 6.9). BCA and BCS are small ligands able to bind to Cu(I) in a 2:1 ratio (BCS_2/BCA_2), giving visible absorbance peaks at 562 nm and 483 nm, respectively. This absorbance decreases once Cu(I) is removed from the chelator by a competing protein. A steeper decrease in absorbance is related to a protein with a higher affinity for Cu(I). BCS has a stronger affinity for Cu(I) compared to BCA and the use of both enables the buffering of Cu(I) over the range of 10^{-12} – 10^{-19} M [403] to be studied.

Varying the concentrations added of apo-CopZ-1317 and CopZ-3079 into $[Cu(I)(BCA)_2]^{3-}$ resulted in a linear decrease in absorbance at 562 nm until a break point was reached at a ratio of ~ 1 CopZ/Cu(I) (Figure 6.10 A. and B). This indicates that BCA cannot compete for Cu(I) with either CopZ protein and is consistent with 1 equivalent of Cu(I) bound per monomer of CopZ. Therefore, to determined the Cu(I) affinity BCS was used and competition with the apo-CopZ proteins was observed. Calculations using Equation 6.4 and a β_2 of $10^{19.8}$ M^{-2} for $[Cu(I)(BCS)_2]^{3-}$ gave average K_dCu(I) values of 2.1×10^{-17} M for CopZ-1317 and K_dCu(I) of 3.7×10^{-18} M for CopZ-3079 (Figure 6.10.C and D). Simulations of the data (solid lines, Figure 6.10.C and D) using the K_{eq} values were calculated from Equation 6.4 confirmed the accuracy of the data. Notably, CopZ-3079 appears to have a 5-fold higher affinity for Cu(I) than CopZ-1317 (Table 6.8).

The H22G and Y71F CopZ-1317 variants also displayed a clear linear decrease in absorbance with BCA until a break point at a ratio of ~ 1 CopZ/Cu(I) was reached (Figure 6.11.A and B). Competition assays with BCS gave average K_dCu(I) values of 6.0×10^{-17} M and 1.8×10^{-17} M for H22G and Y71F, respectively, with simulations of the data (solid lines, Figure 6.11.C and D) using the K_{eq} values calculated from Equation 6.4 confirming the accuracy of the data. Therefore a ~ 3 fold decrease in affinity for Cu(I) is observed for H22G, whereas the Y71F affinity is essentially identical to WT CopZ-1317. The values reported here for the variants and the WT CopZ proteins are comparable to those determined for other Atx1 family members (K_dCu(I) 10^{-18} M) (see Table 6.8) [372, 401, 404, 405].

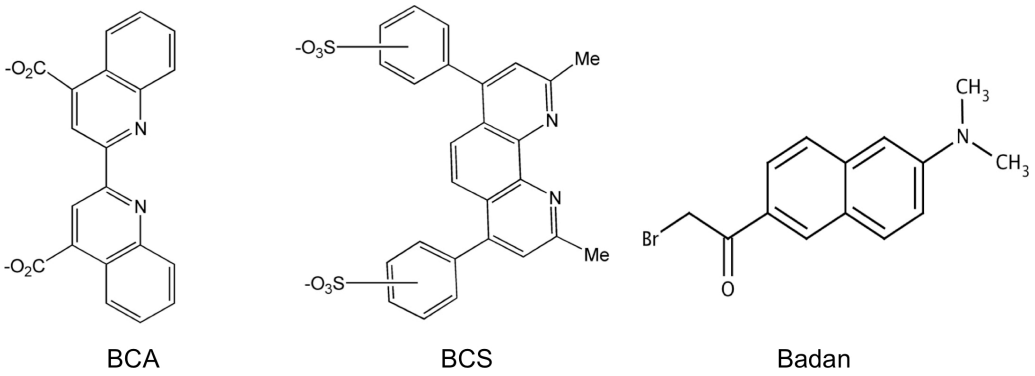


Figure 6.9: Chemical structures of the Cu(I) bidentate chelators BCA and BCS and the fluorescent probe Badan.

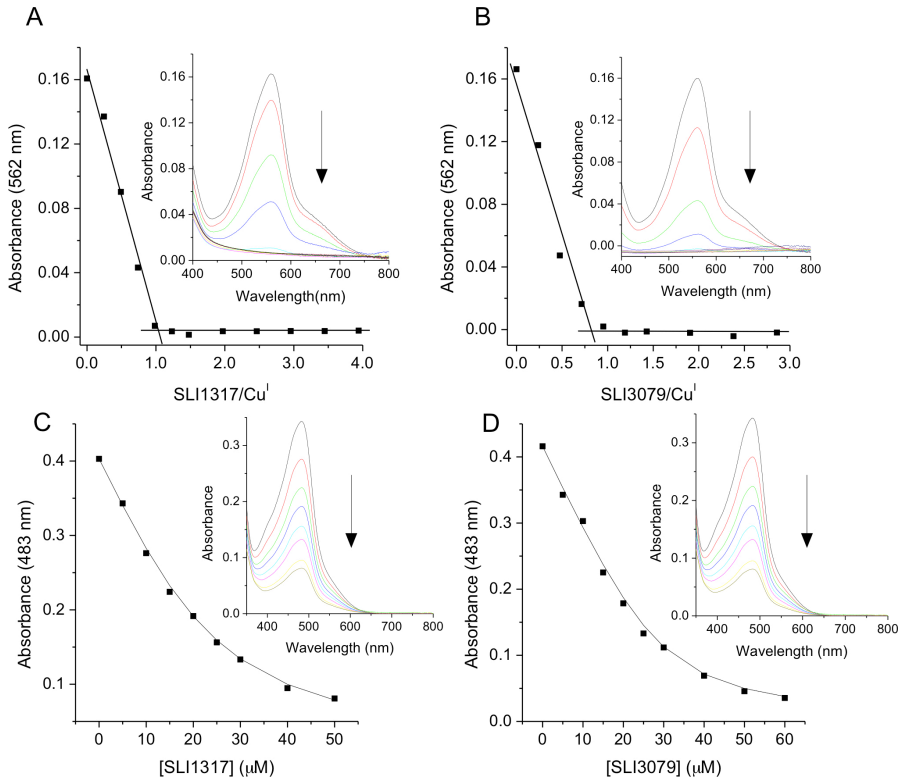


Figure 6.10: Cu(I) binding affinity and stoichiometry of CopZ-1317 and CopZ-3079 using the chromogenic affinity probes BCA and BCS. A and B) Inset, the absorbance at 562 nm of $[Cu(I)(BCA)_2]^{3-}$ decreasing to zero with addition of each apo-CopZ (1317/3079) proteins. Plots of absorbance change at 562 nm as a function of [protein/Cu(I)] indicates a ~ 1:1 stoichiometry for both CopZ proteins based on the intersection of the lines at the start and end of the titration. Conditions are reported in experimental procedures 6.2.11. C and D) Inset, under the Cu-limiting conditions imposed by $[Cu(I)(BCS)_2]^{3-}$ the absorbance at 483 nm in the visible spectrum of $[Cu(I)(BCS)_2]^{3-}$ decreases upon addition of each apo-CopZ. Plots of CopZ concentration versus the absorbance at 483 nm give K_dCu(I) values determined using Equation 6.4. Lines through the data points represents a best fit to the data using a K_dCu(I) of 2.1×10^{-17} M and 3.7×10^{-18} M for CopZ-1317 and CopZ-3079, respectively. Conditions are reported in experimental procedures 6.2.11.

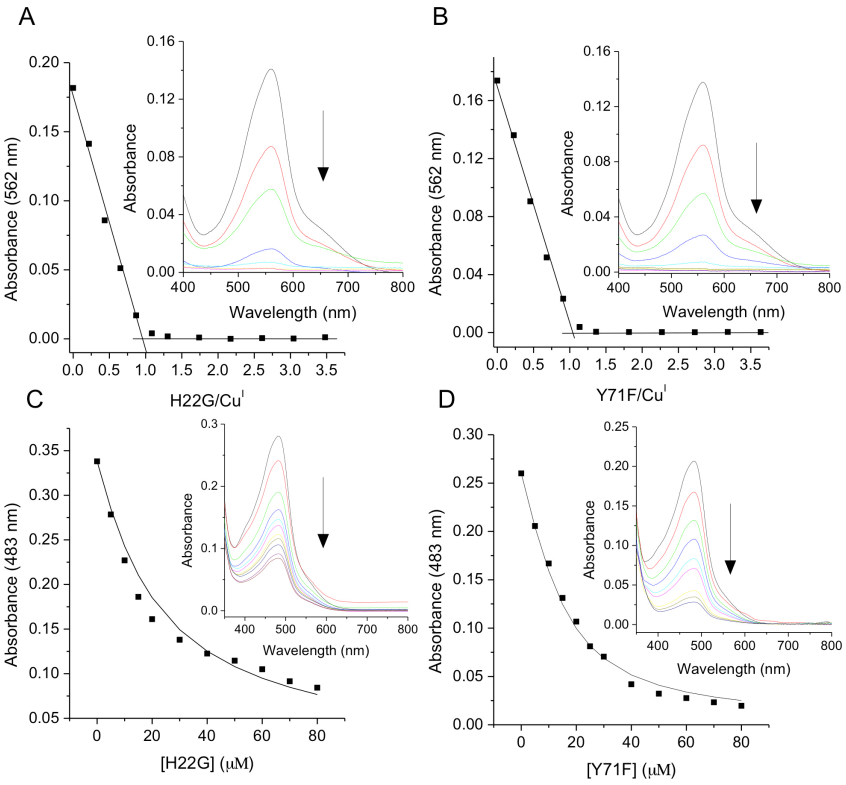


Figure 6.11: Cu(I) binding affinity and stoichiometry of CopZ-1317 variants H22G and Y71F using the chromogenic affinity probes BCA and BCS. A and B) Inset, absorbance at 562 nm of $[Cu(I)(BCA)_2]^{3-}$ decreasing to zero upon increasing additions of each apo-CopZ variant (H22G/Y71F). Plots of absorbance change at 562 nm as a function of [protein/Cu(I)] indicates a ~ 1:1 stoichiometry for both CopZ variants based on the intersection of the lines at the start and end of the titration. Conditions are as reported in experimental procedures 6.2.11. C and D) Inset, under the Cu-limiting conditions imposed by $[Cu(I)(BCS)_2]^{3-}$ the absorbance at 483 nm in the visible spectrum of $[Cu(I)(BCS)_2]^{3-}$ decreases upon addition of each apo-CopZ variant and the K_dCu(I) determined using Equation 6.4. The lines through the data points represents a best fit to the data using a K_dCu(I) of 6.0×10^{-17} M and 1.8×10^{-17} M for H22G and Y71F, respectively. Conditions are as reported in experimental procedures 6.2.11.

6.3.8. Determination of Cys pK_a values of CopZ-1317, CopZ-3079, and the H22G and Y71F variants

The ionization properties of the Cys Cu(I) binding residues within the $MX_1CX_2X_3C$ motif are important in facilitating Cu binding and transfer to a cognate partner [38, 372]. Ionization properties of Cys residues within CopZ-1317 and CopZ-3079 were monitored by following the absorbance at 240 nm (ΔA_{240}) (formation of Cu(I)-S(Cys) bond) as a function of pH and as well as using the fluorescent alkylating agent Badan [406]. Both CopZ proteins contain Tyr residues that deprotonate at high pH values and therefore to avoid their influence data was obtained at pH values of 10 and below [407]. The ΔA_{240} over the pH range employed for both CopZs could best be fit to Equation 6.6 that describes a single ionization process (Figure 6.12.A and B). Average pK_a values of 7.5 and 7.8 with $\Delta\varepsilon_{240}$ values of 4.4 and 3.5 were determined for CopZ-1317 and CopZ-3079, respectively (Table 6.8). The $\Delta\varepsilon_{240}$ values are consistent with an absorbance change for the formation of a single Cys thiolate ($\Delta\varepsilon_{240}$ 3 - 6 mM^{-1} cm^{-1}) [406]. Therefore, assigning an equivalent pK_a to both Cys residues is improbable. It is therefore unclear what the second Cys pK_a value is but has been reported for other CopZs to be very acidic (pH < 4) [404]. To validate these results pK_a values were measured over the same pH range using the fluorescent alkylating agent Badan (structure shown in Figure 6.9). Reaction of Cys side chains with alkylating agents has been well documented and can only occur once the Cys residue becomes deprotonated forming an ionized thiolate anion [408]. Measuring the rate of alkylation as a function of pH can be used to calculate the pK_a values of Cys thiol groups in proteins, where the observed rate constant is proportional to thiol deprotonation at that pH [409]. Badan acts as an alkylating agent creating a thioether bond with the thiolate leading to a significant increase in fluorescence intensity at 540 nm (excitation 391 nm). Analysis of the data obtained for CopZ-1317 and CopZ-3079 was again consistent with a single Cys ionization process (Equation 6.8), with pK_a values of 7.4 for CopZ-3079 and 7.8 for CopZ-1317 essentially identical to those determined from the ΔA_{240} measurements (Table 6.8 and Figure 6.13).

The ΔA_{240} for the H22G and Y71F variants were also monitored over the pH range 4 to 9.5 and the data fitted to Equation 6.6 giving pK_a and $\Delta\varepsilon_{240}$ values again consistent with the ionization of a single Cys thiolate (Figure 6.12.C and Table 6.8). The, H22G variant displays a pK_a value ~ 0.5 pH units higher than WT CopZ-1317, whereas, the pK_a value of Y71F remains within the error of the WT-1317.

Table 6.8: A comparison of the affinity constants and pK_a's for the WT *S. lividans* CopZ pair, the CopZ-1317 variants H22G and Y71F and, *S. lividans* CsoR, with other species relevant to this work also reported. The Cu(I) affinity constants were determined by BCS at pH 7.5, and the pK_a and $\Delta\varepsilon_{240}$ values determined using the ΔA_{240}.

Protein	pK_a		$\Delta\varepsilon_{240}$ mM^{-1} cm^{-1}		K_dCu(I) M
CopZ-1317	7.5 (0.2)		4.4 (0.1)		2.0 (0.2) x 10^{-17}
CopZ-3079	7.8 (0.2)		3.5 (0.25)		3.9 (0.5) x 10^{-18}
1317-H22G	8.0 (0.2)		3.7 (0.1)		6.0 (0.1) x 10^{-17}
1317-Y71F	7.8 (0.1)		3.4 (0.1)		1.7 (0.8) x 10^{-17}
	pK_{a1}	pK_{a2}	$\Delta\varepsilon_{(1)240}$	$\Delta\varepsilon_{(2)240}$	
[a]*Bs*-CopZ	6.1 (0.1)	< 4	-	-	~ 10^{-18}
[b]HAH1	8.9 (0.1)	5.5 (0.1)	[f]7.9 (0.5)	3.6 (0.2)	1.8 x 10^{-18}
[b]HAH1-K60A	8.9 (0.1)	7.0 (0.2)	[f]7.7 (0.5)	3.6 (0.4)	5.5 x 10^{-18}
[c]*Syn*-Atx1	nd	nd	nd	nd	7.0 x 10^{-19}
[d]*Sc*-Atx1	nd	nd	nd	nd	2.0 x 10^{-18}
[b]MNK1	9.2 (0.2)	7.0 (0.1)	4.0 (0.6)	3.3 (0.2)	2.8 x 10^{-18}
NmerA	9.0 (0.5)	6.4 (0.2)	2.6	5.5	na
NmerA-Y62F	9.3 (0.1)	6.9 (0.1)	3.3	2.8	na
[e]CsoR	9.3 (0.3)	6.7 (0.2)	1.9 (0.3)	2.6 (0.4)	2.6 x 10^{-18}

[a]*B. subtilis* CopZ K_dCu(I) determined at pH 7.5 and pK_a values determined using Badan [404]. [b]*H. sapiens* HAH1 and the MBD of the Menkes protein (MNK1) K_dCu(I) determined at pH 7.0 [372]; [c]*Synechocystis* PCC 6803 Atx1 K_dCu(I) determined at pH 7.5 [405]; [d]*S. cerevisiae* Atx1 K_dCu(I) determined at pH 7 [401]; [e]*S. lividans* CsoR K_dCu(I) determined at pH 7.5 [373]. In addition to the Cys residues in the MXCXXC motif a third Cys residue is present in HAH1, which based on the $\Delta\varepsilon_{(1)240}$ value suggests ionization in the same pH region as the N-terminal Cys (CysN pK_{a1}) [372].

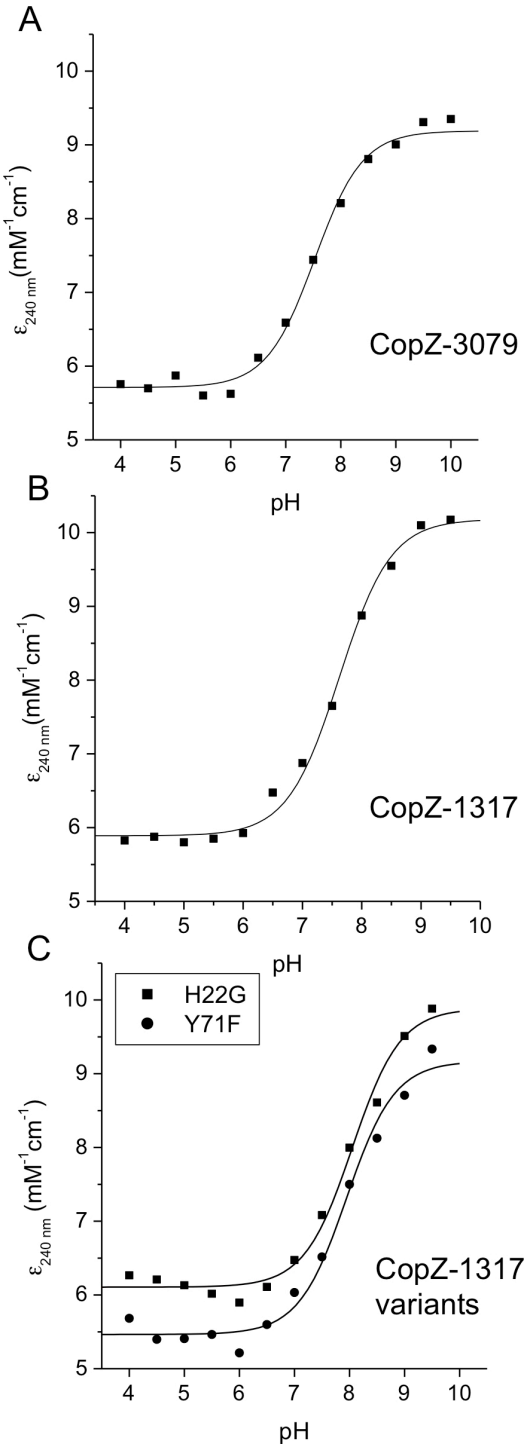


Figure 6.12: Determination of the Cys pK_a values in CopZ-3079, CopZ-1317 and the CopZ-1317 variants H22G and Y71F. A) A plot of $\varepsilon_{240\,nm}$ versus pH for CopZ-3079. B) A plot of $\varepsilon_{240\,nm}$ versus pH for CopZ-1317. C) Plots of $\varepsilon_{240\,nm}$ versus pH for the H22G and Y71F variants of CopZ-1317. The lines show fits of the data to Equation 6.6 and give pK_a and $\Delta\varepsilon_{240}$ values reported in Table 6.8. A series of mixed buffer systems containing 10 mM each of potassium acetate, MES, MOPS, Tris and 200 mM KCl with pH values of 4 to 10 with increments of 0.5 were used and apo-protein concentrations of ~ 40 μM.

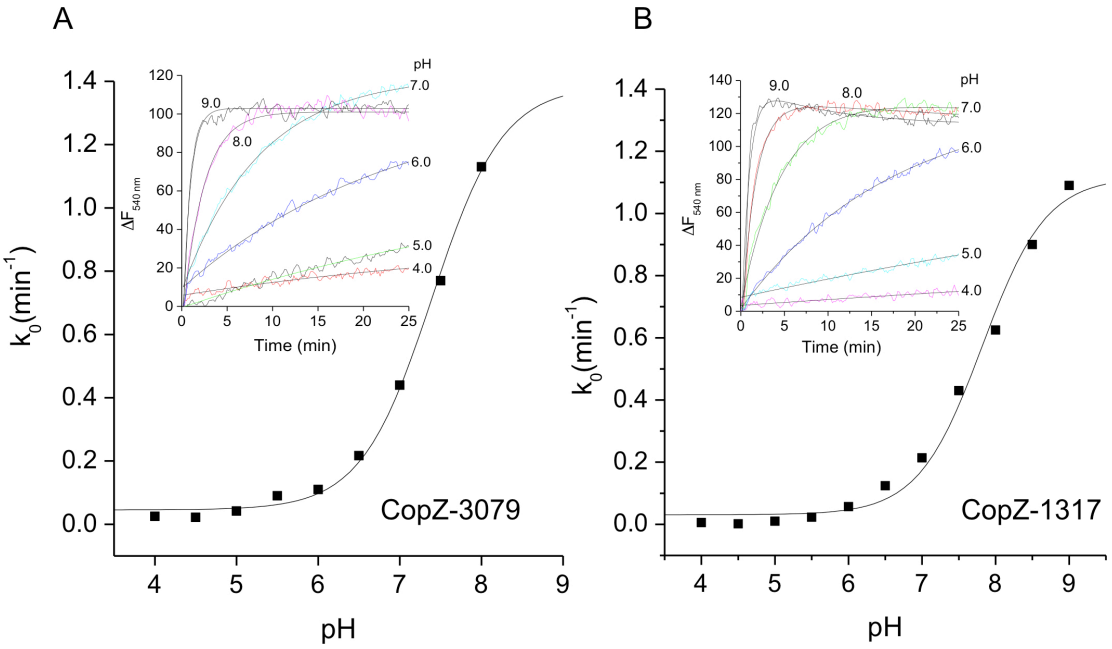


Figure 6.13: Determining the Cys pK_a values of CopZ-3079 and CopZ-1317 using the fluorescent alkylating agent Badan. Fluorescence intensity traces at 540 nm as a function of time following the addition of apo-CopZ-3079 (A) *(inset)* and apo-CopZ-1317 (B) *(inset)* at different pH values, as indicated. Lines represent a fit to an exponential function to obtain a pseudo first-order rate constant (k_0) at the different pH values, and plotted in (A) for CopZ-3079 and (B) for CopZ-1317. The lines show the fits to Equation 6.8 to give values reported in the text.

6.3.9. Determination of the Cys pK_a values of S. lividans CsoR

CsoR binds Cu(I) via Cys$^{75'}$, His^{100}, and Cys^{104} residues in a trigonal coordination with the two Cys ligands coming from two separate protomers of the tetramer assembly (Figure 6.1.B and C) [373]. The ionization properties of the Cu(I) Cys ligands in CsoR were assessed by monitoring the ΔA_{240} in the pH range 5 to 9.5 (Figure 6.14.A). Two ionization processes are clearly apparent and fit well to Equation 6.7 to give pK_a values of 9.3 and 6.7 and $\Delta\varepsilon_{240}$ values of 1.9 and 2.6 (Table 6.8). These values were confirmed using the fluorescent probe Badan also producing two ionization processes giving pK_a values of 8.6 and 6.7, thus similar to those determined by the ΔA_{240} measurements (Figure 6.14.B).

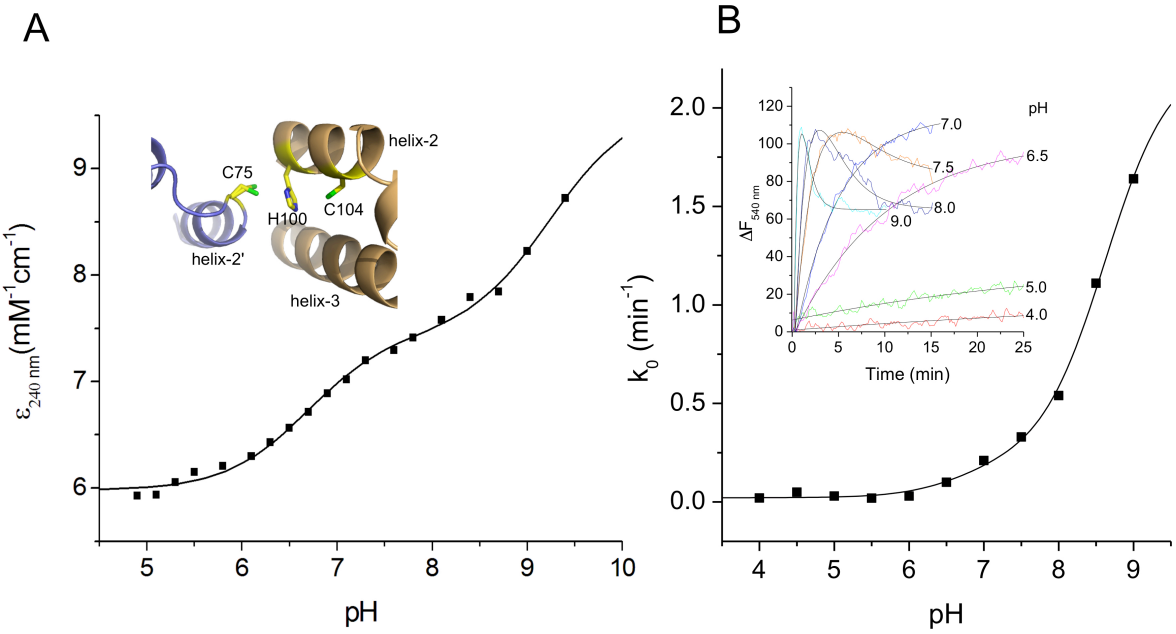


Figure 6.14: Cys pK_a values for *S. lividans* CsoR. A) A plot of $\varepsilon_{240\ nm}$ against pH with the line representing a fit of the data to Equation 6.7 to give pK_a and $\Delta\varepsilon_{240}$ values reported in Table 6.8. Inset, a zoomed-in view of the Cu(I) ligands in the X-ray structure (PDB 4ADZ) [373] of *S. lividans* apo-CsoR, Cys^{75}, Cys^{104} and His^{100} shown as sticks. B) pK_a properties of the Cys thiols of CsoR using the fluorescent alkylating reagent Badan. Inset, fluorescence intensity traces at 540 nm as a function of time following the addition of apo-CsoR at different pH values, as indicated. Lines represent a fit to an exponential function to obtain pseudo first-order rate constants (k_0) at the different pH values, and plotted. The lines show the fits to Equation 6.9, with pK_a values reported in the text.

