Structure and properties of earthworm metallothionein-2

by

Gregory Richard Kowald

A thesis submitted in partial fulfilment of the requirements for the degree of

Doctor of Philosophy in Chemistry

University of Warwick, Department of Chemistry

September 2012

Table of Contents

List of Figures

List of Tables

Acknowledgements

It is true to say, that without the faith of my supervisor, I would still be mixing paints right now. So my sincere thanks to Dr Claudia Blindauer for taking me away from man-made polymers and reintroducing me to biological polymers. I express my thanks to Dr Stephen Stürzenbaum and research group, without their contributions this research would not have been possible. Without the continued support of the Blindauer group, my PhD experience would have been very different, and significantly less enjoyable. Special thanks are due to those within the group who have worked alongside me for many years: Dr Jin Lu, Jie Chu, Frances Kondrat and Esther Martin. During my PhD, I have tried to live up to the high standards set by Dr Oksana Leszczyszyn. I hope that I have helped others at least half as much as you did. Thanks Oksana!

For collaboration during this project, my thanks to Maike Hansen, for her assistance in the lab and her enthusiasm when working with me! Within the department, a number of people have been instrumental. My thanks to Anne Smith, Dr Lijiang Song, Dr Ivan Prokes, Phil Aston and Edward Tunnah. I would also like to thank Magnus Lewis-Smith, I dare not imagine what Chapter 4 would have looked like without his help.

My thanks to my Mum and Dad, for listening to me talk about my work for at least half an hour before your eyes glazed over. I could write another 250 pages on the ways you have helped me out in the last four years, so my heartfelt thanks. Not forgetting my brother, for distracting me after long days in the lab, and keeping me ~~sane~~ saner.

Finally, my thanks to Emma, for listening to me talk about my work for **at least** five minutes before your eyes glazed over, but still listening for much, much, *much* longer. I couldn't have done this without you. This thesis is dedicated to you.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. It has been composed by myself and has not been submitted in any previous application for any degree. The work presented (including data generated and data analysis) was carried out by the author except in the cases outlined below:

Data obtained in the purification of wMT-1 was performed in collaboration with Ms M. Hansen. These experiments were used to generate Figures 3.11, 3.12, 3.13A&B, 3.14, 3.17, 3.18, 3.21 in this thesis.

This research was funded by the EPSRC. The ÄKTA FPLC used in this research was obtained through Birmingham Science City: Innovative Uses for Advanced Materials in the Modern World with support from Advantage West Midlands (AWM) and part funded by the European Regional Development Fund (ERDF).

Abstract

Since their discovery in the 1950s, metallothioneins (MTs) have been isolated from an array of organisms, from sea urchins to sheep, and pigs to peas. With the advent of nuclear Overhauser effect nuclear magnetic resonance (NMR) experiments, a powerful route for MT structure determination became available. The contribution of structural data obtained by NMR has enabled identification of hitherto unknown metallothionein motifs, thought to be significant in the 'tuning' of MT for specific roles or functions. As the MT superfamily is so broad, there is a wealth of information still to be uncovered for MTs from some of the lesser-studied organisms. Although many *in vivo* genomics-based investigations have been performed within the invertebrate family, proteomics-based investigations, especially with respect to terrestrial invertebrates, are sparse.

This research has investigated the metal-binding behaviour and determined the solution structure of earthworm metallothionein-2 (wMT-2) from *Lumbricus rubellus*. In addition, a method for purifying both wMTs expressed in adult earthworms is presented; in contrast to previous studies the purification method includes the removal of the N-terminal S●tag. Using modified purification methods, the novel protein wMT-3 was also able to be isolated and characterised. Electrospray-ionisation mass spectrometry (ESI-MS) revealed that all three wMTs bind 7 equivalents of divalent metal ions as the major species at pH ≈ 8. An NMR investigation of cleaved Cd-wMT-2 indicated that metal-binding is localised in two discrete clusters. Conversely to vertebrate MTs, but similar to some invertebrate MTs, the stoichiometry of the domains is of an M_4Cys_{11}- N-terminal cluster, followed by an M_3Cys_9- cluster. When challenged with protons, demetallation of cleaved Cd-wMT-2 occurred in a step-wise manner: a 3 metal loss being followed by a 4 metal loss; behaviour which is unique to the cadmium form. A comparison of the NMR TOCSY spectra of cleaved Cd- and Zn-wMT-2 showed significant differences, indicating that wMT-2 has a strong metal-binding preference for cadmium *in vitro*.

In conclusion, the studies performed within this thesis have enabled the calculation of the first 3-dimensional solution structure of an MT from a terrestrial invertebrate, wMT-2. In addition, the novel protein wMT-3 was successfully isolated and purified.

Abbreviations

5F-BAPTA	5,5'-Difluoro-1,2-bis(2-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid
Å	Ångstrom
AI	Auto-induction
AnCat	Coupled anion/cation exchange
ATP	Adenosine triphosphate
BBO	Broad band observe
bp	base pairs
CCPN	Collaborative Computing Project for NMR
CDF	Cation diffusion facilitator
CING	Common Interface for NMR structure Generation
COSY	Correlation spectroscopy
cv	colume volumes
CYANA	Combined assignment and dynamics algorithm for NMR applications
DRI	Dietary reference intake
DTNB	5,5'-dithiobis (2-nitrobenzoic acid)
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
ESI-MS	Electrospray ionisation - mass spectrometry
FID	Free induction decay
FPLC	Fast protein liquid chromatography
HSQC	Heteronuclear single quantum coherence
ICP-OES	Inductively coupled plasma-optical emission spectroscopy
IEX	Ion exchange chromatography
IPTG	Isopropyl β-D-1-thiogalactopyranoside
K_{MT}	apparent metal-binding constant for metallothionein
LB	Luria-Bertani
LC-MS	Liquid chromatography–mass spectrometry
mAU	milli-absorbance unit
MCS	Multiple cloning site
MRE(s)	Metal response element(s)
MS	Mass spectrometry
MT(s)	Metallothionein(s)
MWCO	Molecular weight cut-off
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser effect spectroscopy
PCR	Polymerase chain reaction
PDB	Protein Data Bank
p*I*	Isoelectric point
$pK_a^{1/2}$	pH of half dissociation
RCSB	Research Collaboratory for Structural Bioinformatics
RMSD	root-mean-square deviation
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SEC	Size-exclusion chromatography
TOCSY	Total correlation spectroscopy
UV	Ultraviolet
V_o	Void volume
wcw	Wet cell weight
wMT(s)	worm metallothionein(s)
wrt	with respect to

Amino Acids

1-Letter code	Amino Acid	Abbreviation
A	Alanine	Ala
C	Cysteine	Cys
D	Aspartic acid	Asp
E	Glutamic acid	Glu
F	Phenylalanine	Phe
G	Glycine	Gly
H	Histidine	His
I	Isoleucine	Ile
K	Lysine	Lys
L	Leucine	Leu
M	Methionine	Met
N	Asparagine	Asn
P	Proline	Pro
Q	Glutamine	Gln
R	Arginine	Arg
S	Serine	Ser
T	Threonine	Thr
V	Valine	Val
W	Tryptophan	Trp
Y	Tyrosine	Tyr

1

Introduction

1.1 Metal ions in biology

Environmental conditions influence the evolution of organisms; this can be said with some certainty based on the work of Lenski and research group [Michigan State University, USA], and their experiments with *E. coli* for over 20 years [1, 2]. Hence to some extent, organisms are a product of their environment. The classic examples being Galápagos finches, products of the distinctive ecosystems across the Galápagos Islands [3, 4].

Over time, organisms have developed dependencies on certain minerals, present in abundance, for correct cellular function [5]. Within the human body there are ten essential trace elements, eight of which are metals: iron, zinc, manganese, copper, molybdenum, chromium, cobalt and selenium [6]. **Figure 1.1** illustrates the relationship between the requirements of the essential trace metals (**Figure 1.1**, **PURPLE**), and the maximum tolerable doses before severe side-effects are observed (**Figure 1.1**, **ORANGE**). With regard to the highest tolerable amounts of essential trace elements, selenium is the least terrestrially abundant*,

* Kenneth Barbalace. Periodic Table of Elements. EnvironmentalChemistry.com. 1995 - 2011. Accessed on-line: 11/3/2011 http://EnvironmentalChemistry.com/yogi/periodic/

and becomes toxic at the lowest dose [7], whereas iron is most terrestrially abundant, and becomes toxic at the highest dose. So the data shown below indicate that the rarer essential trace metals are toxic in comparatively low doses [8]. However, the lack of certain metals is equally deadly to organisms [9].

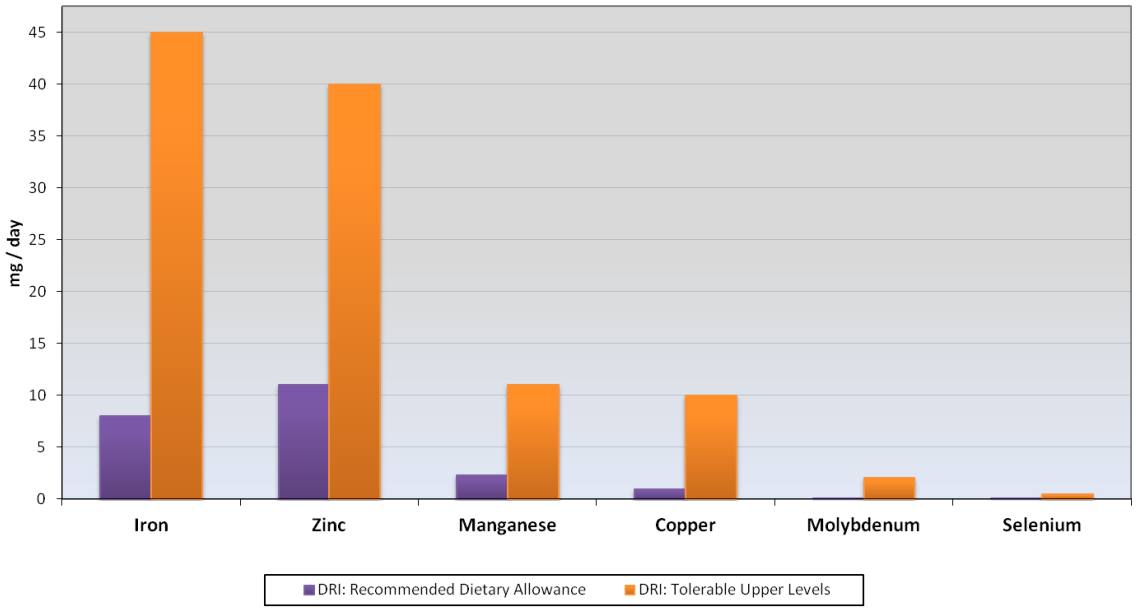


Figure 1.1: The six essential trace elements of which there is dose-based data, highlighting the requirements, and ability of the human body to deal with excesses of these minerals. Recommended daily requirements are highlighted in **PURPLE**; maximum tolerable doses (before side-effects observed) highlighted in **ORANGE**. Chromium and cobalt have incomplete data at this time, and are not included.

Data compiled from Food and Nutrition Board, Institute of Medicine, National Academies: 'Dietary Reference Intakes (DRIs): Tolerable Upper Intake Levels, Elements'; Dietary Reference Intakes for Calcium, Phosphorous, Magnesium, Vitamin D, and Fluoride (1997); Dietary Reference Intakes for Vitamin C, Vitamin E, Selenium, and Carotenoids (2000); Dietary Reference Intakes for Vitamin A, Vitamin K, Arsenic, Boron, Chromium, Copper, Iodine, Iron, Manganese, Molybdenum, Nickel, Silicon, Vanadium, and Zinc (2001); Dietary Reference Intakes for Water, Potassium, Sodium, Chloride, and Sulfate (2005); and Dietary Reference Intakes for Calcium and Vitamin D (2011).

Metals are ubiquitous within the human body, as indicated by their prevalence in structures as diverse as bones and the brain [10]. The reason for this pervasiveness is the ability of certain metal ions to act as co-factors in enzymes, forming specialist catalytic sites such as those found in alkaline phosphatase and carbonic anhydrase [10, 11].

However, metal ions that are generally toxic - cadmium, mercury and lead - serve no known biological function in mammals [12].

1.2 Biological use of cadmium in a simple organism

Although toxic to mammals, there is one organism that can exploit cadmium for a biological process. In general, the surface of the ocean is an environment that is very nutrient poor, especially with respect to zinc (< 1 μM) [13]. In this environment, the coastal diatom *Thalassiosira weissflogii* has adapted to utilise more abundant cadmium as a metal centre within carbonic anhydrase, a role classically filled by zinc [14].

In order to utilise metal ions and maintain them within non-toxic concentrations, means of handling metal ions have evolved [15]. This evolution appears to have been guided by a focus on metal ions in greatest contact with the organism [16]. Continuous adaptation of these metal ion control mechanisms has generated the complexity of present metal ion homeostasis [17].

1.3 Metal ion homeostasis

For most modern organisms to survive, tightly controlled supply and partitioning of metal ions must be maintained [18]; this is the heart of metal ion homeostasis [9]. To paraphrase Paracelsus, the founder of toxicology [19], 'All things are poison, and nothing is without poison;

only the dose permits something not to be poisonous' [20]. Therefore to control the toxicity of a metal ion, we must control the availability (or 'dose') of these ions. The evolution of metal ion homeostasis has been influenced by two main pressures: the propensity for metal ions to participate in redox reactions [21], critical for biological processes such as photosynthesis; and metal ion cytotoxicity when uncontrolled [22, 23].

Metal ion homeostasis has given rise to a specific range of allowable metal ion concentrations within cells [18] before detrimental effects are observed: some metals are controlled to very low concentrations (Cu^{2+}, 10^{-21} M) [24], others are controlled at much higher concentrations (Mn^{2+}, 10^{-5} M) [24]. The level is proportional to their toxicity, with highly toxic copper maintained at lower concentrations than less toxic zinc (Refer to **Figure 1.1**). When metal ions are present in concentrations outside of the allowable range, a variety of organism-wide toxic effects are observed [22, 25, 26].

When homeostasis goes wrong

Although zinc is toxic at very high [27] or low [28] quantities (see **Figure 1.1**), small quantities of cadmium will cause significant cellular damage [29, 30]. This is due to cadmium acting as a chemical doppelgänger, affecting zinc and calcium metabolism and targeting the kidneys and bones [31]. In the kidneys, cadmium perturbs calcium

absorption by disrupting vitamin D metabolism [31]. The end result is that cadmium, not calcium, is incorporated into bone structure causing soft bones [32, 33].

At the cellular level, as cadmium is a soft Lewis acid (large ionic radius, easily polarised) and therefore thiophilic, it will displace zinc from sites within metalloproteins, such as cysteine rich Metallothionein (MT) [29, 30]. The effect on an organism of MT-bound zinc being replaced by cadmium has been the focus of numerous studies [31, 36, 37]. These studies have shown that although Cd-MT within the liver and kidneys is non-toxic, Cd-MT in the blood will cause significant kidney damage before excretion [31].

Another mechanism through which cadmium exerts cellular toxicity is through the replacement of zinc [30, 38] in either structural sites (ie Cys_4 - zinc fingers [36]), or enzyme active sites (ie $His_2Glu(H_2O)$ - alcohol dehydrogenase [38]). The susceptibility for cadmium replacement is proportional to the number of co-ordinating cysteine residues, with Cys_4 structural sites showing the highest affinity for metal ion replacement [30]. Abnormalities in homeostasis of one metal ion can also cause cascading effects throughout an entire organism. For example, an excess of zinc within an organism can induce copper deficiency [27] through competitive inhibition of gastrointestinal absorption [39].

In an attempt to maintain metal ions under homeostatic control, cells obtain their quotient of metal ions through varied and complex pathways [22, 26, 40-42]. The entire system of metal ion homeostasis can be broadly classified into three processes: absorption of metal ions (influx); storage of metal ions; and excretion of metal ions (efflux) [8, 22].

Influx & efflux of metal ions

There is a general problem with passing divalent metal ions through cell membranes. Whereas neutral gases (O_2, CO_2) and anions with small charge-densities can move directly through cell membranes, most cations with high charge-densities must be actively transported [43]. There are three main groups of transporters that facilitate cation entry into cells: those driven by chemical energy (ATP) [22]; those driven by electrochemical gradients (H^+ and other ions) [22]; and gated channels (specialised proteins that open pores in cell membranes) [44, 45]. Some systems, such as those found in yeast, wheat, and certain species of fish, show a combination of these transporters [46]. Generally speaking, a metal ion transporter can be classified as either high affinity, or low affinity for particular metal ions [47, 48]. A high affinity transporter (**Figure 1.2. RED**) is required when cells are metal deficient, with the expression of transporter genes usually under the control of a Metal Response Element (MRE) and a metal sensor protein. On the other hand, low affinity transporters (**Figure 1.2. PINK**) are not usually under the control of MREs and are present when cells are metal-replete [48, 49].

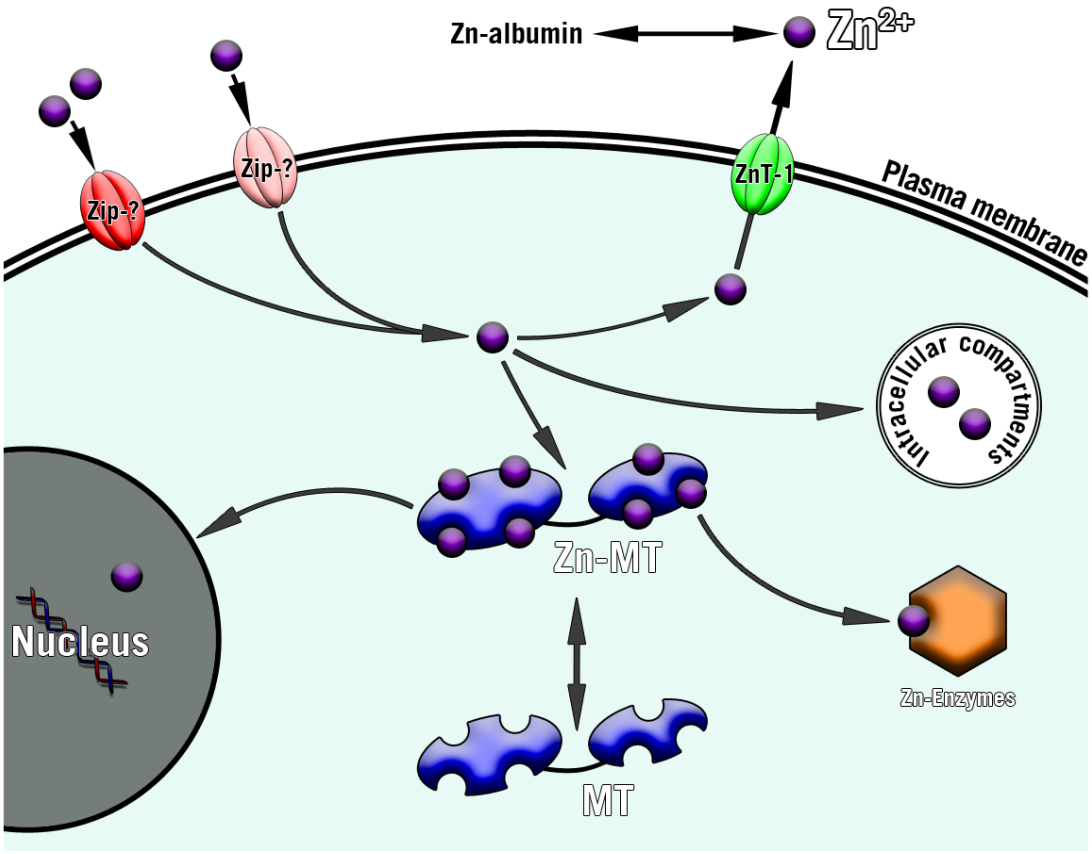


Figure 1.2: Schematic of zinc homeostasis in mammalian cells. Simplified and combined from [50, 51]. Metal ion influx (**RED**, PINK) controlled by Zip- proteins. **RED** Indicates a high-affinity influx transporter, PINK indicates a low affinity influx transporter. Metal ion efflux (**GREEN**) effected by ZnT- proteins. Storage proteins (such as metallothionein) coloured **BLUE**, with zinc-containing enzymes coloured **ORANGE**.

Using zinc homeostasis in humans as a model, we observe two classes of proteins that control influx and efflux between the extracellular space and the cytosol. The Zip (Zrt- and Irt-like protein) family control the influx of zinc [52]; the cation diffusion facilitator (CDF) ZnT (zinc transporter) family of proteins control the efflux (See **Figure 1.2**). Once within the cell, metal ions must be either safely stored until required, or expelled from the cytosol using efflux mechanisms. To maintain a sufficient intracellular supply of metal ions, both storage and transport proteins are required [53].

Storage and transport of metal ions

Within the cell, storage and transport proteins are used for shuttling metal ions to compartments such as the endoplasmic reticulum and Golgi apparatus [50] (for protein synthesis and packaging), or to metal-requiring enzymes [54]. The most widely studied system is intracellular copper transport by a specialised class of delivery proteins, metallochaperones [55]. Metallochaperones mediate protein-protein metal transfer, for example human copper chaperone (CCS) is used to transport Cu(I) to superoxide dismutase 1 (SOD1), thereby making it functionally active [56]. This system avoids toxicity from the uncontrolled redox activity of copper ions [57]. Another family of proteins, induced by excess zinc and cadmium (and copper), also participates in the role of metal ion storage and delivery: Metallothioneins (MTs, **Figure 1.2. BLUE**) [58-60]. Similar to storage proteins such as ferritin (induced by excess iron [26]), MTs are strongly expressed when cells are under metal-induced stress [60].

1.4 Metallothioneins

Metallothioneins (MTs) are small metal-binding proteins (< 100 residues), and when fully metallated have ≈ 10-20 % of their final mass contributed by metal and sulphur [62]. In general, MTs lack significant regions of secondary structure (α-helices and β-sheets) [63], and therefore in the absence of bound metal ions are thought to exist as random coils [21]. MTs are present at high concentrations in the cytosol, with small amounts present in the nucleus [64]. The role of MTs within

the nucleus is not yet known [65], but being present within the nucleus during development does make MT immune to cytosolic lysosomal degradation [66]. There are also low concentrations of MT in extracellular fluids such as plasma, bile and urine [65]. This has enabled a novel test for cadmium exposure in humans: cadmium exposure induces MT; changes in Cd-MT level can be detected by SEC-ICP-MS [67]. Other postulated roles for MTs include: metal ion transport, metal ion detoxification, free-radical scavenging, storage of metal ions, metabolism, immune response and genotoxicity / carcinogenicity (Reviews of the proposed roles for MTs can be found in [68] and [69]). This may indicate that there is no 'main' role for MTs in general, but they instead participate in many different roles [65].

Since their discovery in the brains of mammals [70], interest in MT research has increased in profile. Although the mechanism of action is not currently know, MTs have been implicated in aiding the repair of central nervous system damage [71], and in helping to protect the brain from injury [72]. The upregulation of MT is also a consequence of many human neurological diseases [73] (Alzheimer's disease [74, 75], short-course Creutzfeld-Jakob disease [76], and multiple sclerosis [77]) as well as in some types of cancer cells [78].

Classification

Since the discovery of MT in horse kidneys [79] over 50 years ago, there have been two widely accepted classification systems for MTs [69, 80-82]. The first classification system, (Nordberg & Kojima 1979 [83]; Modified Fowler 1987 [84]) contained three classes of MT:

- Class I: proteins with a high degree of cysteine congruity to equine renal MT;
- Class II: proteins with only a minor degree of cysteine congruity to equine renal MT;
- Class III: atypical or enzymatically synthesised polypeptides (phytochelatins).

The old system was soon deemed inflexible in characterising the growing number of MT sequences being reported in following years. As a result, a new convention was suggested by Binz & Kägi in 1999 [85]. Similar to the first system, the MT superfamily as a whole is defined as comprising all polypeptides which with high primary sequence homology to equine renal MT [84]. However, the new system grouped similar MT sequences into families (many of these families are shown as similar sequences in **Appendix 1**). Each MT sequence can be classified as belonging to only one family, with each family further subdivided into subfamilies, subgroups and isoforms [85, 86]. All members of an MT family share a high degree of primary sequence identity, and as such their sequences can be aligned with each other [85]. This system provides more flexibility in MT classification, as new MT sequences can be compared to any other classified sequence, and assigned a family based on the shared

characteristics of those sequences. The current system for classification is summarised in **Figure 1.3**[†], which highlights both the ubiquity of MTs throughout the biosphere, and the lack of their corresponding experimentally determined structures.

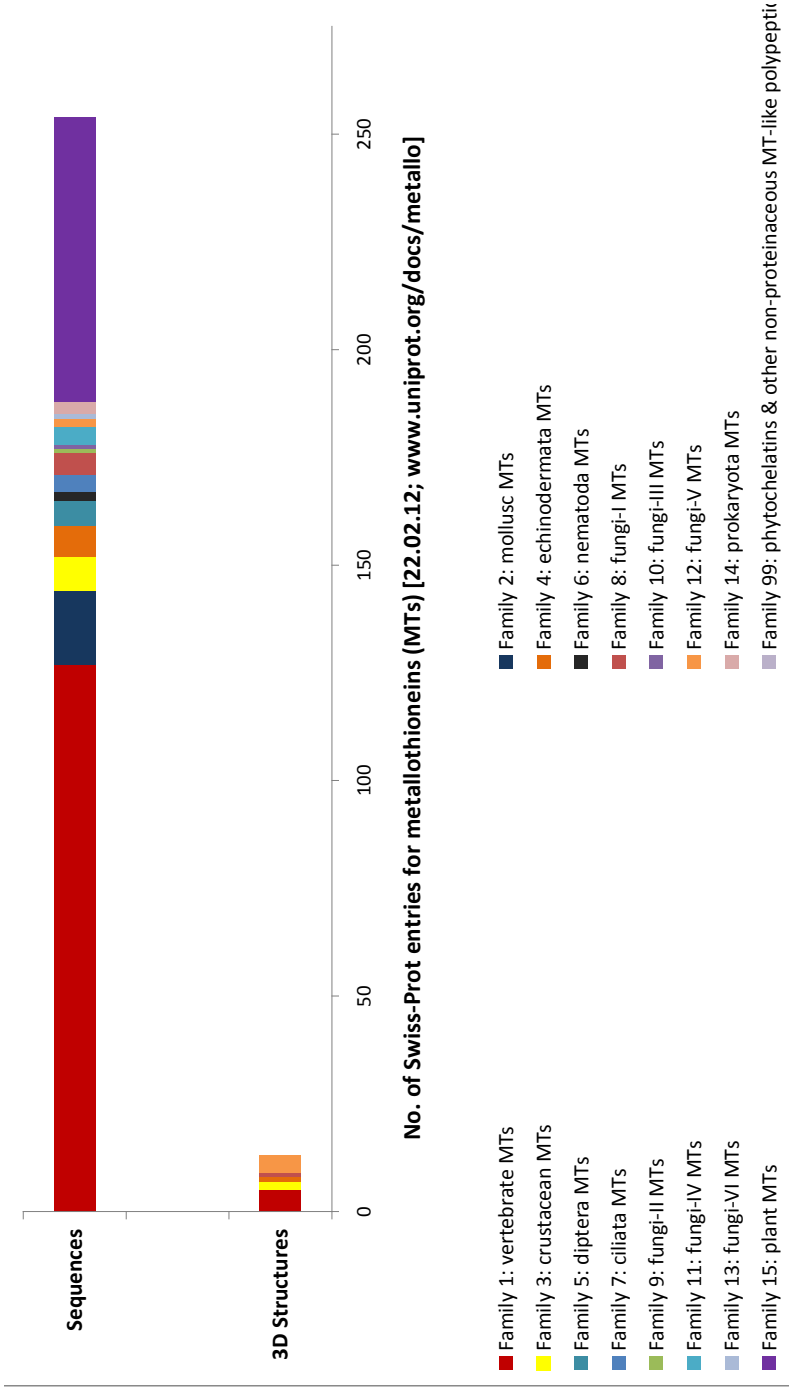


Figure 1.3: Comparison of all metallothionein entries in the Swiss-Prot database. Each family is represented by at least one sequence, apart from Family 99 which currently contains no sequences. Although metallothionein research has been ongoing since the late 1950s, there are still relatively few complete 3-dimentional structures. 3D structures do not necessarily have to be unique, as there are a number of MT sequences with 2 or more 3D structure entries. This means that the overall proportion of sequences compared to 3-dimensional structures is likely smaller than depicted here. Raw data obtained from: http://www.uniprot.org/docs/metallo.txt.

[†] Data obtained from: http://www.uniprot.org/docs/metallo.txt

Although initial studies focussed on MTs from vertebrates, recently a wider array of organisms [87, 88] have been investigated. These later studies have shown that MTs exhibit significant structural diversity between organisms [62].

Structure

Although MTs contain a large proportion of sulphur-containing residues, due to their use in metal chelation, they usually contain no disulphide bonds [68]. Motifs such as **C**x**C** / **C**xx**C** are highly conserved in vertebrate MT sequences [89], with the emergence of similar motifs such as **CC** / **CCC** within other families thought to be caused by convergent evolution influenced by the requirement to form effective metal-binding clusters [90, 91]. **Appendix 1** shows the ubiquity of MTs in a number of organisms, ranging from mice to mouse-ear cress, and the emergence of specialised motifs such as the –**C**P**C**- motif found in human MT-3, determined by mutagenesis to be crucial for functional activity in the brain [89].

X-ray crystallography [92] and NMR with spectroscopically active metal ions ($^{111/113}$Cd [93]/ ^{59}Co [94]) have shown that MTs bind metal ions in clusters exhibiting tetrahedral co-ordination environments [95]. As there are fewer cysteine residues than would be required for individual coordination of each metal ion, MTs utilise bridging sulphur ligands within these clusters (**Figure 1.4**) [96]. Further studies of some MT sequences

(Families 1-4 & 6, **Figure 1.3**) indicate that metal ions are bound in two discrete clusters joined by a linker region [97] (roughly adopting a 'dumb-bell' shape [98] (**Figure 1.4**)). For MTs where two clusters are proposed, the linker region between the two domains varies greatly in length and amino acid composition; with the length of the linker region directly influencing protein stability [99].

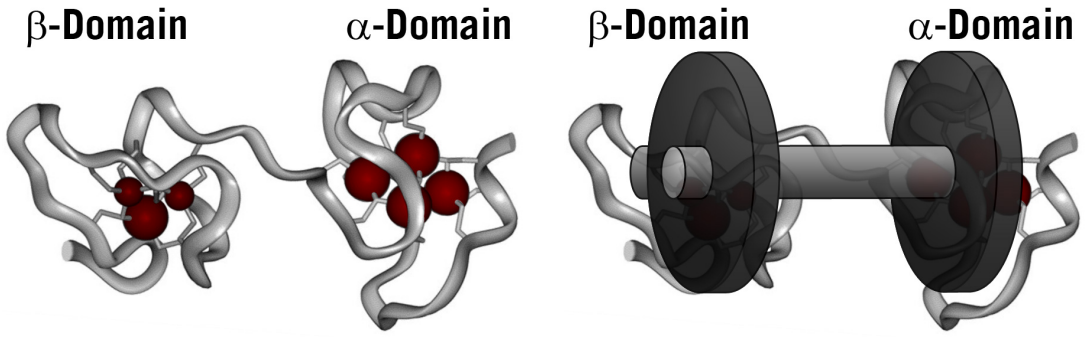


Figure 1.4: Diagram showing the similarity of the two-domain metallothionein to a 'dumb-bell' shape. Overlaid Rat MT-2 modified from [91].

Cysteine residues within a cluster are oriented so that their sulphur atoms are in close proximity, allowing the adoption of stable 6-membered ring conformations (**Figure 1.5**) [100]. Naming convention (based on initial work with vertebrate MTs) designates the clusters as: M_4Cys_{11}, α-domain; M_3Cys_9, β-domain [95] (M representing a divalent metal ion such as Zn^{2+} or Cd^{2+}).

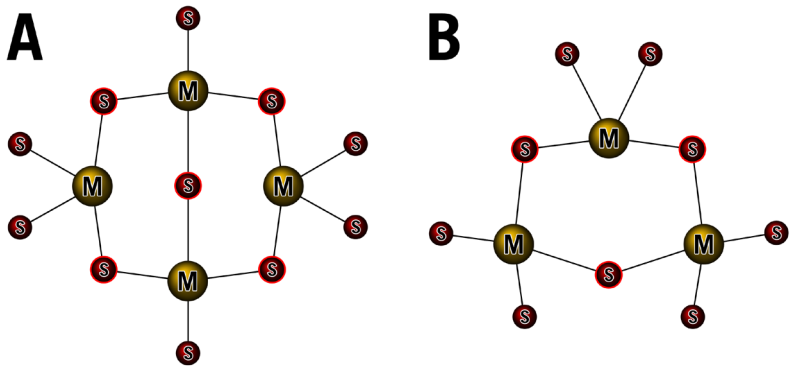


Figure 1.5: Schematic of (**A**) the M_4Cys_{11}-cluster; and (**B**) the M_3Cys_9-cluster found in vertebrate MT [101]. Sulphur atoms indicated in **RED**, Metal ions shown in **YELLOW**. Bridging sulphur ligands shown with an exterior **RED OUTLINE**.

In vertebrate MTs, metal binding is very ordered within the domains [102], providing the metals are physiologically relevant (ie. zinc, copper and cadmium). Metal binding proceeds in a co-operative manner, with preferential filling of sites within the 4-metal α-domain, followed by the 3-metal β–domain [103]. Binding in the α-domain was also found to be more thermodynamically stable than that of the 3-metal β–domain, with the complete release of metal ions in the β-domain before demetallation of the α-domain occurred [91].

Induction

MT gene transcription can be induced by a wide range of stimuli, with detectable intracellular concentrations being accumulated within 1 hour of exposure [65]. As well as zinc, copper and cadmium, metal ions such as silver and mercury are also potent inducers of MT genes [104, 105]. To facilitate a rapid response to changing cellular conditions, MT genes have a number of cis-regulatory elements present in nearby upstream/downstream of the MT-gene promoter [91], including MREs [106] and antioxidant responsive elements [107].

For an MRE to be utilised, a metal sensor protein is required for activation. In mice, this metal sensor was found to be MTF-1 (Metal-regulatory Transcription Factor-1). MTF-1 silencing led to no MT-I/II accumulation in cells, even when challenged with concentrations of metal ions known to induce MT expression in wild-type mice [91]. MTF-1 is a

large protein, comprising six zinc-finger domains [108]. When these sites are populated with zinc, transcription of downstream MT genes is enabled [30]; behaviour which is proposed to be self-regulating [109] (**Figure 1.6**).

As functional MREs have yet to be found in many invertebrates, it is thought that there is less reliance on the control of MT transcription by metal-regulatory transcription factors [110]. This makes the regulation of MTs in invertebrates not easily comparable to that of MT regulation in vertebrates [68].

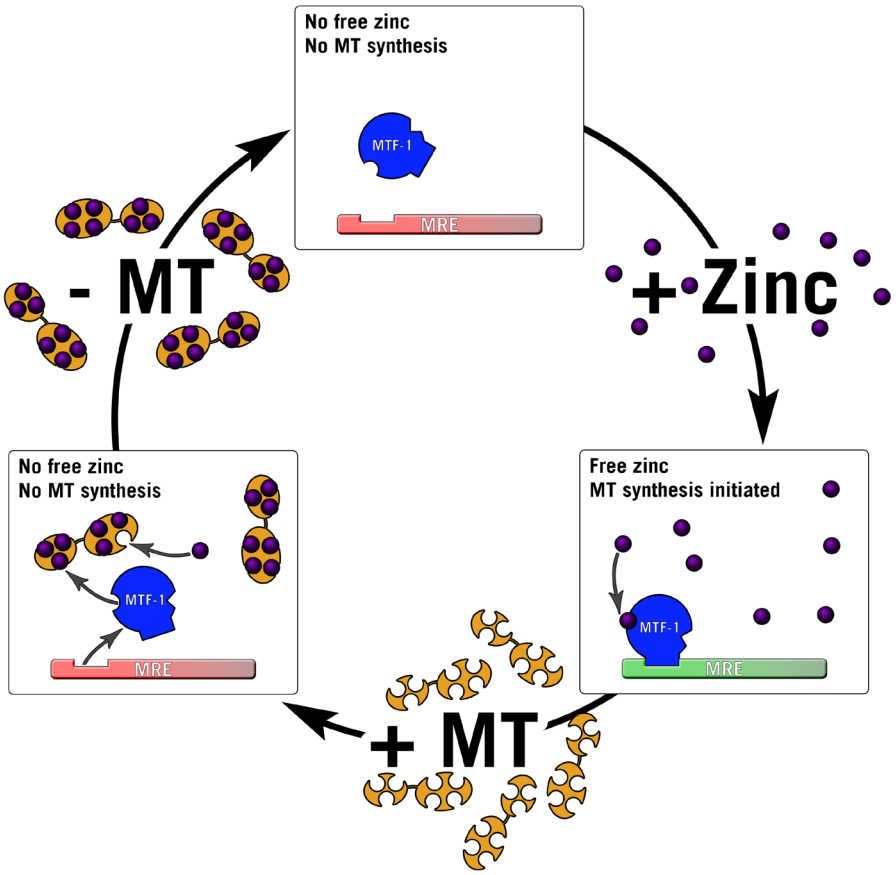


Figure 1.6. Postulated mechanism for control of MT synthesis, modified from [109]. MTF-1 (**BLUE**), Metal-regulatory Transcription Factor-1; MRE, Metal Response Element (**RED** / **GREEN**), Zinc (**PURPLE**), and MT (**ORANGE**). Zinc promotes MTF-1 binding to the MRE DNA sequence, inducing MT synthesis. When available zinc is bound by MT, excess apo-MT removes zinc from MTF-1, causing MTF-1 dissociation from the MRE, and halting MT expression.

1.5 Metallothioneins in invertebrates

Approximately 16 % of MT entries in the ExPASy[‡] database correspond to 'invertebrates' – organisms in Kingdom Animalia that do not contain backbones or spinal columns. The various invertebrate clades for which there are reviewed MT sequences are: mollusca (including snails and mussels), pancrustacea (including the fruit fly, lobsters and crabs), echinacea (including sea urchins), annelida[§] (segmented worms) and nematoda (*C. elegans*). Whereas vertebrate MTs show little variation in primary sequence, there is significant diversity within invertebrate MTs (**Appendix 1**).

Adaptation of invertebrates to their environment

The presence of invertebrates in soil and water-sources continuously exposes them to a wide variety of contaminants, including toxic levels of metal ions [112]. Invertebrates have a large capacity for sequestering contaminant metal ions, for example the earthworm *Lumbricus rubellus* can maintain body burdens of cadmium up to 1 mg/g dry weight [113]; it is thought that the metal-binding ability of MT is crucial for this behaviour [112]. As cadmium excretion proceeds slowly, invertebrates have utilised MTs as an efficient storage mechanism: compartmentalising toxic metals – rendering them biologically inactive - until excretion is complete [114].

[‡] http://www.expasy.org/

[§] Annelids owe their name to the latin *anellus* 'little ring', and are segmented worms found in both aquatic and terrestrial environments [111].

The high metal tolerance observed in invertebrates is postulated to have arisen through two mechanisms: the enhancement of constitutive MT expression; or MT genes adapting to chelate toxic metal ions more efficiently [110]. An example of the former was discovered in the aquatic worm *Limnodrilus hoffmeisteri*. Over the course of 30 years, this worm responded to cadmium pollution (> 500 μg/g Cd in soil) by upregulating expression of an MT-like protein [115]. This trait proved to be genetic in nature, as worms remained resistant to high levels of cadmium after two generations in control soil (≈ 19 μg/g Cd in soil) [116]. If cadmium tolerance across 1-4 generations could have evolved so rapidly in one invertebrate [116], what further insights into metal handling could be learned from studying other invertebrates?

Mollusca MTs

Research into MTs from molluscs has been divided into two groups: snails, such as *Helix pomatia* [117] and *Arianta arbustorum* [118]; and mussels / oysters, such as *Mytilus edulis* [119] and *Crassostrea virginica* [120]. The most widely studied of the terrestrial gastropods mentioned is the land snail *Helix pomatia*. Three distinct MT isoforms have been isolated, not thought to be under the control of an MTF-1 analogue [121]: one responsible for cadmium resistance (CdMT); one for copper homeostasis (CuMT); and a third having a smaller contribution to metal metabolism (Cd/CuMT) [121]. When isolated from the organism, it was found that CdMT (localised largely in the midgut gland) bound 6 equivalents of cadmium, with CuMT binding 12 equivalents of copper(I)

[122]. These two MTs were found to be inert to protein-protein metal exchange, leading to the hypothesis that CdMT protects CuMT by acting as a cadmium sink [122].

With respect to the aquatic molluscs, the *Mytilus* mussel family expresses two subfamilies of isoforms [119], the variable MT-10 (induced by Zn, Cu and Cd) and the less variable MT-20 (induced by Cd) [123]. Each group, characterised by mass, binds seven divalent metal ions [124, 125]. In mussels, MT-10 has a postulated metal metabolism role, and MT-20 that of cadmium detoxification [126]. Of note is the dimerisation of the MT-20 isoform obtained from exposed organisms [127]. The dimerisation is postulated to occur through unique monomer-bridging S-Cd-S bonds between two monomers, which are stable in the presence of DTT [127].

Although mollusca MTs contain histidine residues [121] and a higher percentage of glycine residues than vertebrate MTs [119], their gene structure is closer to vertebrate MTs than some other invertebrate MTs [124].

Pancrustacea MTs

The pancrustcea group contains members from the crustacean and diptera families [128] (**Figure 1.3**). This divides the group into: crustaceans, such as *Homarus americanus* [129] and *Callinectes sapidus* [130]; and diptera, such as *Drosophila melanogaster* [131] and *Orchesella cincta* [132]. *Homarus americanus* (American lobster)

expresses three copper-binding MTs in the hepatopancreas, CuMT-1, CuMT-2 and CuMT-3 [129]. Of interest are the differences in primary sequence between CuMT-1/CuMT-2 and CuMT-3. It is thought that these differences cause CuMT-3 to be the only isoform capable of protein-protein transfer of Cu^+ [129]. The solution structure has been determined for lobster CuMT-1 (PDB 1j5l & 1j5m) indicating the presence of two non-identical β-domains joined by a linker region [133]. In contrast to most invertebrate MTs, but similar to that of vertebrate MTs (PDB 2mrt, 1m0j, 2f5h), there is an α-helix present in the β_C-domain (PDB 1j5l) [134]. With respect to metal-binding, MTs from the lobster subfamily [134] bind 6 equivalents of divalent metal ions, or 6 Cu^+ ions with 1 zinc ion [134].

Since the sequencing of the *Drosophila melanogaster* genome [135], genes encoding five MTs have been discovered: *MtnA*, *MtnB*, *MtnC*, *MtnD* [136] and *MtnE* [137]. The Mtn family of MTs are postulated to be single-cluster proteins, and are somewhat similar to vertebrate and crab MT α-domains [138]. *MtnA* is unique among the five, as it is preferentially induced by copper, not cadmium [139]. Whereas all Mtn isoforms show significant preference for copper binding [140], MtnE is functionally distinct and shows metal-binding preferences that are much more diverse than the other Mtn genes [137]. MtnB, MtnC and MtnD are more closely related to each other (> 60 % sequence identity), and when isolated showed cadmium to be the major species [141]. MtnB appears to be most crucial for cadmium detoxification, as *MtnB*-knockout organisms

become highly susceptible to cadmium toxicity [139]. Unlike many other invertebrates, Mtn genes A-D are under control of dMTF-1 [141], the diptera equivalent of vertebrate MTF-1.

Also within the terrestrial subfamily of pancrustacea is the springtail *Orchesella cincta*. This insect exhibits MT post-translational modification: one encoded peptide product is proteolytically cleaved to form two metal-binding peptides [132].

Echinacea MTs

The Echinacea family includes MTs from sea urchins, with a 3-dimensional solution structure available for *Strongylocentrotus purpuratus* (PDB 1qjk & 1qjl) [87]. Exhibiting the most in common with vertebrate MT regulation, echinoderms express two MTs: MTA (heavily regulated by upstream MREs [142]) and MTB [143]. Both MTs bind 7 divalent metal ions [87], with MTA being obtained from the organism containing cadmium [142]. Of note is the orientation of the domains in MTA, with an N-terminal α-domain being followed by a C-terminal β-domain [87] – the opposite of vertebrate MT (ie [144]). MTA contains a sole phenylalanine aromatic residue [145], uncommon for vertebrate MTs, but not so for invertebrate MTs.

Nematoda MTs

The model organism *C. elegans* expresses two distinct MT isoforms, CeMT1 and CeMT2. Similar to many invertebrates, both CeMT1 and CeMT2 (regulated by a single MRE) are induced heavily in the intestines on ingestion of cadmium [146]. Although both isoforms natively bind 6 divalent metal ions [147, 148], CeMT1 will adopt a Zn_1Cd_6- stoichiometry when cadmium is used to replace bound zinc [149]. Differentially from most invertebrate MTs, CeMT sequences contain histidine residues: CeMT2 containing 1; CeMT1 containing 4 [149]. The hypothesis is that the site in CeMT1 resistant to *in vitro* cadmium exchange is a result of a His_3Cys_1 coordination environment [149] showing zinc preference [150].

CeMT2 shows an *in vitro* [151] and *in vivo* [148] binding preference for cadmium. Knockouts of both *mtl-1* and *mtl-2* genes (encoding CeMT1 and CeMT2 respectively) reduced reproductive fitness of the population (even in the absence of cadmium), and significantly increased their sensitivity to cadmium [152, 153]. These findings have led to the hypothesis that CeMT1 has a role in essential metal metabolism, whereas CeMT2 confers resistance to metal toxicity [149].

Annelida MTs

Stürzenbaum *et al*. [99, 112, 114] identified two earthworm metallothionein (wMT) isoforms in the adult earthworm *Lumbricus rubellus*: wMT-1, highly expressed *in vivo* following exposure to zinc and

copper; and wMT-2, the sole *in vivo* cadmium-responsive MT [113]. A third MT isoform, wMT-3, was identified through *in silico* analysis of expressed sequence tags, obtained from cDNA, reverse-transcribed from cloned mRNA [112].

As is the case with most MT research, the roles and functions of the three isoforms remain enigmatic. However, wMT-1 appears to influence reproduction, as a multivariate statistical analysis showed a positive correlation between wMT-1 concentration and cocoon production rate [154]. wMT-1 is also hypothesised to be better adapted to act as a carrier and donor of essential metals, as its metal clusters are more labile than those of its counterpart, wMT-2 [113, 155]. Localised within the gut epithelium, and thought to act following soil ingestion, wMT-2 is found to accumulate in chloragogenous tissue (analogous to mammalian liver), before being excreted via the nephridium (analogous to mammalian kidneys) [112].

Although the promoter region upstream of *wMT-2* contains 3 MREs, indirect evidence suggests that they are non-functional [156]. Therefore, unlike MTs in *D. melanogaster*, *wMT-2* is unlikely to be under translational control of an MTF analogue [68]. Two N-glycosylation sites have been identified within the wMT-2 sequence, suggesting the ability for post-translational modifications. These could assist in transport of the protein to discrete areas within the cell, called 'cadmosomes' [113], further differentiating between wMT-1 as a carrier and donor of zinc/copper; and wMT-2 as a scavenger/immobiliser of cadmium. wMT-3

sequences were obtained predominantly from libraries created from embryonic tissue, showing high abundance and functional activity during early embryonic development [112]. One sequence, however, was obtained from a library created from adult tissue exposed to cadmium [112]. It is not known how significant this finding may be.

It appears that the two domain vertebrate MT structure (βα) is in reverse orientation for earthworm MTs (**Figure 1.7**): N-terminal α-domain (M_4Cys_{11}-cluster); C-terminal β-domain (M_3Cys_9-cluster) [157]. Although wMT-1 and wMT-2 exhibit > 75 % sequence identity, there is significant variation in the composition of the linker region between the two proposed clusters [113]. wMT-2 exhibits a smaller (4 residue) linker region than wMT-1 (6 residue). A shorter linker region seems to correspond to increased stability, reflected in the greater pH range of metal retention within wMT-2 [113].

1.6 Diversity of annelida MTs

Although few 3-dimensional MT structures have been determined in comparison with known sequences, the role of structure in elucidation of function is well recognised [158-160]. **Figure 1.7** highlights the high cysteine conservation in MT primary sequences across the annelid family. The earthworms can be classified as one of three ecotypes, based on their chosen habitat: Epigeic (1–2.5 mm diameter, **BLUE**), live on the soil surface, feeding on leaf-litter, rarely burrowing and ingesting soil;

Anecic (4–8 mm diameter, **ORANGE**), live in vertical semi-permanent burrows, feeding on the soil surface and leaf-litter; Endogeic (2–4.5 mm diameter, **GREEN**), live in topsoil digging extensive systems of horizontal burrows, feeding on ingested soil [161]. As such, whether an earthworm spends most of its life above (epigeic) or below (endogeic/anecic) the soil surface will affect the extent of interactions with contaminant metals [162, 163]. The highest proportion of sequence variation is found within the proposed linker region. Within the linker region, most variability is seen between the burrowing worms (**GREEN**), compared to those that live on the soil surface.

Figure 1.7: Alignment of all annelid sequences in the non-redundant protein database, from protein BLAST [NCBI]. Cysteine residues indicated in **RED**. Identity of all sequences to *L. rubellus* wMT-2 > 51 %. The three colours represent the earthworm ecotypes: **EPIGEIC**, **ENDOGEIC**, **ANECIC.** The linker region is present between the two domains.

Effect of the linker region between MT domains

The 77 residue wMT-2 contains a linker region between the two domains which is 2 residues shorter than that of wMT-1 (4 residues compared to 6 residues). Experiments varying the length of the linker region between the two domains of vertebrate MT-2 indicated that isoforms engineered with short linker regions are most effective in protecting against cadmium toxicity [164]. Therefore the length of the linker region in

wMT-2 may be crucial for its effectiveness in protection from cadmium toxicity.

There may also be a connection between the susceptibility to proteolytic cleavage and linker region length [165], such as that exhibited by *Eisenia fetida* (**Figure 1.7**, last entry) and *Orchesella cincta*. Through post-translational modification, the *E. fetida* MT-2 forms a stable single-domain protein of 41 residues [166], whilst *O. cincta* generates two single-domain proteins of 42 residues and 30 residues from a single gene product [132]. This evidence supports the theory that the longer the linker region between the two domains, the less stable the protein [91]. It should be noted that plant MT sequences exhibit much longer linker regions (≈ 40 residues [167]) without similar cleavage [168], therefore there is likely another factor effecting whether or not an MT is cleaved *in vivo*.

1.7 Interest in invertebrates for pollution monitoring

One of the motivations for research into the metal-handling of invertebrates is in their use for monitoring pollution, and remediation of pollution [169]. The use of *in situ* organisms for pollution monitoring has several advantages, as they consider both bioavailability and environmentally relevant combinations of toxicants [170]. Organisms that can be used for this sort of pollution quantification are termed

'biomarkers' [169]. The most widely studied invertebrate biomarkers being mussels and earthworms [60, 171, 172].

The largest benefit of using living organisms as biomarkers is the ability to observe changes at molecular, cellular, tissue/organ, and organism level, when contaminants are present at sub-lethal concentrations [173-175]. Earthworms are good biomarkers as they are widely available in almost all soil types and conditions, and have precedent in ecotoxological testing for metals and pesticides [176]. Using biomarkers for pollution monitoring gives very consistent data, with invertebrates shown to be highly responsive to stress conditions [177]. However some difficulties arise when local populations start to adapt to their environment, so care must be taken to screen for genetic differentiation in the populations used for *in situ* monitoring [178]. Currently, a genomics-based approach measures levels of gene expression, enabling quantification of the effects of certain pollutants [179, 180]. The utilisation of biomarkers allows a constitutive overview of the condition of the local ecosystem, allowing both early detection of imbalances within the ecosystem, and monitoring of ongoing ecosystem changes [169, 170]. MT expression within earthworms has been proposed to be an effective biomarker for pollution monitoring [181, 182], however there is no structural information, and little metal-binding information, with respect to the earthworm MTs involved in these genomics-based assays.

1.8 Research motivation and objectives

The primary aim of the research presented in this thesis is to determine the solution structure of wMT-2. Given the lack of 3-dimensional MT structures, especially with respect to invertebrates, significant insights into the conformation of effective metal-binding proteins may be gained from obtaining the first solution structure of an annelid MT. A secondary aim is to recombinantly express and purify wMT-3, the third MT to be identified from earthworms, which is currently uncharacterised. A final aim is to characterise the metal-binding dynamics of wMT-2, with comparison to the other two wMTs where possible. This may help with the elucidation of particular characteristics which make wMTs suited to their proposed *in vivo* roles [112]. It is hoped that information about wMT *in vitro* metal-binding behaviour can be used as an indicator of *in vivo* properties (similar to the systems found in *C. elegans* [151] and *H. pomatia* [121]).

2

Experimental Methods

Chemicals and materials

Unless specified within the text, all reagents were of the highest grade available, containing the lowest contamination from metal ions. The majority of reagents used were obtained from either Sigma-Aldrich (UK) or Fisher Scientific (UK). Ultrapure water was obtained using a Milli-Q™ Integral Water Purification System [Millipore, UK], with a resistance of 18.2 $M\Omega \cdot cm^{-1}$ (25 °C).

The wMT-1, wMT-2 and wMT-3 plasmids

Cloning of all three genes was performed by the Stürzenbaum group (King's College London, UK), generating pET vector plasmids containing the target genes [Novagen]. The chosen vector was pET-29a-c(+) which carries optional N-terminal S•Tag™ / C-terminal His•Tag® sequences[*]. For plasmids containing wMT-1 and wMT-2, both proteins expressed as N-terminal S•Tag fusion proteins. wMT-3 expressed as full length protein, without either tag.

[*] TB076 – pET-29a-c(+) Vectors, http://www.merckmillipore.com/

2.1 Modification of *wMT-1* plasmid

Antibiotic-selective agar plates

Sterilised agar/broth mixture (agar flakes 15 g/L, LB powder 25 g/L, in dH_2O pH 7.5) was supplemented with either kanamycin (50µg/mL final concentration) and/or chloramphenicol (34 µg/mL final concentration) when cool. Kanamycin selected cells that contained the *Kan* resistance gene on the pET-29a plasmid, retaining cells containing the target wMT gene; chloramphenicol selected cells that contained the Cam^R resistance gene, retaining cells containing the pRARE / pLysS plasmids found in Rosetta 2 [Merck Chemicals, UK] competent cells.

Transformation

Competent cells were transformed following the procedure outlined in the Novagen Competent Cell User Protocol. Briefly, 50 µL DH5α library-efficiency[†] [Invitrogen, UK] competent cells were transformed with 2 µL plasmid (immediately once cells thawed). The mixture was cooled in ice for 5 mins, underwent heat-shock for 30 s at 42 °C, and was then cooled in ice for a further 2 mins. 200 µL room-temperature super optimal broth with catabolite repression (SOC) [Invitrogen, UK] was added to the cells, and the entire mixture shaken for 45 mins, 180 rpm, 37 °C. 50-150 µL cells from the outgrowth were spread on each plate, and incubated at 37 °C for 16-18 hrs.

[†] $1x10^6$ to > $1x10^9$ transformants/µg plasmid DNA

Sequencing

Prior to sequencing, individual colonies were picked and incubated in approximately 8 mL antibiotic selective LB media for 16 hrs. Plasmids were extracted using a QIAprep Spin Miniprep Kit [Qiagen, UK] according to manufacturer's instructions. Sequencing was performed by the University of Warwick Molecular Biology Service, on an ABI PRISM 3130xl Genetic Analyser [Applied Biosciences, UK]. Samples to be sequenced contained 9 µL DNA (purified using Qiagen Miniprep Kit), and 1 µL of T7 promoter primer (5'-TAATACGACTCACTATAGGG-3') [Invitrogen, UK].

Correction of wMT-1 point mutation by QuikChange®

To correct the single base mutation in the DNA sequence of wMT-1, a QuikChange® II XL Site-Directed Mutagenesis Kit [Invitrogen, UK] was used with the supplied protocol, using primers:

QC_F 5'-GAA TGT TCG CCA AAC TGC AGG AAG CTT TGC TGT GCT GAT TCC-3';

QC_R 5'-GGA ATC AGC ACA GCA AAG CTT CCT GCA GTT TGG CGA ACA TTC-3'.

T_m 83 °C, site of mutation underlined. Primers were 42 bases long, with 50 % GC content and 3 % sequence mismatch. The method proceeded with an initial 60 s denaturation step (98 °C), 30 cycles of: 5 s denaturation (98 °C); 10 s annealing (55 °C); 20 s elongation (72 °C), and a final elongation step of 60 s (72 °C). Aliquots were transformed into DH5α cells before plasmid extraction and sequencing.

Polymerase chain reaction (PCR) for subcloning of wMT-1

After plasmid extraction, the primers used for the amplification of *wMT-1*,

removing the N-terminal S•Tag were:

Nde1_F 5'-CAT GAA CAT ATG GCT GAT GCA-3';

BamH1_R 5'-CAT GGA TCC TTA GTC ACC ACA-3'.

T_m 58 °C, enzyme sequence recognition sites underlined. Primers were 21 bases long, with 43 % GC content.

PCR was performed according to the New England Biolabs PCR Protocol[‡]. Briefly, 2 µM of each primer and 3 µL of DNA (from Miniprep) were mixed with dNTPs (200 µM final conc.) and 5*x* Phusion® HF buffer, for a 49 µL reaction. 1 µL Phusion® DNA Polymerase [NEB, UK] was added to initiate the reaction. The PCR method proceeded with an initial 300 s denaturation step (94 °C), 35 cycles of: 60 s denaturation (94 °C); 30 s annealing (55 °C); 45 s elongation (72 °C), and a final elongation step of 420 s (72 °C).

DNA gel visualisation

Results were visualised on a 1 % SB-agarose gel following the protocol reported by Brody *et al*. [183]. For visualisation of DNA, GelRed™ [Biotium, UK] was added to each gel before casting. PCR product with DNA loading buffer (5*x*, containing bromophenol blue) was mixed and loaded onto gel. The gel was run in 1x SB buffer at 200 V, 150 mA for 20 mins. DNA bands were visualised in a benchtop UV transilluminator at 302 nm.

[‡] PCR Protocol (M0530), http://www.neb.com/nebecomm/products/protocol631.asp

Digestion and ligation to create subcloned wMT-1

For digestion, information from New England Biolabs[§] was used to determine the optimum conditions for the two chosen enzymes, Nde1 and BamH1. The PCR insert was digested for 3 hrs: 10 µL PCR insert, 10 µL Buffer 3 (10*x*) [New England Biolabs, UK], 2.5 µL BSA (0.8 µg/µL), 2.5 µL Nde1 and 2 µL BamH1, diluted to 100 µL with dH_2O. The vector DNA was digested for 3 hrs: 5 µL pET-29a (from Miniprep), 5 µL Buffer 3 (10*x*) [New England Biolabs, UK], 1.25 µL BSA (0.8 µg/µL), 1.25 µL Nde1 and 1 µL BamH1, diluted to 50 µL with dH_2O.

DNA was precipitated before further processing by mixing with 0.1 vol 3 M sodium acetate, 2 vol 100 % ethanol and cooled at -20 °C for 1 hr. The mixture was centrifuged (20 mins, 4 °C, 13,000 rpm) washed with 70 % ethanol, and centrifuged for a second time (10 mins, 4 °C, 13,000 rpm). The pellet was then left to dry for ≈ 10 mins at room temperature, before being resuspended in 10-50 µL 10 mM Tris-Cl, pH 7.5.

The ligation reaction was performed using a Quick Ligation™ Kit[**] [New England Biolabs, UK]: 7 µL Quick Ligation™ buffer (2*x*), 5 µL cleaned PCR insert, 1 µL cleaned pET-29a vector and 1 µL T4 ligase, were mixed and incubated for 2 hrs at 25 °C. DH5α cells were transformed with the ligated plasmid, and propagated on kanamycin-selective agar plates

[§] Double Digests, http://www.neb.com/nebecomm/tech_reference/restriction_enzymes/double_digests.asp

[**] Quick Ligation, http://www.neb.com/nebecomm/products/protocol2.asp

before outgrowth in selective media and plasmid extraction. The resulting plasmid was confirmed by sequencing to contain the *wMT-1* gene in the correct orientation, and encoding the correct protein sequence.

Preparation of expression strain (Rosetta 2 λDE3 pLysS pRARE2)

Expression strain Rosetta 2 cells [Merck Millipore, UK] were transformed with the sequenced plasmid as previously outlined, onto Kan/Cam-selective agar plates before expression.

2.2 Expression of wMTs

Overnight cultures

Before expression, overnight cultures were grown from single bacterial colonies on the Kan/Cam-selective plates. For these cultures, 100 mL sterile selective LB broth (pH 7.5), with kanamycin (50 µg/mL) and chloramphenicol (34 µg/mL) was inoculated with a single colony, and shaken at 37 °C, 180 rpm for 16-18 hrs.

Expression of unlabelled wMTs

The expression medium used was either standard LB medium, or auto-induction (AI) medium. For the purpose of this thesis, this medium will be referred to as 'AI'. For AI medium, the recipe reported by Studier *et al*. [184] for ZYM-5052 expression media was followed. For the complex medium, three components were required with the following compositions: tryptone/yeast extract (10 g tryptone [Oxoid, UK]; 5 g yeast extract [Oxoid, UK], in 900 mL dH_2O); 20*x* sugar mix (100 g glycerol [Sigma-Aldrich, UK]; 10 g glucose [Fisher Scientific, UK]; 40 g

α-lactose monohydrate [Sigma-Aldrich, UK], to 1000 mL with dH_2O); and 20*x* phosphates mix (0.5 M Na_2HPO_4; 0.5 M KH_2PO_4; 1 M NH_4Cl; 0.1 M Na_2SO_4 [BDH, UK]; 40 mM $MgSO_4$; in 700 mL dH_2O).

For 1 L of AI medium, 900 mL sterile tryptone/yeast extract, 50 mL sterile 20*x* sugar mix, 50 mL 20*x* sterile phosphate mix and 100 µL 1 M $FeCl_3$ were combined. The mixture was then supplemented with antibiotics (50 µg/mL final conc., kanamycin; 34 µg/mL final conc., chloramphenicol), metal ions (500 µM $ZnSO_4$; or 200 µM $CdCl_2$), and 10 mL overnight culture.

For 1 L LB medium, 1 L sterile LB, pH 7.5 was supplemented with antibiotics (as above) and 10 mL overnight culture. Cultures grown in LB medium were induced manually, at OD_{600} 0.6-0.8 with 1 mM IPTG (Isopropyl-D-thiogalactopyranoside). Metal ions (500 µM $ZnSO_4$; or 200 µM $CdCl_2$) were also added at this time.

Standard expression conditions were to shake 400 mL media in 2 L baffled flasks at 180 rpm, 37 °C. Cultures were harvested after 9-10 hrs by centrifugation (5,000 x *g*, 10 mins, 4 °C), drained of supernatant and the cell pellets stored at -20 °C (< 90 days) or -80 °C (> 90 days).

Sonication

Cell pellets were resuspended in 4-8 mL/g wcw sonication buffer (50 mM Tris-Cl; 0.1 M KCl; 3 mM β-mercaptoethanol; in 100 mL dH_2O, pH 8.5), with 1 % TWEEN-20 to aid in solubilisation of cell membranes. To

stabilise the proteins during sonication, 1 mM $ZnSO_4$ / $CdCl_2$ was added to each suspension. The resuspended cell pellet was sonicated with a large-tipped Vibra-Cell sonicator [Sonics, UK] using 1 s pulses requiring 15 watts. Cycles of 30 s with 30 s rest were repeated until mixture was translucent, with no visible cell aggregates. The lysate was then separated from cell debris by centrifugation for 30 mins at 30,000 x *g*, 4 °C.

Expression of $^{13}C/^{15}N$-labelled wMT-2

For production of $^{13}C/^{15}N$-labelled wMT-2, a system using sterile M9 salts and minimal medium was used. A 5*x* stock of M9 salts was produced, requiring $Na_2HPO_4.7H_2O$ (60 g), KH_2PO_4 (30 g) and NaCl (5 g) dissolved in 1 L of Milli-Q water. Additional solutions were prepared using Milli-Q water, with the following compositions: 10 % (w/v) yeast extract, 15 mg/mL $FeCl_3$ solution, 1 M $MgSO_4$, 0.5 M $CaCl_2$, 2 mg/mL thiamine-HCl solution, and 10 mg/mL biotin solution. A metal mix was also required, consisting of: 0.115 g of $ZnSO_4.7H_2O$, 0.0169 g of $MnSO_4.5H_2O$, 0.0290 g of H_3BO_3, 0.0175 g of $CuSO_4.5H_2O$, made to 100 mL with Milli-Q water. For unlabelled stocks of NH_4Cl and glucose, 1 g NH_4Cl was dissolved in 10 mL Milli-Q water; with 20 g D-glucose dissolved in 100 mL Milli-Q water. For labelled stocks of NH_4Cl and glucose, 1 g $^{15}NH_4Cl$ [99 % ^{15}N, Cambridge Isotope Laboratories Inc., UK] was dissolved in 10 mL Milli-Q water; with 6 g D-glucose [99 % ^{13}C, Cambridge Isotope Laboratories Inc., UK] dissolved in 100 mL Milli-Q water.

For (unlabelled) minimal media plates, 3 g of agar dissolved in 80 mL dH_2O and 20 mL 5*x* M9 salts (pH 7.5) was sterilised. When sterile, 200 µL $MgSO_4$, 5 mL glucose, 20 µL $CaCl_2$, 1 mL NH_4Cl, 100 µL $FeCl_3$, 100 µL biotin, 100 µL metal mix, 50 µL thiamine, 100 µL yeast extract, 1 mL NH_4Cl, 5 mL glucose were added. Before the plates were poured, 30 µg/mL kanamycin and 34 µg/mL chloramphenicol were added for antibiotic selectivity.

For the unlabelled overnight media, 2 g of agar dissolved in 73.37 mL dH_2O and 20 mL 5*x* M9 salts (pH 7.5) was sterilised. When sterile, 200 µL $MgSO_4$, 5 mL glucose, 20 µL $CaCl_2$, 1 mL NH_4Cl, 100 µL $FeCl_3$, 100 µL metal mix, 100 µL biotin, 50 µL thiamine, 100 µL yeast extract were added to the medium.

For 1 L of $^{13}C/^{15}N$ labelled minimal media, 2 g of agar dissolved in 733.70 mL dH_2O and 200 mL 5*x* M9 salts (pH 7.5) was sterilised. When sterile, 2 mL $MgSO_4$, 50 mL ^{13}C glucose, 200 µL $CaCl_2$, 10 mL NH_4Cl, 1 mL $FeCl_3$, 1 mL metal mix, 1 mL biotin, 500 µL thiamine, 1 mL yeast extract were added to the media. Antibiotics were added the same manner as for minimal media plates.

Colonies from unlabelled agar plates were grown for 5 hrs in unlabelled LB medium. The resulting culture was then streaked on minimal media plates, and incubated overnight at 37 °C. 100 mL unlabelled overnight M9 medium was inoculated with a single isolated colony from the streaked minimal media plates, and incubated at 37 °C, 14 hrs, 180 rpm.

Cells cultured overnight were harvested at 3,000 x *g* for 10 mins at 4 °C, and washed with 50 mL fresh labelled medium. 1 L of labelled minimal medium was inoculated with the entire washed suspension, and cells were grown at 37 °C, 180 rpm. Cultures were induced manually at OD_{600} 0.8-1.0 with 1 mM IPTG, and supplemented with 200 µM $CdCl_2$.

Cultures were harvested after 10-12 hrs by centrifugation (3,500 x *g*, 20 mins, 4 °C), drained of supernatant and cell pellets were stored at -20 °C (< 30 days). Lysis was performed identically to unlabelled samples.

2.3 Purification of wMTs

Fast protein liquid chromatography (FPLC)

An ÄKTA™ Explorer10 [Amersham Pharmacia] FPLC System was used for chromatography, with columns purchased from GE Healthcare (UK), unless identified in the text. Before chromatography, crude cell lysate was passed through either a 0.2 µm or 0.45 µm syringe-filter [Millipore, UK] to remove small particulates and cell debris. Depending on final sample translucency, the mixture was diluted by addition of 4-8 mL of 20 mM ammonium bicarbonate buffer before separation.

Size-exclusion chromatography (SEC)

To maintain consistency in the retention time of eluates, the injection volume for all SEC experiments was fixed at 4 mL. The separation was monitored at 220 and 280 nm, with eluate collected in 4 mL fractions. The flowrate was reduced for the duration of sample injection to reduce

the increase of pressure caused by the superloop. Standard conditions are indicated in the table below:

Table 2.1. Standard conditions employed in size-exclusion chromatography.

Flowrate (mL/min)	1.0
Injection flowrate (mL/min)	0.3
Injection volume (mL)	4.0
Fraction volume (mL)	4.0
Maximum system pressure (MPa)	0.3
Mobile phase; Buffer	20 mM NH_4HCO_3
Column	HiLoad 1660 Superdex 75pg 120 mL

Ion exchange chromatography

For anion/cation exchange, the columns were used as per manufacturer's instructions [GE Healthcare, UK]. For AnCat chromatography, the anion exchange column was coupled directly to the cation exchange column, giving an effective column volume of 10 mL. Standard conditions are indicated in the table below:

Table 2.2. Standard conditions employed in ion exchange chromatography.

Flowrate (mL/min)	5.0
Injection flowrate (mL/min)	0.3
Injection volume (mL)	Indicated in text
Fraction volume (mL)	4.0 (Flowthrough > 4 mL)
Maximum system pressure (MPa)	0.5
Mobile phase; Buffer A	20 mM NH_4HCO_3
Mobile phase; Buffer B	20 mM NH_4HCO_3, 0.5 M NaCl
Columns	HiTrap Q XL 5 mL **Anion** or HiTrap SP XL 5 mL **Cation**

SDS-PAGE and visualisation by silver-stain enhancement

Two different systems were used during this project: NuPAGE® Novex 4-12 % Bis-Tris minigels [Invitrogen, UK] with SeeBlue® Plus2 Pre-Stained Standard [Invitrogen, UK] and 4*x* NuPAGE® LDS Sample Buffer [Invitrogen, UK]; or Mini-Protean® 4-20 % TGX™ minigels [Biorad, UK] with the Precision Plus Protein™ Dual Colour Standard [Biorad, UK] and 1*x* Laemmli Sample Buffer [Biorad, UK]. The protocol employed is indicated in the figure accompanying the experiment.

Samples of 10 µL were prepared by mixture with sample buffer. Gels were run for either: 35 mins at 200 V (Invitrogen); or 45 mins at 150 V [Biorad, UK].