6.3.10. Cu(I) transfer monitored through EMSAs

CopZ proteins have the ability to traffic Cu(I) to cytosolic targets such as the MBD of the ATPase CopA containing $MX_1CX_2X_3C$ motifs but they can also traffic Cu(I) to other targets such as Zn(II)CopY in *E. hirae* containing a CXCXXXXCXC motif [375, 410]. EMSAs were conducted to establish if there is a Cu(I) trafficking pathway between the CopZ proteins and CsoR in *S. lividans*. Apo-CsoR has previously been shown to bind to three DNA operator sites with the sequences for the individual targets determined (Figure 6.2 and 6.15) [373]. Two apo-CsoR tetramers bind to the operator DNA target, resulting in a low-mobility CsoR:DNA complex as observed by a retardation of the DNA in the EMSA (Figure 6.15) [373, 395]. However, when the DNA is incubated with Cu(I)-loaded CsoR there is no band shift, with only a high-mobility band corresponding to the free DNA observed (Figure 6.15). Controls containing Cu(I) loaded or apo-CopZ-1317 and CopZ-3079 with the DNA targets did not affect the mobility confirming their inability to bind to the operator DNA target sequences. Samples containing the DNA target, apo-CsoR and 1 or 10 molar equivalents of Cu(I)-loaded CopZ-1317 or CopZ-3079 were prepared under anaerobic conditions. Strikingly, the EMSAs no longer display the low mobility band corresponding to the apo-CsoR:DNA complex, but instead a high mobility band is present corresponding to the free DNA (Figure 6.15.A). This was observed for all three operator targets and suggests that Cu(I) is transferred from both CopZ proteins to apo-CsoR causing dissociation (derepression) from the operator DNA (Figure 6.15.C). The reverse experiment was conducted in which 1 to 10 molar equivalents of either apo-CopZ protein were incubated with a mix of Cu(I)-CsoR and DNA under anaerobic conditions (Figure 6.15.B). All three operator DNA show identical results with only the high mobility band being observed indicating that Cu(I) has not been transferred from Cu(I)-CsoR to either apo-CopZ proteins (Figure 6.15.C). These EMSA studies therefore infer that Cu(I) transfer is unidirectional from either CopZ to CsoR.

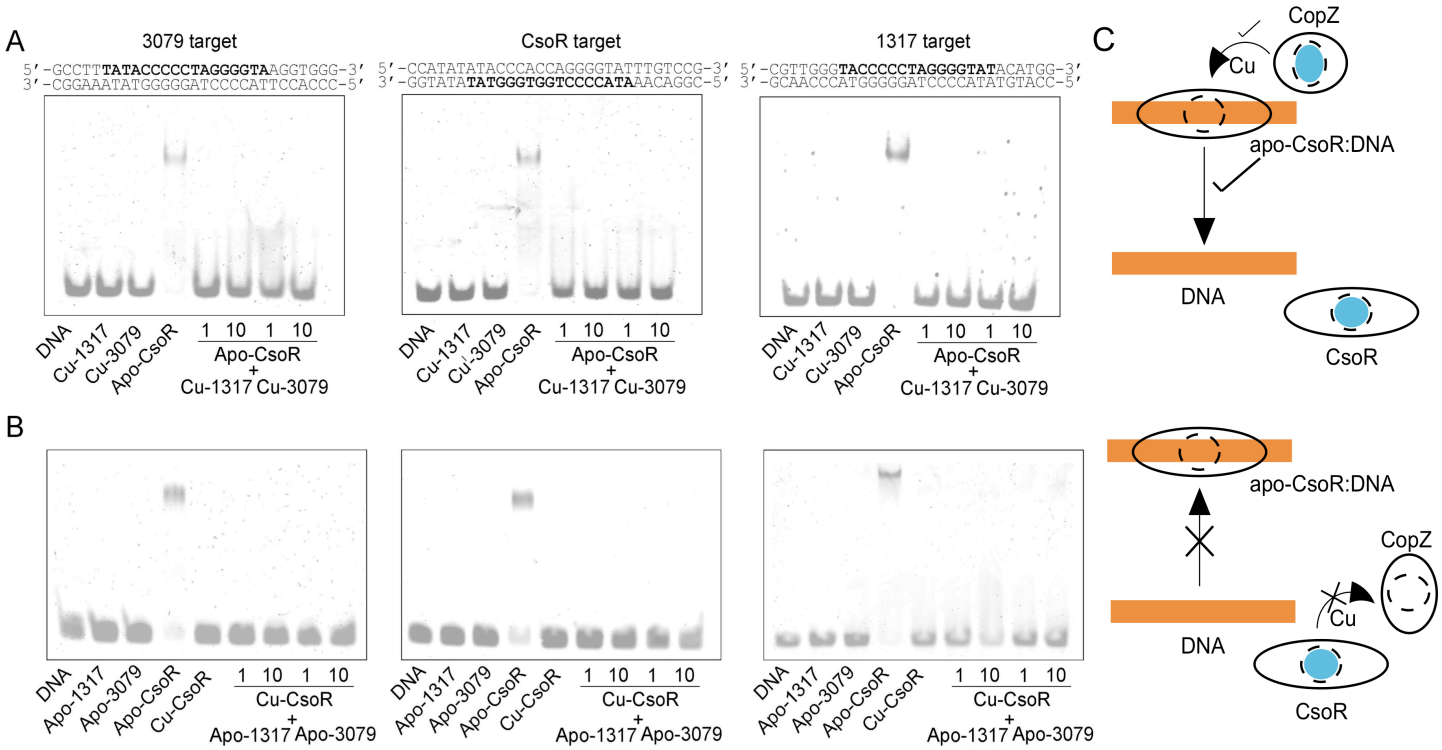


Figure 6.15: Cu(I) transfer between *S. lividans* CsoR and CopZ-3079/1317 probed by EMSA. Experiments were performed with three CsoR operator sequences (*1317*, *3079* and *csoR*) with the individual CsoR operator sites used indicated above the corresponding (6 %) EMSA gel, with the consensus sequence for the individual targets shown in bold. A) 1 or 10 molar equivalent of Cu(I)-loaded CopZ proteins were incubated with apo-CsoR. B) 1 or 10 molar equivalent of apo-CopZ (3079/1317) proteins were mixed with Cu(I)-loaded CsoR. The components present in each lane of the gels are indicated with the concentrations as follows: 0.5 μM [DNA], 4 μM [apo-CsoR-monomer], 4 μM [Cu(I)-CsoR-monomer] (equivalent to 0.5 μM 2 x CsoR tetramer), 4 and 40 μM [apo-CopZ] or [Cu(I)-CopZ] (equivalent to a 1:1 and a 1:10 CsoR-monomer:CopZ). C) The scheme illustrates that Cu (blue sphere) can be transferred from Cu(I)-CopZ (1317/3079) to apo-CsoR bound to the DNA, leading to CsoR dissociating from the DNA. It also shows that Cu(I)-CsoR cannot transfer Cu(I) to either apo-CopZ (1317/3079) protein, therefore not allowing it to bind to the DNA.

6.3.11. Cu(I) transfer using analytical gel filtration

To corroborate Cu(I) transfer results from EMSA experiments, analytical gel filtration experiments were conducted using CopZ-3079 and CsoR. Cu(I) transfer by monitoring CopZ-1317 by gel filtration was not feasible due to the elution volume of Cu(I)-CopZ-1317 overlapping with that of CsoR (~ 9.4 ml for Cu(I)-CopZ-1317 and ~ 9.0 ml for apo- and Cu(I)-CsoR). Cu(I)-CopZ-3079 was incubated with apo-CsoR under anaerobic conditions and the elution profile monitored (Figure 6.16.A). Excitingly, the peak corresponding to Cu(I)-CopZ-3079 (~ 10.8 ml) was no longer observed but in its place two peaks at ~ 8.9 ml and 12.5 ml corresponding to Cu(I)-CsoR and apo-CopZ-3079 are identified (Figure 6.16.A). This demonstrates that Cu(I) has been transferred from CopZ-3079 to the CsoR (Figure 6.16.B). The reverse experiment, whereby Cu(I)-CsoR was mixed with apo-CopZ-3079 under anaerobic conditions displayed two elution peaks indicating that Cu(I)-CsoR and apo-CopZ-3079 still remain intact and Cu(I) has not been transferred from CsoR to CopZ-3079 (Figure 6.16.C and D). Combining these observation with those from EMSA experiments, it can be concluded that *in vitro* Cu(I) is transferred in a unidirectional manner from CopZ proteins to CsoR.

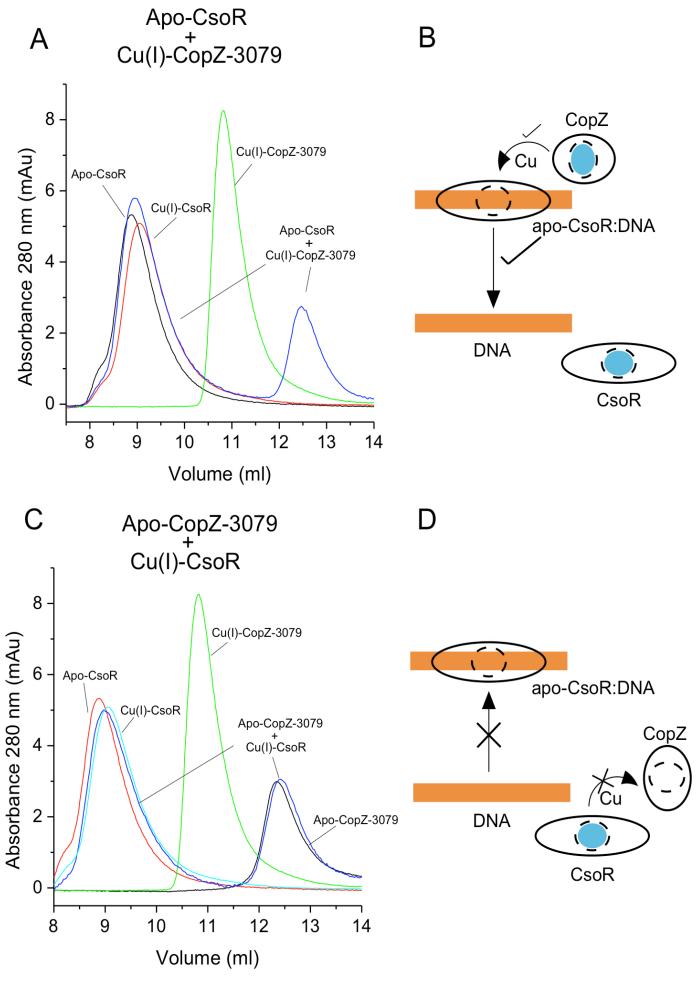


Figure 6.16: Analytical gel filtration monitoring transfer of Cu(I) between CopZ-3079 and CsoR. The absorbance at 280 nm is monitored as a function of the elution volume with samples of CopZ-3079 and CsoR prepared anaerobically in 10 mM MOPS, 150 mM NaCl, pH 7.5. A) Apo-CsoR mixed with Cu(I)-CopZ-3079 (blue), and controls of Cu(I)-CopZ-3079 (green), apo-CsoR (black) and Cu(I)-CsoR (red). B) A scheme illustrating the transfer of Cu(I) from CopZ-3079 to CsoR. C) Apo-CopZ-3079 mixed with Cu(I)-CsoR (blue) and controls of apo-CsoR (red), Cu(I)-CsoR (cyan) and Cu(I)-CopZ-3079 (green). D) A scheme showing the inability of Cu(I)-CsoR to transfer Cu(I) to CopZ-3079.

6.4. Discussion

6.4.1. Three out of four copZ genes in S. lividans are under CsoR control

From transcription analysis using the *S. lividans* genome, four *copZ* and five P-type ATPase (*copA*) encoding genes have been identified (Table 6.6). Four of the P-type ATPase genes are immediately downstream of a *copZ* gene but one P-type ATPase (0845) has been identified exclusive of a gene encoding a CopZ (Table 6.6). In the *csoR* mutant (*ΔcsoR*) even though all four *copZ*/ATPase couples contain a CsoR binding site only three of these become induced (Table 6.6). Therefore, the 0895-0896 operon would appear not to be under the control of SLI4375 (CsoR), raising the possibility of a second CsoR-like transcriptional repressor that uses an analogous binding site. The presence of another Cu-sensing repressor could explain why induction levels in the *ΔcsoR* mutant are lower than in cultures with added Cu. Moreover, the large number of proteins involved in buffering and removing cytoplasmic Cu(I) may provide evidence why *S. lividans* depends on higher Cu levels to fully initiate development compared with other *Streptomyces* species [4, 5, 24].

6.4.2. Higher order assemblies in CopZ-1317 and CopZ-3079

Homologous CopZ proteins have been well studied and reveal different assemblies in solution upon Cu(I) binding. Atx1 from *S. cerevisiae* and CopZ from *E. hirae* have both been identified as monomers in solution [410-412], while, HAH1 (Atox1) forms homo-dimers [370, 413] and Atx1 from *Synechocystis sp*. PCC 6803 can adopt multiple oligomeric states when metalated [414]. CopZ from *B. subtilis* shares ~ 30 % sequence identity to CopZ proteins from *S. lividans* and displays a number of discrete Cu(I)-bound forms, including $Cu(CopZ)_2$, $Cu_2\ (CopZ)_2$, $Cu_3\ (CopZ)_2$, $Cu_3\,(CopZ)_3$ and $Cu_4\,(CopZ)_2$ [387-389]. The latter two were identified through high resolution crystal structures but initial reports were based on experimental evidence similar to that described here. This described three species, in which, species A is a dimer formed between 0-0.5 Cu(I)/protein whereby a Cu(I) ion is shared at the monomer interface. Species B is a dimer formed at 1 Cu(I)/protein leading to a dinuclear Cu(I) site between two monomer, and species C is formed at 1.5 Cu(I)/protein inferred to consist of a dimer with a trinuclear Cu(I) site [388]. CopZ-3079 displays clear similarities to these findings from *B. subtilis*. A species consistent with a dimer was identified from the elution volume using analytical size exclusion chromatography upon 1 equivalent Cu(I) loading to CopZ-3079 (Figure 6.8.B). The Cu(I) bidentate chelator BCA identified one Cu(I) ion bound per monomer of CopZ-3079 (Figure 6.10.B), also confirmed by the increase in absorbance at 240 and 300 nm up to 1 Cu(I)/CopZ (Figure 6.7.C). This suggests a species

comparable to species B ($Cu_2(CopZ)_2$) from *B. subtilis* with a dinuclear Cu(I) site formed between two monomers of CopZ-3079, however may depend upon the amount of Cu loading (Figure 6.8.B, inset). In contrast, the behaviour of CopZ-1317 in solution is more difficult to interpret. Gel filtration data displays a monomeric species at 1 Cu(I)/CopZ-1317 and a higher order assembly predicted as a tetramer observed at 1.5 - 2 Cu(I)/CopZ (Figure 6.8.C). This coincides with two clearly visible absorbance transitions between 0 - 1 and 1.5 - 2 Cu(I)/CopZ-1317 at 240 nm upon Cu(I) titration. A multiphasic absorbance increase is identified at 335 nm indicative of Cu(I) cluster formation confirming the presence of a higher order assembly (Figure 6.7.B). However, the Cu(I) chelator BCA identified a binding stoichiometry of 1 Cu(I) ion per CopZ-1317 and should be upheld in higher order speciation (Figure 6.10.A). It can therefore be envisaged that CopZ-1317 upon Cu(I) addition forms a tetramer in solution with four Cu(I) ions present ($Cu_4(CopZ)_4$) not previously reported (Figure 6.8.C, inset).

Gel filtration experiments also demonstrate that in the presence of excess DTT the formation of the higher order species is no longer viable and both CopZs remain monomeric (Figure 6.8.A). The ability of DTT to alter the assembly states of CopZ from *B. subtilis* has previously been reported, clearly showing that DTT is able to compete with the cysteine thiols for Cu(I) [388]. DTT has an affinity K_dCu(I) ~10^{-15} M at pH 7, so while this is a weaker affinity than the CopZ proteins, once in large excess DTT can either outcompete for Cu(I) or coordinate the Cu(I) preventing the formation of higher oligomeric states [401]. The cytosolic environment of bacteria is extremely reducing, maintained by mycothiol (MSH) in Actinomycetes, bacillithiol (BSH) and cysteine in *Bacillus* and *Staphylococcus* species, in contrast to glutathione (GSH) in eukaryotes and gram negative bacteria [415-418]. Thus, under these reducing conditions as demonstrated using DTT as the reductant the CopZ proteins remain monomeric. It is therefore unclear whether the formation of these higher order assemblies and the differences identified between CopZ-1317 and CopZ-3079 have any physiological importance within the cytosol of *S. lividans* [381, 389].

6.4.3. CopZ affinities for Cu(I)

Competition for Cu(I) between apo-CopZ proteins and $[Cu(I)(BCS)_2]^{3-}$ was observed allowing the Cu(I) binding affinities to be determined. Average K_dCu(I) values for CopZ-1317 and CopZ-3079 of 2 x 10^{-17} M and 3.9 x 10^{-18} M, were determined at pH 7.5 (Figure 6.10 and Table 6.8). Both display high affinities for Cu(I) cementing their suggested roles as chaperones within the cytoplasm preventing the presence of unchelated Cu(I) in *S. lividans*. Recent reports using BCS at physiological pH have determined attomolar affinities (K_dCu(I) 10^{-18} M) for Atx1 family members and are therefore comparable to those reported for CopZ-1317 and CopZ-3079 [372, 401, 404, 405]. Notably, CopZ-3079 appears to have a 5-fold

higher affinity for Cu(I) than CopZ-1317 and thus lies closer to the affinities reported for CopZ from *B. subtilis*, HAH1 and the Atx1 chaperones from yeast and *Synechocystis* PCC6803 (Table 6.8) [372, 401, 404, 405]. CopZ-1317 contributes a total transcription level close to zero under homeostasis conditions and less than 2 % under Cu stress conditions, thus the lower Cu(I) affinity observed does not have a direct physiological implication. In comparison, CopZ-3079 contributes significantly to the total transcription level with a 9-fold increase upon Cu addition (Table 6.6). It would therefore be biologically advantageous that CopZ-3079 posses a higher Cu(I) binding affinity enabling Cu(I) to be buffered during homeostasis. Under Cu stress conditions once buffering capacity has become exceeded, Cu(I) may be sensed directly by CsoR to trigger a transcriptional response, or alternatively, CopZ-3079 may act to chaperone Cu(I) to apo-CsoR bound to its operator DNA.

The H22G and Y71F variants of CopZ-1317 also displayed competition for Cu(I) with BCS giving average K_dCu(I) values of 6 x 10^{-17} M and 1.7 x 10^{-17} M respectively (Table 6.8) (Figure 6.11). Y71F presents an affinity for Cu(I) essentially identical to the WT protein, whereas, removal of His^{22} leads to a ~ 3-fold decrease in affinity (Table 6.8). An explanation of this decrease in Cu(I) affinity could be due to a decrease in the concentration of the thiolate form of the Cys residues at pH 7.5 (pK_a = 8 for H22G, Table 6.8), supporting the concept that low Cys proton affinity (low pK_a values) correlates to an increase in Cu(I) affinity [402, 404].

6.4.4. Single Cys ionization processes are detected for CopZ proteins

Ionization is determined by the pK_a of the amino acid side group and the surrounding environment. This can result from interactions of the side chain with other residues and the polarity of the surrounding medium [419]. To assess the ionization properties of the Cys residues within the $MX_1CX_2X_3C$ motif of the CopZ proteins, two probes (Badan and ΔA_{240}) were used to measure the pK_a values. Analysis of the data produced single Cys ionization processes with pK_a values of 7.5 and 7.8 for CopZ-1317 and CopZ-3079, respectively (Table 6.8 and Figure 6.12). The data is consistent with a single Cys ionization observed for CopZ from *B. subtilis* in the pH range 4 to 7 with a calculated pK_a of 6.1 (Table 6.8) [404]. However, using Badan in conjunction with mass spectrometry, a second pK_a < 4 for CopZ from *B. subtilis* was observed [404]. Both these pK_a values are significantly lower than for a free Cys (pK_a ~ 9) but infers that the one with the pK_a < 4 has its thiolate form stabilized (relative to the thiol form) to a greater extent compared to the Cys with a pK_a of 6.1 [404].

Thioredoxin-like proteins also contain a CXXC motif and the acid-base properties of the Cys residues provide useful comparisons [409, 420]. The thioredoxin DsbA from *E. coli* displays a low pK_a of 3.5 assigned to the N-terminal Cys (Cys^N). A His residue located at the X_2 position of the CX_1X_2C motif is believed to interact electrostatically with the N-terminal

Cys thiolate, stabilizing it and resulting in the low pK_a value [421]. A His residue at the equivalent position (X_3 of the $MX_1CX_2X_3C$ motif) can be identified in CopZ-1317 and CopZ-3079 and also in *E. hirae* and *B. subtilis* CopZs, but is a Gly in eukaryotic proteins (* Figure 6.3). To determine whether this His residue is a contributing factor in modulating the pK_a of a Cys to < 4 within the *S. lividans* CopZ pair, H22G was prepared. The ΔA_{240} was monitored over the same pH range as WT CopZ-1317 giving a pK_a of 8 and $\Delta\varepsilon_{240}$ of 3.7 (Table 6.8) (Figure 6.12) consistent with the ionization of a single Cys residue. Thus replacing His^{22} with an inert Gly residue does not result in the observation of a second ionization process, meaning that His^{22} is not stabilising the thiolate of a Cys residue creating an acidic pK_a (< 4). However, the single pK_a value observed is ~ 0.5 pH units higher than WT CopZ-1317, suggesting His^{22} does have a role in modulating the pK_a of the observable Cys in the pH range studied (Table 6.8). This increase in pK_a correlates to an observed ~ 3-fold decrease in affinity due to a decrease in thiolate (compared to thiol) at pH 7.5 as predicted (*vide supra*) (Figure 6.12 and Table 6.8).

In contrast to CopZ-1317 and CopZ-3079, HAH1 consists of two well resolved ionization processes, with pK_a values of 5.5 and 8.9 (Table 6.8) [372]. Through structural and biochemical studies the acidic pK_a was assigned to the C-terminal (Cys^C) residue and the more basic pK_a to the N-terminal (Cys^N), the opposite to the thioredoxin proteins [372]. The acidic pK_a value is assignable to a hydrogen bond between the amino group of Lys^{60} and the Sγ of Cys^{15} (Cys^C) (* Figure 6.3) [372, 419, 421]. A similar H-bond has been identified in the structure of Atx1 from yeast (* Figure 6.3) [356]. This polar interaction stabilizes the thiolate form of Cys^C compared to Cys^N, decreasing the pK_a and decreasing the nucleophilic potential of the Cu(I) ligand (better leaving group) [372]. The K60A mutant of HAH1 confirmed the affect of this basic residue, resulting in an increase in the pK_a of Cys^C to 7 due to the thiolate no longer being stabilized. This increase in the pK_a value was in combination with a successive 3-fold decrease in Cu(I) affinity relative to WT as expected (Table 6.8) [372]. This behaviour has also been reported for the N-terminal heavy metal-associated domain (NmerA) of MerA, a Hg^{2+} reductase which contains Tyr^{62} at the equivalent position of Lys^{60} in HAH1 [407]. Two ionization processes are observed for NmerA with pK_a values of 6.4 and 9 assigned to Cys^C and Cys^N, respectively. They were assigned on the basis that the thiolate of Cys^C is stabilized through a H-bond to Tyr^{62} [407]. Mutation of this Tyr to a non-polar Phe residue increased the pK_a of both Cys residues to 6.9 and 9.3 confirming the stabilizing role of the polar Tyr residue (Table 6.8) [407].

CopZ-1317 displays a Tyr residue at the equivalent position (* Figure 6.3) so in correlation with NmerA and HAH1 it could be expected to play a role in lowering the pK_a value of Cys^C. To investigate this theory Tyr^{71} in CopZ-1317 was mutated to a non-polar Phe residue (Y71F). Nevertheless, Y71F revealed a single ionization process with a pK_a value

within the error of the WT CopZ-1317 protein (Table 6.8) (Figure 6.12). Therefore, replacing Tyr^{71} with a non-H-bonding Phe residue did not result in the observation of a second pK_a or affect the single observed ionization event. Additionally, a Phe residue is present at the equivalent position in CopZ-3079 again producing a single observable ionization event (Figure 6.12). Remarkably, a Phe residue is located at the equivalent position in the MBDs of ATPases such as MNK1 serving to increase the pK_a value of Cys^C (Table 6.8). The deficiency of the stabilizing H-bond increases the nucleophilicity of Cys^C in the MBD making it a better acceptor in the transfer of Cu(I) from HAH1 [372].