To enable visualisation of the protein bands after separation, silver-stain enhancement was used. After removal from the gel enclosure, the gel was rinsed thoroughly with dH_2O. The gel was then transferred into fixing solution for 15 mins. The fixing solution comprised 60 mL 50 % acetone, 1.5 mL 50 % TCA and 25 µL formaldehyde (37 %). After fixing, the gel was rinsed and washed for 2 mins in dH_2O, followed by washing for 5 mins in 60 mL 50 % acetone. To enhance the sensitivity of the silver-stain, the gel was transferred into a solution containing 10 mg sodium thiosulphate in 60 mL dH_2O for 1 minute. To stain the bands, the gel was incubated, in the dark, for 8 mins in a solution containing 160 mg silver nitrate and 600 µL formaldehyde (37 %) in 60 mL dH_2O. After rinsing the gel in dH_2O, the developing solution was added to visualise the protein bands. This solution contained 1.2 g sodium carbonate, 25 µL

formaldehyde (37 %) and 2.5 mg sodium thiosulphate in 60 mL dH_2O. To halt the developing process, the gel was quickly drained of all developing solution, and the reaction stopped using a 1 % acetic acid solution.

Inductively coupled plasma-optical emission spectrometry (ICP-OES)

Inductively coupled plasma-optical emission spectrometry or ICP-OES is used to measure the concentration of certain elements within a sample. Samples are analysed following nebulisation and introduction into the centre of a plasma [185]. The plasma is an inert ionised gas (such as argon) at temperatures in excess of 7000 K [186]. The extreme temperature of the plasma will provide enough energy to atomise the sample, dissociating all chemical bonds. Following excitation within the plasma, the individual atomised atoms and ions emit light at wavelengths characteristic of the elements being analysed. The intensity of these lines, when compared to standards of known concentration, can be used to determine the concentration of elements within a sample [185].

Samples for ICP-OES and the calibration standards were prepared using analytical grade 70 % nitric acid stock, diluted to a working concentration of 0.1 M HNO_3 with Milli-Q water. All elemental standards were obtained from Sigma-Aldrich (UK), supplied in 2 % nitric acid. ICP-OES analysis was performed on an Optima 5300DV [Perkin-Elmer, UK], and data processed using the supplied WinLab32 software [Perkin-Elmer, UK].

Mixed element calibration standards were prepared gravimetrically in the range 0.2 ppm - 2.0 ppm for zinc (213.857 nm), sulphur (180.669 nm), cadmium (228.802 nm) and copper (324.752 nm). Six standards were used during this project with metal concentrations of: 0 ppm (blank), 0.2 ppm, 0.5 ppm, 0.7 ppm, 1.0 ppm and 2.0 ppm. This gave a working range broad enough to cover the expected concentrations of these elements in the samples from FPLC.

Samples were prepared either volumetrically or gravimetrically (depending on the level of precision required). For FPLC samples of up to 8 mL, 1 mL of FPLC fraction was diluted with 4 mL 0.1 M HNO_3. For samples with high 220 nm absorbance from FPLC (ie flow-through fractions) the volume of FPLC fraction used was halved.

Cysteine assay

As an alternative to full elemental analysis, the cysteine assay provided a quick way of calculating protein concentration in large numbers of samples. The assay allows the determination of the concentration of free thiols in the sample by UV-vis spectrophotometry. 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB, **Figure 2.1**) reacts specifically with free thiols in solution, quantitatively releasing the nitro-thiobenzoate anion, which exhibits a characteristic absorption at 412 nm (at pH > 8).

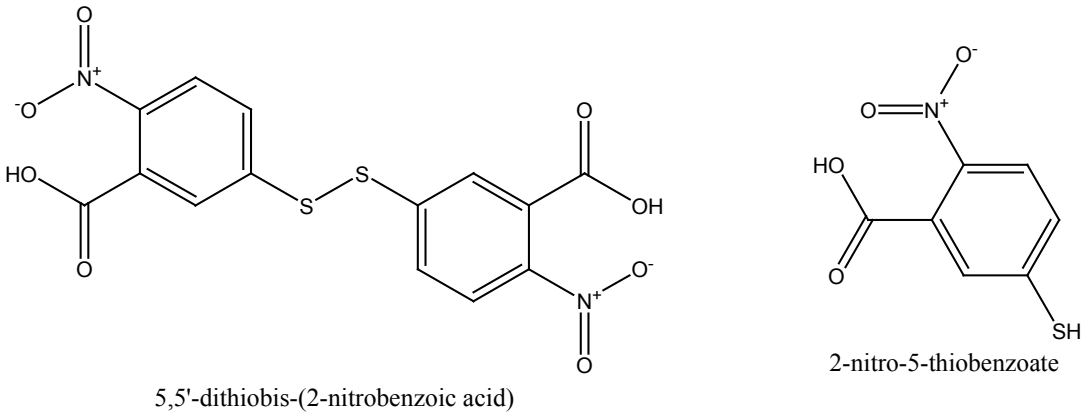


Figure 2.1. Chemical structure of 5,5'-dithiobis-(2-nitrobenzoic acid (DTNB) which after reacting with free thiols in solution, splits to form the (yellow) nitro-thiobenzoate ion.

The concentration of thiols in a sample may be derived from either the molar extinction coefficient of reduced nitrothiobenzene (14,150 $M^{-1}cm^{-1}$ [187]), or a set of standard solutions with known concentrations of cysteine. It was chosen to use a set of standards, as this will correct for any contributions in error from equipment used for the measurements. Standard solutions with known concentration of cysteine were used, and their absorbances at 412 nm plotted with a linear fit [185] as per Equation 1:

$$[Cys] = \frac{A_{412} \bullet D}{S} \qquad \text{Equation I}$$

[Cys] = concentration of free thiols (cysteines), D = dilution factor, S = gradient of linear fit standard solutions.

For the cysteine assay, three stock solutions were required, each prepared with Milli-Q water: 800 µM L-cysteine with 1 mM EDTA; 0.1 M Tris-Cl (pH 7.05) with 1 mM EDTA; 2.5 mM DTNB in 50 mM ammonium acetate (pH 5.00) with 1 mM EDTA. In the case of metallothioneins, high

concentrations of EDTA are required in each solution to fully remove MT-bound metal ions, and generate the measurable free thiol groups.

From the stock solutions, 5 calibration standards were created with cysteine concentrations of 0.0 µM (blank), 13.3 µM, 26.7 µM, 40.0 µM and 53.3 µM. Standards were diluted to 3 mL, and contained 0.167 mM DTNB (final conc.). Samples were prepared in a similar manner, comprising 2.60 mL Tris-Cl solution, 200 µL DTNB solution and 200 µL sample (3 mL total volume). Samples of high concentration were diluted with dH_2O to maintain the sample volume of 200 µL. Samples were incubated for 10-15 mins before their absorbances were measured. Protein concentration in pure samples could be estimated using the linear plot of the absorbances from the calibration standards, and dividing the calculated free thiol concentration by the number of sulphur-containing residues within the protein sequence (see **Appendix 2**).

Thrombin cleavage

For thrombin cleavage, samples were concentrated to approximately 1 mg/mL protein, the optimum concentration as specified by the supplied protocol [Sigma-Aldrich, UK]. Protein concentration was performed using an Amicon® Ultra-4 (3 kDa MWCO) centrifugal filter [Millipore, UK]. Samples were incubated with 110 % of the required NIH units of thrombin for 3 hrs at 25 °C. To remove the cleaved protein from thrombin and other contaminants, samples were syringe-filtered, diluted

to 4 mL with 20 mM ammonium bicarbonate buffer and separated by SEC.

2.4 Mass spectrometry (MS) of wMTs

Samples of wMT were prepared for MS to have concentrations of approximately 30-50 µM. This was performed using either Amicon® Ultra-4 centrifugal filters (3 kDa MWCO) [Millipore, UK], or Vivaspin® centrifugal concentrators (5 kDa MWCO) [Sigma-Aldrich, UK]. To maintain sample volatility, 20 mM ammonium bicarbonate was supplemented with 10 % (v/v) HPLC-grade methanol. Samples were analysed on an HCTultra ETD II ion-trap mass spectrometer [Bruker Daltonics, UK] equipped with an electrospray ionisation (ESI) source, by direct infusion at 240 µL/hr. To obtain apo-protein, samples were acidified with either acetic acid or formic acid to approximately pH 2.0.

Resulting mass spectra were analysed using the supplied Data Analysis Suite [Bruker Daltonics, UK], an average mass spectrum being generated from 0.5-1 min of analysis time. Unless indicated in the text, no smoothing algorithms have been used to generate any of the figures. Deconvoluted spectra were generated from the most intense peak in the raw spectrum.

2.5 Additional purification of wMT-2

Chemical precipitation

After sonication, the crude lysate was mixed with a 10 % streptomycin sulphate solution (0.375 mL/g wcw pellet mass). The mixture was centrifuged at 5,000 x *g* for 20 mins at 4 °C. The supernatant was decanted and cooled in an ice bath whilst volumes of ice cold ethanol:chloroform (100:8) were added dropwise. In the case of wMT-2, three separations of ¾ vol, 1¾ vol and 3 vol were performed. After the required volume of ethanol:chloroform had been added, the solution was incubated at -20 °C for 12 hrs before centrifugation at 5,000 x *g* for 20 mins at 4 °C. The supernatant containing unprecipitated protein from the centrifugation was placed back into the ice bath, and the procedure repeated until the required number of separations had been performed. Precipitated protein at the various volumes were resuspended in 20 mM ammonium bicarbonate buffer for further analysis.

Salt precipitation

A stock of 4 M ammonium sulphate solution was required for salt precipitation. Post sonication, the required concentration of $(NH_4)_2SO_4$ was achieved through additions of small amounts of the stock. After incubation for no more than 5 mins, separation of precipitate was performed at 14,000 rpm for 10 mins at 4 °C on a MiniSpin® centrifuge [Eppendorf, UK]. As the aim was to keep wMT in solution, only contaminant proteins were precipitated. Therefore only the supernatant was required for FPLC separation.

2.6 Nuclear magnetic resonance (NMR) spectroscopy of wMT-2

^{1}H 1-dimensional NMR spectroscopy

The foundation of NMR lies in the magnetic properties of an atomic nucleus, characterised by a spin quantum number (I) [188]. Of the NMR active nuclei, the most commonly used is the spin-½ hydrogen nucleus. Nuclei can exist in $2I+1$ possible spin states, for the spin-½ hydrogen nucleus this means 2 spin states (denoted ½ and -½) can exist (**Figure 2.2**), similar to the two opposing poles of small bar-magnet [189].

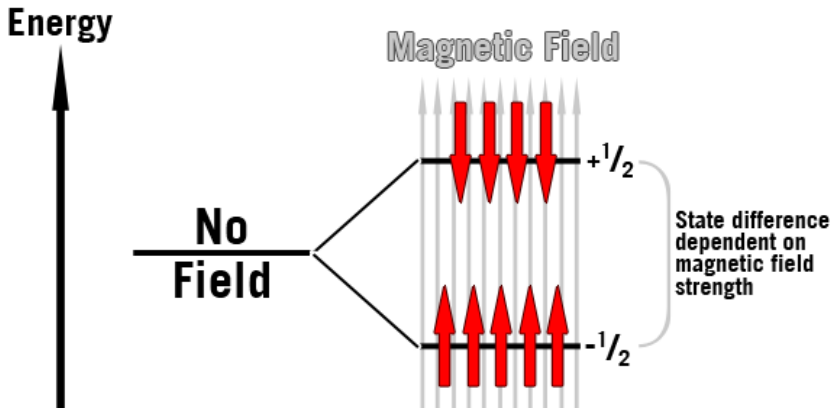


Figure 2.2. Quantum model of NMR, visualising the two spin states as being with/against the externally applied magnetic field. In an externally applied magnetic field, slightly more than half of the nuclei will exist in the lower energy state. Modified from [186].

At thermal equilibrium half of the nuclei will be in the lower energy state, and half will be in the higher energy state. However the equilibrium will shift when the nuclei orient themselves with/against an externally applied magnetic field, as a greater population of nuclei will reside in the lower energy state, aligned with the external magnetic field. It is then possible to excite nuclei using a radio-frequency pulse from the lower energy state to the higher energy state, before equilibrium is re-established through relaxation [189]. The gap between the two energy states is

directly proportional to the experienced magnetic field at each nucleus, and the nuclei resonate between states at characteristic frequencies. The quantum model can be simplified to give a more classical view of the nucleus as a charged particle spinning in a magnetic field, with the ½ and -½ energy states represented as pointing "down" and "up" respectively (**Figure 2.3**) [188].

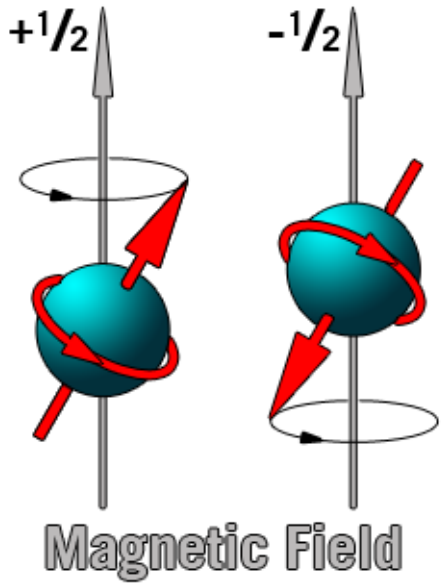


Figure 2.3. Vector model of NMR, visualising the two spin states as pointing "up" or "down" in the externally applied magnetic field.

This use of the vector model allows the visualisation of the magnetic field effect on a number of nuclei. The spinning nucleus causes the direction of magnetisation to rotate around the direction of the magnetic field, this effect is called precession [190]. If we consider the precession of a number of nuclei in the magnetic field, we can see that the average direction of the magnetic field (bulk magnetisation vector) is in the same direction as the external magnetic field (**Figure 2.4**) [190]. When an external radio-frequency pulse is applied, the direction of bulk magnetisation will orient in the x-y plane, before returning to the vertical position. During the return to equilibrium, the rotating bulk

magnetisation vector induces a voltage in the radio-frequency coils and the NMR spectrum is recorded as a free induction decay (FID) signal [189].

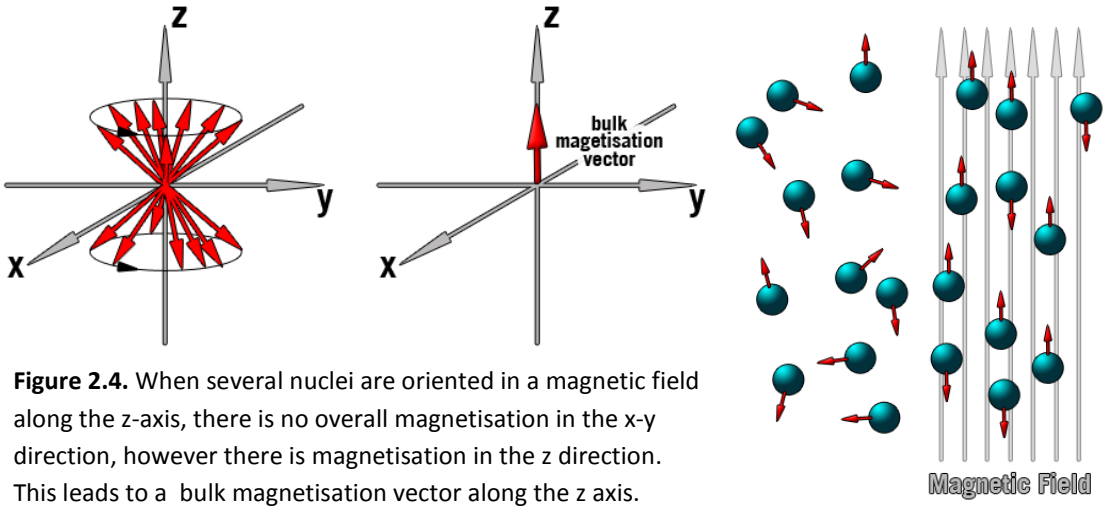


Figure 2.4. When several nuclei are oriented in a magnetic field along the z-axis, there is no overall magnetisation in the x-y direction, however there is magnetisation in the z direction. This leads to a bulk magnetisation vector along the z axis.

As each nucleus possesses its own intrinsic magnetic field, the proximity to other nuclei will slightly affect the strength of the effective magnetic field of neighbouring nuclei. This causes each nucleus to experience a unique magnetic field creating one individual signal for each hydrogen atom [189]. Although each signal will be itself unique, the differences between the effective magnetic field strength are relatively small, and high-resolution NMR instruments are needed to resolve as many individual signals as possible.

All ^{1}H 1-dimensional NMR experiments of wMT-2 (500 µM) were performed on an AV II 700 [Bruker Biospin, UK] equipped with a TCI cryoprobe and VT module. Sample conditions were 20 mM ammonium bicarbonate, pH 6.9, 10 % D_2O, in a 3 mm NMR tube. Spectral conditions

were: zgesgp pulse program; 298 K; 256 scans; pulse length ≈ 8.00 µs (p1); 13.0 ppm spectral width centred on the water peak ≈4.6 ppm.

Variable Temperature ^{1}H 1-dimensional NMR spectroscopy

Cd-wMT-2 was incubated for 5 mins at temperatures between 278-308 K on an AV II 700 [Bruker Biospin, UK]. Sample conditions were identical to those used for ^{1}H 1-dimensional NMR experiments.

[^{1}H, ^{1}H] 2-dimensional NMR spectroscopy

To create a 2-dimensional NMR spectrum, a number of 1-dimensional NMR experiments are performed with different periods of time between excitation and detection [190]. After excitation, the nuclei are allowed to spin freely (evolution time), labelling the magnetisation with the frequency of the nucleus. The magnetisation can then be transferred via either scalar coupling or dipolar interaction (NOE) [189]. Before detection, the magnetisation is labelled with the frequency of the second nucleus. The experiment will generate a cross-peak at the position in each dimension of the two nuclei that exchanged magnetisation (**Figure 2.5**) [189].

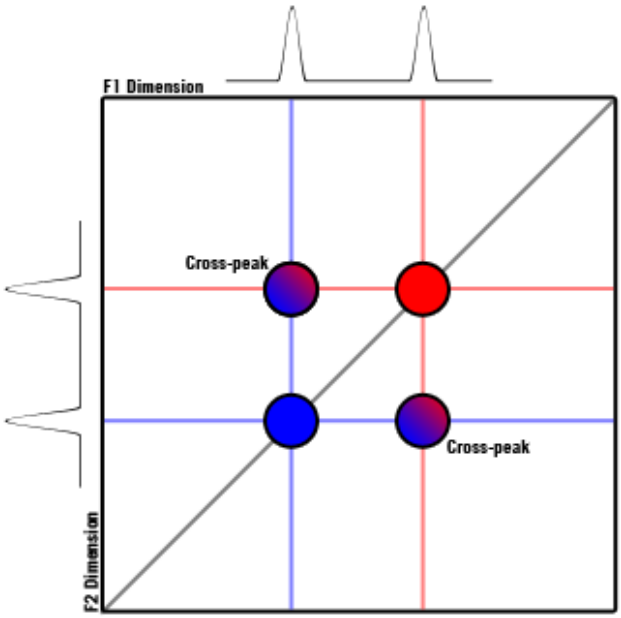


Figure 2.5. Simplified schematic of 2-dimensional NMR. Magnetisation is transferred between the two peaks, generating two cross-peaks at the specific frequencies of the two nuclei.

The benefit of 2-dimensional NMR is the spreading of the spectrum in both the x and y direction, helping to overcome the problems of overlapping signals found in 1-dimensional NMR experiments [189].

All ^{1}H 2-dimensional NMR experiments of wMT-2 (500 µM) were performed using the same instrument as for the ^{1}H 1-dimensional NMR experiments. TOCSY experiments were performed with a mixing time of 60 ms (d9), and NOESY experiments were performed with a mixing time of 60-120 ms (d8). Sample conditions were 20 mM ammonium bicarbonate, pH 6.9, 10 % D_2O, in a 3 mm NMR tube. Spectral conditions for TOCSY and NOESY were: 298 K; 48 scans; pulse length ≈ 8.00 µs (p1); 16.0 ppm spectral width centred on the water peak ≈4.6 ppm. The pulse programs used were noesyesgpph [191] for NOESY, and mlevesgpph [192] for TOCSY.

^{13}C and ^{15}N NMR spectroscopy

Before NMR, the labelled protein was buffer exchanged using Amicon® Ultra-4 centrifugal filters (3 kDa MWCO) [Millipore, UK], with five steps of concentration and regeneration of volume with 20 mM ammonium bicarbonate, pH 6.8. ^{13}C and ^{15}N 2-dimensional and 3-dimensional NMR experiments (HNCA, HN(CO)CA, HNHA, [^{1}H, ^{1}H, ^{15}N] TOCSY-HSQC, [^{1}H, ^{15}N] HSQC) of double-labelled wMT-2 (500 µM) were performed on an AV II 700 [Bruker Biospin, UK]. For 3-dimensional NMR experiments, a number of 2-dimensional NMR experiments are performed. Whilst the nuclei are spinning freely (evolving) during a 2-dimensional NMR experiment, the magnetisation is transferred to a third nucleus and labelled by a third frequency [188]. The third dimension, containing 2-dimensional 'slices' of an NMR spectrum allows further reduction of signal overlapping problems with a spectrum (**Figure 2.6**). As there is an additional evolution time, a significant amount of time is required to perform 3-dimensional NMR. Sample conditions were 20 mM ammonium bicarbonate, pH 6.8, 10 % D_2O, in a 3 mm NMR tube.

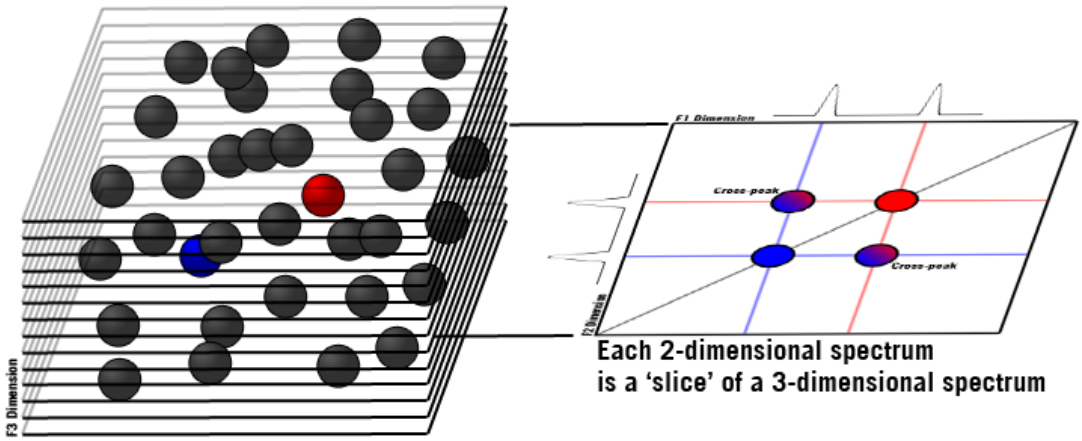


Figure 2.6. Simplified schematic of 3-dimensional NMR. Using a third dimension allows 2-dimensional 'slices' to be taken of a spectrum, minimising signal overlap.

Spectral widths for HNCA and HN(CO)CA experiments were: 13.0 ppm (centred at 4.6 ppm), 36 ppm (centred at 121 ppm), 64 ppm (centred at 59.5 ppm) in the hydrogen, nitrogen and carbon dimensions respectively. The HNCA spectrum was acquired using the hncagpwg3d [193] pulse program for a total of 16 scans, with 2,048 *x* 36 *x* 64 data points in the H, N, C dimensions. The HN(CO)CA spectrum was acquired using the hncocagpwg3d [194] pulse program for a total of 32 scans, with 2,048 *x* 36 *x* 64 data points in the H, N, C dimensions.

Spectral widths for HNHA experiments were: 13.0 ppm (centred at 4.6 ppm), 13 ppm (centred at 4.6 ppm), 36 ppm (centred at 121 ppm) in the H, H (2^{nd}) and N dimensions, respectively. The HNHA spectrum was acquired using the hnhagp3d [195] pulse program for a total of 16 scans, with 2,048 *x* 128 *x* 40 data points in the H, H (2^{nd}), N dimensions.

Spectral widths for [^{1}H, ^{15}N] HSQC experiments were: 13.0 ppm (centred at 4.6 ppm) and 45 ppm (centred at 121 ppm) in the H and N dimensions, respectively. The [^{1}H, ^{15}N] HSQC spectrum was acquired using the hsqcetf3gpsi [196] pulse program for a total of 16 scans.

Spectral widths for [^{1}H, ^{1}H, ^{15}N] TOCSY-HSQC experiments were: 13.0 ppm (centred at 4.6 ppm), 36 ppm (centred at 121 ppm), 13 ppm (centred at 4.6 ppm) in the H, N and H (2^{nd}) dimensions, respectively. The [^{1}H, ^{1}H, ^{15}N] TOCSY-HSQC spectrum was acquired using the mlevhsqcetf3gp3d [197] pulse program for a total of 16 scans, with 2,048 *x* 40 *x* 128 data points in the H, N, H (2^{nd}) dimensions.

^{111}Cd 1-dimensional NMR spectroscopy

A stock solution of labelled cadmium was produced from solid 96.4% ^{111}CdO [Cambridge Isotope Laboratories Inc., UK], and dissolved in 1 M HCl. 500 µM wMT-2 was incubated with 10 mM DTT (final conc.) for 1 hr in an inert atmosphere (N_2). The sample pH was reduced to 1 with 1 M HCl, and applied to a PD-10 column [GE Healthcare, UK] equilibrated with 0.1 M HCl. The first 3.5 mL of eluate was collected (under constant flow of N_2), into a vessel containing 10 equivalents of $^{111}CdCl_2$ solution, to ensure that all metal–binding sites were filled with metal ions. The pH was raised to pH 7.5 with 1 M deuterated Tris base, and the solution was concentrated by centrifugation to 500 µL. 10 % D_2O was added to the sample before analysis in a 5 mm Shigemi tube [Shigemi, Japan] on a

DRX-500 spectrometer [Bruker Biospin, UK]. A broad band observe (BBO) probe allowed direct observation of ^{111}Cd nuclei. Spectra were acquired for a total of 98,304 scans, with a spectral width of 264 ppm centred at 475 ppm. Concentrations of the elements in the sample were confirmed after NMR data acquisition by ICP-OES.

All data was processed using TopSpin [Bruker Biospin, UK]. 1-dimensional NMR were processed Gaussian window multiplication, underwent Fourier transformation with 64k real data points and followed by phase correction and baseline correction. 2-dimensional NMR were processed using the QSINE window multiplication, underwent Fourier transformation with 2k *x* 2k real data points (^{15}N-HSQC with 2k *x* 1k real data points). 3-dimensional NMR were processed using the QSINE window multiplication, underwent Fourier transformation with 2k *x* 256 *x* 128 real data points (HNHA with 2k *x* 512 *x* 512 data points; TOCSY-HSQC with 2k *x* 128 *x* 512 real data points).

2.7 Structure calculations

The CcpNmr Analysis ver. 2.2.2 from The Collaborative Computing Project for NMR (CCPN) [University of Cambridge, UK] was used to display and assign all spectra. From the manually assigned NOESY spectrum, upper and lower distance restraints were obtained through CcpNmr Analysis, before exporting to Combined assignment and

dynamics algorithm for NMR applications (CYANA, v2.1) format. To ensure the maximum amount of unambiguous distance restraints, the NOESY peaks were also exported and automatically assigned by CYANA [198, 199]. These two lists underwent a 'distance modify' function, to remove redundant restraints. Torsional angles were generated from the assigned data, with the respective ^{3}J-HNHA module in CcpNmr Analysis. For metal restraints, a limit of 4.05 - 8.5 Å (lower distance limit – upper distance limit) for cysteine sulphur atoms (within a cluster), and 4.50 - 4.70 Å for the cadmium ions within a cluster were used. A 2.60 Å bond length was set for Cd-S bonds.

From 200 calculated structures, the 10 containing the lowest target function were taken forward. Distance restraints were refined until the maximum deviation of non-metal restraints was < 0.6 Å, and the target function was < 10.00 Å.

2.8 Metal-binding studies

Acid competition reactions

Mass spectrometry conditions were identical to those outlined in Section 2.4. Aliquots of stock wMT solutions were diluted to working concentrations of between 30-50 µM with 20 mM ammonium bicarbonate, before being acidified with either acetic acid or formic acid. pH was measured after mass spectrometry to avoid contamination of the sample with K^{+} from the pH probe. Approximately 100 µL of sample was

required to record a mass spectrum for approximately 2 minutes, with the remainder exceeding the minimum volume requirements for the pH probe (≈ 50 µL).

pH of ½ dissociation ($pK_a^{1/2}$)

For the analysis by mass spectrometry, the estimations were performed through analysis of the magnitude of the deconvoluted peaks of the 5+ charge state. For analysis by ICP-OES, sample aliquots were allowed to equilibrate for 1 hr at the desired pH (between pH 2.0-8.5). After equilibration, wMT-bound metal ions were separated from free metal ions using a pre-equilibrated PD-10 column [GE Healthcare, UK] at the same pH. The protein fraction (3.5 mL) and salt fraction (6.0 mL) were collected. Due to the potential for low concentrations, all samples were diluted by 1.5 to generate the minimum required volume for ICP-OES. The percentages of metal ions bound to wMT at each pH were determined with respect to the total metal present in both fractions.

Cadmium exchange

Purified Zn-wMT-2 was incubated for 1 hr with 0-7 and 10 equivalents of $CdCl_2$. Free metal was separated from protein-bound metal using a PD-10 desalting column [GE Healthcare, UK]. The first 3.5 mL of eluate (protein fraction) and the next 6.0 mL of eluate (salt fraction) were collected. Analysis by ICP-OES gave the proportions of each metal remaining either bound to wMT-2 (protein fraction), or displaced and in

solution (salt fraction). A sigmoidal fit from the data was performed using Origin Pro 8 [OriginLab Corporation].

Reaction with EDTA

A sample of wMT-2 was concentrated to approximately 40 µM before addition of 40 µM EDTA (1:1 protein, 1:7 metal). The sample was injected over the course of 90 mins, with raw spectra being averaged ± 30 s at each time period, to enable averaging over 60 s total. Mass spectrometry conditions were identical to those reported in Section 2.4.

Reaction with 5F-BAPTA

Samples of wMT-2 were buffer exchanged into 10 mM Tris-Cl (pH 8.1) by three steps of concentration and redilution using Amicon® Ultra-4 centrifugal filters (3 kDa MWCO) [Millipore, UK]. A stock of 30 mM 5F-BAPTA [Molecular Probes, UK] was used for competition experiments, stored at 4 °C. 450 µL of sample was incubated with 50 µL D_2O and 50 µL 5F-BAPTA stock (3 mM final conc.) for 12 hrs. Direct observe ^{19}F 1-dimensional spectroscopy was performed on a DRX-400 spectrometer [Bruker Biospin, UK], fitted with a QNP probe-head operating at 375.91 MHz. Proton decoupled spectra were acquired for a total of 24,576 scans, with a spectral width of 200.229 ppm centred at -50 ppm.

Concentrations of the elements in the sample were determined by ICP-OES after NMR data acquisition. Calculations of apparent stability constants (K_{MT}) were performed using the published procedure [146, 200].

2.9 Creation of Figures

Figures based on PDB coordinates were created using: RasMol 2.7.5 (atoms and bonds), JMol 12.2 (cartoon backbone conformation), and MolMol 2K.2 ('sausage' plots). NMR experimental data were exported from CcpNmr Analysis ver 2.2.2.

3

Isolation and identification of wMT-1, wMT-2 & wMT-3

3.1 Introduction

The first DNA sequences of earthworm metallothioneins (wMTs) were reported in 1998 by Stürzenbaum, Kille and Morgan [99]. Genes for two distinct isoforms were sequenced, nominally wMT-1 (**Figure 3.1**) and wMT-2 (**Figure 3.2**). Electron probe x-ray emission analysis of 'cadmosomes' (discrete vesicular compartments enclosed within a lipid bilayer) within earthworm cells, revealed a Cd:S ratio of 1:3. Native protein was obtained from earthworms by size exclusion chromatography, coupled with anion exchange chromatography at pH 8.1 [99].

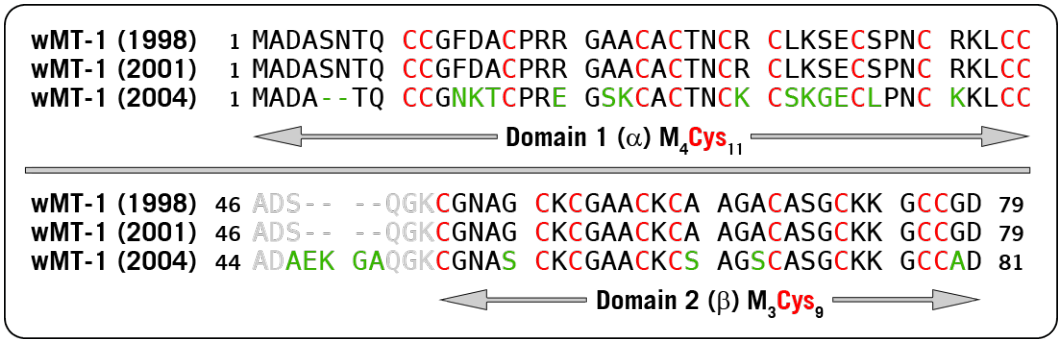


Figure 3.1. Aligned sequences of wMT-1 from all reports publishing wMT sequences [99, 112, 113, 157]. The proposed two domain structure is indicated, with potential metal chelating cysteine residues coloured **RED**, linker region is in GREY. Residues that are divergent to the construct in this thesis are coloured **GREEN**. Greater variation is found in sequences from the 2004 study, as they were deduced from expressed sequence tags.

In 2001, genes for both wMT-1 and wMT-2 were cloned into pET-29a vector DNA. This construct recombinantly expressed a fusion protein:

wMT with an N-terminal S•Tag [113]. Through titrating equivalents of cadmium ions into a solution containing only apo-protein (monitored by UV-vis spectrophotometry at A_{250}, the S-Cd absorbance wavelength), the stoichiometry for both S•Tag-wMT-1 and S•Tag-wMT-2 was estimated to be 6 cadmium ions per protein molecule [113]. Recombinant S•Tag-wMTs were isolated by an initial heating step, followed by cation exchange chromatography at pH 7.4 and size-exclusion chromatography.

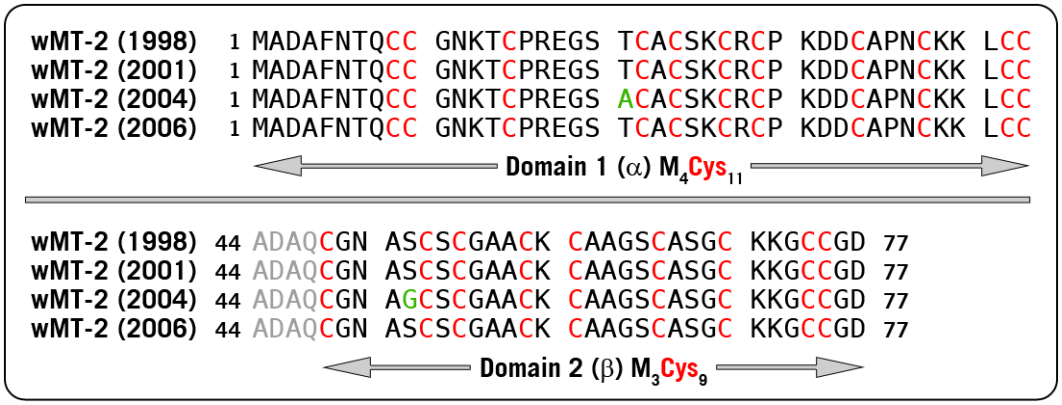


Figure 3.2. Aligned sequences of wMT-2 from all reports publishing wMT sequences [99, 112, 113, 157]. The proposed two domain structure is indicated, with potential metal chelating cysteine residues coloured **RED**, linker region is in GREY. Residues that are divergent to the construct in this thesis are coloured **GREEN**. Greater variation is found in sequences from the 2004 study, as they were deduced from expressed sequence tags.

A third earthworm MT gene, *wMT-3* (**Figure 3.3**), was reported in a 2004 study [112]. However, the corresponding wMT-3 protein was not isolated. X-ray microanalysis showed that both cadmium and sulphur were co-localised within earthworm cells, and held in compartments that were devoid of calcium, phosphorus, zinc and lead. This indicated that within the earthworm, cadmium could be bound solely to MT [112].

Figure 3.3. Derived sequence of wMT-3 [112], reported in 2004. Cysteine residues coloured **RED**.

In 2006, Stillman *et al*. published the first studies of Cd-S•Tag-wMT-2 using mass spectrometry [157]. S•Tag-wMT-2 was determined to bind seven cadmium ions as the major species, revising observations from 2001. S•Tag-wMT-2 was recombinantly expressed as a fusion protein from similar plasmids to those used by Kille *et al.* in 2001. In the 2006 study, the sole technique used for purification was size-exclusion chromatography [157].

In 2008, work within the Blindauer group (performed by Ms Ilka Thiel) was partially successful in the recombinant expression and purification of S•Tag-wMT-2 in the presence of cadmium. The construct was obtained from the Stürzenbaum group, exhibiting the N-terminal S•Tag used in previous work by Stillman *et al*. [157]. The S•Tag was utilised during purification, by passing crude cell lysate through an S•protein affinity column. This proved to be very inefficient, as a significant amount of S•Tag-wMT-2 was not retained and passed through the column (determined by Western blot against S•Tag polyclonal antibodies – data not shown).

Attempts to express S•Tag-wMT-2 solely in the presence of zinc were shown by mass spectrometry to be unsuccessful, as contamination by cadmium during workup caused mixed metal species to be observed. LC-MS was used to obtain mass spectra of both S•Tag-wMT-2 and, for the first time, wMT-2 with the tag cleaved. However, approximately half of the protein was observed with up to five N-terminal residues (-GSMAD-)

not present, making unambiguous identification of individual metal species problematic. NMR experiments were performed on the S•Tag-wMT-2 mixed metal species, however due to significant signal overlap in the crucial fingerprint region (8.6-7.8 ppm), the quality of the sample was deemed insufficient for a structural investigation. Although this work met with limited success, it was thought that with optimisation of the purification protocol, sufficient protein could be isolated to enable structural studies of wMT-2.

3.2 General expression strategy

The creation of a protocol for obtaining high yields of pure untagged wMT was a major goal of this project. Although previous studies had obtained wMTs both recombinantly (with S•Tag) and natively from earthworms, recombinant expression was perceived as the only viable way of isolating sufficient protein for structural studies. It should be noted that some experiments described within this Chapter concerning wMT-1 were performed in collaboration with Ms Maike Hansen. Where appropriate, credit has been given for the figures created from collaborative experiments. All work with wMT-2 and wMT-3 was performed independently by myself. The schematic in **Figure 3.4** provides a graphical representation of the work presented in this chapter.

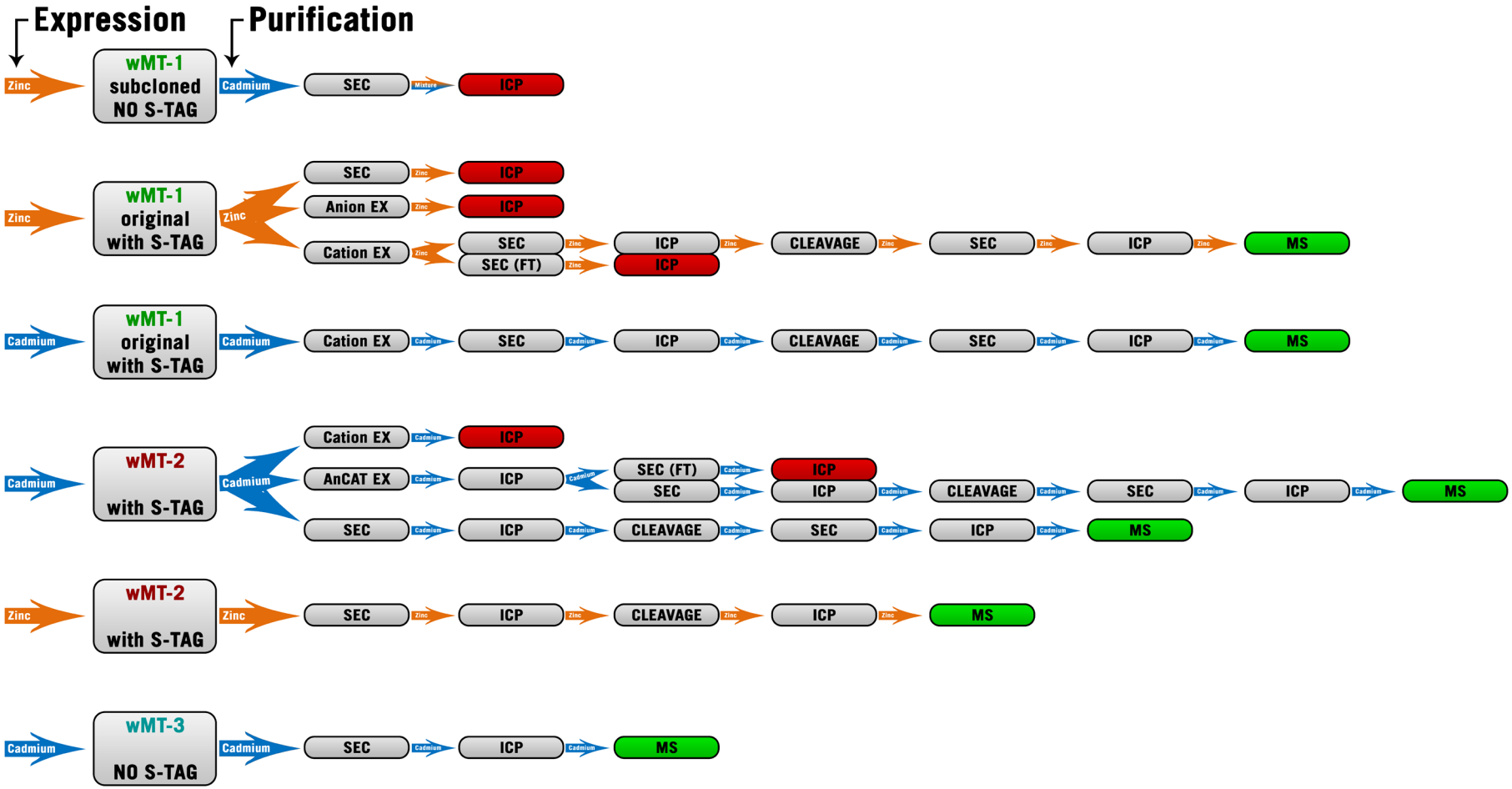


Figure 3.4. Purification procedures used to obtain purified earthworm metallothionein. The metals supplemented at expression and purification are indicated where appropriate. The coloured stages at the end of each step are the point at which that route was not pursued further. SDS-PAGE gels were usually run as a final confirmation, after each chromatography step.

3.3 Working with tags

The original constructs containing the *wMT-1* and *wMT-2* genes used in the present work (obtained from the Stürzenbaum group) were within a pET vector construct; each protein being expressed with an N-terminal S•Tag. Both wMT-1 and wMT-2 sequences were identical to that reported in 1998 and 2001, but different from the sequence deduced from expressed sequence tags in 2004 (see **Figure 3.1**).

Tagged proteins are thought to have many advantages, from manufacturers' claims of 80 % protein recovery [201], to increasing solubility and stabilising protein folding [202]. The S•Tag does not aid in protein solubility, but is mainly used for affinity purification. However, the quantity of pure wMT-2 obtained by Ms Thiel through exploitation of the S•Tag during purification was insufficient for metal-binding or structural studies due to significant losses of protein during cleavage.

For removal of the S•Tag, the protease thrombin was required [201, 202]. Cleavage of protein introduces an undesirable further purification step, with potential for non-specific cleavage. This procedure was followed by Ms Thiel, with limited success, as yields of cleaved protein were insufficient for metal-binding studies. Leaving the construct unaltered was considered, however for metal-binding studies, and structural work, there is a strong preference to work with full length untagged protein. This is due to the difficulty in determining if the S•Tag

is influencing metal binding behaviour; the presence of 27 additional residues also makes NMR assignment more difficult. Furthermore, previous work within the group using full-length untagged MTs had proved to be effective for metal binding studies [151, 168]. Taking the previous work into account, it was decided to subclone the *wMT* genes to remove the S•Tag sequence, starting with wMT-1.

3.4 Sequencing and subcloning of wMT-1

The plasmid containing S•Tag-wMT-1 was propagated in DH5α cells and sequenced using T7 promoter primers. After sequencing, a single nucleotide point mutation within the sequence for wMT-1 was discovered at position 65 (**Figure 3.5.** CYAN).

```
wMT-1   1 MKETAAAKFE RQHMDSPDLG TLVPRGSMAD ASNTQCCGFD ACPRRGAACA CTNCRCLKSE
wMT-1  61 CSPNYRKLCC ADSQGKCGNA GCKCGAACKC AAGACASGCK KGCCGD  106
             C
```

Figure 3.5. Obtained sequencing data of wMT-1. Cysteine residues are coloured **RED**, linker region is in GREY. The point mutation, encoding a tyrosine residue is circled in CYAN with the cysteine of the correct sequence indicated below. The entire S•Tag sequence is coloured GREEN.

The single mutation created a T**A**C codon, encoding for tyrosine, rather than the reported cysteine (encoded by a T**G**C codon) [99, 113]. This point mutation was significant, as this cysteine is thought to be utilised by the protein in metal-binding and cluster formation.

Quikchange Mutagenesis was performed using the protocol supplied, outlined in **Chapter 2.1**. When the corrected sequence was confirmed, the

plasmid was ready for sub-cloning. The next consideration was which restriction enzyme cleavage sites to engineer within the PCR primers. The S•Tag-wMT-1 construct already had an Nde1 restriction site at the start of the S•Tag DNA sequence, however the C-terminus of the *wMT-1* gene provided more flexibility. Based on the downstream sequence, and the enzymes previously employed successfully within the Blindauer group, it was elected to use high fidelity BamH1 as the second restriction enzyme. PCR was performed to amplify the *wMT-1* gene, before double-digestion of both the vector and insert with Nde1 and BamH1. DNA was separated by electrophoresis on a 1 % agarose sodium boric acid gel [183] supplemented with GelRed for visualisation under UV, shown in **Figure 3.6**.

The PCR insert (the *wMT-1* gene with restriction sites) ran at approximately 260 bp, which corresponds with the 239 bp elongation expected from the primer design and *in silico* analysis using Serial Cloner v2.1. The double-digested plasmid ran as a clean single band between 2,100 bp and 8,000 bp, in agreement with the calculated size of the digested vector (5,274 bp).

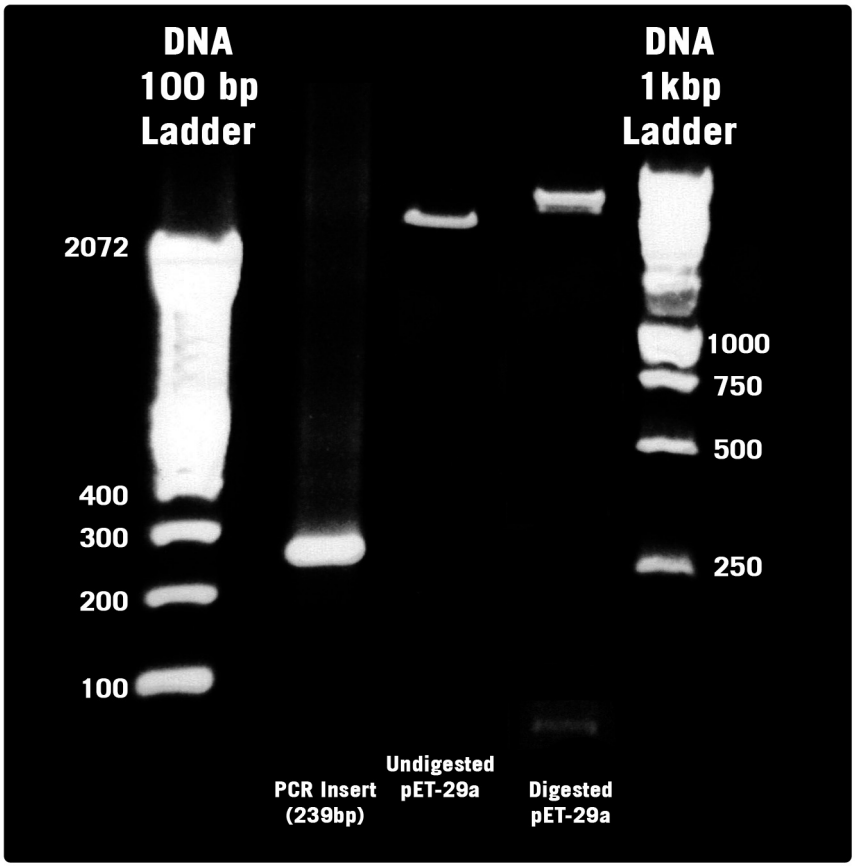


Figure 3.6. 1 % agarose DNA Gel ($Na_2B_4O_7$, pH 8.8) of: the wMT-1 PCR product, the undigested pET-29a DNA, and the digested pET-29a (Nde1/BamH1). A faint band below 100 bp in the digested plasmid is the cut MCS between Nde1 and BamH1. 1 % agarose gel visualised using GelRed.

There was also a faint band, < 100 bp, which could correspond to the removed portion of vector (97 bp), between the Nde1 and BamH1 restriction sites. It should be noted that the undigested plasmid ran slightly faster than the digested plasmid due to being supercoiled and therefore not linearised.

After ligation, the plasmid was propagated in DH5α cells. The cells were then grown on kanamycin-selective LB-agar plates, before plasmid extraction. Plasmids containing the corrected *wMT-1* gene without N-terminal S•Tag were confirmed by subsequent sequencing

(**Appendix 3**). The plasmid was extracted as before, and transformed into the *E. coli* expression strain Rosetta 2 (λDE3) pLysS.

A schematic outlining the general characteristics of the pET expression system is shown in **Appendix 4**. The Rosetta 2 strain is classified as: λDE3, containing a chromosomal copy of the T7 RNA polymerase controlled through the *lacUV5* promoter on the *E. coli* genome (**Appendix 4A**), which allows the induction of a target gene by IPTG; and pLysS, which contains the T7 lysozyme sequence, suppressing basal expression of the target gene caused by small amounts of T7 RNA polymerase present prior to induction by IPTG (**Appendix 4C**).

Rosetta 2 host strains also contain an additional pRARE2 plasmid, which supplies tRNAs for 6 rare codons rarely used in *E. coli*. The availability of these tRNAs, provided by pRARE2, is thought to enhance the expression of eukaryotic proteins which often require codons seldom used by *E. coli*. **Figure 3.7** indicates the position and number of rare codons in the DNA sequences. The wMT-1 sequence (**Figure 3.7A**) contains 3 rare arginine codons (2 AGA, 1 AGG); the wMT-2 sequence (**Figure 3.7B**) contains 2 rare arginine codons (1 AGA, 1 AGG). Due to the ability of Rosetta 2 cells to provide these rare codons, they were used to express all of the wMTs.

A wMT-1 DNA Sequence (**M** - START, * - FINISH)

```
  1 atg aaa gaa acc gct gct gct aaa ttc gaa cgc cag cac atg gac agc cca gat ctg ggt
 61 acc ctg gtg cca cgc ggt tcc atg gct gat gca tcg aat act cag tgc tgt gga ttt gat
121 gca tgc cca aga cgg gga gct gcg tgt gct tgc acc aac tgc aga tgc ctg aaa agt gaa
181 tgt tcg cca aac tgc agg aag ctt tgc tgt gct gat tcc caa gga aaa tgt gga aat gca
241 ggc tgc aag tgc ggt gct gcc tgc aaa tgt gcc gct ggt gca tgt gcc tct gga tgc aag
301 aag gga tgc tgt ggt gac taa 321
```

B wMT-2 DNA Sequence (**M** - START, * - FINISH)

```
  1 atg aaa gaa acc gct gct gct aaa ttc gaa cgc cag cac atg gac agc cca gat ctg ggt
 61 acc ctg gtg cca cgc ggt tcc atg gct gat gca ttc aac act cag tgc tgt gga aat aag
121 aca tgc cca aga gag gga tca aca tgt gcc tgc tcc aaa tgc agg tgt ccg aaa gat gac
181 tgc gcg cca aac tgc aag aag ctt tgc tgt gct gat gca caa tgt gga aat gca agc tgc
241 agt tgc ggg gct gca tgc aag tgc gca gct ggt tca tgc gcc tca gga tgc aag aaa gga
301 tgc tgc ggt gac taa 315
```

Figure 3.7. DNA sequences of wMT-1 (**A**) and wMT-2 (**B**) obtained from sequenced plasmids (using T7 promoter primer). Frame 1 contains entire expressed protein sequence (including S•Tag). Rare codons highlighted in **RED**. First amino acid encoded is N-terminal methionine (atg), C-terminal residue proceeded by stop codon (taa). Rare codons identified using the Rare Codon Calculator (RaCC) http://nihserver.mbi.ucla.edu/RACC/ [NIH MBI Laboratory for Structural Genomics and Proteomics].

3.5 Subcloned wMT-1, expressed with zinc, purified with cadmium

The subcloned wMT-1 construct was expressed in the presence of zinc, but purified with the addition of cadmium on recommendation from M. Stillman. Previous work expressing wMTs had indicated that expression in the presence of zinc gave higher yields of protein than expression in the presence of cadmium [113]. However, cadmium is a useful indicator in ICP-OES, as it is not present in medium or buffer (unlike zinc).

After expression, cells were lysed in the presence of 1 mM $CdCl_2$, before being separated by size-exclusion chromatography (SEC). The results of the SEC showed a peak in the A_{220} chromatogram around Fraction 10

(**Figure 3.8A**). There is no corresponding peak in the A_{280} chromatogram as there is only one contributing amino acid (phenylalanine). As this column was calibrated with known protein standards (BSA 66 kDa; myoglobin, 17 kDa; cytochrome c 12 kDa), a retention volume of approximately 72 mL (Fraction 9, **Figure 3.8A**) corresponds to a protein of approximately 8-15 kDa. This was supported by the SDS-PAGE gel, with a protein of approximately 10 kDa being present in Fractions 7-9 (**Figure 3.8B**, **GREEN**). Another set of bands presenting at slightly higher mass (approximately 20 kDa) may have also indicated the presence of wMT-1, as it is not uncommon for MTs to run more slowly through gels and present at a higher molecular mass [203].

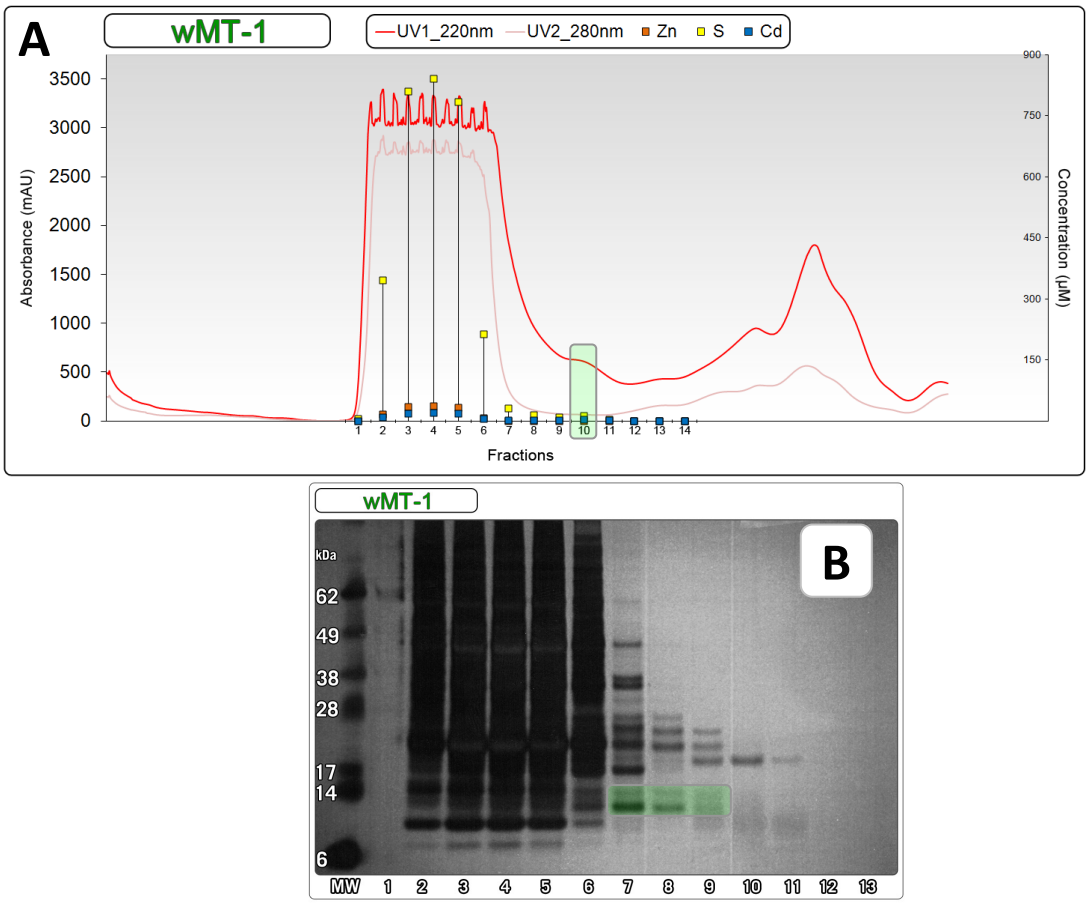


Figure 3.8A. FPLC purification by SEC of subcloned wMT-1. UV absorbances measured at **220 nm** and **280 nm**. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fraction 10 (**GREEN**) showed a peak which could contain wMT-1. **B.** shows an SDS-PAGE gel developed using silver stain enhancement (Invitrogen SeeBlue Plus2 Prestained Standard, designated **MW**), of the Fractions in (**A**). The bands corresponding to Fractions 7-9 (**GREEN**) indicate a protein of a similar molecular mass to wMT-1, with Fractions 9-11 showing another potential protein of slightly higher molecular mass. Fraction volumes are 4mL, with Fraction 1(**A**) beginning at a retention volume of 40 mL.

All fractions after the initial void volume ($V_o \approx 40$ mL for HiLoad 1660 Superdex 75pg 120 mL) were collected and their Zn, S, Cd and Cu content measured by ICP-OES (**Figure 3.8A**, Cu readings below detection limit). Of note are the small quantities of sulphur, cadmium and zinc in Fractions 7-10 (**Figure 3.8A**). Fractions 7-8 and 9-11 were pooled separately and concentrated, the limit of concentration being the minimum volume required for identification by MS (≈ 150 μL).

Neither sample generated a mass spectrum of a protein that was discernible from background noise. The fact that intact wMT-1 was not detectable by ESI-MS, from a 400 mL culture, suggested that either protein expression levels were very low or that the protein had degraded during purification. Subsequent expressions and purifications with approximately three times the starting culture did not generate a clear mass spectrum of wMT-1. To rule out problems caused by expression or the new construct, it was decided to compare the expression profiles of S•Tag-wMT-1 with that of subcloned wMT-1.

3.6 Comparing subcloned wMT-1 with S•Tag-wMT-1

To compare the expression efficiency of the two plasmids (with / without S•Tag), cultures of transformed *E.* coli Rosetta 2 cells were grown in the presence of zinc, and induced after 3 hrs (OD_{600} 0.5-0.7). Aliquots of each culture were removed every hour for the first 6 hrs, and again after 16 hrs. Analysis of the gels (**Figure 3.9**), showed significant differences in the expression profiles of the two constructs. Although the samples were normalised to slightly different optical densities, there was a significant increase in the strength of a band at around 12 kDa in the samples for S•Tag-wMT-1. This is in the correct region for wMT-1, as S•Tag-wMT-1 should present at 10,762 Da (subcloned wMT-1 should present at 7,807 Da).

Although this band (**Figure 3.9, GREEN**) coincides with that of another protein expressed in *E. coli* (hence a 12 kDa band present in subcloned wMT-1, and before induction in S•Tag-wMT-1), there was clear overexpression of protein in the S•Tag-wMT-1 culture. This was not the case for subcloned wMT-1, suggesting that although the construct was correct as indicated by sequencing, the lack of observable protein by MS is likely due to expression problems.

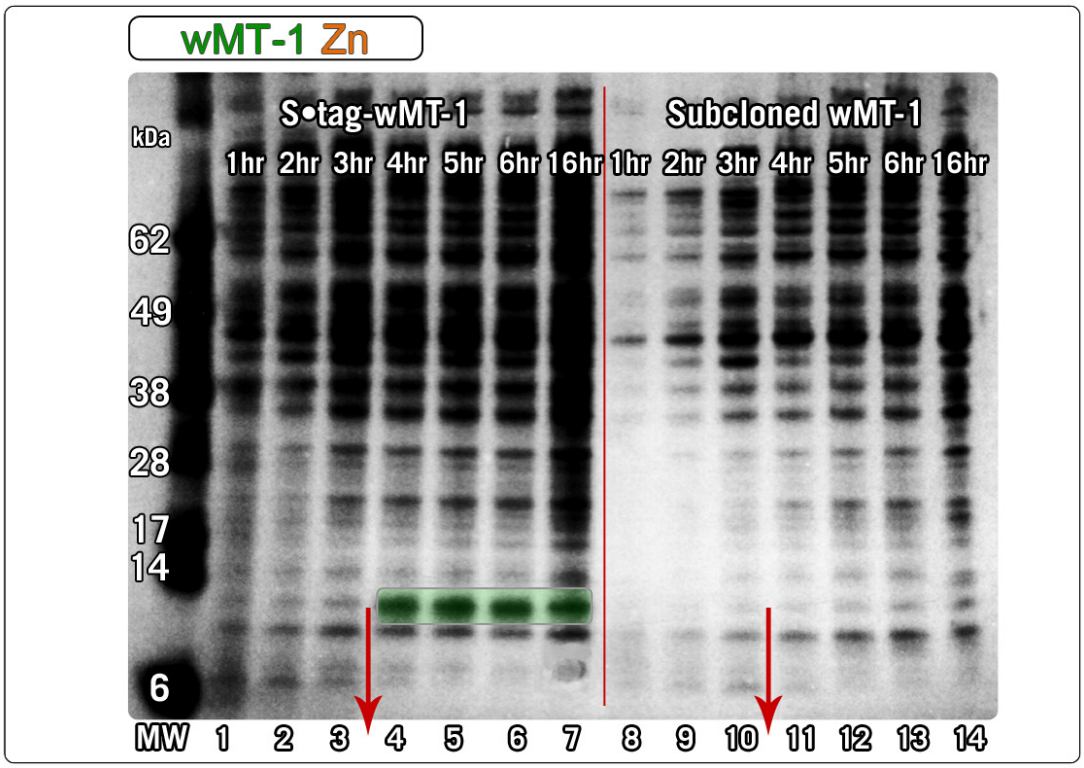


Figure 3.9. SDS-PAGE gel developed using silver stain enhancement (Invitrogen SeeBlue Plus2 Prestained Standard, designated **MW**). Fractions 1-7 are of the S•Tag-wMT-1 plasmid. Fractions 8-14 are of the subcloned wMT-1. The **ARROW** indicates the point at which the cultures were induced. The bands in Fractions 4-7 (**GREEN**), highlight the difference in expression profile between the two constructs. Samples 2-6 were normalised to A_{600} 0.550, samples 9-13 were normalised to A_{600} 0.425. samples 1,8 A_{600} 0.175; samples 7,14 A_{600} 0.950.

This prompted work to cease with the subcloned wMT-1, and work to begin using the original S•Tag-wMT-1 construct (with corrected mutation), and removal of the S•Tag during purification.

3.7 S•Tag-wMT-1, expressed with zinc, purified with zinc

The construct was expressed and purified using the same protocol as for subcloned wMT-1. However, to generate homo-metallated species, 1 mM $ZnSO_4$ was added during sonication. There were three main reasons for not introducing cadmium: it did not prove fruitful in previous work with subcloned wMT-1; wMT-1 is postulated to be a zinc/copper responsive MT, so it may be more stable in the presence of these metals; mass spectra of mixed-metal wMT-1 would be harder to analyse, with cadmium being close to double the molecular mass of zinc. The purification progressed, as before, with sonication and filtration of the crude lysate. The lysate was then separated by SEC which generated the chromatogram in **Figure 3.10**.

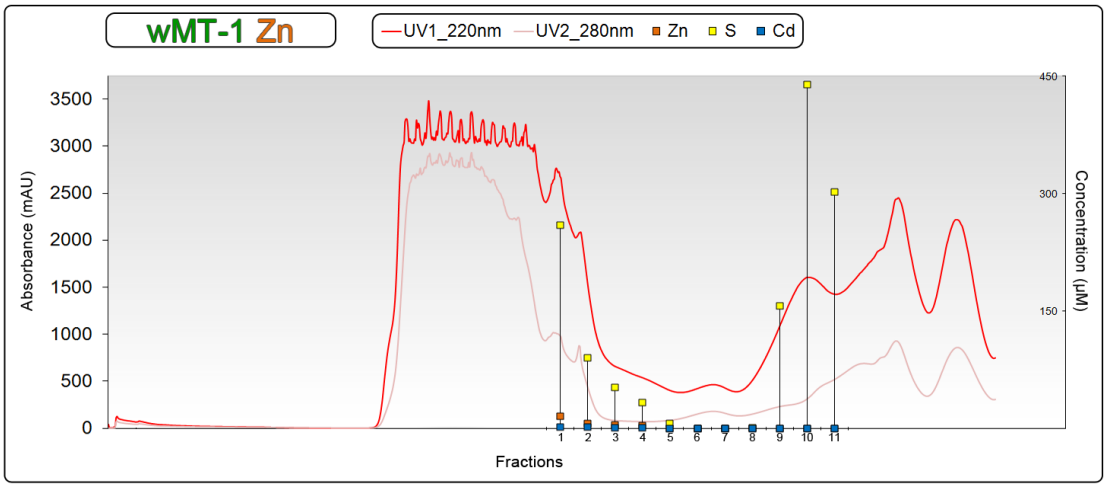


Figure 3.10. FPLC purification by SEC of S•Tag-wMT-1. UV absorbances measured at **220 nm** and 280 nm. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fraction volumes are 4 mL, with Fraction 1 beginning at a retention volume of 64 mL.

This chromatogram was significantly different to that of the subcloned wMT-1 (**Figure 3.8A**). Of note are the fractions at retention volumes between 64-72 mL (**Figure 3.8A**, Fractions 7-9 and **Figure 3.10**, Fractions 1-3). In these fractions, there is a significant amount of sulphur co-

eluting with zinc, as expected for fractions containing MT. However, from the appearance of the A_{220} chromatogram, we can assume that there are a number of proteins eluting within those fractions. Although there is little metal present, the small molecule fractions in Fractions 9-11 (**Figure 3.10**) are particularly high in sulphur, therefore the presence of degraded MT in these fractions cannot be ruled out. It would appear that the quantity of material loaded onto the column caused the UV detector to become overloaded, decreasing the overall resolution of the chromatogram. In combating the problem of detector overload, smaller injections, or more dilute injections were not thought to be viable due to significant increases in purification time. Therefore it was decided to attempt other chromatographic methods such as anion exchange, or cation exchange chromatography, which had been successful in past purifications of wMTs. The following work with both Cd- and Zn- S•Tag-wMT-1 was performed in collaboration with Ms Maike Hansen.

Ion exchange chromatography

One of the difficulties in purifying metal-binding proteins by ion exchange chromatography is ascertaining whether their actual isoelectric point (*pI*) is close to their theoretical isoelectric point. The theoretical p*I* for S•Tag-wMT-1 is 8.49[*], with previous separations being performed at pH 8.1 with anion exchange [99], and pH 7.4 with cation exchange [113] chromatography. The fact that both of these values are within 1 pH unit, but are for techniques which require the protein to carry opposite

[*] http://web.expasy.org/cgi-bin/protparam/protparam

charges, highlights the difficulty of working with protein theoretical p*I*. The ambiguity could be decreased by experimentally determining p*I* through isoelectric focussing experiments, however these often involve the use of specialised equipment. The theoretical method of p*I* determination involves averaging the p*I* of the constituent amino acids of a protein based on work performed by Bjellqvist *et al*. [204]. Making the theoretical p*I* approximation more relevant for MTs, we may consider that divalent metals carry a +2 charge, and thiols are thought to be deprotonated (contributing a -1 charge).

The S•Tag-wMT-1 sequence contains 9 residues which are expected to be negatively charged at neutral pH (Aspartic acid, Glutamic acid); and 15 residues which should be positively charged (Arginine, Lysine), leaving 6 positive charges at neutral pH. If we consider that the cysteine residues are negatively charged, and that the divalent metal ions will contribute a +2 charge, there are 6 additional negative charges contributing to the overall charge on the protein at a given pH, which may decrease the p*I*. There appears, however, to be no information available as to how the charges on the buried metal-thiolate cluster affect affinity to ion exchange columns. Manufacturer's guidelines [GE Healthcare, UK] suggest that optimum conditions for ion exchange chromatography consist of mobile phase buffers at least 1 pH unit above the p*I* of a protein for anion exchange, and 1 pH unit below the p*I* of a protein for cation exchange. Therefore it was decided to attempt both cation exchange and anion exchange, with modification to the previous

protocols. For anion exchange chromatography the pH was increased slightly; for cation exchange the pH was decreased slightly. Another consideration was to keep the pH as high as possible, as low pH has the tendency to strip metal ions from metalloproteins.

Anion exchange chromatography

The cell pellet for anion exchange chromatography was sonicated, before being separated at pH 8.94 on a HiTrap Q XL 5mL column [GE Healthcare, UK]. All fractions were collected and analysed for their S, Zn and Cd content (**Figure 3.11**). It should be noted that the flow-through fraction also appears to contain these elements, however due to just the initial portion being collected, the determined concentration of these elements is likely to be an underestimate. A proportion of S•Tag-wMT-1 remaining in the flow-through therefore, cannot be ruled out. Even so, the raw sulphur concentrations obtained by ICP-OES of the samples collected during the elution were very high (> 1000 μM) when compared to previous experiments (≈ 600 μM).

The separation was ultimately deemed inadequate for subsequent experiments, as significant amounts of contaminant proteins also bound to the anion exchange column. This was a consideration during preparation, as high pH could generate a significant amount of proteins carrying a negative charge. It appears that anion exchange is a valid

technique for capturing protein, especially those with large quantities of sulphur and zinc, which may indicate the presence of MT.

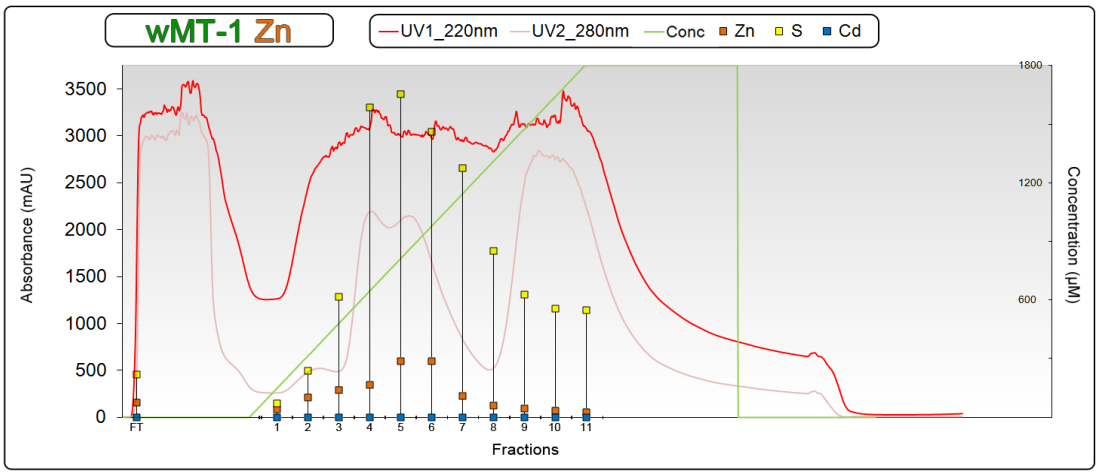


Figure 3.11. FPLC purification by anion exchange chromatography. UV absorbances measured at **220 nm** and 280 nm, salt gradient indicated as a (GREEN) line. ICP-OES concentrations are overlaid (ZINC, SULPHUR, CADMIUM). Elution was over 10 cv (≈ 50 mL), salt gradient of 1 M NaCl from 0-100 %. Mobile phase at pH 9.0. Fraction FT is the 'flow-through' from the column. Experiment performed by Ms Hansen.

However, as the aim of this method was to increase the yield of pure S•Tag-wMT-1, it was decided not to optimise this technique further. However, as ion exchange chromatography could still prove useful for purification, the next stage was to compare this separation to that obtained by cation exchange chromatography.

Cation exchange chromatography

The cell pellet for cation exchange chromatography was prepared in an identical manner to that used for anion exchange. Separation was performed at pH 6.53 on a HiTrap SP XL 5mL column [GE Healthcare, UK]. Fractions were collected through the entire experiment (including flow-through), however only flow-through and fractions with A_{220} > 100 mAU were analysed by ICP-OES (**Figure 3.12**).

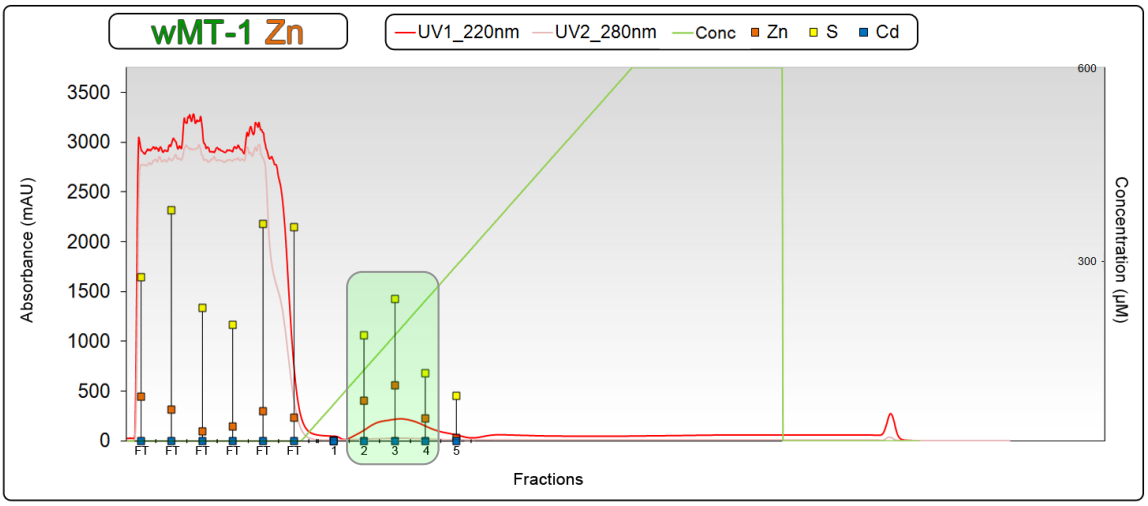


Figure 3.12. FPLC purification by cation exchange chromatography. UV absorbances measured at **220 nm** and 280 nm, salt gradient indicated as a (GREEN) line. ICP-OES concentrations are overlaid (ZINC, SULPHUR, CADMIUM). The peak in Fractions 2-4 highlighted in GREEN was thought likely to contain wMT-1, Elution was over 10 cv (≈ 50 mL), salt gradient of 1 M NaCl from 0-100 %. Mobile phase at pH 6.5. Fractions FT are the 'flow-through' from the column. Experiment performed by Ms Hansen.

The separation by cation exchange chromatography showed one small peak centred at ≈ 18 % Buffer B (**Figure 3.12**, **GREEN**). Although there was a significant amount of sulphur and zinc present in the flow-through fractions, a small amount was retained by proteins which bound to the column. Both the flow-through fractions and Fractions 2-4 were pooled and their volumes reduced for SEC. Due to the large volume of flow-through, to obtain a 4 mL sample for SEC a significant amount of concentration was required. This caused visible protein aggregation within the centrifugal filter, decreasing the final yield of protein. There was no difficulty with concentrating pooled Fractions 2-4 however, so minimal losses in yield were expected. The two samples were filtered and loaded onto an SEC column for separation (**Figure 3.13**).

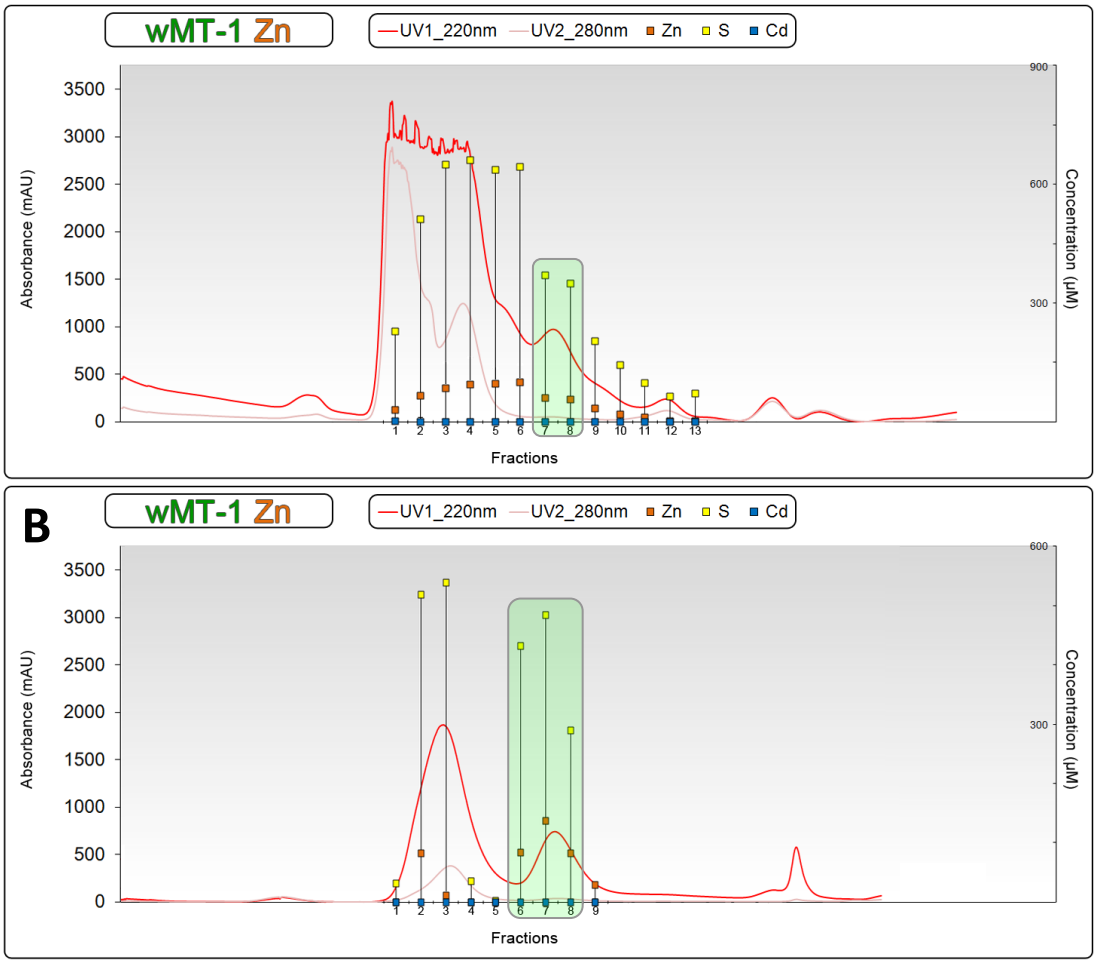


Figure 3.13A. FPLC purification by SEC of pooled and concentrated FT Fractions from cation exchange chromatography. **B.** SEC of pooled and concentrated Fractions 2-4 from cation exchange chromatography. UV absorbances measured at **220 nm** and 280 nm. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). **GREEN** fractions are those proposed to contain S•Tag-wMT-1. Fractions start at a retention volume of 44 mL in both A and B. Both experiments performed by Ms Hansen.

The two chromatograms above indicate that although there appeared to be a significant amount of protein present (indicated by high A_{220} and high sulphur concentration), zinc was present in highest concentration in Fractions 6-8 in the SEC from pooled Fractions 2-4 (**Figure 3.13B**). Although there is more than one protein in the fractions obtained from elution from cation exchange chromatography, it would appear that using this protocol as a first step yields a protein with MT-like characteristics (low A_{280}, high Zn and S content, elution at 64 mL). Fractions 7-8 from **Figure 3.13A** (from cation exchange flow-through), and Fractions 6-9

from **Figure 3.13B** (broad peak in cation exchange, Fractions 2-4) were pooled and concentrated for thrombin cleavage. It should be noted that at this point, a mass spectrum of the putative Zn-S•Tag-wMT-1 was not recorded.

Previously Ms Thiel had some success in the cleavage of the S•Tag from wMT-2, using the Thrombin CleanCleave™ Kit [Sigma Aldrich, UK], the main component being immobilised bovine thrombin on agarose beads. The benefits of using the kit were the ease at which the thrombin could be removed, the speed of the protocol, and the reusability of the enzyme. However, the kit also provided some disadvantages: one being the presence of salt in the sample buffer; the other was the very low protein recovery. The kit also came with a significant cost disadvantage.

After cleavage, thrombin must be removed by some means. The benefit of using SEC for this purpose was in the removal of contaminants eluting at similar retention volumes to S•Tag-wMT-1. When the S•Tag is cleaved, the size of the resulting protein is reduced, potentially solving the problem of low sample purity. Thrombin is priced much more competitively, and for a similar cost to the kit, approximately four times as much protein could be cleaved. Considering the cost and potential purification benefits, and considering the difficulties experienced by Ms Thiel during her work with the CleanCleave™ Kit, it was decided to use a two-step thrombin method, rather than using the premade kit.

The optimum protein concentration for thrombin cleavage is approximately 1 mg/mL, achieved by centrifugal filtration. Approximately 110 % of the required NIH units (used to measure thrombin activity) were added to the concentrated S•Tag-wMT-1 solution. The mixture was incubated with gentle agitation at room temperature for 3 hours. The crude cleavage reaction was then filtered and loaded onto an SEC column (**Figure 3.14**).

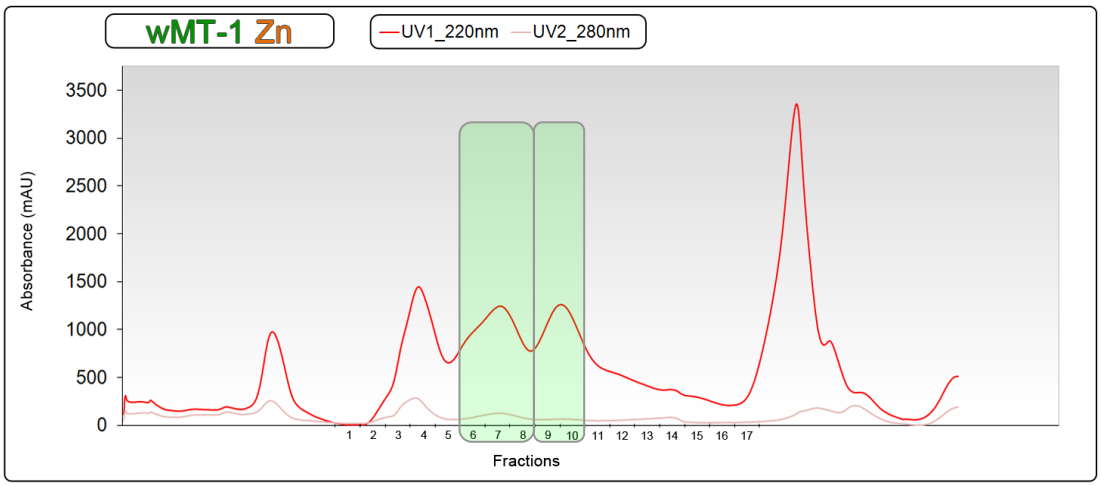


Figure 3.14. FPLC purification by SEC of the pooled and concentrated Fractions 6-9 from cation exchange chromatography. UV absorbances measured at **220 nm** and 280 nm. Pooled fractions for further analysis highlighted in **GREEN**. Fractions start at a retention volume of 40mL. Experiment performed by Ms Hansen.

Fractions 9 and 10 were analysed by ICP-OES, and were found to contain significant amounts of zinc and sulphur, but no cadmium. An SDS-PAGE gel was used to try and discriminate between pooled fractions that may or may not contain cleaved wMT-1 (**Figure 3.15**, below). Although the pool of Fractions 6-8 did contain a possible band for cleaved wMT-1, there was a significant amount of contamination. However in pool Fractions 9 and 10, there appeared to be very few contaminants (although some higher molecular mass proteins at 37 kDa and 25 kDa

are present), and a strong band, corresponding to a molecular mass of approximately 10-11 kDa (cleaved wMT-1 apo- mass is ≈ 8 kDa).

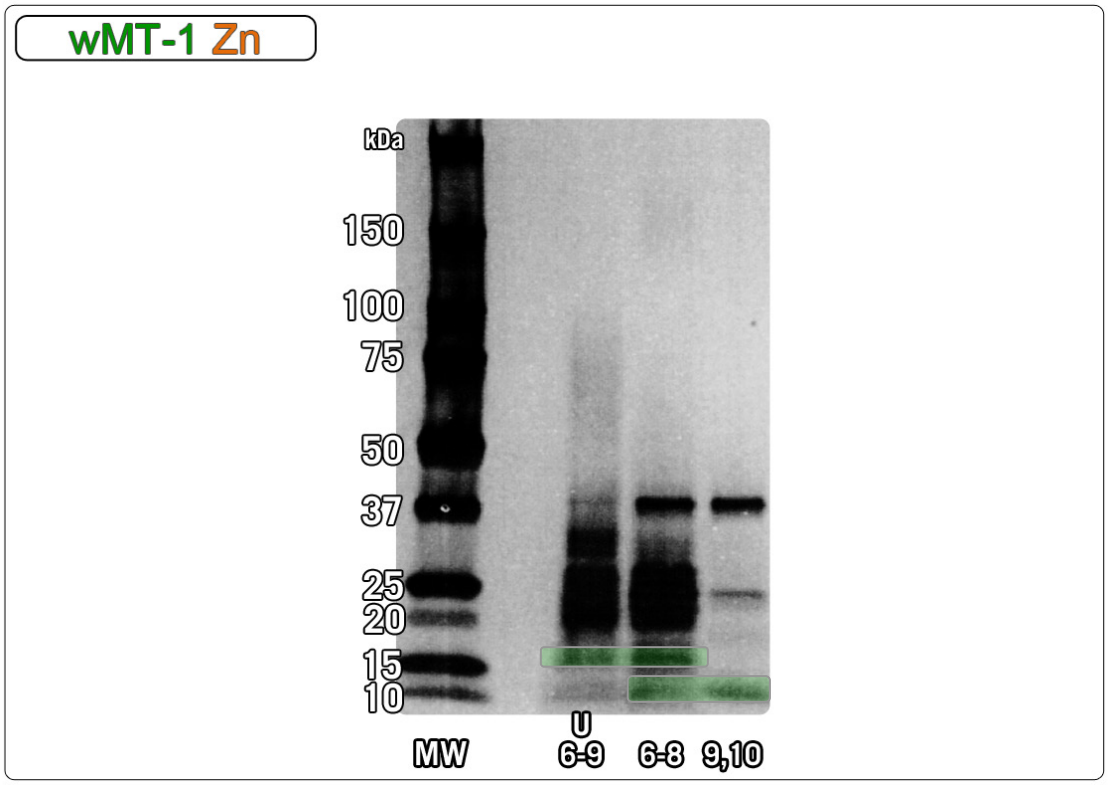


Figure 3.15. SDS-PAGE gel developed using silver stain enhancement comparing uncleaved pooled Fractions 6-9 from SEC of cation exchange flow-through (**Figure 3.13A**), with cleaved pooled Fractions 6-8 and 9,10 from SEC of cation exchange peak (**Figure 3.14**). Biorad Dual Colour Prestained Standard, designated **MW**. Notice the **GREEN** bands highlighting the approximate positions of S•Tag-wMT-1 (U6-9; 6-8) and cleaved wMT-1 (6-8; 9,10).

There is the possibility that with < 100 % cleavage efficiency, some S•Tag-wMT-1 remained in Fractions 6-8 from the separation, so at this point the identity of the band remained ambiguous. The concentrations of the pooled Fractions 9 & 10 were determined to be 486.9 µM S and 121.7 µM Zn, giving a S:Zn ratio of approximately 4.0:1. If we consider that cleaved wMT-1 contains 21 sulphur atoms, and can bind seven metal ions, we would expect a S:Zn ratio of 3:1 in a pure sample of cleaved Zn_7-wMT-1.

The mass spectrum is shown in **Figure 3.16**, with the raw spectrum indicating the peak used for the deconvoluted spectra. The overall quality and the signal/noise ratio (< 10) of the mass spectrum was poor, indicating that the actual protein concentration was significantly less than expected. These losses could have occurred during concentration either through protein permeating the filter membrane or by aggregation of protein. Another explanation could be that cleaved wMT-1 has low ionising efficiency, and therefore a higher concentration of protein is required to achieve an acceptable signal/noise ratio.

Deconvolution of the peaks in **Figure 3.16C** gave the cleaved Zn-wMT-1 species: Zn_7-wMT-1 8,394.3 Da (8,394.9 Da theoretical); Zn_6-wMT-1 8,328.5 Da (8,331.5 Da theoretical); Zn_5-wMT-1 8,260.1 Da (8,268.1 Da theoretical); Zn_1-wMT-1 8,031.3 Da (8,077.9 Da theoretical); apo-wMT-1 7,942.9 Da (7,951.2 Da theoretical). This spectrum shows recombinant cleaved Zn-wMT-1 for the first time. Following the methodology of Capdevila *et al*. [205], we can assume that the intensities of protein species in a mass spectrum correspond to their actual abundance, therefore the major species at pH 8.50 is Zn_7-wMT1. Decreasing the pH to 2.6 allowed observation of the apo-mass. The raw mass spectrum (**Figure 3.16B**), indicates that although the 6+ and 5+ charge states were used for deconvolution, there are a number of contaminant peaks present which do not correspond to known Zn-wMT-1 species (**Figure 3.16**, '*'). The major raw contaminant peaks (**Figure 3.16B** at 1401.7, 1456.3, 1546.2 m/z), when manually deconvoluted assuming a

5+ charge state gave masses which are below the mass of apo-wMT-1. When manually deconvoluted assuming a 6+ charge state, the contaminant at 1401.7 m/z gave a mass of 8,404.2 Da, which is similar to the Zn_7-wMT-1 theoretical mass of 8,394.9 Da. However, it is unlikely that wMT-1 would contain seven zinc ions below pH 3. Knowledge gained from this successful purification of S•Tag-Zn-wMT-1 was then used in the purification of S•Tag-Cd-wMT-1.

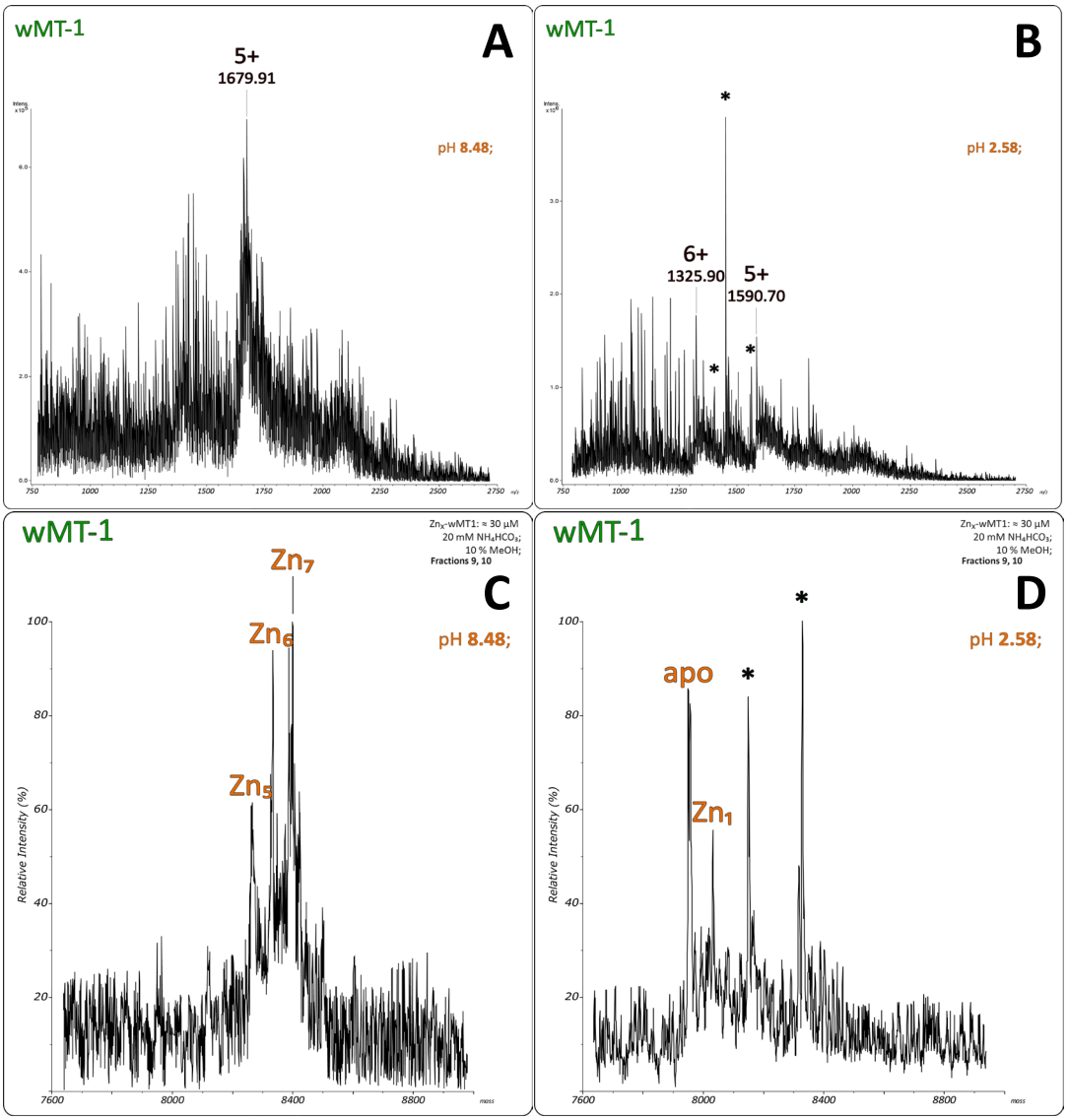


Figure 3.16. ESI-MS of cleaved Zn-wMT-1 of Fractions 9,10 from **Figure 3.14**. Cleaved wMT-1 concentration ≈ 30 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is the raw spectrum of wMT-1 at pH 8.48, (**B**) is the raw ESI-MS after addition of 2 % acetic acid, pH 2.58.'*' peaks indicate possible contaminants, some contributing to the spurious data when deconvoluted. (**C**) is the deconvoluted spectrum of (**A**), (**D**) is the deconvoluted spectrum of (**B**) at pH 2.58, which enabled the observation of the apo- mass of the protein. '*' peaks indicate possible contaminants. A mass list is presented in **Appendix 4**.

3.8 S•Tag-wMT-1 expressed with cadmium, purified with cadmium

Cd-S•Tag-wMT-1 was obtained using a similar protocol to that for Zn-S•Tag-wMT-1. However, there exist two published strategies for the use of cadmium in *E. coli* recombinant protein expression: either multiple additions of 50 μM $CdCl_2$ [113]; or addition of 500 μM $CdCl_2$ at induction [151]. As cadmium is not required before induction, and 500 μM $CdCl_2$ could have caused retardation of cell growth, the accepted protocol within the group was to use a lower concentration of 200 μM $CdCl_2$ at induction. In an attempt to stabilise the target protein, a further 1 mM $CdCl_2$ was added before sonication. The crude lysate, after centrifugation and filtration, was purified by cation exchange chromatography at pH 7.34. In an attempt to build on lessons learned from purifying Zn-S•Tag-wMT-1 (**Figure 3.12**), the chromatography proceeded in a step-wise manner (**Figure 3.17**), at a slightly higher pH. This was to attenuate protein binding to the cation exchange column, whilst maintaining full metallation of the bound proteins. The elution proceeded with step-wise gradients with respect to column volume (cv): 4 cv at 12 % NaCl; 4 cv at 24% (expected S•Tag-wMT-1 elution stage); 4 cv at 50% and finally 4 cv at 100%.

Performing cation exchange chromatography with approximately twice the amount of cellular material (compared to the previous purification of Zn-S•Tag-wMT-1) meant that a significant amount of protein bound to the column. The UV detector was also heavily overloaded, hence

crenulations at the top of the chromatograms. Although it cannot be ruled out that cadmium was in the first 10 mL of flow-through, or in the final 100 % gradient fractions (which were not analysed for cadmium), a large quantity of cadmium presented during the 12-24 % gradient step, which was a good indication of Cd-S•Tag-wMT-1 being present.

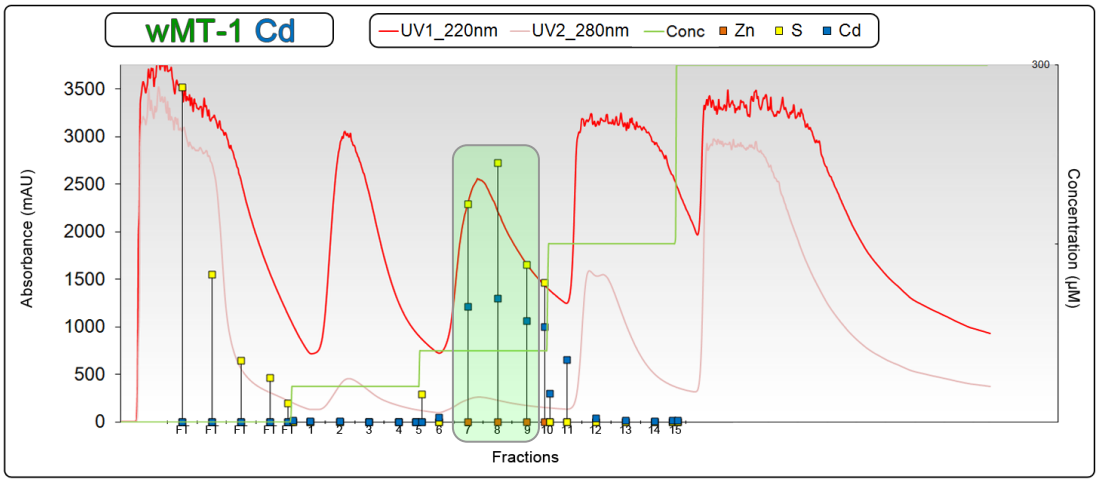


Figure 3.17. FPLC purification by cation exchange chromatography. UV absorbances measured at **220 nm** and **280 nm**, salt gradient indicated as a (**GREEN**) line. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fractions though to contain wMT-1 shown in **GREEN**. Elution was over 20 cv (≈ 100 mL), salt gradient of 1 M NaCl performed stepwise: 0-12 %; 12-24 %; 24-50 %; 50-100 %. Mobile phase at pH 7.3. FT Fractions are the 'flow-through' from the column. Performed by Ms Hansen. For Fractions 1-4 and 12-15 the sulphur concentration was > 400 µM (which put it outside of the ICP-OES calibration curve).

Fractions 7-10 were pooled, concentrated by centrifugation, and applied to an SEC column (**Figure 3.18**). The separation of the fractions from cation exchange chromatography showed a large number of proteins were present in the 24 % NaCl step, indicating a large proportion of other proteins co-eluting with Cd-S•Tag-wMT-1.

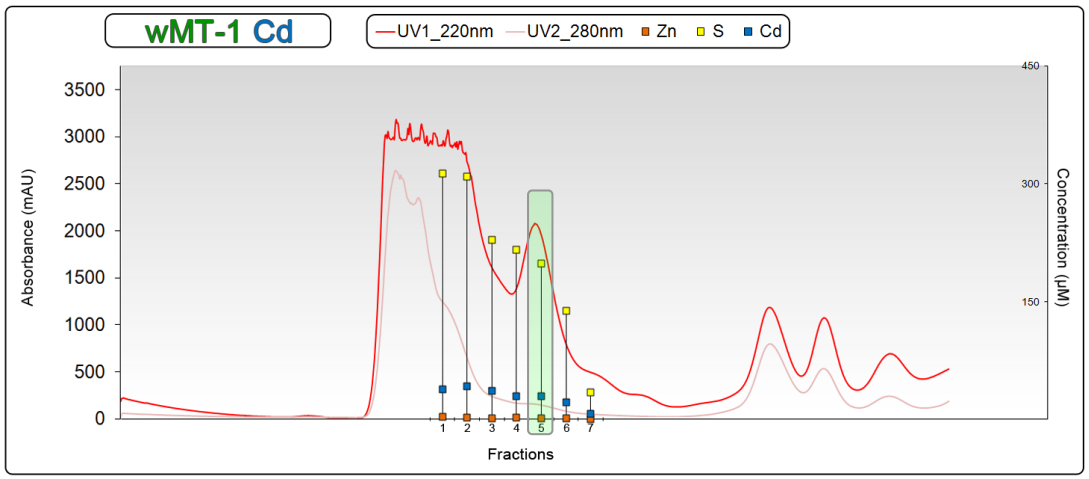


Figure 3.18. FPLC purification by SEC of the pooled and concentrated Fractions 7-10. UV absorbances measured at **220 nm** and 280 nm. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fractions start at a retention volume of 52 mL. **GREEN** Fractions were determined to contain Cd-S•Tag-wMT-1 by MS. Experiment performed by Ms Hansen.

For mass spectrometry, Fraction 5 (**Figure 3.18, GREEN**) provided a compromise between protein yield and purity (estimated from ICP-OES cadmium concentration as ≈ 52 % pure), and was pooled and concentrated. Contaminants in the sample, whilst contributing to the sulphur content, are unlikely to be cadmium-binding. Therefore the cadmium concentration was used as a better indicator of wMT-1 concentration than sulphur alone.

Few species were observed in the mass spectrum of Cd-S•Tag-wMT-1 (**Figure 3.19**), although the analyses were performed at the extremes of the pH range when compared to past experiments with Zn-S•Tag-wMT-1. The major species at pH 8.21 was Cd_7-S•Tag-wMT-1, with a significant peak for a Cd_8-S•Tag-wMT-1 species also present. This was unexpected, although overmetallalation of MTs with cadmium does have some precedent [206, 207]. It would also appear that the Cd_7-S•Tag-wMT-1 species is stable, as it is the major metal species (ie. partially

demetallated species are not observed as with Zn-S•Tag-wMT-1, **Figure 3.16**). Although the raw mass spectrum (**Figure 3.20**) did show some contamination, the peaks of greatest magnitude corresponded to Cd-S•Tag-wMT-1.

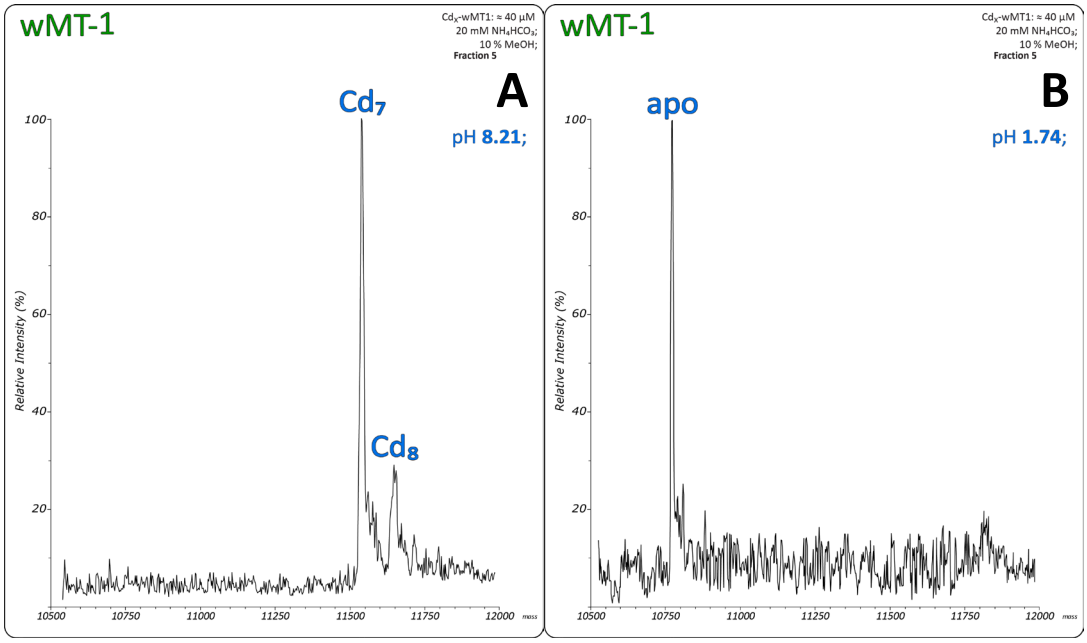


Figure 3.19. ESI-MS of Cd-S•Tag-wMT-1 from **Figure 3.18**, Fraction 5. S•Tag-wMT-1 concentration ≈ 40 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is at pH 8.21, spectrum (**B**) is after addition of 5 % formic acid, pH 1.74, which enabled the observation of the apo- mass of the protein. A mass list is presented in **Appendix 5**.

The contaminant proteins, either unassigned, or indicated '*' in **Figure 3.20** have a calculated molecular mass outside of the ± 1 kDa window used by the deconvolution algorithm of Mass Analysis [Bruker Daltonics, UK], as they did not affect the quality of the deconvoluted mass spectrum in **Figure 3.19**.

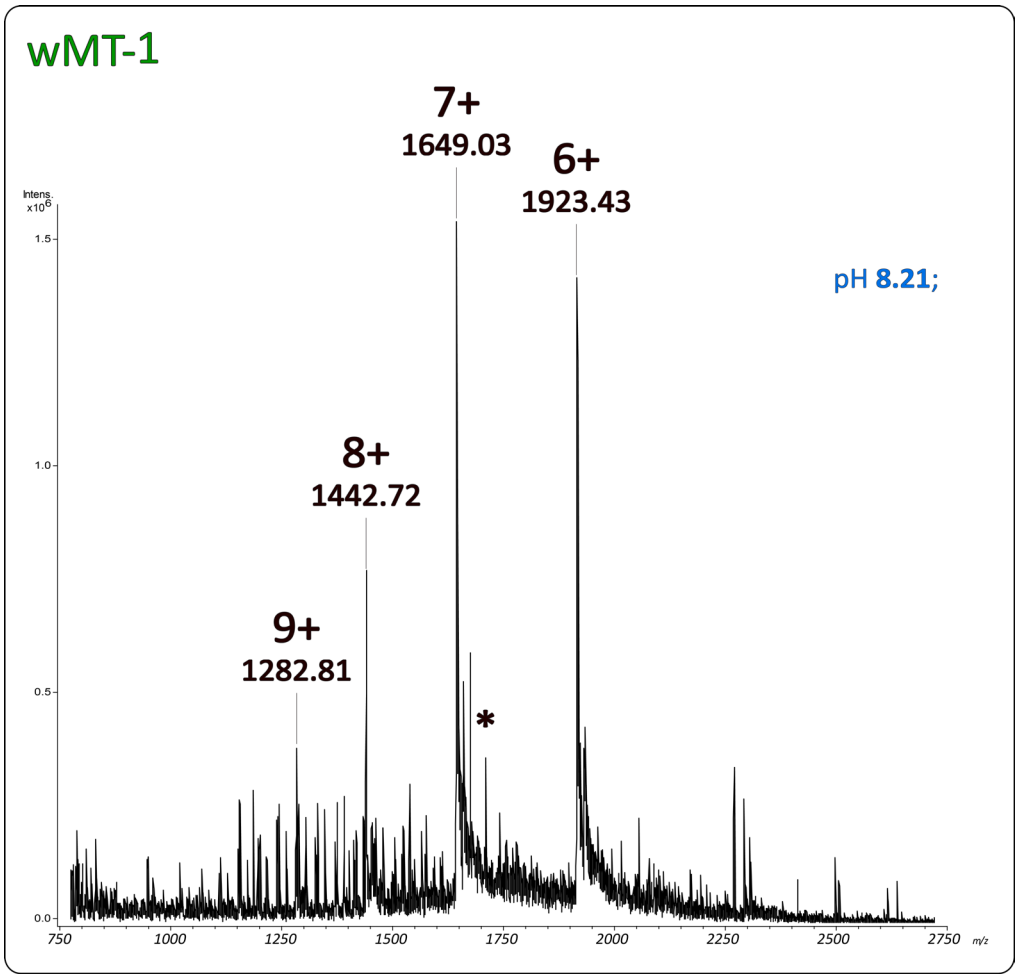


Figure 3.20. Raw ESI-MS of Cd-S•Tag-wMT-1 at pH 8.21.'*' peak (and surrounding unlabelled peaks) indicate possible contaminants, although they did not affect the deconvolution of Cd-wMT-1 peaks.

Realising that contaminant proteins could be removed after thrombin cleavage and SEC, Fractions 3-6 (**Figure 3.18**) were pooled and concentrated to 1 mg/mL protein. The SEC chromatogram of the cleaved fractions showed that the cleavage efficiency was either low, or there was significant contamination from other proteins, as a large peak remained in Fractions 1 and 2 (**Figure 3.21**).

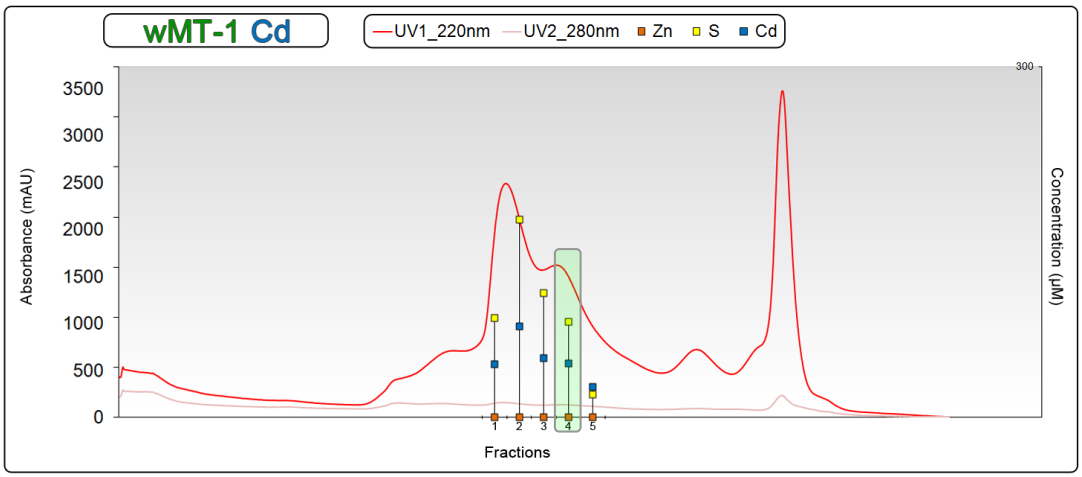


Figure 3.21. FPLC purification by SEC of the pooled and concentrated Fractions 3-6 from **Figure 3.18**. UV absorbances measured at **220 nm** and 280 nm. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fractions start at a retention volume of 60 mL. **GREEN** Fractions were determined to contain cleaved Cd-wMT-1 by MS. Experiment performed by Ms Hansen.

Analysing the metal content of these samples gave low sulphur:cadmium ratios, with approximately half of the expected sulphur present. Fraction 4 (**GREEN**, **Figure 3.21**) was concentrated (as of the five fractions, this was the most likely to contain cleaved Cd-wMT-1 with the least contamination, based on retention volume and elemental concentrations). As sulphur (≈ 1 ppm detection limit) is less sensitive than cadmium (≈ 1 ppb detection limit), it is most likely that the sulphur concentrations were underestimated.

SDS-PAGE (**Figure 3.22**) showed some bands present after cleavage (Fractions 1-3, **Figure 3.21**) not present previously (Fractions 3-6, **Figure 3.18**). The unidentified bands could be either thrombin (one subunit is approximately 36 kDa), or created through specific / non-specific cleavage of proteins other than S•Tag-wMT-1.

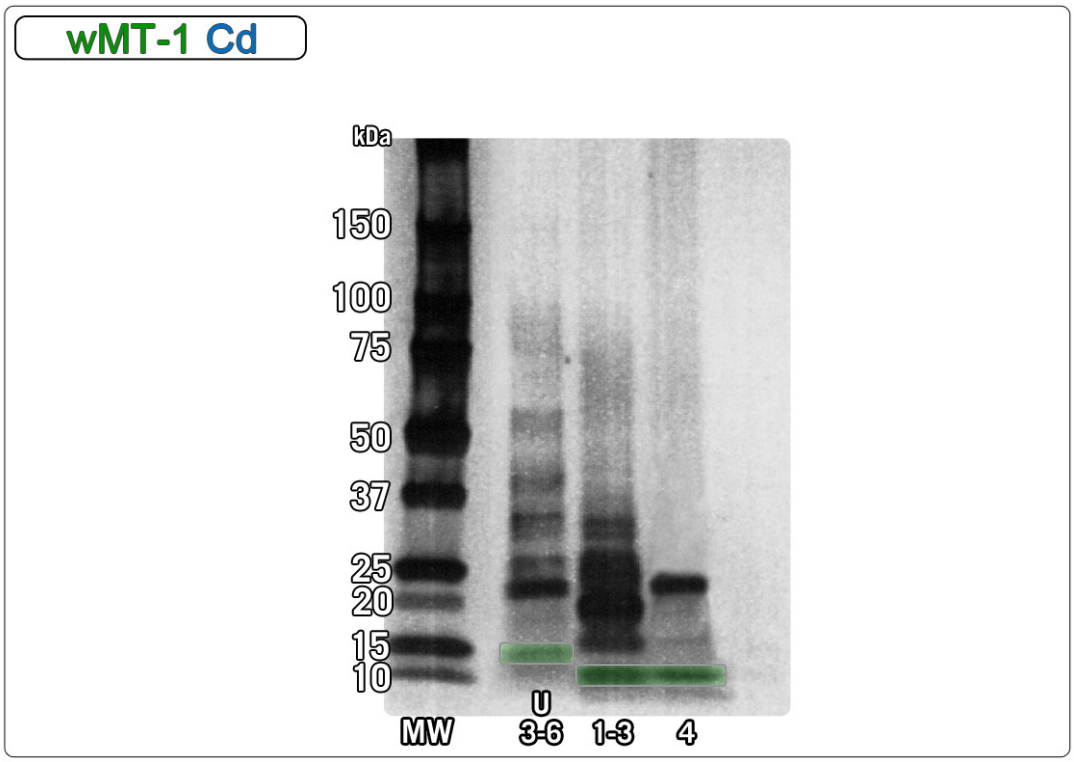


Figure 3.22. SDS-PAGE gel developed using silver stain enhancement comparing uncleaved pooled Fractions 3-6 obtained from cation exchange and SEC (**Figure 18**), with cleaved pooled Fractions 1-3 and 4 (**Figure 21**). Biorad Dual Colour Prestained Standard, designated **MW**. Notice the **GREEN** bands highlighting the approximate positions of S•Tag-wMT-1 (U3-6) and cleaved wMT-1 (1-3; 4).

Mass spectrometry of Fraction 4 (**Figure 3.23**) were recorded with a lower signal/noise ratio than for the Cd-S•Tag-wMT-1 (**Figure 3.19**). This could be due to the sample concentration being lower than was approximated from the cadmium concentration. Although the spectra of cleaved Cd-wMT-1 and Cd-S•Tag-wMT-1 > pH 8 showed the same species, there appear to be more metallated species at lower pH (**Figure 3.23B**) for cleaved Cd-wMT-1. This may be due to the pH being ≈ 0.5 pH units higher for the data shown in **Figure 3.23**.

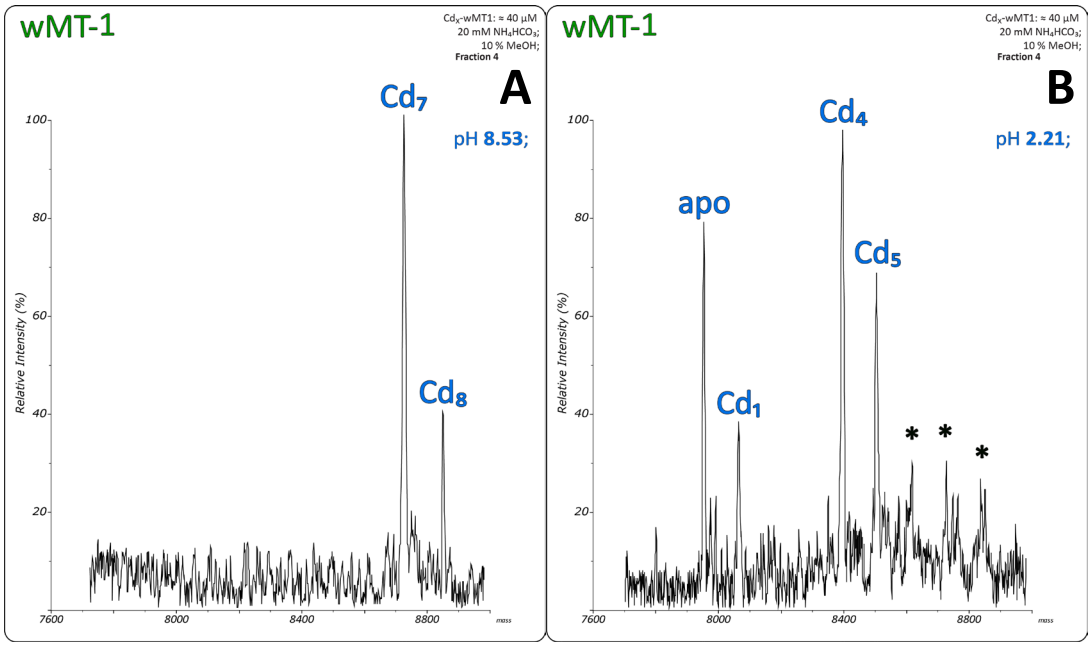


Figure 3.23. ESI-MS of cleaved Cd-wMT-1 from **Figure 3.21**, Fraction 4. Cleaved wMT-1 concentration ≈ 40 µM, Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is at pH 8.53, spectrum (**B**) is after addition of 2 % acetic acid, pH 2.21, which enabled the observation of the apo- mass of the protein. Contaminant peaks indicated with '*'. A mass list is presented in **Appendix 6**.

From **Figure 3.23A**, the major species shows seven cadmium ions bound per cleaved wMT-1 molecule. This is in reasonable agreement with the seven metal ions expected to be bound by wMT-1, based on primary sequence analysis. The SDS-PAGE gel (**Figure 3.22**) indicated that even after the second round of SEC, some contaminant proteins remained, somewhat confirmed by the observation of contaminant peaks in the spectrum at low pH (**Figure 3.23B**, '*'). Further optimisation would have to be performed, with a focus on obtaining pure protein. With the success of the now established procedure, the next objective was to attempt to optimise each step of protein expression and purification.

3.9 Optimising expression

The largest change in expression was to trial auto-induction (AI) medium [184]. One of the biggest advantages of AI medium is that it can support growth to higher cell densities [184], which may enable cells to produce higher quantities of the target protein. Additionally, expression is under metabolic control of *E. coli*, initiating the change in priority from cell replication, to production of target protein at the optimum time.

Metabolic control is through carbon catabolite repression (CCR) allowing rapid adaptation to a preferred energy source in a medium [208]. The AI medium contained a mixture of glucose, glycerol and lactose. Unless the medium becomes acidic (< pH 5), *E. coli* will continue to utilise glucose as its sole carbon source until depletion. Other sources of carbon (ie glycerol and lactose) will be utilised thereafter [184], with glycerol supporting growth at a similar rate to glucose, but suppressing the effect of other carbon sources much less efficiently. This allows lactose to be metabolised inducing expression of target protein (analogous to IPTG). To ensure that induction was occurring during the exponential phase of growth, growth curves were plotted (**Figure 3.24**). The experiment was performed with S•Tag-wMT-1, expressed in the presence of cadmium (200 µM $CdCl_2$ and 1 mM IPTG added when LB culture induced; 200 µM $CdCl_2$ and 0.20 % lactose present from the start in AI medium). To maintain uniformity in results, and direct comparison between each set, 30 °C was chosen. Low temperature expression has been a proven strategy for increasing the solubility, and decreasing the propensity for

misfolding of recombinant proteins [209], a potential cause of the inability to isolate subcloned wMT-1.

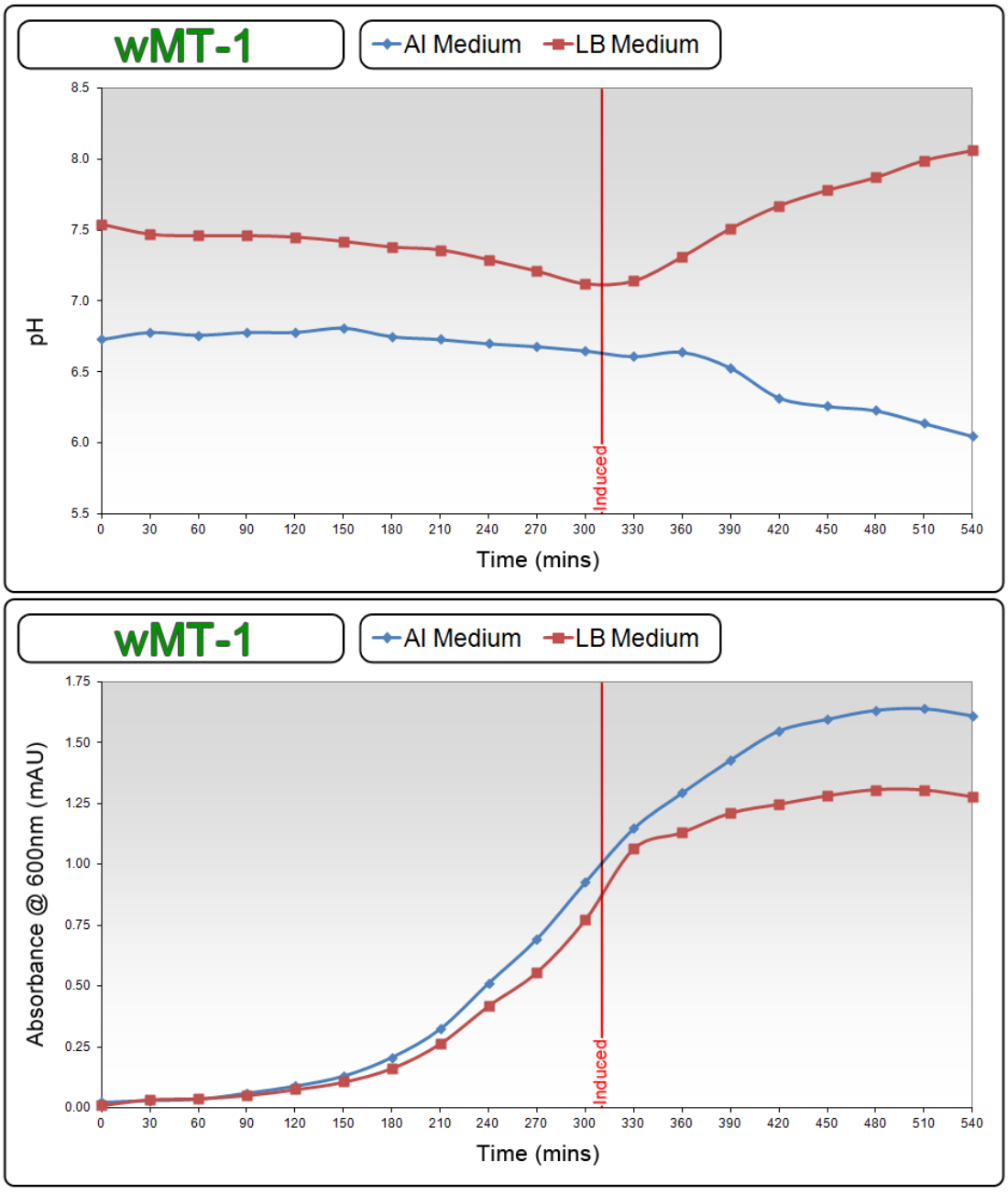


Figure 3.24. Comparison of pH (**TOP**) and OD_{600} (**BOTTOM**) of AI (BLUE) and LB (**RED**) medium in Rosetta 2(DE3) pLysS cells. The LB culture was induced with 1 mM IPTG and supplemented with 0.2 mM $CdCl_2$ at 310 mins. Readings taken every 30 minutes until death phase observed at 540 mins. Experiment performed by Ms Hansen. Cultures shaken at 180 rpm at 30 °C.

Figure 3.24 highlights a difference between the two media, as a significant divergence of pH values between the two experiments is observed after induction. Decreasing pH is often equated with respiration

by-products - increased amounts of aqueous CO_2, generating more hydrogen carbonate in the medium. A steady decrease in pH was observed in the 'unbuffered' LB culture before induction, caused by the build-up of acidic products from respiration and metabolism of sugar. **Figure 3.24** indicates that AI medium can support cultures at higher cell densities, with OD_{600} > 120 % higher when grown in AI medium. The data also confirmed that induction of LB media at OD_{600} ≈ 0.7 is well within the exponential phase of *E. coli* growth, the optimal time for induction with respect to protein yield.

There was a large lag time in both experiments, known to be caused by the addition of chloramphenicol to select for the pRARE and pLysS plasmids [Invitrogen User Manual]. A distinct pH increase after IPTG induction in LB could indicate expression of the target protein, a trend observed by groups working with similar metal-binding proteins [210]. It may be that the removal of metal ions from solution causes the overall pH to increase slightly, although the magnitude of the effect would indicate a contribution from another source.

Another function of the Rosetta 2 cell line is a proportional response to IPTG concentration [Novagen Manual]. Increasing the concentration between 0.1-1 mM IPTG will induce protein expression at different levels. This can be useful if the protein is prone to form inclusion bodies, or is sparingly soluble in the medium. However, as the temperature was already decreased (from 37 °C to 30 °C), the effects of a low

concentration of inducing agent was not investigated at this time; this may have also compromised the aim of generating maximum yield of the target protein. From comparing the expressions between the two media, AI medium did allow greater cell densities to be achieved, but translating this into higher yields of pure protein proved to be rather more difficult. The subsequent experiments were carried out with the S•Tag-wMT-2 construct, since time constraints required the development of an adequate purification procedure for both proteins.

3.10 S•Tag-wMT-2 expressed with cadmium, purified with cadmium

With the perceived benefits of using AI medium over LB medium, a comparison experiment was undertaken using the S•Tag-wMT-2 construct (**Figure 3.25**). With AI medium able to support cells grown to higher cell densities than in LB medium, larger pellets were produced in each culture, with some increases of mass in excess of 400 %. **Figure 3.25** also shows the comparison between cultures being induced and uninduced. It would be expected that if the culture is not induced then the cell priority would never diverge from replication, causing more cells to be present during harvesting. In all cases the pellet size was increased when uninduced, and smallest for induced LB cultures. However, whilst the average pellet masses for AI medium were in the order of four times larger, this did not necessarily correspond to a fourfold increase in yield of wMT.

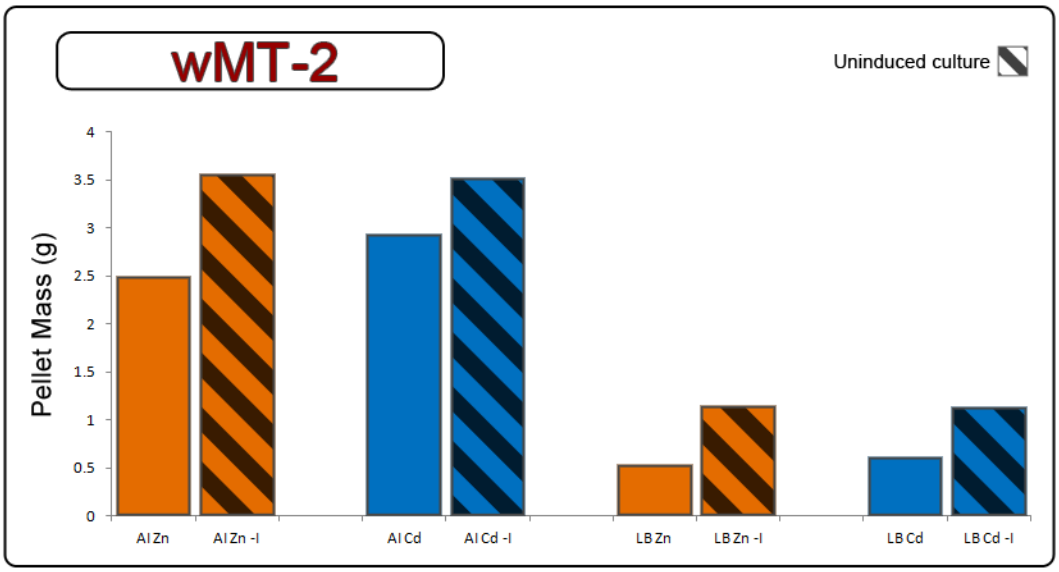


Figure 3.25. Compiled chart of S•Tag-wMT-2 pellet masses (wcw) from both LB and AI media. All conditions were kept constant, other than the addition of IPTG (1 mM) and metal (0.5 mM $ZnSO_4$; 0.2 mM $CdCl_2$) when LB cultures induced. AI medium without inducing agent was created without lactose. '-I' designates cultures without inducing agent (shown as striped bars).

Cation exchange chromatography

After sonication (supplemented with 1 mM $CdCl_2$) and filtration, cation exchange chromatography was performed with a mobile phase of 20 mM ammonium bicarbonate at pH 7.19, and Buffer B containing 1 M NaCl (**Figure 3.26**). Although step-wise elution was successful for Cd-S•Tag-wMT-1, a gradient of 0-100 % over 30 cv was chosen, as Cd-S•Tag-wMT-2 had not been previously separated by ion exchange chromatography. The S•Tag-wMT-2 construct contained a possible 14 positively-charged residues, one less than S•Tag-wMT-1, and 9 negatively-charged residues. This gave an overall 5+ charge at neutral pH. Although sulphur concentrations were very high, there was no appreciable amount of cadmium present in eluate fractions. Unfortunately the likely location of Cd-S•Tag-wMT-2, the flow-through, was not collected during this experiment.

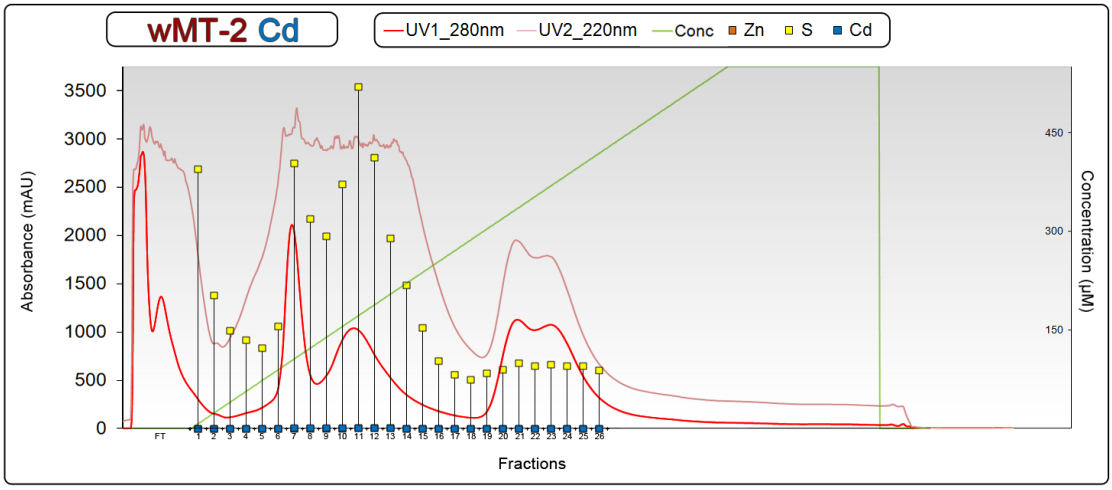


Figure 3.26. FPLC purification of Cd-S•Tag-wMT-2 by cation exchange chromatography; UV absorbances measured and 220 nm at **280 nm** (Note that these colours are reversed from other experiments). ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Note that the colours are reversed from other experiments. Salt gradient (**GREEN**) performed 0-100 % Buffer B over 30 cv. Both buffers at pH ≈ 7.4.

Previous experiments with S•Tag-wMT-1 separated by anion exchange chromatography (**Figure 3.11**) had shown a significant amount of protein binding at pH ≈ 9, which made it unsuitable for wMT isolation. However, considering the high p*I* of S•Tag-wMT-1, and the ability of both cation and anion exchange columns to capture large amounts of contaminant proteins, an alternative technique was tested. This was to perform concurrent anion exchange and cation exchange chromatography, referred to as 'AnCat' exchange chromatography.

AnCat exchange chromatography

For the AnCat exchange column, an anion exchange column, and a cation exchange column were connected in series, forming a hybrid column with a cv of 10 mL. The experiment was performed with Buffer A, 20 mM ammonium bicarbonate, with Buffer B containing 1 M NaCl (**Figure 3.27**).

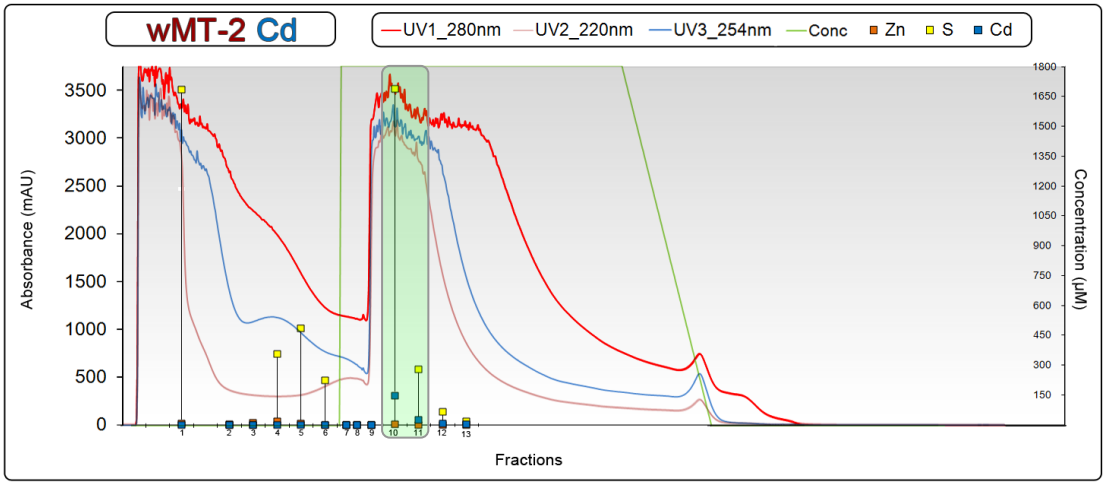


Figure 3.27. FPLC purification of Cd-S•Tag-wMT-2 by AnCat chromatography; UV absorbances measured at **220 nm**, **280 nm** and **254 nm**. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Stepwise elution of 5 cv 0 % Buffer B, 5 cv of 100 % Buffer B; followed by gradient elution to 0 % Buffer B (**GREEN**). Both buffers at pH ≈ 8.4. S•Tag-wMT-2 containing fractions indicated in **GREEN**.

For AnCat exchange chromatography, the mobile phase was at pH 8.29, approximately the theoretical *pI* of S•Tag-wMT-2 (the point at which there is no net charge on the protein). For optimal column retention of the target protein, the mobile phase should have a pH ± 1 unit from the theoretical p*I.* Therefore a pH at the protein p*I* should cause the target protein to pass through the columns, with contaminant proteins being bound to the columns. It was hoped that the low affinity of S•Tag-wMT-2 for the exchange columns would make this a very efficient step for removing contaminants. The FPLC chromatogram (**Figure 3.27**) shows two broad peaks at the start of each step (either 0 % Buffer B, or 100 % Buffer B). A third UV wavelength, 254 nm, was used during this experiment, as it can be somewhat useful in identifying fractions containing S-Cd bonds (although there is some overlap with 220 nm and 280 nm). The fractions with the highest A_{254} were Fractions 1-2 (in the flow-through) and also Fractions 9-11. However, only Fraction 10 contained a significant amount of cadmium as determined by ICP-OES

(147 µM). A significant amount of sulphur was present in Fraction 10 (> 1mM), a good indicator that MT may be present. This was not expected, as at pH 8.4, the protein was assumed not to have affinity for either column, based on theoretical p*I*. However, the presence of S•Tag-wMT-2 eluting when the salt-containing buffer was applied was supported by an SDS-PAGE gel (**Figure 3.28**).

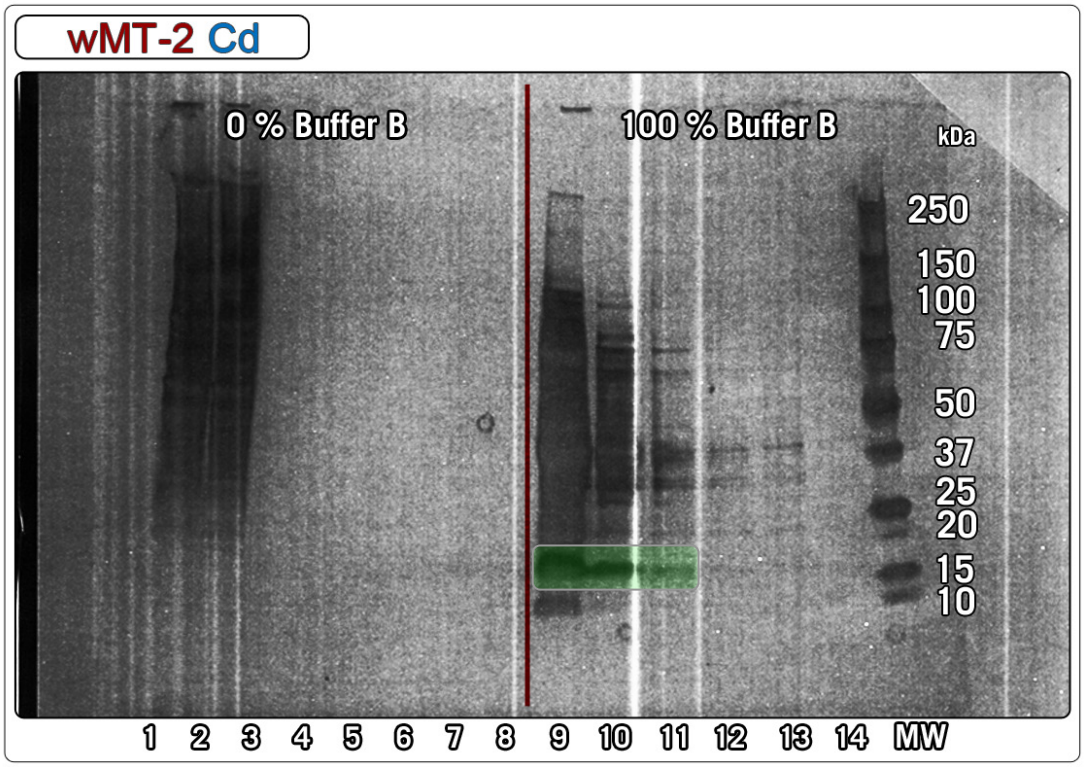


Figure 3.28. SDS-PAGE gel of Fractions 1-14 from AnCat exchange chromatography. Biorad Dual Colour Prestained Standard, designated **MW**. The **RED** line divides the gel into: 0 % Buffer B (**left**) and 100 % Buffer B (**right**). S•Tag-wMT-2 containing fractions are highlighted **GREEN**.

The SDS-PAGE gel was rather ambiguous for Fractions 1-8, with no real identifiable bands presenting in the 10-15 kDa range. However, it was decided to pool Fractions 1-8, and concentrate by centrifugation at slow speed (2,000 x *g*, to avoid protein aggregation). The supernatant was filtered and separated by SEC (**Figure 3.29**). The amount of cadmium present in all fractions was very low, with no discernible peak present. There was also distortion of the peak area around the expected retention

volume for S•Tag-wMT-2 (Fractions 4-6, **Figure 3.29**). All fractions were analysed by SDS-PAGE (**Figure 3.30**).

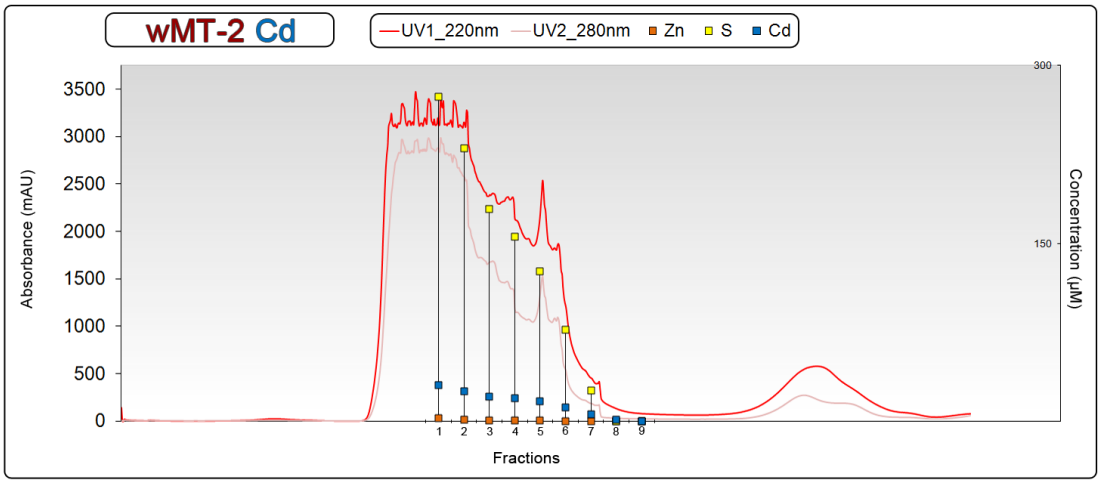


Figure 3.29. SEC of pooled 0 % Buffer B fractions from AnCat chromatography; UV absorbances measured at **220 nm**, **280 nm** and **254 nm**. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). Fractions begin at 52 mL retention volume.

There appeared to be some potential candidates for S•Tag-wMT-2, as there were some bands present in the 10-15 kDa range. However, these bands were present across the separation, and the resolution between proteins in the sample was insufficient.