Therefore, based on these experiments and those reported for CopZ from *B. subitlis* [404] it can be speculated that bacterial CopZ proteins retain a more acidic (low proton affinity) Cys pK_a value compared with eukaryotes (Table 6.8). For the *S. lividans* CopZ pair only a single ionization event could be determined with the prediction of an acidic pK_a < 4 in correlation with *B. subtilis*. Through mutational studies, two residues (His^{22} and Tyr^{71}) thought to play a role in modifying the ionization potential were replaced. However, neither of these mutations led to a second detectable ionization process, thus, other mechanisms stabilizing the thiolate of Cys^C must be present in bacterial CopZ proteins. In Atx1 members and thioredoxins, one of the Cys residues is positioned at the N-terminus of α-helix 1. The electrostatic effect of having a net partial positive charge at the N-terminus of the helix (macrodipole) can account for a slight stabilization of the Cys thiolate (~ 1.5 units) [419, 422]. Other mechanisms such as H-bonding with water molecules due to the solvent exposure of the Cys residues, and cationic interactions with neighbouring amino acid side chains can also act to decrease the apparent pK_a [419]. In human thioredoxin the increased stabilization of Cys^N S^- is due to a hydrogen bond with the backbone amide of Cys^C and in glutaredoxin it is because of a H-bond with Cys^C SH, an equivalent interaction could exist in the CopZs from *S. lividans* [422, 423]. In thioredoxins, H-bond interactions alone can contribute to lowering the Cys pK_a by up to 3 units and cationic interactions with the macrodipole of the N-terminal helix by up to 5.6 units [419]. Therefore, the acidic pK_a value of Cys^C in bacterial CopZ proteins must be due to a combination of the N-terminal macrodipole of α-helix 1 and H-bond and/or cationic interactions other than with His^{22} or Tyr^{71}. However, although H22G did not result in the observation of a second ionization process, the remaining pK_a value was increased by ~ 0.5 pH units assigned to Cys^N. This may explain why bacterial proteins possess more acidic Cys^N pK_a values compared to eukaryotes due to His^{22} stabilizing the thiolate, which is a non-cationic Gly residue in eukaryotes (Figure 6.3 and Table 6.8). Therefore, when a CopZ protein is bound to Cu(I) the Cys^N will act as a better leaving group (less nucleophilic) due to the lower pK_a value (more stabilized) compared to the acceptor which has a higher Cys^N pK_a value (more nucleophilic), thus favouring transfer [372].

6.4.5. Cu(I) transfer is unidirectional from CopZ to CsoR

Cu(I) transfer from CopZ to the MBD of CopA P-type ATPase has been identified in *B. subitlis* and *E. hirae* which retain the $MX_1CX_2X_3C$ motif present within Atx1-like domains [375, 386, 389]. Cu(I) transfer has also been identified from CopZ to the Zn(II)-CopY repressor in *E. hirae,* which in contrast features a CXCXXXXCXC motif [375, 410]. Furthermore, Cu(I) transfer has been reported for Atx1 in *Synechocystis* PCC6803 to plastocyanin [424]. Therefore, transfer from CopZ-1317 and CopZ-3079 to the Cys_2His motif of CsoR was explored. EMSA and analytical gel filtration experiments revealed that both Cu(I)-CopZ proteins are capable of facile Cu(I) transfer to the transcriptional regulator CsoR (Figure 6.15 and 6.16). This transfer is unidirectional with Cu(I)-CsoR unable to transfer Cu(I) to either CopZ protein. The observation of transfer suggests a complex likely to be transient between the proteins with Cu(I) transfer occurring via a ligand exchange mechanism. Shallow but favourable thermodynamic affinity gradients exist between CsoR and the CopZ pair (Table 6.8). A K_dCu(I) of 2.6 x 10^{-18} M has been reported for CsoR, thus, creating a very shallow gradient with CopZ-3079 (2.9 x 10^{-18} M) and a more favourable gradient with CopZ-1317 (2 x 10^{-17} M) [373] (Table 6.8). However, the Cu(I) affinity for CsoR bound to DNA has not yet been determined and is likely to be a weaker affinity due the thermodynamic linkage of CsoR being in equilibrium with its two ligands, the DNA operator target and the Cu(I) metal ion [377, 425]. This would create an unfavourable thermodynamic affinity gradient, but nevertheless Cu(I) trafficking has been identified in such systems [405]. Therefore, kinetic factors must play a role in Cu(I) trafficking pathways and the reactivity of the ligands involved will be paramount to whether transfer can occur. To determine whether this is the case for CsoR in *S. lividans* the ionization properties of the coordinating Cys residues were examined using two probes (ΔA_{240} and Badan). Two ionization processes are apparent with pK_a values of ~9 and 6.6 (Figure 6.14). From analysis of the crystal structure no H-bond interactions with either Cys are apparent, with the exception of water molecules due to their high solvent exposure. Cys^{75} is positioned on a loop between α-helix 1 and 2 towards the N-terminus of α-helix 2, with its side chain modelled in two orientations illustrating the increased flexibility in this area. Cys^{104} is present towards the C-terminal end of α-helix 2 of an opposite protomer, and consequently does not experience any affects from a helix macrodipole [373]. Therefore, the more acidic pK_a of 6.6 could be assigned to Cys^{75} whereby the thiolate is stabilized by the α-helix 2 macrodipole. Crucially, with the absence of any stabilizing H-bonds, both Cys residues can potentially act as strong nucleophiles. At physiological pH the more acidic Cys residue would be predominately deprotonated (~ 83 % thiolate form at pH 7.4). Thus, this coupled to the increased electrophilic behaviour of Cys^N (and possibly Cys^C) in the CopZ proteins may lower the activation barrier for Cu(I) transfer. Therefore allowing the observed unidirectional Cu(I) transfer from both CopZ proteins to the

transcriptional repressor CsoR, irrespective of whether the thermodynamic gradient is unfavourable.

6.4.6. The role of CopZ proteins within the CsoR regulon in S. lividans

The investigation on the pair of CopZ proteins from *S. lividans* so far has revealed their roles as cytosolic Cu(I) chaperones and identified their ability to traffic Cu(I) in a unidirectional manner to CsoR. In addition to the expanded knowledge of the *S. lividans* genome a model has been determined for Cu(I) handling (Figure 6.17). The higher affinity of CopZ-3079 for Cu(I) compared to CopZ-1317 and the observed basal levels of this CopZ/CopA pair fits well with this CopZ acting as a buffer to maintain Cu homeostasis. Expression levels of 0895/0896 and 1063/1064 pairs were also identified during homeostasis, inferring a similar role to CopZ-3079 creating a large pool including apo-CsoR capable of maintaining homeostasis. Upon Cu stress conditions, data here suggests that CopZ-1317 and CopZ-3079 transfer Cu(I) to CsoR, which have their Cys residues optimized for ligand exchange. Firstly, CopZ-3079/1063/0895 could transfer Cu(I) to the DNA bound CsoR, releasing it from the DNA resulting in derepression of the three *copZ/copA* efflux systems. Levels of CopZ could then continue to buffer Cu(I) levels, eventually aided by CopZ-1317 (Figure 6.17 steps 1-3). The Cu(I) loaded CopZ proteins can then unload their Cu(I) to their cognate CopA P-type ATPase for its removal from the cytoplasm (Figure 6.17 step 4). Partial derepression of the *csoR* gene can then occur as identified by RNA-seq data and confirmed through data indicating different thermodynamic binding properties of this target operator site [395] (Figure 6.17 step 5). Once the Cu(I) levels have reduced, it is essential that the system resets back to homeostasis. CopZ proteins can traffic Cu(I) to their cognate ATPase renewing the supply of apo-CopZ. Transcription deactivation of the Cu(I)-bound CsoR will likely involve the unbinding of a promoter-bound Cu(I)-CsoR from the DNA followed by the binding of an apo-CsoR repressor, rather than the dissociation of the Cu(I). A recent *in vivo* study with CueR and ZntR has provided evidence for a concentration dependent unbinding from chromosomal recognition sites that involves either assisted dissociation or direct substitution by an apo-regulator from the cytoplasm [426]. Whether a similar mechanism occurs with CsoR is not known. It should be noted however, that unlike CueR or ZntA, CsoR does not possess a structurally recognisable DNA binding motif, which is present in both CueR and ZntA, and plays a significant role in distorting the DNA structure, which strongly influences the concentration dependent binding and unbinding to the DNA [426, 427]. Although a clear Cu(I) trafficking model can be elucidated for *S. lividans,* variations in the functions of CopZ proteins can exists across bacterial species. As illustration of this the CopZ chaperone present in the pathogenic bacterium *Listeria monocytogenes* is not required for the

Cu induction of its CsoR and in fact deletion of CopZ increases the transcriptional response [428]. It can be concluded that in the cytoplasm of *S. lividans* the CopZ pair, CopZ-1317 and CopZ-3079 are indeed Cu(I) metallochaperones important for shuttling intracellular cuprous ions to their intended targets, thus preventing deleterious effects of inappropriate Cu binding.

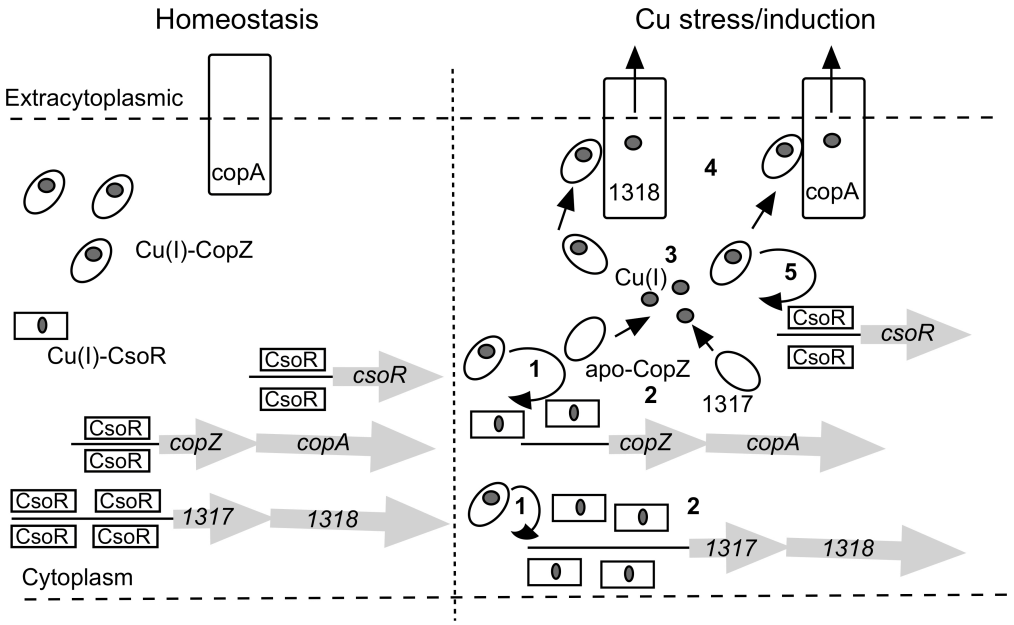


Figure 6.17: A model of the CsoR regulon in *S. lividans*. The *copZ/copA* arrows represent the highly expressed operons 1063/1064 and 3079/3080, with the Cu(I)-CopZ and CopA representative of these expressed genes. The *copZ/copA* couple 0895/0896 is not included in the model since the control mechanism is not known. However, it does contribute to a significant amount of CopZ and CopA both under homeostasis and upon Cu stress/induction. Under homeostasis conditions CopZ-3079/0895/1063 maintain buffering Cu(I) levels. Upon elevated Cu(I) levels CopZs facilitate transfer of Cu(I) to CsoR causing derepression of three CopZ/CopA efflux systems (3079/1063/0895) allowing continued buffering from these CopZs eventually being assisted by CopZ-1317 (steps 1-3). The Cu(I)-CopZs then remove Cu(I) by transferring it directly to the CopA pumps (step 4). Derepression of the *csoR* gene is only partial but occurs in the final step (step 5).

Chapter Seven

Summary

Streptomycetes are crucial for their ability to produce a surplus of secondary metabolites, to naturally secrete an array of important enzymes and as potential production hosts for the heterologous expression of biomolecules in bioreactors [5, 8, 10, 11, 35, 36]. These important features can be exploited further if a greater knowledge on the proteins, enzymes and pathways involved can be elucidated and characterised. The switch between substrate mycelium to aerial hyphae in the life cycle of these actinobacteria is concomitant with the secretion of enzymes and secondary metabolites. *Streptomyces lividans*, the strain studied here, displays a distinct dependence on the bioavailability copper (Cu) to fully initiate morphological development [5, 24]. The majority of this thesis has been focused on metalloenzymes initially hypothesised to be involved in the Cu-dependent morphological development in *S. lividans*. In particular a link between the Cu-trafficking *sco* operon and the *cslA/glxA* locus has been investigated. The characterisation of several proteins, GlxA, *Sli*LPMO10E and DtpA associated within these two genome environments has been undertaken. A summary of the findings in this thesis is presented below and coupled with *in vivo* data provided from collaborators, enables for a model of the Cu-dependent morphogenesis implicating glycan processing in the hyphal tips of *S. lividans* to be proposed.

The *glxA* gene is the distal gene in an operon with *cslA*, which encodes the CslA protein belonging to family 2 glycosyltransferases [62, 65]. *In vivo* work has shown that GlxA and CslA are co-localized at the hyphal tips in *S. lividans,* and are required for glycan synthesis or attachment [65, 72, 166]. Also *glxA* and *cslA* null-mutants drastically change morphology on both solid substrates and liquid cultures, illustrating an important role in the life cycle of this actinobacteria. GlxA was biochemically and structurally characterised (Chapter 2) and found to be membrane associated and to contain an active site consisting of a mononuclear Cu and a cross-linked Cys-Tyr co-factor. The domain organisation of the tertiary structure defines GlxA as a new structural member of the Cu oxidase family, with Cu coordination geometry similar to, but spectroscopically distinct from fungal galactose oxidase (Gox). Canonical Gox substrates were not readily turned over by GlxA with the natural substrate of GlxA still remaining unknown. In addition the Δ*glxA* mutant was shown to lose

the stimulation of development by addition of Cu, supporting a model whereby the enzymatic action of GlxA on the glycan is required for development and morphology. EPR spectroscopy revealed a protein radical residing on the Cys-Tyr cross-link upon oxidation of cupric GlxA using the mild oxidant $[Fe(CN)_6]^{3-}$. It was shown in Chapter 3 through mutagenesis studies that two residues, Cys^{121} and Trp^{288} modulate the distinct spectroscopic and radical behaviour of GlxA. The second coordination sphere residue, Trp^{288}, plays a major role in tuning the electronic properties with its removal abolishing the protein radical and perturbing the spectroscopic properties. Removal of the Cys^{121} residue, forming the Cys-Tyr cross-link prevents cross-link formation, changes the spectroscopic properties but remarkably does not abolish protein radical formation. X-ray crystallography coupled with western blotting analysis revealed the maturation of GlxA through the Cys-Tyr cross-link formation, which displays a distinct dependence upon Cu.

Despite neither a natural or 'good' substrate for GlxA being known, it is determined here to function as a Cu oxidase resulting in the production of the potentially harmful reactive oxygen species H_2O_2. The discovery that DtpA, part of the *sco* operon, is involved in GlxA maturation [171], could suggest a role for this dye-type peroxidase A (DtpA) acting to 'mop-up' the H_2O_2 produced by GlxA. DtpA has been shown here to function as a peroxidase with the ability to oxidise a range of substrates including classical peroxidase substrates and an AQ dye. Crucially, DtpA can function as a peroxidase in the presence of GlxA by utilizing the H_2O_2 produced by GlxA during turnover of glycolaldehyde (best substrate to have activity with GlxA - Chapter 2), thus signifying an *in vitro* relationship between these two enzymes. This relationship however does not explain its role in GlxA maturation.

Once the CslA/GlxA pair creates the β-(1-4)-glycan it is likely that during the life cycle it will eventually need degrading. Lytic polysaccharide monooxygenases (LPMOs) and endoglucanases act to depolymerise recalcitrant polysaccharides in nature [292]. *Sli*LPMO10E is present within the *cslA/glxA* operon in *S. lividans* and encodes for an AA10 LPMO. *Sli*LPMO10E was found to exclusively be active on chitin and has a tertiary structure and Cu coordination sphere consistent with activity confined to oxidising the C1 ring position of its glycan substrate.

A model for the Cu-dependent morphogenesis and glycan processing in the hyphal tips of *S. lividans* is presented in Figure 7.1. An emerging paradigm in Cu homeostasis and trafficking is that Cu is routed via the cytosol before being released into the periplasm or extracellular space for 'collection' by a designated chaperone [293, 429]. This may be the case in this system whereby the putative Cu transporter SLI_4212 could be involved in shuttling Cu ions over the membrane from the cytoplasm to the extracellular environment. Previous work has identified a Cu-trafficking pathway involving the extracellular chaperones ECuC and Sco (Figure 7.1) [38]. It may be that ECuC receives Cu(I) from SLI_4212 and is then trafficked to Sco, which delivers Cu to the Cu_A site of the aa_3-type cytochrome *c*

oxidase, and to the immature GlxA [171]. On accepting Cu, GlxA matures, through forming the Cys-Tyr crosslink and acts cooperatively with the cellulose synthase-like protein CslA in the formation and deposition of an extracellular chitin-like glycan required for aerial hyphae and pellet formation in solution cultures. A role for DtpA in absorbing the H_2O_2 produced by GlxA can be envisaged, but its role in the maturation of GlxA is more difficult to discern. It could be proposed that DtpA plays a role in regulating the redox state of the Cu bound to Sco. ECuC is highly specific for Cu(I) and from previous work a Cu(II) form could not be generated [38]. This led to the suggestion that Cu(I) is the oxidation state accepted by Sco [38]. It may be that GlxA requires Cu(II) for maturation and that DtpA acts to oxidise Cu(I)-Sco using an alternative H_2O_2 source to that generated by GlxA. Once the chitin-like glycan is synthesized by CslA and oxidized by GlxA it is proposed that degradation could involve the constitutively expressed *Sli*LPMO10E. This oxygenase can instigate chain breaks into the polysaccharide that would enable for the hydrolytic endoglucanase CslZ, to gain access and degrade the glycan into smaller saccharide subunits.

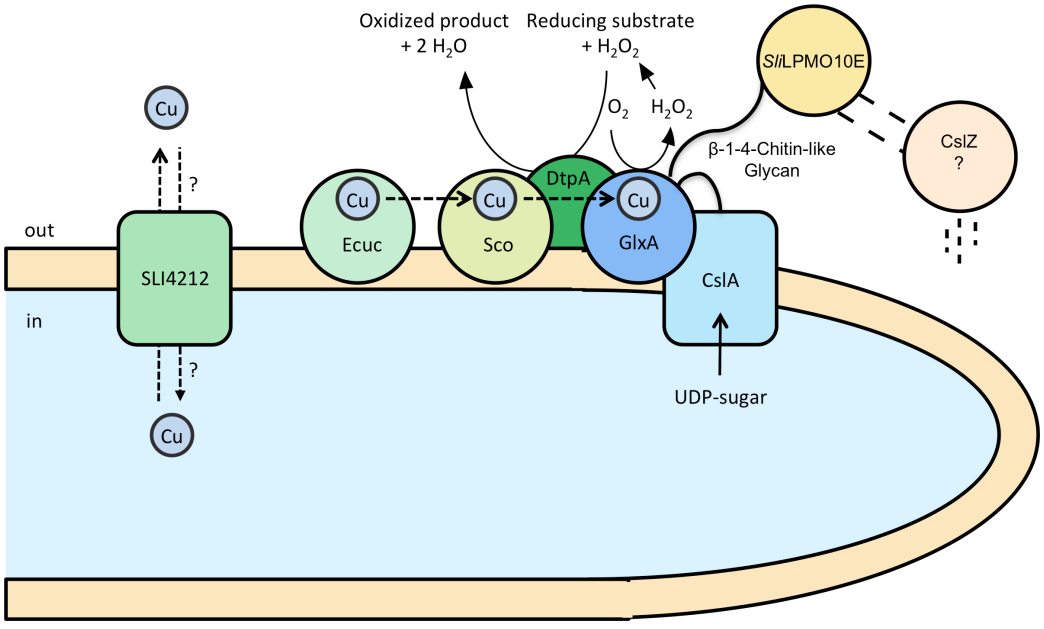


Figure 7.1: Proposed model for the Cu-dependent morphogenesis pathway in the hyphal tips of *S. lividans*. Inside is the cytosol (light blue) and outside is the extracellular environment. With the exception of *Sli*LPMO10E and CslZ all other proteins are membrane inserted (SLI4212 and CslA) or membrane tethered aiding therefore with co-localisation.

Whilst the majority of this thesis has been focused on extracellular Cu enzymes, the importance of Cu regulation within the cytoplasm of *S. lividans* has also been investigated. In the cytoplasmic environment of bacteria Cu homeostasis is tightly controlled to enable the delicate balance between minimizing toxicity of unbound Cu and ensuring the metabolic needs of the cell for Cu are met. In *S. lividans* the action of the metalloregulator CsoR (Cu sensitive operon repressor) is imperative for controlling a homeostatic Cu level [373]. Cytosolic Cu metallochaperones also play a key role in maintaining cellular Cu levels by tightly binding cuprous ions and trafficking it to target proteins or membrane transporters [343, 351-354, 430]. From genome analysis four *copZ* Cu-metallochaperone and five P-type ATPase (*copA*) encoding genes are present in *S. lividans* 1326 [4]. Under Cu stress, Cu(I) binds to apo-CsoR resulting in the transcriptional derepression of genes encoding for Cu-efflux systems involving these CopZ-like Cu-chaperones and their cognate CopA-like P-type ATPases. In this thesis two CopZ proteins (CopZ-1317 and CopZ-3079) under the transcriptional control of a CsoR (SLI4375) were investigated (Chapter 6) [431]. The results indicate that both CopZ-proteins have Cys residues in the Cu(I) binding $MX_1CX_2X_3C$ motif with acid/base properties that are modulated for a high cuprous ion affinity and favourable Cu(I)-exchange with a target protein. EMSAs and size-exclusion chromatography were used to show that Cu(I) is transferred in a unidirectional manner from the CopZ proteins to the CsoR. Using RNA-seq coupled with the mechanistic insights of Cu(I) transfer between CopZ and CsoR *in vitro*, a Cu-trafficking pathway for the CsoR regulon could be proposed that initially involves the buffering of cytosolic Cu by three of the four CopZ chaperones followed by transfer of Cu(I) to CsoR to illicit a transcriptional response as proposed in the model depicted in Figure 6.17 (Chapter 6) [431].

As pointed out above there is a suggestion that Cu acquisition of many cuproenzymes and cuproproteins in bacteria is achieved through routing the Cu through the cytosol [429]. Whilst the CsoR/CopZ/CopA system is highly conserved throughout the bacterial kingdom, it is doubtful that constitutively expressed CopZ proteins would provide enough buffering capacity to handle the storage of Cu required for the metabolic needs of the organism. There must therefore be another system operating, possibly similar to that recently reported for the handling of Cu in the methanotroph *Methylosinus trichosporium* and proposed for the metallation of particulate methane monooxygenase [432]. This Cu storage protein in *M. trichosporium* is periplasmic, but a homolog has been identified in *S. lividans* with no export signal sequence suggesting a cytosolic location.