This evidence would suggest that S•Tag-wMT-2 did bind in some way to the AnCat column at pH 8.4, most likely to the anion exchange column, indicating that using the theoretical p*I* for S•Tag-wMT-2 is indeed not reliable. However the use of anion exchange with S•Tag-wMT-1 proved to be problematic. In addition, using the theoretical p*I* as a starting point, and increasing by 1 pH unit (for optimal separation) gives a very high pH (pH 9.4). Compromising at pH 9.0 meant that too many proteins are bound to the column, therefore anion exchange chromatography is unsuitable as a first fractionation step for S•Tag-wMT-2.

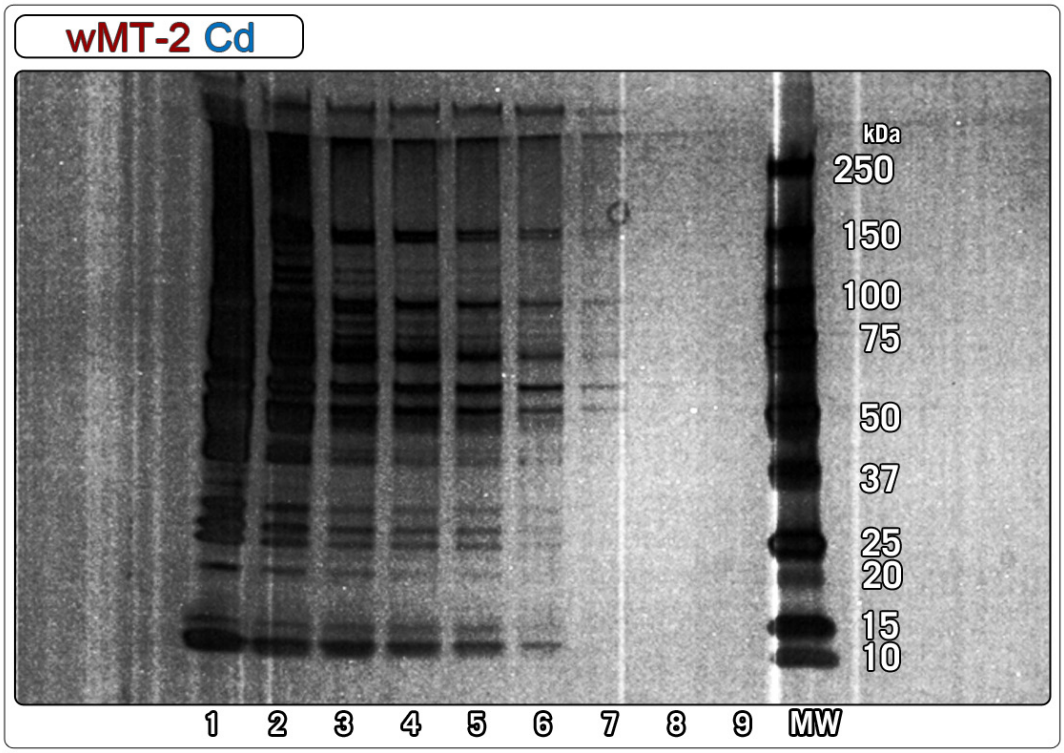


Figure 3.30. SDS-PAGE gel of Cd-S•Tag-wMT-2 SEC Fractions 1-9 from pooled 0 % Buffer B obtained by AnCat chromatography (**Figure 3.29**). Biorad Dual Colour Prestained Standard, designated **MW**.

With the absence of S•Tag-wMT-2 confirmed in the fractions of 0 % Buffer B, Fractions 10 and 11 from **Figure 3.27** were concentrated by slow centrifugation (2,000 x *g*). In an attempt to decrease the amount of salt within the sample, an additional dilution (with 20 mM ammonium bicarbonate) and centrifugation was performed before FPLC. The supernatant was filtered and loaded onto the SEC column as one injection (**Figure 3.31**). A_{254} was used again, to assist the determination of those fractions containing S-Cd bonds. This was helpful, as the detector at 220 nm was heavily overloaded. However, ICP-OES showed that Fraction 5 contained a significant amount of cadmium, and SDS-PAGE (**Figure 3.32**) supported the presence of S•Tag-wMT-2.

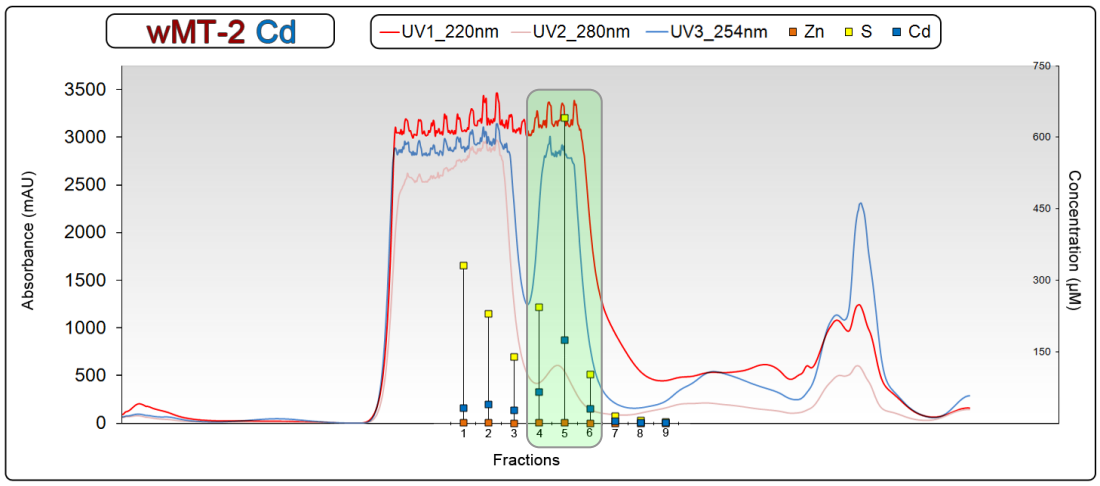


Figure 3.31. SEC purification of the 100 % Buffer B fractions obtained from AnCat chromatography (**Figure 3.27**); UV absorbances measured at **220 nm**, **280 nm** and **254 nm**. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). S•Tag-wMT-2 is indicated in **GREEN**. Samples begin at 52 mL retention volume.

As Fraction 5 contains the least contamination, it was chosen for mass spectrometry. Concentrations estimated by ICP-OES were 588.0 µM S, 176.4 µM Cd, giving an S:Cd ratio of ≈ 3.3:1, which was close to the ratio of 3:1 S:Cd for Cd_7-S•Tag-wMT-2.

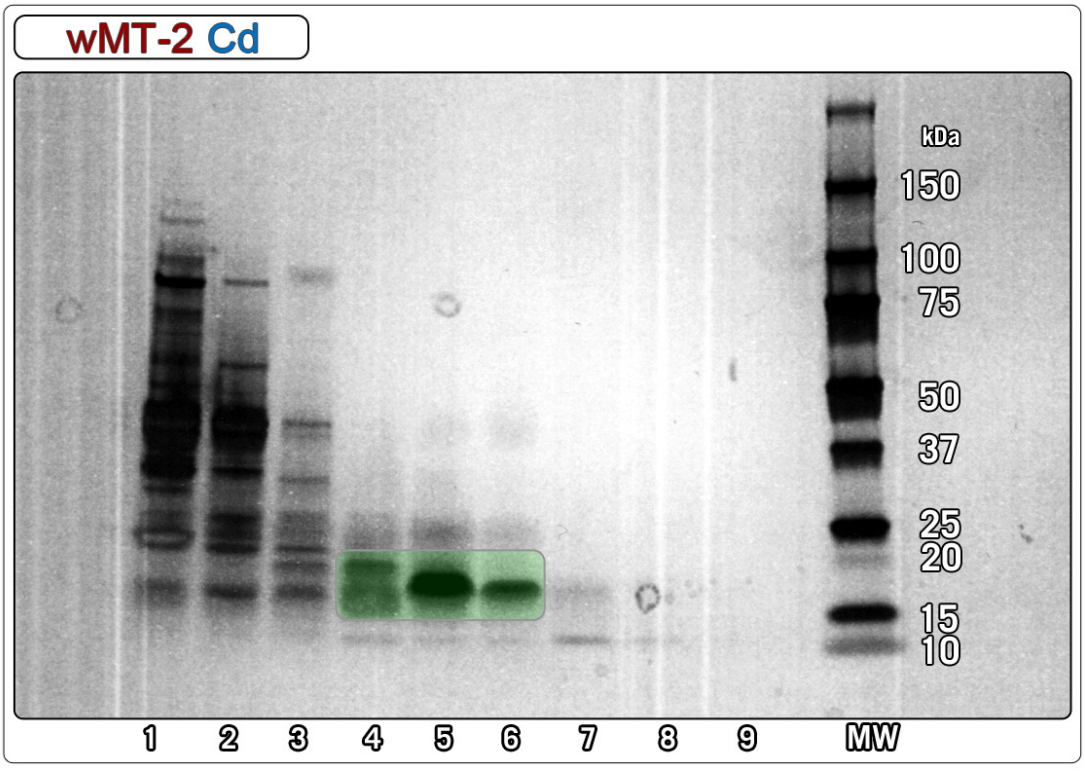


Figure 3.32. SDS-PAGE gel of 100% Buffer B Fractions from **Figure 3.27**, separated by SEC in **Figure 3.31**. Biorad Dual Colour Prestained Standard, designated **MW**. S•Tag-wMT-2 is indicated in **GREEN**.

To remove the maximum amount of salt before mass spectrometry, Fraction 5 was applied to a PD-10 column as per manufacturer's instructions [GE Healthcare], causing a slight dilution of the sample.

S•Tag-wMT-2 was then concentrated by centrifugation at a slow speed (3,000 x *g*), to achieve an MS sample of ≈ 40 μM. Clearly, Cd-S•Tag-wMT-2 has been isolated from the 100 % Buffer B fractions (**Figure 3.33**). However, even after 2 centrifugation steps, an SEC column, and a PD-10 desalting column, there was a significant amount of sodium remaining, causing sodium adducts in the spectrum (**Figure 3.33B**). The salt also appears to have caused a slight unfolding of the protein, as high charge states (> 6+) were not observed for cleaved Cd-wMT-2 purified without AnCat. This suggests that the salt has caused the protonation of non-surface sites on the protein, which could affect metal-binding studies if performed on these samples.

Using ion exchange chromatography in this way was somewhat successful – it did allow the purification of a pure sample of S•Tag-wMT-2. Whilst salt from exchange chromatography did not appear to be a problem for the purification of Cd-S•Tag-wMT-1 (the mass spectrum showed few sodium adducts), for Cd-S•Tag-wMT-2 the quantity of salt carried through from ion exchange chromatography makes it unsuitable for mass spectrometry. Therefore a purification method was attempted without the introduction of sodium ions.

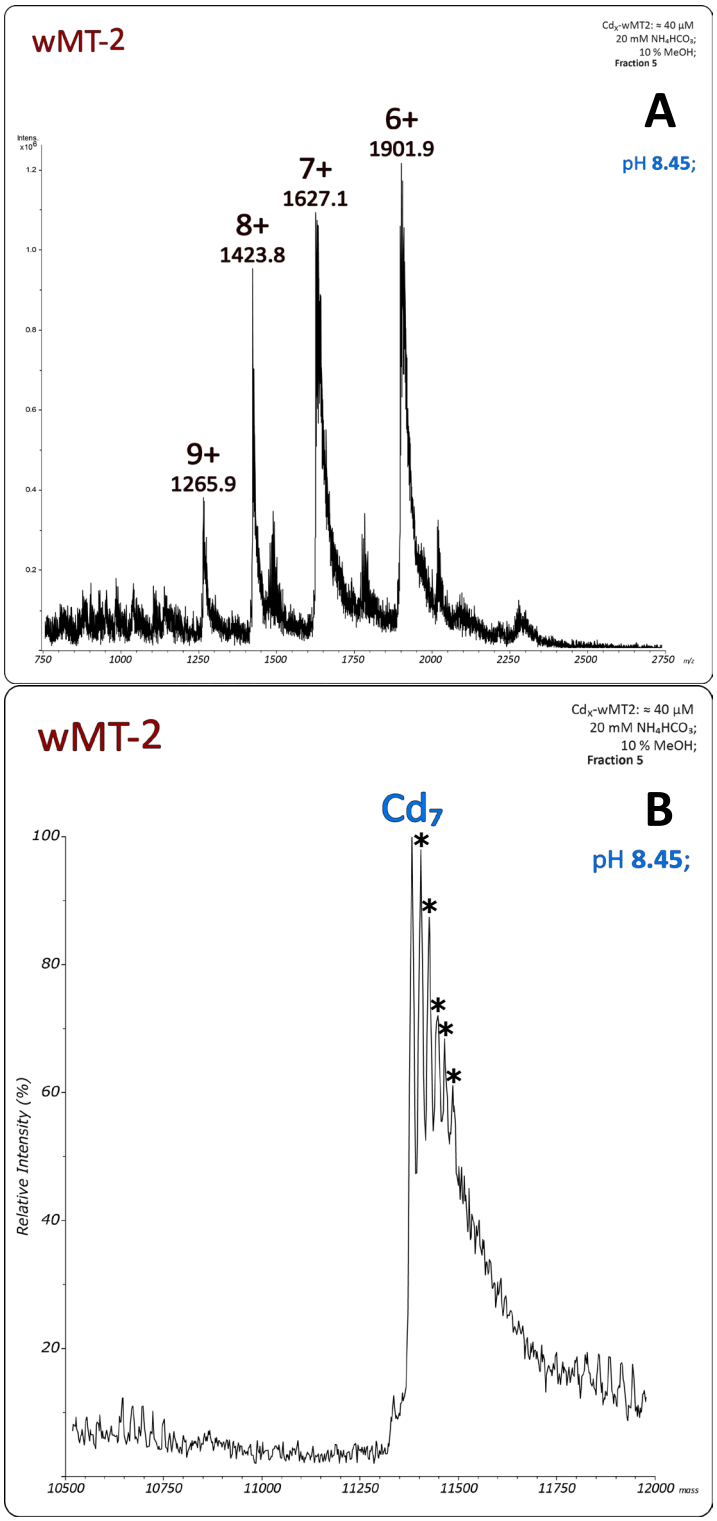


Figure 3.33A. Raw mass spectrum of S•Tag-wMT-2 indicating the charge states used for deconvolution, **B.** shows the deconvoluted mass spectrum of the above peaks. S•Tag-wMT-2 concentration ≈ 40 µM, pH 8.45. Sample in 20 mM NH_4HCO_3, and 10 % MeOH. Single metallospecies were observed, with significant amounts of sodium adducts (+23 mass units, *). Peak masses shown in **Appendix 7**.

Chemical precipitation

Chemical precipitation is used to either concentrate a target protein, or to remove contaminants and other proteins. Attractive and repulsive forces between proteins are partially negated by solvation in water. Displacement of the water solvation layer that forms around proteins by organic solvents (causing the water molecules to form layers around the organic solvents instead) will cause attractive electrostatic and dipole forces to lead to aggregation of protein [211]. Organic solvents lower the dielectric constant of water, allowing similar charges to exist in close proximity, promoting protein aggregation. The temperature is kept close to 0 °C, to minimise protein denaturation [211]. Initially three separations were performed based on the volume of lysate, with addition of ¾ vol, 1¾ vol and 3 vol of ethanol:chloroform (100:8). To ensure that proteins could precipitate quantitatively, each solution was stored at -20 °C for > 12 hrs. SDS-PAGE was used to identify which volumes contained potential candidates for S•Tag-wMT-2 (**Figure 3.34**).

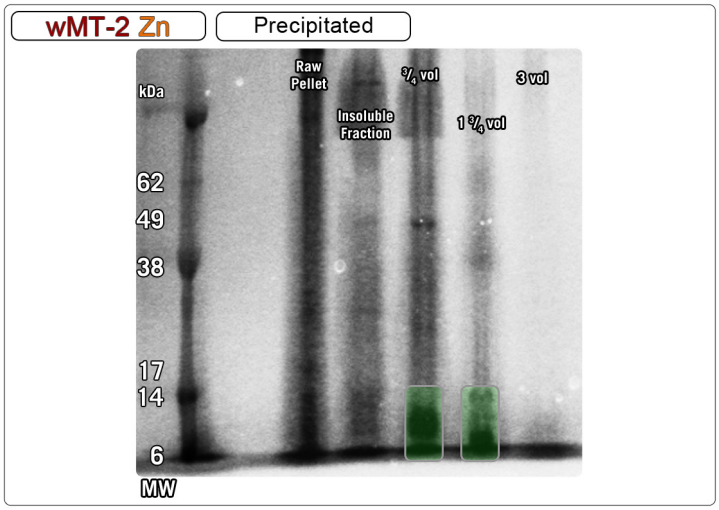


Figure 3.34. SDS-PAGE gel of the fractions from chemical precipitation. Biorad Dual Colour Prestained Standard, designated **MW**. Fractions thought to contain S•Tag-wMT-2 are indicated in **GREEN**. The poor resolution is thought to be due to difficulty in resuspending the precipitate.

Whilst the gel is not well resolved, there does appear to be some separation of protein within the mixture. Due to the long incubation times, the total time for purification was approximately 2 days from the start of purification. Although the additional time would have been acceptable if the separation of S•Tag-wMT-2 from contaminant proteins had been significant, the size-exclusion chromatograms of the three fractions showed considerable overlap in the region thought to contain S•Tag-wMT-2 (**Figure 3.35**).

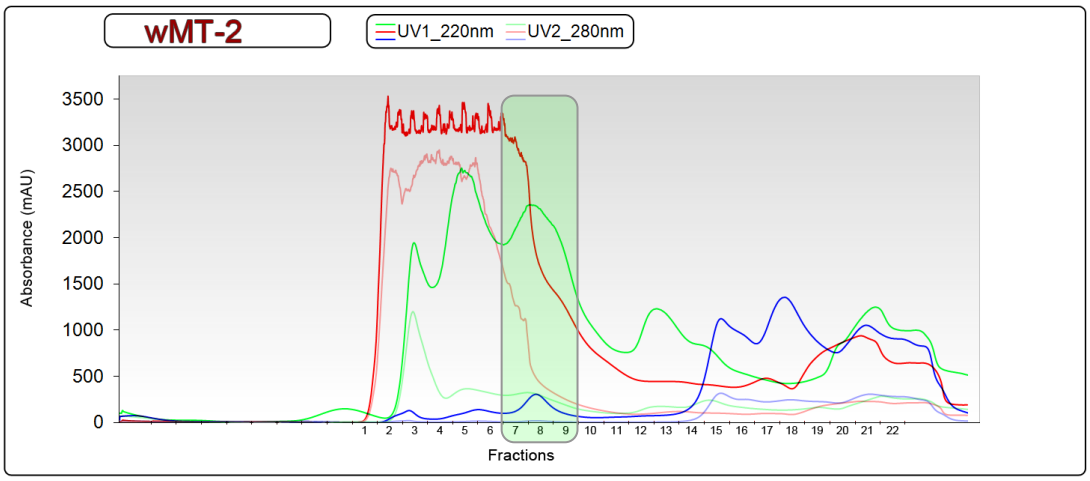


Figure 3.35. FPLC purification by SEC of volumes from ethanol:chloroform precipitation. UV absorbances measured for ¾ vol (**220 nm**, 280 nm); 1¾ vol (**220 nm**, 280 nm); 3 vol (**220 nm**, 280 nm). Fractions thought to contain S•Tag-wMT-2 highlighted in **GREEN**.

As this generated > 50 samples, sulphur concentration was determined using a cysteine assay. All samples were analysed in this manner, with the region expected to contain S•Tag-wMT-2 showing a protein concentration of 1-2 μM (**Table 3.1**).

Table 3.1. Table indicating the fractions for which the total protein concentration is over 2 µM for the fractions in **Figure 3.35**. *Df* is the dilution factor (40x) used to measure the samples.

Samples	A_{412}	[Cys] µM	[Prot] µM	[Prot] x *df* µM
0.75 vol *2*	*0.007*	*1.32*	*0.07*	*2.6*
8	*0.017*	*1.89*	*0.09*	*3.8*
1.75 vol *4*	*0.007*	*1.32*	*0.07*	*2.6*
5	*0.015*	*1.77*	*0.09*	*3.5*
8	*0.009*	*1.44*	*0.07*	*2.9*
10	*0.014*	*1.72*	*0.09*	*3.4*
13	*0.001*	*0.98*	*0.05*	*2.0*

Fraction 8 (**Figure 3.35**) in both ¾ and 1¾ volumes showed a relatively high concentration of sulphur (> 1 µM sulphur), however it would appear that the yield of S•Tag-wMT-2 was significantly lower than previous experiments without precipitation. Considering the purpose of the experiment, to increase the yield of pure protein, it would seem that ethanol:chloroform precipitation is not suitable. The next course of action was to attempt a salt precipitation.

Salt precipitation

Another common way of removing bulk contaminants or precipitating target protein is the use of ammonium sulphate. The theory is similar to that of chemical precipitation, as protein solubility within the sample is altered through dehydration and increasing protein-protein interactions. Protein solubility is heavily influenced by the ionic strength of the solution; addition of a highly water soluble salt, such as ammonium sulphate, will increase the ionic strength of the solution to the point where most proteins precipitate. After salt precipitation, salt must be removed by a technique such as SEC. SDS-PAGE of sequential increases

in 0.5 M ammonium sulphate concentration were poorly resolved (data not shown), however, S•Tag-wMT-2 was estimated to precipitate between 2.0-2.5 M ammonium sulphate concentration.

In an attempt to maintain S•Tag-wMT-2 in solution, only 1.5 M final ammonium sulphate concentration was used in the purification. If S•Tag-wMT-2 was resistant to resolubilisation, then maintaining it in solution should increase the final yield of S•Tag-wMT-2. The pellet was divided, and a parallel purification either with or without salt precipitation were compared (**Figure 3.36**).

For the salt precipitated sample, the FPLC chromatograms were not well resolved (**Figure 3.36B**), with ICP-OES suggesting a decrease of ≈ 20 % in protein yield after precipitation; and SDS-PAGE indicating a reduction of fraction purity (**Figure 3.36C** & **D**). These experiments showed that there is no major benefit in using a salt precipitation step at this stage of the purification. A comparison with the cell pellet that was not precipitated with salt showed that moderately pure S•Tag-wMT-2 could be obtained through one sole SEC separation. If thrombin cleavage was utilised, then this protocol may be sufficient to obtain pure S•Tag-wMT-2.

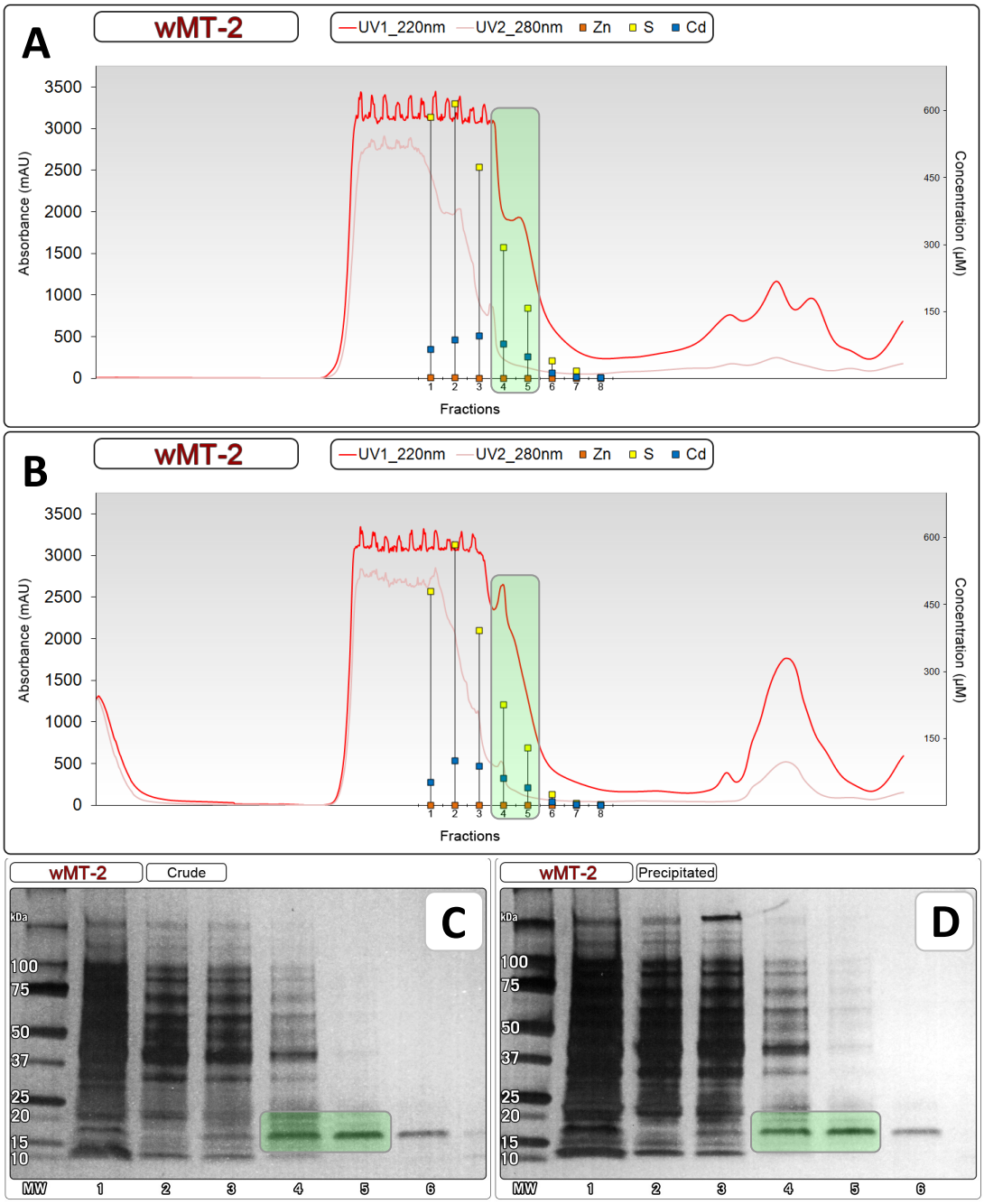


Figure 3.36. FPLC purification by SEC; **A** is without salt precipitation, **B** is with salt precipitation. UV absorbances measured at **220 nm** and 280 nm . ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). SDS-PAGE gels **C** & **D** shown with Biorad Dual Colour Prestained Standard, designated **MW**. S•Tag-wMT-2 highlighted in **GREEN**, notice in well 5 that there are more contaminants present in **D**, after salt precipitation.

Using SEC as the initial purification step

Sonication buffer was supplemented with 1 mM $CdCl_2$, and the sample sonicated as before. The sonicated mixture was filtered before separation by SEC. **Figure 3.37A** shows a distinct peak in the A_{220} chromatogram of the Cd-S•Tag-wMT-2 chromatogram, which is in the region where S•Tag-wMT-1 eluted (≈ 64 mL). SDS-PAGE (**Figure 3.37B**) also shows that fractions within **Figure 3.37A** thought to contain S•Tag-wMT-2 (Fractions 6-9) were moderately pure. Fractions 7-8 had the highest proportion of S•Tag-wMT-2, with a small amount in Fractions 6 and 9. Measured S:Cd ratios for Fractions 6-9 were 3.1:1, 3.0:1, 2.8:1 and 2.6:1 respectively. Fraction 7 had the highest sulphur concentration, and had the closest S:Cd ratio to that expected for pure Cd_7-S•Tag-wMT-2 (3.0:1). Fraction 7 was concentrated to approximately 35 µM S•Tag-wMT-2, and its identity was confirmed by mass spectrometry (**Figure 3.38**). The MS results showed that at neutral pH or above, there was only one observable species, Cd_7-S•Tag-wMT-2. Acidification to pH ≈ 2 generated the apo- mass and a number of other species, varying from Cd_6- to apo-S•Tag-wMT-2. Of note is the good agreement between the calculated and observed masses shown in **Appendix 8**.

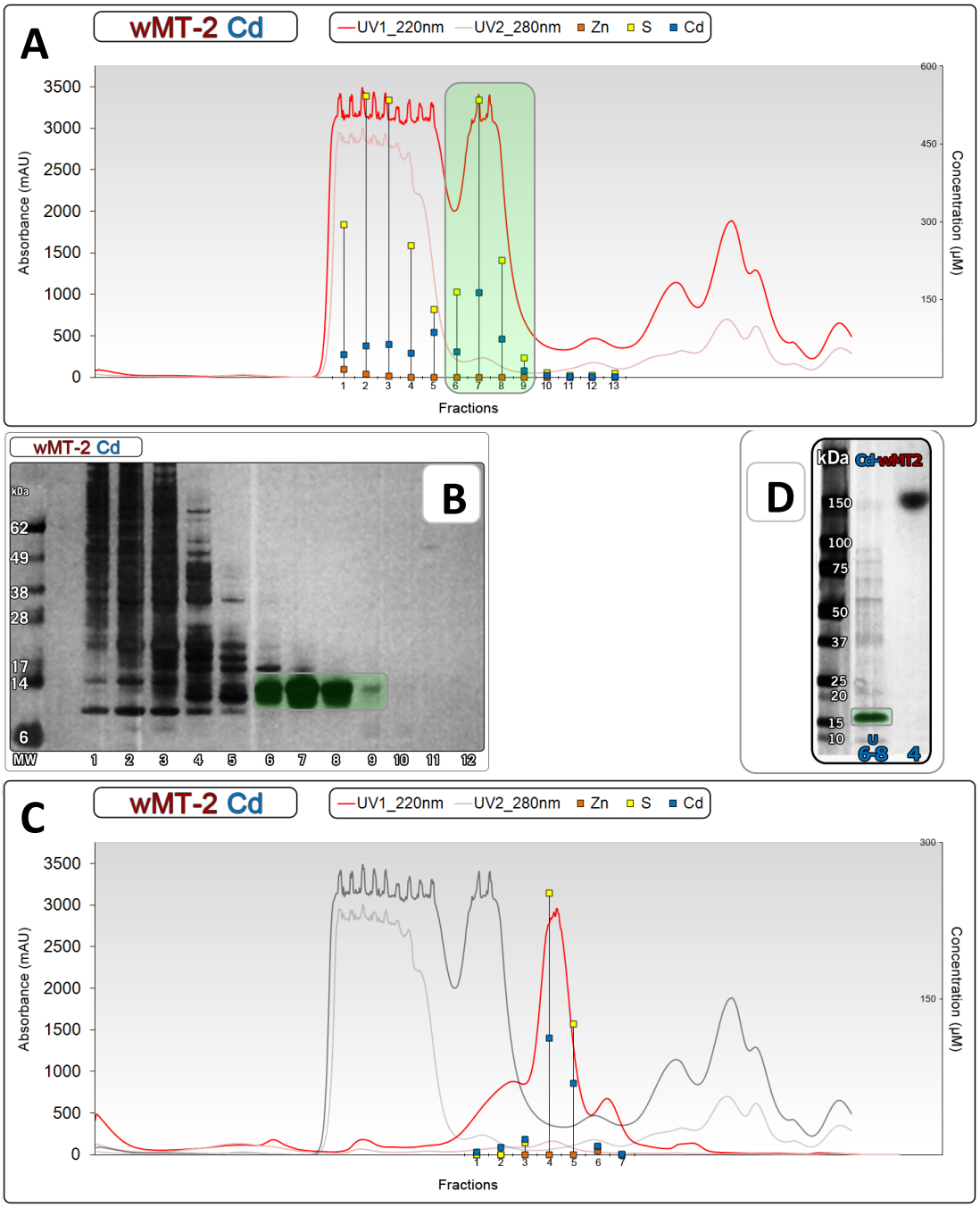


Figure 3.37A. SEC of S•Tag-wMT-2 & **C** SEC of cleaved wMT-2. UV absorbances measured at **220 nm** and 280 nm . ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). The GREY chromatography lines indicate the previous spectrum of (**A**). **B** is an SDS-PAGE gel of Cd-S•Tag-wMT-2 from (**A**). **D** shows a comparison of selected Fractions from (**A**) and (**C**). The bands in Fractions 6-9 (**GREEN**), indicate a protein of a molecular mass similar to S•Tag-wMT-2. Fractions in (**A**) begin at a retention volume of 40 mL, Fractions in (**C**) begin at a retention volume of 64 mL.

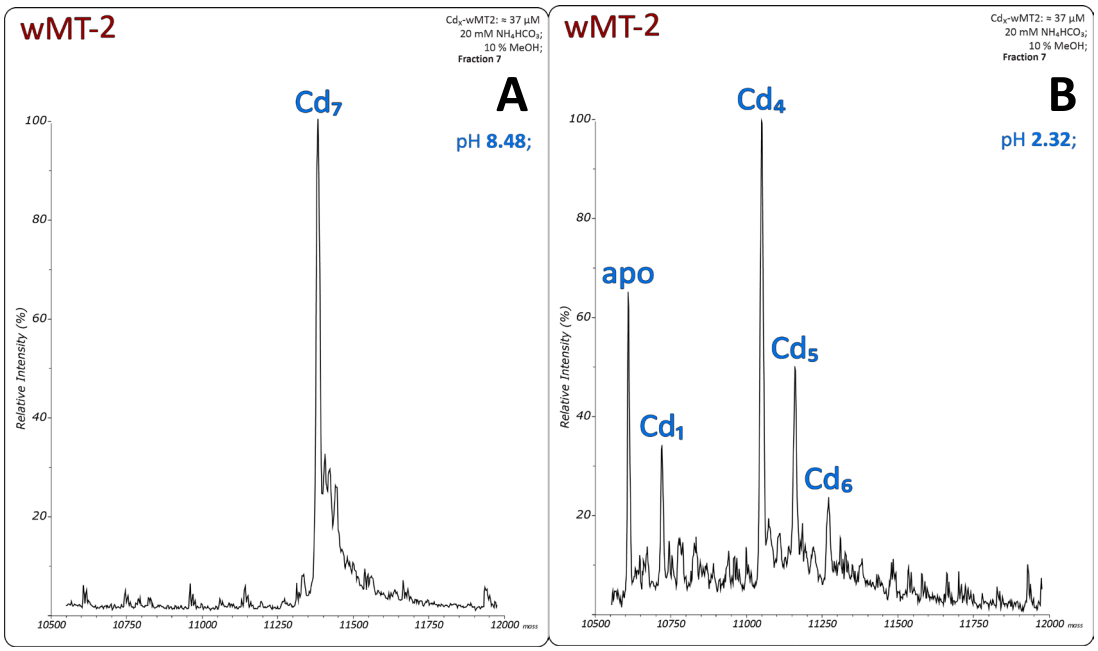


Figure 3.38. ESI-MS of Cd-S•Tag-wMT-2, Fraction 7 from SEC (**Figure 3.37A**). S•Tag-wMT-2 concentration ≈ 37 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is at pH 8.48, spectrum (**B**) is after addition of 20 % acetic acid, which enabled the observation of the apo- mass of the protein at pH 2.32.

After confirmation of Cd_7-S•Tag-wMT-2 by mass spectrometry, S•Tag-wMT-2 was concentrated to 1 mg/mL before thrombin cleavage. After cleavage, the protein was separated by SEC. **Figure 3.37C** shows that there was, as expected, a shift in the wMT-2 peak to the right (from 68 mL to 76 mL), indicating a protein of a smaller hydrodynamic volume. This is expected, as the S•Tag is approximately 3 kDa, or $^1/_3$ of the total size of the fusion protein. By reducing the size of wMT-2, trace contaminants were removed by relocating the peak (through the decrease in size) to a contaminant-free area of the elution profile. Fractions 7 & 8 were pooled and concentrated for further analysis, generating a pure final sample, as estimated by SDS-PAGE. Curiously, cleaved Cd-wMT-2 presented at approximately 150 kDa in the gel (**Figure 3.37D**). This is ten times higher than expected, but even so, this single band corresponded unmistakably to cleaved wMT-2, as shown by

the mass spectrum in **Figure 3.39A**. There could be two explanations for this behaviour: oligomerisation/aggregation of the protein; or incomplete denaturation and linearization of the protein in the reducing buffer. Before mass spectrometry, Fractions 4 and 5 were pooled and concentrated to approximately 35 µM with respect to sulphur concentration, or 52 µM with respect to cadmium concentration, as determined by ICP-OES. Inaccuracy in ICP-OES measurements caused some ambiguity in the concentration and stoichiometry for wMT-2. Stoichiometry by MS (**Figure 3.39**) was determined to be 6.4 cadmium ions per wMT-2 molecule, caused by the quantity of Cd_4- present in the cleaved spectrum at neutral pH. The major species however, is still Cd_7-wMT-2 showing the ability of cleaved wMT-2 to bind seven divalent metal ions. For the cleaved spectrum, the concentration was approximately the same as before thrombin cleavage, but the observed species were decidedly different.

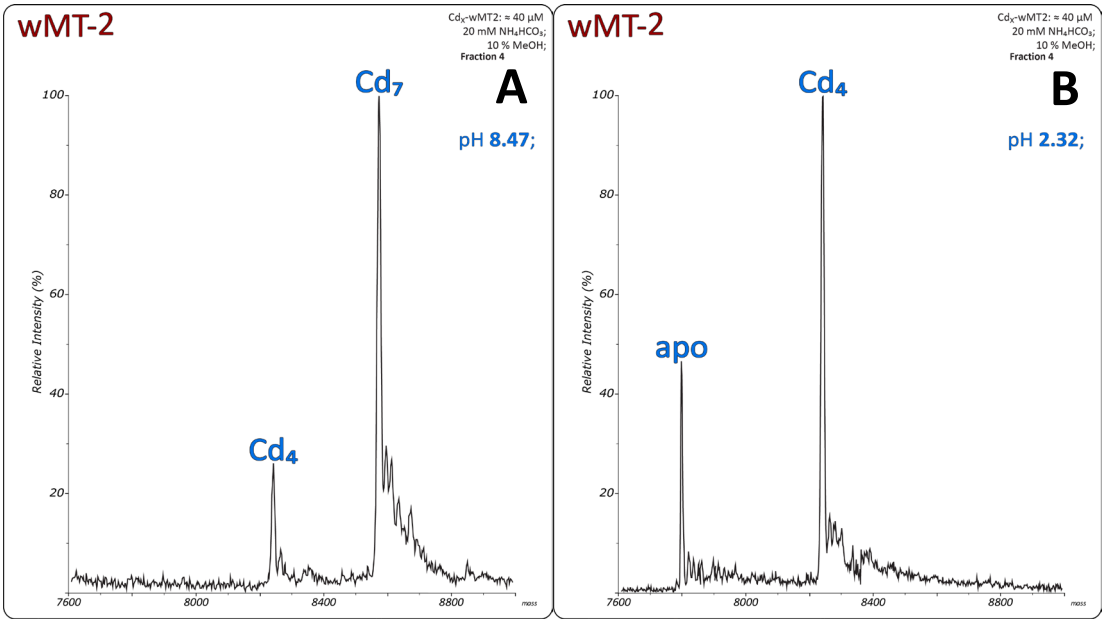


Figure 3.39. ESI-MS of cleaved Cd-wMT-2, Fractions 4-5 from SEC (**Figure 3.37C**). wMT-2 concentration ≈ 35 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is cleaved wMT-2 at pH 8.47, spectrum (**B**) is cleaved wMT-2 after addition of 20 % acetic acid, which enabled the observation of the apo- mass of the protein, also at pH 2.32.

The S•Tag may be stabilising the fully-loaded species, as at a similar pH, these intermediate species were not present in the S•Tag-wMT-2 spectrum (**Figure 3.38**). The observed masses were very close to that of the theoretical masses, given in **Appendix 9**: Cd_7-wMT-2 8,570.3 (8,570.8 Da theoretical); Cd_4-wMT-2 8,240.3 (8,239.6 Da theoretical); apo-wMT-2 7,796.4 (7,797.9 Da theoretical). The final yield of protein purified in this way was approximately 2-2.25 mg/L of medium, meaning that the method described for purifying wMT-2 was successful in obtaining sufficient quantities of pure cleaved Cd-wMT-2. This same method was then utilised for the successful production of cleaved Zn-wMT-2.

3.11 S•Tag-wMT-2 expressed with zinc, purified with zinc

S•Tag-wMT-2 was purified as before with the addition of 1 mM $ZnSO_4$ before sonication. Crude lysate was filtered and applied to an SEC column (**Figure 3.40A**). A well defined peak (Fractions 7 and 8), distinct from the main bulk of contaminants thought to contain Zn-S•Tag-wMT-2 was observed in the spectrum. There was also a peak in the concentration of zinc, around Fraction 7 (**Figure 3.40A**), matched by a spike in the sulphur concentration. This was in a similar location as Cd-S•Tag-wMT-2 (**Figure 3.37**).

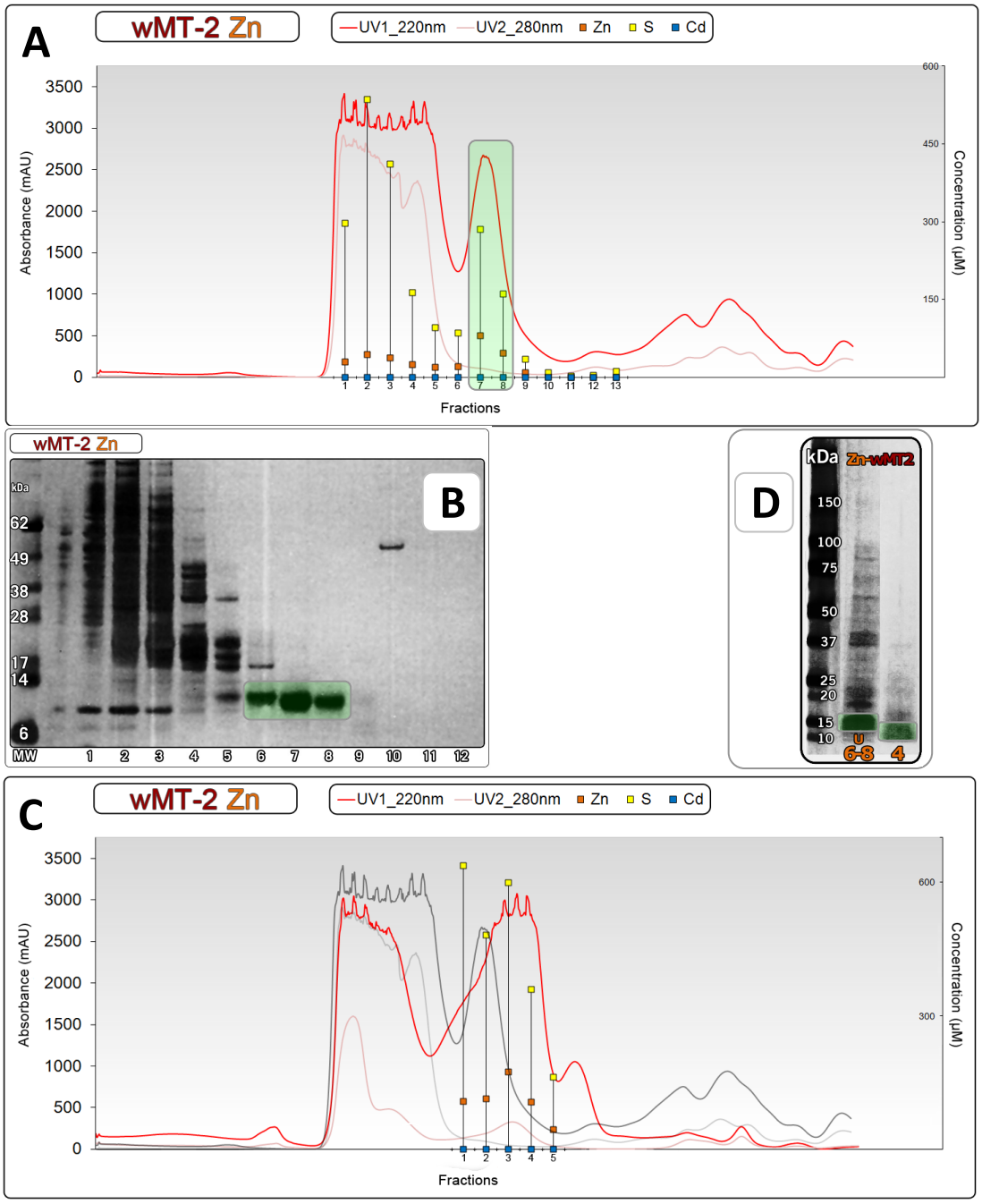


Figure 3.40A. SEC of S•Tag-wMT-2 & **C** SEC of cleaved wMT-2. UV absorbances measured at **220 nm** and 280 nm . ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). The GREY chromatography lines indicate the previous spectrum of (**A**). **B** is an SDS-PAGE gel of Zn-S•Tag-wMT-2 from (**A**). **D** shows a comparison of selected Fractions from (**A**) and (**C**). The bands in Fractions 6-9 (**GREEN**), indicate a protein of a molecular mass similar to S•Tag-wMT-2. Fractions in (**A**) begin at a retention volume of 40 mL, Fractions in (**C**) begin at a retention volume of 60 mL.

The magnitude of the peak was also similar, but slightly smaller than that obtained from purifying Cd-S•Tag-wMT-2. This indicates that *E. coli* will express S•Tag-wMT-2 equally well, if not better, when cultured with

cadmium, rather than with zinc. This finding has been reported in other studies with MTs [212, 213].

SDS-PAGE (**Figure 3.40B**) shows Fractions 7 & 8 to be purest, with a small contamination around 17 kDa in Fraction 6. The S:Zn ratios of Fractions 6-8 were 3.6:1, 3.2:1 and 3.1:1 respectively, the optimum being an S:Zn ratio of 3:1 corresponding to Zn_7-wMT-2. To obtain a mass spectrum of Zn-S•Tag-wMT-2, Fraction 7 was concentrated to ≈ 35 µM, with a S:Zn ratio of 3.2:1.

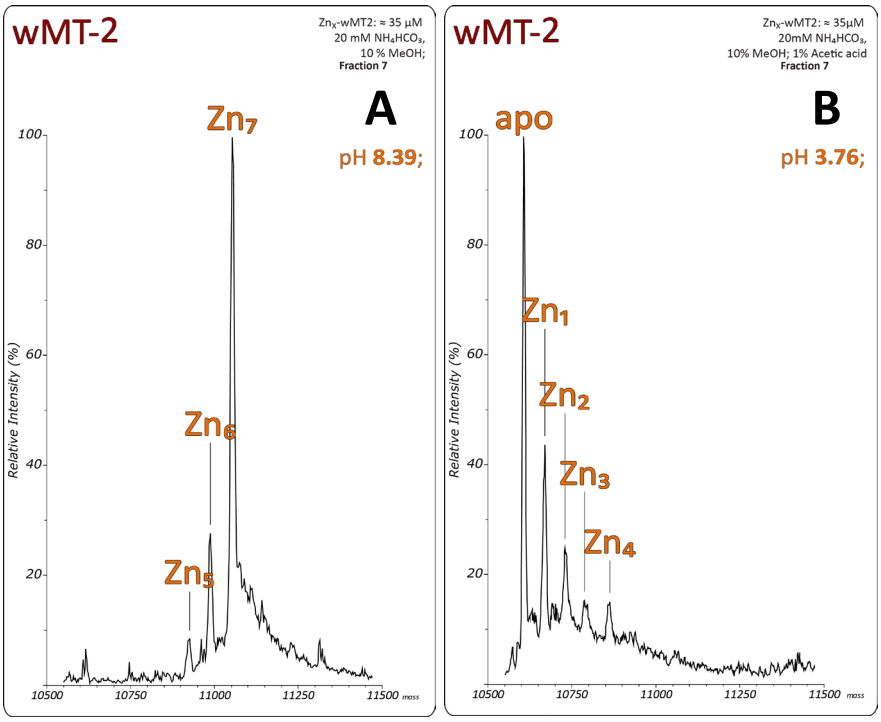


Figure 3.41. ESI-MS of Zn-S•Tag-wMT-2, Fraction 7 from SEC (**Figure 3.40A**). S•Tag-wMT-2 concentration ≈ 35 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is at pH 8.39, spectrum (**B**) is after addition of 1 % acetic acid, which enabled the observation of the apo- mass of the protein at pH 3.76. Accurate masses provided in **Appendix 10**.

Analysis by MS shows significantly more species than seen with cadmium, as at high pH (> pH 8) there were 3 species. The major

species however, is Zn_7-S•Tag-wMT-2. This could indicate differences in metal-binding affinities for cadmium and zinc.

To obtain cleaved mass spectra, the remainder of Fractions 7 and 8 were pooled and concentrated by centrifugation. Some precipitate was visible after centrifugation, which can be assumed from SDS-PAGE (**Figure 3.40B**) to be aggregated wMT-2. Cleavage was performed identically to that of Cd-S•Tag-wMT-2, and filtered before being separated by SEC (**Figure 3.40C**). Similar to the behaviour of cleaved Cd-wMT-2, a defined shift in the cleaved Zn-wMT-2 peak to the right (as expected from a reduction in molecular mass) was observed. However, it does appear that there was a certain amount of protein remaining in Fractions 1 & 2 (**Figure 3.40C**) that was left uncleaved. The column was also overloaded at this point, but even so, the S:Zn ratios for Fractions 3-5 were 3.4:1, 3.4:1 and 3.7:1 - close to the S:Zn ratio of 3:1 for Zn_7-wMT-2. The purification appeared to work well, as Fraction 4 appeared almost without contamination (**Figure 3.40D**).

Concentration to 35 µM and subsequent mass spectrometry (**Figure 3.42**) showed a significant increase in baseline noise over that of Zn-S•Tag-wMT-2 (**Figure 3.41**). This could be due to the higher ionisation efficiency of S•Tag-wMT-2. Another possibility is that the estimated 35 µM wMT-2 concentration was incorrect, and/or a proportion of protein was lost to aggregation at the concentration stage.

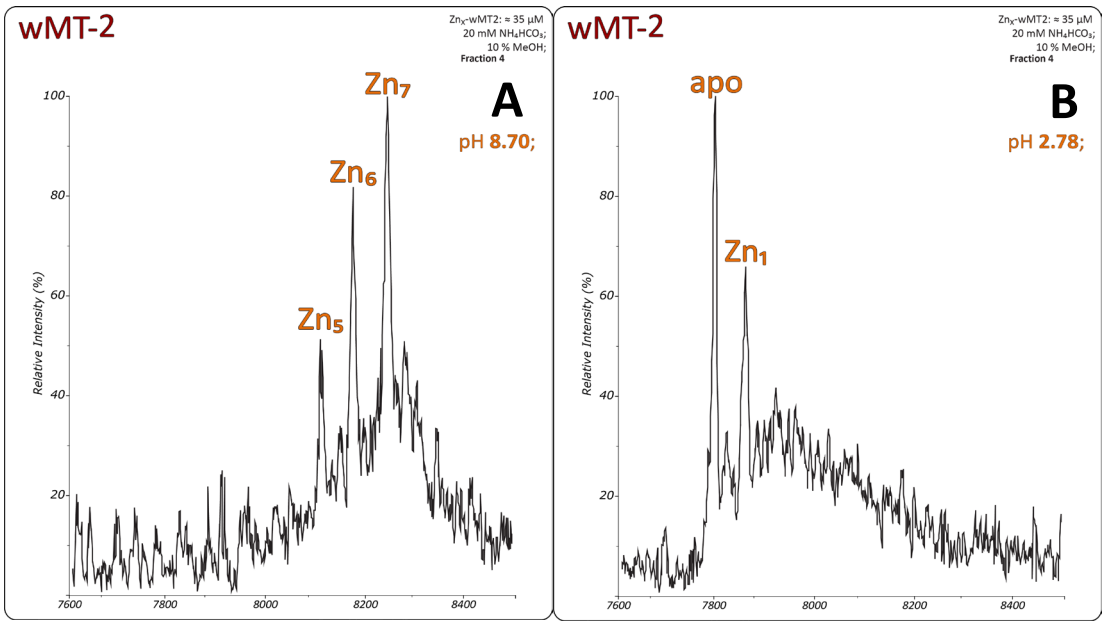


Figure 3.42. ESI-MS of cleaved Zn-wMT-2, Fraction 4 from SEC (**Figure 3.40C**). wMT-2 concentration ≈ 35 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is cleaved wMT-2 at pH 8.70, spectrum (**B**) is cleaved wMT-2 after addition of 2 % acetic acid, which enabled the observation of the apo- mass of the protein at pH 2.78. Masses provided in **Appendix 11**.

Similar species were observed to those in the Zn-S•Tag-wMT-2 spectrum (**Figure 3.41**), with Zn_7- to Zn_5- being major species. If all species in the mass spectrum of cleaved Zn-wMT-2 are analysed (**Figure 3.42**), a value of approximately 6.2 zinc ions per wMT-2 molecule is calculated, slightly lower than that reported for Zn-S•Tag-wMT-2. Interestingly there are 3 major species, cleaved Zn_7-, Zn_6- and Zn_5-wMT-2, which is similar to that observed with cleaved Zn-wMT-1 (**Figure 3.16**). A further consideration is that proteins observed by mass spectrometry often contain multiple species if they contain a metal that is not 'native' [214]. For example, wMT-2 is postulated as the sole cadmium responsive MT in earthworms, hence hypothesised to be a cadmium-binding MT, and might therefore bind zinc in a less well-defined manner.

Observed masses for the cleaved spectra were somewhat more divergent from the calculated masses (**Appendix 11**), the main source of error being the low signal/noise ratio reducing the accuracy of the mass data.

3.12 wMT-3 expressed with cadmium, purified with cadmium

The plasmid encoding the *wMT-3* gene was obtained from the Stürzenbaum group. Initial attempts to express the construct were successful, but a metal-binding product approximately 2 kDa larger than Cd-wMT-3 was isolated. Attempts to sequence the plasmid by both myself, and within the Stürzenbaum group, could not generate acceptable data. A second clone, however, was received which was confirmed to contain the *wMT-3* gene of the correct sequence. However, the resultant protein sequence was slightly different to that first reported in 2004 (**Figure 3.43**). This is due to the 2004 sequence being derived from crude consensus sequences from earthworms in multiple locations (S. Stürzenbaum, personal communication). The new construct was created from earthworms in a single location, with the most abundant sequence being cloned.

```
wMT-3 (2004)    1 MADAAVPCNK LTKCCGKTSC PREGSKCACT NCKCVKGECL PNCDKDCCGA
wMT-3 (Seq.)    1 MADAAVPCNK LTKCCGKTSC PREGSKCVCT NCKCVKGECL PNCDKDCCGG

wMT-3 (2004)   51 TEQCASKCGN PNCKCGADCK CAPGQCTTEC AKGCCE 86
wMT-3 (Seq.)   51 TEQCASKCGN ANCKCGADCK CAPGQCTTEC AKGCCE 86
```

Figure 3.43. The wMT-3 sequences obtained from DNA sequencing of the construct (**Seq.**) compared to reported in 2004 [111]. Cysteine residues are coloured **RED**. Residues that are different to those reported in the 2004 study coloured **GREEN**.

The main difference in the purification of wMT-3 is that the construct obtained from the Stürzenbaum group expressed full-length native

protein, without N-terminal S•Tag. This makes this construct similar to the subcloned wMT-1 plasmid. With the success of expressing and purifying in the presence of cadmium, for both S•Tag-wMT-2 and S•Tag-wMT-1, it was chosen to express wMT-3 in LB, in the presence of cadmium. The first stage was to sonicate the cell pellet in sonication buffer, with the addition of 1 mM $CdCl_2$. The crude lysate was filtered before separation by SEC (**Figure 3.44**).

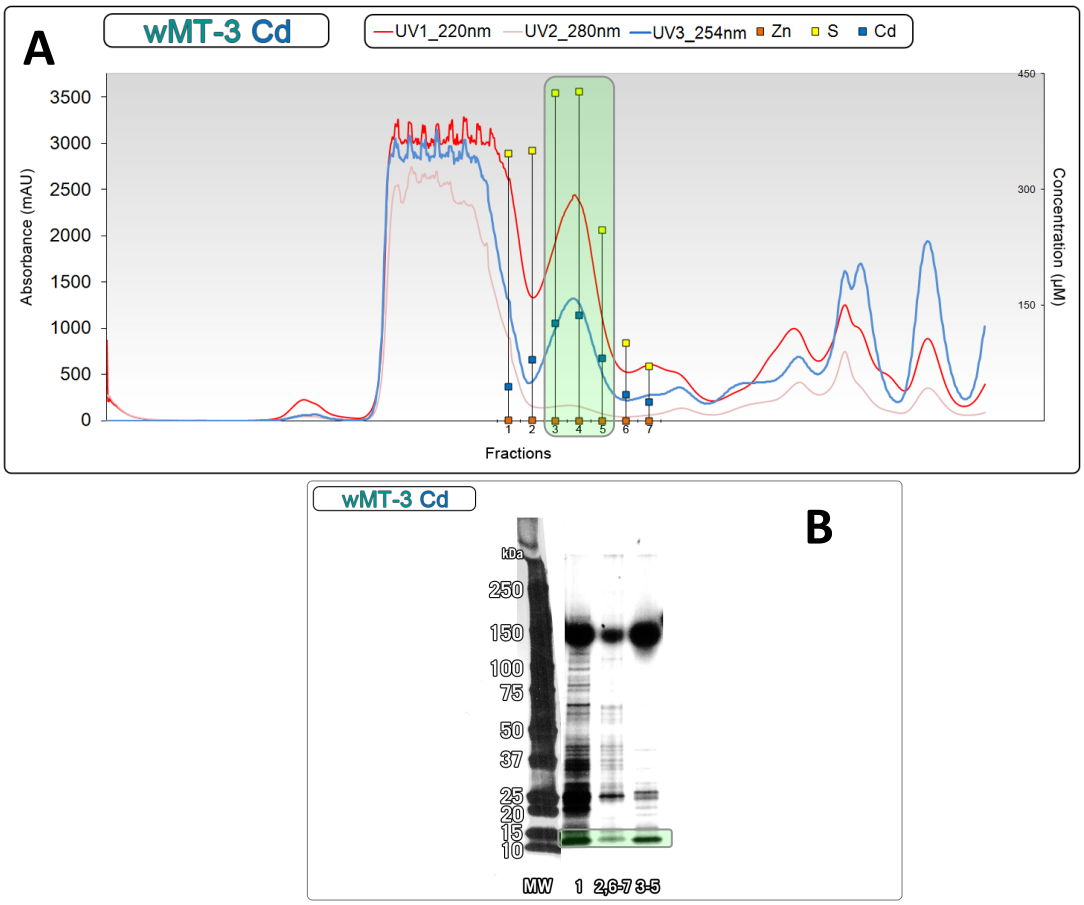


Figure 3.44A. FPLC purification by SEC of wMT-3. UV absorbances measured at **220 nm**, **280 nm** and **280 nm**. ICP-OES concentrations are overlaid (**ZINC**, **SULPHUR**, **CADMIUM**). **B** is an SDS-PAGE gel of pooled Fractions from (**A**), shown with Biorad Dual Colour Prestained Standard, designated **MW**. The pool containing Fractions 3-5 (**GREEN**), indicate a protein of a molecular mass similar to wMT-3. A band present at approximately 150 kDa is similar to that observed with cleaved Cd-wMT-2. Fractions in(**A**) begin at a retention volume of 58 mL.

The chromatogram from SEC indicates a well resolved peak, distinct to the main bulk of contaminants (those eluting before Fraction 2,

Figure 3.44A). The absorbtion at 254 nm was also measured for this separation, to potentially observe the effect of S-Cd bonds, bearing in mind there could be some interference from neighbouring wavelengths. The peak proposed for wMT-3 (Fractions 3-5, **Figure 3.44A**) showed strong absorbance at this wavelength, potentially indicating many S-Cd bonds. SDS-PAGE (**Figure 3.44B**) used combined fractions for ease of analysis, of most interest are the two bands observed across the gel: One at approximately 12 kDa; a second at approximately 150 kDa.

The lower of these two bands is close to the expected apo- molecular mass of wMT-3 (8,779 Da), the upper may reflect similar behaviour to that observed with cleaved Cd-wMT-2, with an apparent molecular mass of around 150 kDa. Although there appear to be many contaminant proteins, S:Cd ratios for Fractions 3-5 were 3.2:1, 3.0:1 and 2.9:1 respectively. The expected S:Cd ratio for Cd_7-wMT-3 is 3.1:1, as wMT-3 contains 22 cysteine residues. Mass spectrometry (**Figure 3.45**) showed the observed mass corresponded to full-length wMT-3 lacking the initiation N-terminal methionine, so this was not included in calculation of theoretical mass (**Appendix 12**). This is different to both cleaved wMT-1 and cleaved wMT-2 due to their N-termini containing residues remaining from thrombin cleavage (-GSM-). Pooled Fractions 3-5 showed the highest proportion of the band at 10 kDa, so were concentrated by centrifugation to 50 µM, filtered, and analysed by mass spectrometry.

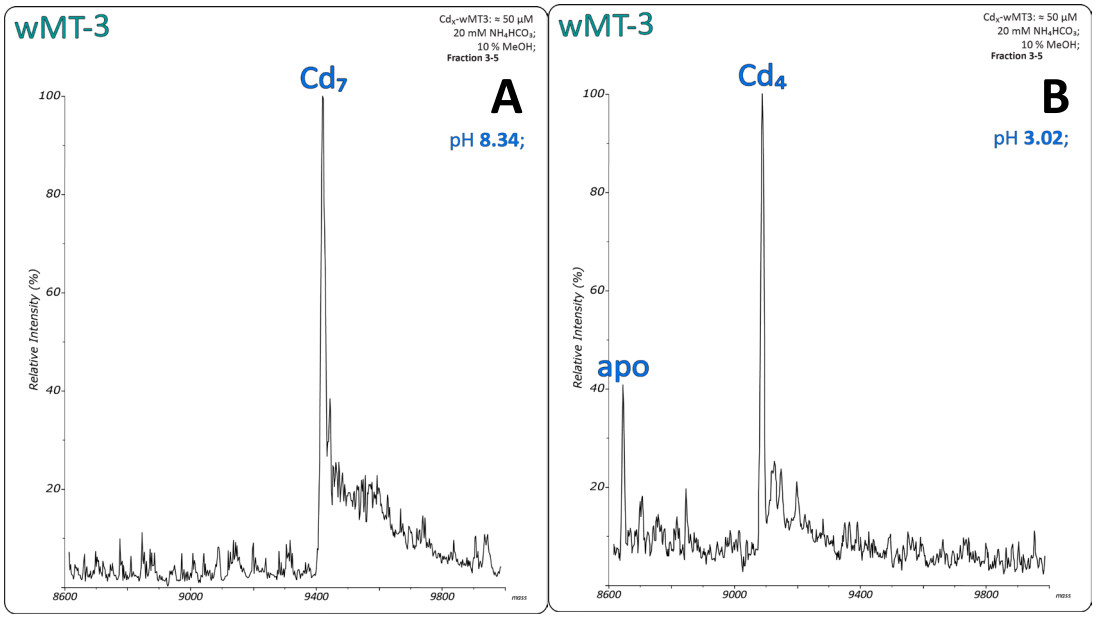


Figure 3.45. ESI-MS of Cd-wMT-3, Pooled Fractions 3-5 from SEC (**Figure 3.44**). wMT-3 concentration ≈ 50 µM. Samples in 20 mM NH_4HCO_3, and 10 % MeOH. Spectrum (**A**) is at pH 8.34, spectrum (**B**) is after addition of 2 % formic acid, which enabled the observation of the apo- mass of the protein at pH 3.02. Masses presented in **Appendix 12**.

The mass spectrum (**Figure 3.45**) is fairly conclusive, as the major species is of Cd_7-wMT-3, showing that wMT-3 has been isolated for the first time. There is a small amount of noise to the right of the Cd_7- peak, which could potentially conceal a Cd_8- peak, but this is not reflected in the S:Cd ratios observed by ICP-OES. Decreasing the pH allowed observation of the apo-mass, and also the Cd_4- species. Interestingly this behaviour was similar to that of cleaved Cd-wMT-2 .

The construct used in this expression was designed to produce wMT-3 without an S•Tag. Initial data demonstrate that significant quantities (≈ 1 mg/L) of well-defined, full-length wMT-3 can be obtained using this approach. For future studies, a further purification step will be required to separate the protein from remaining contaminants. In addition, expression in the presence of zinc should also be attempted.

3.13 Summary

In this chapter I have outlined the development of successful expression and purification methods for obtaining, close to their native sequence, all three currently known wMTs in *Lumbricus rubellus*. In the case of cleaved wMT-1, some further refinement is required, as the purity of the fractions showed contamination in both MS and SDS-PAGE experiments (**Figures 3.16 / 3.23**). However this did not hinder the recording of MS experiments that showed fully-metallated cleaved wMT-1 with both zinc and cadmium. For S•Tag-wMT-2, the protocol was modified to minimise the loss of protein from multiple purification steps. This method is however, unsuitable for purifying S•Tag-wMT-1, with the protein peak not well resolved by performing SEC without ion exchange chromatography (**Figure 3.10**). Fortunately this is not the case for S•Tag-wMT-2, and the problematic ion exchange step could be omitted. This has enabled purification of both cleaved Zn-wMT-2 and cleaved Cd-wMT-2, with good purity, as observed by very few contaminants in MS and SDS-PAGE (**Figures 3.39 / 3.42**). Using the general parts from the purification of S•Tag-wMT-2, wMT-3 (without S•Tag) was able to be isolated in the presence of cadmium (**Figure 3.45**). At this time, purification of Zn-wMT-3 has not been performed.

With the success of wMT-3 expression in a construct without S•Tag, questions are raised as to why this would not be the case for subcloned wMT-1. There is no doubt that the chromatograms obtained for wMT-3 and wMT-1 from SEC of their crude lysate are significantly different, with the absence of a significant peak at 60-70 mL retention volume in the subcloned wMT-1 chromatogram (**Figure 3.8**). With the difficulty in expressing the first wMT-3 construct, it is assumed that the problem with subcloned wMT-1 was more likely at the expression stage, rather than at the purification stage. **Table 3.2** (below) shows a summary of results within this Chapter, including the purification route, and the major species observed when purified with zinc or cadmium.

Table 3.2. Tabulated summary of results from Chapter3. wMT-1 data from **Figure 3.16** (Zn) and **Figure 3.23** (Cd); wMT-2 data from **Figure 3.39** (Cd) and **Figure 3.42** (Zn); wMT-3 data from **Figure 3.45** (Cd), no Zn-wMT-3 data recorded at this time.

Isoform	Purification method	Zn Major Species		Cd Major Species	
		pH >8	pH <4	pH >8	pH <4
wMT-1	**Cation Exchange -> SEC -> Cleavage -> SEC**	**Zn_7**, Zn_6, Zn_5	**apo**, Zn_1	Cd_8, **Cd_7**	Cd_5, **Cd_4**, Cd_1, apo
wMT-2	**SEC -> Cleavage -> SEC**	**Zn_7**, Zn_6, Zn_5	**apo**, Zn_1	**Cd_7**, Cd_4	**Cd_4**, apo
wMT-3	**SEC**	-	-	**Cd_7**	**Cd_4**, apo

Moving forwards, the purification of the isoform found to be isolated in the greatest quantity, Cd-S•Tag-wMT-2, was expanded to larger culture volumes (approximately 1,200 mL). This enabled purification of sufficient quantities of cleaved Cd-wMT-2 required for solution NMR structure determination, reported in Chapter 4. This is followed by a comparison of the metal-binding dynamics of the three isoforms, and the effect of competitive metal chelators such as EDTA and 5F-BAPTA on cleaved Cd-wMT-2 in Chapter 5.

4

Structural Investigation of cleaved Cd-wMT-2

4.1 Introduction

The link between the structure and function of proteins is now well established, and as such the structural determination of proteins is a quickly expanding field. The first protein structure deposited into the Research Collaboratory for Structural Bioinformatics (RCSB) protein data bank was in 1976, obtained by X-ray crystallography. It took until 1989 before the first protein structure was determined by solution NMR (a 43 residue protein [215]), coinciding with the availability of powerful 2-dimensional NMR experiments such as COSY and NOESY. These NMR experiments were pioneered at the end of the 1970s, meaning that in just 10 years, a viable alternate route to crystallography for structure determination was available. Nowadays there are significantly more powerful pulse sequences available for solution structure determination. Whilst X-ray crystallography is still popular, of the current 83,983 protein structures in the protein data bank, 9,586 have been determined by solution NMR[*]. X-ray crystallography has been used widely in protein structure determination, however due to their flexibility and lack of overall rigid structure MTs are very difficult to crystallise. There are disadvantages to using solution NMR for structure determination, such as the significant time required for analysis of data. However, due to the

[*] http://www.pdb.org/pdb/static.do?p=general_information/pdb_statistics/index.html

inherent flexibility of the linker region between the two domains, MT exhibits a dynamic non-rigid structure [216, 217]. Solution NMR becomes an invaluable tool in the structural investigation of MTs.

The nomenclature used during the discussion of NMR experiments will be in relation to the standard protein model in **Figure 4.1**. This defines the carbon atoms on the backbone as C_O and C_α, with subsequent carbon atoms labelled incrementally. Hydrogen atoms are labelled with respect to the nitrogen / carbon atom they are bonded to, ie all hydrogen atoms bonded to the C_β atom are labelled H_β. There are also four specific angles which are utilised for structure calculations: ϕ (phi) is defined as: C_O-**N-C_α**-C_O; ψ (psi) is defined as: N-**C_α-C_O**-N ; ω (omega) is defined as: C_α-**C_O-N**-C_α; τ (tau) is defined as: N-**C_α**-C_O.

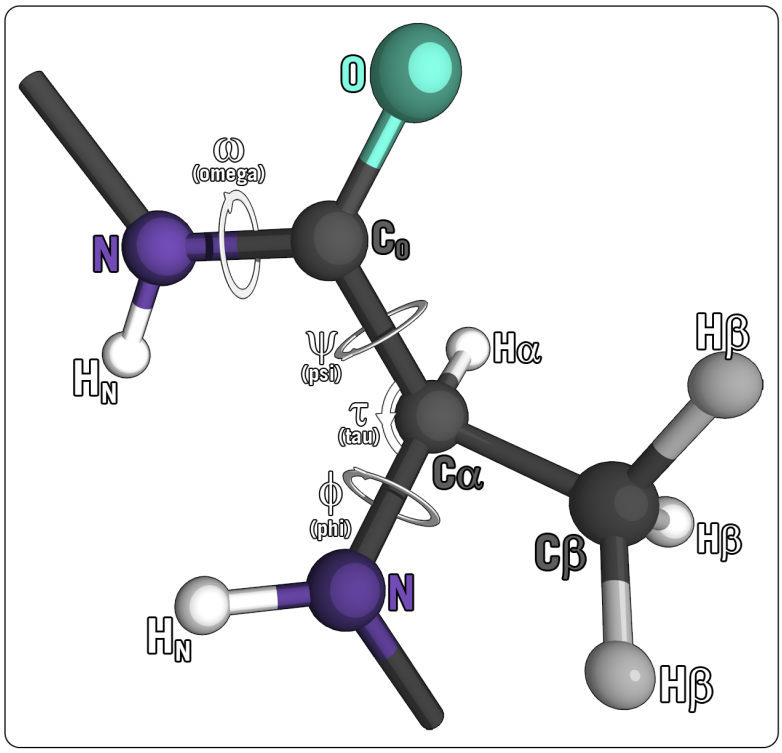


Figure 4.1. Cartoon of a protein backbone showing the names of angles and atoms within an amino acid. Note that the view is centred around an alanine residue, with an additional nitrogen atom from a subsequent amino acid present to include the ω angle.