References

1 Chater, K. F. (2006) Streptomyces inside-out: a new perspective on the bacteria that provide us with antibiotics. Philos Trans R Soc Lond B Biol Sci. **361**, 761-768

2 Embley, T. M. and Stackebrandt, E. (1994) The molecular phylogeny and systematics of the actinomycetes. Annu Rev Microbiol. **48**, 257-289

3 Hodgson, D. A. (2000) Primary metabolism and its control in streptomycetes: a most unusual group of bacteria. Adv Microb Physiol. **42**, 47-238

4 Cruz-Morales, P., Vijgenboom, E., Iruegas-Bocardo, F., Girard, G., Yanez-Guerra, L. A., Ramos-Aboites, H. E., Pernodet, J. L., Anne, J., van Wezel, G. P. and Barona-Gomez, F. (2013) The genome sequence of Streptomyces lividans 66 reveals a novel tRNA-dependent peptide biosynthetic system within a metal-related genomic island. Genome Biol Evol. **5**, 1165-1175

5 Worrall, J. A. R. and Vijgenboom, E. (2010) Copper mining in Streptomyces: enzymes, natural products and development. Natural product reports. **27**, 742-756

6 Mayfield, C. I., Williams, S. T., Ruddick, S. M. and Hatfield, H. L. (1972) Studies on the ecology of actinomycetes in soil. IV.Observation on the form and growth of streptomycetes in soil. Soil Biology and Biochemistry. **4**, 79-91

7 McCarthy, A. J. and Williams, S. T. (1992) Actinomycetes as agents of biodegradation in the environment--a review. Gene. **115**, 189-192

8 Hopwood, D. A. (2007) Streptomyces in Nature and Medicine: The Antibiotic Makers. Oxford University Press, New York

9 Waksman, S. A. and Schatz, A. (1945) Strain Specificity and Production of Antibiotic Substances: VI. Strain Variation and Production of Streptothricin by Actinomyces Lavendulae. Proc Natl Acad Sci U S A. **31**, 208-214

10 Chater, K. F., Biro, S., Lee, K. J., Palmer, T. and Schrempf, H. (2010) The complex extracellular biology of Streptomyces. FEMS Microbiol Rev. **34**, 171-198

11 Paradkar, A., Trefzer, A., Chakraburtty, R. and Stassi, D. (2003) Streptomyces genetics: a genomic perspective. Crit Rev Biotechnol. **23**, 1-27

12 Noda, S., Kawai, Y., Miyazaki, T., Tanaka, T. and Kondo, A. (2013) Creation of endoglucanase-secreting Streptomyces lividans for enzyme production using cellulose as the carbon source. Appl Microbiol Biotechnol. **97**, 5711-5720

13 Jing, D. (2010) Improving the simultaneous production of laccase and lignin peroxidase from Streptomyces lavendulae by medium optimization. Bioresour Technol. **101**, 7592-7597

14 Amore, A., Pepe, O., Ventorino, V., Birolo, L., Giangrande, C. and Faraco, V. (2012) Cloning and recombinant expression of a cellulase from the cellulolytic strain Streptomyces sp. G12 isolated from compost. Microb Cell Fact. **11**, 164

15 Sharma, M. (2014) Actinomycetes: Source, Identification, and Their Applications. Int.J.Curr.Microbiol.App.Sci. **3**, 801-832

16 Meena, B., Anburajan, L., Sathish, T., Vijaya Raghavan, R., Dharani, G., Vinithkumar, N. V. and Kirubagaran, R. (2015) L-Asparaginase from Streptomyces griseus NIOT-VKMA29: optimization of process variables using factorial designs and molecular characterization of L-asparaginase gene. Sci Rep. **5**, 12404

17 Dube, E., Shareck, F., Hurtubise, Y., Daneault, C. and Beauregard, M. (2008) Homologous cloning, expression, and characterisation of a laccase from Streptomyces coelicolor and enzymatic decolourisation of an indigo dye. Appl Microbiol Biotechnol. **79**, 597-603

18 Da Vinha, F. N., Gravina-Oliveira, M. P., Franco, M. N., Macrae, A., da Silva Bon, E. P., Nascimento, R. P. and Coelho, R. R. (2011) Cellulase production by Streptomyces viridobrunneus SCPE-09 using lignocellulosic biomass as inducer substrate. Appl Biochem Biotechnol. **164**, 256-267

19 Kar, S., Ray, R. C. and Mohapatra, U. B. (2008) Alpha-amylase production by Streptomyces erumpens MTCC 7317 in solid state fermentation using response surface methodology (RSM). Pol J Microbiol. **57**, 289-296

20 Gupta, R., Saxena, R. K., Chaturvedi, P. and Virdi, J. S. (1995) Chitinase production by Streptomyces viridificans: its potential in fungal cell wall lysis. J Appl Bacteriol. **78**, 378-383

21 Prianishnikova, N. I., Al-Nuri, M. A., Aslanian, R. R. and Egorov, N. S. (1984) [Natural variability of Streptomyces spheroides--a producer of extracellular proteolytic enzymes possessing fibrinolytic action]. Mikrobiologiia. **53**, 768-771

22 Noda, S., Miyazaki, T., Tanaka, T., Chiaki, O. and Kondo, A. (2013) High-level production of mature active-form Streptomyces mobaraensis transglutaminase via pro-transglutaminase processing using Streptomyces lividans as a host. Biochemical Engineering Journal. **74**, 76-80

23 Ueda, K., Tomaru, Y., Endoh, K. and Beppu, T. (1997) Stimulatory effect of copper on antibiotic production and morphological differentiation in Streptomyces tanashiensis. J Antibiot (Tokyo). **50**, 693-695

24 Keijser, B. J. F., van Wezel, G. P., Canters, G. W., Kieser, T. and Vijgenboom, E. (2000) The ram-dependence of Streptomyces lividans differentiation is bypassed by copper. Journal of molecular microbiology and biotechnology. **2**, 565-574

25 Fujimoto, M., Yamada, A., Kurosawa, J., Kawata, A., Beppu, T., Takano, H. and Ueda, K. (2012) Pleiotropic role of the Sco1/SenC family copper chaperone in the physiology of Streptomyces. Microbial biotechnology. **5**, 477-488

26 Flardh, K. and Buttner, M. J. (2009) Streptomyces morphogenetics: dissecting differentiation in a filamentous bacterium. Nat Rev Microbiol. **7**, 36-49

27 Shi, W. and Zusman, D. R. (1993) Fatal attraction. Nature. **366**, 414-415

28 Manteca, A., Fernandez, M. and Sanchez, J. (2005) A death round affecting a young compartmentalized mycelium precedes aerial mycelium dismantling in confluent surface cultures of Streptomyces antibioticus. Microbiology. **151**, 3689-3697

29 Manteca, A., Mader, U., Connolly, B. A. and Sanchez, J. (2006) A proteomic analysis of Streptomyces coelicolor programmed cell death. Proteomics. **6**, 6008-6022

30 Manteca, A., Claessen, D., Lopez-Iglesias, C. and Sanchez, J. (2007) Aerial hyphae in surface cultures of Streptomyces lividans and Streptomyces coelicolor originate from viable segments surviving an early programmed cell death event. FEMS Microbiol Lett. **274**, 118-125

31 Manteca, A. and Sanchez, J. (2009) Streptomyces development in colonies and soils. Appl Environ Microbiol. **75**, 2920-2924

32 Chater, K. F. (1993) Genetics of differentiation in Streptomyces. Annu Rev Microbiol. **47**, 685-713

33 Guijarro, J., Santamaria, R., Schauer, A. and Losick, R. (1988) Promoter determining the timing and spatial localization of transcription of a cloned Streptomyces coelicolor gene encoding a spore-associated polypeptide. J Bacteriol. **170**, 1895-1901

34 Elliot, M. A., Karoonuthaisiri, N., Huang, J., Bibb, M. J., Cohen, S. N., Kao, C. M. and Buttner, M. J. (2003) The chaplins: a family of hydrophobic cell-surface proteins involved in aerial mycelium formation in Streptomyces coelicolor. Genes Dev. **17**, 1727-1740

35 Demain, A. L. (2000) Small bugs, big business: the economic power of the microbe. Biotechnology advances. **18**, 499-514

36 van Wezel, G. P. and McDowall, K. J. (2011) The regulation of the secondary metabolism of Streptomyces: new links and experimental advances. Nat Prod Rep. **28**, 1311-1333

37 Blundell, K. L., Wilson, M. T., Svistunenko, D. A., Vijgenboom, E. and Worrall, J. A. (2013) Morphological development and cytochrome c oxidase activity in Streptomyces lividans are dependent on the action of a copper bound Sco protein. Open Biol. **3**, 120163

38 Blundell, K. L., Hough, M. A., Vijgenboom, E. and Worrall, J. A. (2014) Structural and mechanistic insights into an extracytoplasmic copper trafficking pathway in Streptomyces lividans. Biochem J. **459**, 525-538

39 Widdick, D. A., Dilks, K., Chandra, G., Bottrill, A., Naldrett, M., Pohlschroder, M. and Palmer, T. (2006) The twin-arginine translocation pathway is a major route of protein export in Streptomyces coelicolor. Proc Natl Acad Sci U S A. **103**, 17927-17932

40 Colpa, D. I., Fraaije, M. W. and van Bloois, E. (2014) DyP-type peroxidases: a promising and versatile class of enzymes. J Ind Microbiol Biotechnol. **41**, 1-7

41 Sugano, Y., Sasaki, K. and Shoda, M. (1999) cDNA cloning and genetic analysis of a novel decolorizing enzyme, peroxidase gene dyp from Geotrichum candidum Dec 1. J Biosci Bioeng. **87**, 411-417

42 Singh, R. and Eltis, L. D. (2015) The multihued palette of dye-decolorizing peroxidases. Arch Biochem Biophys. **574**, 56-65

43 Vrancken, K. and Anné, J. (2009) Secretory production of recombinant proteins by Streptomyces. Future Microbiol. **4**, 181-188

44 Anné, J., Maldonado, B., Van Impe, J., Van Mellaert, L. and Bernaerts, K. (2012) Recombinant protein production and streptomycetes. J Biotechnol. **158**, 159-167

45 Binnie, C., Cossar, J. D. and Stewart, D. I. (1997) Heterologous biopharmaceutical protein expression in Streptomyces. Trends Biotechnol. **15**, 315-320

46 Díaz, M., Ferreras, E., Moreno, R., Yepes, A., Berenguer, J. and Santamaria, R. (2008) High-level overproduction of Thermus enzymes in Streptomyces lividans. Appl Microbiol Biotechnol. **79**, 1001-1008

47 Nakashima, N., Mitani, Y. and Tamura, T. (2005) Actinomycetes as host cells for production of recombinant proteins. Microb Cell Fact. **4**, 7

48 Posta, K., Beki, E., Wilson, D. B., Kukolya, J. and Hornok, L. (2004) Cloning, characterization and phylogenetic relationships of cel5B, a new endoglucanase encoding gene from Thermobifida fusca. J Basic Microbiol. **44**, 383-399

49 Sianidis, G., Pozidis, C., Becker, F., Vrancken, K., Sjoeholm, C., Karamanou, S., Takamiya-Wik, M., van Mellaert, L., Schaefer, T., Anne, J. and Economou, A. (2006) Functional large-scale production of a novel Jonesia sp. xyloglucanase by heterologous secretion from Streptomyces lividans. J Biotechnol. **121**, 498-507

50 van Dissel, D., Claessen, D. and van Wezel, G. P. (2014) Morphogenesis of Streptomyces in submerged cultures. Adv Appl Microbiol. **89**, 1-45

51 Nielsen, J. (1996) Modelling the morphology of filamentous microorganisms. Trends Biotechnol. **14**, 438-443

52 Sevillano, L., Diaz, M. and Santamaria, R. I. (2013) Stable expression plasmids for Streptomyces based on a toxin-antitoxin system. Microb Cell Fact. **12**, 39

53 Cox, P. W., Paul, G. C. and Thomas, C. R. (1998) Image analysis of the morphology of filamentous micro-organisms. Microbiology. **144 (Pt 4)**, 817-827

54 Kim, Y. M. and Kim, J. H. (2004) Formation and dispersion of mycelial pellets of Streptomyces coelicolor A3(2). J Microbiol. **42**, 64-67

55 Whitaker, A. (1992) Actinomycetes in submerged culture. Appl Biochem Biotechnol. **32**, 23-35

56 Celler, K., Picioreanu, C., van Loosdrecht, M. C. and van Wezel, G. P. (2012) Structured morphological modeling as a framework for rational strain design of Streptomyces species. Antonie Van Leeuwenhoek. **102**, 409-423

57 van Dissel, D., Claessen, D., Roth, M. and van Wezel, G. P. (2015) A novel locus for mycelial aggregation forms a gateway to improved Streptomyces cell factories. Microb Cell Fact. **14**, 44

58 van Wezel, G. P., Krabben, P., Traag, B. A., Keijser, B. J., Kerste, R., Vijgenboom, E., Heijnen, J. J. and Kraal, B. (2006) Unlocking Streptomyces spp. for use as sustainable industrial production platforms by morphological engineering. Appl Environ Microbiol. **72**, 5283-5288

59 Traag, B. A. and van Wezel, G. P. (2008) The SsgA-like proteins in actinomycetes: small proteins up to a big task. Antonie Van Leeuwenhoek. **94**, 85-97

60 Noens, E. E., Mersinias, V., Willemse, J., Traag, B. A., Laing, E., Chater, K. F., Smith, C. P., Koerten, H. K. and van Wezel, G. P. (2007) Loss of the controlled localization of growth stage-specific cell-wall synthesis pleiotropically affects developmental gene expression in an ssgA mutant of Streptomyces coelicolor. Mol Microbiol. **64**, 1244-1259

61 van Wezel, G. P., van der Meulen, J., Kawamoto, S., Luiten, R. G., Koerten, H. K. and Kraal, B. (2000) ssgA is essential for sporulation of Streptomyces coelicolor A3(2) and affects hyphal development by stimulating septum formation. J Bacteriol. **182**, 5653-5662

62 Petrus, M. L. and Claessen, D. (2014) Pivotal roles for Streptomyces cell surface polymers in morphological differentiation, attachment and mycelial architecture. Antonie Van Leeuwenhoek. **106**, 127-139

63 Claessen, D., Rozen, D. E., Kuipers, O. P., Søgaard-Andersen, L. and van Wezel, G. P. (2014) Bacterial solutions to multicellularity: a tale of biofilms, filaments and fruiting bodies. Nature Reviews Microbiology. **12**, 115-124

64 van Veluw, G. J., Petrus, M. L., Gubbens, J., de Graaf, R., de Jong, I. P., van Wezel, G. P., Wosten, H. A. and Claessen, D. (2012) Analysis of two distinct mycelial populations in liquid-grown Streptomyces cultures using a flow cytometry-based proteomics approach. Appl Microbiol Biotechnol. **96**, 1301-1312

65 Xu, H., Chater, K. F., Deng, Z. and Tao, M. (2008) A cellulose synthase-like protein involved in hyphal tip growth and morphological differentiation in Streptomyces. Journal of bacteriology. **190**, 4971-4978

66 de Jong, W., Wosten, H. A., Dijkhuizen, L. and Claessen, D. (2009) Attachment of Streptomyces coelicolor is mediated by amyloidal fimbriae that are anchored to the cell surface via cellulose. Mol Microbiol. **73**, 1128-1140

67 Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. and Henrissat, B. (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. **42**, D490-495

68 Römling, U. (2002) Molecular biology of cellulose production in bacteria. Res Microbiol. **153**, 205-212

69 Ross, P., Mayer, R. and Benziman, M. (1991) Cellulose biosynthesis and function in bacteria. Microbiol Rev. **55**, 35-58

70 Avigad, G., Asensio, C., Amaral, D. and Horecker, B. L. (1961) Galacto-dialdose production with an enzyme from the mold Polyporus circinatus. Biochem Biophys Res Commun. **4**, 474-477

71 Avigad, G., Amaral, D., Asensio, C. and Horecker, B. (1962) The D-galactose oxidase of Polyporus circinatus. Journal of Biological Chemistry. **237**, 2736-2743

72 Liman, R., Facey, P. D., van Keulen, G., Dyson, P. J. and Del Sol, R. (2013) A laterally acquired galactose oxidase-like gene is required for aerial development during osmotic stress in Streptomyces coelicolor. PloS one. **8**, e54112

73 Aisaka, K. and Terada, O. (1982) Purification and properties of galactose oxidase from Gibberella fujikuroi. Agricultural and Biological Chemistry. **46**, 1191-1197

74 Amaral, D., Bernstein, L., Morse, D. and Horecker, B. L. (1963) Galactose oxidase of Polyporus circinatus: a copper enzyme. J Biol Chem. **238**, 2281-2284

75 Cooper, J. A., Smith, W., Bacila, M. and Medina, H. (1959) Galactose oxidase from Polyporus circinatus, Fr. J Biol Chem. **234**, 445-448

76 Ito, N., Phillips, S. E., Stevens, C., Ogel, Z. B., McPherson, M. J., Keen, J. N., Yadav, K. D. and Knowles, P. F. (1991) Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase.

77 Koroleva, O., Rabinovich, M., Buglova, T. and Iaropolov, A. (1982) [Properties of Fusarium graminearum galactose oxidase]. Prikladnaia biokhimiia i mikrobiologiia. **19**, 632-637

78 Kersten, P. J. (1990) Glyoxal oxidase of Phanerochaete chrysosporium: its characterization and activation by lignin peroxidase. Proc Natl Acad Sci U S A. **87**, 2936-2940

79 Kersten, P. J. and Kirk, T. K. (1987) Involvement of a new enzyme, glyoxal oxidase, in extracellular H2O2 production by Phanerochaete chrysosporium. J Bacteriol. **169**, 2195-2201

80 Whittaker, M. M., Kersten, P. J., Nakamura, N., Sanders-Loehr, J., Schweizer, E. S. and Whittaker, J. W. (1996) Glyoxal oxidase from Phanerochaete chrysosporium is a new radical-copper oxidase. Journal of Biological Chemistry. **271**, 681-687

81 Whittaker, M. M. and Whittaker, J. W. (1988) The active site of galactose oxidase. Journal of Biological Chemistry. **263**, 6074-6080

82 Whittaker, M. M. and Whittaker, J. W. (1990) A tyrosine-derived free radical in apogalactose oxidase. Journal of Biological Chemistry. **265**, 9610-9613

83 Whittaker, M. M. and Whittaker, J. W. (1993) Ligand interactions with galactose oxidase: mechanistic insights. Biophysical journal. **64**, 762-772

84 Ito, N., Phillips, S. E., Yadav, K. D. and Knowles, P. F. (1994) Crystal structure of a free radical enzyme, galactose oxidase. Journal of molecular biology. **238**, 704-814

85 Lee, Y.-K., Whittaker, M. M. and Whittaker, J. W. (2008) The Electronic Structure of the Cys-Tyr• Free Radical in Galactose Oxidase Determined by EPR Spectroscopy†. Biochemistry. **47**, 6637-6649

86 Whittaker, M. M. and Whittaker, J. W. (2003) Cu(I)-dependent biogenesis of the galactose oxidase redox cofactor. J Biol Chem. **278**, 22090-22101

87 Rogers, M. S., Hurtado-Guerrero, R., Firbank, S. J., Halcrow, M. A., Dooley, D. M., Phillips, S. E., Knowles, P. F. and McPherson, M. J. (2008) Cross-Link Formation of the Cysteine 228−Tyrosine 272 Catalytic Cofactor of Galactose Oxidase Does Not Require Dioxygen†‡. Biochemistry. **47**, 10428-10439

88 Vijgenboom, E., Woudt, L. P., Heinstra, P. W., Rietveld, K., van Haarlem, J., van Wezel, G. P., Shochat, S. and Bosch, L. (1994) Three tuf-like genes in the kirromycin producer Streptomyces ramocissimus. Microbiology. **140 (Pt 4)**, 983-998

89 Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods. **9**, 671-675

90 Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol. **305**, 567-580

91 Sonnhammer, E. L., von Heijne, G. and Krogh, A. (1998) A hidden Markov model for predicting transmembrane helices in protein sequences. Proc Int Conf Intell Syst Mol Biol. **6**, 175-182

92 von Heijne, G. (1992) Membrane protein structure prediction. Hydrophobicity analysis and the positive-inside rule. J Mol Biol. **225**, 487-494

93 Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M. R., Appel, R. D. and Bairoch, A. (2005) Protein identification and analysis tools on the ExPASy server. In The proteomics protocols handbook. pp. 571-607, Springer

94 Battye, T. G., Kontogiannis, L., Johnson, O., Powell, H. R. and Leslie, A. G. (2011) iMOSFLM: a new graphical interface for diffraction-image processing with MOSFLM. Acta crystallographica. Section D, Biological crystallography. **67**, 271-281

95 Evans, P. R. and Murshudov, G. N. (2013) How good are my data and what is the resolution? Acta crystallographica. Section D, Biological crystallography. **69**, 1204-1214

96 Stein, N. (2008) CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. **41**, 641-643

97 Perrakis, A., Morris, R. and Lamzin, V. S. (1999) Automated protein model building combined with iterative structure refinement. Nat Struct Biol. **6**, 458-463

98 Murshudov, G. N., Vagin, A. A. and Dodson, E. J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica. Section D, Biological crystallography. **53**, 240-255

99 McCoy, A. J., Grosse-Kunstleve, R. W., Adams, P. D., Winn, M. D., Storoni, L. C. and Read, R. J. (2007) Phaser crystallographic software. J. Appl. Crystallogr. **40**, 658-674

100 Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography. **60**, 2126-2132

101 Davis, I. W., Leaver-Fay, A., Chen, V. B., Block, J. N., Kapral, G. J., Wang, X., Murray, L. W., Arendall, W. B., 3rd, Snoeyink, J., Richardson, J. S. and Richardson, D. C. (2007) MolProbity: all-atom contacts and structure validation for proteins and nucleic acids. Nucleic Acids Res. **35**, W375-383

102 Yang, H., Guranovic, V., Dutta, S., Feng, Z., Berman, H. M. and Westbrook, J. D. (2004) Automated and accurate deposition of structures solved by X-ray diffraction to the Protein Data Bank. Acta Crystallogr D Biol Crystallogr. **60**, 1833-1839

103 Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C. and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr. **66**, 213-221

104 Terwilliger, T. C. (2003) Statistical density modification using local pattern matching. Acta Crystallogr D Biol Crystallogr. **59**, 1688-1701

105 Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2007) Clustal W and Clustal X version 2.0. Bioinformatics. **23**, 2947-2948

106 Kozlikova, B., Sebestova, E., Sustr, V., Brezovsky, J., Strnad, O., Daniel, L., Bednar, D., Pavelka, A., Manak, M., Bezdeka, M., Benes, P., Kotry, M., Gora, A., Damborsky, J. and Sochor, J. (2014) CAVER Analyst 1.0: graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. Bioinformatics. **30**, 2684-2685

107 van Keulen, G., Jonkers, H. M., Claessen, D., Dijkhuizen, L. and Wosten, H. A. (2003) Differentiation and anaerobiosis in standing liquid cultures of Streptomyces coelicolor. J Bacteriol. **185**, 1455-1458

108 Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M. and Söding, J. (2011) Fast, scalable generation of high‐quality protein multiple sequence alignments using Clustal Omega. Molecular systems biology. **7**

109 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. and Barton, G. J. (2009) Jalview Version 2--a multiple sequence alignment editor and analysis workbench. Bioinformatics. **25**, 1189-1191

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

111 Whittaker, M. M. and Whittaker, J. W. (2006) Streptomyces coelicolor oxidase (SCO2837p): A new free radical metalloenzyme secreted by Streptomyces coelicolor A3 (2). Archives of biochemistry and biophysics. **452**, 108-118

112 Stivala, A., Wybrow, M., Wirth, A., Whisstock, J. C. and Stuckey, P. J. (2011) Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics. **27**, 3315-3316

113 Drozdetskiy, A., Cole, C., Procter, J. and Barton, G. J. (2015) JPred4: a protein secondary structure prediction server. Nucleic Acids Res. **43**, W389-394

114 Antholine, W. E. (2005) Low frequency EPR of Cu(II) in proteins. Biological Magnetic Resonance. **23**, 417-454

115 Peisach, J. and Blumberg, W. E. (1974) Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins. Arch Biochem Biophys. **165**, 691-708

116 Bereman, R. D. and Kosman, D. J. (1977) Stereoelectronic properties of metalloenzymes. 5. Identification and assignment of ligand hyperfine splittings in the electron spin resonance spectrum of galactose oxidase. Journal of the American Chemical Society. **99**, 7322-7325

117 Cleveland, L., Coffman, R., Coon, P. and Davis, L. (1975) An investigation of the role of the copper in galactose oxidase. Biochemistry. **14**, 1108-1115

118 Berliner, L. J., Khramtsov, V., Fujii, H. and Clanton, T. L. (2001) Unique in vivo applications of spin traps. Free radical biology & medicine. **30**, 489-499

119 Kirima, K., Tsuchiya, K., Sei, H., Hasegawa, T., Shikishima, M., Motobayashi, Y., Morita, K., Yoshizumi, M. and Tamaki, T. (2003) Evaluation of systemic blood NO dynamics by EPR spectroscopy: HbNO as an endogenous index of NO. American journal of physiology. Heart and circulatory physiology. **285**, H589-596

120 Sillitoe, I., Cuff, A. L., Dessailly, B. H., Dawson, N. L., Furnham, N., Lee, D., Lees, J. G., Lewis, T. E., Studer, R. A., Rentzsch, R., Yeats, C., Thornton, J. M. and Orengo, C. A. (2013) New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures. Nucleic Acids Res. **41**, D490-498

121 McNicholas, S., Potterton, E., Wilson, K. S. and Noble, M. E. (2011) Presenting your structures: the CCP4mg molecular-graphics software. Acta Crystallogr D Biol Crystallogr. **67**, 386-394

122 Baron, A. J., Stevens, C., Wilmot, C., Seneviratne, K. D., Blakeley, V., Dooley, D. M., Phillips, S., Knowles, P. F. and McPherson, M. J. (1994) Structure and mechanism of galactose oxidase. The free radical site. Journal of Biological Chemistry. **269**, 25095-25105

123 Rogers, M. S., Tyler, E. M., Akyumani, N., Kurtis, C. R., Spooner, R. K., Deacon, S. E., Tamber, S., Firbank, S. J., Mahmoud, K., Knowles, P. F., Phillips, S. E., McPherson, M. J. and Dooley, D. M. (2007) The stacking tryptophan of galactose oxidase: a second-coordination sphere residue that has profound effects on tyrosyl radical behavior and enzyme catalysis. Biochemistry. **46**, 4606-4618

124 Chovancova, E., Pavelka, A., Benes, P., Strnad, O., Brezovsky, J., Kozlikova, B., Gora, A., Sustr, V., Klvana, M., Medek, P., Biedermannova, L., Sochor, J. and Damborsky, J. (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Comput Biol. **8**, e1002708

125 Whittaker, M. M., Kersten, P. J., Cullen, D. and Whittaker, J. W. (1999) Identification of catalytic residues in glyoxal oxidase by targeted mutagenesis. J Biol Chem. **274**, 36226-36232

126 Saysell, C. G., Barna, T., Borman, C. D., Baron, A. J., McPherson, M. J. and Sykes, A. G. (1997) Properties of the Trp290His variant of Fusarium NRRL 2903 galactose oxidase: interactions of the GOasesemi state with different buffers, its redox activity and ability to bind azide. JBIC Journal of Biological Inorganic Chemistry. **2**, 702-709

127 Sun, L., Bulter, T., Alcalde, M., Petrounia, I. P. and Arnold, F. H. (2002) Modification of Galactose Oxidase to Introduce Glucose 6‐Oxidase Activity. ChemBioChem. **3**, 781-783

128 Arends, I. W. C. E., Gamez, P. and Sheldon, R. A. (2006) Green oxidation of alcohols using biomimetic Cu complexes and Cu enzymes as catalysts. Advances in Inorganic Chemistry. **58**, 235-279

129 Hamilton, G. A., Dyrkacz, G. R. and Libby, R. D. (1976) The involvement of superoxide and trivalent copper in the galactose oxidase reaction. Adv Exp Med Biol. **74**, 489-504

130 Okeley, N. M. and van der Donk, W. A. (2000) Novel cofactors via post-translational modifications of enzyme active sites. Chem Biol. **7**, R159-171

131 Stubbe, J. A. (1989) Protein radical involvement in biological catalysis? Annu Rev Biochem. **58**, 257-285

132 Frey, P. A. (2001) Radical mechanisms of enzymatic catalysis. Annu Rev Biochem. **70**, 121-148

133 Stubbe, J. and van Der Donk, W. A. (1998) Protein Radicals in Enzyme Catalysis. Chem Rev. **98**, 705-762

134 Voet, D., Voet, J., G., and Pratt, C., W. (2006) Fundamentals of biochemistry : life at the molecular level Wiley, Hoboken, NJ

135 Mure, M. (2004) Tyrosine-derived quinone cofactors. Acc Chem Res. **37**, 131-139

136 Janes, S. M., Mu, D., Wemmer, D., Smith, A. J., Kaur, S., Maltby, D., Burlingame, A. L. and Klinman, J. P. (1990) A new redox cofactor in eukaryotic enzymes: 6-hydroxydopa at the active site of bovine serum amine oxidase. Science. **248**, 981-987

137 Ruggiero, C. E., Smith, J. A., Tanizawa, K. and Dooley, D. M. (1997) Mechanistic studies of topa quinone biogenesis in phenylethylamine oxidase. Biochemistry. **36**, 1953-1959

138 Kim, M., Okajima, T., Kishishita, S., Yoshimura, M., Kawamori, A., Tanizawa, K. and Yamaguchi, H. (2002) X-ray snapshots of quinone cofactor biogenesis in bacterial copper amine oxidase. Nat Struct Biol. **9**, 591-596

139 Samuels, N. M. and Klinman, J. P. (2006) Investigation of Cu(I)-dependent 2,4,5-trihydroxyphenylalanine quinone biogenesis in Hansenula polymorpha amine oxidase. J Biol Chem. **281**, 21114-21118

140 Wang, S. X., Mure, M., Medzihradszky, K. F., Burlingame, A. L., Brown, D. E., Dooley, D. M., Smith, A. J., Kagan, H. M. and Klinman, J. P. (1996) A crosslinked cofactor in lysyl oxidase: redox function for amino acid side chains. Science. **273**, 1078-1084

141 Yamada, Y., Fujiwara, T., Sato, T., Igarashi, N. and Tanaka, N. (2002) The 2.0 A crystal structure of catalase-peroxidase from Haloarcula marismortui. Nat Struct Biol. **9**, 691-695

142 Ghiladi, R. A., Knudsen, G. M., Medzihradszky, K. F. and de Montellano, P. R. O. (2005) The Met-Tyr-Trp Cross-link in Mycobacterium tuberculosis Catalase-peroxidase (KatG) Autocatalytic formation and effect on enzyme catalysis and spectroscopic properties. Journal of Biological Chemistry. **280**, 22651-22663

143 Smulevich, G., Jakopitsch, C., Droghetti, E. and Obinger, C. (2006) Probing the structure and bifunctionality of catalase-peroxidase (KatG). Journal of inorganic biochemistry. **100**, 568-585

144 Ostermeier, C., Harrenga, A., Ermler, U. and Michel, H. (1997) Structure at 2.7 A resolution of the Paracoccus denitrificans two-subunit cytochrome c oxidase complexed with an antibody FV fragment. Proc Natl Acad Sci U S A. **94**, 10547-10553

145 Yoshikawa, S., Shinzawa-Itoh, K., Nakashima, R., Yaono, R., Yamashita, E., Inoue, N., Yao, M., Fei, M. J., Libeu, C. P., Mizushima, T., Yamaguchi, H., Tomizaki, T. and Tsukihara, T. (1998) Redox-coupled crystal structural changes in bovine heart cytochrome c oxidase. Science. **280**, 1723-1729

146 Proshlyakov, D. A., Pressler, M. A., DeMaso, C., Leykam, J. F., DeWitt, D. L. and Babcock, G. T. (2000) Oxygen activation and reduction in respiration: involvement of redox-active tyrosine 244. Science. **290**, 1588-1591

147 Schnell, R., Sandalova, T., Hellman, U., Lindqvist, Y. and Schneider, G. (2005) Siroheme- and [Fe4-S4]-dependent NirA from Mycobacterium tuberculosis is a sulfite reductase with a covalent Cys-Tyr bond in the active site. J Biol Chem. **280**, 27319-27328

148 McCoy, J. G., Bailey, L. J., Bitto, E., Bingman, C. A., Aceti, D. J., Fox, B. G. and Phillips, G. N., Jr. (2006) Structure and mechanism of mouse cysteine dioxygenase. Proc Natl Acad Sci U S A. **103**, 3084-3089

149 Simmons, C. R., Liu, Q., Huang, Q., Hao, Q., Begley, T. P., Karplus, P. A. and Stipanuk, M. H. (2006) Crystal structure of mammalian cysteine dioxygenase. A novel mononuclear iron center for cysteine thiol oxidation. J Biol Chem. **281**, 18723-18733

150 Rogers, M. S., Baron, A. J., McPherson, M. J., Knowles, P. F. and Dooley, D. M. (2000) Galactose oxidase pro-sequence cleavage and cofactor assembly are self-processing reactions. Journal of the American Chemical Society. **122**, 990-991

151 Johnson, J. M., Halsall, H. B. and Heineman, W. R. (1985) Redox activation of galactose oxidase: thin-layer electrochemical study. Biochemistry. **24**, 1579-1585

152 Kim, S. and Barry, B. A. (1998) The protein environment surrounding tyrosyl radicals D. and Z. in photosystem II: a difference Fourier-transform infrared spectroscopic study. Biophys J. **74**, 2588-2600

153 Narvaez, A. J., Kalman, L., LoBrutto, R., Allen, J. P. and Williams, J. C. (2002) Influence of the protein environment on the properties of a tyrosyl radical in reaction centers from Rhodobacter sphaeroides. Biochemistry. **41**, 15253-15258

154 Barrows, T. P., Bhaskar, B. and Poulos, T. L. (2004) Electrostatic control of the tryptophan radical in cytochrome c peroxidase. Biochemistry. **43**, 8826-8834

155 Pérez-Boada, M., Ruiz-Duenas, F. J., Pogni, R., Basosi, R., Choinowski, T., Martinez, M. J., Piontek, K. and Martinez, A. T. (2005) Versatile peroxidase oxidation of high redox potential aromatic compounds: site-directed mutagenesis, spectroscopic and crystallographic investigation of three long-range electron transfer pathways. J Mol Biol. **354**, 385-402

156 Gerez, C., Elleingand, E., Kauppi, B., Eklund, H. and Fontecave, M. (1997) Reactivity of the tyrosyl radical of Escherichia coli ribonucleotide reductase -- control by the protein. Eur J Biochem. **249**, 401-407

157 Ormo, M., Regnstrom, K., Wang, Z., Que, L., Jr., Sahlin, M. and Sjoberg, B. M. (1995) Residues important for radical stability in ribonucleotide reductase from Escherichia coli. J Biol Chem. **270**, 6570-6576

158 Xie, J., Yikilmaz, E., Miller, A. F. and Brunold, T. C. (2002) Second-sphere contributions to substrate-analogue binding in iron(III) superoxide dismutase. J Am Chem Soc. **124**, 3769-3774

159 Zoidakis, J., Sam, M., Volner, A., Han, A., Vu, K. and Abu-Omar, M. M. (2004) Role of the second coordination sphere residue tyrosine 179 in substrate affinity and catalytic activity of phenylalanine hydroxylase. J Biol Inorg Chem. **9**, 289-296

160 Miranda, F. F., Kolberg, M., Andersson, K. K., Geraldes, C. F. and Martinez, A. (2005) The active site residue tyrosine 325 influences iron binding and coupling efficiency in human phenylalanine hydroxylase. J Inorg Biochem. **99**, 1320-1328

161 Chen, K., Tilley, G. J., Sridhar, V., Prasad, G. S., Stout, C. D., Armstrong, F. A. and Burgess, B. K. (1999) Alteration of the reduction potential of the [4Fe-4S](2+/+) cluster of Azotobacter vinelandii ferredoxin I. J Biol Chem. **274**, 36479-36487

162 Tomchick, D. R., Phan, P., Cymborowski, M., Minor, W. and Holman, T. R. (2001) Structural and functional characterization of second-coordination sphere mutants of soybean lipoxygenase-1. Biochemistry. **40**, 7509-7517

163 Rogers , M. S., Knowles , P. F., Baron , A. J., McPherson , M. J. and Dooley , D. M. (1998) Characterization of the active site of galactose oxidase and its active site mutational variants Y495F/H/K and W290H by circular dichroism spectroscopy. Inorganica Chimica Acta, 275-276/175-181

164 Reynolds, M. P., Baron, A. J., Wilmot, C. M., Vinecombe, E., Stevens, C., Phillips, S. E., Knowles, P. F. and McPherson, M. J. (1997) Structure and mechanism of galactose oxidase: catalytic role of tyrosine 495. JBIC Journal of Biological Inorganic Chemistry. **2**, 327-335

165 Studier, F. W. (2005) Protein production by auto-induction in high-density shaking cultures. Protein expression and purification. **41**, 207-234

166 Chaplin, A. K., Petrus, M. L., Mangiameli, G., Hough, M. A., Svistunenko, D. A., Nicholls, P., Claessen, D., Vijgenboom, E. and Worrall, J. A. (2015) 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. **469**, 433-444

167 Svistunenko, D. A. and Cooper, C. E. (2004) A new method of identifying the site of tyrosyl radicals in proteins. Biophys J. **87**, 582-595

168 Potterton, E., Briggs, P., Turkenburg, M. and Dodson, E. (2003) A graphical user interface to the CCP4 program suite. Acta Crystallogr D Biol Crystallogr. **59**, 1131-1137

169 Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A. and Wilson, K. S. (2011) Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr. **67**, 235-242

170 Rogers, M. S. and Dooley, D. M. (2001) Posttranslationally modified tyrosines from galactose oxidase and cytochrome c oxidase. Adv Protein Chem. **58**, 387-436

171 Petrus, M. L., Vijgenboom, E., Chaplin, A. K., Worrall, J. A., van Wezel, G. P. and Claessen, D. (2016) The DyP-type peroxidase DtpA is a Tat-substrate required for GlxA maturation and morphogenesis in Streptomyces. Open Biol. **6**

172 Kieser, T. and Hopwood, D. A. (1991) Genetic manipulation of Streptomyces: integrating vectors and gene replacement. Methods Enzymol. **204**, 430-458

173 Banci, L. (1997) Structural properties of peroxidases. J Biotechnol. **53**, 253-263

174 van Bloois, E., Torres Pazmino, D. E., Winter, R. T. and Fraaije, M. W. (2010) A robust and extracellular heme-containing peroxidase from Thermobifida fusca as prototype of a bacterial peroxidase superfamily. Appl Microbiol Biotechnol. **86**, 1419-1430

175 Welinder, K. G., Mauro, J. M. and Norskov-Lauritsen, L. (1992) Structure of plant and fungal peroxidases. Biochem Soc Trans. **20**, 337-340

176 Loughran, N. B., O'Connor, B., O'Fagain, C. and O'Connell, M. J. (2008) The phylogeny of the mammalian heme peroxidases and the evolution of their diverse functions. BMC Evol Biol. **8**, 101

177 Finn, R. D., Tate, J., Mistry, J., Coggill, P. C., Sammut, S. J., Hotz, H. R., Ceric, G., Forslund, K., Eddy, S. R., Sonnhammer, E. L. and Bateman, A. (2008) The Pfam protein families database. Nucleic Acids Res. **36**, D281-288

178 Passardi, F., Theiler, G., Zamocky, M., Cosio, C., Rouhier, N., Teixera, F., Margis-Pinheiro, M., Ioannidis, V., Penel, C., Falquet, L. and Dunand, C. (2007) PeroxiBase: the peroxidase database. Phytochemistry. **68**, 1605-1611

179 Kim, S. J., Ishikawa, K., Hirai, M. and Shoda, M. (1995) Characteristics of a newly isolated fungus, Geotrichum candidum Dec 1, which decolorizes various dyes. J. Ferment. Bioeng. **79**, 601–607

180 Kim, S. J. and Shoda, M. (1999) Purification and characterization of a novel peroxidase from Geotrichum candidum dec 1 involved in decolorization of dyes. Appl Environ Microbiol. **65**, 1029-1035

181 Zamocky, M. and Obinger, C. (2010) Molecular phylogeny of heme peroxidases. In: Biocatalysis Based on Heme Peroxidases. Springer Verlag, Berlin, Germany

182 Scheibner, M., Hulsdau, B., Zelena, K., Nimtz, M., de Boer, L., Berger, R. G. and Zorn, H. (2008) Novel peroxidases of Marasmius scorodonius degrade beta-carotene. Appl Microbiol Biotechnol. **77**, 1241-1250

183 Sturm, A., Schierhorn, A., Lindenstrauss, U., Lilie, H. and Bruser, T. (2006) YcdB from Escherichia coli reveals a novel class of Tat-dependently translocated hemoproteins. J Biol Chem. **281**, 13972-13978

184 Ogola, H. J., Kamiike, T., Hashimoto, N., Ashida, H., Ishikawa, T., Shibata, H. and Sawa, Y. (2009) Molecular characterization of a novel peroxidase from the cyanobacterium Anabaena sp. strain PCC 7120. Appl Environ Microbiol. **75**, 7509-7518

185 Koua, D., Cerutti, L., Falquet, L., Sigrist, C. J., Theiler, G., Hulo, N. and Dunand, C. (2009) PeroxiBase: a database with new tools for peroxidase family classification. Nucleic Acids Res. **37**, D261-266

186 Sugano, Y., Muramatsu, R., Ichiyanagi, A., Sato, T. and Shoda, M. (2007) DyP, a unique dye-decolorizing peroxidase, represents a novel heme peroxidase family: ASP171 replaces the distal histidine of classical peroxidases. J Biol Chem. **282**, 36652-36658

187 Fawal, N., Li, Q., Savelli, B., Brette, M., Passaia, G., Fabre, M., Mathe, C. and Dunand, C. (2013) PeroxiBase: a database for large-scale evolutionary analysis of peroxidases. Nucleic Acids Res. **41**, D441-444

188 Jongbloed, J. D., Grieger, U., Antelmann, H., Hecker, M., Nijland, R., Bron, S. and van Dijl, J. M. (2004) Two minimal Tat translocases in Bacillus. Mol Microbiol. **54**, 1319-1325

189 Chen, C., Shrestha, R., Jia, K., Gao, P. F., Geisbrecht, B. V., Bossmann, S. H., Shi, J. and Li, P. (2015) Characterization of Dye-decolorizing Peroxidase (DyP) from Thermomonospora curvata Reveals Unique Catalytic Properties of A-type DyPs. J Biol Chem. **290**, 23447-23463

190 Ahmad, M., Roberts, J. N., Hardiman, E. M., Singh, R., Eltis, L. D. and Bugg, T. D. (2011) Identification of DypB from Rhodococcus jostii RHA1 as a lignin peroxidase. Biochemistry. **50**, 5096-5107

191 Santos, A., Mendes, S., Brissos, V. and Martins, L. O. (2014) New dye-decolorizing peroxidases from Bacillus subtilis and Pseudomonas putida MET94: towards biotechnological applications. Appl Microbiol Biotechnol. **98**, 2053-2065

192 Uchida, T., Sasaki, M., Tanaka, Y. and Ishimori, K. (2015) A Dye-Decolorizing Peroxidase from Vibrio cholerae. Biochemistry

193 Li, X., Beeson, W. T. t., Phillips, C. M., Marletta, M. A. and Cate, J. H. (2012) Structural basis for substrate targeting and catalysis by fungal polysaccharide monooxygenases. Structure. **20**, 1051-1061

194 Zubieta, C., Joseph, R., Krishna, S. S., McMullan, D., Kapoor, M., Axelrod, H. L., Miller, M. D., Abdubek, P., Acosta, C., Astakhova, T., Carlton, D., Chiu, H. J., Clayton, T., Deller, M. C., Duan, L., Elias, Y., Elsliger, M. A., Feuerhelm, J., Grzechnik, S. K., Hale, J., Han, G. W., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kozbial, P., Kumar, A., Marciano, D., Morse, A. T., Murphy, K. D., Nigoghossian, E., Okach, L., Oommachen, S., Reyes, R., Rife, C. L., Schimmel, P., Trout, C. V., van den Bedem, H., Weekes, D., White, A., Xu, Q., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. and Wilson, I. A. (2007) Identification and structural characterization of heme binding in a novel dye-decolorizing peroxidase, TyrA. Proteins. **69**, 234-243

195 Contreras, H., Joens, M. S., McMath, L. M., Le, V. P., Tullius, M. V., Kimmey, J. M., Bionghi, N., Horwitz, M. A., Fitzpatrick, J. A. and Goulding, C. W. (2014) Characterization of a Mycobacterium tuberculosis nanocompartment and its potential cargo proteins. J Biol Chem. **289**, 18279-18289

196 Brown, M. E., Barros, T. and Chang, M. C. (2012) Identification and characterization of a multifunctional dye peroxidase from a lignin-reactive bacterium. ACS Chem Biol. **7**, 2074-2081

197 Strittmatter, E., Liers, C., Ullrich, R., Wachter, S., Hofrichter, M., Plattner, D. A. and Piontek, K. (2013) First crystal structure of a fungal high-redox potential dye-decolorizing peroxidase: substrate interaction sites and long-range electron transfer. J Biol Chem. **288**, 4095-4102

198 Salvachúa, D., Prieto, A., Martinez, A. T. and Martinez, M. J. (2013) Characterization of a novel dye-decolorizing peroxidase (DyP)-type enzyme from Irpex lacteus and its application in enzymatic hydrolysis of wheat straw. Appl Environ Microbiol. **79**, 4316-4324

199 Liers, C., Pecyna, M. J., Kellner, H., Worrich, A., Zorn, H., Steffen, K. T., Hofrichter, M. and Ullrich, R. (2013) Substrate oxidation by dye-decolorizing peroxidases (DyPs) from wood- and litter-degrading agaricomycetes compared to other fungal and plant heme-peroxidases. Appl Microbiol Biotechnol. **97**, 5839-5849

200 Johjima, T., Ohkuma, M. and Kudo, T. (2003) Isolation and cDNA cloning of novel hydrogen peroxide-dependent phenol oxidase from the basidiomycete Termitomyces albuminosus. Appl Microbiol Biotechnol. **61**, 220-225

201 Sato, T., Hara, S., Matsui, T., Sazaki, G., Saijo, S., Ganbe, T., Tanaka, N., Sugano, Y. and Shoda, M. (2004) A unique dye-decolorizing peroxidase, DyP, from Thanatephorus cucumeris Dec 1: heterologous expression, crystallization and preliminary X-ray analysis. Acta Crystallogr D Biol Crystallogr. **60**, 149-152

202 Yoshida, T., Tsuge, H., Hisabori, T. and Sugano, Y. (2012) Crystal structures of dye-decolorizing peroxidase with ascorbic acid and 2,6-dimethoxyphenol. FEBS Lett. **586**, 4351-4356

203 Liu, X., Du, Q., Wang, Z., Zhu, D., Huang, Y., Li, N., Wei, T., Xu, S. and Gu, L. (2011) Crystal structure and biochemical features of EfeB/YcdB from Escherichia coli O157: ASP235 plays divergent roles in different enzyme-catalyzed processes. J Biol Chem. **286**, 14922-14931

204 Roberts, J. N., Singh, R., Grigg, J. C., Murphy, M. E., Bugg, T. D. and Eltis, L. D. (2011) Characterization of dye-decolorizing peroxidases from Rhodococcus jostii RHA1. Biochemistry. **50**, 5108-5119

205 Zubieta, C., Krishna, S. S., Kapoor, M., Kozbial, P., McMullan, D., Axelrod, H. L., Miller, M. D., Abdubek, P., Ambing, E., Astakhova, T., Carlton, D., Chiu, H. J., Clayton, T., Deller, M. C., Duan, L., Elsliger, M. A., Feuerhelm, J., Grzechnik, S. K., Hale, J., Hampton, E., Han, G. W., Jaroszewski, L., Jin, K. K., Klock, H. E., Knuth, M. W., Kumar, A., Marciano, D., Morse, A. T., Nigoghossian, E., Okach, L., Oommachen, S., Reyes, R., Rife, C. L., Schimmel, P., van den Bedem, H., Weekes, D., White, A., Xu, Q., Hodgson, K. O., Wooley, J., Deacon, A. M., Godzik, A., Lesley, S. A. and Wilson, I. A. (2007) Crystal structures of two novel dye-decolorizing peroxidases reveal a beta-barrel fold with a conserved heme-binding motif. Proteins. **69**, 223-233

206 Gajhede, M., Schuller, D. J., Henriksen, A., Smith, A. T. and Poulos, T. L. (1997) Crystal structure of horseradish peroxidase C at 2.15 A resolution. Nat Struct Biol. **4**, 1032-1038

207 Goblirsch, B., Kurker, R. C., Streit, B. R., Wilmot, C. M. and DuBois, J. L. (2011) Chlorite dismutases, DyPs, and EfeB: 3 microbial heme enzyme families comprise the CDE structural superfamily. J Mol Biol. **408**, 379-398

208 Linde, D., Ruiz-Duenas, F. J., Fernandez-Fueyo, E., Guallar, V., Hammel, K. E., Pogni, R. and Martinez, A. T. (2015) Basidiomycete DyPs: Genomic diversity, structural-functional aspects, reaction mechanism and environmental significance. Arch Biochem Biophys. **574**, 66-74

209 Singh, R., Grigg, J. C., Armstrong, Z., Murphy, M. E. and Eltis, L. D. (2012) Distal heme pocket residues of B-type dye-decolorizing peroxidase: arginine but not aspartate is essential for peroxidase activity. J Biol Chem. **287**, 10623-10630

210 Poulos, T. L. and Kraut, J. (1980) The stereochemistry of peroxidase catalysis. J Biol Chem. **255**, 8199-8205

211 Sugano, Y. (2009) DyP-type peroxidases comprise a novel heme peroxidase family. Cell Mol Life Sci. **66**, 1387-1403

212 Mendes, S., Catarino, T., Silveira, C., Todorovica, S. and Martinsa, L., O (2015) The catalytic mechanism of A-type dye-decolourising peroxidase BsDyP: neither aspartate nor arginine is individually essential for peroxidase activity. Catal. Sci. Technol

213 Stolz, A. (2001) Basic and applied aspects in the microbial degradation of azo dyes. Appl Microbiol Biotechnol. **56**, 69-80

214 McMullan, G., Meehan, C., Conneely, A., Kirby, N., Robinson, T., Nigam, P., Banat, I. M., Marchant, R. and Smyth, W. F. (2001) Microbial decolourisation and degradation of textile dyes. Appl Microbiol Biotechnol. **56**, 81-87

215 Robinson, T., McMullan, G., Marchant, R. and Nigam, P. (2001) Remediation of dyes in textile effluent: a critical review on current treatment technologies with a proposed alternative. Bioresour Technol. **77**, 247-255

216 Zille, A., Gornacka, B., Rehorek, A. and Cavaco-Paulo, A. (2005) Degradation of azo dyes by Trametes villosa laccase over long periods of oxidative conditions. Appl Environ Microbiol. **71**, 6711-6718

217 Abadulla, E., Tzanov, T., Costa, S., Robra, K. H., Cavaco-Paulo, A. and Gubitz, G. M. (2000) Decolorization and detoxification of textile dyes with a laccase from Trametes hirsuta. Appl Environ Microbiol. **66**, 3357-3362

218 Chivukula, M. and Renganathan, V. (1995) Phenolic Azo Dye Oxidation by Laccase from Pyricularia oryzae. Appl Environ Microbiol. **61**, 4374-4377

219 Bafana, A., Chakrabarti, T., Muthal, P. and Kanade, G. (2009) Detoxification of benzidine-based azo dye by E. gallinarum: time-course study. Ecotoxicol Environ Saf. **72**, 960-964

220 Dey, S., Maiti, T. K. and Bhattacharyya, B. C. (1994) Production of some extracellular enzymes by a lignin peroxidase-producing brown rot fungus, Polyporus ostreiformis, and its comparative abilities for lignin degradation and dye decolorization. Appl Environ Microbiol. **60**, 4216-4218

221 Ollikka, P., Alhonmaki, K., Leppanen, V. M., Glumoff, T., Raijola, T. and Suominen, I. (1993) Decolorization of Azo, Triphenyl Methane, Heterocyclic, and Polymeric Dyes by Lignin Peroxidase Isoenzymes from Phanerochaete chrysosporium. Appl Environ Microbiol. **59**, 4010-4016

222 Verma, P. and Madamwar, D. (2002) Production of ligninolytic enzymes for dye decolorization by cocultivation of white-rot fungi Pleurotus ostreatus and phanerochaete chrysosporium under solid-state fermentation. Appl Biochem Biotechnol. **102-103**, 109-118

223 Christian, V. V., Shrivastava, R., Novotny, C. and Vyas, B. R. (2003) Decolorization of sulfonphthalein dyes by manganese peroxidase activity of the white-rot fungus Phanerochaete chrysosporium. Folia Microbiol (Praha). **48**, 771-774

224 Yang, Q., Yang, M., Pritsch, K., Yediler, A., Hagn, A., Schloter, M. and Kettrup, A. (2003) Decolorization of synthetic dyes and production of manganese-dependent peroxidase by new fungal isolates. Biotechnol Lett. **25**, 709-713

225 Kamitsuji, H., Honda, Y., Watanabe, T. and Kuwahara, M. (2004) Production and induction of manganese peroxidase isozymes in a white-rot fungus Pleurotus ostreatus. Appl Microbiol Biotechnol. **65**, 287-294

226 Kawai, S., Umezawa, T. and Higuchi, T. (1988) Degradation mechanisms of phenolic beta-1 lignin substructure model compounds by laccase of Coriolus versicolor. Arch Biochem Biophys. **262**, 99-110