4.2 Investigation of temperature effects by NMR

To ascertain the stability, and optimise the NMR conditions for wMT-2, a variable temperature experiment was performed to establish whether temperature caused a significant sharpening (or broadening) effect of NMR proton signals due to correlation between lower relaxation time in the transverse plane (x-y) with increased temperature [218].

It was a concern that high temperature would significantly decrease the lifetime of the protein, therefore it was of benefit to use the lowest possible temperature whilst still achieving well-resolved spectra. To monitor the effect of temperature, spectra from 278 K to 308 K (in increments of 10 K) were recorded. To determine an optimum temperature, a peak in the downfield region > 9.5 ppm was used. The isolated nature of this peak made it a useful probe for monitoring the effect of changing temperature. Due to the difficulties in establishing a comparative baseline over the four spectra, determining peak integral area proved to be problematic. Therefore the peak width at half height was chosen as the parameter to be compared. Although there were significant changes in the entire fingerprint region as temperature increased from 278 K (**Figure 4.2A**), further increase in temperature above 298 K appeared to give no significant benefit in spectral dispersion, resolution or peak intensity. The measured peak widths at half height for each peak in **Figure 4.2B** were: 0.6885 Hz, 0.7138 Hz, 0.6025 Hz, 0.6499 Hz for 278-308 K respectively; a lower value indicating a sharper peak. Examining peak intensity and peak width at

half height, the ideal temperature for experiments with Cd-wMT-2 was close to 298 K.

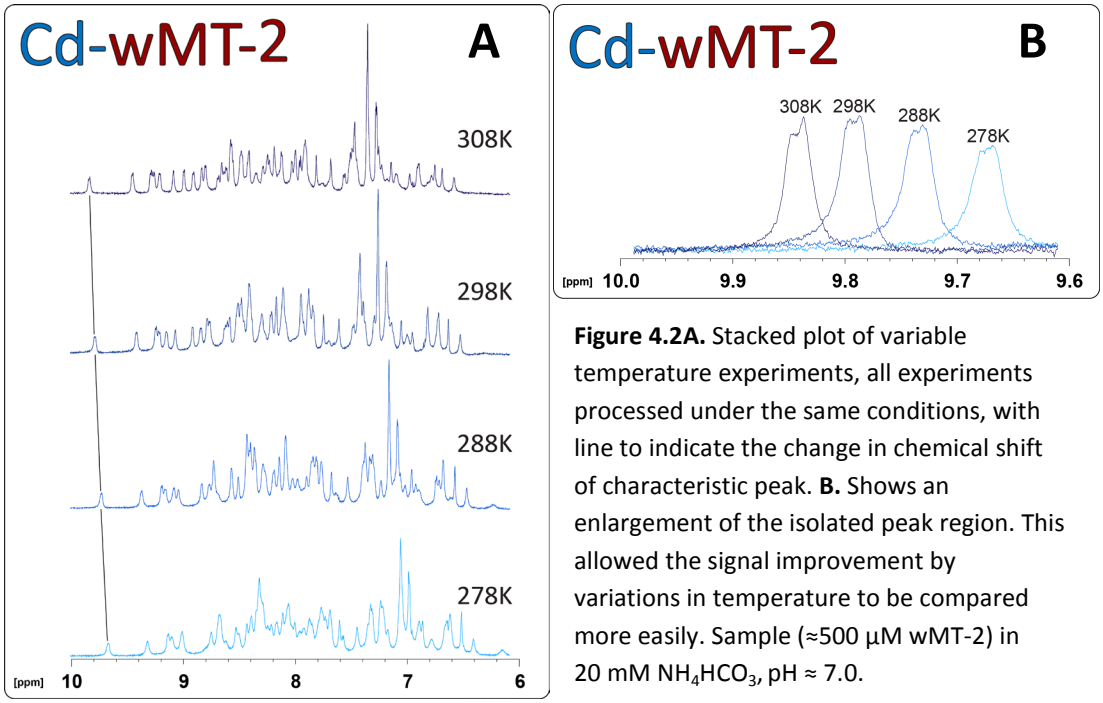


Figure 4.2A. Stacked plot of variable temperature experiments, all experiments processed under the same conditions, with line to indicate the change in chemical shift of characteristic peak. **B.** Shows an enlargement of the isolated peak region. This allowed the signal improvement by variations in temperature to be compared more easily. Sample (≈500 µM wMT-2) in 20 mM NH_4HCO_3, pH ≈ 7.0.

4.3 Sequential assignment

[^{1}H, ^{1}H] TOCSY, NOESY

The TOCSY (Total correlation spectroscopy) experiment is an improvement over the original COSY (Correlation spectroscopy) experiment, which enables the correlation of hydrogen atoms within an amino acid residue. Whereas COSY generates crosspeaks for hydrogen atoms generated by 3-bond (^{3}J) scalar coupling (within 3 bonds - ie H_N to Hα, Hα to Hβ and H_N), TOCSY can in principle show the crosspeaks of *all* hydrogen atoms in a residue (ie H_N to Hα, Hβ, Hγ; Hα to H_N, Hβ, Hγ) due to many ^{3}J transfers occurring during a mixing step which keeps all spins together (also called spin-lock) [188]. The schematics of the two experiments are shown in **Figure 4.3**.

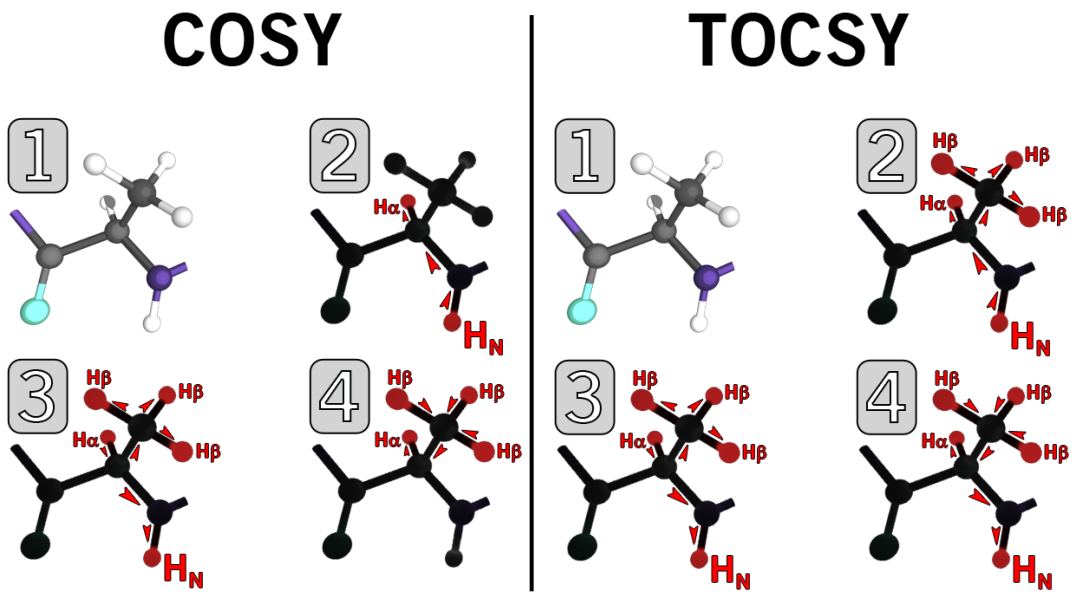


Figure 4.3. Schematic of the COSY (**LEFT**) and TOCSY (**RIGHT**) experiments. **RED** indicates the direction of magnetisation transfer. In the COSY spectrum, the H_N residue can form a cross-peak with the Hα atom, through transfer of magnetisation over 3 bonds (Pane 2). The Hα can transfer magnetisation to the H_N, and also the three Hβ atoms (Pane 3). The Hβ atoms can only transfer back to the Hα (Pane 4). For TOCSY, crosspeaks between every hydrogen atom on the residue are possible.

Assignment of the Cd-wMT-2 [^{1}H, ^{1}H] TOCSY experiment proceeded using the two stage method, first reported by Wüthrich in 1982 [219, 220]. As magnetisation is transferred through-bonds to all of the hydrogen atoms within an amino acid, the crosspeaks of different amino acids will form characteristic patterns. These patterns can be grouped as belonging to one of 5 broad types of spin system. The first stage was to identify the spin system type (**Figure 4.4**): either J-type (S, D, N, C, W, F, Y, H), U-type (K, R, E, M, Q, and P), A/T-type, V/I/L-type or glycine. Due to the amino acid composition of wMT-2, there are significantly more J-type spin systems than U-type, with: 37 J-type; 17 U-type, 15 A/T-type, 1 leucine and 9 glycines. The majority of the analysis of the NMR spectra utilised the fingerprint region of the spectrum (6-10 ppm in the

F2 dimension), however peaks in the aliphatic portion (Hα-Hβ and Hα-CH_3) were also assigned where possible.

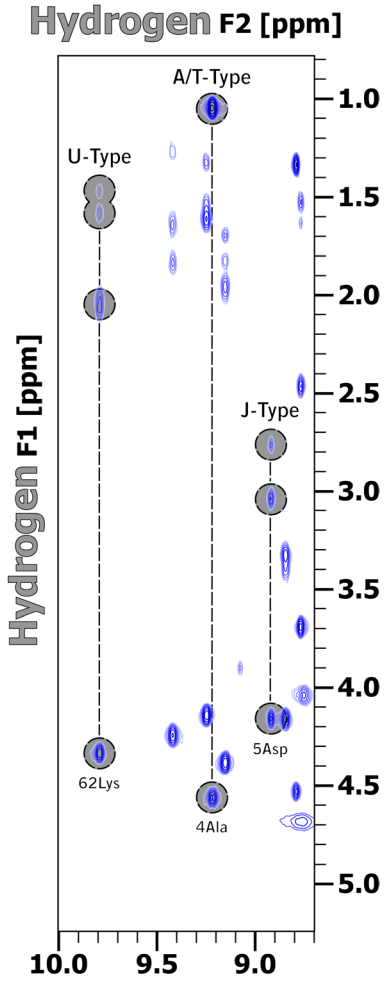


Figure 4.4. A section of the TOCSY spectrum between 8.7-10.0 in the F2 dimension of Cd-wMT-2. This region is downfield of the bulk of peaks in the fingerprint region of the spectrum. Shown are 3 types of spin systems, with their characteristic patterns (residue-specific patterns in **Appendix 13**). The U-type spin system has an Hα crosspeak ≈ 4.5 ppm, with contributions from any Hβ, Hγ, Hδ and Hε in the region between 1.0-3.0 ppm in the F1 dimension. The A/T-type spin system has a Hα crosspeak ≈ 4.5 ppm, with Hβ ≈ 1 ppm for alanine, and Hα / Hβ ≈ 4.5 ppm with Hγ ≈ 1 ppm for threonine. The J-type spin system has an Hα crosspeak ≈ 4.5 ppm, with the two Hβs in the region between 2.5-4.0 ppm.

The second stage was to assign the spin systems to a particular residue in the amino acid sequence using nuclear Overhauser effect spectroscopy (NOESY [221]) data, which gives information on the type of neighbouring spin systems. In NOESY experiments, magnetisation is transferred through-space between hydrogen atoms and is detectable for interactions within ≈ 5 Å (**Figure 4.5**) [188]. Interactions between nuclear spins in close proximity cause an intensity enhancement through transfer of spin polarisation via mutual relaxation of two nuclei (called the nuclear Overhauser effect) [188]. NOESY peaks are most often found

between backbone H_N and the Hα of neighbouring residues, but depending on backbone conformation, longer range NOEs can be observed.

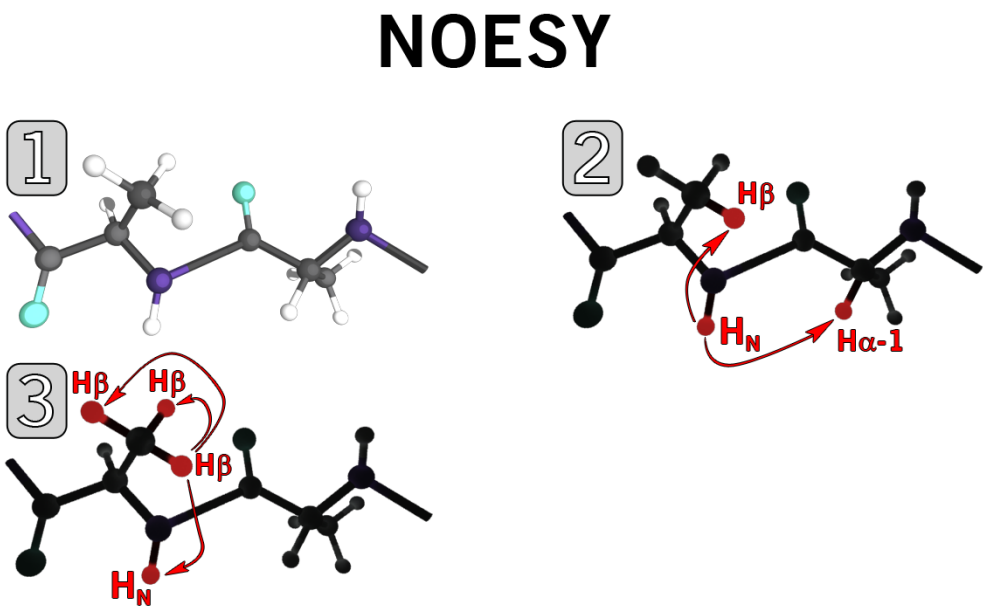


Figure 4.5. Schematic of a NOESY NMR experiment. **RED** indicates the direction of magnetisation transfer. In the NOESY spectrum, crosspeaks are formed between atoms in close spatial proximity. Crosspeaks from neighbouring residues appear in NOESY spectra (Pane 2), therefore based on the spin system type of nearby residues, a picture of the entire backbone can be created, aiding unambiguous assignment of residues.

Using the protein primary sequence, unique motifs can begin to be identified. One of the starting points for this assignment was the unique leucine (43Leu) residue, which had a characteristic spin pattern (**Appendix 4**). A backbone walk can be performed by selecting either a TOCSY crosspeak or NOESY crosspeak. Using a TOCSY crosspeak as a starting point, a corresponding NOESY crosspeak can be located in the same hydrogen dimension as either the H_N or H_α. Once a NOESY crosspeak is found, a corresponding TOCSY crosspeak in the second hydrogen dimension can be identified. The NOESY peak linking these two hydrogen chemical shifts indicates that the hydrogen atoms of the two residues are within 5 Å of each other. Examining **Figure 4.7**, the TOCSY/NOESY backbone walk revealed a stretch of 5 amino acids. The

exact sequence is determined to be U-type - U-type - L-type - J-type - J-type. There are four sections within the wMT-2 sequence with two subsequent U-types together, and ten sections with two adjacent J-types (**Figure 4.6**).

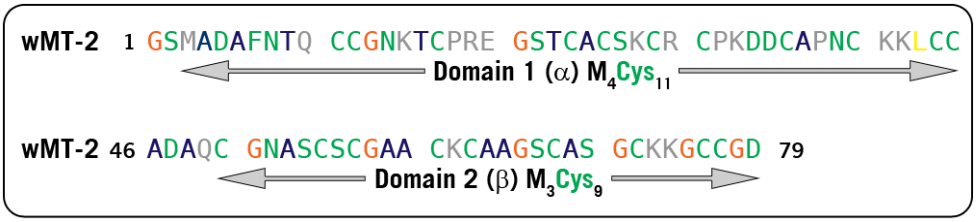


Figure 4.6. The sequence of cleaved wMT-2, with colours indicating the different spin systems: **J-type** ; **U-type** ; **G** ; **A/T-type** ; **L**.

However utilising the unique leucine residue, these five residues were unambiguously assigned to the -**KKLCC**- block of the sequence, corresponding to residues 41-45. The right hand side of **Figure 4.7** shows all assigned NOEs for each residue (not just those used in the TOCSY/NOESY backbone walk). Of note are some of the longer range NOEs in this portion of the sequence - for instance some of the interactions between the backbone H_N of 42Lys, and the Hβs of 45Cys. Although many motifs could be identified from the TOCSY and NOESY spectrum, there were many peaks that could not be unambiguously assigned from these spectra alone. To aid in assignment, a wMT-2 sample was expressed in medium containing ^{13}C and ^{15}N isotopes.

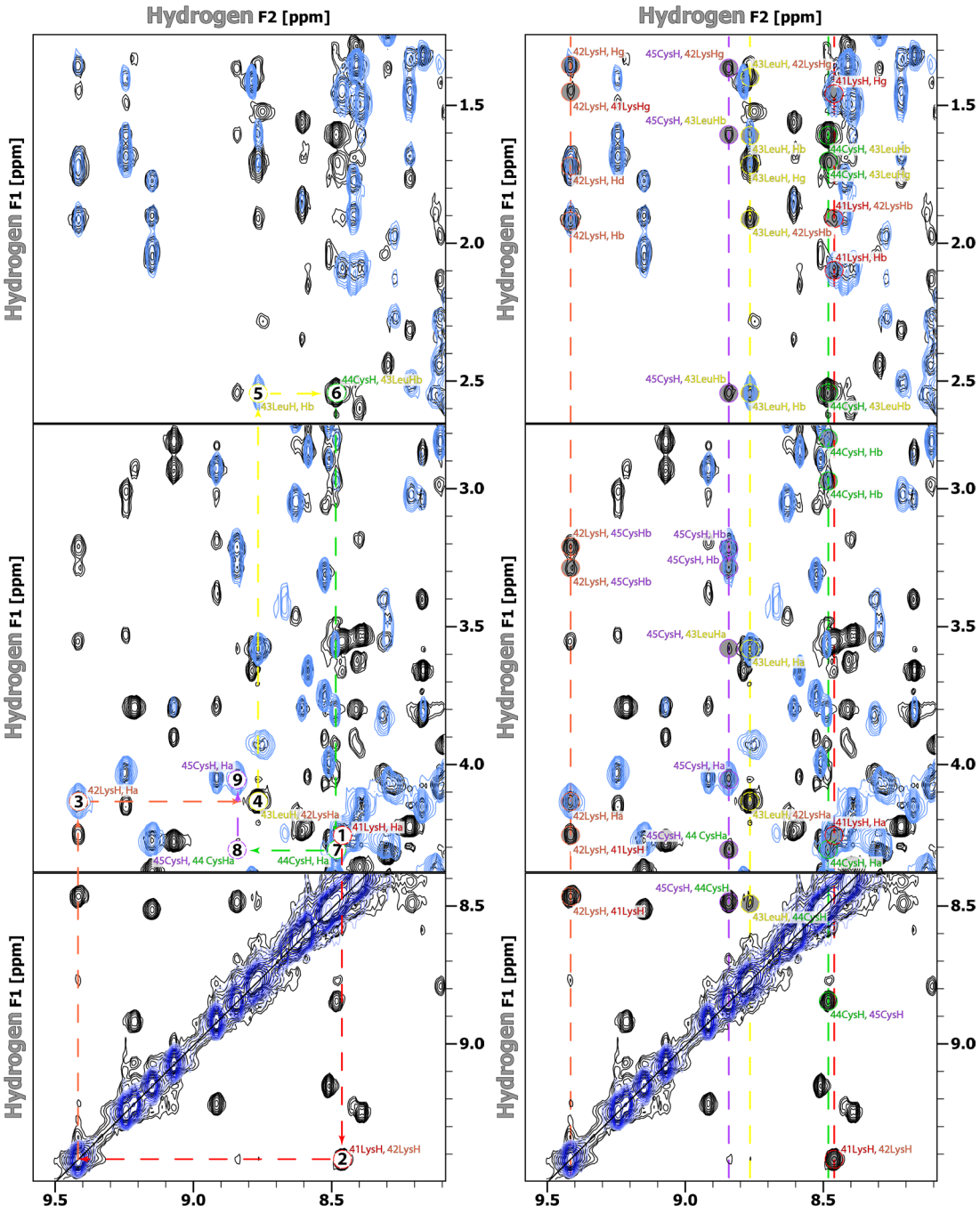


Figure 4.7. LEFT. Schematic of the backbone walk procedure, TOCSY spectrum in **BLUE**, NOESY spectrum in **BLACK**. The initial starting point (**1**) was a U-type H_N - Hα crosspeak (determined afterwards to be 41LysH, Hα) from the TOCSY spectrum. From interactions through-space, there is a NOESY (**2**) crosspeak between the spatially close backbone hydrogen from the proceeding residue. Following this NOESY peak, we reach another TOCSY crosspeak between the backbone nitrogen and the Hα from another U-type. Proceeding in this manner, the spin-types for a stretch of about 5 amino acids were elucidated. **RIGHT**. All NOEs indicated for the five residues, indicating long-range interactions between residues.

4.4 Generation of $^{13}C/^{15}N$ double-labelled cleaved Cd-wMT-2

As there was significant ambiguity and peak overlap in the unlabelled spectra, even when the peaks were spread over two dimensions, the next stage was to generate an isotopically labelled protein. This allows the separation of peaks over a greater number of dimensions, and the assignment of carbon and nitrogen atoms within the protein backbone. The strategy employed for generating isotopically labelled protein was to express in minimal medium supplemented with M9 salts. The expression protocol was similar to that of expression of unlabelled Cd-S●tag-wMT-2 in LB. The two main differences were: a higher ratio of initial overnight culture used (1 mL per 10 mL medium), larger than the 1 mL per 100 mL medium ratio required for expression in LB; the minimal medium cultures were grown to OD_{600} ≈ 0.8 before induction, 33 % higher than expression in LB. The route to purify the labelled protein was identical to that reported for Cd-S●tag-wMT-2 from LB cultures (see **Chapter 2.2**).

Both ^{15}N and ^{13}C isotopes were 99 % enriched, increasing the theoretical average molecular mass of cleaved Cd_7-wMT-2 from 8,570.8 Da to 8,955.0 Da[†]. From **Figure 4.8**, the experimentally determined neutral mass of the predominant species is 8,950.3 Da, confirming the presence of double-labelled cleaved Cd_7-wMT-2. The slight difference in mass (< 5 Da) may be accounted for by considering a lower than reported

[†] Calculated using the number of atoms reported by the ProtParam tool, http://web.expasy.org/protparam/

99 % isotopic fidelity of both ^{15}N and ^{13}C, with an average of 97.79 % giving a theoretical mass which matches that of the mass spectrum.

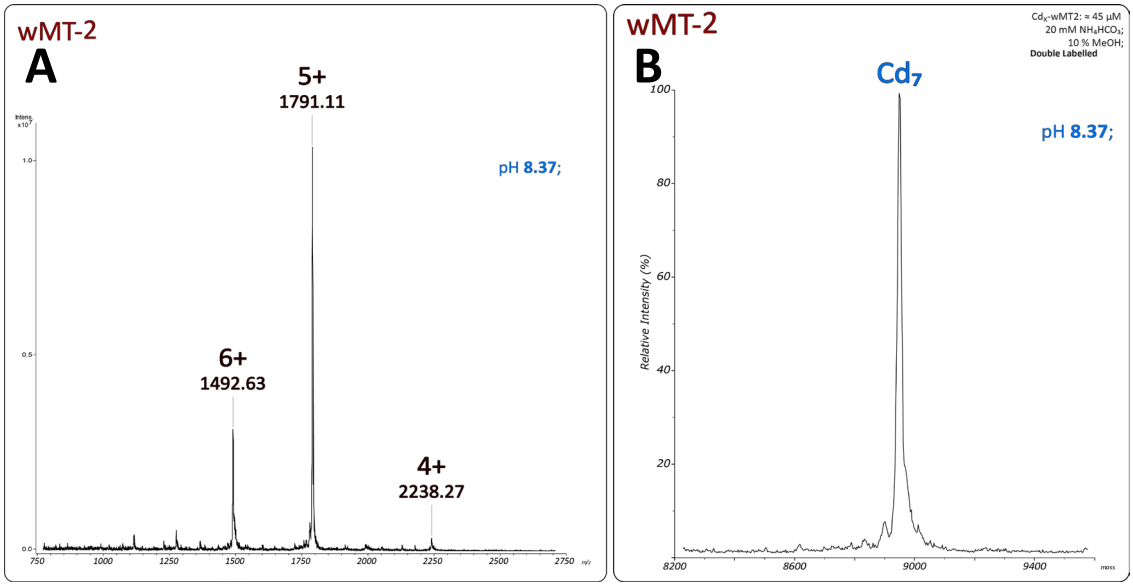


Figure 4.8A. Raw mass spectrum of double-labelled cleaved Cd_7-wMT-2. **B.** Deconvoluted mass spectrum of the charge states indicated in (**A**). wMT-2 concentration 44.7 µM, pH 8.37. Sample in 20 mM NH_4HCO_3, and 10 % MeOH. A sole Cd_7- (**BLUE**) metalloform is observed.

As the mass spectrum showed pure cleaved Cd_7-wMT-2, the resulting protein was concentrated for multinuclear NMR experiments. To maintain the maximum concentration of wMT-2, the buffer exchange and concentration steps were performed concurrently, with the final NMR buffer comprising 20 mM ammonium bicarbonate buffer (made with Milli-Q water), pH 6.8. The final concentration of the double-labelled cleaved wMT-2 sample used in the remaining sections in the chapter was determined by ICP-OES to be 542.4 µM.

4.5 Multinuclear NMR experiments for structure determination

[^{1}H, ^{15}N] HSQC

Combining the chemical shifts of backbone H_N atoms from the TOCSY/NOESY spectra, and the corresponding backbone N chemical shift from the HNCA/HN(CO)CA spectra, the next spectrum to be assigned was the [^{1}H, ^{15}N] HSQC. The [^{1}H, ^{15}N] HSQC experiment (**Figure 4.9**) correlates the chemical shifts of the nitrogen chemical shift (F1 dimension), with that of the chemical shifts of any attached hydrogen atoms (F2 dimension) in the form of a crosspeak in the 2-dimensional spectrum.

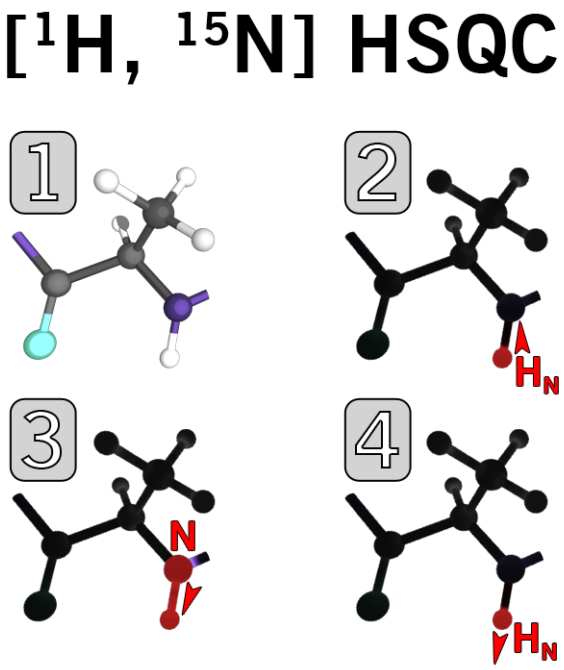


Figure 4.9. Schematic of the [^{1}H, ^{15}N] HSQC experiment. **RED** indicates the direction of magnetisation transfer. In Panes 2-4, magnetisation is transferred from the H_N for evolution on the backbone nitrogen, before being transferred back to the H_N for detection.

Although the main objective of the experiment was to assign the backbone nitrogen atoms, some nitrogen-containing sidechains were also present in the upper right quadrant of the spectrum. For wMT-2, four of six asparagine / glutamine residues showed peaks in this area (**Figure 4.10**). Most peaks have a well-defined and narrow peak shape.

Peaks are well dispersed, indicating a well-folded protein, with some clustering around the centre of the spectrum.

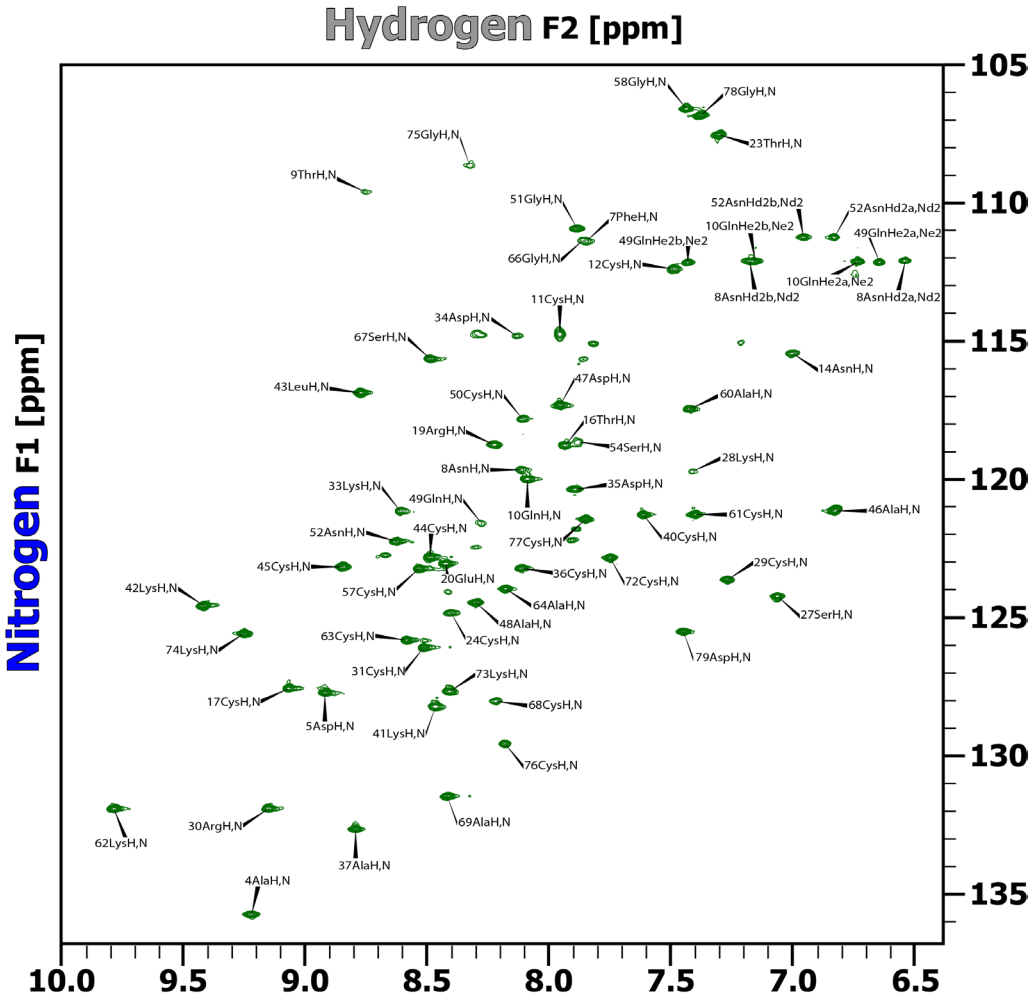


Figure 4.10. [^{1}H, ^{15}N] HSQC of double-labelled cleaved Cd-wMT-2. Buffer was 20mM ammonium bicarbonate, pH ≈ 6.8 with 10 % D_2O. NMR conditions: 700 MHz, 298 K. Approximately 80 % of the residues were able to be assigned, based on information from the NMR experiments reported in this section.

Triple-resonance experiments: HNCA, HN(CO)CA

Combined experimental data from both the HNCA and HN(CO)CA experiments were required for a C_α backbone walk. The HNCA experiment generates a spectrum with two peaks per residue, the stronger of the two usually corresponding to that of the intraresidual N-C_α transfer, the weaker to transfer to the preceding C_α-1 (i-1). The

HN(CO)CA experiment generates a spectrum of one peak from the preceding C_α-1 (i-1) residue, labelled with the hydrogen and nitrogen shifts of the i residue (**Figure 4.11**). When the two experiments are overlaid, the two overlapping peaks correspond to the C_α-1 of the i-1 residue, the lone peak corresponding to the C_α of the i residue. As the experiments separate peaks in 3-dimensions, the problem of overlapping peaks was significantly reduced. Therefore using the limited unambiguous information from the 2D TOCSY experiment, further neighbouring residues could be unambiguously assigned. The backbone nitrogen chemical shift for residues could also be tentatively assigned, based on the chemical shift of the 2-dimensional slice of the 3-dimensional spectrum in the nitrogen dimension containing the backbone C_α - H_N crosspeak. **Figure 4.12** shows a C_α backbone walk from an unambiguously assigned residue, 44Cys. The next residue was also unambiguously assigned from the TOCSY (45Cys), confirming that the peak for the C_α is at the expected carbon chemical shift with respect to the hydrogen dimension.

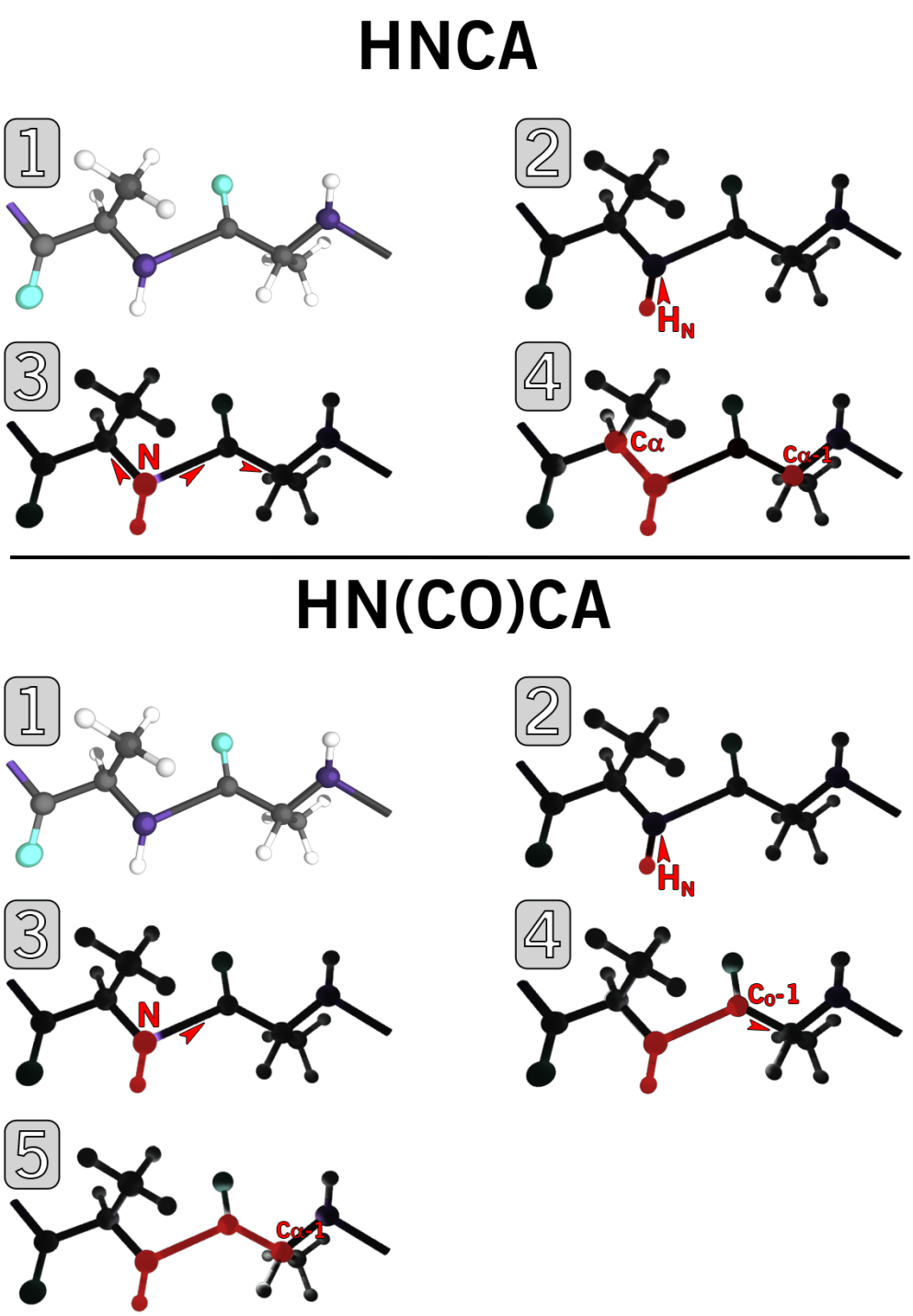


Figure 4.11. Schematic of the HNCA experiment (**TOP**) and HN(CO)CA experiment (**BOTTOM**). **RED** indicates the direction of magnetisation transfer. For the HNCA experiment, magnetisation is transferred from the H_N to the backbone N, before being passed to both the intraresidual C_α, and the C_α–1 of the i-1 residue. Magnetisation is then transferred back through the backbone N, to the H_N for detection (not shown). For HN(CO)CA, magnetisation is transferred from the H_N to the backbone nitrogen, before being passed to the i-1 C_α-1 through the backbone carbonyl group (C=O). Magnetisation is transferred identically to HNCA for detection (not shown).

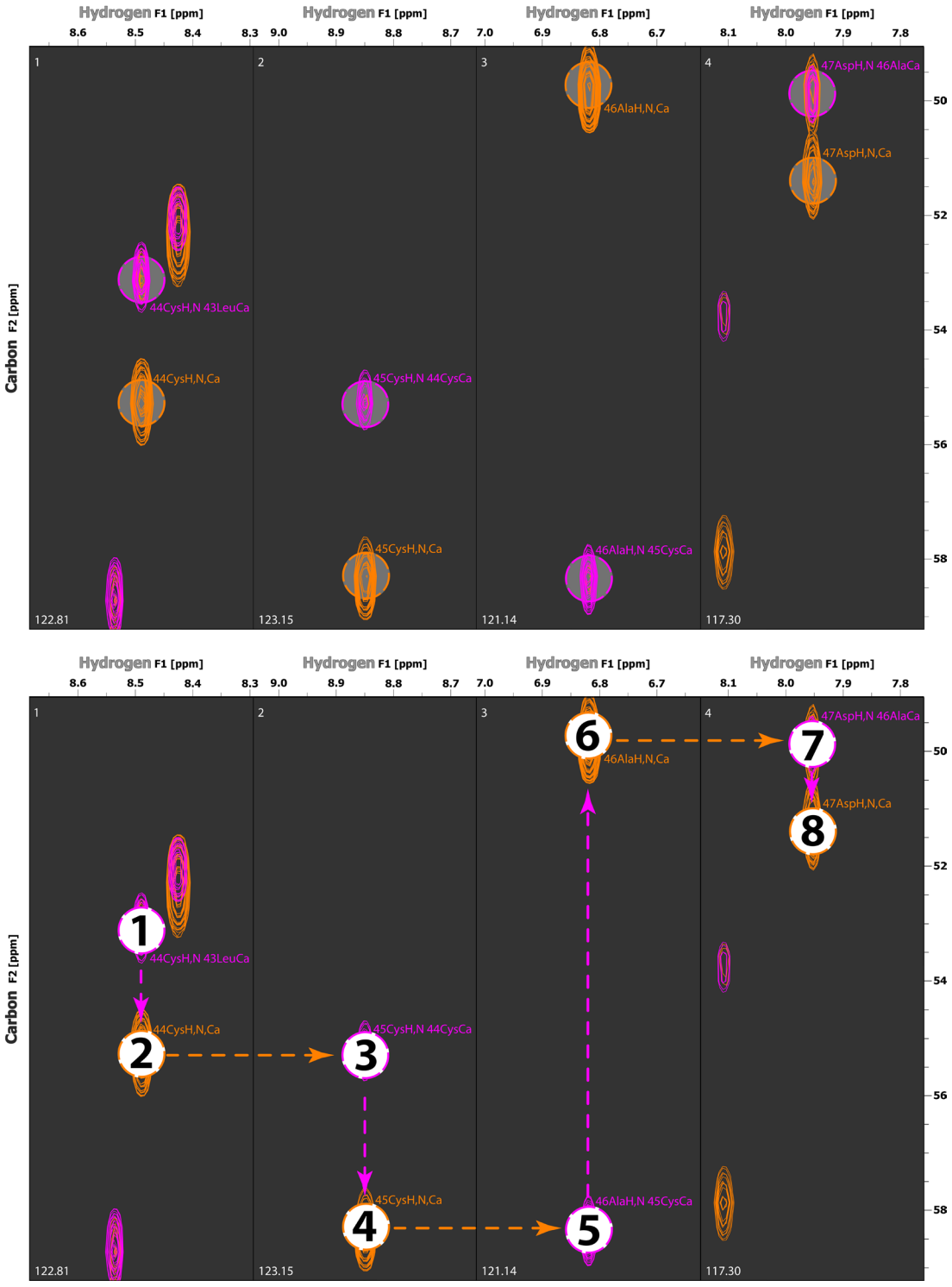


Figure 4.12. Schematic of the backbone walk procedure for HNCA (**ORANGE**) and HN(CO)CA (**PINK**) spectra. The initial starting point (**1**) is the backbone H shift of a known residue, here 44Cys (**1**). Combined with the assigned nitrogen shift (from ^{15}N-HSQC), there is one possibility. The HN(CO)CA indicates the preceding residue, moving from HN(CO)CA to HNCA means we are walking forwards along the backbone. Moving from the 44Cys HNCA (**2**), we reach a peak in the HN(CO)CA spectrum with the same carbon shift (**3**). This shift is at the same hydrogen shift as the Cα of the next residue (**4**). The backbone walk proceeds in this manner. Note that the second peak in the HNCA spectrum is underneath the peak in the HN(CO)CA spectrum.

[^{1}H, ^{1}H, ^{15}N] TOCSY-HSQC

The [^{1}H, ^{1}H, ^{15}N] TOCSY-HSQC experiment proceeds initially like a standard TOCSY experiment, ie magnetisation is transferred to all hydrogen atoms within an amino acid residue, however the magnetisation is then transferred onto the backbone nitrogen, and finally detected on the H_N. This adds a third (nitrogen) dimension which is invaluable for assigning overlapping spin systems in the 2-dimensional TOCSY experiment. This was the case for 67Ser, with three possibilities for assignment of the H-Hα crosspeak (**Figure 4.13**). The TOCSY-HSQC allowed unambiguous assignment of the H-Hα crosspeak, and aided in the assignment of the entire overlapping region. Although a lower resolution in each nitrogen slice of the 3-dimensional spectrum is acceptable due to the separation of peaks along a third axis, the low intensities of some peaks meant that many slices were not used in the sequential assignment. Sequence coverage in the TOCSY-HSQC was ≈ 24 %.

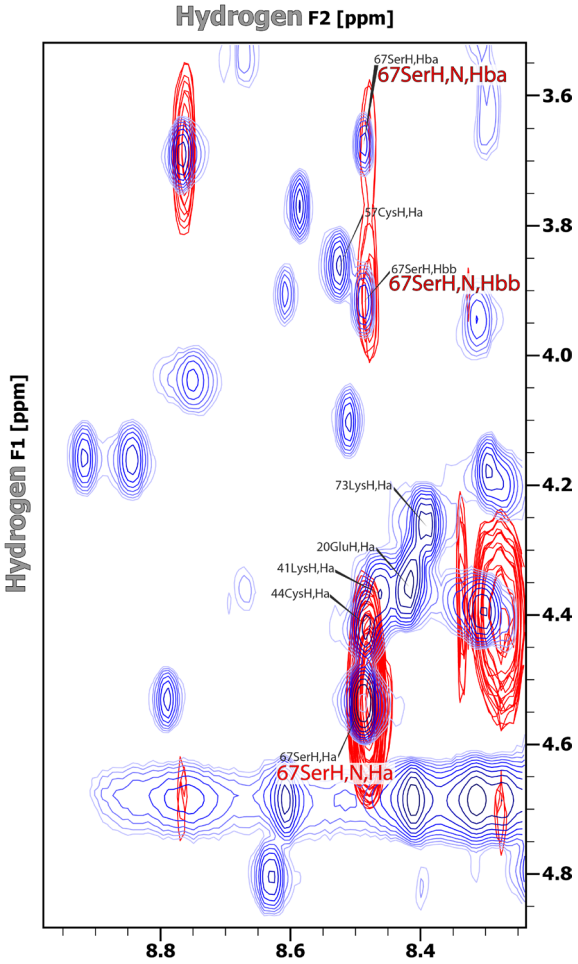


Figure 4.13. Experimental data from a TOCSY-HSQC experiment (**RED**), overlaid with the original TOCSY experimental data (**BLUE**). Although the resolution is poorer than the TOCSY due to the requirement of keeping the number of scans per slice low, it is sufficient to determine which crosspeak in the overlapping region of the spectrum corresponds to that of 67SerH, Hα. The Hβ crosspeaks can also be seen in the TOCSY-HSQC, overlapping with the original TOCSY peaks, for additional confirmation.

Summary of wMT-2 assignment

Combining information from all of the experiments in this section, it was possible to unambiguously assign 63 of the 79 residues in the cleaved wMT-2 sequence. This included within the unambiguous assignments four residues (24Cys, 39Arg, 40Cys, 72Cys) which presented solely in the NOESY spectrum, although absent from the TOCSY spectrum. In addition, partial information (some hydrogen atoms not present) was available for 5 residues, allowing their assignment as: 6Ala, 18Pro, 32Pro, 38Pro, 56Ser. Considering the two surplus amino acids at the N-terminus remaining from cleavage of the S●tag, there was 88 % coverage of the wMT-2 sequence in the TOCSY/NOESY spectra. In the

HN(CO)CA/HNCA spectra, 72 residues had their Cα atoms assigned, giving wMT-2 sequence coverage of 94 %. In the HSQC experiments, 66 residues had their nitrogen atoms assigned, giving 86 % wMT-2 sequence coverage. The chemical shifts of the assigned atoms within the residues of wMT-2 are tabulated in **Appendix 17**. From the assigned spectrum, distance restraints were generated for structure calculations. Additionally some backbone torsional angles were determined experimentally using the HNHA experiment.

HNHA

The HNHA experiment (**Figure 4.14**) can be used to experimentally determine phi (ϕ) protein backbone dihedral angles (C_O-**N-Cα**-C_O), using the Karplus relationship [222]. From an assigned spectrum, the relative intensities of the H_N and H-H_α crosspeak heights (or volumes) are required for the equation. Due to the pulse sequence used for the experiment, crosspeaks for the H-H_α chemical shifts appear as positive peaks, with the backbone H_N appearing as negative peaks. After manual assignment, values for torsional angles were calculated using the CcpNmr Analysis ^{3}J H-Hα Coupling module. This output was subsequently exported as an angle file to be used in CYANA calculations. Based on the assigned spectrum, angles were able to be calculated for 57 % of the protein sequence due to peaks being either missing or unable to be unambiguously assigned.

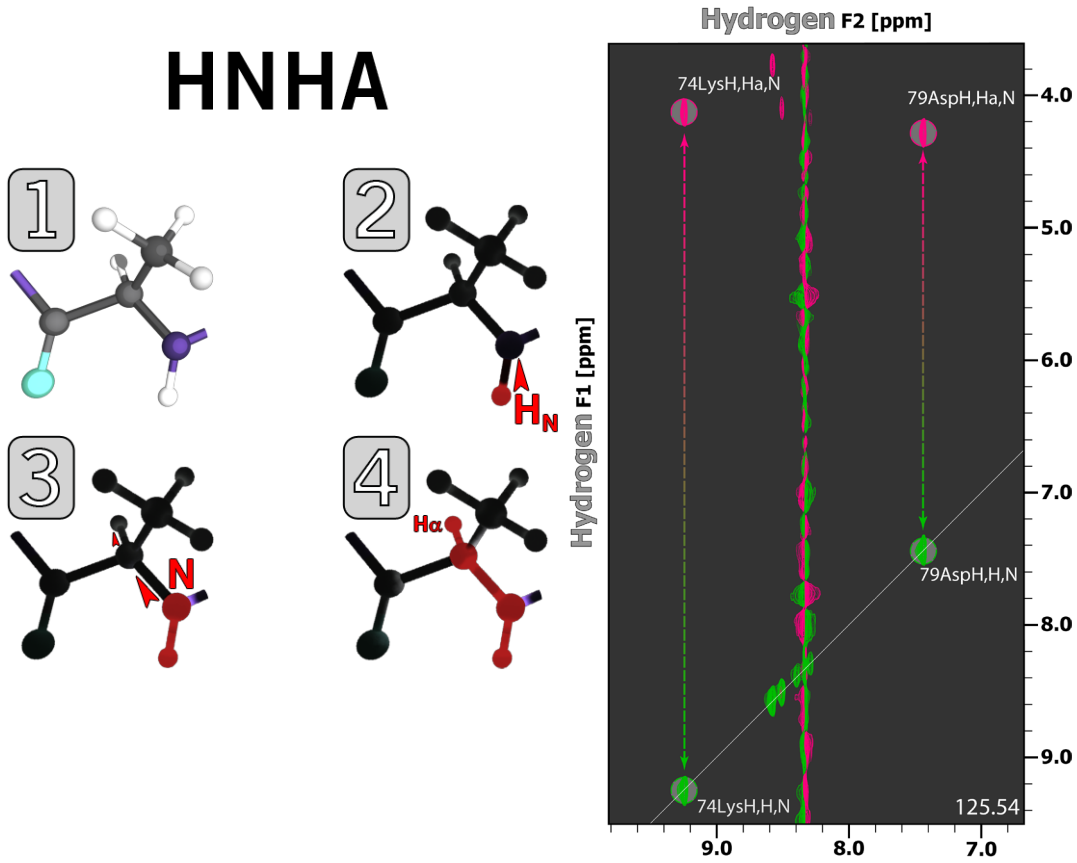


Figure 4.14. LEFT.Schematic of the HNHA experiment. **RED** indicates the direction of magnetisation transfer. Magnetisation is transferred from the H_N to the backbone N, before being passed through the intraresidual Cα to the Hα atom. Magnetisation is then transferred back the same way for detection (not shown). **RIGHT.** HNHA experimental data is presented of two isolated residues (74Lys and 79Asp). Peaks in the positive spectrum (**PINK**) correspond to that of the backbone H, Hα crosspeak. This can be compared to the negative peak (**GREEN**) on the diagonal (corresponding to the backbone H) using the Karplus relationship. This will generate the dihedral backbone angles for the residue.

4.6 ^{111}Cd 1-dimensional NMR Spectroscopy

To determine cluster assignments for specific cadmium-cysteine connectivities, attempts to prepare samples reconstituted with the ^{111}Cd isotope were undertaken [93]. The specific isotope used was 96.4 % ^{111}Cd, obtained as solid CdO [Cambridge Isotope Laboratories Inc., UK], and dissolved in HCl. Due to the high cost of isotopically enriched metals, wMT-2 was expressed first in the presence of unlabelled cadmium, and then reconstituted with labelled cadmium just before spectroscopy. Initial

attempts at this procedure caused precipitation of the sample, possibly due to protein not refolding correctly around the newly supplied metal, or the apo-protein chains aggregating at low pH. A subsequent purification met with similar losses in protein, the best spectrum being obtained from a sample of approximately 100 µM cleaved ^{111}Cd-wMT-2, with a stoichiometry of 6.8 cadmium ions per wMT-2 molecule. The low concentration meant that a large number of scans (96k) were required for the experiment, and only three of the seven possible environments emerged significantly from background noise (**Figure 4.15**).

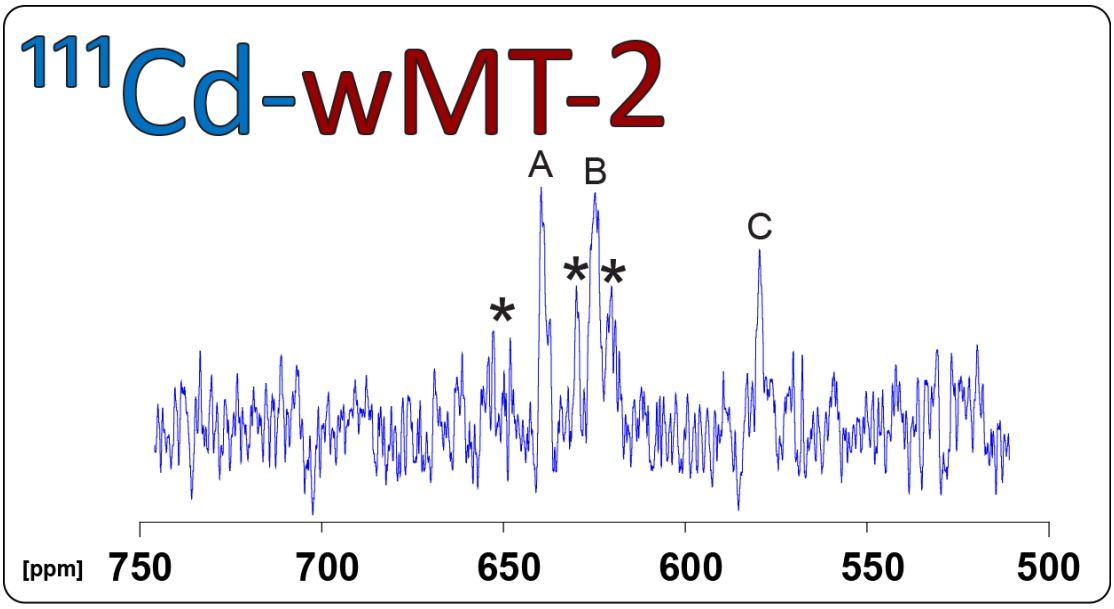


Figure 4.15. ^{111}Cd 1-dimensional NMR of reconstituted ^{111}Cd-wMT-2. wMT-2 concentration approximately 100 µM. Sample buffer 50 mM [D_{11}]Tris-Cl, pH 7.5. Spectrum acquired at 298 K, 96k scans. There are 3 distinct peaks are labelled A-C, with potential additional peaks indicated with '*'.

Work performed by Hemmingsen *et al*. [223] indicated a correlation in the type and number of ligands, and the observed chemical shift in $^{111/113}$Cd NMR. The S_4Cd tetrahedral coordination environment can theoretically present anywhere from 622-791 ppm [223], however, experimentally determined chemical shifts for other MTs indicate shifts in the 610-672 ppm range being most likely [223]. Peaks A and B are well

within this range, with three additional potential peaks designated '*' (**Figure 4.15**). Peak C, however appears at an unexpectedly low chemical shift (580 ppm), which may suggest that the metal ion is either not coordinated by 4 sulphur ligands, or experiences shielding usually observed with nitrogen ligands. This could be a consequence of the harsh reconstitution conditions significantly affecting the ability of wMT-2 to reform the correct coordination environment. To enable full characterisation of the binding sites, further ^{111}Cd experiments will be required.

4.7 Structure Calculations using CYANA

The basis for structure calculations is the retrieval of restraints from NOE experiments. Restraints are values imposed on the calculation based on experimental data, for example the magnitude of an NOE crosspeak generates an approximate distance (distance restraint) between the two hydrogen atoms. The magnitude of the assigned NOE peaks gives information on the approximate through-space distances of the respective atoms. The structure determination of two-domain MTs by NMR usually requires the separate treatment of the two domains, as the data tend to offer no information on the mutual orientation of the domains. This is due to the rigid domains being joined by a flexible linker region. Flexible regions in NMR are problematic for structural determination, as they often have few or weak NOEs due to peak broadening from multiple conformers. This method [87] was followed for wMT-2 and separate calculations were performed for each domain. In the

following experiments, 200 structures were calculated, and the 10 with the lowest target function used in backbone root-mean-square deviation (RMSD) calculations. The target function is defined as being zero when no restraints are violated by the calculation; target functions < 10 are generally acceptable.

Linker Region

There are two potential candidates for linker regions in wMT-2. Although the linker region thus far has been postulated to consist of residues 46-49 (-ADAQ-), a -KK- motif (found in vertebrate MT linker regions [224, 225]) is also present in the wMT-2 sequence. To further add to the linker region conundrum, the two candidates either divide the wMT-2 sequence into Cys_{11} / Cys_9 clusters (not uncommon for invertebrate MTs), or Cys_9 / Cys_{11} clusters (similar to vertebrate MTs) (**Figure 4.16**).

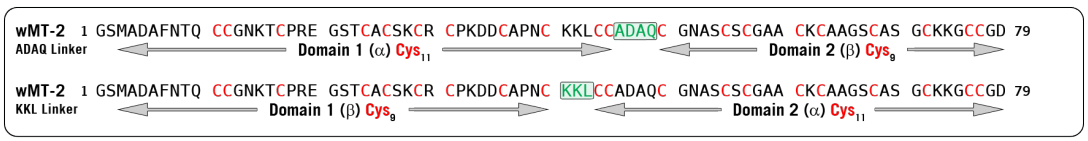


Figure 4.16. Sequences of cleaved wMT-2 with the two proposed linker regions highlighted in **GREEN**. Cysteines coloured **RED**, with the domains shown below the respective sequence.

In an attempt to experimentally ascertain the correct domain boundary, two sets of calculations were performed. Both were of Domain 2 without Cys-Cys or Cys-Cd restraints. The two calculations were performed for: residues 50-79 (Cys_9 cluster, **Figure 4.16 TOP**); or residues 44-79 (Cys_{11} cluster, **Figure 4.16 BOTTOM**). The first calculation generated an RMSD result of 2.56 ± 0.53 Å, whereas the second calculation (including

-ADAQ-) gave an RMSD of 5.97 ± 1.49Å. As a consequence of MT linker regions being flexible and dynamic, the residues involved are likely to be lacking in NOE peaks due to peak broadening, and therefore distance restraints. As inclusion of the –ADAQ- portion caused a significantly higher RMSD value in the second structure calculation, it will be assumed that this corresponds to the correct linker region. In addition, some long range NOE crosspeaks corresponding to the lysine doublet (41-42) with residues 37Ala and 31Cys are present. These crosspeaks would be unlikely to be present if the lysine residues were part of a linker region. This reinforces the hypothesis that the linker region consists of residues 46-49.

Deriving metal-cysteine connectivities

In the absence of unambiguous labelled cadmium data, the connectivities between specific cysteine sulphur and cadmium atoms had to be approximated by other means. Due to the lack of deposited MT structures, a sequence-specific search of the individual domain sequences against those with Protein Data Bank (PDB) coordinate information was performed. Usually the measure for determining a 'real' hit from a 'chance' hit is the E value of the match. The E value is the 'Expect value', the higher this value, the more likely a hit is due to chance matches of part of the sequence. Generally speaking, an E value between matched sequences > 10 is purely due to chance, with an E value > 1 being unlikely to be related. When working with short sequences however, the E values are sometimes misleading and the

'Total Score' value is more effective for determining 'real' hits. The 'Total Score' reflects the homology in all portions of the sequence, even if they are non-contiguous (ie there are gaps in the sequence), with higher values indicating a better match. The search for Domain 1 (CCGNKTCPREGSTCACSKCRCPKDDCAPNCKKLCC) generated no significant hits (E value > 5). Performing the same search with Domain 2 however (CGNASCSCGAACKCAAGSCASGCKKGCC) generated 2 potentially significant hits (E value < 1), with the β-domain of sea urchin MTA covering 96 % of the sequence. Relaxing the E value parameters to accept potential 'chance' matches generated 4 additional hits with E value < 1.5, all from the vertebrate MT family (**Figure 4.17**). This is probably due to the disproportionately large number of vertebrate MT sequences which have had their three-dimensional structures determined.

		Total Score	E value
wMT-2 (Earthworm)	-CGNASCSCGA---ACKCAAGSCASGCKKGCCGD--	N/A	N/A
MTA (Sea urchin)	XCTNAACKCAN---GCKCGSG-CS--CTEGNCAC--	86.3	0.043
MT-2 (Human)	--XDPNCSCAA-GDSCTCAGS-CK--CKECKCTSCK	67.7	1.5
MT-2 (Rat)	--XDPNCSCAT-DGSCSCAGS-CK--CKQCKCTSCK	47.9	1.1
E_c (Wheat)	CGCGEHCGCNP----CACG|-|CS--CGAACNCASC	38.8	7.00E-5
MT-2 (Rabbit)	--XDPNCSCAAAGDSCTCANS-CT--CKACKCTSCK	26.5	1.1
MT-1 (Mouse)	--MDPNCSCST-GGSCTCTSS-CA--CKNCKCTSCK	26.1	1.5

Figure 4.17. Results from a PDB search of wMT-2 Domain 2. Note that all domains are β-domain. The Horizontal lines in the Wheat E_c sequence indicate a section of 14 residues which have been removed for the alignment only.

Unexpectedly, there was a hit from the β_E-domain of the plant MT, Wheat E_c. Contrary to previous experimentally determined metal connectivities, a modelling approach was taken for the β_E-domain of Wheat E_c, with the connectivity with the lowest RMSD published as the expected connectivity [226]. This required a calculation for every permutation of every cysteine

in the cluster, for in excess of 7,000 calculations. This was not deemed a suitable method for deriving the connectivities for wMT-2.

To determine the possible structure of wMT-2 Domain 2, two approaches were taken. The first was to transpose the connectivities from the similar β-domain of sea urchin MTA to the sequence of wMT-2 [87]. The approach is somewhat similar to that of homology modelling, however only the metal-cysteine connectivities are required, and no other structural inferences are made between the two protein sequences. This is possible due to the highly conserved positions of the cysteine residues in this area of both proteins (**Appendix 18**). The second approach was to refine a structure from the starting point of a structure determined without metal restraints. Distances measured between the sulphur atoms of different cysteine residues in a cluster were sufficient for determining likely cluster formations caused by four thiols being in close spatial proximity.

Domain 2 (Residues 50-77)

Structure calculations included all cysteine residues for the proposed second domain, leaving the flexible termini out of the calculation. **Figure 4.18** shows the restraints generated for Domain 2 by CCPN from a starting 98 NOESY peaks, including both upper and lower distance limits.

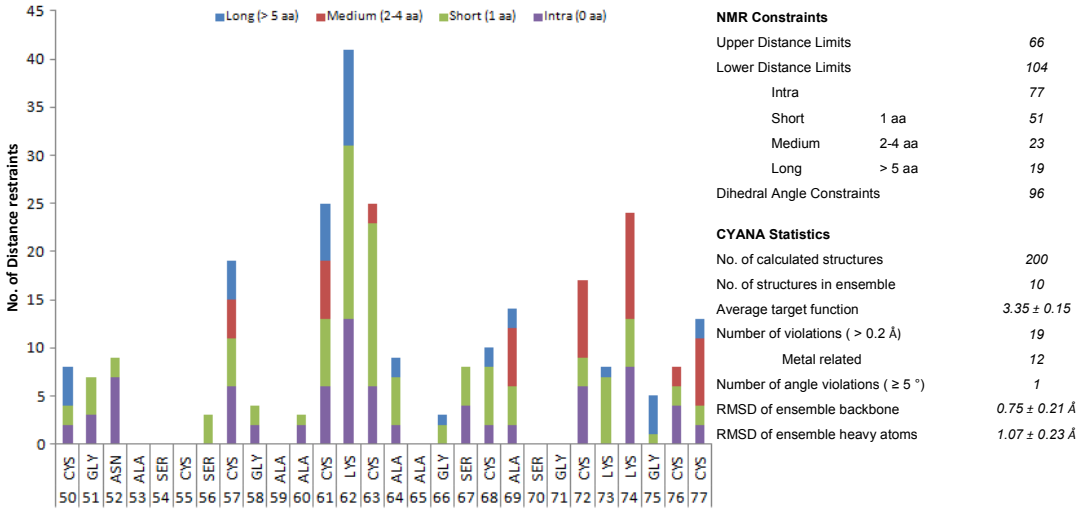


Figure 4.18. LEFT Plot of the number of upper and lower distance limits per residue in wMT-2 Domain 2 (note that metal restraints are not included). Intra-residual NOEs shown in **PURPLE**, Short-range (sequential) NOEs shown in **GREEN**, Medium-range (2-4 residues) NOEs shown in **RED**, Long-range (> 5 residues) NOEs shown in **BLUE**. **RIGHT** Summary of the NMR constraints, and statistics obtained from CYANA for Domain 2 of wMT-2.

Domain 2 is not well defined without metal restraints, and gave an RMSD of 2.56 ± 0.53 Å. (**Figure 4.19**). This lack of definition is reflected in the 'sausage' plot, with thick lines indicating that within the 20 structures within the ensemble, multiple conformations were present. This is in keeping with secondary structure analysis using the PDBsum Web Interface, which indicates no secondary structure elements are present.

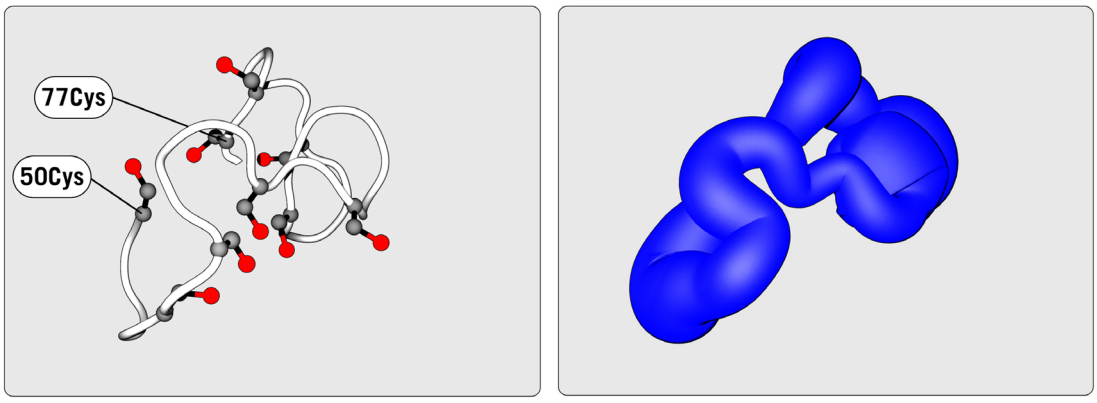


Figure 4.19. LEFT. Representative structure with lowest target function (T = 9.50E-04) of wMT-2 Domain 2 without metal restraints. **RIGHT.** The ensemble of 20 structures calculated by CYANA, indicating the flexibility of the structures. Thinner lines represent lower variability of the structure.

To improve the quality of the structure, an additional 51 metal restraints were included in the next calculation: the upper and lower distance limits were fixed to 2.60 Å between S-Cd - defined as the value for non-strained Cd-S bonds [227], Cd-Cd distances between metals in the cluster were fixed to between 4.50-4.70 Å, and cysteine S-S distances in the cluster were fixed to 4.05-9.00 Å[‡]. Transposing the relative positions of the cysteine residues in the MTA β-domain to wMT-2 Domain 2 generated a structure with 18 violated distance restraints (≥ 0.2 Å), with a sum of the mean of these violations of 7.13 ± 0.13 Å. There was also 1 violated angle restraint (≥ 5 °). The structure determined using these metal constraints had a backbone RMSD of 0.85 ± 0.17 Å, which indicates good resolution. The 'sausage' plot showed that the backbone was significantly higher resolution than without metal restraints (**Figure 4.20**), whilst still resembling the original shape.

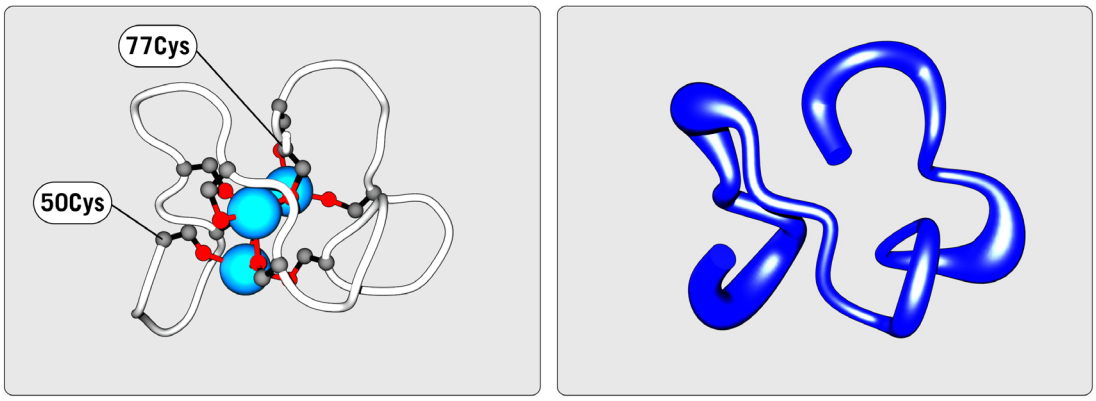


Figure 4.20. LEFT. Representative structure with lowest target function (T = 3.65) of wMT-2 Domain 2, utilising the connectivities from sea urchin MTA. **RIGHT.** The ensemble of 20 structures calculated by CYANA, indicating the flexibility of the structures. Thinner lines represent lower variability of the structure.

Visualising the structure with the lowest target function in SWISS PDB Viewer enabled other permutations of the metal cluster to be

[‡] These numbers were derived from a survey of known MT PDB structures, 1dft, 1qjl, 2kak, 2mhu, 2mrb, 4mt2

investigated. The investigation was performed on the PDB structure with the lowest target function (the target function is zero when all restraints are met), calculated in the absence of metal restraints. From this structure, the distances between the sulphur atoms of each cysteine were measured, and tabulated (**Figure 4.21**).

DOMAIN 2 - S-S distances

	50	55	57	61	63	68	72	76	77
50		8.88	9.29	8.31	10.47	15.04	13.03	7.92	5.98
55	8.88		3.20	4.89	10.26	6.43	15.17	13.34	10.78
57	9.29	3.20		3.92	8.31	6.15	15.41	11.89	11.10
61	8.31	4.89	3.92		7.01	2.96	11.53	9.87	7.99
63	10.47	10.26	8.31	7.01		6.08	14.02	13.70	13.51
68	15.04	6.43	6.15	2.96	6.08		10.23	11.64	8.62
72	13.03	15.17	15.41	11.53	14.02	10.23		12.07	7.10
76	7.92	13.34	11.89	9.87	13.70	11.64	12.07		2.46
77	5.98	10.78	11.10	7.99	13.51	8.62	7.10	2.46	

Figure 4.21. Cysteine S-S distances (in Å) of wMT-2 Domain 2, measured from the structure with lowest target function from the ensemble, values lower than the mean (9.40 Å) are indicated in **BLUE**.

Although it was thought that residues with > 3 short interresidual distances[§] (ie 61Cys has another 6 cysteine residues in close proximity) were more likely to be bridging cysteine residues, this was not necessarily the case. A total of ten permutations were attempted, based on possible connectivity patterns from their inter-residual S-S distances, the results are tabulated in **Table 4.1**.

[§] short distances are distances less than the mean distance of 9.40 Å for Domain 2

Table 4.1. The outcome of performing structure calculations based on ten possible cluster compositions for Domain 2. Results are ranked by RMSD of backbone atoms. **BLUE** indicates bridging cysteine residues.

	Cd^{I}				Cd^{II}				Cd^{III}				Target	RMSD ± Å		Violated constraints	
																distance	angle
Alternate Connectivity 1	50	55	57	76	57	61	63	77	68	72	76	77	3.35	0.75	± 0.21	19	1
Sea Urchin Connectivity	50	55	63	68	55	72	76	77	57	61	63	77	4.19	0.85	± 0.17	25	3
Alternate Connectivity 2	50	55	57	72	55	61	76	77	57	61	63	68	5.57	1.06	± 0.32	27	4
Alternate Connectivity 3	50	61	76	77	55	57	61	68	63	68	72	77	7.82	1.41	± 0.33	23	3
Alternate Connectivity 4	50	55	57	63	50	72	76	77	57	61	68	77	6.02	1.41	± 0.33	31	4
Alternate Connectivity 5	50	61	76	77	55	57	61	68	63	68	72	77	7.61	1.43	± 0.33	24	3
Alternate Connectivity 6	50	61	63	77	55	57	61	68	68	72	76	77	6.64	1.43	± 0.93	27	2
Alternate Connectivity 7	50	61	63	68	50	55	57	77	68	72	76	77	4.88	2.70	± 1.01	24	5
Alternate Connectivity 8	50	55	63	77	55	57	61	68	61	72	76	77	8.81	2.74	± 0.66	39	5
Alternate Connectivity 9	50	55	57	63	57	72	76	77	61	63	68	72	8.85	3.37	± 0.97	35	8

The permutation with the lowest RMSD score (Alternate Connectivity 1, **Table 4.1**) was used for further work in this section. From analysis of the domain composition, both connectivity options look very similar, especially in the coordination of Cd^{II} and Cd^{III}, although the nature (either bridging or terminal) of the cysteine residues was different. The most significant difference was in the Cd^{I} coordination cluster, at the start of the sequence (**Figure 4.22**).

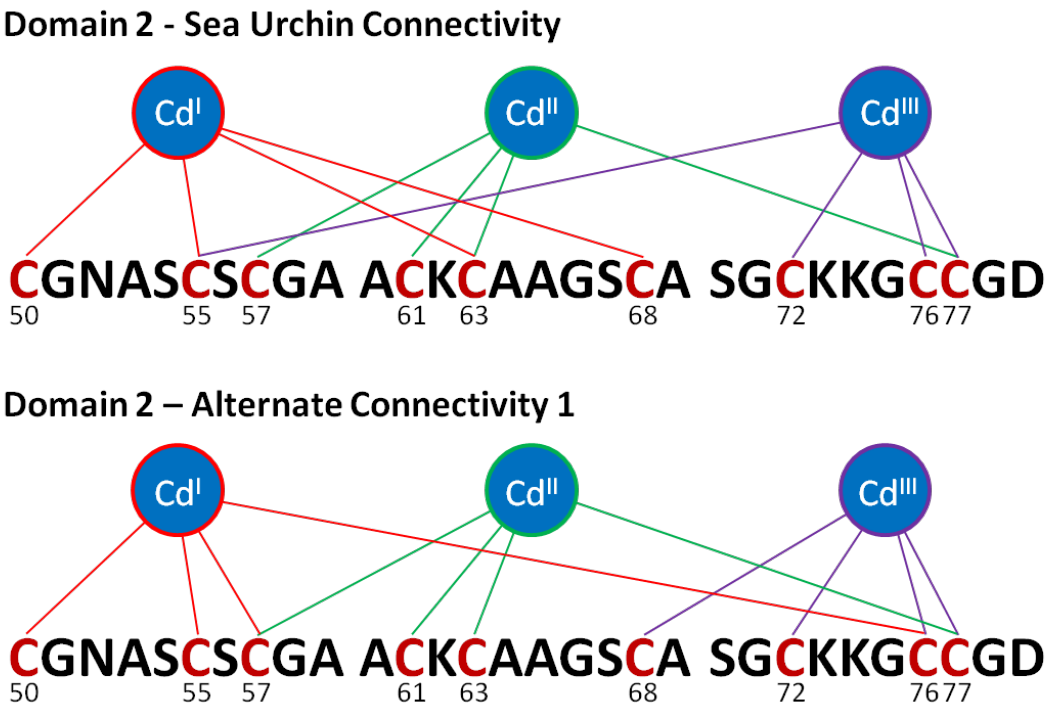


Figure 4.22. Comparison of the wMT-2 Domain 2 connectivities obtained by transposition of the β-domain of sea urchin MTA (**TOP**), and the permutation with lowest RMSD from PDB analysis (**BOTTOM**).

The 'Alternate 1' domain connectivity generated a structure with an RMSD of 0.75 ± 0.21 Å (**Figure 4.23**), approximately 10 % smaller than with 'Sea Urchin' domain connectivity. The sausage plot in **Figure 4.23** shows a slightly different backbone conformation, although the general motif of a large clockwise turn moving into a set of three anti-clockwise turns (as oriented in **Figure 4.19** & **Figure 4.20**) was conserved.

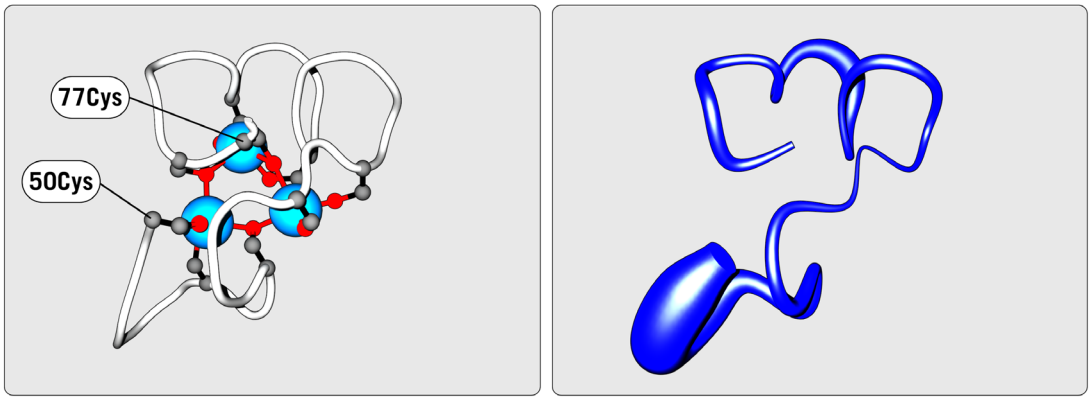


Figure 4.23. LEFT. Representative structure with lowest target function (T = 3.11) of wMT-2 Domain 2 utilising 'Alternate 1' metal restraints. **RIGHT.** The ensemble of 20 structures calculated by CYANA, indicating the flexibility of the structures. Thinner lines represent lower variability of the structure.

The changes in structure are due in part to the scarcity of long-range connectivities, making the metal ions and their connectivities the largest contributor to the structure of Domain 2. Within the structure, 19 distance restraints were mildly violated (0.2 - 0.7 Å), more than half of these were from the inclusion of perhaps overly restrictive metal restraints. With that in mind, the magnitude of the violations was similar to literature [144], with no violation exceeding a mean of 0.67 Å, with the sum of all mean violations 6.14 ± 0.01 Å. To estimate whether the values for violations are acceptable, they were compared to the values reported for sea urchin MTA (sum of all mean violations of 1.8 ± 0.4 Å).

Analysis of the protein ensemble of 20 structures with lowest target function was performed using the Common Interface for NMR structure Generation (CING) Web Interface[**] for analysing PDB files. The CING interface outputs general information on both the quality of the structure, and how well the structure conforms to current structures within the database (**Table 4.2**).

Table 4.2. Summary of the output from CING for the two connectivities, compared with that of sea urchin MTA β-domain. Structure Z-scores generally give an indication of the quality of the structure, based on currently available structures, the more positive the score the better. The RMS Z-scores generally give an indication of how well the model conforms to current structural data, the closer to 1.0 the score the better.

Structure Z-scores	**wMT-2 Sea Urchin**	**wMT-2 Alternate 1**	**MTA**
Packing quality	*-3.644 ± 0.691*	*-3.566 ± 0.858*	*-6.192*
Ramachadran plot	*-7.802 ± 0.275*	*-7.719 ± 0.572*	*-5.581*
Backbone conformation	*-2.459 ± 1.083*	*-1.834 ± 0.381*	*-5.744*
RMS Z-scores			
Bond lengths	*0.948 ± 0.001*	*0.948 ± 0.001*	*1.065*
Bond angles	*0.147 ± 0.000*	*0.147 ± 0.000*	*0.688*
Improper dihedral distribution	*0.154 ± 0.002*	*0.154 ± 0.002*	*0.601*

For gauging the quality of the data, the two connectivity options were compared with the β-domain Cys_9-cluster of MTA. Packing scores are an evaluation of the interaction between buried amino acids and interactions between exposed amino acids. Generally, proteins pack tightly, with the expected distance (or packing) between a particular amino acid and its neighbours in a sequence being experimentally determined [226]. Therefore low packing quality scores are probably due to the metal-binding clusters of MT causing unique conformations of the protein

[**] https://nmr.cmbi.ru.nl/icing/iCing.html

backbone. The very low Ramachadran score may be of some concern, with 1.9 % (sea urchin MTA β-domain connectivity) or 3.8 % ('Alternate 1' connectivity) of residues exhibiting angles in the disallowed region (**Figure 4.24**). Glycine residues have been shown for reference, but as they may adopt any backbone torsional angle, they are not of concern. Although the structure of the sea urchin MTA β-domain has no angles in disallowed areas, the presence of angles in disallowed regions is not uncommon for MTs. For example, the accepted structure of rat MT-II β-domain (PDB 1mrt) has 3.8 % of residues with angles in disallowed areas.

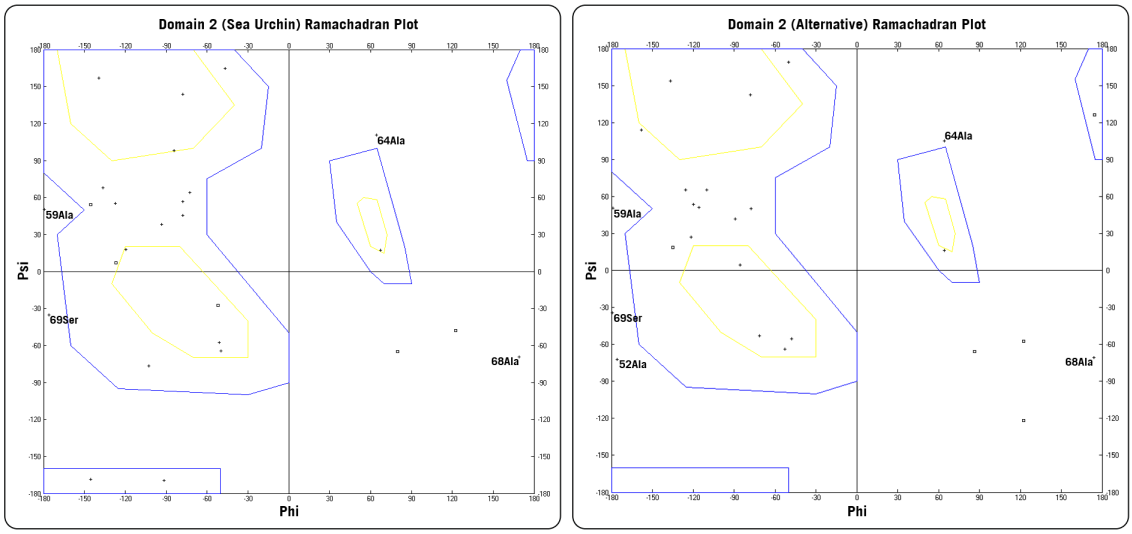


Figure 4.24. Ramachadran plot of the torsional angles in Domain 2 of wMT-2 with either sea urchin MTA β-domain or 'Alternate 1' connectivities. Crosses within **YELLOW** regions are favourable regions, crosses within **BLUE** regions are generously allowed. Residues that are not glycine and are outside of these regions are labelled, and show abnormal torsional angles.

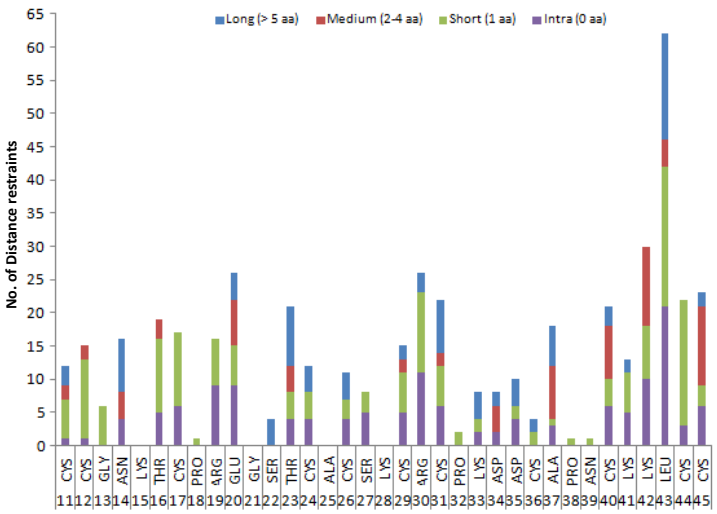
The backbone conformation score looks at the backbone formed by a residue, and its direct neighbours either side. Based on the conformation of these triplets, a score is derived, with low scores indicating that residues are part of an unusual loop, or there are issues with the angles formed by those particular residues. Comparing the RMS Z-scores from

the actual sea urchin MTA β-domain with those of the two wMT-2 connectivities suggests that both options for Domain 2 are conformationally less unusual than the β domain of MTA. This may indicate that the angles imposed by CcpNmr Analysis and CYANA on the wMT-2 calculations are more tightly restrained than in the older program DYANA, which was used for calculating the MTA domain structures. This led to slightly less variation in bond angles than would be expected with respect to the protein backbone composition for wMT-2 Domain 2. Considering the ambiguous nature of the assignment of coordinating residues within wMT-2 Domain 2, the exact metal connectivity cannot be determined without heteronuclear [^{1}H, $^{111/3}$Cd]-correlation spectroscopy [87].

The packing quality and Ramachadran scores are better for the 'Alternate 1' connectivity than any other tested, and again comparable to those reported in accepted MT PDB structures. Although the percentage of angles in disallowed regions is slightly worse for the 'Alternate 1' connectivity than for 'Sea Urchin MTA β-domain' connectivity, the sum of the mean of all violated distance restraints is significantly lower (≈ 15 %). Considering that the general shape of the molecule without metal restraints matched fairly closely with the refined structure utilising the connectivity with lowest backbone RMSD ('Alternate 1'), it is likely that this connectivity is close to the real structure.

Domain 1 (Residues 11-45)

The determination of the structure of Domain 1 proceeded in a similar manner to that of Domain 2. Structure calculations included all cysteine residues for the proposed first domain, leaving the flexible termini out of the calculation. **Figure 4.25** shows the restraints available for Domain 1, generated by CCPN from a starting 161 NOESY peaks, including both upper and distance limits.



NMR Constraints		
Upper Distance Limits		111
Lower Distance Limits		192
Intra		136
Short	1 aa	85
Medium	2-4 aa	37
Long	> 5 aa	45
Dihedral Angle Constraints		128

CYANA Statistics	
No. of calculated structures	200
No. of structures in ensemble	10
Average target function	8.68 ± 0.16
Number of violations (> 0.2 Å)	31
Metal related	16
Number of angle violations (≥ 5 °)	3
RMSD of ensemble backbone	0.52 ± 0.12 Å
RMSD of ensemble heavy atoms	1.13 ± 0.24 Å

Figure 4.25. LEFT Plot of the number of relevant distance restraints per residue in Domain 1 of wMT-2 (note that metal restraints are not included). Intra-residual NOEs shown in **PURPLE**, Short-range (sequential) NOEs shown in **GREEN**, Medium-range (2-4 residues) NOEs shown in **RED**, Long-range (> 5 residues) NOEs shown in **BLUE**. **RIGHT** Summary of the NMR constraints, and statistics obtained from CYANA for Domain 1 of wMT-2.

Without metal restraints, Domain 1 had a high RMSD of 4.58 ± 0.46 Å (**Figure 4.26**), significantly higher than for Domain 2 without metal restraints (**Figure 4.19**). As investigation of similar structures was not possible (there were no hits in the database), and the primary sequence of the α-domain of sea urchin MTA is significantly different to that of wMT-2 Domain 1 (**Appendix 18**), the method of deducing connectivity permutations by measuring S-S distances between cysteine residues in the cluster was attempted (**Figure 4.27**).

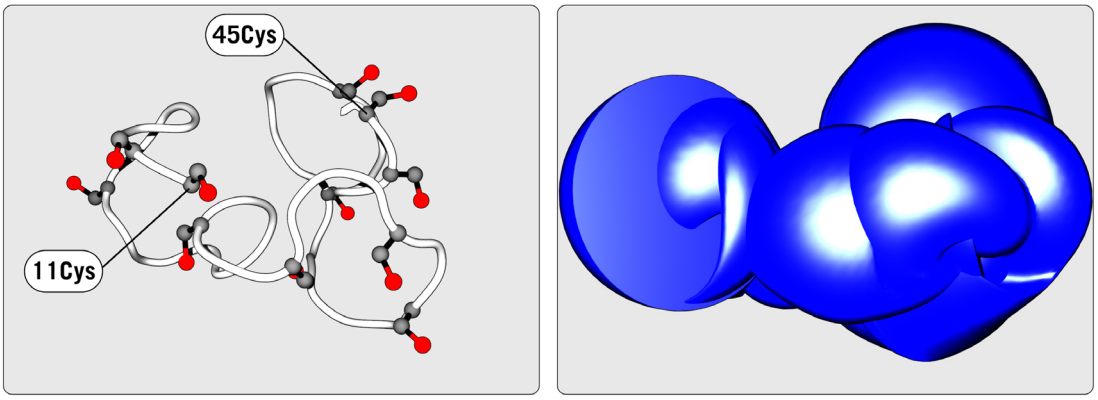


Figure 4.26. LEFT. Representative structure with lowest target function (T = 0.22) of wMT-2 Domain 1, without metal ion restraints. **RIGHT.** The ensemble of 20 structures calculated by CYANA, indicating the flexibility of the structures. Thinner lines represent lower variability of the structure.

DOMAIN 1 - S-S distances

	11	12	17	24	26	29	31	36	40	44	45
11		5.28	17.38	8.17	8.82	11.33	15.21	19.94	15.61	14.32	14.50
12	5.28		14.92	9.40	12.87	16.50	20.14	22.47	18.29	18.77	18.10
17	17.38	14.92		12.00	17.59	24.09	25.80	18.72	20.89	23.12	23.14
24	8.17	9.40	12.00		6.54	12.82	14.85	15.36	16.39	14.46	16.52
26	8.82	12.87	17.59	6.54		6.69	8.48	13.07	13.52	8.70	12.42
29	11.33	16.50	24.09	12.82	6.69		4.65	16.60	14.74	7.43	11.87
31	15.21	20.14	25.80	14.85	8.48	4.65		16.20	17.31	8.88	14.75
36	19.94	22.47	18.72	15.36	13.07	16.60	16.20		12.00	11.29	13.73
40	15.61	18.29	20.89	16.39	13.52	14.74	17.31	12.00		8.73	4.07
44	14.32	18.77	23.12	14.46	8.70	7.43	8.88	11.29	8.73		6.40
45	14.50	18.10	23.14	16.52	12.42	11.87	14.75	13.73	4.07	6.40	

Figure 4.27. Cysteine S-S distances (in Å) of wMT-2 Domain 1, measured from the structure with lowest target function from the ensemble, values lower than the mean (14.11 Å) are indicated in **BLUE**.

Similarly to Domain 2, the resolution of the structure was improved with the addition of 77 metal restraints, but due to the number of proposed bridging cysteine residues in Domain 1, the rules regarding the connectivities were much stricter. Therefore a total of five permutations were attempted, the results are tabulated in **Table 4.3**.

Table 4.3. The outcome of performing the structure calculations based on five possible cluster compositions for wMT-2 Domain 1. Results are ranked by RMSD of backbone atoms. **BLUE** indicates bridging cysteine residues.

	Cd^{IV}				Cd^{V}				Cd^{VI}				Cd^{VII}				Target	RMSD ± Å	Violated constraints	
																			distance	angle
Alternate Connectivity 1	11	12	24	44	11	17	36	40	24	26	31	36	29	31	40	45	8.68	0.52 ± 0.12	44	5
Alternate Connectivity 2	11	17	24	31	26	29	31	36	11	36	40	44	12	24	44	45	16.96	0.70 ± 0.29	61	4
Alternate Connectivity 3	11	12	24	26	11	29	45	17	26	29	31	44	36	40	44	45	18.99	1.10 ± 0.54	59	10
Alternate Connectivity 4	11	12	24	26	17	24	29	36	26	29	31	44	36	40	44	45	20.77	1.33 ± 0.27	69	6
Alternate Connectivity 5	12	17	24	31	26	29	31	36	17	36	40	44	11	24	44	45	23.38	1.96 ± 0.55	65	8

The permutation with the lowest target function, 'Alternate 1', showed a backbone RMSD value of 0.52 ± 0.12 Å (**Figure 4.28**). This structure causes the violation of 31 distance restraints ≥ 0.2 Å (44 total distance violations), with a sum of the mean of these violations 12.34 ± 0.16 Å; there were also 3 violated angle restraints ≥ 5 ° (5 total angle violations). Due to the number of violated restraints, the average target function was higher than for Domain 2, at T = 8.68.

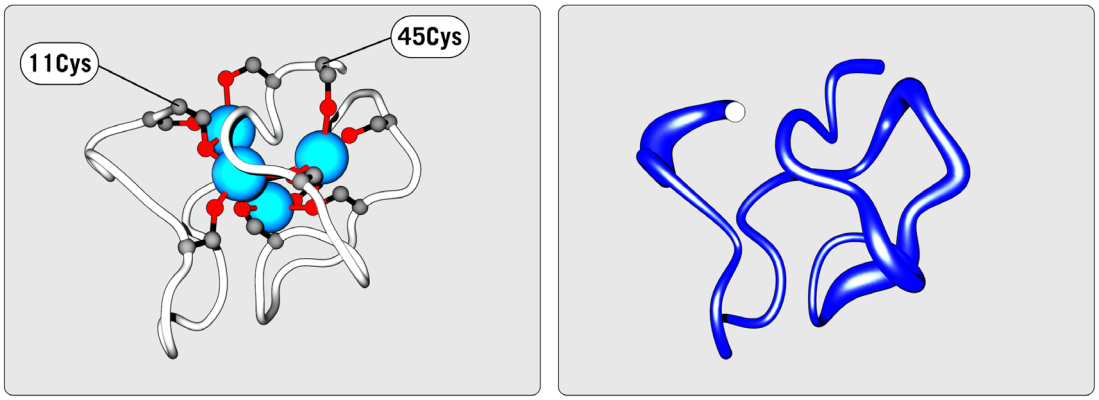


Figure 4.28. LEFT. Representative structure with lowest target function (T = 8.22) of wMT-2 Domain 1, with 'Alternative 1' metal restaints. **RIGHT.** The ensemble of 20 structures calculated by CYANA, indicating the definition of the structures. Thinner lines represent lower variability of the structure.

Although the primary sequences of MTA α-domain and wMT-2 Domain 1 are divergent, some insight into the quality of the structure can still be gathered by comparing their respective values. The Z-scores in **Table 4.4** are similar to those for the accepted structure of MTA α-domain from sea urchin, suggesting that the calculation has generated a good potential structure. In a similar fashion to Domain 2, Domain 1 exhibits a slightly lower Ramachadran score.

Table 4.4. Summary of the output from CING for the two connectivities, compared with that of sea urchin MTA α-domain. Structure Z-scores generally give an indication of the quality of the structure, based on currently available structures, the more positive the score the better. The RMS Z-scores generally give an indication of how well the model conforms to the current restraints, the closer to 1.0 the score the better.

Structure Z-scores	**wMT-2 Alternate 1**	**MTA**
Packing quality	*-4.019 ± 0.837*	*-6.041*
Ramachadran plot	*-7.534 ± 0.487*	*-7.202*
Backbone conformation	*-3.189 ± 0.394*	*-3.559*
RMS Z-scores		
Bond lengths	*0.966 ± 0.000*	*1.11*
Bond angles	*0.191 ± 0.000*	*0.686*
Improper dihedral distribution	*0.248 ± 0.002*	*0.642*

The Ramachandran plot (**Figure 4.29**) indicated 9.3 % of residues with backbone torsional angles in disallowed regions. Although this is a very high percentage, it is similar to that of the α-domain of rat MT-II (PDB 2mrt) which has 8.0 % of residues with backbone angles in disallowed regions, but higher than sea urchin MTA α-domain, which has 3.4 % in disallowed regions.

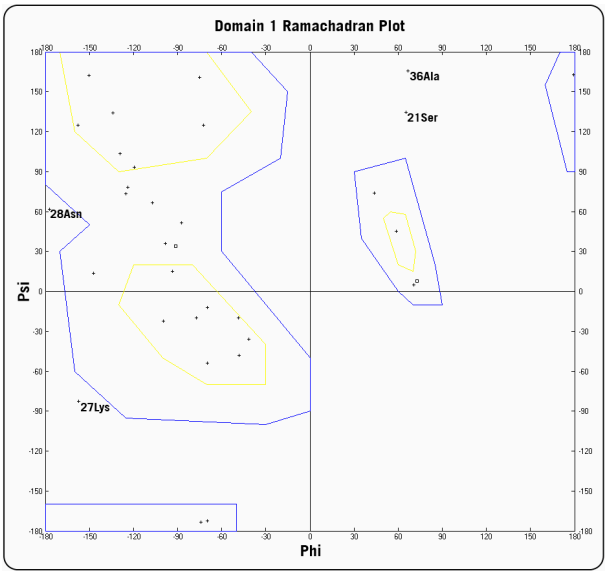


Figure 4.29. Ramachadran plot of the torsional angles in Domain 1 of wMT-2. Crosses within YELLOW regions are favourable regions, crosses within BLUE regions are generously allowed. Residues that are not glycine and are outside of these regions are labelled, and show abnormal torsional angles.

In general, the RMS Z-scores for Domain 1 (**Table 4.4**) are slightly better than for Domain 2 (**Table 4.2**), with the low variation in bond lengths and bond angles indicative of tight restrictions within CYANA. As the overall target function is very much higher than for Domain 2, with a similar increase in the number of distance and angle violations, more refinement of the Domain 1 structure may be required. However, the structure generated thus far exhibits low variability within the ensemble, as shown by the thin lines in the 'sausage' plot in **Figure 4.28**. With this in mind, and without the aid of experimentally determined metal-cysteine connectivities, this structure is acceptable.

4.8 Summary

The proposed structures for Domain 1 and Domain 2 of wMT-2 were experimentally determined through the use of 2-dimensional and 3-dimensional NMR techniques. When an acceptable 1-dimensional spectrum is recorded, 2D [^{1}H, ^{111}Cd] HSQC can be performed, and with knowledge of the ^{1}H chemical shifts assigned previously, the through-bond couplings between the hydrogen atoms of the coordinating residue and the specific cadmium ions [93] can be used for unambiguous cluster assignment. Without the aid of experimentally determined cadmium-cysteine connectivities, a sequence similarity search was performed for each domain, and distances between every cysteine in a domain were measured from a structure without metal restraints. Based on the inter-residual cysteine distances (55 for Domain 1, 36 for Domain 2), likely permutations of cysteine residues about the metal ions were then attempted. Domain 2 was found to be similar in sequence to sea urchin MTA (≈ 50 %). When the domain connectivities were transposed, the backbone RMSD was calculated to fulfil all criteria for an acceptable MT structure. However, based on a permutation generated from inter-residual cysteine distances a structure was calculated with lower backbone RMSD. This was accepted in preference, as likely to be representative of the correct backbone conformation. The search with respect to Domain 1 yielded no similar sequences, but cysteine-cysteine connectivities generated a structure which also showed all the characteristics of an acceptable MT structure.

The next step towards a structure/function relationship for wMT-2 was to investigate the *in vitro* metal-binding properties. Where possible, these properties are compared to the other members of the earthworm MT family, wMT-1 and wMT-3.

5

Metal-binding studies of wMTs

5.1 Introduction

In order to gather information about the metal-binding properties of earthworm MTs, a series of experiments to ascertain their responses to zinc and cadmium ions were undertaken. As with the majority of the work presented in this thesis, the focus was on wMT-2. However, due to the successful purification of wMT-1 and wMT-3, some general comparisons could be made between all three wMTs. Mass spectrometry measurements were performed on a Bruker HCTultra ion-trap mass spectrometer. Sample conditions were 20 mM ammonium bicarbonate (pH indicated in the text), with 10 % HPLC grade MeOH. NMR measurements were performed on either a Bruker AV II 700 spectrometer or a Bruker AV-400 fitted with a quadruple nucleus probe (QNP) able to measure ^{19}F nuclei.

5.2 Competition with protons

Like virtually all metal-binding proteins, MTs exhibit pH dependence within their metal-binding clusters [229]. Some pH dependent behaviour of Cd-S●tag-wMT-2 had emerged in a 2006 study performed by Stillman *et al*. [157]. This study observed only the Cd_7-, Cd_4- and apo- forms of S●tag-wMT-2 during acidification at three pH values (6.80, 3.46, 2.30).

An investigation with more pH data points is required to ascertain whether other species might be present at intermediate pH values, and if observations with S●tag-wMT-2 were applicable to cleaved wMT-2.

Behaviour of wMT-2

A pH titration followed by mass spectrometry was undertaken for both the Zn- (13 readings) and Cd- (21 readings) forms of cleaved wMT-2 (**Figure 5.1**). The general trends for the two metalloforms were found to be significantly different. For example at pH 4.0 cleaved Zn-wMT-2 showed species from Zn_7- to apo-, whereas only fully metallated Cd_7- was observed as the major species. A Zn_8- species became apparent around neutral pH, indicating lability within the zinc clusters. This indicates that wMT-2 can either form divalent metal adducts, or the backbone can adopt a conformation that will accept an extra zinc ion (possibly due to the ionic radius of the Zinc(II) ion being approximately 80 % smaller than Cadmium(II)). The alternate backbone conformation of Zn-wMT-2 may influence the lability of zinc within the clusters, and therefore allow the observation of multiple Zn-wMT-2 species. The zinc spectra at pH 4.2-3.8 suggest that four of the remaining zinc ions may be bound within the proposed M_4Cys_{11}- cluster, as acidification to pH 3.8 caused a significant change in the species observed.

If we consider that cadmium is harder to remove than zinc from MT, we might expect the behaviour of Zn-wMT-2 to be shown by Cd-wMT-2 but

at a lower pH. This was not the case, as Cd_4- became the major species (at pH 3.5), followed by the emergence of the apo- form (at pH 2.5). There were very few intermediate species observed, even though

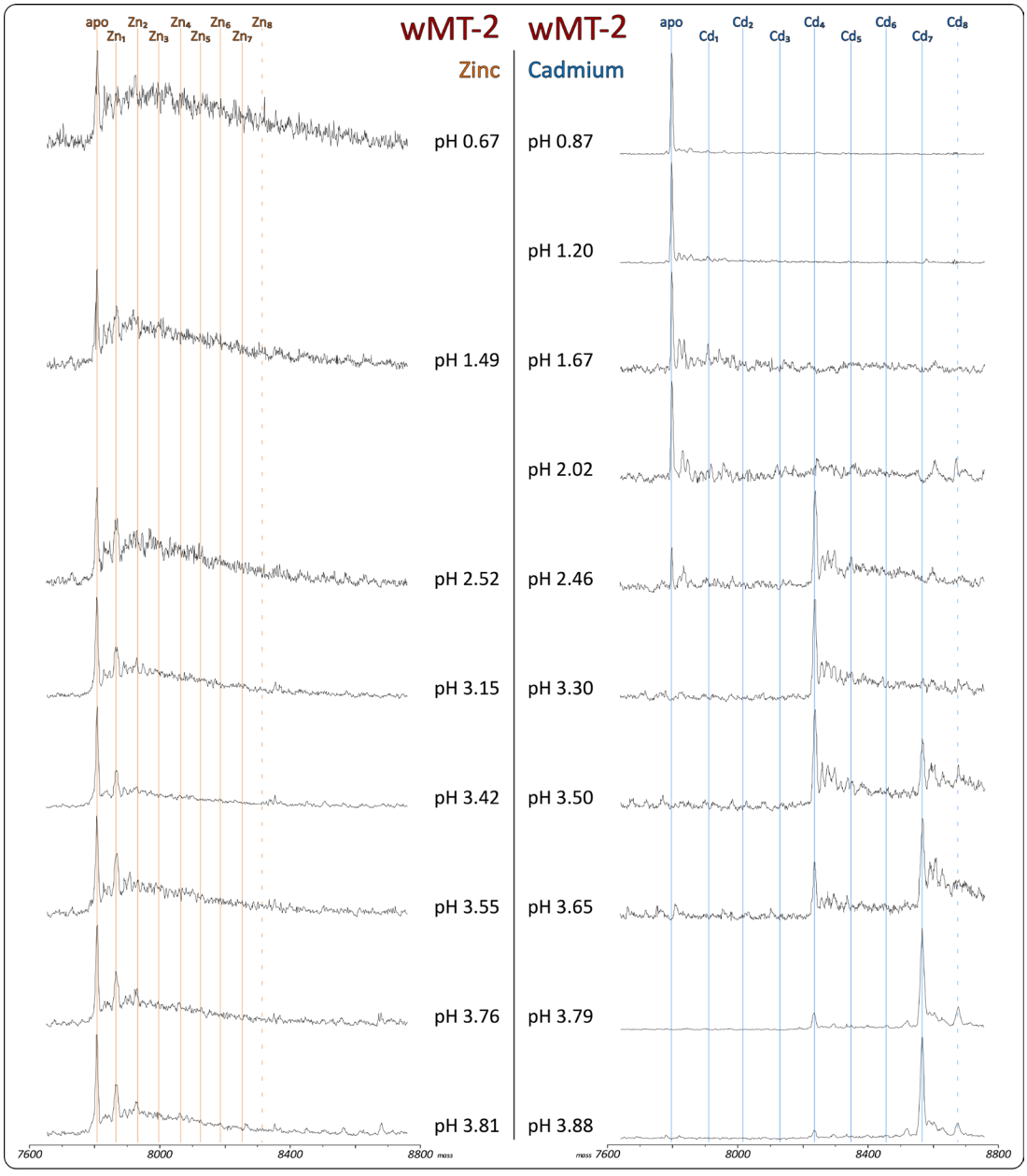


Figure 5.1. Stacked plot of the deconvoluted MS of cleaved wMT-2 with either zinc (**LEFT**) or cadmium (**RIGHT**). Similar pH conditions are juxtaposed to allow direct comparison of the spectra. The calculated mass of the metallated species are indicated with lines through the spectra, with supermetalated species indicated with dashed lines. Cd-wMT-2, 6.7 cadmium ions per wMT-2 molecule (48 µM protein); Zn-wMT-2, 6.1 zinc ions per wMT-2 molecule (42 µM protein) in 20 mM ammonium bicarbonate, 10 % MeOH. Further analysis in **Appendix 14**. **FIGURE CONTINUED OVER THE PAGE**.

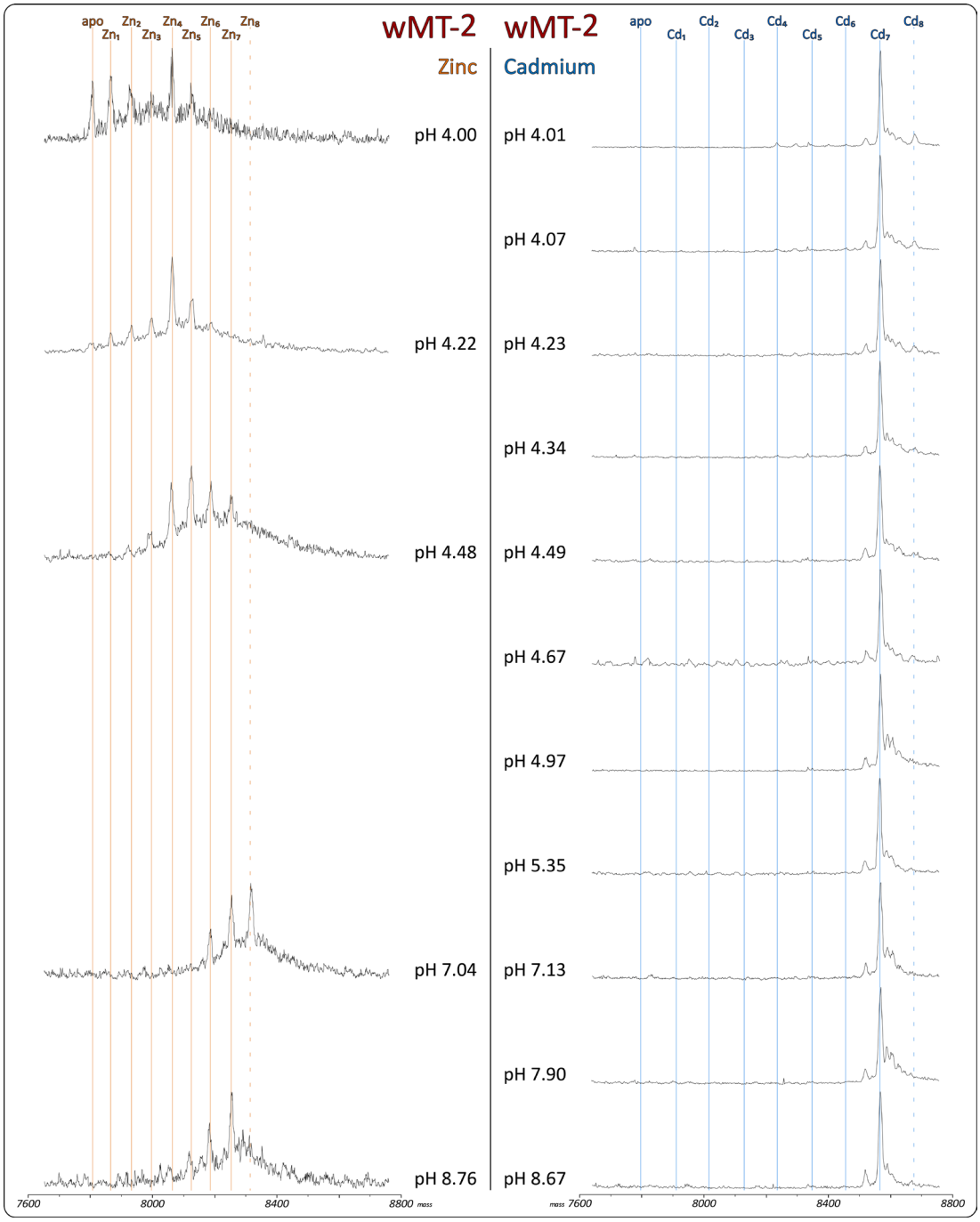


Figure 5.1 (continued). Stacked plot of the deconvoluted MS of cleaved wMT-2 with either zinc (**LEFT**) or cadmium (**RIGHT**). Further analysis in **Appendix 14**.

pH steps of < 0.5 units were used in this study. Essentially, only the Cd_7-, Cd_4-, and apo- wMT-2 forms were observed. This supports the hypothesis that, at least with respect to H^+ induced metal loss, Cd^{2+} ions are cooperatively bound in Cd-wMT-2. The three major species (Cd_7-, Cd_4- and apo-) show a 3 metal ion loss, followed by a 4-metal ion loss.

These may correspond to complete losses from the likely C-terminal M_3Cys_9- cluster, followed by the N-terminal M_4Cys_{11}-cluster. This is similar to previous studies with metal *addition* to vertebrate MT [217]. Even though the orientation of the clusters in earthworm MTs (M_4Cys_{11}- / M_3Cys_9-) are reversed when compared to vertebrate MTs (M_3Cys_9- / M_4Cys_{11}-) the orientation does not seem to effect the cooperative behaviour of the MT. Experiments with hybrid M_3Cys_9-/ M_3Cys_9- MTs [230] and reversal of the entire vertebrate MT sequence [231] also suggest that metal-binding is independent of domain orientation.

A large reduction in pH (below pH 4.0) was required for initial cadmium ion removal from wMT-2. This is significantly different to the behaviour of Zn-wMT-2, which shows undermetallation above neutral pH. Due to the presence of multiple species present in the zinc spectrum (ie. Zn_7-, Zn_6-, Zn_5-), there is a lower initial stoichiometry of Zn-wMT-2 compared to Cd-wMT-2. Without further experiments, it cannot be determined whether the relative instability of Zn-wMT-2 is magnified by the presence of species which are not fully metallated, causing non-chelating cysteine residues to be more susceptible to protonation. Stability to cadmium removal at low pH could be crucial for the proposed storage role of wMT-2. This would support the hypothesis that Cd-wMT-2 is secreted into 'cadmosomes', cellular compartments similar to lysosomes [232]. Lysosomes have an internal acidity of ≈ pH 5 [233], and if this is similar for 'cadmosomes', then stability at this pH may be crucial for ameliorating the toxic effects of cadmium ions within the earthworm.

Above pH 5.0, there were no discernible changes in the spectra for Cd-wMT-2. However, a small proportion of a supermetalated species (Cd_8-) appeared concurrently with the emergence of a small Cd_4- peak. It is difficult to ascertain whether these species would exist *in vivo*, or if they are a consequence of the imposed *in vitro* experimental conditions. As metals are not removed from the solution during the pH titration, it could be that the appearance of Cd_8- and Cd_4- at the same time indicates that some of the Cd^{2+} liberated from the Cd_3Cys_9- cluster may form an adduct with Cd_7-wMT-2. As this behaviour is seen for both zinc and cadmium forms, the propensity for supermetalation may be a specific advantage for a protein whose role is to quickly chelate large influxes of toxic metal ions. There have been certain cases of known supermetalation, either of individual clusters [206, 207], or of entire MTs [234], but this has not been previously seen in wMTs.

Behaviour of wMT-1 and wMT-3

Due to time constraints and low protein yields, mass spectra for wMT-1 (Zn and Cd forms) and wMT-3 (Cd form only) were recorded at fewer pH intervals than wMT-2 (**Figure 5.2**). However, initial trends did start to emerge, even with the comparatively limited data.

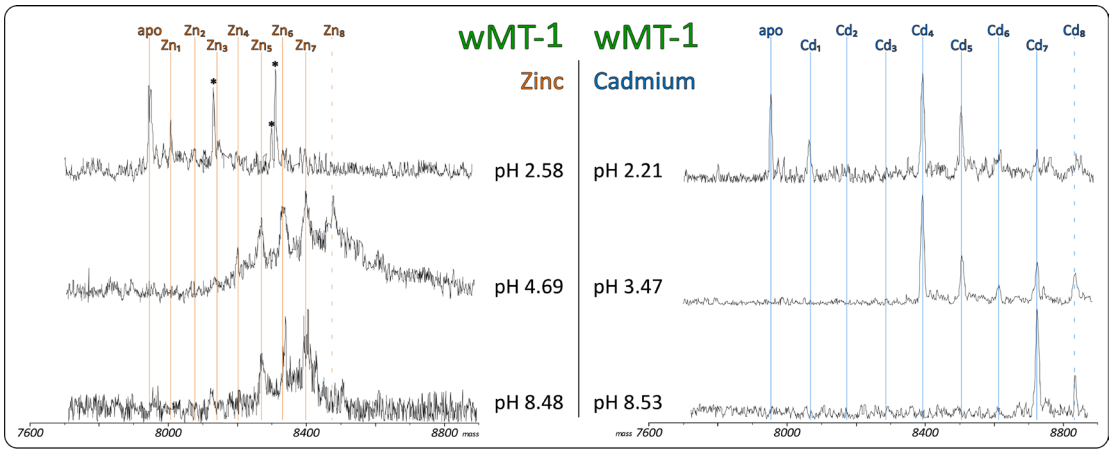


Figure 5.2. Deconvoluted mass spectra of cleaved wMT-1 with either zinc or cadmium. Samples in 20 mM NH_4HCO_3, with 10 % MeOH. Contaminant peaks indicated with '*'. Relevant concentration and metal:sulphur ratios were: Cd-wMT-1 (35.5 µM, 2.9 S:Cd); Zn-wMT-1 (23.2 µM, 4.0 S:Zn). Ideal ratios for binding seven metal ions were: Cd_7-wMT-1, 3.0 S:Cd; Zn_7-wMT-1, 3.0 S:Zn.

For cleaved wMT-1 at pH 2.5 (**Figure 5.2**), although there were some contaminants present, all zinc ions were stripped. This is identical to that observed at the same pH for Zn-wMT-2 (**Figure 5.1**). Similarly, at ≈ pH 8.5, a mixture of metal species was observed, with Zn_7- to Zn_5-wMT-1 comparable to the Zn_7- to Zn_5-wMT-2 species present at this pH. At an intermediate pH value (pH 4.7), multiple species were observed, ranging from Zn_8- to Zn_4-wMT-1. This behaviour was largely similar to that of Zn-wMT-2, which at pH 4.5 showed species from Zn_7- to Zn_2-wMT-2 (with Zn_8-wMT-2 observed at pH 7.0).

Cd-wMT-1 showed supermetalation at much higher pH (> 8) than Zn-wMT-1. The Cd_8- species was still present at pH 3.5, at a similar pH to the observation of Cd_8-wMT-2. At pH 3.5, Cd_4- was the major species (≈ 50 %) for both wMTs, but there was a higher percentage of Cd_8- to Cd_7- in wMT-1 (76 %) compared to wMT-2 (63 %). At low pH, there were similar amounts of apo-wMT-1 and apo-wMT-2, with major species

of Cd_4- and apo- for both MTs. It might be expected that wMT-2 would retain cadmium ions over a wider pH range, as it is a postulated to be a native cadmium binding protein [112], however the sequence similarity between wMT-1 and wMT-2 may make the magnitude of these differences very small.

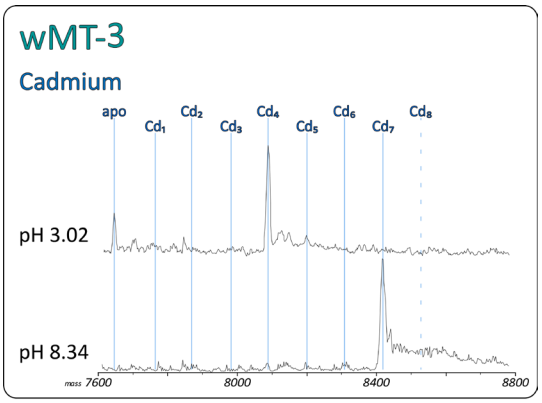


Figure 5.3. Deconvoluted mass spectra of Cd-wMT-3. Samples in 20 mM NH_4HCO_3, with 10 % MeOH. Cd-wMT-3 concentration and metal:sulphur ratios were 46.1 µM. 3.0 S:Cd. Ideal ratios for binding seven metal ions were Cd_7-wMT-3, 3.1 S:Cd.

Cd-wMT-3 at pH > 8 exhibited a sole Cd_7- species (**Figure 5.3**), which could indicate properties similar to Cd-wMT-2, rather than Cd-wMT-1. It is not known if this is significant, as comparison of the character of wMT-3 to wMT-1 or wMT-2 would also require experimental data for the zinc form of wMT-3. At low pH, no comparison could be drawn between Cd-wMT-3 and Cd-wMT-1, as the pH differences between the experiments were too great. However, the presence of apo-wMT-3 at pH 3.0 contrasts with Cd-wMT-2, where apo-wMT-2 was first observed at pH 2.5. The major species at pH 3.0 was Cd_4-wMT-3, but the presence of the apo- form suggested that the Cd-thiolate clusters in Cd-wMT-3 are more susceptible to protonation than Cd-wMT-2. This may be consistent with the proposed role of wMT-3, as a 'native' zinc carrier [112].

pH of ½ dissociation ($pK_a^{1/2}$)

Previous work performed by Stürzenbaum *et al*. provided a good reference point for the pH of ½ dissociation (the pH at which the wMT-2 had lost half of its bound metal ions), as the $pK_a^{1/2}$ for Cd-S●tag-wMT-2 was determined to be 2.8 [113]. Using the data obtained from the pH titration observed by mass spectrometry, a rough $pK_a^{1/2}$ can be calculated. The quality of the data obtained in this way was validated with Cd-wMT-2 by performing elemental analysis as reported by Stürzenbaum [113] and Tommey *et al*. [235]. Briefly, samples were incubated at different pH values, before the MT-bound and free metal were separated by gel filtration and analysed by ICP-OES or a similar technique. The ratio of MT-bound to free metal was then calculated. Both methods of analysis generated data that showed a sigmoidal shape (**Figure 5.4**), which is similar to titration curves for other MTs [168].

Investigation into the single clusters of MTs has highlighted that M_3Cys_9- clusters often have a higher $pK_a^{1/2}$ than M_4Cys_{11}- clusters. One study performed on the C-terminal M_3Cys_9- cluster of recombinant rabbit MT-II, displayed a $pK_a^{1/2}$ of 3.3 (compared to a $pK_a^{1/2}$ of 3.0 for the entire sequence) [236]. Similar experiments repeated with the M_4Cys_{11}- and M_3Cys_9- clusters from Human MT-II, reported similar findings, but with a difference of 0.5 pH units for the M_3Cys_9- cluster ($pK_a^{1/2}$ 3.6), compared to the M_4Cys_{11}- cluster ($pK_a^{1/2}$ 3.1) [231]. The overall $pK_a^{1/2}$ determined for human MT-II was determined to be $pK_a^{1/2}$ 3.03, slightly lower than the

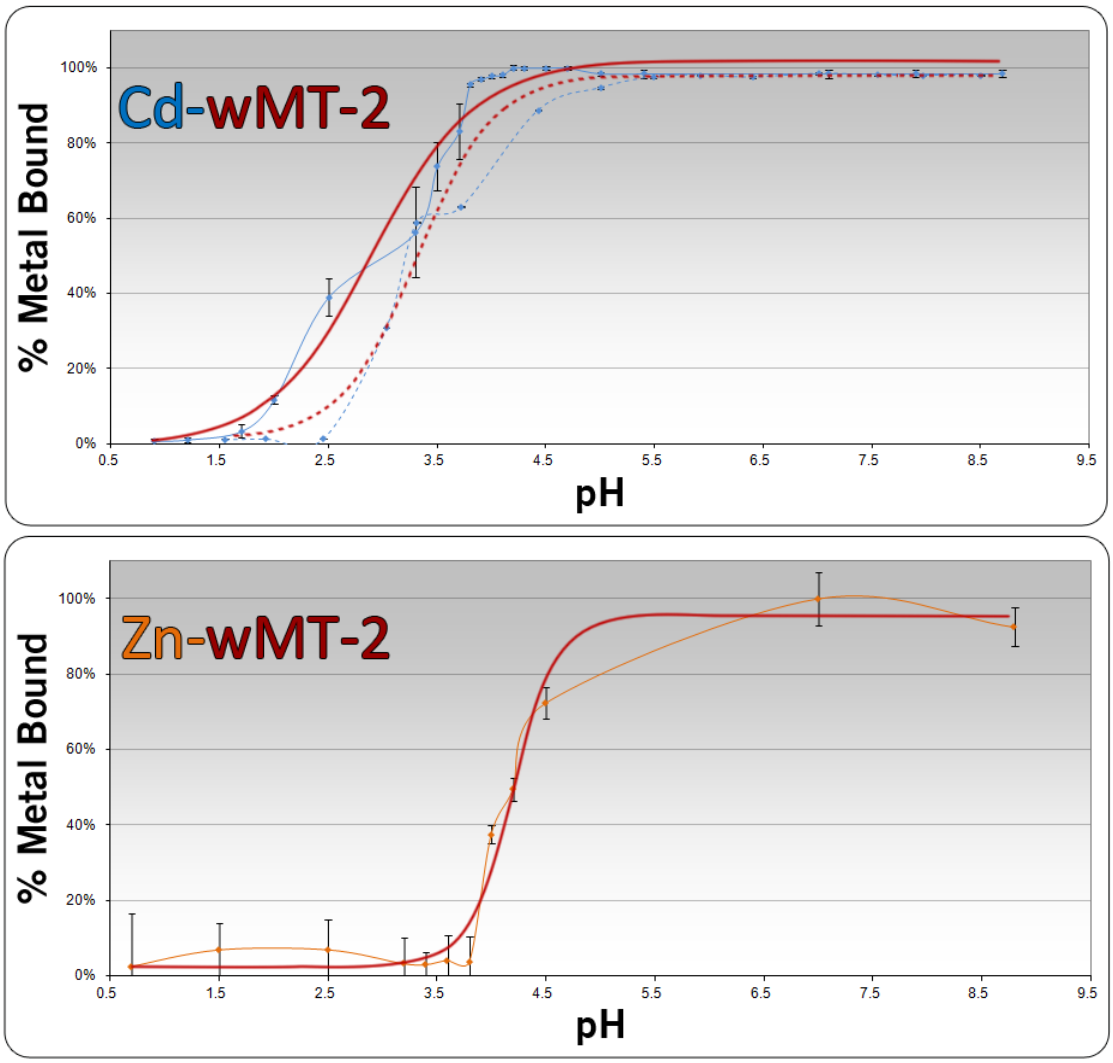


Figure 5.4. pH titration of cadmium and zinc wMT-2. Cadmium (**TOP**), solid **BLUE** line showing species estimated by MS, error bars show the measurement uncertainty between spectra. Dotted **BLUE** line corresponds to the analysis using ICP-OES. Zinc (**BOTTOM**) solid **ORANGE** line showing species estimated from MS. Note the error bars are significantly larger as the spectrum baseline is significantly more variable. **RED** lines are sigmoidal fitted curves (calculated in Origin Pro 8.5).

M_4Cys_{11}- cluster, suggesting some inter-domain influence on metal binding characteristics [237]. If this is the case, then additional amino acids from purification and expression tags could also influence these inter-domain effects. This validates efforts expended on the observation of the wMTs unencumbered by tags.

In the Cd-wMT-2 experiment, at pH 3.5 there was a large deviation from the sigmoidal curve. These data suggest a two-step demetallation process, corresponding to the complete removal of metals from one cluster, before the metals are removed from the other. This would support the data observed in the MS pH titration for Cd-wMT-2 (**Figure 5.1**). Considering the reported higher $pK_a^{1/2}$ for M_3Cys_9- clusters in other MTs, this data also supports the hypothesis that Cd-wMT-2 demetallation is cluster specific, with the M_3Cys_9- cluster demetallating first.

To allow comparison with previous MT data, a sigmoidal fit was performed with only one constant. Although fitting a biphasic curve requires two constants, in some cases experimental data is insufficient for this, therefore a single constant was used to fit all of the data reported. The $pK_a^{1/2}$ for Cd-wMT-2 as determined by MS was $pK_a^{1/2}$ 2.92 ± 0.10, and by ICP-OES was $pK_a^{1/2}$ 3.28 ± 0.07. The value obtained by MS was very close to the value reported by Stürzenbaum *et al*. ($pK_a^{1/2}$ 2.8) for Cd-S●tag-wMT-2 [113]. However, the value obtained by ICP-OES lies within the range of values reported for vertebrate MTs ($pK_a^{1/2}$ 3.0-3.5) [238, 239], but lower than those reported for plant MTs ($pK_a^{1/2}$ 3.9-5.8) [235, 240].

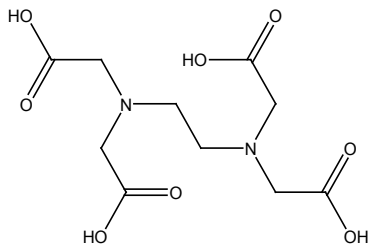
Further work is required to validate the use of mass spectrometry to determine $pK_a^{1/2}$ values, however it does seem to generate data that are

comparable to those obtained by elemental analysis. The $pK_a^{1/2}$ for Zn-wMT-2 was not determined by ICP-OES, however utilising the mass spectra in **Figure 5.1** gave a value of $pK_a^{1/2}$ 4.20 ± 0.05. As expected, this is higher than that observed for Cd-wMT-2, and similar to that of equine renal MT-II ($pK_a^{1/2}$ 4.5) [235] and rabbit MT-II ($pK_a^{1/2}$ 4.6) [239].

5.3 Competition Reactions

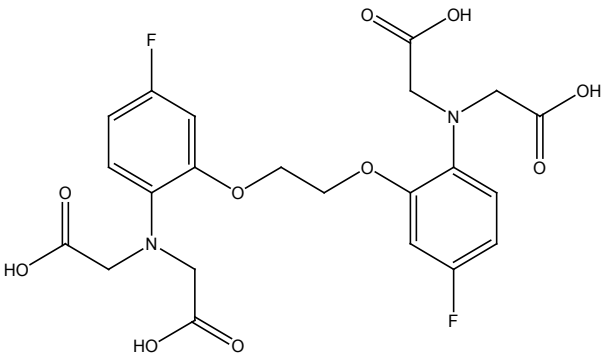
Reaction with EDTA

To probe the lability of cadmium ions bound to wMT-2 at neutral pH, the competitive divalent metal chelator EDTA (Ethylenediaminetetraacetic acid, **Figure 5.5 LEFT**) was used. Fully metallated cleaved Cd-wMT-2 (6.9 cadmium ions per wMT-2 molecule) was incubated with one equivalent (with respect to metal) of EDTA, and incubated for 60 mins (**Appendix 15**). A low concentration of EDTA was used in this initial experiment to overcome the propensity for EDTA to quickly strip metal ions from MT, and therefore only observe apo-wMT-2 [241].



EDTA

Ethylenediaminetetraacetic acid



5F-BAPTA

5,5'-Difluoro-1,2-bis(2-aminophenoxy)ethane-N,N,N',N'-tetraacetic acid

Figure 5.5. The structure of the two metal chelators used in the following section, EDTA (**LEFT**) and 5F-BAPTA (**RIGHT**). Standard IUPAC names are displayed below the respective structures.

Figure 5.6 shows the appearance of $EDTA_xCd_7$-wMT-2 complexes after just 30 seconds, highlighting the propensity for metal-binding proteins to form EDTA adducts [242, 243]. Spectra taken at further time-intervals did not show significant differences, indicating that wMT-2 was possibly inert to cadmium removal by low concentrations of EDTA within 60 min (**Appendix 15**).

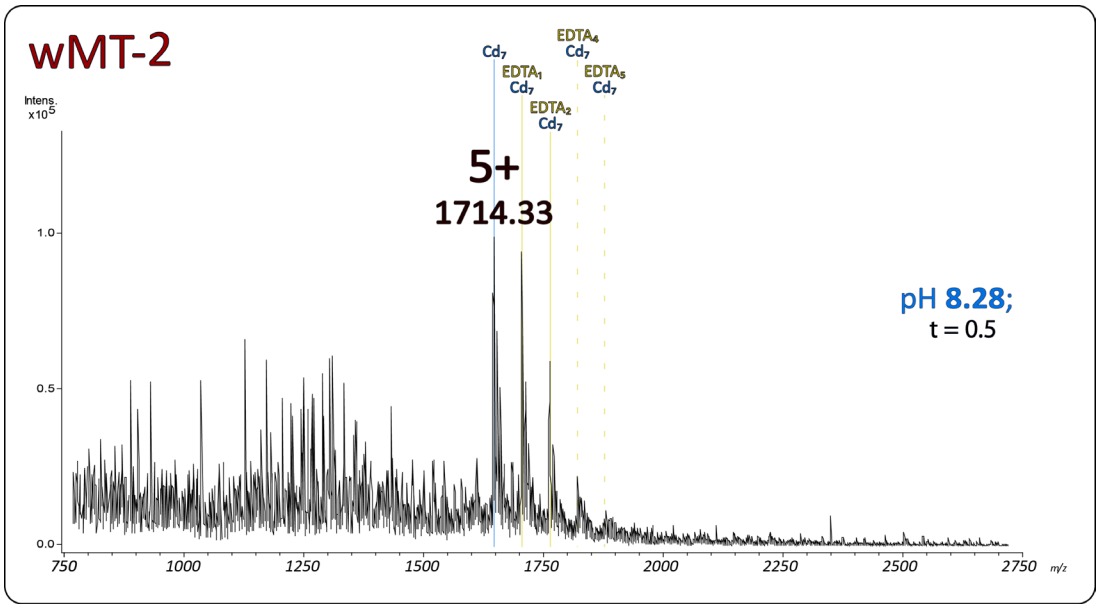


Figure 5.6. Raw mass spectrum at 30 seconds (0.5 mins) of cleaved Cd_7-wMT-2 incubated with 1 eq of EDTA (wrt cadmium concentration).

Although only a single equivalent of EDTA was added, species up to $EDTA_5Cd_7$-wMT-2 were observed in the raw spectrum. Cd-wMT-2 contains 20 cysteine residues which are postulated to chelate metals, 8 bridging ligands (contributing 2 bonds), and 6 terminal ligands. Although EDTA usually forms 6 multi-dentate bonds with divalent metal ions, this is not always the case. This suggests that EDTA is able to form adducts, possibly by displacing the terminal ligands, and adopting a bidentate conformation, whilst still maintaining seven metals bound. This would cause the appearance of the particular adducts (of $EDTA_1$- to $EDTA_6$-) in

MS, but no higher states ($EDTA_7$- upwards). Alternatively, the EDTA adducts may be with wMT-2 itself, and not involve interaction with metal ions. It may be inferred from the raw spectra (**Appendix 15**) that Cd-wMT-2 is remarkably stable with respect to cadmium removal by low concentrations of EDTA, similar to studies performed with Cd-MT by Petering *et al*. [244]. However, later NMR experiments showed that rabbit MT-II may react in an 'all-or-nothing' way with low concentrations of EDTA [242]. There is no peak corresponding to apo-wMT-2 in the raw mass spectra however, so this is unlikely to be the case.

Reaction with 5F-BAPTA

Another competitive chelator is 5F-BAPTA (**Figure 5.5 RIGHT**). The use of 5F-BAPTA allows the observation of changes in the environment of the fluorine atoms of the chelator by ^{19}F 1-dimensional NMR spectroscopy [200]. The experiments were performed using the technique first reported by Smith *et al*. [245], with suggested modifications by Benters *et al*. [246] for experiments involving cadmium ions. 5F-BAPTA forms 1:1 complexes with divalent metal ions, each complex having a distinctive chemical shift in the ^{19}F spectrum (**Figure 5.7**). Using the integral values of the peaks for free 5F-BAPTA (≈ -121 ppm) and cadmium-bound 5F-BAPTA (≈ -116 ppm) / zinc-bound 5F-BAPTA (≈ -117 ppm) in **Equation II** [200], allowed the overall metal-binding affinity (apparent metal-binding constant, K_{wMT}) for wMT-2 to be calculated.

$$K_{wMT} = K_{BAPTA} \left(\left(\frac{[M_x\text{-}wMT] \; x \; \beta}{\alpha^2 \; x \; [BAPTA_0]} \; x \; (\alpha + \beta) \right) - \frac{\alpha}{\beta} \right)$$ Equation II

Equation II. K_{wMT}, apparent metal-binding constant of wMT; K_{BAPTA}, equilibrium binding constant of 5F-BAPTA with metal ion; [Mx-wMT], concentration of metal filled sites of MT; [$BAPTA_0$], initial 5F-BAPTA concentration; α, M-5F-BAPTA integral; β, Free 5F-BAPTA integral. From [200].

Using the conditional equilibrium binding constant of 5F-BAPTA with cadmium ($K_{Cd\text{-}BAPTA} = 5.62 \times 10^{11}\ M^{-1}$) and zinc ($K_{Zn\text{-}BAPTA} = 8.13 \times 10^{9}\ M^{-1}$) for the particular ionic strength and pH of the experiment [151], the calculated apparent metal-binding constants of wMT-2 under these conditions were: $K_{Cd\text{-}wMT\text{-}2} = 3.88 \times 10^{13}\ M^{-1}$ and $K_{Zn\text{-}wMT\text{-}2} = 8.26 \times 10^{10}\ M^{-1}$.

The spectra in **Figure 5.7** show that, although potentially inert to metal removal by EDTA, 5F-BAPTA will remove both zinc and cadmium from wMT-2. As the chelating residues of EDTA and 5F-BAPTA are very similar (**Figure 5.5**), this behaviour may be due to 5F-BAPTA being present in a much higher concentration in this reaction (approximately 42*x* higher). Comparing these values to the apparent metal-binding constants for other MTs, derived under similar conditions (**Figure 5.8**), we can see that whilst the value for $K_{Zn\text{-}wMT\text{-}2}$ seems to be neither particularly high nor low, the value for $K_{Cd\text{-}wMT\text{-}2}$ is comparatively low. In comparison, the major cadmium binding protein CeMT2 from *C. elegans* [151], has a $K_{Cd\text{-}CeMT2}$ approximately 1.5 orders of magnitude greater. However, this

could be a result of CeMT2 binding all metals strongly, exhibiting a $K_{Zn\text{-}CeMT2}$ approximately 1.1 orders of magnitude greater than wMT-2. If we compare the behaviour of Cd-wMT-2, $pK_a^{1/2} \approx 2.8\text{-}3.4$, $\log(K_{Cd\text{-}wMT\text{-}2}) = 13.6$, with that of human Cd-MT-II, $pK_a^{1/2} = 3.0$, $\log(K_{Cd\text{-}MT\text{-}II}) = 14.8$, these values appear to be reasonable. However, with wMT-2 being the postulated native cadmium binding protein in earthworms [112], one might expect either a higher $\log(K_{Cd\text{-}wMT\text{-}2})$ or lower $pK_a^{1/2}$ than their experimentally determined values.

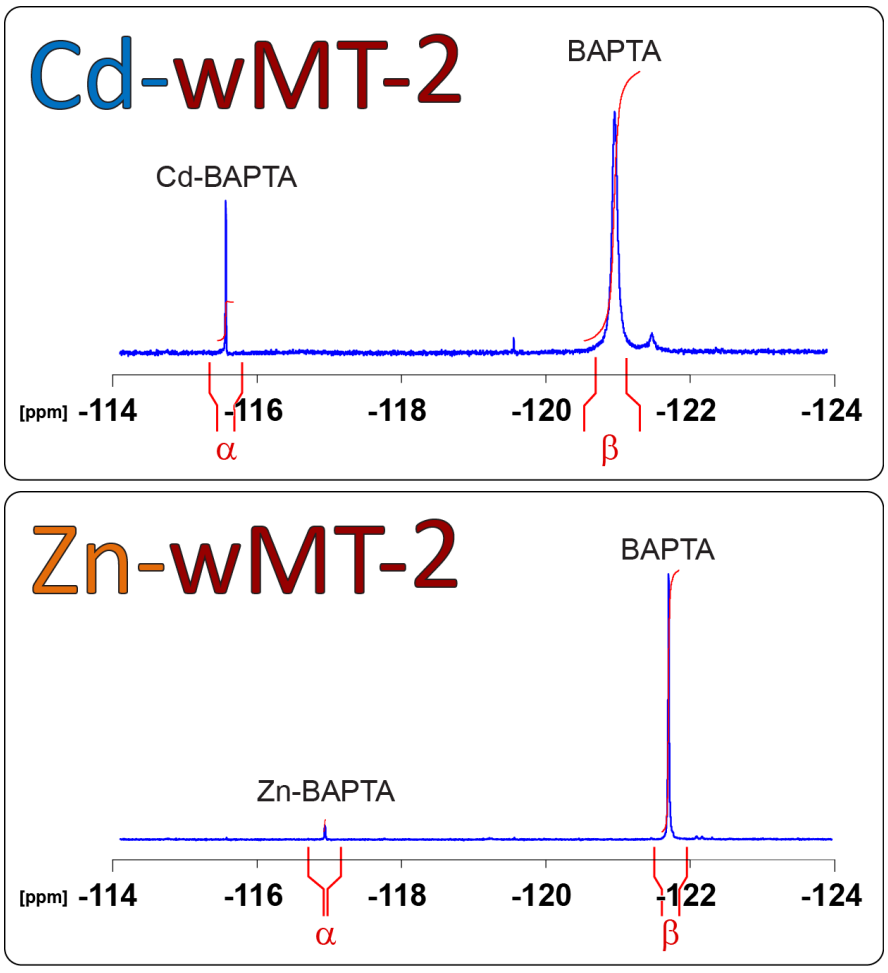


Figure 5.7. ^{19}F NMR spectroscopy of the cadmium and zinc form of wMT-2. Notice the approximate areas and magnitudes of α and β peaks. These were integrated before being utilised in Equation II. Both samples were approximately 500 µM wrt metal ion concentration, in 10 mM Tris-Cl, pH 8.1; Ionic strength of the solution was 4 mM.

The most crucial characteristics of metal-binding proteins within a system are thought to be their relative affinities for metal ions when compared with other MTs and metal-binding proteins [247]. Whilst wMT-2 has an apparent metal-binding constant approximately 3 orders of magnitude higher for cadmium over zinc, further experiments with wMT-1 would be necessary for this relationship to be investigated fully.

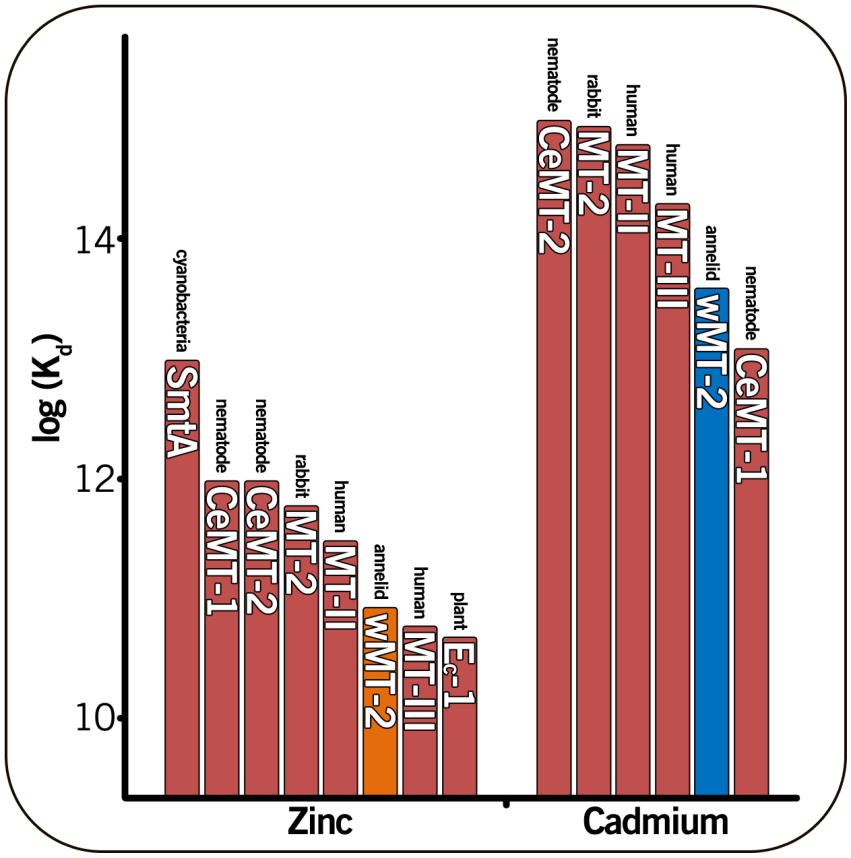


Figure 5.8. Comparison of derived apparent metal-binding constants of MTs derived in 10 mM Tris-Cl pH 8.1, < 5 mM ionic strength. Values expressed as log(K_{MT}) values. Human MT-II and MT-III [200], SmtA & E_c-1 [168], CeMT1 and CeMT2 [151].