227 Wariishi, H., Valli, K., Renganathan, V. and Gold, M. H. (1989) Thiol-mediated oxidation of nonphenolic lignin model compounds by manganese peroxidase of Phanerochaete chrysosporium. J Biol Chem. **264**, 14185-14191

228 Liers, C., Bobeth, C., Pecyna, M., Ullrich, R. and Hofrichter, M. (2010) DyP-like peroxidases of the jelly fungus Auricularia auricula-judae oxidize nonphenolic lignin model compounds and high-redox potential dyes. Appl Microbiol Biotechnol. **85**, 1869-1879

229 Sugano, Y., Matsushima, Y., Tsuchiya, K., Aoki, H., Hirai, M. and Shoda, M. (2009) Degradation pathway of an anthraquinone dye catalyzed by a unique peroxidase DyP from Thanatephorus cucumeris Dec 1. Biodegradation. **20**, 433-440

230 Kaur, A., Van, P. T., Busch, C. R., Robinson, C. K., Pan, M., Pang, W. L., Reiss, D. J., DiRuggiero, J. and Baliga, N. S. (2010) Coordination of frontline defense mechanisms under severe oxidative stress. Mol Syst Biol. **6**, 393

231 Kong, L., Guo, D., Zhou, S., Yu, X., Hou, G., Li, R. and Zhao, B. (2010) Cloning and expression of a toxin gene from Pseudomonas fluorescens GcM5-1A. Arch Microbiol. **192**, 585-593

232 Letoffe, S., Heuck, G., Delepelaire, P., Lange, N. and Wandersman, C. (2009) Bacteria capture iron from heme by keeping tetrapyrrol skeleton intact. Proc Natl Acad Sci U S A. **106**, 11719-11724

233 Sutter, M., Boehringer, D., Gutmann, S., Gunther, S., Prangishvili, D., Loessner, M. J., Stetter, K. O., Weber-Ban, E. and Ban, N. (2008) Structural basis of enzyme encapsulation into a bacterial nanocompartment. Nat Struct Mol Biol. **15**, 939-947

234 Beers, R. F., Jr. and Sizer, I. W. (1952) A spectrophotometric method for measuring the breakdown of hydrogen peroxide by catalase. J Biol Chem. **195**, 133-140

235 Whitmore, L. and Wallace, B. A. (2004) DICHROWEB, an online server for protein secondary structure analyses from circular dichroism spectroscopic data. Nucleic Acids Res. **32**, W668-673

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

237 Compton, L. A. and Johnson Jr, W. C. (1986) Analysis of protein circular dichroism spectra for secondary structure using a simple matrix multiplication. Analytical biochemistry. **155**, 155-167

238 Manavalan, P. and Johnson Jr, W. C. (1987) Variable selection method improves the prediction of protein secondary structure from circular dichroism spectra. Analytical biochemistry. **167**, 76-85

239 Sreerama, N., Venyaminov, S. Y. and Woody, R. W. (2000) Estimation of protein secondary structure from circular dichroism spectra: inclusion of denatured proteins with native proteins in the analysis. Anal Biochem. **287**, 243-251

240 Provencher, S. W. and Glockner, J. (1981) Estimation of globular protein secondary structure from circular dichroism. Biochemistry. **20**, 33-37

241 van Stokkum, I. H. (1990) Modelling the response of auditory midbrain neurons in the grassfrog to temporally structured monaural stimuli. Hear Res. **43**, 231-250

242 Lees, J. G., Miles, A. J., Wien, F. and Wallace, B. (2006) A reference database for circular dichroism spectroscopy covering fold and secondary structure space. Bioinformatics. **22**, 1955-1962

243 Mao, D., Wachter, E. and Wallace, B. A. (1982) Folding of the mitochondrial proton adenosinetriphosphatase proteolipid channel in phospholipid vesicles. Biochemistry. **21**, 4960-4968

244 Linde, D., Pogni, R., Canellas, M., Lucas, F., Guallar, V., Baratto, M. C., Sinicropi, A., Saez-Jimenez, V., Coscolin, C., Romero, A., Medrano, F. J., Ruiz-Duenas, F. J. and Martinez, A. T. (2015) Catalytic surface radical in dye-decolorizing peroxidase: a computational, spectroscopic and site-directed mutagenesis study. Biochem J. **466**, 253-262

245 Cole, C., Barber, J. D. and Barton, G. J. (2008) The Jpred 3 secondary structure prediction server. Nucleic Acids Res. **36**, W197-201

246 Fernandez-Fueyo, E., Linde, D., Almendral, D., Lopez-Lucendo, M. F., Ruiz-Duenas, F. J. and Martinez, A. T. (2015) Description of the first fungal dye-decolorizing peroxidase oxidizing manganese(II). Appl Microbiol Biotechnol

247 Min, K., Gong, G., Woo, H. M., Kim, Y. and Um, Y. (2015) A dye-decolorizing peroxidase from Bacillus subtilis exhibiting substrate-dependent optimum temperature for dyes and beta-ether lignin dimer. Sci Rep. **5**, 8245

248 Dunford, H. B. (1999) Heme Peroxidase. John Wiley & Sons, Inc., New York

249 Wariishi, H., Akileswaran, L. and Gold, M. H. (1988) Manganese peroxidase from the basidiomycete Phanerochaete chrysosporium: spectral characterization of the oxidized states and the catalytic cycle. Biochemistry. **27**, 5365-5370

250 Renganathan, V., Miki, K. and Gold, M. H. (1986) Role of molecular oxygen in lignin peroxidase reactions. Arch Biochem Biophys. **246**, 155-161

251 Wariishi, H., and M. H. Gold. . (1989) Lignin peroxidase compound III: formation, inactivation, and conversion to the native enzyme. FEBS Lett. **243**, 165–168

252 Valli, K., Wariishi, H. and Gold, M. H. (1990) Oxidation of monomethoxylated aromatic compounds by lignin peroxidase: role of veratryl alcohol in lignin biodegradation. Biochemistry. **29**, 8535-8539

253 Wariishi, H. and Gold, M. H. (1990) Lignin peroxidase compound III. Mechanism of formation and decomposition. J Biol Chem. **265**, 2070-2077

254 Ogola, H. J., Hashimoto, N., Miyabe, S., Ashida, H., Ishikawa, T., Shibata, H. and Sawa, Y. (2010) Enhancement of hydrogen peroxide stability of a novel Anabaena sp. DyP-type peroxidase by site-directed mutagenesis of methionine residues. Appl Microbiol Biotechnol. **87**, 1727-1736

255 Poulos, T. L., Edwards, S. L., Wariishi, H. and Gold, M. H. (1993) Crystallographic refinement of lignin peroxidase at 2 A. J Biol Chem. **268**, 4429-4440

256 Sundaramoorthy, M., Kishi, K., Gold, M. H. and Poulos, T. L. (1994) The crystal structure of manganese peroxidase from Phanerochaete chrysosporium at 2.06-A resolution. J Biol Chem. **269**, 32759-32767

257 Finzel, B. C., Poulos, T. L. and Kraut, J. (1984) Crystal structure of yeast cytochrome c peroxidase refined at 1.7-A resolution. J Biol Chem. **259**, 13027-13036

258 Yoshida, T., Tsuge, H., Konno, H., Hisabori, T. and Sugano, Y. (2011) The catalytic mechanism of dye-decolorizing peroxidase DyP may require the swinging movement of an aspartic acid residue. FEBS J. **278**, 2387-2394

259 Linde, D., Coscolin, C., Liers, C., Hofrichter, M., Martinez, A. T. and Ruiz-Duenas, F. J. (2014) Heterologous expression and physicochemical characterization of a fungal dye-decolorizing peroxidase from Auricularia auricula-judae. Protein Expr Purif. **103**, 28-37

260 Trent, J. T., 3rd and Hargrove, M. S. (2002) A ubiquitously expressed human hexacoordinate hemoglobin. J Biol Chem. **277**, 19538-19545

261 Reeder, B. J., Svistunenko, D. A. and Wilson, M. T. (2011) Lipid binding to cytoglobin leads to a change in haem co-ordination: a role for cytoglobin in lipid signalling of oxidative stress. Biochem J. **434**, 483-492

262 Johansson, T. and Nyman, P. O. (1993) Isozymes of lignin peroxidase and manganese(II) peroxidase from the white-rot basidiomycete Trametes versicolor. I. Isolation of enzyme forms and characterization of physical and catalytic properties. Arch Biochem Biophys. **300**, 49-56

263 Paszczynski, A., Huynh, V. B. and Crawford, R. (1986) Comparison of ligninase-I and peroxidase-M2 from the white-rot fungus Phanerochaete chrysosporium. Arch Biochem Biophys. **244**, 750-765

264 Tian, Y. S., Xu, H., Peng, R. H., Yao, Q. H. and Wang, R. T. (2014) Heterologous expression and characterization of laccase 2 from capable of decolourizing different recalcitrant dyes. Biotechnol Biotechnol Equip. **28**, 248-258

265 Ruiz-Dueñas, F. J., Morales, M., Garcia, E., Miki, Y., Martinez, M. J. and Martinez, A. T. (2009) Substrate oxidation sites in versatile peroxidase and other basidiomycete peroxidases. J Exp Bot. **60**, 441-452

266 Morales, M., Mate, M. J., Romero, A., Martinez, M. J., Martinez, A. T. and Ruiz-Duenas, F. J. (2012) Two oxidation sites for low redox potential substrates: a directed mutagenesis, kinetic, and crystallographic study on Pleurotus eryngii versatile peroxidase. J Biol Chem. **287**, 41053-41067

267 Strittmatter, E., Wachter, S., Liers, C., Ullrich, R., Hofrichter, M., Plattner, D. A. and Piontek, K. (2013) Radical formation on a conserved tyrosine residue is crucial for DyP activity. Arch Biochem Biophys. **537**, 161-167

268 Doyle, W. A., Blodig, W., Veitch, N. C., Piontek, K. and Smith, A. T. (1998) Two substrate interaction sites in lignin peroxidase revealed by site-directed mutagenesis. Biochemistry. **37**, 15097-15105

269 Floudas, D., Binder, M., Riley, R., Barry, K., Blanchette, R. A., Henrissat, B., Martinez, A. T., Otillar, R., Spatafora, J. W., Yadav, J. S., Aerts, A., Benoit, I., Boyd, A., Carlson, A., Copeland, A., Coutinho, P. M., de Vries, R. P., Ferreira, P., Findley, K., Foster, B., Gaskell, J., Glotzer, D., Gorecki, P., Heitman, J., Hesse, C., Hori, C., Igarashi, K., Jurgens, J. A., Kallen, N., Kersten, P., Kohler, A., Kues, U., Kumar, T. K., Kuo, A., LaButti, K., Larrondo, L. F., Lindquist, E., Ling, A., Lombard, V., Lucas, S., Lundell, T., Martin, R., McLaughlin, D. J., Morgenstern, I., Morin, E., Murat, C., Nagy, L. G., Nolan, M., Ohm, R. A., Patyshakuliyeva, A., Rokas, A., Ruiz-Duenas, F. J., Sabat, G., Salamov, A., Samejima, M., Schmutz, J., Slot, J. C., St John, F., Stenlid, J., Sun, H., Sun, S., Syed, K., Tsang, A., Wiebenga, A., Young, D., Pisabarro, A., Eastwood, D. C., Martin, F., Cullen, D., Grigoriev, I. V. and Hibbett, D. S. (2012) The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science. **336**, 1715-1719

270 Mauro, J. M., Fishel, L. A., Hazzard, J. T., Meyer, T. E., Tollin, G., Cusanovich, M. A. and Kraut, J. (1988) Tryptophan-191----phenylalanine, a proximal-side mutation in yeast cytochrome c

peroxidase that strongly affects the kinetics of ferrocytochrome c oxidation. Biochemistry. **27**, 6243-6256

271 Klapper, M. H. and Faraggi, M. (1979) Applications of pulse radiolysis to protein chemistry. Q Rev Biophys. **12**, 465-519

272 Martinez, A. T., Ruiz-Duenas, F. J., Martinez, M. J., Del Rio, J. C. and Gutierrez, A. (2009) Enzymatic delignification of plant cell wall: from nature to mill. Curr Opin Biotechnol. **20**, 348-357

273 Guallar, V. and Wallrapp, F. (2008) Mapping protein electron transfer pathways with QM/MM methods. J R Soc Interface. **5 Suppl 3**, S233-239

274 Miethke, M., Monteferrante, C. G., Marahiel, M. A. and van Dijl, J. M. (2013) The Bacillus subtilis EfeUOB transporter is essential for high-affinity acquisition of ferrous and ferric iron. Biochim. Biophys. Acta - Mol. Cell Res, 1833, 2267–2278

275 Younes, I. and Rinaudo, M. (2015) Chitin and chitosan preparation from marine sources. Structure, properties and applications. Mar Drugs. **13**, 1133-1174

276 de Souza, P. M. and de Oliveira Magalhaes, P. (2010) Application of microbial alpha-amylase in industry - A review. Braz J Microbiol. **41**, 850-861

277 Muralikrishna, G. and Nirmala, M. (2005) Cereal α-amylases – an overview. Carbohydrate Polymers **60**, 163-173

278 Peplow, M. (2014) Cellulosic ethanol fights for life. Nature. **507**, 152-153

279 Krieger, K. (2014) Renewable energy: Biofuels heat up. Nature. **508**, 448-449

280 Dimarogona, M., Topakas, E. and Christakopoulos, P. (2012) Cellulose degradation by oxidative enzymes. Comput Struct Biotechnol J. **2**, e201209015

281 Lynd, L. R., Weimer, P. J., van Zyl, W. H. and Pretorius, I. S. (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev. **66**, 506-577, table of contents

282 Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J. and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. Journal of the American Chemical Society. **135**, 6069-6077

283 Wilson, D. B. (2012) Processive and nonprocessive cellulases for biofuel production--lessons from bacterial genomes and structural analysis. Appl Microbiol Biotechnol. **93**, 497-502

284 Yakovlev, I., Vaaje-Kolstad, G., Hietala, A. M., Stefanczyk, E., Solheim, H. and Fossdal, C. G. (2012) Substrate-specific transcription of the enigmatic GH61 family of the pathogenic white-rot fungus Heterobasidion irregulare during growth on lignocellulose. Appl Microbiol Biotechnol. **95**, 979-990

285 Horn, S. J., Sorlie, M., Varum, K. M., Valjamae, P. and Eijsink, V. G. (2012) Measuring processivity. Methods Enzymol. **510**, 69-95

286 Ray, A., Saykhedkar, S., Ayoubi-Canaan, P., Hartson, S. D., Prade, R. and Mort, A. J. (2012) Phanerochaete chrysosporium produces a diverse array of extracellular enzymes when grown on sorghum. Appl Microbiol Biotechnol. **93**, 2075-2089

287 Quinlan, R. J., Sweeney, M. D., Leggio, L. L., Otten, H., Poulsen, J.-C. N., Johansen, K. S., Krogh, K. B., Jørgensen, C. I., Tovborg, M. and Anthonsen, A. (2011) Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proceedings of the National Academy of Sciences. **108**, 15079-15084

288 Wilson, D. B. (2009) Cellulases and biofuels. Curr Opin Biotechnol. **20**, 295-299

289 Sweeney, M. D. and Xu, F. (2012) Biomass Converting Enzymes as Industrial Biocatalysts for Fuels and Chemicals: Recent Developments. Catalyst. **2**, 244-263

290 Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V. and Henrissat, B. (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for Glycogenomics. Nucleic Acids Res. **37**, D233-238

291 Vaaje-Kolstad, G., Westereng, B., Horn, S. J., Liu, Z., Zhai, H., Sorlie, M. and Eijsink, V. G. (2010) An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science. **330**, 219-222

292 Beeson, W. T., Vu, V. V., Span, E. A., Phillips, C. M. and Marletta, M. A. (2015) Cellulose degradation by polysaccharide monooxygenases. Annu Rev Biochem. **84**, 923-946

293 Hemsworth, G. R., Johnston, E. M., Davies, G. J. and Walton, P. H. (2015) Lytic Polysaccharide Monooxygenases in Biomass Conversion. Trends Biotechnol. **33**, 747-761

294 Harris, P. V., Welner, D., McFarland, K., Re, E., Navarro Poulsen, J.-C., Brown, K., Salbo, R., Ding, H., Vlasenko, E. and Merino, S. (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry. **49**, 3305-3316

295 Langston, J. A., Shaghasi, T., Abbate, E., Xu, F., Vlasenko, E. and Sweeney, M. D. (2011) Oxidoreductive cellulose depolymerization by the enzymes cellobiose dehydrogenase and glycoside hydrolase 61. Appl Environ Microbiol. **77**, 7007-7015

296 Horn, S. J., Vaaje-Kolstad, G., Westereng, B. and Eijsink, V. G. (2012) Novel enzymes for the degradation of cellulose. Biotechnology for biofuels. **5**, 1-13

297 Eibinger, M., Ganner, T., Bubner, P., Rosker, S., Kracher, D., Haltrich, D., Ludwig, R., Plank, H. and Nidetzky, B. (2014) Cellulose surface degradation by a lytic polysaccharide monooxygenase and its effect on cellulase hydrolytic efficiency. J Biol Chem. **289**, 35929-35938

298 Vermaas, J. V., Crowley, M. F., Beckham, G. T. and Payne, C. M. (2015) Effects of lytic polysaccharide monooxygenase oxidation on cellulose structure and binding of oxidized cellulose oligomers to cellulases. J Phys Chem B. **119**, 6129-6143

299 Hu, J., Chandra, R., Arantes, V., Gourlay, K., van Dyk, J. S. and Saddler, J. N. (2015) The addition of accessory enzymes enhances the hydrolytic performance of cellulase enzymes at high solid loadings. Bioresour Technol. **186**, 149-153

300 Dimarogona, M., Topakas, E., Olsson, L. and Christakopoulos, P. (2012) Lignin boosts the cellulase performance of a GH-61 enzyme from Sporotrichum thermophile. Bioresour Technol. **110**, 480-487

301 Levasseur, A., Drula, E., Lombard, V., Coutinho, P. M. and Henrissat, B. (2013) Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels. **6**, 41

302 Agger, J. W., Isaksen, T., Varnai, A., Vidal-Melgosa, S., Willats, W. G., Ludwig, R., Horn, S. J., Eijsink, V. G. and Westereng, B. (2014) Discovery of LPMO activity on hemicelluloses shows the importance of oxidative processes in plant cell wall degradation. Proc Natl Acad Sci U S A. **111**, 6287-6292

303 Isaksen, T., Westereng, B., Aachmann, F. L., Agger, J. W., Kracher, D., Kittl, R., Ludwig, R., Haltrich, D., Eijsink, V. G. and Horn, S. J. (2014) A C4-oxidizing lytic polysaccharide monooxygenase cleaving both cellulose and cello-oligosaccharides. J Biol Chem. **289**, 2632-2642

304 Forsberg, Z., Vaaje-Kolstad, G., Westereng, B., Bunaes, A. C., Stenstrom, Y., MacKenzie, A., Sorlie, M., Horn, S. J. and Eijsink, V. G. (2011) Cleavage of cellulose by a CBM33 protein. Protein Sci. **20**, 1479-1483

305 Hemsworth, G. R., Henrissat, B., Davies, G. J. and Walton, P. H. (2014) Discovery and characterization of a new family of lytic polysaccharide monooxygenases. Nat Chem Biol. **10**, 122-126

306 Lo Leggio, L., Simmons, T. J., Poulsen, J. C., Frandsen, K. E., Hemsworth, G. R., Stringer, M. A., von Freiesleben, P., Tovborg, M., Johansen, K. S., De Maria, L., Harris, P. V., Soong, C. L., Dupree, P., Tryfona, T., Lenfant, N., Henrissat, B., Davies, G. J. and Walton, P. H. (2015) Structure and boosting activity of a starch-degrading lytic polysaccharide monooxygenase. Nat Commun. **6**, 5961

307 Vu, V. V., Beeson, W. T., Span, E. A., Farquhar, E. R. and Marletta, M. A. (2014) A family of starch-active polysaccharide monooxygenases. Proc Natl Acad Sci U S A. **111**, 13822-13827

308 Vaaje-Kolstad, G., Horn, S. J., van Aalten, D. M., Synstad, B. and Eijsink, V. G. (2005) The non-catalytic chitin-binding protein CBP21 from Serratia marcescens is essential for chitin degradation. J Biol Chem. **280**, 28492-28497

309 Beeson, W. T., Phillips, C. M., Cate, J. H. and Marletta, M. A. (2012) Oxidative cleavage of cellulose by fungal copper-dependent polysaccharide monooxygenases. J Am Chem Soc. **134**, 890-892

310 Vaaje-Kolstad, G., Bohle, L. A., Gaseidnes, S., Dalhus, B., Bjoras, M., Mathiesen, G. and Eijsink, V. G. (2012) Characterization of the chitinolytic machinery of Enterococcus faecalis V583 and high-resolution structure of its oxidative CBM33 enzyme. J Mol Biol. **416**, 239-254

311 Aachmann, F. L., Sørlie, M., Skjåk-Bræk, G., Eijsink, V. G. and Vaaje-Kolstad, G. (2012) NMR structure of a lytic polysaccharide monooxygenase provides insight into copper binding, protein dynamics, and substrate interactions. Proceedings of the National Academy of Sciences. **109**, 18779-18784

312 Wu, M., Beckham, G. T., Larsson, A. M., Ishida, T., Kim, S., Payne, C. M., Himmel, M. E., Crowley, M. F., Horn, S. J., Westereng, B., Igarashi, K., Samejima, M., Stahlberg, J., Eijsink, V. G. and Sandgren, M. (2013) Crystal structure and computational characterization of the lytic polysaccharide monooxygenase GH61D from the Basidiomycota fungus Phanerochaete chrysosporium. J Biol Chem. **288**, 12828-12839

313 Forsberg, Z., Mackenzie, A. K., Sorlie, M., Rohr, A. K., Helland, R., Arvai, A. S., Vaaje-Kolstad, G. and Eijsink, V. G. (2014) Structural and functional characterization of a conserved pair of bacterial cellulose-oxidizing lytic polysaccharide monooxygenases. Proc Natl Acad Sci U S A. **111**, 8446-8451

314 Forsberg, Z., Rohr, A. K., Mekasha, S., Andersson, K. K., Eijsink, V. G., Vaaje-Kolstad, G. and Sorlie, M. (2014) Comparative study of two chitin-active and two cellulose-active AA10-type lytic polysaccharide monooxygenases. Biochemistry. **53**, 1647-1656

315 Vu, V. V., Beeson, W. T., Phillips, C. M., Cate, J. H. and Marletta, M. A. (2014) Determinants of regioselective hydroxylation in the fungal polysaccharide monooxygenases. J Am Chem Soc. **136**, 562-565

316 Gudmundsson, M., Kim, S., Wu, M., Ishida, T., Momeni, M. H., Vaaje-Kolstad, G., Lundberg, D., Royant, A., Stahlberg, J., Eijsink, V. G., Beckham, G. T. and Sandgren, M. (2014) Structural and electronic snapshots during the transition from a Cu(II) to Cu(I) metal center of a lytic polysaccharide monooxygenase by X-ray photoreduction. J Biol Chem. **289**, 18782-18792

317 Phillips, C. M., Beeson, W. T., Cate, J. H. and Marletta, M. A. (2011) Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by Neurospora crassa. ACS Chem Biol. **6**, 1399-1406

318 Kim, S., Stahlberg, J., Sandgren, M., Paton, R. S. and Beckham, G. T. (2014) Quantum mechanical calculations suggest that lytic polysaccharide monooxygenases use a copper-oxyl, oxygen-rebound mechanism. Proc Natl Acad Sci U S A. **111**, 149-154

319 Crawford, D. L. (1978) Lignocellulose decomposition by selected streptomyces strains. Appl Environ Microbiol. **35**, 1041-1045

320 Goodfellow, M. and Williams, S. T. (1983) Ecology of actinomycetes. Annu Rev Microbiol. **37**, 189-216

321 Adams, A. S., Jordan, M. S., Adams, S. M., Suen, G., Goodwin, L. A., Davenport, K. W., Currie, C. R. and Raffa, K. F. (2011) Cellulose-degrading bacteria associated with the invasive woodwasp Sirex noctilio. ISME J. **5**, 1323-1331

322 Takasuka, T. E., Book, A. J., Lewin, G. R., Currie, C. R. and Fox, B. G. (2013) Aerobic deconstruction of cellulosic biomass by an insect-associated Streptomyces. Sci Rep. **3**, 1030

323 Vaaje-Kolstad, G., Houston, D. R., Riemen, A. H., Eijsink, V. G. and van Aalten, D. M. (2005) Crystal structure and binding properties of the Serratia marcescens chitin-binding protein CBP21. Journal of Biological Chemistry. **280**, 11313-11319

324 Nakagawa, Y. S., Kudo, M., Loose, J. S., Ishikawa, T., Totani, K., Eijsink, V. G. and Vaaje-Kolstad, G. (2015) A small lytic polysaccharide monooxygenase from Streptomyces griseus targeting alpha- and beta-chitin. FEBS J. **282**, 1065-1079

325 Paspaliari, D. K., Loose, J. S., Larsen, M. H. and Vaaje-Kolstad, G. (2015) Listeria monocytogenes has a functional chitinolytic system and an active lytic polysaccharide monooxygenase. FEBS J. **282**, 921-936

326 Hemsworth, G. R., Taylor, E. J., Kim, R. Q., Gregory, R. C., Lewis, S. J., Turkenburg, J. P., Parkin, A., Davies, G. J. and Walton, P. H. (2013) The copper active site of CBM33 polysaccharide oxygenases. J Am Chem Soc. **135**, 6069-6077

327 Stigbrand, T., Malmstrom, B. G. and Vanngard, T. (1971) On the state of copper in the blue protein umecyanin. FEBS Lett. **12**, 260-262

328 Gewirth, A. A., Cohen, S. L., Schugar, H. J. and Solomon, E. I. (1987) Spectroscopic and Theoretical- Studies of the Unusual Electron-Paramagnetic-Res Parameters of Distorted Tetrahedral Cupric Sites - Correlations to X-Ray Spectral Features of Core Levels. . Inorg Chem. **7**, 1133-1146

329 LaCroix, L. B. (1998) Spectroscopic and geometric variations in perturbed blue copper centers: Electronic structures of stellacyanin and cucumber basic protein. J Am Chem Soc. **37**, 9621-9631

330 Cox, J. C., Aasa, R. and Malmstrom, B. G. (1978) EPR studies on the blue copper protein, rusticyanin: a protein involved in Fe2+ oxidation at pH 2.0 in Thiobacillus ferro-oxidans. FEBS Lett. **93**, 157-160

331 Knowles, P. F., Brown, R. D., Koenig, S. H., Wang, S., Scott, R. A., McGuirl, M. A., Brown, D. E. and Dooley, D. M. (1995) Spectroscopic Studies of the Active Site of Galactose Oxidase. Inorg. Chem. **15**, 3895–3902

332 Buffoni, F., Corte, L. D. and Knowles, P. F. (1968) The nature of copper in pig plasma benzylamine oxidase. Biochem J. **106**, 575-576

333 Mondovi, B., Rotilio, G., Costa, M. T., Finazzi-Agro, A., Chiancone, E., Hansen, R. E. and Beinert, H. (1967) Diamine oxidase from pig kidney. Improved purification and properties. J Biol Chem. **242**, 1160-1167

334 Wishnick, M., Lane, M. D., Scrutton, M. C. and Mildvan, A. S. (1969) The presence of tightly bound copper in ribulose diphosphate carboxylase from spinach. J Biol Chem. **244**, 5761-5763

335 McCord, J. M. and Fridovich, I. (1969) Superoxide dismutase. An enzymic function for erythrocuprein (hemocuprein). J Biol Chem. **244**, 6049-6055

336 Tsukihara, T., Aoyama, H., Yamashita, E., Tomizaki, T., Yamaguchi, H., Shinzawa-Itoh, K., Nakashima, R., Yaono, R. and Yoshikawa, S. (1995) Structures of metal sites of oxidized bovine heart cytochrome c oxidase at 2.8 A. Science. **269**, 1069-1074

337 Duff, A. P., Cohen, A. E., Ellis, P. J., Kuchar, J. A., Langley, D. B., Shepard, E. M., Dooley, D. M., Freeman, H. C. and Guss, J. M. (2003) The crystal structure of Pichia pastoris lysyl oxidase. Biochemistry. **42**, 15148-15157

338 Holmberg, C. G. and Laurell, C. B. (1948) Histaminolytic activity of a copper protein in serum. Nature. **161**, 236

339 Dollwet, H. H. A., and Sorenson, J. R. J. . (1985) History uses of copper compounds in medicine. Trace Elements Med. **2**, 80-87

340 Hodgkinson, V. and Petris, M. J. (2012) Copper homeostasis at the host-pathogen interface. Journal of Biological Chemistry. **287**, 13549-13555

341 Grass, G., Rensing, C. and Solioz, M. (2011) Metallic copper as an antimicrobial surface. Applied and environmental microbiology. **77**, 1541-1547

342 Irving, H. and Williams, R. (1948) Order of stability of metal complexes. Nature. **162**, 746-747

343 Robinson, N. J. and Winge, D. R. (2010) Copper metallochaperones. Annual review of biochemistry. **79**, 537-562

344 Waldron, K. J. and Robinson, N. J. (2009) How do bacterial cells ensure that metalloproteins get the correct metal? Nature Reviews Microbiology. **7**, 25-35

345 Chevion, M. (1988) A site-specific mechanism for free radical induced biological damage: the essential role of redox-active transition metals. Free Radical Biology and Medicine. **5**, 27-37

346 Imlay, J. A. (2002) How oxygen damages microbes: oxygen tolerance and obligate anaerobiosis. Advances in microbial physiology. **46**, 111-153

347 Halliwell, B. and Gutteridge, J. M. (1984) Oxygen toxicity, oxygen radicals, transition metals and disease. Biochem J. **219**, 1-14

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

349 Chillappagari, S., Seubert, A., Trip, H., Kuipers, O. P., Marahiel, M. A. and Miethke, M. (2010) Copper stress affects iron homeostasis by destabilizing iron-sulfur cluster formation in Bacillus subtilis. J Bacteriol. **192**, 2512-2524

350 Rae, T. D., Schmidt, P. J., Pufahl, R. A., Culotta, V. C. and O'Halloran, T. V. (1999) Undetectable intracellular free copper: the requirement of a copper chaperone for superoxide dismutase. Science. **284**, 805-808

351 Pufahl, R., Singer, C., Peariso, K., Lin, S.-J., Schmidt, P., Fahrni, C., Culotta, V. C., Penner-Hahn, J. and O'halloran, T. (1997) Metal ion chaperone function of the soluble Cu (I) receptor Atx1. Science. **278**, 853-856

352 O'Halloran, T. V. and Culotta, V. C. (2000) Metallochaperones, an intracellular shuttle service for metal ions. Journal of Biological Chemistry. **275**, 25057-25060

353 Huffman, D. L. and O'Halloran, T. V. (2001) Function, structure, and mechanism of intracellular copper trafficking proteins. Annu Rev Biochem. **70**, 677-701

354 Harrison, M. D., Jones, C. E., Solioz, M. and Dameron, C. T. (2000) Intracellular copper routing: the role of copper chaperones. Trends in biochemical sciences. **25**, 29-32

355 Lin, S.-J., Pufahl, R. A., Dancis, A., O'Halloran, T. V. and Culotta, V. C. (1997) A role for the Saccharomyces cerevisiae ATX1 gene in copper trafficking and iron transport. Journal of Biological Chemistry. **272**, 9215-9220

356 Rosenzweig, A. C., Huffman, D. L., Hou, M. Y., Wernimont, A. K., Pufahl, R. A. and O'Halloran, T. V. (1999) Crystal structure of the Atx1 metallochaperone protein at 1.02 A resolution. Structure. **7**, 605-617

357 Lin, S.-J. and Culotta, V. C. (1995) The ATX1 gene of Saccharomyces cerevisiae encodes a small metal homeostasis factor that protects cells against reactive oxygen toxicity. Proceedings of the National Academy of Sciences. **92**, 3784-3788

358 Dancis, A., Yuan, D. S., Haile, D., Askwith, C., Eide, D., Moehle, C., Kaplan, J. and Klausner, R. D. (1994) Molecular characterization of a copper transport protein in S. cerevisiae: an unexpected role for copper in iron transport. Cell. **76**, 393-402

359 Askwith, C., Eide, D., Van Ho, A., Bernard, P. S., Li, L., Davis-Kaplan, S., Sipe, D. M. and Kaplan, J. (1994) The FET3 gene of S. cerevisiae encodes a multicopper oxidase required for ferrous iron uptake. Cell. **76**, 403-410

360 Yuan, D. S., Stearman, R., Dancis, A., Dunn, T., Beeler, T. and Klausner, R. D. (1995) The Menkes/Wilson disease gene homologue in yeast provides copper to a ceruloplasmin-like oxidase required for iron uptake. Proc Natl Acad Sci U S A. **92**, 2632-2636

361 Klomp, L. W. J., Lin, S. J., Yuan, D. S., Klausner, R. D., Culotta, V. C. and Gitlin, J. D. (1997) Identification and functional expression of HAH1, a novel human gene involved in copper homeostasis. Journal of Biological Chemistry. **272**, 9221-9226

362 Vulpe, C., Levinson, B., Whitney, S., Packman, S. and Gitschier, J. (1993) Isolation of a candidate gene for Menkes disease and evidence that it encodes a copper-transporting ATPase. Nat Genet. **3**, 7-13

363 Chelly, J., Tumer, Z., Tonnesen, T., Petterson, A., Ishikawa-Brush, Y., Tommerup, N., Horn, N. and Monaco, A. P. (1993) Isolation of a candidate gene for Menkes disease that encodes a potential heavy metal binding protein. Nat Genet. **3**, 14-19

364 Yamaguchi, Y., Heiny, M. E. and Gitlin, J. D. (1993) Isolation and characterization of a human liver cDNA as a candidate gene for Wilson disease. Biochem Biophys Res Commun. **197**, 271-277

365 Bull, P. C., Thomas, G. R., Rommens, J. M., Forbes, J. R. and Cox, D. W. (1993) The Wilson disease gene is a putative copper transporting P-type ATPase similar to the Menkes gene. Nat Genet. **5**, 327-337

366 Tanzi, R. E., Petrukhin, K., Chernov, I., Pellequer, J. L., Wasco, W., Ross, B., Romano, D. M., Parano, E., Pavone, L., Brzustowicz, L. M. and et al. (1993) The Wilson disease gene is a copper transporting ATPase with homology to the Menkes disease gene. Nat Genet. **5**, 344-350

367 de Bie, P., Muller, P., Wijmenga, C. and Klomp, L. W. (2007) Molecular pathogenesis of Wilson and Menkes disease: correlation of mutations with molecular defects and disease phenotypes. J Med Genet. **44**, 673-688

368 Himelblau, E., Mira, H., Lin, S. J., Culotta, V. C., Peñarrubia, L. and Amasino, R. M. (1998) Identification of a functional homolog of the yeast copper homeostasis gene ATX1 from Arabidopsis. Plant Physiology. **117**, 1227-1234

369 Solioz, M. and Stoyanov, J. V. (2003) Copper homeostasis in Enterococcus hirae. FEMS Microbiol Rev. **27**, 183-195

370 Wernimont, A. K., Huffman, D. L., Lamb, A. L., O'Halloran, T. V. and Rosenzweig, A. C. (2000) Structural basis for copper transfer by the metallochaperone for the Menkes/Wilson disease proteins. Nat Struct Biol. **7**, 766-771

371 Arnesano, F., Banci, L., Bertini, I., Huffman, D. L. and O'Halloran, T. V. (2001) Solution structure of the Cu(I) and apo forms of the yeast metallochaperone, Atx1. Biochemistry. **40**, 1528-1539

372 Badarau, A. and Dennison, C. (2011) Copper trafficking mechanism of CXXC-containing domains: Insight from the pH-dependence of their Cu (I) affinities. Journal of the American Chemical Society. **133**, 2983-2988

373 Dwarakanath, S., Chaplin, A. K., Hough, M. A., Rigali, S., Vijgenboom, E. and Worrall, J. A. R. (2012) Response to Copper Stress in Streptomyces lividans Extends beyond Genes under Direct Control of a Copper-sensitive Operon Repressor Protein (CsoR). Journal of Biological Chemistry. **287**, 17833-17847

374 Odermatt, A. and Solioz, M. (1995) Two trans-acting metalloregulatory proteins controlling expression of the copper-ATPases of Enterococcus hirae. J Biol Chem. **270**, 4349-4354

375 Cobine, P., Wickramasinghe, W. A., Harrison, M. D., Weber, T., Solioz, M. and Dameron, C. T. (1999) The Enterococcus hirae copper chaperone CopZ delivers copper (I) to the CopY repressor. FEBS letters. **445**, 27-30

376 Strausak, D. and Solioz, M. (1997) CopY is a copper-inducible repressor of the Enterococcus hirae copper ATPases. Journal of Biological Chemistry. **272**, 8932-8936

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

378 Ma, Z., Jacobsen, F. E. and Giedroc, D. P. (2009) Metal transporters and metal sensors: How coordination chemistry controls bacterial metal homeostasis. Chemical reviews. **109**, 4644

379 Banci, L., Bertini, I., Del Conte, R., Markey, J. and Ruiz-Duenas, F. J. (2001) Copper trafficking: the solution structure of Bacillus subtilis CopZ. Biochemistry. **40**, 15660-15668

380 Banci, L., Bertini, I., Ciofi-Baffoni, S., D'Onofrio, M., Gonnelli, L., Marhuenda-Egea, F. C. and Ruiz-Dueñas, F. J. (2002) Solution structure of the N-terminal domain of a potential copper-translocating P-type ATPase from Bacillus subtilis in the apo and Cu (I) loaded states. Journal of molecular biology. **317**, 415-429

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

382 Wu, C. C., Rice, W. J. and Stokes, D. L. (2008) Structure of a copper pump suggests a regulatory role for its metal-binding domain. Structure. **16**, 976-985

383 González-Guerrero, M., Hong, D. and Argüello, J. M. (2009) Chaperone-mediated Cu+ Delivery to Cu+ Transport ATPases. Journal of Biological Chemistry. **284**, 20804-20811

384 Gourdon, P., Liu, X. Y., Skjorringe, T., Morth, J. P., Moller, L. B., Pedersen, B. P. and Nissen, P. (2011) Crystal structure of a copper-transporting PIB-type ATPase. Nature. **475**, 59-64

385 Padilla-Benavides, T., McCann, C. J. and Arguello, J. M. (2013) The mechanism of Cu+ transport ATPases: interaction with CU+ chaperones and the role of transient metal-binding sites. J Biol Chem. **288**, 69-78

386 Banci, L., Bertini, I., Ciofi-Baffoni, S., Del Conte, R. and Gonnelli, L. (2003) Understanding copper trafficking in bacteria: interaction between the copper transport protein CopZ and the N-terminal domain of the copper ATPase CopA from Bacillus subtilis. Biochemistry. **42**, 1939-1949

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

388 Kihlken, M. A., Leech, A. P. and Le Brun, N. E. (2002) Copper-mediated dimerization of CopZ, a predicted copper chaperone from Bacillus subtilis. Biochemical Journal. **368**, 729

389 Singleton, C., Hearnshaw, S., Zhou, L., Le Brun, N. and Hemmings, A. (2009) Mechanistic insights into Cu (I) cluster transfer between the chaperone CopZ and its cognate Cu (I)-transporting P-type ATPase, CopA. Biochem. J. **424**, 347-356

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

391 Ward, S. K., Hoye, E. A. and Talaat, A. M. (2008) The global responses of Mycobacterium tuberculosis to physiological levels of copper. J Bacteriol. **190**, 2939-2946

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

393 Chang, F. M., Coyne, H. J., Cubillas, C., Vinuesa, P., Fang, X., Ma, Z., Ma, D., Helmann, J. D., Garcia-de los Santos, A., Wang, Y. X., Dann, C. E., 3rd and Giedroc, D. P. (2014) Cu(I)-mediated allosteric switching in a copper-sensing operon repressor (CsoR). J Biol Chem. **289**, 19204-19217

394 Porto, T. V., Hough, M. A. and Worrall, J. A. (2015) Structural insights into conformational switching in the copper metalloregulator CsoR from Streptomyces lividans. Acta Crystallogr D Biol Crystallogr. **71**, 1872-1878

395 Tan, B. G., Vijgenboom, E. and Worrall, J. A. (2014) Conformational and thermodynamic hallmarks of DNA operator site specificity in the copper sensitive operon repressor from Streptomyces lividans. Nucleic Acids Res. **42**, 1326-1340

396 Hiard, S., Maree, R., Colson, S., Hoskisson, P. A., Titgemeyer, F., van Wezel, G. P., Joris, B., Wehenkel, L. and Rigali, S. (2007) PREDetector: a new tool to identify regulatory elements in bacterial genomes. Biochem Biophys Res Commun. **357**, 861-864

397 Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. **5**, 621-628

398 Ellman, G. L. (1959) Tissue sulfhydryl groups. Arch Biochem Biophys. **82**, 70-77

399 Xiao, Z., Donnelly, P. S., Zimmermann, M. and Wedd, A. G. (2008) Transfer of copper between bis(thiosemicarbazone) ligands and intracellular copper-binding proteins. insights into mechanisms of copper uptake and hypoxia selectivity. Inorg Chem. **47**, 4338-4347

400 Xiao, Z., Loughlin, F., George, G. N., Howlett, G. J. and Wedd, A. G. (2004) C-terminal domain of the membrane copper transporter Ctr1 from Saccharomyces cerevisiae binds four Cu(I) ions as a cuprous-thiolate polynuclear cluster: sub-femtomolar Cu(I) affinity of three proteins involved in copper trafficking. J Am Chem Soc. **126**, 3081-3090

401 Xiao, Z., Brose, J., Schimo, S., Ackland, S. M., La Fontaine, S. and Wedd, A. G. (2011) Unification of the Copper (I) Binding Affinities of the Metallo-chaperones Atx1, Atox1, and Related Proteins DETECTION PROBES AND AFFINITY STANDARDS. Journal of biological chemistry. **286**, 11047-11055

402 Zhou, L., Singleton, C., Hecht, O., Moore, G. R. and Le Brun, N. E. (2012) Cu(I)- and proton-binding properties of the first N-terminal soluble domain of Bacillus subtilis CopA. FEBS Journal. **279**, 285-298

403 Xiao, Z. and Wedd, A. G. (2010) The challenges of determining metal-protein affinities. Natural product reports. **27**, 768-789

404 Zhou, L., Singleton, C. and Le Brun, N. (2008) High Cu (I) and low proton affinities of the CXXC motif of Bacillus subtilis CopZ. Biochem. J. **413**, 459-465

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

406 Benesch, R. E. and Benesch, R. (1955) The Acid Strength of the -SH Group in Cysteine and Related Compounds. Journal of the American Chemical Society. **77**, 5877-5881

407 Ledwidge, R., Hong, B., Dotsch, V. and Miller, S. M. (2010) NmerA of Tn501 mercuric ion reductase: structural modulation of the pKa values of the metal binding cysteine thiols. Biochemistry. **49**, 8988-8998

408 Jocelyn, P. C. (1972) Biochemistry of the SH Group. Academic Press, London and New York

409 Nelson, J. W. and Creighton, T. E. (1994) Reactivity and ionization of the active site cysteine residues of DsbA, a protein required for disulfide bond formation in vivo. Biochemistry. **33**, 5974-5983

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

411 Arnesano, F., Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Huffman, D. L. and O'Halloran, T. V. (2001) Characterization of the binding interface between the copper chaperone Atx1 and the first cytosolic domain of Ccc2 ATPase. Journal of Biological Chemistry. **276**, 41365-41376

412 Wimmer, R., Herrmann, T., Solioz, M. and Wüthrich, K. (1999) NMR structure and metal interactions of the CopZ copper chaperone. Journal of Biological Chemistry. **274**, 22597-22603

413 Tanchou, V. r., Gas, F., Urvoas, A., Cougouluègne, F. o., Ruat, S., Averseng, O. and Quéméneur, E. (2004) Copper-mediated homo-dimerisation for the HAH1 metallochaperone. Biochemical and biophysical research communications. **325**, 388-394

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

415 Jothivasan, V. K. and Hamilton, C. J. (2008) Mycothiol: synthesis, biosynthesis and biological functions of the major low molecular weight thiol in actinomycetes. Nat Prod Rep. **25**, 1091-1117

416 Newton, G. L., Buchmeier, N. and Fahey, R. C. (2008) Biosynthesis and functions of mycothiol, the unique protective thiol of Actinobacteria. Microbiol Mol Biol Rev. **72**, 471-494

417 Newton, G. L., Rawat, M., La Clair, J. J., Jothivasan, V. K., Budiarto, T., Hamilton, C. J., Claiborne, A., Helmann, J. D. and Fahey, R. C. (2009) Bacillithiol is an antioxidant thiol produced in Bacilli. Nat Chem Biol. **5**, 625-627

418 Fahey, R. C. (2013) Glutathione analogs in prokaryotes. Biochim Biophys Acta. **1830**, 3182-3198

419 Harris, T. K. and Turner, G. J. (2002) Structural basis of perturbed pKa values of catalytic groups in enzyme active sites. IUBMB Life. **53**, 85-98

420 Lewin, A., Crow, A., Oubrie, A. and Le Brun, N. E. (2006) Molecular basis for specificity of the extracytoplasmic thioredoxin ResA. Journal of Biological Chemistry. **281**, 35467-35477

421 Kortemme, T., Darby, N. J. and Creighton, T. E. (1996) Electrostatic interactions in the active site of the N-terminal thioredoxin-like domain of protein disulfide isomerase. Biochemistry. **35**, 14503-14511

422 Forman-Kay, J. D., Clore, G. M. and Gronenborn, A. M. (1992) Relationship between electrostatics and redox function in human thioredoxin: characterization of pH titration shifts using two-dimensional homo- and heteronuclear NMR. Biochemistry. **31**, 3442-3452

423 Nordstrand, K., Aslund, F., Meunier, S., Holmgren, A., Otting, G. and Berndt, K. D. (1999) Direct NMR observation of the Cys-14 thiol proton of reduced Escherichia coli glutaredoxin-3 supports the presence of an active site thiol-thiolate hydrogen bond. FEBS Lett. **449**, 196-200

424 Tottey, S., Rondet, S. A., Borrelly, G. P., Robinson, P. J., Rich, P. R. and Robinson, N. J. (2002) A copper metallochaperone for photosynthesis and respiration reveals metal-specific targets, interaction with an importer, and alternative sites for copper acquisition. J Biol Chem. **277**, 5490-5497

425 Reyes-Caballero, H., Campanello, G. C. and Giedroc, D. P. (2011) Metalloregulatory proteins: metal selectivity and allosteric switching. Biophys Chem. **156**, 103-114

426 Chen, T. Y., Santiago, A. G., Jung, W., Krzeminski, L., Yang, F., Martell, D. J., Helmann, J. D. and Chen, P. (2015) Concentration- and chromosome-organization-dependent regulator unbinding from DNA for transcription regulation in living cells. Nat Commun. **6**, 7445

427 Philips, S. J., Canalizo-Hernandez, M., Yildirim, I., Schatz, G. C., Mondragon, A. and O'Halloran, T. V. (2015) TRANSCRIPTION. Allosteric transcriptional regulation via changes in the overall topology of the core promoter. Science. **349**, 877-881

428 Corbett, D., Schuler, S., Glenn, S., Andrew, P. W., Cavet, J. S. and Roberts, I. S. (2011) The combined actions of the copper-responsive repressor CsoR and copper-metallochaperone CopZ modulate CopA-mediated copper efflux in the intracellular pathogen Listeria monocytogenes. Molecular microbiology. **81**, 457-472

429 Waldron, K. J., Firbank, S. J., Dainty, S. J., Perez-Rama, M., Tottey, S. and Robinson, N. J. (2010) Structure and metal loading of a soluble periplasm cuproprotein. J Biol Chem. **285**, 32504-32511

430 Banci, L., Bertini, I., McGreevy, K. S. and Rosato, A. (2010) Molecular recognition in copper trafficking. Nat Prod Rep. **27**, 695-710

431 Chaplin, A. K., Tan, B. G., Vijgenboom, E. and Worrall, J. A. (2015) Copper trafficking in the CsoR regulon of Streptomyces lividans. Metallomics. **7**, 145-155

432 Vita, N., Platsaki, S., Basle, A., Allen, S. J., Paterson, N. G., Crombie, A. T., Murrell, J. C., Waldron, K. J. and Dennison, C. (2015) A four-helix bundle stores copper for methane oxidation. Nature. **525**, 140-143

Appendix

Appendix 1

Appendix 1.1: Derivation of Equation 3.3

Scheme 1:

a = GlxA

y = GlxA*Cu(II)

x = GlxA-Cu(II)