5.4 Effects of zinc and cadmium on the backbone of wMT-2

Two separate expressions and purifications were carried out in the presence of either zinc or cadmium. The parent overnight culture of both experiments was identical, in an effort to normalise protein expression

levels. Sample concentrations for each experiment were 524 µM for cleaved Cd-wMT-2 (6.7 cadmium ions per wMT-2 molecule, measured after NMR data acquisition), and 483 µM for cleaved Zn-wMT-2 (5.9 zinc ions per wMT-2 molecule). To minimise loss of protein from purification, the samples were not buffer exchanged by gel filtration, but the buffer replenished during 4 concentration by centrifugation steps (25-35 mins, 4 °C, 3,000 x *g*) with a source of 20 mM ammonium bicarbonate made with Milli-Q water and pH adjusted to ≈ 6.8. This was sufficient to achieve a sample at pH 7.0, with no observable precipitation of protein during centrifugation. **Figure 5.9** illustrates the stark differences between the two samples. Both datasets were acquired at 700 MHz, 298 K, pH 7.0. Although some residues are present in both spectra, the spectrum for cleaved Cd-wMT-2 is significantly better defined; indicative of a well-folded protein [248].

There could be two reasons for the poor quality spectrum of cleaved Zn-wMT-2: wMT-2 being partly folded due to insufficient zinc ions; or the wMT-2 backbone being more flexible when binding zinc. As the stoichiometry of the sample is lower than 7, the implication is that Zn-wMT-2 is undermetallated. As metal ions are vital for MT folding [249], undermetallation may lead to increased flexibility - broadening the NMR crosspeaks to the point where they are no longer observed. However, Zn-wMT-2 samples exhibiting stoichiometry measured by ICP-OES between 5.5-6.3 zinc ions per wMT-2 (data from across 12 preparations) still showed Zn_7-wMT-2 as the major species by mass

spectrometry (**Figure 3.42**). The difference between the spectra is likely due to the backbone fold being more flexible in the zinc form, or the sampling of multiple conformations of the protein caused by incomplete metallation of the protein. The generally lower quality of both MS and NMR spectra of Zn-wMT-2 would support the hypothesis that the 'native' metal for wMT-2 is cadmium [112].

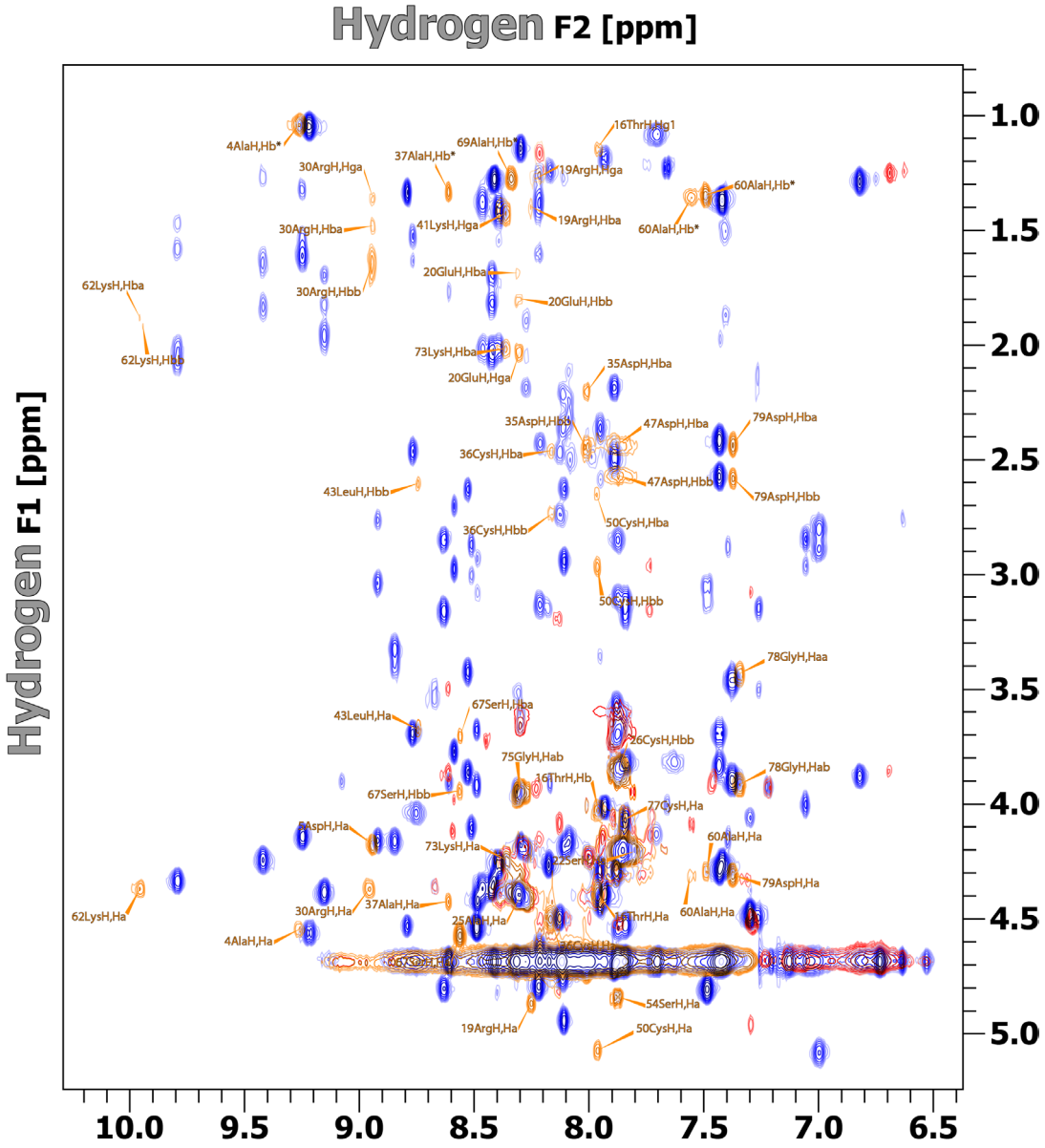


Figure 5.9. 2D TOCSY NMR of 500 µM cleaved wMT-2 with: Cadmium (**BLUE**) or Zinc (**ORANGE**). Picked peaks in the zinc spectrum (in **ORANGE**) corresponded closely (ppm < 0.2) with those observed in the assigned cadmium TOCSY (assignments not shown). Buffer was 20mM ammonium bicarbonate, pH ≈ 7.0 with 10 % D_2O. **RED** peaks are peaks observed in the zinc spectrum which were unable to be unambiguously assigned from the cadmium spectrum.

Residues which were able to be unambiguously assigned in the spectra of cleaved Zn-wMT-2 were compared the assignment of cleaved Cd-wMT-2. The partial assignment of Zn-wMT-2 indicated that only 11 of the 45 residues in the M_4Cys_{11}- cluster were assignable, with 11 of 30 residues in the M_3Cys_9- cluster assignable (not including the proposed linker region) (**Figure 5.10**). Overall, this meant that a similar number of crosspeaks were assigned in both domain clusters (M_3Cys_9- ≈ 37 %; M_4Cys_{11}- cluster ≈ 36 %).

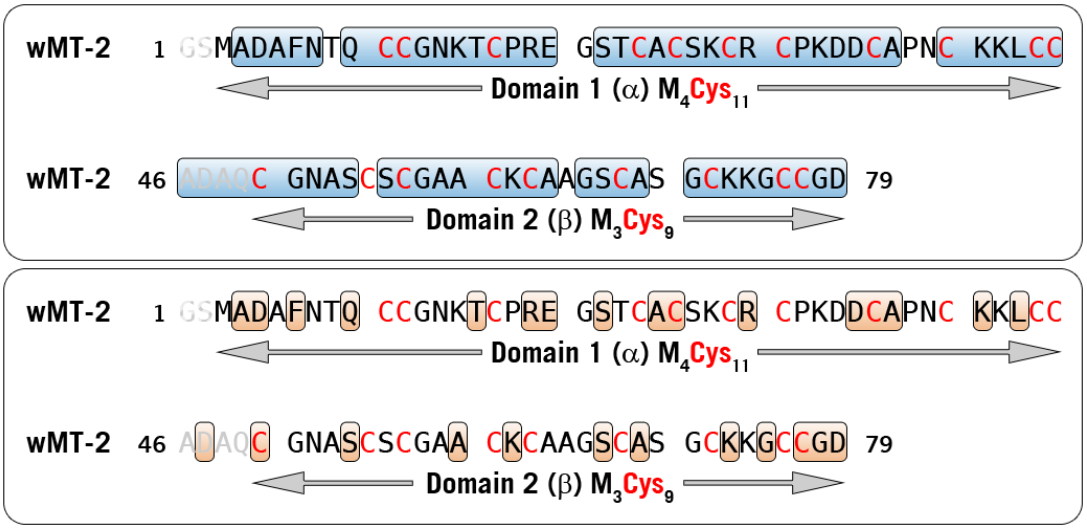


Figure 5.10. Comparison of the residues identified in 2D TOCSY experiments. Scheme indicates: Cadmium (**BLUE**) or Zinc (ORANGE). The proposed linker region is shown in GREY, with the N-terminal -GS- residues remaining from the removal of the S●tag also highlighted.

Of the unassigned residues in **Figure 5.9**, 4 spin systems exhibited a J-type pattern. The difficulty in assigning these ambiguous J-type spin systems in the spectra of Zn-wMT-2 could indicate that to accommodate zinc ions the protein has to adopt an alternative fold to that of the cadmium form. This would position the cysteine residues in different environments to those observed in Cd-wMT-2, significantly altering their chemical shifts. The assigned residues are spread throughout the entire sequence of wMT-2 (**Figure 5.10**), with gaps of no more than four amino

acids between assigned amino acids, indicating that both clusters are at least partially folded.

There is clustering of assignable crosspeaks around cysteine residues, as might be expected from their structure-forming ability. This being considered, the apparent absence of resolved cysteine residues in the zinc form would suggest that the ambiguous J-type spin systems could be the coordinating cysteine residues absent from **Figure 5.10** (ie. Cys17, Cys24, Cys44, Cys61, Cys68, Cys76). This suggests that while some MTs do show domain preferences for metal ions [250], this appears not to be the case for wMT-2.

5.5 Metal exchange reactions

Cadmium exchange reaction

To investigate the propensity for wMT-2 to preferentially bind cadmium ions over zinc ions, aliquots of Zn-wMT-2 were incubated with increasing equivalents of Cd^{2+} for 60 mins, before free metal was separated from protein-bound metal by gel filtration. Both the 'protein' fraction and the 'salt' fraction were analysed for metal composition (**Figure 5.11**).

The replacement system appears to be stoichiometric up until 5 equivalents of cadmium. The deviation when 6 equivalents of cadmium are added is likely due to measurement uncertainty. However on addition

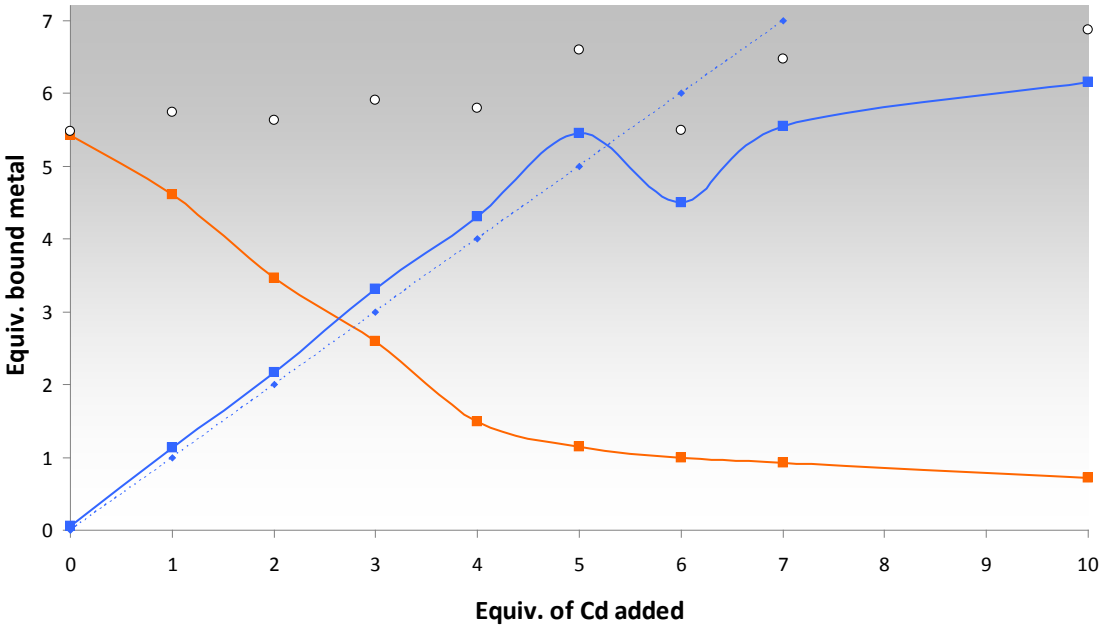


Figure 5.11. Titration experiment with increasing equivalents of Cd^{2+} added to a solution of cleaved Zn-wMT-2. Total metals bound to cleaved wMT-2 shown as open circles, zinc content shown in **ORANGE**, cadmium content shown in **BLUE**. The dashed **BLUE** line indicates the values expected for a 1:1 stoichiometric replacement of zinc by cadmium.

of > 7 equivalents, incomplete exchange (gradually increasing to 90 % total cadmium) was observed, possibly as thermodynamic equilibrium is reached. During the experiment, the initial stoichiometry of 5.5 zinc ions per wMT-2 gradually increased to 6.9 metal ions per wMT-2 (and a final stoichiometry of 6.2 cadmium ions with 0.7 zinc ions bound). This final stoichiometry indicates that a stable conformation of wMT-2 with 7 metal ions bound can be achieved with contributions from both cadmium and zinc. Similar behaviour was observed by Capdevila *et al*. in their paper investigating the metal preference of different MTs. It was found that MTs with a preference for divalent metal ions showed a clear reluctance for *in vitro* Zn/Cd exchange [214]. As wMT-2 is postulated to bind divalent cadmium *in vivo* [113], even though the measured $K_{Cd\text{-}wMT\text{-}2}$ value is 3 orders of magnitude higher than for $K_{Zn\text{-}wMT\text{-}2}$, complete zinc exchange is not observed.

Zinc displacement by cadmium during purification

The stability of the M_7-wMT-2 species was reinforced with data obtained from a sample expressed in the presence (excess > 500 µM) of zinc, and purified in the presence (excess > 1 mM) of cadmium (**Figure 5.12**).

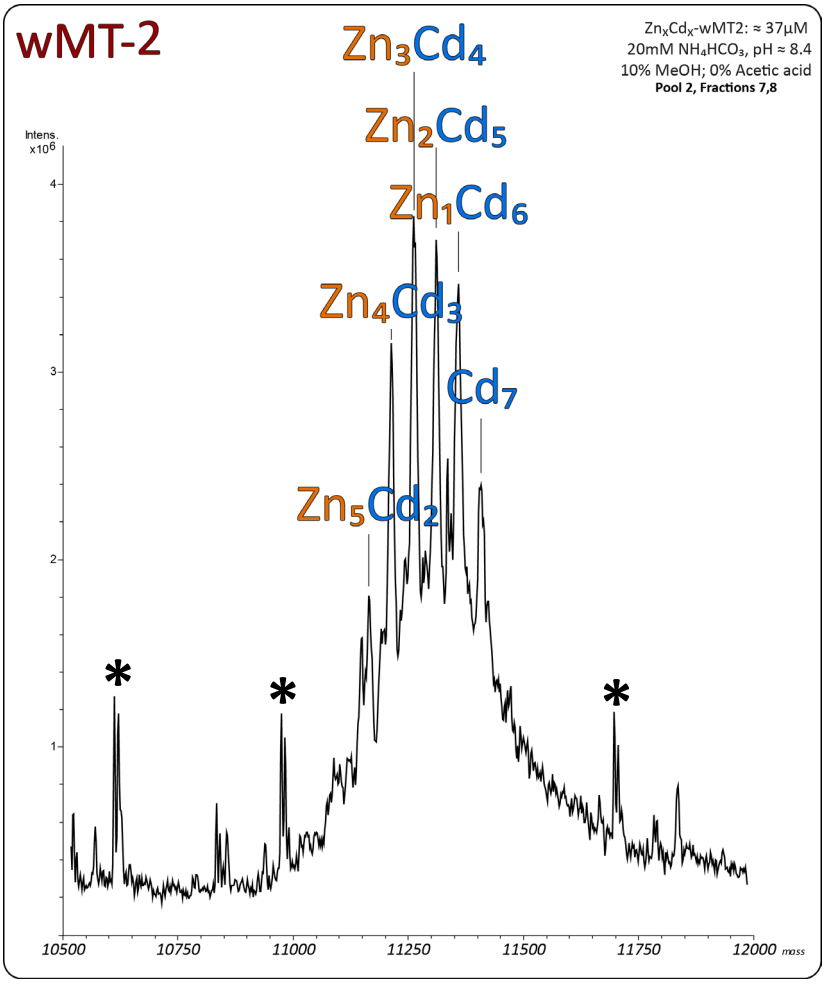


Figure 5.12. Deconvoluted mass spectrum of mixed-metal species in S●tag-wMT-2. S●tag-wMT-2 concentration 36.6 µM, pH 8.39. Sample in 20 mM NH_4HCO_3, and 10 % MeOH. Zinc (**ORANGE**) and cadmium (**BLUE**) forms are indicated, '*' indicates a contaminant peak, or deconvolution artefact. Peak masses and percentages of species are summarised in **Appendix 16**.

When purified, there were 6 major species, all of which contained seven metal ions. Due to the significant excess of cadmium ions being present during purification (the protein concentration during purification is in the tens of µM), complete displacement and formation of a Cd_7- species was

observed. This indicated that although complete metal removal could not be achieved in the incubation experiment, it is likely that the incubation produced a sample containing a similar mixture of species. The species with the highest abundance was Zn_3Cd_4-, this could indicate the presence of a Cd_4- M_4Cys_{11}- cluster and a Zn_3- M_3Cys_9- cluster. However without isolation of the individual domains, it is not known at this point if this is the case. It is known however that in vertebrate apo-MT, cadmium begins to fill the M_4Cys_{11}- cluster first [251].

The stability as mixed-metal species may be important for wMT-2, which exists in an environment containing a mixture of zinc and cadmium ions. Were there insufficient cadmium to completely fill metal sites in wMT-2, zinc could be temporarily utilised to slow the degradation rate of the protein *in vivo*. If concentrations of cadmium were subsequently increased, a slow metal exchange appears to be possible, relinquishing the bound zinc ions. The ability to form a stable structure when partially filled by cadmium ions may be of benefit for a wMT-2 whose proposed role is to protect the earthworm from excesses of cadmium [113].

5.6 Summary

wMT-2 exhibits differential behaviour to acidification depending on the metal bound. MS has shown that demetallation of Cd-wMT-2 is cooperative, whereas demetallation of zinc proceeds non-cooperatively. Both wMT-1 and wMT-2 indicate a propensity to become supermetalated,

with the appearance of a transient Cd_8- species. This was not observed for the novel protein, wMT-3. The stability of Cd-wMT-2 was highlighted by exhibiting no transfer to EDTA, and showed high structural stability from well-defined 2-dimensional spectra. The inability for complete metal exchange (behaviour not uncommon for this type of MTs, even considering the > 3 orders of magnitude preference for cadmium) is therefore likely due to wMT-2 being a native divalent metal-binding MT.

6

Conclusions from structural and metal-binding studies

The 'holy grail' of working with proteins is in ascertaining a proteins role *in vivo*, and the characteristics which make it suitable for that function [252]. As such, a wealth of information about biological function has been gained from identifying protein characteristics from experimentally determining their 3-dimensional structure [253]. There are two barriers to experimental work with proteins: the ability to recombinantly express or obtain native protein in milligram quantities; and the ability to purify the protein to an acceptable level. This chapter consists of a general summary of the work presented in this thesis, a discussion of the more important results, and concludes with some avenues for further research.

6.1 Validating the purification methods used for obtaining *Lumbricus rubellus* wMTs

The genes encoding the three wMTs were confirmed by sequencing plasmids received from the Stürzenbaum group. In the case of wMT-1, the plasmid sequence was corrected to match that reported in the literature [99]. Two approaches for the removal of the N-terminal S●tag were investigated: either through subcloning or cleavage. Using a combination of ion exchange and size exclusion chromatography, sufficient yields of pure wMT-2 and wMT-3 were isolated and

characterised by mass spectrometry. Lower yields of impure wMT-1 were also able to be characterised by MS. Both expression and purification protocols were optimised with wMTs in mind.

Generally speaking, obtaining yields of MT in the region of 1 mg/L expression culture is acceptable for structural studies [124], with the levels for wMTs determined to be 0.8-1.7 mg/L. However, whilst the purification protocol was employed for production of pure non-labelled wMT-3 and unlabelled and double-labelled wMT-2, further optimisation is required to obtain a similar quantity of pure wMT-1. All three MTs when expressed recombinantly with either zinc or cadmium were detected by mass spectrometry with seven divalent metal ions bound as the major species. This is the first time that wMT-1 and wMT-2 have been isolated (as homometallated species) without the encumbrance of an S●tag, and the first time that the novel wMT-3 has been isolated. With respect to purification, wMT-2 was obtained in highest yields when expressed in the presence of cadmium, and slightly lower yields when expressed in the presence of zinc. This pattern was mirrored in the expression of wMT-1; with comparable data for wMT-3 not available at this time. The difference in MT expression level in the presence of cadmium is not uncommon, for example higher yields of MT from rainbow trout and monkey were obtained in the presence of cadmium rather than zinc [212, 213].

6.2 Structural investigation of wMT-2

For the structural investigation, a Cd-wMT-2 sample labelled with ^{13}C and ^{15}N was produced. From the combination of multidimensional NMR experiments, and the use of both unlabelled and the double-labelled cleaved Cd-wMT-2 samples, a spectral assignment was undertaken. Using restraints generated from the assigned spectra, structure calculations for the two domains of wMT-2 were then performed. However, in the absence of meaningful restraints from ^{111}Cd NMR experiments, the metal-cysteine connectivities were approximated by minimising the backbone violations of restraints with different permutations of connectivities.

Although currently the main route for structure determination is through experimental means, there are a growing number of computational approaches based on homology modelling from protein primary sequences [254, 255]. Due to the nature of homology modelling, the quality of the resultant model is directly proportional to the quantity of data available, ie the number of experimentally determined 3-dimensional protein structures within a library [256]. Therefore for an acceptable structure to be modelled, at least one 'template' experimentally solved structure is required with high primary sequence similarity [255].

Although vertebrate MTs show significant similarity, MT-3 exhibits differential cluster preferences, with the α-domain (M_4Cys_{11}- cluster) containing copper, not zinc [257]. The novel -CPCP- sequence in the α-domain of MT-3 is thought to tune the protein to be biologically active, by the preferential binding of copper ions [258]. Although also somewhat active with zinc bound, when the–C**P**C**P**- motif is mutated to the -C**S**C**A**- sequence found in MT-1 and MT-2, the functional activity of MT-3 is lost [257].

In the case of MTs, the difficulty in using a homology-based approach for structure determination is that there are simply no available 'template' sequences for the majority of Class II MTs (ie, those with only some similarity to equine renal MT). To fulfil the criterion of high primary sequence similarity, the sequences must also share an evolutionary relationship; there is no such relationship between vertebrate and invertebrate MT sequences (**Appendix 1**). This reduces the quantity of potential 'template' sequences for wMT-2 to just 3 (lobster, crab, sea urchin). Even when an evolutionary relationship is present, such as for MT-3 and MT-1/MT-2, a homology modelling approach would have been problematic, due to the unique motif present in MT-3.

The evolutionary diversity and the surge of current research on MTs, has meant that novel folds have been recently characterised in cyanobacteria (SmtA [259]), and plants (wheat E_c [260]). Although wMT-2 Domain 1 showed little similarity to other available MT structures, the similarity of

wMT-2 Domain 2 to the β-domain of MTA [87] was used as a starting point for the determination of cysteine-metal connectivities. However, this connectivity pattern did not have the lowest backbone RMSD of the permutations investigated (see **Chapter 4.6**). As metal-cysteine connectivities had not been experimentally determined, it was hypothesised that the cluster formation with the lowest backbone RMSD was correct, as per Peroza *et al*. [226]. The structure obtained using 'Alternate 1' connectivities fits this criterion. Given the number and magnitude of violated restraints, and backbone RMSD values, the structure of the two domains of wMT-2 presented in this thesis show the same characteristics as other accepted MT structures [87, 144, 224, 261, 262].

Comparison of the resultant structure of wMT-2 Domain 2 (Cys_9- cluster) showed only slight similarity to the backbone conformation of the Cys_9-cluster of sea urchin MTA (**Figure 6.1**).

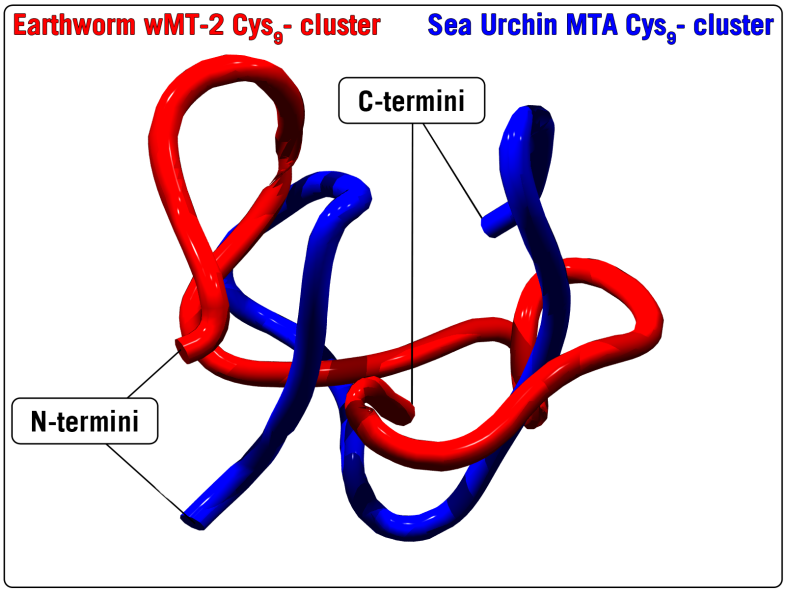


Figure 6.1. 3-dimensional structure of wMT-2 Cys_9- cluster 'Alternate 1' (**RED**) fitted to sea urchin MTA Cys_9- cluster (**BLUE**). Although the general form is similar, the fit is relatively poor.

When the backbone conformation of residues 37Ile-64Cys of sea urchin MTA and 50Cys-77Cys of wMT-2 were compared, an RMSD of 8.08 Å* between the two structures was calculated. If only the conserved positions of the cysteine residues at the N-terminus of the Cys_9- cluster (**Appendix 18**) were considered, this value was reduced to 5.84 Å. This value remains higher than might be expected from the similarity in primary sequence of the two MTs, suggesting the two sequences show little evolutionary relationship. The apparent lack of similarity in 3-dimensional structure between the Cys_9- cluster of wMT-2 and sea urchin MTA may rationalise why the metal-cysteine connectivities generated from transposing the connectivities found in sea urchin MTA to the Cys_9-cluster of wMT-2 did not exhibit the lowest backbone RMSD of the permutations attempted.

Although the sea urchin MTA Cys_9- cluster shows highest sequence identity to the wMT-2 (45.2 %†), representatives from both the vertebrate MT family (41.4 % identity for rat MT-2, PDB 2mrt [259]) and the invertebrate MT family (37.5 % identity for crab MT-1 (C-terminus), PDB 1dmc [263]) were investigated for structural similarities. The results of fitting the respective 3-dimensional structures to the wMT-2 Cys_9-cluster are shown in **Figure 6.2**.

* obtained from SWISS PDB Viewer, calculated for 28 residues

† GENESTREAM Align Query Tool http://xylian.igh.cnrs.fr/bin/align-guess.cgi

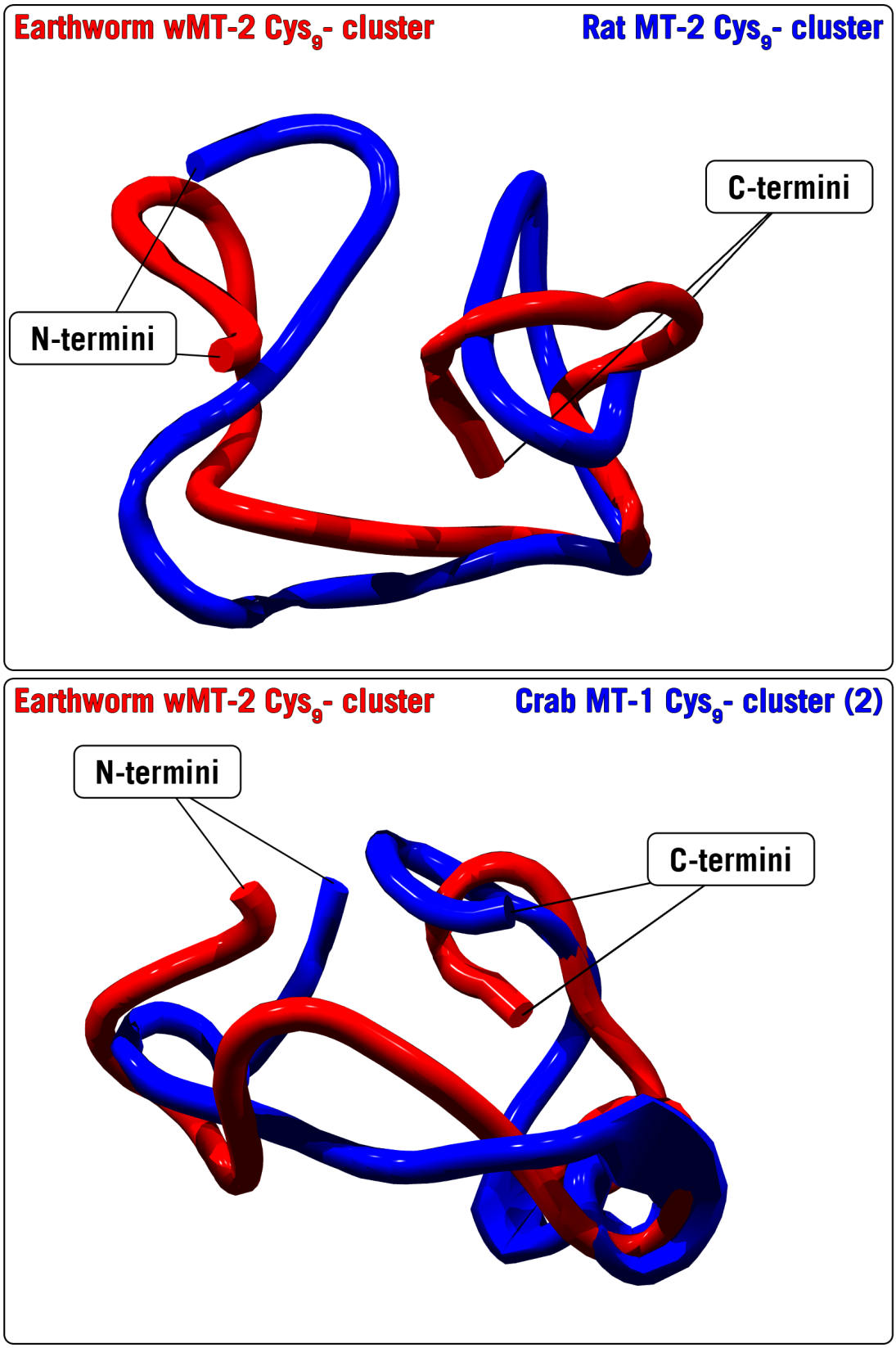


Figure 6.2. 3-dimensional structures of wMT-2 Cys_9- cluster 'Alternate 1' (**RED**) fitted to crab MT-1 Cys_9- cluster (C-terminus) (**BLUE**, **TOP**), and rat MT-2 Cys_9- cluster (**BLUE**, **BOTTOM**). The backbone of rat MT-2 and crab MT-1 both conform generally to the shape of wMT-2, and show greater similarity than sea urchin MTA.

Both crab and rat structures showed significant similarity to that of wMT-2. Values for backbone RMSD comparisons were lower than that for sea urchin MTA (8.08 Å), with rat MT showing a backbone RMSD difference of 5.72 Å, and crab MT showing a backbone RMSD difference of 4.50 Å.

This comparison suggests that the currently identified MT with closest evolutionary link to earthworm wMT-2 may be crab MT-1. However, although the backbone conformations were similar, the positions of cysteine residues in the primary sequences of earthworm and crab MT are not conserved (hence the low sequence identity). The structural features of the specific backbone conformation adopted by earthworm wMT-2, crab MT-1 and rat MT-2 may give the first indication of a 'cadmium' specific fold utilised in MT β-domains/Cys_9- clusters, as all three MTs are differentially expressed when the organism is challenged with cadmium [113, 130, 264].

6.3 Metal-binding studies with wMTs

An investigation into isoform specific metal-binding behaviour was performed, with an emphasis on wMT-2. Using MS, a large-scale pH titration was performed for both cleaved Zn- and cleaved Cd-wMT-2. Smaller-scale pH titrations for cleaved Zn- and cleaved Cd-wMT-1, and Cd-wMT-3 were also performed. For cleaved wMT-2, the $pK_a^{1/2}$ was then calculated using two techniques (MS and ICP-OES). In addition to

competition with protons, the metal chelator 5F-BAPTA was introduced to samples of cleaved wMT-2, which allowed the calculation of the apparent metal-binding constants for the cadmium and zinc form ($K_{Cd\text{-}wMT\text{-}2}$ and $K_{Zn\text{-}wMT\text{-}2}$). The metal chelator EDTA was also incubated with a sample of cleaved Cd-wMT-2, and monitored for an hour. NMR was performed on both cleaved Zn- and cleaved Cd-wMT-2 to observe the effect of metal binding on the backbone fold of wMT-2. Finally, the inertness of cleaved wMT-2 to zinc displacement was investigated by incubation with cadmium.

To determine the behaviour of MTs, extensive metal-binding studies are usually undertaken. The *in vitro* metal-binding studies can, in some cases, give indications of *in vivo* properties [122, 146]. Although the trend for MT metal-binding affinities generally follows the Irving-Williams series (Zn < Co < Pb < Cd < Cu / Ag < Hg) [247], metal ion preferences have begun to be elucidated for a number of MTs from different organisms. It was found that for the two MTs in *C. elegans*, CeMT1 and CeMT2, it was not the direct affinity for metal ions that was important, but the relative affinities of each isoform for each metal ion [151]. Studies on metal preferences *in vitro*, coupled with *in vivo* studies. proposed the hypothesis that although both CeMTs are utilised with respect to zinc handling, CeMT2 had the major role in cadmium accumulation and detoxification [151]. The two MTs have over 50 % sequence identity, but show different metal-binding preferences, with CeMT2 favouring cadmium significantly more than CeMT1 [146].

Similarly a system in the invertebrate *Helix pomatia* (Roman snail) has been elucidated, resulting in the identification of two distinct isoforms CuMT or CdMT [122, 247]. The two isoforms in snails also exhibit over 50 % sequence identity, but show differential metal-binding characteristics.

There are two explanations for this behaviour. The first is that MTs themselves can differentiate between metal ions, as in the case of CeMT1 *in vitro* [146]. Based on this idea, a system was proposed for classifying MTs is based solely on their metal-binding preferences. The system suggested by Capdevila *et al*. [214] showed that generally MTs exhibit characteristics of either Zn-thionein or a Cu-thionein. As the character is reflected during *in vitro* experiments, the major contributor to the metal-binding preference of MT is proposed to be the protein fold. Zn-wMT-2 was found to be inert to complete exchange with excess cadmium, which characterises it as being a Zn-thionein (ie preference for divalent metal ions [214]). Considering the significant spectral differences indicated by comparison of the TOCSY spectra of Zn- and Cd- form, wMT-2 may be characterised specifically as being a Cd-thionein.

The second explanation is that the cell controls selection of metals that can be bound [247]. This is the case for the metal-binding protein MncA, which binds manganese over the more competitive copper [247, 265]. In the case of MncA, folding in the cytoplasm provides an environment where comparatively high concentrations of manganese are present

[265]. Therefore it may be that differential expression causes the apparent metal-binding preferences of MT *in vivo*, governing which metals are found bound to MT when isolated from native sources. However as all wMTs in this thesis were obtained from recombinant expression, it is difficult to speculate if this may be case.

In Chapter 5, wMT-2 was determined to have a higher apparent metal-binding affinity for cadmium than zinc. However the value was somewhat lower than that of vertebrate MT [200]. It would be interesting to see how the apparent metal-binding affinities of wMT-1, the postulated zinc-handling protein [112], compared to wMT-2. These two sequences have over 70 % identity, yet appear to have different metal-binding properties (**Figures 5.1 & 5.2**). Comparing the properties in a more comprehensive manner may lead to new insights into the specific function of the isoforms within the earthworm.

6.4 Further avenues for research

The proposed structure of wMT-2 is presented within this thesis. However, for unambiguous metal connectivities, 2-dimensional heteronuclear labelled-cadmium experiments are required, as per Frey *et al*. [93]. This will remove any ambiguity from the assigned connectivities, and increase the confidence in the model proposed in this thesis. Although reconstitution of a sample with ^{111}Cd was attempted twice during the course of the project, significant protein losses were incurred

after demetallation in both cases. In the future, performing the entire experiment with rigorous exclusion of oxygen may be more successful than elution under a constant flow of nitrogen. If losses of protein are unavoidable however, starting the reconstitution procedure with a larger quantity of protein (> 1mg) may generate an acceptable spectrum.

The majority of the work in this thesis has focussed on only a part of the three wMT system in earthworms: wMT-2. Therefore as further work, it would be interesting to compare the metal-binding properties of wMT-1 and wMT-2 in more detail. The results may help to answer such questions as: 'what makes the two isoforms different?', and 'how do the differences in metal-binding properties between wMT-1 and wMT-2 effect their function in adult earthworms?'.

Mutation studies involving the linker region between the two domains may provide further insight into the metal ion preferences of wMTs. The longer linker region in wMT-1 may be crucial for this role. Therefore it may be possible to emulate the metal-binding properties of wMT-2 with the creation of a wMT-1 deletion mutant, removing 2 residues in the linker region.

Splitting wMT-2 into two domains either by cloning, or by limited proteolysis, may provide information about domain-specific metal-binding characteristics. Comparisons to *Eisenia foetida* MT may then be

possible, which exists as a stable α-domain (M_4Cys_{11}- cluster), and is similarly induced by cadmium [166]. It is also within Domain 1 that wMT-1 and wMT-2 show higher sequence disparity, so it may be that the metal-binding preferences of the M_4Cys_{11}-cluster in wMTs are significantly more diverse than that of the M_3Cys_9-cluster. Isolation of the individual domains would also enable comparison with full-length protein, and determine if there are inter-domain effects influencing the characteristics of wMTs, similar to that observed for vertebrate MT [266].

Additionally, structural determination of wMT-3 may reveal a novel fold, as although the cysteine residues are highly conserved in the three wMTs, there is lower sequence identity between wMT-3 and wMT-1/wMT-2. As it is present in earthworm cocoons [112], wMT-3 is expected to be functionally distinct from wMT-1 and wMT-2. Therefore a structural investigation and comprehensive metal-binding study may shed new light on the function of wMT-3.

6.5 Conclusion

In conclusion, it is believed that the investigations presented in this thesis have furthered the understanding of wMT-2, a member of the earthworm MT family within *Lumbricus rubellus*. Furthermore, data have been presented for the first MT 3-dimensional structure from the phylum annelida. The best outcome for this research is that it will pave the way

for further work with wMTs, ultimately leading to the determination of a structure/function relationship for all three earthworm metallothioneins.

References

1. Dawkins, R., *The Greatest Show on Earth: The Evidence for Evolution*. First Edition ed. 2009, London, UK: Bantam Press.
2. Lenski, R.E., *Evolution in action: a 50,000-generation salute to Charles Darwin.* Microbe, 2011. **6**: p. 30-33.
3. Begon, M., Townsend, C.R., and Harper, J.L., *Ecology: From Individuals to Ecosystems*. 4th Edition ed. 2006, Hoboken, New Jersey: John Wiley & Sons.
4. Sulloway, F.J., "Darwin and His Finches: The Evolution of a Legend." *Journal of the History of Biology*, 15 (1982):1-53. Reprinted in *Darwin's Finches: Readings in the Evolution of a Scientific Paradigm*, edited by Kathleen Donohue (Chicago: University of Chicago Press, 2011, pp. 44-97).
5. Hengeveld, R. and Fedonkin, M.A., *Bootstrapping the energy flow in the beginning of life.* Acta Biotheor., 2007. **55**(2): p. 181-226.
6. Burtis, C.A., Ashwood, E.R., Border, B., and Tietz, N.W., *Tietz Fundamentals of Clinical Chemistry, Chapter 29*. 5th ed. 2001, Philadelphia: W.B. Saunders. xxv, 1091 p.
7. Schroeder, H.A., Frost, D.V., and Balassa, J.J., *Essential trace metals in man: selenium.* J. Chronic Dis., 1970. **23**(4): p. 227-243.
8. Nielsen, F.H., *Evolutionary events culminating in specific minerals becoming essential for life.* Eur. J. Nutr., 2000. **39**(2): p. 62-66.
9. Kabata-Pendias, A., *Trace Elements in Soils and Plants*. Third Edition ed. 2000: CRC Press.
10. McCall, K.A., Huang, C., and Fierke, C.A., *Function and mechanism of zinc metalloenzymes.* J Nutr, 2000. **130** (5S Suppl): p. 1437S-1446S.
11. Matthews, R.G. and Goulding, C.W., *Enzyme-catalyzed methyl transfers to thiols: the role of zinc.* Curr. Opin. Chem. Biol., 1997. **1**(3): p. 332-339.
12. Chowdhury, B.A. and Chandra, R.K., *Biological and health implications of toxic heavy metal and essential trace element interactions.* Prog. Food Nutr. Sci., 1987. **11**(1): p. 55-113.
13. Ellwood, M.J. and Van den Berg, C.M.G., *Zinc speciation in the Northeastern Atlantic Ocean.* Mar. Chem., 2001. **68**: p. 295–306.
14. Lane, T.W., Saito, M.A., George, G.N., Pickering, I.J., Prince, R.C., and Morel, F.M., *Biochemistry: a cadmium enzyme from a marine diatom.* Nature, 2005. **435**: p. 42.
15. Coombs, J.M. and Barkay, T., *New findings on evolution of metal homeostasis genes: evidence from comparative genome analysis of bacteria and archaea.* Appl. Environ. Microbiol., 2005. **71**(11): p. 7083-7091.
16. Williams, R.J.P., *My past and a future role for inorganic biochemistry.* J Inorg Biochem, 2006. **100**(12): p. 1908-1924.
17. Jørgensen, S.E., Patten, B.C. and Straškraba, M., *Ecosystems emerging: toward an ecology of complex systems in a complex future.* Ecol. Model., 1992. **62**(1-3): p. 1-27.
18. Waldron, K.J., Rutherford, J.C., Ford, D., and Robinson, N.J., *Metalloproteins and metal sensing.* Nature, 2009. **460**(7257): p. 823-830.

19. Shibamoto, T. and Bjeldanes, L.F., *Introduction to Food Toxicology, Chapter 1: Principles of Toxicology*. 2nd ed. Food science and technology international series. 2009, Amsterdam ; Boston: Academic Press. x, 309 p.
20. Holsapple, M.P. and Wallace, K.B., *Dose response considerations in risk assessment--an overview of recent ILSI activities.* Toxicol Lett, 2008. **180**(2): p. 85-92.
21. Williams, R.J.P. and Frausto Da Silva, J.J.R., *The distribution of elements in cells.* Coord. Chem. Rev., 2000. **200-202**: p. 247-348.
22. Nelson, N., *Metal ion transporters and homeostasis.* EMBO J., 1999. **18**(16): p. 4361-4371.
23. Halliwell, B., *Reactive species and antioxidants. Redox biology is a fundamental theme of aerobic life.* Plant Physiol., 2006. **141**(2): p. 312-322.
24. Waldron, K.J. and Robinson, N.J., *How do bacterial cells ensure that metalloproteins get the correct metal?* Nat. Rev. Microbiol., 2009. **7**(1): p. 25-35.
25. Adlard, P.A. and Bush, A.I., *Metals and Alzheimer's disease.* J. Alzheimers Dis., 2006. **10**(2-3): p. 145-163.
26. Finney, L.A. and O'Halloran, T.V., *Transition metal speciation in the cell: insights from the chemistry of metal ion receptors.* Science, 2003. **300**(5621): p. 931-936.
27. Fosmire, G.J., *Zinc toxicity.* Am. J. Clin. Nutr., 1990. **51**(2): p. 225-227.
28. Beach, R.S., Gershwin, M.E., and Hurley, L.S., *Gestational zinc deprivation in mice: persistence of immunodeficiency for three generations.* Science, 1982. **218**(4571): p. 469-471.
29. Huang, M., Choi, S.J., Kim, D.W., Kim, N.Y., Park, C.H., Yu, S.D., Kim, D.S., Park, K.S., Song, J.S*., et al.*, *Risk assessment of low-level cadmium and arsenic on the kidney.* J. Toxicol. Environ. Health A, 2009. **72**(21-22): p. 1493-1498.
30. Moulis, J.M., *Cellular mechanisms of cadmium toxicity related to the homeostasis of essential metals.* BioMetals, 2010. **23**(5): p. 877-896.
31. Goyer, R.A., *Toxic and essential metal interactions.* Annu. Rev. Nutr., 1997. **17**: p. 37-50.
32. Kjellström, T., *Mechanism and epidemiology of bone effects of cadmium.* IARC Sci. Publ., 1992(118): p. 301-310.
33. Hood, E., *Putting a load on your bones: low-level cadmium exposure and osteoporosis.* Environ Health Perspect, 2006. **114**(6): p. 369-370.
34. Funk, A.E., Day, F.A., and Brady, F.O., *Displacement of zinc and copper from copper-induced metallothionein by cadmium and by mercury:* in vivo *and* ex vivo *studies.* Comp. Biochem. Physiol. C, 1987. **86**(1): p. 1-6.
35. Day, F.A., Funk, A.E., and Brady, F.O., In vivo *and* ex vivo *displacement of zinc from metallothionein by cadmium and by mercury.* Chem.-Biol. Interact., 1984. **50**(2): p. 159-174.
36. Hartwig, A., *Zinc Finger Proteins as Potential Targets for Toxic Metal Ions: Differential Effects on Structure and Function* Antioxid. Redox Signaling, 2001. **3**(4): p. 625-634.
37. Mead, M.N., *Cadmium confusion: do consumers need protection?* Environ Health Perspect, 2010. **118**(12): p. a528-534.
38. Auld, D.S., *Zinc coordination sphere in biochemical zinc sites.* BioMetals, 2001. **14**(3-4): p. 271-313.
39. Greenberg, S.A. and Briemberg, H.R., *A neurological and hematological syndrome associated with zinc excess and copper deficiency.* J. Neurol., 2004. **251**(1): p. 111-114.
40. Hall, J.L., *Cellular mechanisms for heavy metal detoxification and tolerance.* J. Exp. Bot., 2002. **53**(366): p. 1-11.

41. Eide, D.J., *The molecular biology of metal ion transport in* Saccharomyces cerevisiae*.* Annu. Rev. Nutr., 1998. **18**: p. 441-469.
42. Gadd, G.M. and Griffiths, A.J., *Microorganisms and Heavy Metal Toxicity.* Microb. Ecol. Health Dis., 1978. **4**: p. 303-317.
43. García-Giménez, E., Alcaraz, A., and Aguilella, V.M., *Divalent Metal Ion Transport across Large Biological Ion Channels and Their Effect on Conductance and Selectivity.* Biochemistry Research International, 2012. **2012**: p. 1-12.
44. Liberman, E.A. and Topaly, V.P., *Selective transport of ions through bimolecular phospholipid membranes.* Biochim Biophys Acta, 1968. **163**(2): p. 125-136.
45. Colvin, R.A., Davis, N., Nipper, R.W., and Carter, P.A., *Zinc transport in the brain: routes of zinc influx and efflux in neurons.* J Nutr, 2000. **130**(5S Suppl): p. 1484S-1487S.
46. John, S.G., Geis, R.W., Saito, M.A., and Boyle, E.A., *Zinc isotope fractionation during high-affinity and low-affinity Zn transport in* T. oceanica*.* Limnol. Oceanogr., 2007. **52**: p. 2710-2714.
47. Zhao, H. and Eide, D., *The yeast* ZRT1 *gene encodes the zinc transporter protein of a high-affinity uptake system induced by zinc limitation.* Proc Natl Acad Sci U S A, 1996. **93**(6): p. 2454-2458.
48. Zhao, H. and Eide, D., *The* ZRT2 *gene encodes the low affinity zinc transporter in* Saccharomyces cerevisiae*.* J Biol Chem, 1996. **271**(38): p. 23203-23210.
49. Eide, D., *Molecular biology of iron and zinc uptake in eukaryotes.* Curr. Opin. Cell Biol., 1997. **9**(4): p. 573-577.
50. Rink, L. and Haase, H., *Zinc homeostasis and immunity.* Trends Immunol., 2007. **28**(1): p. 1-4.
51. Suhy, D.A. and O'Halloran, T.V., *Metal-responsive gene regulation and the zinc metalloregulatory model.* Met. Ions Biol. Syst., 1996. **32**: p. 557-578.
52. Gaither, L.A. and Eide, D.J., *Eukaryotic zinc transporters and their regulation.* BioMetals, 2001. **14**(3-4): p. 251-270.
53. Kramer, U., Talke, I.N., and Hanikenne, M., *Transition metal transport.* FEBS Lett., 2007. **581**(12): p. 2263-2272.
54. Rosenzweig, A.C., *Metallochaperones: bind and deliver.* Chem. Biol., 2002. **9**(6): p. 673-677.
55. O'Halloran, T.V. and Culotta, V.C., *Metallochaperones, an intracellular shuttle service for metal ions.* J Biol Chem, 2000. **275**(33): p. 25057-25060.
56. Bancia, L., Bertinia, I., Cantinia, F., Kozyrevaa, T., Massagnia, C., Palumaad, P., Rubinoa, J.T., and Zovod, K., *Human superoxide dismutase 1 (hSOD1) maturation through interaction with human copper chaperone for SOD1 (hCCS).* Proc Natl Acad Sci U S A, 2012. **109**(34): p. 13555-13560.
57. Gaetke, L.M. and Chow, C.K., *Copper toxicity, oxidative stress, and antioxidant nutrients.* Toxicology, 2003. **189**(1-2): p. 147-163.
58. Canli, M., Stagg, R.M., and Rodger, G., *The induction of metallothionein in tissues of the Norway lobster* Nephrops norvegicus *following exposure to cadmium, copper and zinc: the relationships between metallothionein and the metals.* Environ. Pollut., 1997. **96**(3): p. 343-350.
59. Andrews, G.K., *Regulation of metallothionein gene expression by oxidative stress and metal ions.* Biochem Pharmacol, 2000. **59**(1): p. 95-104.
60. Burgos, M.G., Winters, C., Stürzenbaum, S.R., Randerson, P.F., Kille, P., and Morgan, A.J., *Cu and Cd effects on the earthworm* Lumbricus rubellus *in the laboratory: Multivariate statistical analysis of relationships between exposure,*

biomarkers, and ecologically relevant parameters. Environ. Sci. Technol., 2005. **39**(6): p. 1757-1763.

61. Haq, F., Mahoney, M., and Koropatnick, J., *Signaling events for metallothionein induction.* Mutat. Res., 2003. **533**(1-2): p. 211-226.
62. Blindauer, C.A. and Leszczyszyn, O.I., *Metallothioneins: unparalleled diversity in structures and functions for metal ion homeostasis and more.* Nat. Prod. Rep., 2010. **27**(5): p. 720-741.
63. Cobine, P.A., McKay, R.T., Zangger, K., Dameron, C.T., and Armitage, I.M., *Solution structure of Cu_6 metallothionein from the fungus* Neurospora crassa*.* Eur J Biochem, 2004. **271**(21): p. 4213-4221.
64. Banerjee, D., Onosaka, S., and Cherian, M.G., *Immunohistochemical localization of metallothionein in cell nucleus and cytoplasm of rat liver and kidney.* Toxicology, 1982. **24**(2): p. 95-105.
65. Wolfa, C., Strenziokb, R., and Kyriakopoulosa, A., *Elevated metallothionein-bound cadmium concentrations in urine from bladder carcinoma patients, investigated by size exclusion chromatography-inductively coupled plasma mass spectrometry.* Analytica Chimica Acta, 2009. **631**(2): p. 218-222.
66. Bremner, I. and Beattie, J.H., *Metallothionein and the trace minerals.* Annu. Rev. Nutr., 1990. **10**: p. 63-83.
67. Klaassen, C.D., Liu, J., and Choudhuri, S., *Metallothionein: an intracellular protein to protect against cadmium toxicity.* Annu. Rev. Pharmacol. Toxicol., 1999. **39**: p. 267-294.
68. Nordberg, M. and Nordberg, G.F., *Metallothioneins: Historical Development and Overview*, in *Metallothioneins and Related Chelators*, Astrid Sigel, Helmut Sigel, and Sigel, R.K.O., Editors. 2009, Royal Society of Chemistry: London, UK.
69. Palmiter, R.D., *The elusive function of metallothioneins.* Proc. Natl. Acad. Sci. U. S. A., 1998. **95**(15): p. 8428-8430.
70. Nishimura, N., Nishimura, H., Ghaffar, A., and Tohyama, C., *Localization of metallothionein in the brain of rat and mouse.* J. Histochem. Cytochem., 1992. **40**(2): p. 309-315.
71. Chung, R.S. and West, A.K., *A role for extracellular metallothioneins in CNS injury and repair.* Neuroscience, 2004. **123**(3): p. 595-599.
72. Trendelenburg, G., Prass, K., Priller, J., Kapinya, K., Polley, A., Muselmann, C., Ruscher, K., Kannbley, U., Schmitt, A.O*., et al.*, *Serial analysis of gene expression identifies* metallothionein-II *as major neuroprotective gene in mouse focal cerebral ischemia.* J. Neurosci., 2002. **22**(14): p. 5879-5888.
73. Hidalgo, J., Penkowa, M., Espejo, C., Martinez-Caceres, E.M., Carrasco, J., Quintana, A., Molinero, A., Florit, S., Giralt, M*., et al.*, *Expression of metallothionein-I, -II, and -III in Alzheimer disease and animal models of neuroinflammation.* Exp. Biol. Med., 2006. **231**(9): p. 1450-1458.
74. Duguid, J.R., Bohmont, C.W., Liu, N.G., and Tourtellotte, W.W., *Changes in brain gene expression shared by scrapie and Alzheimer disease.* Proc Natl Acad Sci U S A, 1989. **86**(18): p. 7260-7264.
75. Zambenedetti, P., Giordano, R., and Zatta, P., *Metallothioneins are highly expressed in astrocytes and microcapillaries in Alzheimer's disease.* J. Chem. Neuroanat., 1998. **15**(1): p. 21-26.
76. Kawashima, T., Doh-Ura, K., Torisu, M., Uchida, Y., Furuta, A., and Iwaki, T., *Differential expression of metallothioneins in human prion diseases.* Dement. Geriatr. Cogn. Disord., 2000. **11**(5): p. 251-262.

77. Penkowa, M., Espejo, C., Ortega-Aznar, A., Hidalgo, J., Montalban, X., and Martinez Caceres, E.M., *Metallothionein expression in the central nervous system of multiple sclerosis patients.* Cell. Mol. Life Sci., 2003. **60**(6): p. 1258-1266.
78. Pedersen, M.O., Larsen, A., Stoltenberg, M., and Penkowa, M., *The role of metallothionein in oncogenesis and cancer prognosis.* Prog. Histochem. Cytochem., 2009. **44**(1): p. 29-64.
79. Margoshes, M. and Vallee, B.L., *A Cadmium Protein from Equine Kidney Cortex.* J. Am. Chem. Soc., 1957. **79**(17): p. 4813-4814.
80. Tran, C.D., Cool, J., and Xian, C.J., *Dietary zinc and metallothionein on small intestinal disaccharidases activity in mice.* World J Gastroenterol, 2011. **17**(3): p. 354-360.
81. Kägi, J.H. and Valee, B.L., *Metallothionein: a cadmium- and zinc-containing protein from equine renal cortex.* J Biol Chem, 1960. **235**: p. 3460-3465.
82. Hamer, D.H., *Metallothionein.* Annu. Rev. Biochem., 1986. **55**: p. 913-951.
83. Kägi, J.H.R. and Nordberg, M. *Metallothionein*. in *International Meeting on Metallothionein and Other Low Molecular Weight Metal-binding Proteins*. 1979. Zurich, Switzerland: Basel ; Boston : Birkhäuser.
84. Fowler, B.A., Hildebrand, C.E., Kojima, Y., and Webb, M., *Nomenclature of metallothionein.* Experientia Suppl, 1987. **52**: p. 19-22.
85. Kägi, J.H.R. and Binz, P.A. *Metallothionein: Molecular evolution and classification*. in *Metallothionein 4*, Klaasen, C.D., Editor. 1999, Birkhauser Verlag AG: Basel, Switzerland.
86. Kojima, Y., Binz, P.A., and Kägi, J.H., *Nomenclature of metallothionein: Proposal for a revision*, in *Metallothionein 4*, Klaasen, C.D., Editor. 1999, Birkhauser Verlag AG: Basel, Switzerland.
87. Riek, R., Prêcheur, B., Wang, Y., Mackay, E.A., Wider, G., Güntert, P., Liu, A., Kägi, J.H., and Wüthrich, K., *NMR structure of the sea urchin (*Strongylocentrotus purpuratus*) metallothionein MTA.* J Mol Biol, 1999. **291**(2): p. 417-428.
88. Capasso, C., Carginale, V., Crescenzi, O., Di Maro, D., Parisi, E., Spadaccini, R., and Temussi, P.A., *Solution structure of MT_nc, a novel metallothionein from the Antarctic fish* Notothenia coriiceps*.* Structure, 2003. **11**(4): p. 435-443.
89. Romero-Isart, N. and Vašák, M., *Advances in the structure and chemistry of metallothioneins.* J Inorg Biochem, 2002. **88**(3-4): p. 388-396.
90. Gutierrez, J.C., Amaro, F., and Martin-Gonzalez, A., *From heavy metal-binders to biosensors: ciliate metallothioneins discussed.* BioEssays, 2009. **31**(7): p. 805-816.
91. Klaassen, C.D., *Metallothionein IV*. 1999, Basel ; Boston: Birkhäuser Verlag. xlvi, 645 p.
92. Furey, W.F., Robbins, A.H., Clancy, L.L., Winge, D.R., Wang, B.C., and Stout, C.D., *Crystal structure of Cd,Zn metallothionein.* Experientia Suppl, 1987. **52**: p. 139-148.
93. Frey, M.H., Wagner, G., Vašák, M., Sorensen, O.W., Neuhaus, D., Worgotter, E., Kägi, J.H., Ernst, R.R., and Wüthrich, K., *Polypeptide metal cluster connectivities in metallothionein-2 by novel ^{1}H-^{113}Cd heteronuclear two-dimensional NMR experiments.* J Am Chem Soc, 1985. **107**: p. 6394-6396.
94. Vašák, M., Galdes, A., Hill, H.A., Kägi, J.H., Bremner, I., and Young, B.W., *Investigation of the structure of metallothioneins by proton nuclear magnetic resonance spectroscopy.* Biochemistry, 1980. **19**(3): p. 416-425.
95. Hunt, C.T., Boulanger, Y., Fesik, S.W., and Armitage, I.M., *NMR analysis of the structure and metal sequestering properties of metallothioneins.* Environ Health Perspect, 1984. **54**: p. 135-145.
96. Braun, W., Vašák, M., Robbins, A.H., Stout, C.D., Wagner, G., Kägi, J.H., and Wüthrich, K., *Comparison of the NMR solution structure and the x-ray crystal*

structure of rat metallothionein-2. Proc Natl Acad Sci U S A, 1992. **89**(21): p. 10124-10128.

97. Armitage, I.M., Otvos, J.D., Briggs, R.W., and Boulanger, Y., *Structure elucidation of the metal-binding sites in metallothionein by ^{113}Cd NMR.* Fed. Proc., 1982. **41**(13): p. 2974-2980.
98. Stürzenbaum, S.R., Kille, P., and Morgan, A.J., *The identification, cloning and characterization of earthworm metallothionein.* FEBS Lett., 1998. **431**(3): p. 437-442.
99. Otvos, J.D., Olafson, R.W., and Armitage, I.M., *Structure of an invertebrate metallothionein from* Scylla serrata*.* J Biol Chem, 1982. **257**(5): p. 2427-2431.
100. Otvos, J.D. and Armitage, I.M., *Structure of the metal clusters in rabbit liver metallothionein.* Proc Natl Acad Sci U S A, 1980. **77**(12): p. 7094-7098.
101. Boulanger, Y., Goodman, C.M., Forte, C.P., Fesik, S.W., and Armitage, I.M., *Model for mammalian metallothionein structure.* Proc Natl Acad Sci U S A, 1983. **80**(6): p. 1501-1505.
102. Byrd, J. and Winge, D.R., *Cooperative cluster formation in metallothionein.* Arch. Biochem. Biophys., 1986. **250**(1): p. 233-237.
103. Nielson, K.B. and Winge, D.R., *Order of metal binding in metallothionein.* J Biol Chem, 1983. **258**(21): p. 13063-13069.
104. Lee, Y.H., Shaikh, Z.A., and Tohyama, C., *Urinary metallothionein and tissue metal levels of rats injected with cadmium, mercury, lead, copper or zinc.* Toxicology, 1983. **27**(3-4): p. 337-345.
105. Amiard, J.C., Amiard-Triquet, C., Barka, S., Pellerin, J., and Rainbow, P.S., *Metallothioneins in aquatic invertebrates: their role in metal detoxification and their use as biomarkers.* Aquat. Toxicol., 2006. **76**(2): p. 160-202.
106. Stuart, G.W., Searle, P.F., and Palmiter, R.D., *Identification of multiple metal regulatory elements in mouse metallothionein-I promoter by assaying synthetic sequences.* Nature, 1985. **317**(6040): p. 828-831.
107. Dalton, T., Palmiter, R.D., and Andrews, G.K., *Transcriptional induction of the mouse* metallothionein-I *gene in hydrogen peroxide-treated Hepa cells involves a composite major late transcription factor/antioxidant response element and metal response promoter elements.* Nucleic Acids Res, 1994. **22**(23): p. 5016-5023.
108. Radtke, F., Heuchel, R., Georgiev, O., Hergersberg, M., Gariglio, M., Dembic, Z., and Schaffner, W., *Cloned transcription factor MTF-1 activates the mouse metallothionein I promoter.* EMBO J., 1993. **12**(4): p. 1355-1362.
109. Palmiter, R.D., *Regulation of metallothionein genes by heavy metals appears to be mediated by a zinc-sensitive inhibitor that interacts with a constitutively active transcription factor, MTF-1.* Proc Natl Acad Sci U S A, 1994. **91**(4): p. 1219-1223.
110. Janssens, T.K.S., Roelofs, D., and Van Straalen, N.M., *Molecular mechanisms of heavy metal tolerance and evolution in invertebrates.* Insect Sci., 2009. **16**(1): p. 3-18.
111. Barnes, R.S.K., Calow, P.P., Olive, P.J.W., Golding, D.W., and Spicer, J.I., *The Invertebrates: A Synthesis*. 3rd Edition ed. 2001: Wiley-Blackwell.
112. Stürzenbaum, S.R., Georgiev, O., Morgan, A.J., and Kille, P., *Cadmium detoxification in earthworms: From genes to cells.* Environ. Sci. Technol., 2004. **38**(23): p. 6283-6289.
113. Stürzenbaum, S.R., Winters, C., Galay, M., Morgan, A.J., and Kille, P., *Metal ion trafficking in earthworms - Identification of a cadmium-specific metallothionein.* J. Biol. Chem., 2001. **276**(36): p. 34013-34018.

114. Morgan, A.J., Stürzenbaum, S.R., Winters, C., Grime, G.W., Abd Aziz, N.A., and Kille, P., *Differential metallothionein expression in earthworm (*Lumbricus rubellus*) tissues.* Ecotoxicol. Environ. Saf., 2004. **57**(1): p. 11-19.
115. Klerks, P.L. and Bartholomew, P.R., *Cadmium accumulation and detoxification in a Cd-resistant population of the oligochaete Limnodrilus hoffmeisteri.* Aquat. Toxicol., 1991. **19**: p. 97-112.
116. Klerks, P.L. and Levinton, J.S., *Rapid evolution of metal resistance in a benthic oligochaete inhabiting a metal-polluted site.* Biol. Bull., 1989. **176**: p. 135-141.
117. Dallinger, R., Berger, B., Hunziker, P., and Kägi, J.H., *Metallothionein in snail Cd and Cu metabolism.* Nature, 1997. **388**(6639): p. 237-238.
118. Berger, B., Hunziker, P.E., Hauer, C.R., Birchler, N., and Dallinger, R., *Mass spectrometry and amino acid sequencing of two cadmium-binding metallothionein isoforms from the terrestrial gastropod* Arianta arbustorum*.* Biochem J, 1995. **311**: p. 951-957.
119. Mackay, E.A., Overnell, J., Dunbar, B., Davidson, I., Hunziker, P.E., Kägi, J.H., and Fothergill, J.E., *Complete amino acid sequences of five dimeric and four monomeric forms of metallothionein from the edible mussel* Mytilus edulis*.* Eur J Biochem, 1993. **218**(1): p. 183-194.
120. Roesijadi, G., Kielland, S., and Klerks, P., *Purification and properties of novel molluscan metallothioneins.* Arch. Biochem. Biophys., 1989. **273**(2): p. 403-413.
121. Hockner, M., Dallinger, R., and Stürzenbaum, S.R., *Nematode and snail metallothioneins.* J Biol Inorg Chem, 2011. **16**(7): p. 1057-1065.
122. Palacios, O., Pagani, A., Perez-Rafael, S., Egg, M., Hockner, M., Brandstatter, A., Capdevila, M., Atrian, S., and Dallinger, R., *Shaping mechanisms of metal specificity in a family of metazoan metallothioneins: evolutionary differentiation of mollusc metallothioneins.* BMC Biol., 2011. **9**: p. 4.
123. Leignel, V. and Laulier, M., *Isolation and characterization of* Mytilus edulis *metallothionein genes.* Comp. Biochem. Physiol. C Toxicol. Pharmacol., 2006. **142**(1-2): p. 12-18.
124. Vergani, L., Grattarola, M., Borghi, C., Dondero, F., and Viarengo, A., *Fish and molluscan metallothioneins: structural and functional comparison.* FEBS J, 2005. **272**(23): p. 6014-6023.
125. Grattarola, M., Carloni, M., Dondero, F., Viarengo, A., and Vergani, L., *Expression, purification and preliminary characterization of mussel (*Mytilus galloprovincialis*) metallothionein MT20.* Mol. Biol. Rep., 2006. **33**(4): p. 265-272.
126. Orihuela, R., Domenech, J., Bofill, R., You, C., Mackay, E.A., Kägi, J.H., Capdevila, M., and Atrian, S., *The metal-binding features of the recombinant mussel* Mytilus edulis *MT-10-IV metallothionein.* J Biol Inorg Chem, 2008. **13**(5): p. 801-812.
127. Ivankovic, D., Pavicic, J., Kozar, S., and Raspor, B., *Multiple forms of metallothionein from the digestive gland of naturally occurring and cadmium-exposed mussels, Mytilus galloprovincialis.* Helgol. Mar. Res., 2002. **56**: p. 95-101.
128. Zrzavy, J. and Stys, P., *The basic body plan of arthropods: insights from evolutionary morphology and developmental biology.* J. Evol. Biol., 1997. **10**(3): p. 353-367.
129. Brouwer, M., Winge, D.R., and Gray, W.R., *Structural and functional diversity of copper-metallothioneins from the American lobster* Homarus americanus*.* J Inorg Biochem, 1989. **35**(4): p. 289-303.
130. Brouwer, M., Enghild, J., Hoexum-Brouwer, T., Thogersen, I., and Truncali, A., *Primary structure and tissue-specific expression of blue crab (*Callinectes sapidus*) metallothionein isoforms.* Biochem J, 1995. **311 (Pt 2)**: p. 617-622.

131. Silar, P., Theodore, L., Mokdad, R., Erraiss, N.E., Cadic, A., and Wegnez, M., *Metallothionein* Mto *gene of* Drosophila melanogaster*: structure and regulation.* J Mol Biol, 1990. **215**(2): p. 217-224.
132. Hensbergen, P.J., Donker, M.H., Hunziker, P.E., Van Der Schors, R.C., and Van Straalen, N.M., *Two metal-binding peptides from the insect* Orchesella cincta *(Collembola) as a result of metallothionein cleavage.* Insect Biochem. Mol. Biol., 2001. **31**(11): p. 1105-1114.
133. Munoz, A., Forsterling, F.H., Shaw, C.F., 3rd, and Petering, D.H., *Structure of the* $^{113}Cd_3\beta$ *domains from* Homarus americanus *metallothionein-1: hydrogen bonding and solvent accessibility of sulfur atoms.* J Biol Inorg Chem, 2002. **7**(7-8): p. 713-724.
134. Valls, M., Bofill, R., Gonzalez-Duarte, R., Gonzalez-Duarte, P., Capdevila, M., and Atrian, S., *A new insight into metallothionein (MT) classification and evolution - The in vivo and in vitro metal binding features of* Homarus americanus *recombinant MT.* J. Biol. Chem., 2001. **276**(35): p. 32835-32843.
135. Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A*., et al.*, *The genome sequence of* Drosophila melanogaster*.* Science, 2000. **287**(5461): p. 2185-2195.
136. Maroni, G., Ho, A.S., and Theodore, L., *Genetic control of cadmium tolerance in* Drosophila melanogaste*r.* Environ Health Perspect, 1995. **103**(12): p. 1116-1118.
137. Perez-Rafael, S., Kurz, A., Guirola, M., Capdevila, M., Palacios, O., and Atrian, S., *Is MtnE, the fifth Drosophila metallothionein, functionally distinct from the other members of this polymorphic protein family?* Metallomics, 2012. **4**(4): p. 342-349.
138. Lastowski-Perry, D., Otto, E., and Maroni, G., *Nucleotide sequence and expression of a Drosophila metallothionein.* J Biol Chem, 1985. **260**(3): p. 1527-1530.
139. Egli, D., Domenech, J., Selvaraj, A., Balamurugan, K., Hua, H., Capdevila, M., Georgiev, O., Schaffner, W., and Atrian, S., *The four members of the Drosophila metallothionein family exhibit distinct yet overlapping roles in heavy metal homeostasis and detoxification.* Genes Cells, 2006. **11**(6): p. 647-658.
140. Atrian, S., *Chapter 6: Metallothioneins in Diptera*, in *Metallothioneins and Related Chelators*, Sigel, A., Sigel, H., and Sigel, R.K.O., Editors. 2009, RSC Pub.: Cambridge, UK. p. 514 p.
141. Egli, D., Selvaraj, A., Yepiskoposyan, H., Zhang, B., Hafen, E., Georgiev, O., and Schaffner, W., *Knockout of 'metal-responsive transcription factor' MTF-1 in Drosophila by homologous recombination reveals its central role in heavy metal homeostasis.* EMBO J., 2003. **22**(1): p. 100-108.
142. Wang, Y., Mackay, E.A., Kurasaki, M., and Kägi, J.H., *Purification and characterisation of recombinant sea urchin metallothionein expressed in Escherichia coli.* Eur J Biochem, 1994. **225**(1): p. 449-457.
143. Wilkinson, D.G. and Nemer, M., *Metallothionein genes* MTa *and* MTb *expressed under distinct quantitative and tissue-specific regulation in sea urchin embryos.* Mol. Cell. Biol., 1987. **7**(1): p. 48-58.
144. Messerle, B.A., Schaffer, A., Vašák, M., Kägi, J.H., and Wüthrich, K., *Three-dimensional structure of human* $^{113}Cd_7$ *metallothionein-2 in solution determined by nuclear magnetic resonance spectroscopy.* J Mol Biol, 1990. **214**(3): p. 765-779.
145. Nemer, M., Wilkinson, D.G., Travaglini, E.C., Sternberg, E.J., and Butt, T.R., *Sea urchin metallothionein sequence: key to an evolutionary diversity.* Proc Natl Acad Sci U S A, 1985. **82**(15): p. 4992-4994.
146. Leszczyszyn, O.I., Zeitoun-Ghandour, S., Stürzenbaum, S.R., and Blindauer, C.A., *Tools for metal ion sorting: in vitro evidence for partitioning of zinc and cadmium in* C. elegans *metallothionein isoforms.* Chem. Commun., 2010. **47**(1): p. 448-450.