At t=0, $a=a_0$, y=0 and x=0

At t=t, (a_0-y-x)

$$K_d = \frac{[a][Cu]}{[y]}$$

$$K_d = \frac{[a_0-y-x][Cu]}{[y]}$$

$$K_d[y] = [Cua_0 - Cu_y - Cu_x]$$

$$(K_d[y] + [Cu_y]) = [Cua_0 - Cu_x]$$

$$[y](K_d + [Cu]) = [Cu(a_0 - x)]$$

$$[y] = \left(\frac{[Cu]}{K_d + [Cu]}\right)([a_0 - x])$$

$$\frac{dx}{dt} = k_2[y]$$

$$\frac{dx}{dt}=\frac{k_2[Cu]}{K_d+[Cu]}([a_0-x])$$

$$\int\frac{dx}{[a_0-x]}=\frac{k_2[Cu]}{K_d+[Cu]}t$$

$$M=\frac{k_2[Cu]}{K_d+[Cu]}t$$

$$\ln([a_0-x])=Mt+Q$$

$$\ln\left(\frac{[a_0]}{[a_0-x]}\right)=Mt$$

$$[a_0-x]=e^{Mt}$$

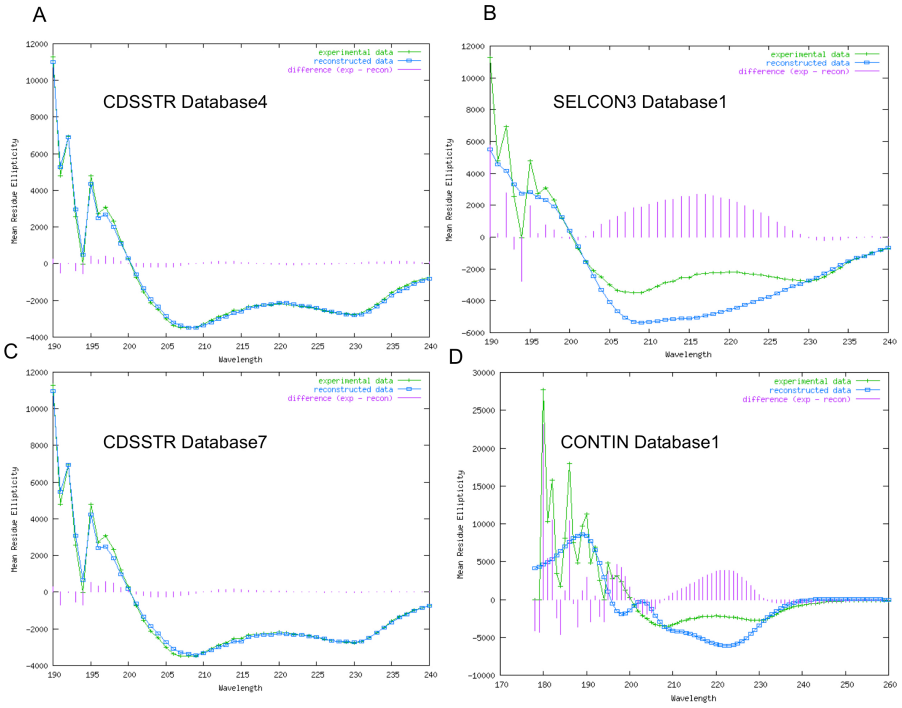
$$[a_0]=[a_0]e^{Mt}-xe^{Mt}$$

$$[x]e^{Mt}=[a_0]e^{Mt}-[a_0]$$

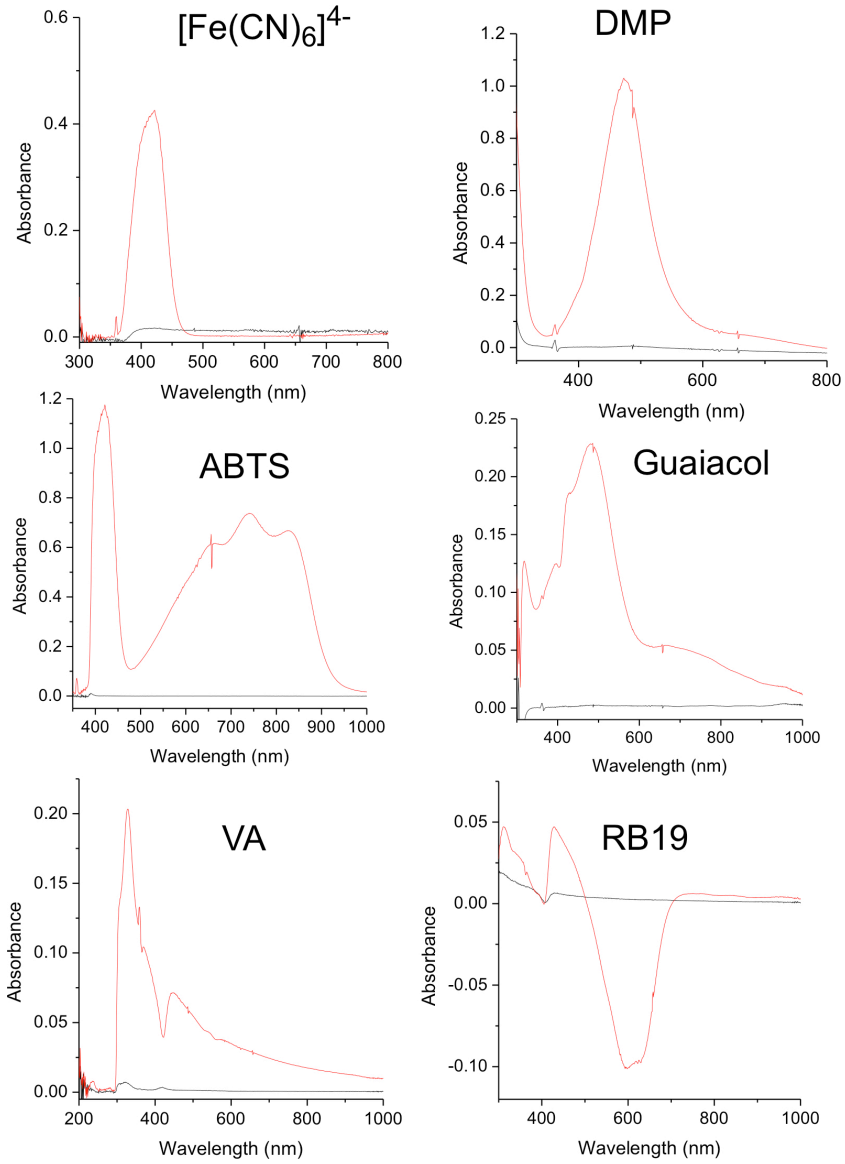
$$[x]=[a_0-a_0]e^{-Mt}=[a_0](1-e^{-Mt})$$

$$x=[a_0](1-e^{\left(\frac{K_2[Cu]}{K_{d+[Cu]}}\right)t}$$

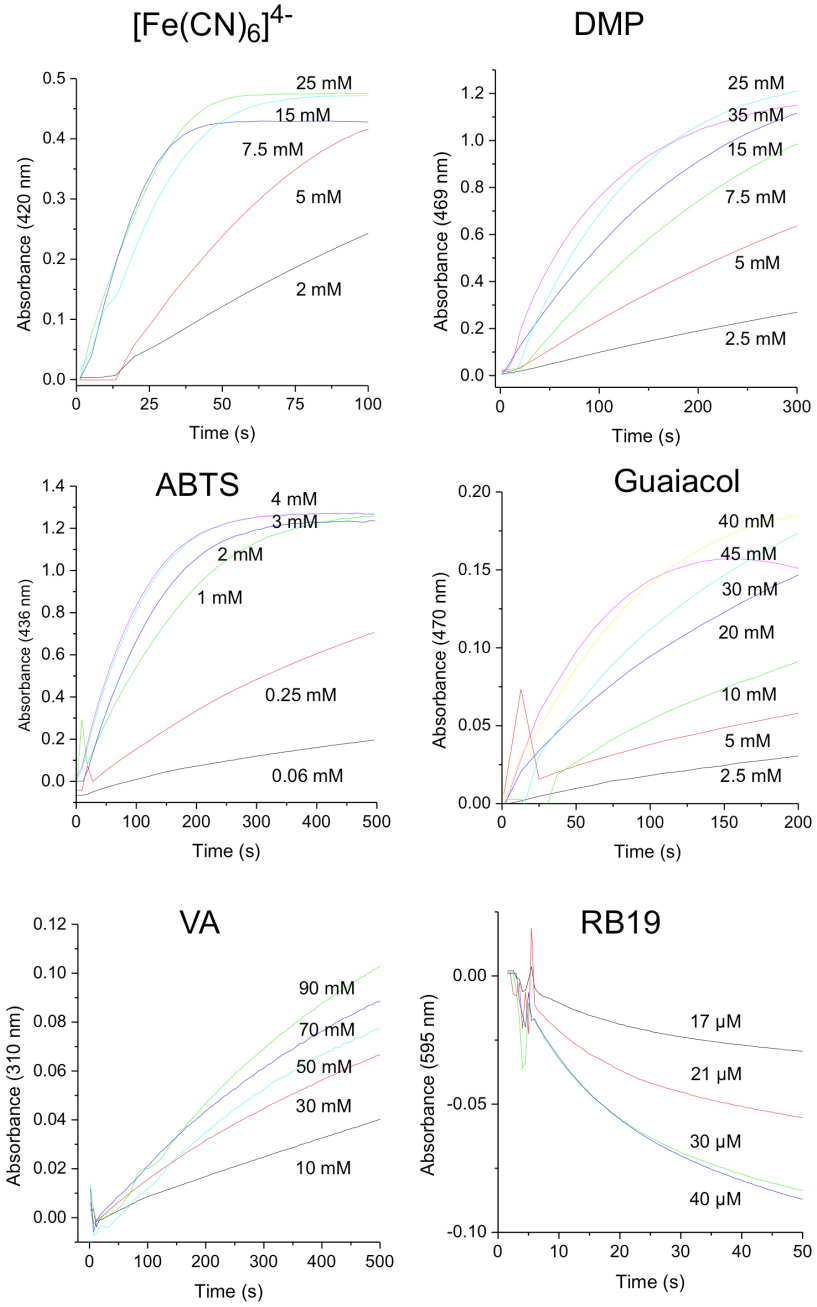
Appendix 2



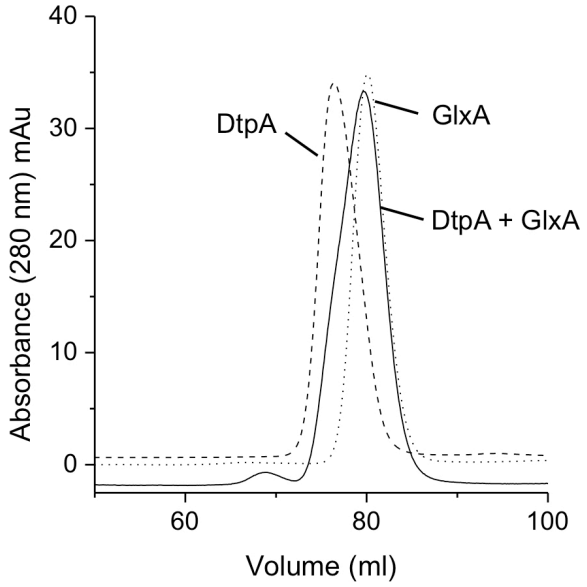
Appendix 2.1: Examples of dichroweb analysis of DtpA CD spectrum. A and C) Examples of programs and databases which gave good fits with NRMSD < 0.1. B and D) Examples of programs and databases which gave bad fits to the data and NRMSD > 0.1.



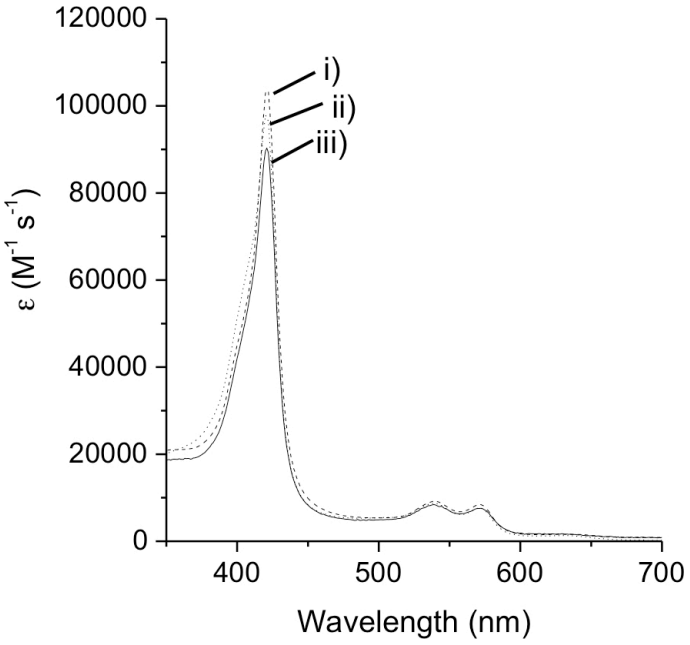
Appendix 2.2: Substrate oxidation by *S. lividans* DtpA. UV-visible absorption spectrum of the substrates depicted during oxidation by DtpA with H_2O_2. $[Fe(CN)_6]^{4-}$ 30 mM pH 5, DMP 30 mM pH 9, ABTS 2.5 mM pH 5, guaiacol 45 mM pH 10.5, VA 90 mM pH 5 and RB19 40 mM pH 4 using 0.2 mM H_2O_2 to initiate the reaction with 30-1000 nM DtpA. Black shows before and red shows after the reaction.



Appendix 2.3: Substrate oxidation by *S. lividans* DtpA. Catalytic turnover of DtpA showing absorbance versus time traces for the substrates and their concentrations depicted. $[Fe(CN)_6]^{4-}$ pH 5, DMP pH 9, ABTS pH 5, guaiacol pH 10.5, VA pH 5 and RB19 pH 4 using 0.2 mM H_2O_2 to initiate the reaction with 30-1000 nM DtpA.

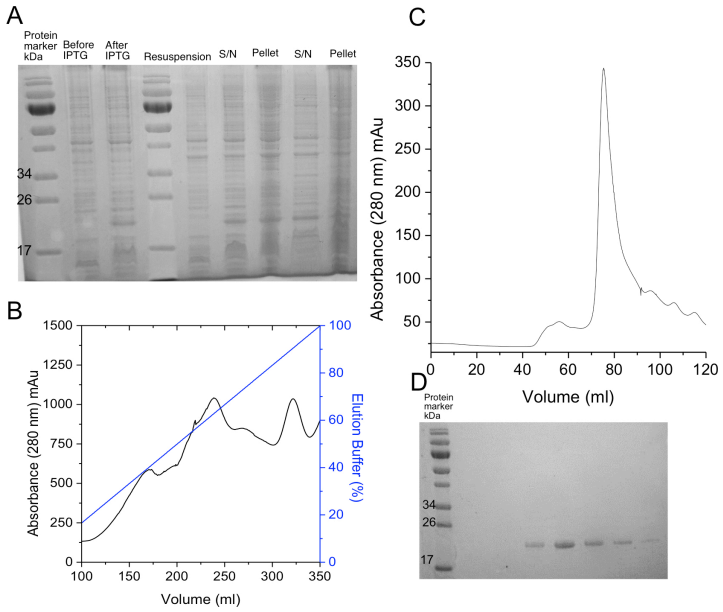


Appendix 2.4: S200 size-exclusion column profiles of GlxA, DtpA and a mixture. An elution profile absorbance 280 nm (mAu) profile of GlxA and DtpA and a 1:1 mixture of the two in 20 mM NaPi, 100 mM NaCl, pH 7 with the absorbance's corrected.

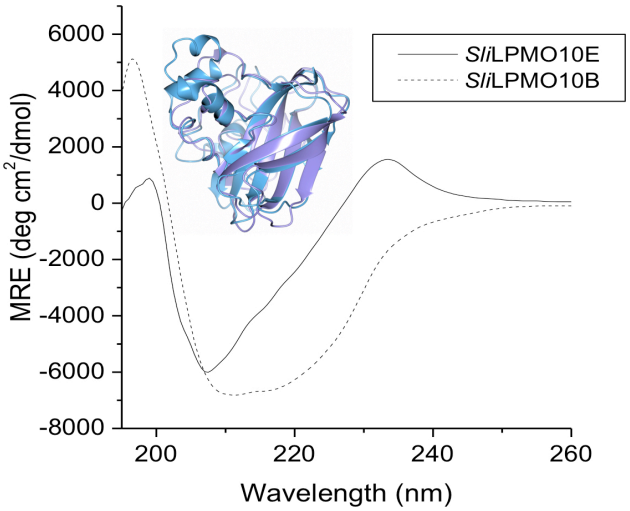


Appendix 2.5. Static UV-visible spectra of oxy-ferrous DtpA (20 mM sodium phosphate pH 7, 100 mM NaCl) generated from i) resting state DtpA, GlxA and 0.2 M glycolaldehyde (GA); ii) resting state DtpA and 0.2 M GA; iii) Compound I generated via addition of 1 equivalent of H_2O_2 followed by addition of 0.2 M GA.

Appendix 3



Appendix 3.1: Over-expression and purification of *Sli*LPMO10B. A) Coomassie stained 15 % SDS PAGE gel analysis of *Sli*LPMO10B during over-expression and osmotic shock treatment, with the content of the lanes depicted in the figure (*Sli*LPMO10B– 20,723 Da). B) DEAE column profiles of *Sli*LPMO10B showing an absorbance peak (mAu) at 280 nm in black at ~80 % buffer B shown in blue for the protein. C) G75 size-exclusion Sephadex column absorbance 280 nm (mAu) profile of *Sli*LPMO10B eluting at ~ 79 ml, consistent with monomeric forms of the protein, in 10 mM sodium acetate pH 5, 150 mM NaCl. D) Coomassie stained 15 % SDS PAGE gel analysis of purified *Sli*LPMO10B displaying a single protein band between 17-26 kDa consistent with the mass of the protein, 20.7 kDa.



Appendix 3.2: Far UV-CD spectra of *Sli*LPMO10E and *Sli*LPMO10B at 20 °C, 10 mM KPi, 50 mM KF, pH 7 with protein concentrations of 20 µM. WT-*Sli*LPMO10E– Solid line and *Sli*LPMO10B– dashed line. Inset, the overlaid crystal structures of *Sli*LPMO10E (purple) and *Sc*LPMO10B from *S. coelicolor* (blue) homologue of *Sli*LPMO10B.

Appendix 4

Appendix 4.1: Simulation calculations

An average K_d value is obtained by using a competition assay to determine the affinity of a protein towards a metal-ion. Binding curves can be simulated for various different metal and chelator concentrations by using the average K_d value. This section describes the method used to generate fits of experimental data and for simulating plots where competition and no competition between the chelator and protein are observed.

For a give L_{tot}, M_{tot} and K_{eq} value, a plot of metal-chelator absorbance against protein concentration can be simulated using Equation 6.4. By using Equation 6.5 and substituting for MP in Equation 6.4, you get:

$$\left(\mathrm{M}^{-1}\right) K_{eq} = \frac{\frac{[\mathrm{P}]_{tot}}{[\mathrm{M}]_{tot}-[\mathrm{ML}_2]}-1}{(\frac{[\mathrm{L}]_{tot}}{[\mathrm{ML}_2]}-2)^2[\mathrm{ML}_2]} \quad \text{Equation 6.10}$$

$$\left(\mathrm{M}^{-1}\right) K_{eq} = \frac{\frac{[\mathrm{P}]_{tot}-[\mathrm{M}]_{tot}+[\mathrm{ML}_2]}{[\mathrm{M}]_{tot}-[\mathrm{ML}_2]}}{(\frac{[\mathrm{L}]_{tot}}{[\mathrm{ML}_2]}-2)^2[\mathrm{ML}_2]} \quad \text{Equation 6.11}$$

On rearranging the terms Equation 6.12 is obtained.

$$\left(\mathrm{M}^{-1}\right) K_{eq} = \frac{\frac{[\mathrm{P}]_{tot}-[\mathrm{M}]_{tot}+[\mathrm{ML}_2]}{[\mathrm{M}]_{tot}-[\mathrm{ML}_2]}}{\frac{([\mathrm{L}]_{tot}-2[\mathrm{ML}_2])^2}{[\mathrm{ML}_2]}} \quad \text{Equation 6.12}$$

By substituting $[\mathrm{P}]_{tot}-[\mathrm{M}]_{tot}$ as Z.

$$\left(\mathrm{M}^{-1}\right) K_{eq} = \frac{\frac{Z+[\mathrm{ML}_2]}{[\mathrm{M}]_{tot}-[\mathrm{ML}_2]}}{\frac{([\mathrm{L}]_{tot}-2[\mathrm{ML}_2])^2}{[\mathrm{ML}_2]}} \quad \text{Equation 6.13}$$

Which can be simplified further to obtain a cubic equation in $[\mathrm{ML}_2]$.

$$-4K_{eq}[ML_2]^3+\{K_{eq}(4[M]_{tot}+4[L]_{tot})-1\}[ML_2]^2-\{K_{eq}(4[M]_{tot}[L]_{tot}+[L]_{tot})^2+Z\}[ML_2]+K_{eq}[M]_{tot}[L]_{tot}^2=0$$

Equation 6.14

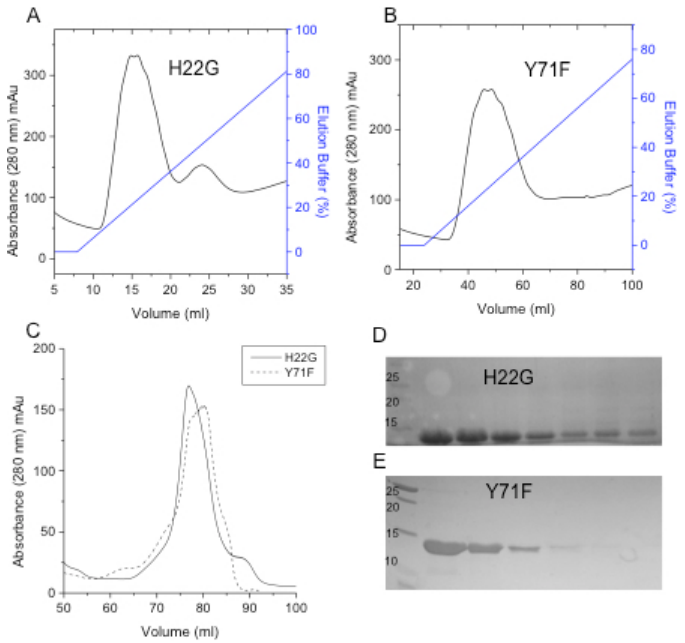
By substituting the coefficients of $[ML_2]$ as A, B, C and D Equation 6.15 is obtained:

$$A[ML_2]^3+B[ML_2]^2+C[ML_2]+D=0$$ Equation 6.15

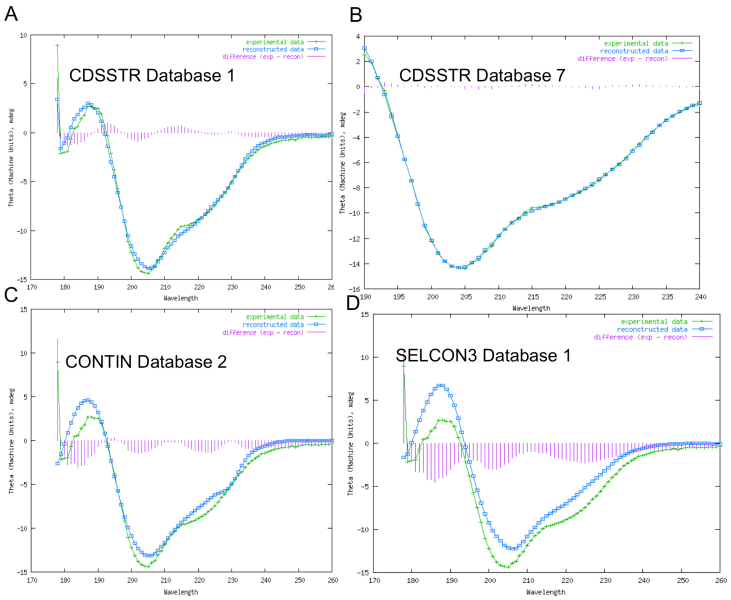
Where,

$$A = -4K_{eq}$$
$$B = K_{eq}(4[M]_{tot} + 4[L]_{tot}) - 1$$
$$C = -K_{eq}(4[M]_{tot}[L]_{tot} + [L]_{tot}^2) - Z$$
$$D = K_{eq}[M]_{tot}[L]_{tot}^2$$

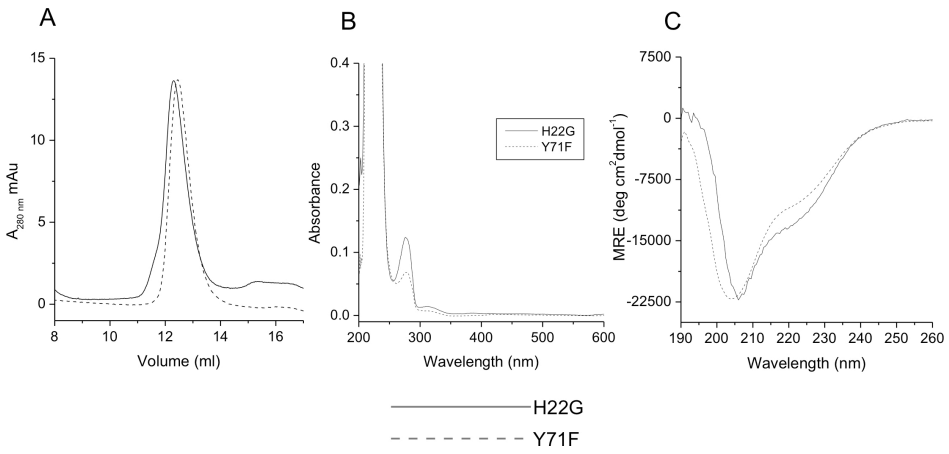
Only the coefficient C contains the term Z that is dependent on the protein concentration. Hence, with increasing protein concentration, coefficient C changes, thus providing a different solution for $[ML_2]$. Thus, for a known concentration of L_{tot} and M_{tot} and a defined K_{eq} value, $[ML_2]$ can be calculated for varying protein concentrations by solving the cubic Equation 6.15. Thus, fits for experimental data can be simulated to determine the accuracy of the binding affinity value obtained from the competition assay. Simulations can also be used to generate curves for various experimental conditions, thus predicting the metal binding affinity behavior of the protein.



Appendix 4.2: Purification of H22G and Y71F CopZ-1317 variants. A) Ni^{2+}-NTA profile of H22G showing an absorbance peak (mAu) at 280 nm in black at ~ 30 % buffer B shown in blue. B) Ni^{2+}-NTA profile of Y71F showing an absorbance peak (mAu) at 280 nm in black at ~ 30 % buffer B shown in blue. C) G75 size-exclusion Sephadex column absorbance 280 nm (mAu) profile of H22G and Y71F eluting at ~ 82 ml and ~ 78 ml, respectively, consistent with monomeric forms of the proteins in 50 mM Tris/HCl, 150 mM NaCl, 2 mM EDTA, 2 mM DTT, pH 8. D) Coomassie stained 15 % SDS PAGE gel analysis of purified H22G displaying a single protein band between 10-15 kDa. E) Coomassie stained 15 % SDS PAGE gel analysis of purified Y71F displaying a single protein band between 10-15 kDa.



Appendix 4.3: Examples of Dichroweb analysis of CopZ-1317. A and B) Examples of programs and databases which give good fits with NRMSD values < 0.1. C and D) Examples of programs and databases which give bad fits and NRMSD values > 0.1.



Appendix 4.4: Preliminary characterization of CopZ-1317 mutants H22G and Y71F. A) Analytical gel filtration chromatography profile for apo-H22G and apo-Y71F in 10 mM MOPS, pH 7.5, 150 mM NaCl, 2 mM DTT. B) UV-visible spectroscopy of CopZ H22G and Y71F in 50 mM Tris, 150 mM NaCl, 2 mM EDTA, 2 mM DTT, pH 8, 20 ^{o}C. C) Far UV-CD spectra for H22G and Y71F at 20 °C, 10 mM KPi, 50 mM KF, pH 7 with protein concentrations of 20 μM.