147. Slice, L.W., Freedman, J.H., and Rubin, C.S., *Purification, characterization, and cDNA cloning of a novel metallothionein-like, cadmium-binding protein from* Caenorhabditis elegans*.* J Biol Chem, 1990. **265**(1): p. 256-263.
148. Imagawa, M., Onozawa, T., Okumura, K., Osada, S., Nishihara, T., and Kondo, M., *Characterization of metallothionein cDNAs induced by cadmium in the nematode* Caenorhabditis elegans*.* Biochem J, 1990. **268**(1): p. 237-240.
149. Bofill, R., Orihuela, R., Romagosa, M., Domenech, J., Atrian, S., and Capdevila, M., Caenorhabditis elegans *metallothionein isoform specificity - metal binding abilities and the role of histidine in CeMT1 and CeMT2.* FEBS J., 2009. **276**(23): p. 7040-7069.
150. Krizek, B.A., Merkle, D.L., and Berg, J.M., *Ligand variation and metal ion binding specificity in zinc finger peptides.* Inorg Chem, 1993. **32**(6): p. 937-940.
151. Zeitoun-Ghandour, S., Charnock, J.M., Hodson, M.E., Leszczyszyn, O.I., Blindauer, C.A., and Stürzenbaum, S.R., *The two* Caenorhabditis elegans *metallothioneins (CeMT1 and CeMT2) discriminate between essential zinc and toxic cadmium.* FEBS J., 2011. **277**(11): p. 2531-2542.
152. Jiang, G.C., Hughes, S., Stürzenbaum, S.R., Evje, L., Syversen, T., and Aschner, M., Caenorhabditis elegans *metallothioneins protect against toxicity induced by depleted uranium.* Toxicol. Sci., 2009. **111**(2): p. 345-354.
153. Hughes, S. and Stürzenbaum, S.R., *Single and double metallothionein knockout in the nematode* C. elegans *reveals cadmium dependent and independent toxic effects on life history traits.* Environ. Pollut., 2007. **145**(2): p. 395-400.
154. Burgos, M.G., Winters, C., Stürzenbaum, S.R., Randerson, P.F., Kille, P., and Morgan, A.J., *Cu and Cd effects on the earthworm* Lumbricus rubellus *in the laboratory: Multivariate statistical analysis of relationships between exposure, biomarkers, and ecologically relevant parameters.* Environ. Sci. Technol., 2005. **39**(6): p. 1757-1763.
155. Maret, W., *Redox biochemistry of mammalian metallothioneins.* J Biol Inorg Chem, 2011. **16**(7): p. 1079-1086.
156. Stürzenbaum, S.R., *Chapter 7: Earthworm and Nematode Metallothioneins*, in *Metallothioneins and Related Chelators*, Sigel, A., Sigel, H., and Sigel, R.K.O., Editors. 2009, RSC Pub.: Cambridge, UK. p. 514 p.
157. Ngu, T., Stürzenbaum, S., and Stillman, M., *Cadmium binding studies to the earthworm* Lumbricus rubellus *metallothionein by electrospray mass spectrometry and circular dichroism spectroscopy.* Biochem. Biophys. Res. Commun., 2006. **351**(1): p. 229-233.
158. Fischer, E.H. and Davie, E.W., *Recent excitement regarding metallothionein.* Proc Natl Acad Sci U S A, 1998. **95**(7): p. 3333-3334.
159. Paoli, M., Marles-Wright, J., and Smith, A., *Structure-function relationships in heme-proteins.* DNA Cell Biol., 2002. **21**(4): p. 271-280.
160. Kennedy, M.L. and Gibney, B.R., *Metalloprotein and redox protein design.* Curr. Opin. Struct. Biol., 2001. **11**(4): p. 485-490.
161. Karaca, A., ed. *Biology of Earthworms*. First Edition ed. Soil Biology. 2011, Springer: New York, USA.
162. Lukkari, T., Taavitsainen, M., Vaisanen, A., and Haimi, J., *Effects of heavy metals on earthworms along contamination gradients in organic rich soils.* Ecotoxicol. Environ. Saf., 2004. **59**(3): p. 340-348.
163. Morgan, J.E. and Morgan, A.J., *Seasonal changes in the tissue-metal (Cd, Zn and Pb) concentrations in two ecophysiologically dissimilar earthworm species: pollution-monitoring implications.* Environ. Pollut., 1992. **82**: p. 1-7.
164. Rhee, I.K., Lee, K.S., and Huang, P.C., *Metallothioneins with interdomain hinges expanded by insertion mutagenesis.* Protein Eng, 1990. **3**(3): p. 205-213.

165. Winge, D.R. and Miklossy, K.A., *Domain nature of metallothionein.* J Biol Chem, 1982. **257**(7): p. 3471-3476.
166. Gruber, C., Stürzenbaum, S., Gehrig, P., Sack, R., Hunziker, P., Berger, B., and Dallinger, R., *Isolation and characterization of a self-sufficient one-domain protein - Cd-metallothionein from* Eisenia foetida*.* Eur. J. Biochem., 2000. **267**(2): p. 573-582.
167. Robinson, N.J., Tommey, A.M., Kuske, C., and Jackson, P.J., *Plant metallothioneins.* Biochem J, 1993. **295 (Pt 1)**: p. 1-10.
168. Leszczyszyn, O.I., Schmid, R., and Blindauer, C.A., *Toward a property/function relationship for metallothioneins: Histidine coordination and unusual cluster composition in a zinc-metal lothionein from plants.* Proteins, 2007. **68**(4): p. 922-935.
169. Cortet, J., Gomot-De-Vauflery, A., Poinsot-Balaguer, N., Gomot, L., Texier, C., and Cluzeau, D., *The use of invertebrate soil fauna in monitoring pollutant effects.* Eur. J. Soil Biol., 1999. **35**(3): p. 115-134.
170. Livingstone, D.R., *Biotechnology and pollution monitoring: use of molecular biomarkers in the aquatic environment.* J Chem Tech Biotechnol, 1993. **57**: p. 195-211.
171. Regoli, F. and Orlando, E., *Seasonal variation of trace metal concentrations in the digestive gland of the Mediterranean mussel* Mytilus galloprovincialis*: comparison between a polluted and a non-polluted site.* Arch. Environ. Contam. Toxicol., 1994. **27**(1): p. 36-43.
172. K. Walsh, R. H. Dunstan, and Murdoch, R.N., *Differential bioaccumulation of heavy metals and organopollutants in the soft tissue and shell of the marine gastropod,* Austrocochlea constricta*.* Arch. Environ. Contam. Toxicol., 1995. **28**(1): p. 35-39.
173. McLoughlin, N., Yin, D., Maltby, L., Wood, R. M. and Yu, H., *Evaluation of sensitivity and specificity of two crustacean biochemical biomarkers.* Environ. Toxicol. Chem., 2000. **19**(8): p. 2085-2092.
174. Au, D.W., *The application of histo-cytopathological biomarkers in marine pollution monitoring: a review.* Mar. Pollut. Bull., 2004. **48**(9-10): p. 817-834.
175. De Luca-Abbott, S., *Biomarkers of sublethal stress in the soft-sediment bivalve* Austrovenus stutchburyi *exposed in-situ to contaminated sediment in an urban New Zealand harbour.* Mar. Pollut. Bull., 2001. **42**(10): p. 817-825.
176. Rodríguez-Castellanos, L. and Sanchez-Hernandez, J.C., *Earthworm biomarkers of pesticide contamination: Current status and perspectives.* J. Pestic. Sci., 2007. **32**(4): p. 360-371.
177. Gastaldi, L., Ranzato, E., Capri, F., Hankard, P., Peres, G., Canesi, L., Viarengo, A., and Pons, G., *Application of a biomarker battery for the evaluation of the sublethal effects of pollutants in the earthworm Eisenia andrei.* Comp Biochem Phys C, 2007. **146**(3): p. 398-405.
178. Morgan, A.J., Kille, P., and Stürzenbaum, S.R., *Microevolution and ecotoxicology of metals in invertebrates.* Environ. Sci. Technol., 2007. **41**(4): p. 1085-1096.
179. Galay-Burgos, M., Spurgeon, D.J., Weeks, J.M., Stürzenbaum, S.R., Morgan, A.J., Kille, P., *Developing a new method for soil pollution monitoring using molecular genetic biomarkers.* Biomarkers, 2003. **8**(3-4): p. 229-239.
180. Sarkar, A., Ray, D., Shrivastava, A.N., and Sarker, S., *Molecular Biomarkers: their significance and application in marine pollution monitoring.* Ecotoxicology, 2006. **15**(4): p. 333-340.
181. Galay-Burgos, M., Spurgeon, D.J., Weeks, J.M., Stürzenbaum, S.R., Morgan, A.J., and Kille, P., *Developing a new method for soil pollution monitoring using molecular genetic biomarkers.* Biomarkers, 2003. **8**(3-4): p. 229-239.

182. Dallinger, R., Berger, B., Gruber, C., Hunziker, P., and Stürzenbaum, S., *Metallothioneins in terrestrial invertebrates: Structural aspects, biological significance and implications for their use as biomarkers.* Cell. Mol. Biol., 2000. **46**(2): p. 331-346.
183. Brody, J.R. and Kern, S.E., *Sodium boric acid: a Tris-free, cooler conductive medium for DNA electrophoresis.* BioTechniques, 2004. **36**(2): p. 214-216.
184. Studier, F.W., *Protein production by auto-induction in high-density shaking cultures.* Protein Expression Purif., 2005. **41**(1): p. 207-234.
185. Nölte, J., *ICP Emission Spectrometry: A Practical Guide*. First Edition ed. 2003, Weinheim, Germany: Wiley VCH.
186. Taylor, H.E., *Inductively Coupled Plasma-Mass Spectrometry: Practices and Techniques*. First Edition ed. 2000, Waltham, USA: Academic Press Inc.
187. Riddles, P.W., Blakeley, R.L., and Zerner, B., *Reassessment of Ellman's Reagent.* Methods in Enzymology., 1983. **91**: p. 49-60.
188. Jacobsen, N.E., *NMR Spectroscopy Explained: Simplified Theory, Applications and Examples for Organic Chemistry and Structural Biology*. First Edition ed. 2007, Hoboken, USA: Wiley-Blackwell.
189. Claridge, T.D.W., *High-Resolution NMR Techniques in Organic Chemistry*. Second Edition ed. 2008, Kidlington, UK: Elsevier Science Ltd.
190. Keeler, J., *Understanding NMR Spectro*scopy. Second Edition ed. 2010, Hoboken, USA: Wiley-Blackwell.
191. Jeener, J., Meier, B.H., Bachmann, P., and Ernst, R.R., *Investigation of exchange processes by two-dimensional NMR spectroscopy.* J Chem Phys, 1979. **71**: p. 4546-4553.
192. Braunschweiler, L., and Ernst, R.R., *Coherence transfer by isotropic mixing: Application to proton correlation spectroscopy*, J. Magn. Reson., 1983. **53**:p. 521.
193. Grzesik, S., and Bax, A., *Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR.* J Magn Res, 1992. **96**: p. 432-440.
194. Bax, A., and Ikura, M., *An efficient 3D NMR technique for correlating the proton and15N backbone amide resonances with the α-carbon of the preceding residue in uniformly15N/13C enriched proteins.* J Biomol NMR, 1991. **1**(1): p. 99-104.
195. Vuister, G.W., and Bax, A., *Quantitative J correlation: a new approach for measuring homonuclear three-bond J (HNHα) coupling constants in ^{15}N-enriched proteins.* JACS, 1993. **115**: p. 7772-7777.
196. Bax, A., Griffey, R.H., and Hawkins, B.L., *Sensitivity-enhanced correlation of nitrogen-15 and proton chemical shifts in natural-abundance samples via multiple quantum coherence.* J Magn Reson, 1983. **55**: p. 301-315.
197. Davis, A.L., Keeler, J., Laue, E.D., and Moskau, D., *Experiments for recording pure-absorption heteronuclear correlation spectra using pulsed field gradients.* J Magn Reson, 1992. **98**: p. 207-216.
198. Güntert, P., Mumenthaler, C., and Wüthrich, K., *Torsion angle dynamics for NMR structure calculation with the new program DYANA.* J Mol Biol, 1997. **273**(1): p. 283-298.
199. Güntert, P., *Automated NMR Structure Calculation With CYANA.* Methods Mol. Biol., 2004. **278**: p. 353-378.
200. Hasler, D.W., Jensen, L.T., Zerbe, O., Winge, D.R., and Vašák, M., *Effect of the two conserved prolines of human growth inhibitory factor (metallothionein-3) on its biological activity and structure fluctuation: comparison with a mutant protein.* Biochemistry, 2000. **39**(47): p. 14567-14575.

201. Lichty, J.J., Malecki, J.L., Agnew, H.D., Michelson-Horowitz, D.J., and Tan, S., *Comparison of affinity tags for protein purification.* Protein Expression Purif., 2005. **41**(1): p. 98-105.
202. Terpe, K., *Overview of tag protein fusions: from molecular and biochemical fundamentals to commercial systems.* Appl. Microbiol. Biotechnol., 2003. **60**: p. 523-533.
203. Cols, N., Romero-Isart, N., Bofill, R., Capdevila, M., Gonzalez-Duarte, P., Gonzalez-Duarte, R., and Atrian, S., *In vivo copper- and cadmium-binding ability of mammalian metallothionein beta domain.* Protein Eng., 1999. **12**(3): p. 265-269.
204. Bjellqvist, B., Hughes, G.J., Pasquali, C., Paquet, N., Ravier, F., Sanchez, J.C., Frutiger, S., and Hochstrasser, D., *The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences.* Electrophoresis, 1993. **14**(10): p. 1023-1031.
205. Perez-Rafael, S., Atrian, S., Capdevila, M., and Palacios, O., *Differential ESI-MS behaviour of highly similar metallothioneins.* Talanta, 2010. **83**(3): p. 1057-1061.
206. Sutherland, D.E., Willans, M.J., and Stillman, M.J., *Single domain metallothioneins: supermetalation of human MT 1a.* J Am Chem Soc, 2012. **134**(6): p. 3290-3299.
207. Sutherland, D.E., Willans, M.J., and Stillman, M.J., *Supermetalation of the beta domain of human metallothionein 1a.* Biochemistry, 2010. **49**(17): p. 3593-3601.
208. Deutscher, J., *The mechanisms of carbon catabolite repression in bacteria.* Curr. Opin. Microbiol., 2008. **11**(2): p. 87-93.
209. Sorensen, H.P. and Mortensen, K.K., *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microb. Cell Fact., 2005. **4**(1): p. 1.
210. Jaganaman, S., Pinto, A., Tarasev, M., and Ballou, D.P., *High levels of expression of the iron-sulfur proteins phthalate dioxygenase and phthalate dioxygenase reductase in* Escherichia coli*.* Protein Expression Purif., 2007. **52**(2): p. 273-279.
211. Van Oss, C.J., *On the mechanism of the cold ethanol precipitation method of plasma protein fractionation.* J. Protein Chem., 1989. **8**(5): p. 661-668.
212. Kille, P., Stephens, P., Cryer, A., and Kay, J., *The expression of a synthetic rainbow trout metallothionein gene in* E. coli*.* Biochim Biophys Acta, 1990. **1048**(2-3): p. 178-186.
213. Yu, W.H., Cai, B., Gao, Y., Xie, Y., and Huang, Z.X., *Expression, characterization, and reaction of recombinant monkey metallothionein-1 and its C33M mutant.* J. Protein Chem., 2002. **21**(3): p. 177-185.
214. Bofill, R., Capdevila, M., and Atrian, S., *Independent metal-binding features of recombinant metallothioneins convergently draw a step gradation between Zn- and Cu-thioneins.* Metallomics, 2009. **1**(3): p. 229-234.
215. Driscoll, P.C., Gronenborn, A.M., Beress, L., and Clore, G.M., *Determination of the three-dimensional solution structure of the antihypertensive and antiviral protein BDS-I from the sea anemone* Anemonia sulcata*: a study using nuclear magnetic resonance and hybrid distance geometry-dynamical simulated annealing.* Biochemistry, 1989. **28**(5): p. 2188-2198.
216. Scudiero, R., Temussi, P.A., and Parisi, E., *Fish and mammalian metallothioneins: a comparative study.* Gene, 2005. **345**(1): p. 21-26.
217. Vašák, M., *Dynamic metal-thiolate cluster structure of metallothioneins.* Environ Health Perspect, 1986. **65**: p. 193-197.
218. Rule, G.S., and Hitchens, T.K., *Fundamentals of Protein NMR Spectr*oscopy. 2006 Edition ed. 2006, Dordrecht, Netherlands: Springer.
219. Wüthrich, K., Wider, G., Wagner, G., and Braun, W., *Sequential resonance assignments as a basis for determination of spatial protein structures by high resolution proton nuclear magnetic resonance.* J Mol Biol, 1982. **155**(3): p. 311-319.

220. Billeter, M., Braun, W., and Wüthrich, K., *Sequential resonance assignments in protein 1H nuclear magnetic resonance spectra. Computation of sterically allowed proton-proton distances and statistical analysis of proton-proton distances in single crystal protein conformations.* J Mol Biol, 1982. **155**(3): p. 321-346.
221. Kumar, A., Ernst, R.R., and Wüthrich, K., *A two-dimensional nuclear Overhauser enhancement (2D NOE) experiment for the elucidation of complete proton-proton cross-relaxation networks in biological macromolecules.* Biochem Biophys Res Commun, 1980. **95**(1): p. 1-6.
222. Vuister, G.W. and Bax, A., *Quantitative J correlation: a new approach for measuring homonuclear three-bond J(HNHa) coupling constants in ^{15}N-enriched proteins.* J Am Chem Soc, 1993. **115**(17): p. 7772-7777.
223. Hemmingsen, L., Olsen, L., Antony, J., and Sauer, S.P., *First principle calculations of ^{113}Cd chemical shifts for proteins and model systems.* J Biol Inorg Chem, 2004. **9**(5): p. 591-599.
224. Arseniev, A., Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vašák, M., Kägi, J.H., and Wüthrich, K., *Three-dimensional structure of rabbit liver Cd_7 metallothionein-2a in aqueous solution determined by nuclear magnetic resonance.* J Mol Biol, 1988. **201**(3): p. 637-657.
225. Braun, W., Wagner, G., Worgotter, E., Vašák, M., Kägi, J.H., and Wüthrich, K., *Polypeptide fold in the two metal clusters of metallothionein-2 by nuclear magnetic resonance in solution.* J Mol Biol, 1986. **187**(1): p. 125-129.
226. Peroza, E.A., Schmucki, R., Güntert, P., Freisinger, E., and Zerbe, O., *The β_E-Domain of Wheat E_c-1 Metallothionein: A Metal-Binding Domain with a Distinctive Structure.* J. Mol. Biol., 2009. **387**(1): p. 207-218.
227. Jensen, K.P. and Rykaer, M., *The building blocks of metallothioneins: heterometallic Zn^{2+} and Cd^{2+} clusters from first-principles calculations.* Dalton Trans, 2010. **39**(40): p. 9684-9695.
228. Pattabiraman, N., Ward, K.B., and Fleming, P.J., *Occluded molecular surface: analysis of protein packing.* J. Mol. Recognit., 1995. **8**(6): p. 334-344.
229. Good, M., Hollenstein, R., Sadler, P.J., and Vašák, M., *^{113}Cd NMR studies on metal-thiolate cluster formation in rabbit Cd(II)-metallothionein: evidence for a pH dependence.* Biochemistry, 1988. **27**(18): p. 7163-7166.
230. Huang, Z.X., Gao, Y., Yu, W.H., Zhang, S.Y., and Yang, P.Y., *Construction of alpha-alpha domains structure in recombinant monkey metallothionein-1.* J Inorg Biochem, 2002. **92**(3-4): p. 183-192.
231. P. Kuan-Yeu Pan, Z. F. Zheng, P. C. Lyu, and Huang, P.C., *Why reversing the sequence of the α domain of human metallothionein-2 does not change its metal-binding and folding characteristics.* FEBS J, 1999. **266**(1): p. 33-39.
232. Spurgeon, D.J., Stürzenbaum, S.R., Svendsen, C., Hankard, P.K., Morgan, A.J., Weeks, J.M., and Kille, P., *Toxicological, cellular and gene expression responses in earthworms exposed to copper and cadmium.* Comp Biochem Phys C, 2004. **138**(1): p. 11-21.
233. Cooper, G.M. and Hausman, R.E., *Chapter 9: Protein Sorting and Transport*, in *The Cell: A Molecular Approach*, Cooper, G.M., Editor. 2000, Sinauer Associates: Sunderland, USA.
234. Sutherland, D.E. and Stillman, M.J., *The "magic numbers" of metallothionein.* Metallomics, 2011. **3**(5): p. 444-463.
235. Tommey, A.M., Shi, J., Lindsay, W.P., Urwin, P.E., and Robinson, N.J., *Expression of the pea gene* PSMTA *in* E. coli*. Metal-binding properties of the expressed protein.* FEBS Lett., 1991. **292**(1-2): p. 48-52.

236. Munoz, A., Laib, F., Petering, D.H., and Shaw, C.F., 3rd, *Characterization of the cadmium complex of peptide 49-61: a putative nucleation center for cadmium-induced folding in rabbit liver metallothionein IIA.* J Biol Inorg Chem, 1999. **4**(4): p. 495-507.
237. Freisinger, E., *Chapter 5: Metallothioneins in plants*, in *Metallothioneins and Related Chelators*, Sigel, A., Sigel, H., and Sigel, R.K.O., Editors. 2009, Royal Society of Chemistry: London, UK.
238. Xiong, Y. and Ru, B., *Purification and characteristics of recombinant mouse metallothionein-I from* Escherichia coli*.* J Biochem, 1997. **121**(6): p. 1102-1106.
239. Fowler, B.A. and Megginson, M.M., *Isolation and partial characterization of a high molecular weight Cd/Zn-binding protein from the kidney of the scallop* Placopecten magellanicus*: preliminary studies.* Environ Health Perspect, 1986. **65**: p. 199-203.
240. Reese, R.N. and Wagner, G.J., *Properties of tobacco (Nicotiana tabacum) cadmium-binding peptide(s). Unique non-metallothionein cadmium ligands.* Biochem J, 1987. **241**(3): p. 641-647.
241. Leszczyszyn, O.I. and Blindauer, C.A., *Zinc transfer from the embryo-specific metallothionein E_C from wheat: a case study.* Phys. Chem. Chem. Phys. **12**(41): p. 13408-13418.
242. Gan, T., Munoz, A., Shaw, C.F., 3rd, and Petering, D.H., *Reaction of $^{111}Cd_7$-metallothionein with EDTA. A reappraisal.* J Biol Chem, 1995. **270**(10): p. 5339-5345.
243. Blindauer, C.A., *Metallothioneins with unusual residues: histidines as modulators of zinc affinity and reactivity.* J Inorg Biochem, 2008. **102**(3): p. 507-521.
244. Li, T.Y., Kraker, A.J., Shaw, C.F. 3rd, and Petering, D.H., *Ligand substitution reactions of metallothioneins with EDTA and apo-carbonic anhydrase.* Proc Natl Acad Sci U S A, 1980. **77**(11): p. 6334-6338.
245. Smith, G.A., Hesketh, R.T., Metcalfe, J.C., Feeney, J., and Morris, P.G., *Intracellular calcium measurements by ^{19}F NMR of fluorine-labeled chelators.* Proc Natl Acad Sci U S A, 1983. **80**(23): p. 7178-7182.
246. Benters, J., Flogel, U., Schafer, T., Leibfritz, D., Hechtenberg, S., and Beyersmann, D., *Study of the interactions of cadmium and zinc ions with cellular calcium homoeostasis using ^{19}F-NMR spectroscopy.* Biochem J, 1997. **322 (Pt 3)**: p. 793-799.
247. Foster, A.W. and Robinson, N.J., *Promiscuity and preferences of metallothioneins: the cell rules.* BMC Biol., 2011. **9**: p. 25.
248. Montelione, G.T., Zheng, D., Huang, Y.J., Gunsalus, K.C., and Szyperski, T., *Protein NMR spectroscopy in structural genomics.* Nat. Struct. Biol., 2000. **7 Suppl**: p. 982-985.
249. Vašák, M., *Advances in metallothionein structure and functions.* J. Trace Elem. Med. Biol., 2005. **19**(1): p. 13-17.
250. Roschitzki, B. and Vašák, M., *Redox labile site in a Zn_4 cluster of Cu_4,Zn_4-metallothionein-3.* Biochemistry, 2003. **42**(32): p. 9822-9828.
251. Vašák, M., *Application of ^{113}Cd NMR to metallothioneins.* Biodegradation, 1998. **9**(6): p. 501-512.
252. Martin, A.C., Orengo, C.A., Hutchinson, E.G., Jones, S., Karmirantzou, M., Laskowski, R.A., Mitchell, J.B., Taroni, C., and Thornton, J.M., *Protein folds and functions.* Structure, 1998. **6**(7): p. 875-884.
253. Bradshaw, R.A., Murray-Rust, J., Ibanez, C.F., McDonald, N.Q., Lapatto, R., and Blundell, T.L., *Nerve growth factor: structure/function relationships.* Protein Sci., 1994. **3**(11): p. 1901-1913.

254. Cheng, J., Randall, A.Z., Sweredoski, M.J., and Baldi, P., *SCRATCH: a protein structure and structural feature prediction server.* Nucleic Acids Res., 2005. **33**(2): p. 72-76.
255. Schwede, T., Kopp, J., Guex, N., and Peitsch, M.C., *SWISS-MODEL: An automated protein homology-modeling server.* Nucleic Acids Res, 2003. **31**(13): p. 3381-3385.
256. Fischer, D., Rice, D., Bowie, J.U., and Eisenberg, D., *Assigning amino acid sequences to 3-dimensional protein folds.* FASEB J., 1996. **10**(1): p. 126-136.
257. Faller, P. and Vašák, M., *Distinct metal-thiolate clusters in the N-terminal domain of neuronal growth inhibitory factor.* Biochemistry, 1997. **36**(43): p. 13341-13348.
258. Vašák, M. and Meloni, G., *Chemistry and biology of mammalian metallothioneins.* J Biol Inorg Chem, 2011. **16**(7): p. 1067-1078.
259. Blindauer, C.A., *A metallothionein containing a zinc finger within a four-metal cluster protects a bacterium from zinc toxicity.* P Natl Acad Sci USA, 2001. **98**(17): p. 9593-9598.
260. Leszczyszyn, O.I., White, C.R., and Blindauer, C.A., *The isolated Cys_2His_2 site in E_C metallothionein mediates metal-specific protein folding.* Mol. BioSyst., 2010. **6**(9): p. 1592-1603.
261. Schultze, P., Worgotter, E., Braun, W., Wagner, G., Vašák, M., Kägi, J.H., and Wüthrich, K., *Conformation of Cd_7-metallothionein-2 from rat liver in aqueous solution determined by nuclear magnetic resonance spectroscopy.* J Mol Biol, 1988. **203**(1): p. 251-268.
262. Zangger, K., Oz, G., Otvos, J.D., and Armitage, I.M., *Three-dimensional solution structure of mouse Cd_7-metallothionein-1 by homonuclear and heteronuclear NMR spectroscopy.* Protein Sci., 1999. **8**(12): p. 2630-2638.
263. Narula, S.S., Brouwer, M., Hua, Y., and Armitage, I.M., *Three-dimensional solution structure of* Callinectes sapidus *metallothionein-1 determined by homonuclear and heteronuclear magnetic resonance spectroscopy.* Biochemistry, 1995. **34**(2): p. 620-631.
264. Lehman-McKeeman, L.D. and Klaassen, C.D., *Induction of* metallothionein-I *and* metallothionein-II *in rats by cadmium and zinc.* Toxicol. Appl. Pharmacol., 1987. **88**(2): p. 195-202.
265. Tottey, S., Waldron, K.J., Firbank, S.J., Reale, B., Bessant, C., Sato, K., Cheek, T.R., Gray, J., Banfield, M.J*., et al.*, *Protein-folding location can regulate manganese-binding versus copper- or zinc-binding.* Nature, 2008. **455**(7216): p. 1138-1142.
266. Jiang, L.J., Vašák, M., Vallee, B.L., and Maret, W., *Zinc transfer potentials of the α- and β-clusters of metallothionein are affected by domain interactions in the whole molecule.* Proc Natl Acad Sci U S A, 2000. **97**(6): p. 2503-2508.

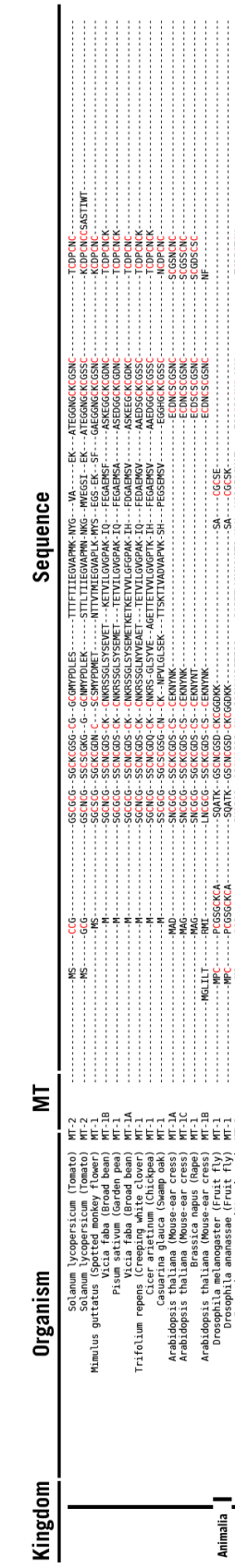
Appendix 1

Kingdom
Organism
MT
Sequence
Chromista
Animalia
Plantae
Fungi
Bacteria
Animalia

Appendix 1

Kingdom
Organism
MT
Sequence
Fungi
Plantae
Mus musculus (Mouse) MT-4
Canis familiaris (Dog) MT-4
Homo sapiens (Human) MT-4
Scyliorhinus torazame (Cloudy catshark) MT
Oryzias latipes (Medaka fish) MT
Liza aurata (Golden grey mullet) MT
Chaenocephalus aceratus (Blackfin icefish) MT-A
Parachaenichthys charcoti (Charcot's dragonfish) MT-A
Chionodraco rastrospinosus (Ocellated icefish) MT-A
Trematomus bernacchii (Emerald rockcod) MT-A
Chionodraco hamatus (Antarctic teleost icefish) MT-A
Chionodraco hamatus (Antarctic teleost icefish) MT-B
Dicentrarchus labrax (European seabass) MT-B
Chaenocephalus aceratus (Blackfin icefish) MT-B
Chionodraco rastrospinosus (Ocellated icefish) MT-B
Thermarces cerberus (Hydrothermal vent eelpout) MT-A
Zoarces viviparus (Viviparous eelpout) MT
Pagrus major (Red sea bream) MT
Pleuronectes platessa (European plaice) MT
Pseudopleuronectes americanus (Winter flounder) MT
Takifugu flavidus (Sansha pufferfish) MT
Oreochromis mossambicus (Mozambique tilapia) MT
Oreochromis aureus (Israeli tilapia) MT
Sparus aurata (Gilthead sea bream) MT-A
Lithognathus mormyrus (Striped seabream) MT
Gadus morhua (Atlantic cod) MT
Gobiomorphus cotidianus (New Zealand common bully) MT
Oncorhynchus keta (Keta salmon) MT-B
Salvelinus alpinus (Arctic char) MT-A
Salmo salar (Atlantic salmon) MT-B
Esox lucius (Northern pike) MT
Plecoglossus altivelis (Ayu) MT
Cyprinus carpio (Common carp) MT-1
Carassius auratus cuvieri (Common carp) MT-B
Rutilus rutilus (Roach) MT
Cyprinus carpio (Common carp) MT-2
Carassius auratus (Goldfish) MT
Carassius cuvieri (Crucian carp) MT-A
Danio rerio (Zebrafish) MT-2
Danio rerio (Zebrafish) MT-1
Ictalurus punctatus (Channel catfish) MT
Cyprinodon sp. (Pupfish) MT-A
Nectria lugdunensis (Aquatic fungus) MT
Colletotrichum gloeosporioides (Anthracnose fungus) MT-CAP5
Scylla serrata (Mud crab) MT-1
Callinectes sapidus (Blue crab) MT-1B
Callinectes sapidus (Blue crab) MT-2B
Carcinus maenas (Common shore crab) MT-1
Homarus americanus (American lobster) MT-1
Astacus fluviatilis (Broad-fingered crayfish) MT
Potamon potamios MT
Triticum aestivum (Wheat) EC I/II
Triticum aestivum (Wheat) EC III
Zea mays (Maize) EC
Oryza sativa subsp. japonica (Rice) MT-1A
Arabidopsis thaliana (Mouse-ear cress) EC I
Arabidopsis thaliana (Mouse-ear cress) EC III
Colletotrichum gloeosporioides (Anthracnose fungus) MT-CAP3
Oryza sativa subsp. japonica (Rice) MT-2C
Oryza sativa subsp. japonica (Rice) MT-2B
Oryza sativa subsp. indica (Rice) MT-1A
Oryza sativa subsp. japonica (Rice) MT-1B
Zea mays (Maize) MT-1
Hordeum vulgare (Barley) MT-1
Oryza sativa subsp. japonica (Rice) MT-4B
Oryza sativa subsp. japonica (Rice) MT-4C
Oryza sativa subsp. japonica (Rice) MT-4A
Festuca rubra (Red fescue) MT-1
Triticum aestivum (Wheat) MT-2
Musa acuminata (Banana) MT-2
Oryza sativa subsp. japonica (Rice) MT-2A
Malus domestica (Apple) MT-2
Actinidia deliciosa (Kiwi) MT-2
Actinidia deliciosa (Kiwi) MT-2B
Cicer arietinum (Chickpea) MT-2
Trifolium repens (Creeping white clover) MT-2
Ricinus communis (Castor bean) MT-2
Solanum lycopersicum (Tomato) MT-2B
Nicotiana glutinosa (Tobacco) MT-2
Brassica napus subsp. napus (Oilseed rape) MT-EIF98
Arabidopsis thaliana (Mouse-ear cress) MT-2A
Brassica juncea (Indian mustard) MT-2
Brassica juncea (Indian mustard) MT-2
Brassica juncea (Indian mustard) MT-2
Thlaspi caerulescens (Alpine penny-cress) MT-2
Arabidopsis thaliana (Mouse-ear cress) MT-2B
Coffea canephora (Robusta coffee) MT-1
Solanum lycopersicum (Tomato) MT-2A
Solanum lycopersicum (Tomato) MT-2

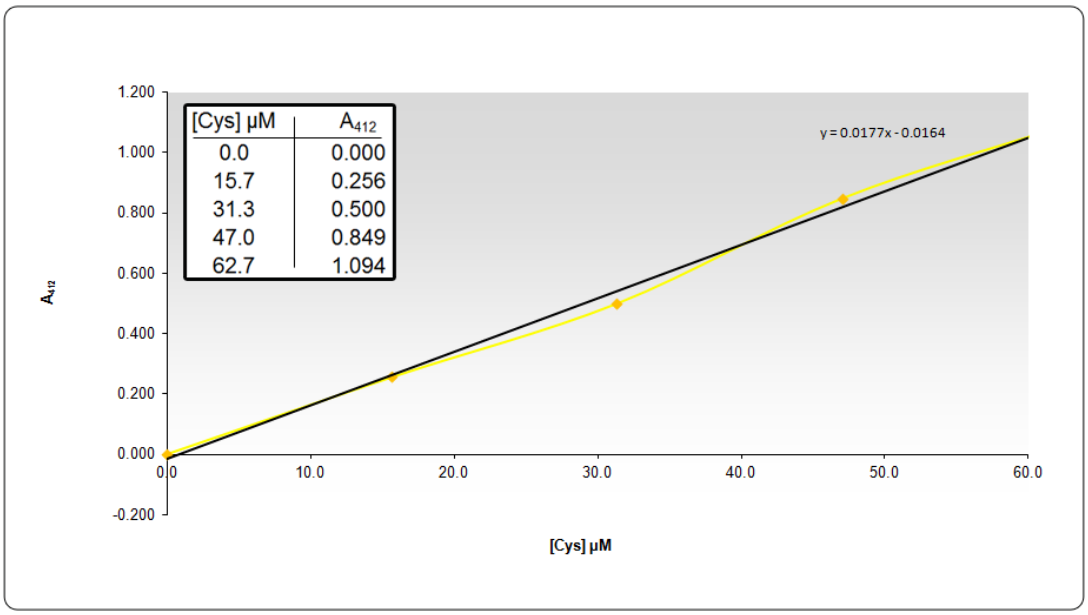
Appendix 1



Appendix 1. Multiple sequence alignment from CLUSTALW2 of metallothionein sequences within the ExPASY database (http://expasy.org/). The sequences have been ordered due to similarity, which tends to group sequences which are thought to be evolutionary related. This shows the great diversity in MT primary sequence between kingdoms. Cysteine residues have been indicated in **RED** for ease of identification.

Appendix 2

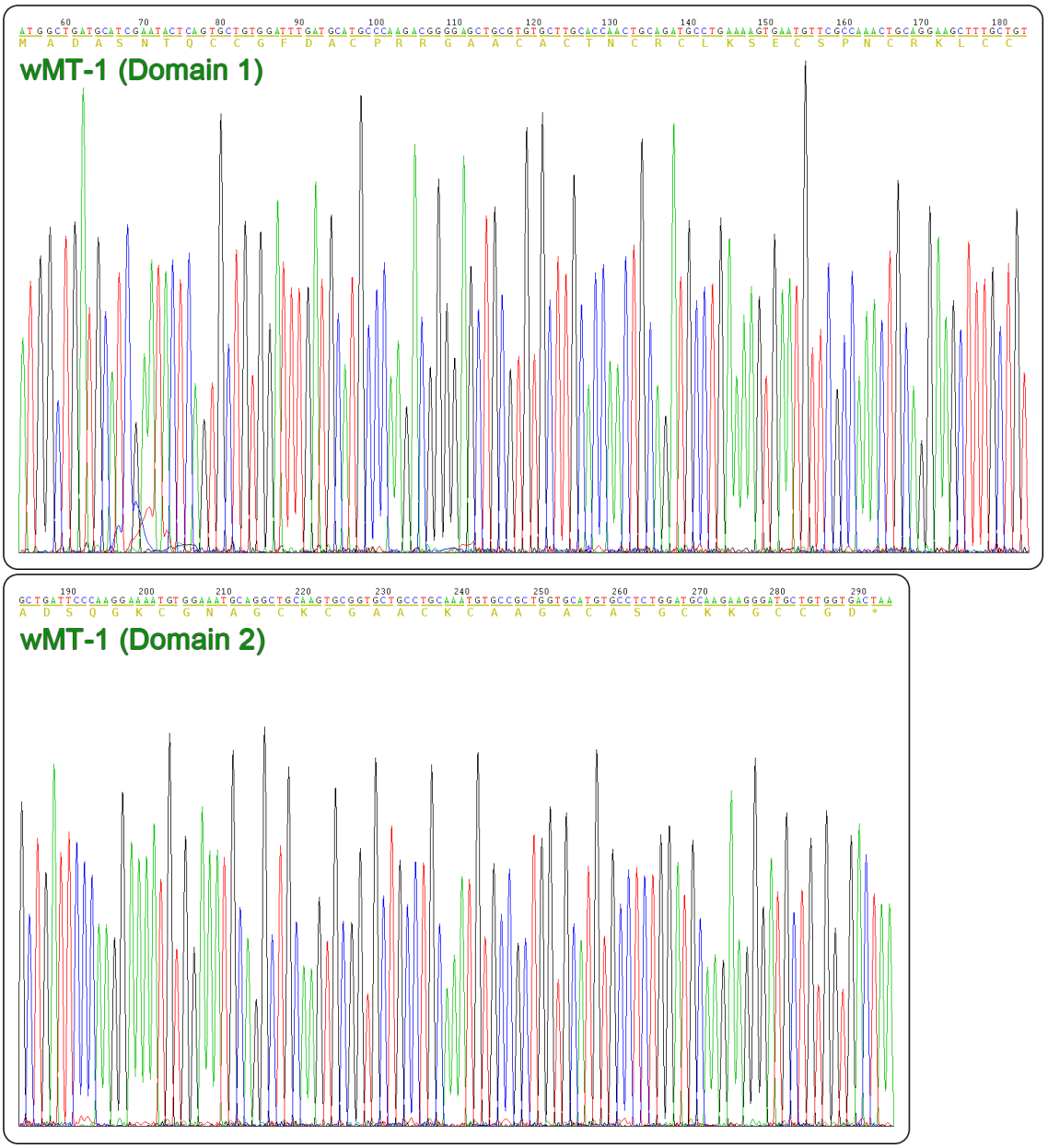
Referred from: Chapter 2, Section 2.3 – *Cysteine Assay*



Appendix 2. Calibration curve used for Cys Assay, showing the standards used, and absorbencies recorded. The wavelength used was 412 nm. A Line of Best Fit was used for concentration calculations, with the equation shown on the graph.

Appendix 3

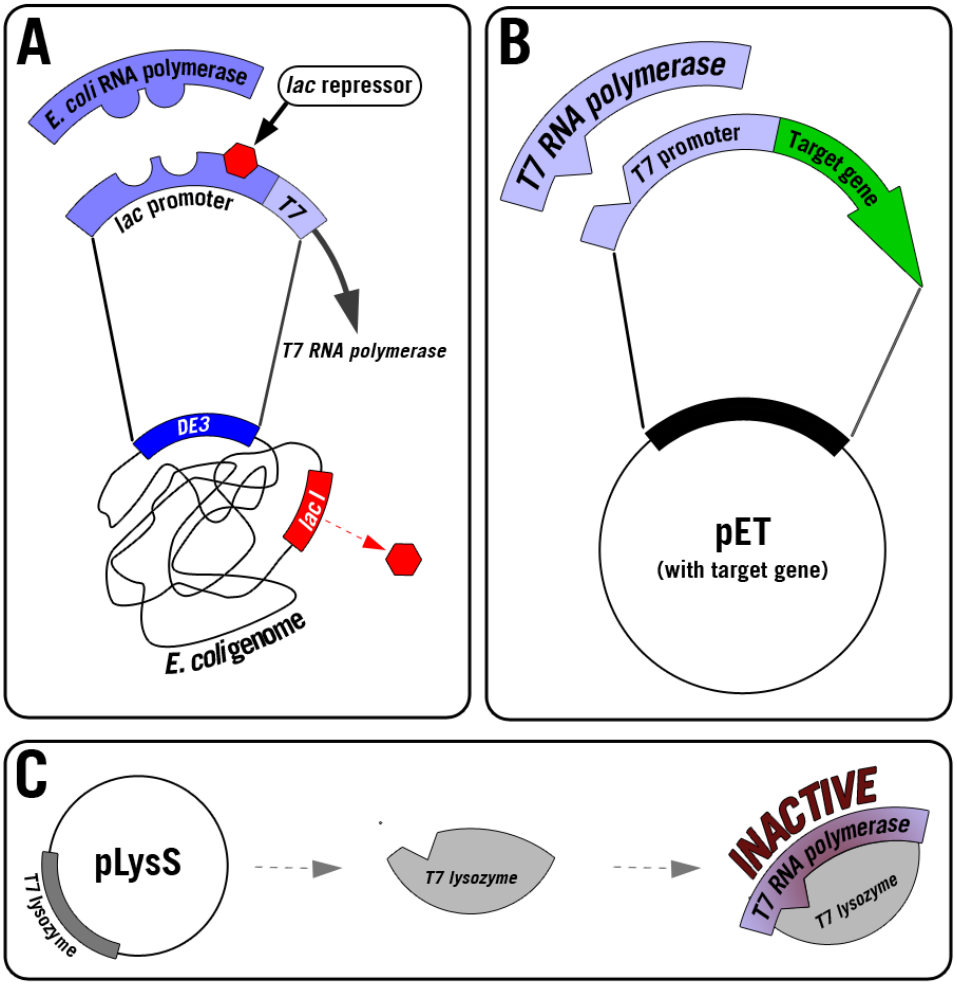
Referred from: Chapter 3, Section 3.5.



Appendix 3. (**TOP**) Sequencing data of Domain 1, (**BOTTOM**) Sequencing data of Domain 2. Peak colours are: Guanine **BLACK**, Thymine **RED**, Cytosine **BLUE**, Adenine **GREEN**. The translated codons are shown in **YELLOW**, below the DNA bases. Sequencing data analysis performed using Chromas Lite [Technelysium Pty Ltd, Australia].

Appendix 4

Referred from: Chapter 3, Section 3.4



Appendix 4. Schematic of the pET expression system, modified from the Novagen pET Sytem Manual. Genes, targets and proteins are indicated in similar colours. (**A**) Shows the control of the production of T7 RNA polymerase by the use of a chromosomal copy of the *lacI* gene, encoding a lac repressor protein. When present, *lacI* does not allow binding of the native *E. coli* RNA polymerase. Addition of IPTG will remove the lac repressor protein, allowing transcription of T7 RNA polymerase. (**B**) Shows the position of the target gene in relation to the T7 promoter region, in the pET vector. (**C**) Shows the action of T7 lysozyme, encoded by the pLysS plasmid. When expressed, T7 lysozyme will inactivate T7 RNA polymerase, inhibiting basal

Appendix 5

Referred from: Chapter 3, Section 3.7

Appendix 5. For the indicated peaks, in Figure 3.16, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.6% of their theoretical counterparts. A, B, C designates the spectrum from which the readings were taken.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
C	Zn_7- wMT-1	8,394.9	8,394.3
C	Zn_6- wMT-1	8,331.5	8,328.5
C	Zn_5- wMT-1	8,268.1	8,260.1
D	*		8,321.0
D	*		8,141.4
D	Zn_1- wMT-1	8,077.9	8,031.3
D	apo- wMT-1	7,951.2	7,942.9

Appendix 6

Referred from: Chapter 3, Section 3.8

Appendix 6. For the indicated peaks, in Figure 3.19, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.11% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	**Cd_8- wMT-1**	11,645.7	11,643.0
A	**Cd_7- wMT-1**	11,535.3	11,535.9
B	**apo- wMT-1**	10,762.4	10,773.3

Appendix 7

Referred from: Chapter 3, Section 3.8

Appendix 7. For the indicated peaks, in Figure 3.23, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.02% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	**Cd_8- wMT-1**	8,834.5	8,835.9
A	**Cd_7- wMT-1**	8,724.0	8,723.2
B	*		8,829.8
B	*		8,718.4
B	*		8,619.0
B	**Cd_5- wMT-1**	8,503.2	8,501.0
B	**Cd_4- wMT-1**	8,392.8	8,393.3
B	**Cd_1- wMT-1**	8,061.6	8,061.9
B	**apo- wMT-1**	7,951.2	7,950.5

Appendix 8

Referred from: Chapter 3, Section 3.10

Appendix 8. For the indicated peaks, in Figure 3.33, the theoretical and observed masses for the single metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. Observed mass within 0.006% of its theoretical counterpart. B designates the spectrum from which the reading was observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
B	**Cd_7- wMT-2**	11,382.0	11,382.6

Appendix 9

Referred from: Chapter 3, Section 3.10

Appendix 9. For the indicated peaks, in Figure 3.38, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.1%, most within 0.02% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	**Cd_7- wMT-2**	11,382.0	11,381.0
B	**Cd_6- wMT-2**	11,271.6	11,271.4
B	**Cd_5- wMT-2**	11,161.2	11,158.8
B	**Cd_4- wMT-2**	11,050.8	11,049.6
B	**Cd_1- wMT-2**	10,719.5	10,719.3
B	**apo- wMT-2**	10,609.1	10,606.8

Appendix 10

Referred from: Chapter 3, Section 3.10

Appendix 10. For the indicated peaks, in Figure 3.39, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.1%, most within 0.02% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	Cd_7- wMT-2	8,570.8	8,570.3
A	Cd_4- wMT-2	8,239.6	8,240.3
B	Cd_4- wMT-2	8,239.6	8,239.1
B	apo- wMT-2	7,797.9	7,796.4

Appendix 11

Referred from: Chapter 3, Section 3.11

Appendix 11. For the indicated peaks, in Figure 3.41, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.11%, most within 0.05% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	**Zn_7- wMT-2**	11,052.8	11,051.0
A	**Zn_6- wMT-2**	10,989.4	10,984.8
A	**Zn_5- wMT-2**	10,926.1	10,923.9
B	**Zn_4- wMT-2**	10,862.7	10,870.6
B	**Zn_3- wMT-2**	10,799.3	10,788.1
B	**Zn_2- wMT-2**	10,735.9	10,731.7
B	**Zn_1- wMT-2**	10,672.5	10,669.5
B	**apo- wMT-2**	10,609.1	10,607.1

Appendix 12

Referred from: Chapter 3, Section 3.11

Appendix 12. For the indicated peaks, in Figure 3.42, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.11%, most within 0.05% of their theoretical counterparts. A, B designates the spectrum from which the readings are observed.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	**Zn_7- wMT-2**	8,241.7	8,238.7
A	**Zn_6- wMT-2**	8,178.3	8,170.4
A	**Zn_5- wMT-2**	8,114.9	8,105.3
B	**Zn_1- wMT-2**	7,861.3	7,857.8
B	**apo- wMT-2**	7,797.9	7,796.4

Appendix 13

Referred from: Chapter 3, Section 3.12

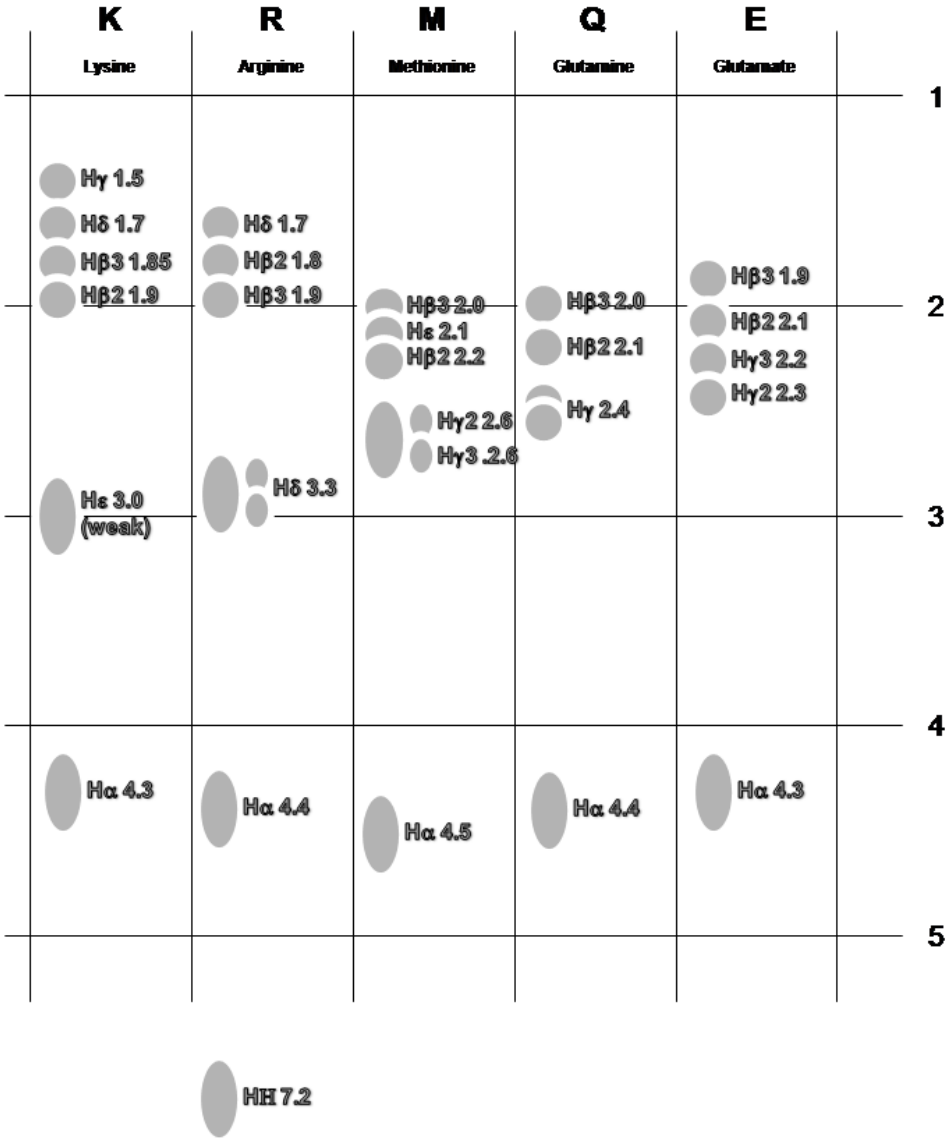
Appendix 13. For the indicated peaks, in Figure 3.44, the theoretical and observed masses for all metallated species. All masses have been corrected from +1 to neutral, by subtracting 1Da. All observed masses are within 0.14% of their theoretical counterparts. A, B, designates the spectrum from which the readings are observed. Calculated masses are without N-terminal methionine.

	Metallospecies	Calculated Mass (Da)	Observed Mass (Da)
A	Cd_7- wMT-3	9,420.9	9,420.7
B	Cd_4- wMT-3	9,089.7	9,088.5
B	apo- wMT-3	8,648.0	8,646.9

Appendix 14

Referred from: Chapter 4, Section 4.2

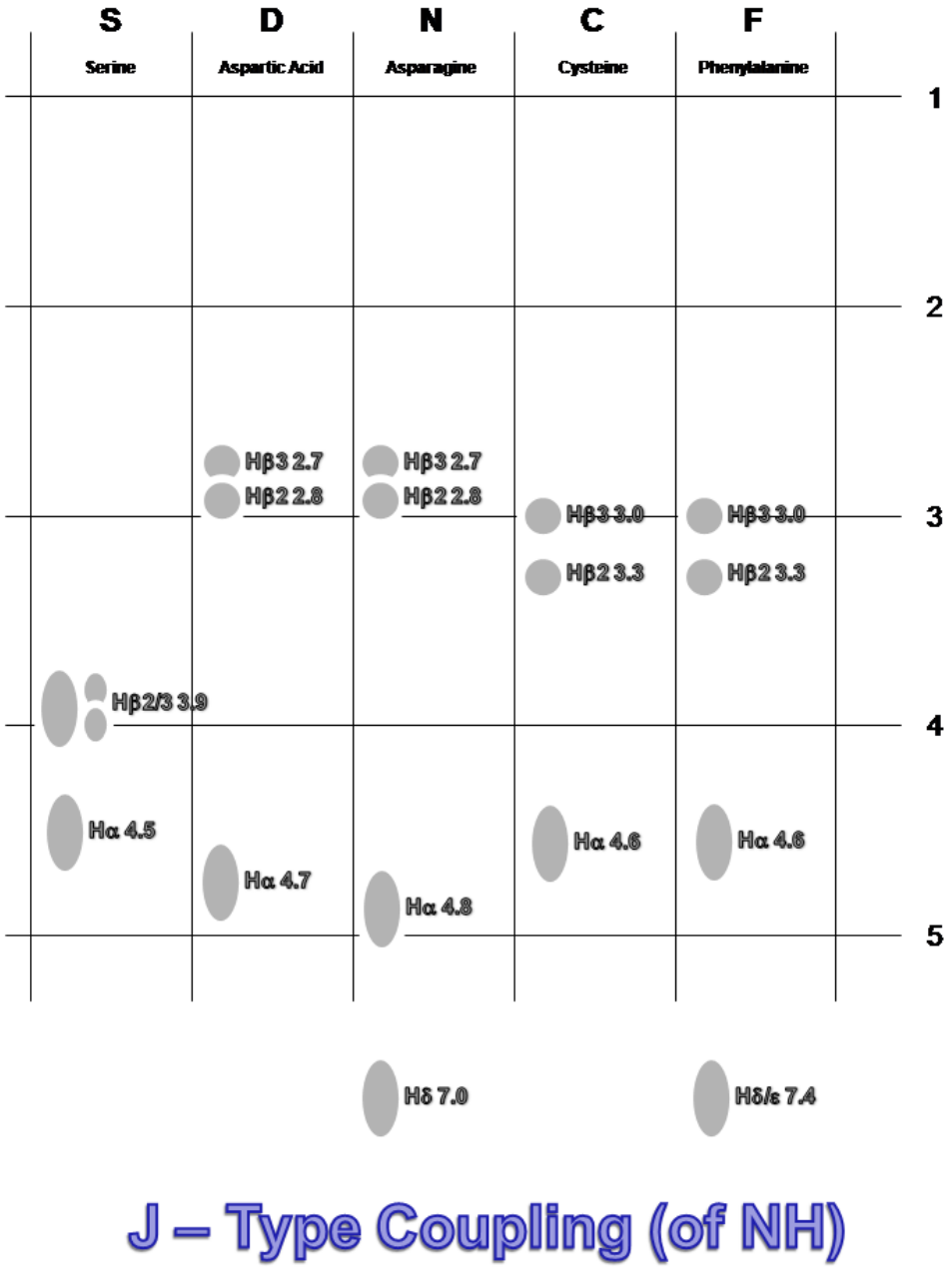
Appendix 14. Spin systems of the U-type amino acids. The proline spin system has been characterised as a 'special' system, and is presented later.



Appendix 14

Referred from: Chapter 4, Section 4.2

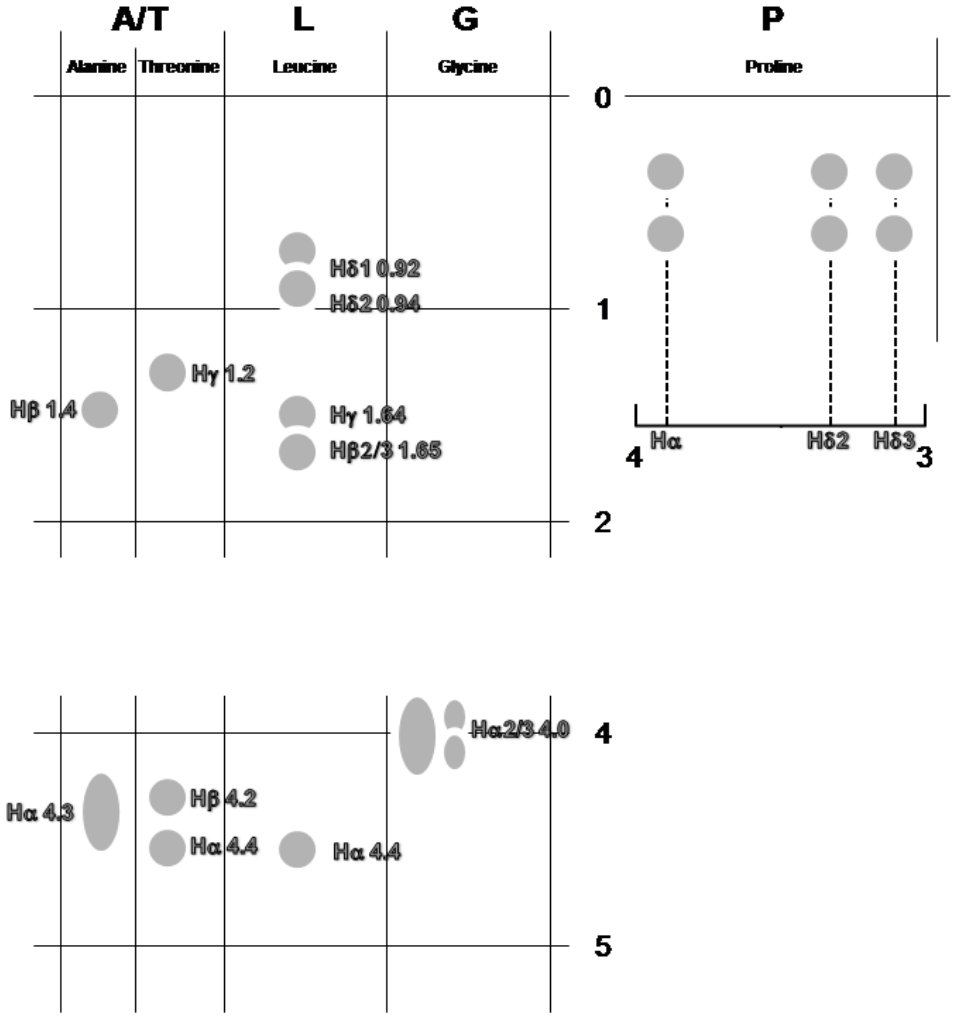
Appendix 14. Spin systems of the J-type amino acids. As H / Y / W residues are not present in the wMT-2 sequence, they have not been included.



Appendix 14

Referred from: Chapter 4, Section 4.2

Appendix 14. Spin systems of the remaining amino acids A/T-, L-, G- and P-type. As V / I residues are not present in the wMT-2 sequence, they have not been included.

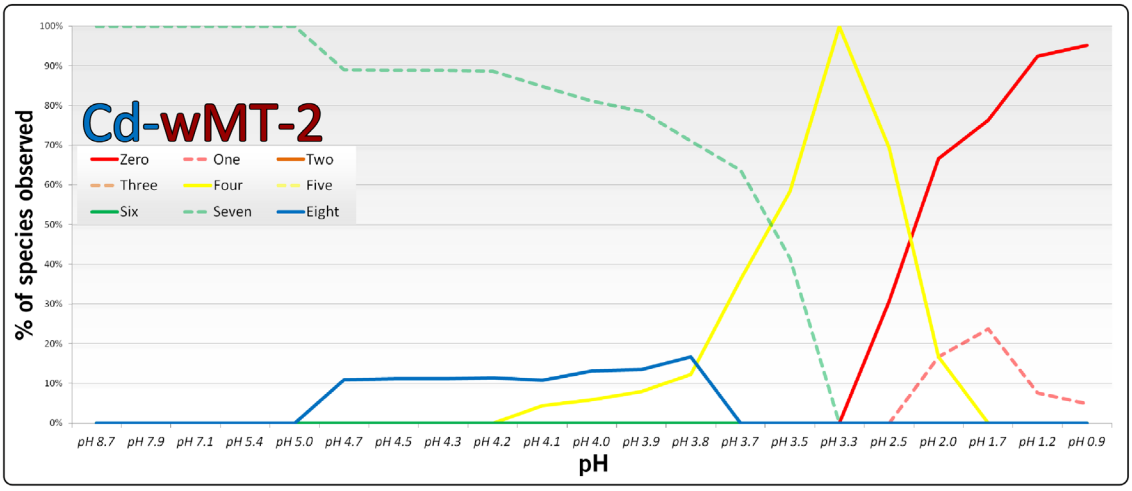


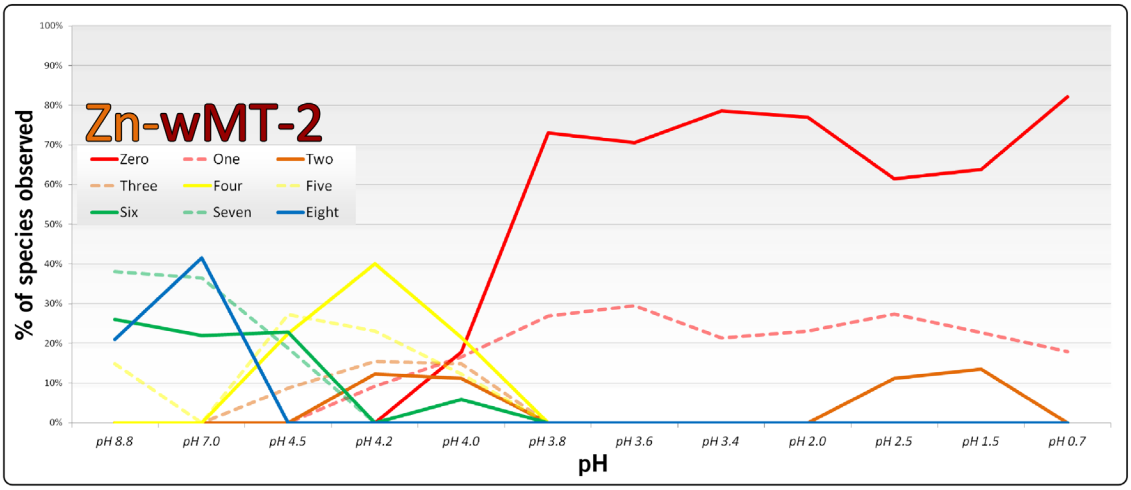
A/T; L; G; P – Type Coupling (of NH)

Appendix 15

Referred from: Chapter 5, Section 5.2

Appendix 15. Comparison of the number of metal ions bound at decreasing pH. Of note are lifetimes of the species present at various pH. **TOP** is the analysis of Cd-wMT-2, **BOTTOM** is the analysis of Zn-wMT-2. Spectra analysed are displayed in Figure 5.1. Number of metal ions bound are indicated in **RAINBOW** colours, for ease of viewing, alternating dashed and solid lines.

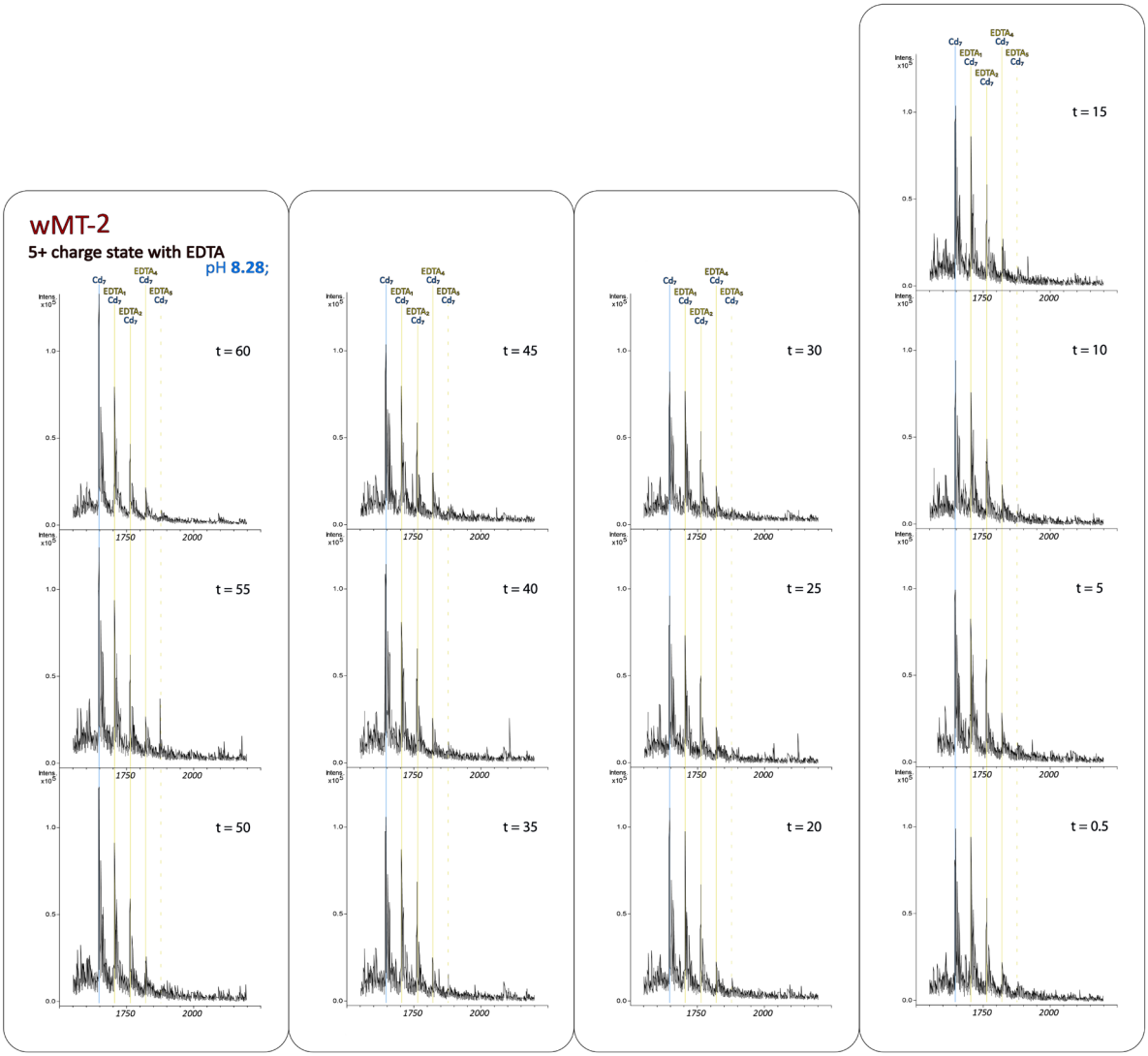




Appendix 16

Referred from: Chapter 5, Section 5.3

Appendix 16. Focussed region between 1,500 *m/z* and 2,500 *m/z*, showing the 5+ charge state of wMT-2. All spectra have been scaled identically, so changes in their peak intensities can be roughly compared. 47.2 µM wMT-2 in 20 mM ammonium bicarbonate, 10 % MeOH. Stoichiometry 6.8:1 cadmium ions per wMT-2 molecule. Times are displayed in minutes from first EDTA addition.



Appendix 17

Referred from: Chapter 5, Section 5.5

Appendix 17. Comparing the theoretical calculated masses of potential states of wMT-2 when metallated with both zinc and cadmium. Observed mass is the experimental mass – 1Da to give a neutral mass. % calculated from the intensity of the respective peaks from MS in Figure 5.9.

Metallospecies	Calculated Mass (Da)	Observed Mass (Da)	% of sample
Cd_7- wMT-2	11,382.1	11,381.4	12.7
Zn_1Cd_6- wMT-2	11,335.0	11,335.1	19.0
Zn_2Cd_5- wMT-2	11,288.0	11,287.0	20.4
Zn_3Cd_4- wMT-2	11,241.0	11,241.0	21.2
Zn_4Cd_3- wMT-2	11,194.0	11,192.6	17.3
Zn_5Cd_2- wMT-2	11,146.9	11,144.6	9.3
Zn_6Cd_1- wMT-2	11,099.9	-	
Zn_7 - wMT-2	11,052.9	-	
apo- wMT-2	10,609.2	-	

Appendix 18

Appendix 18. Chemical shift list of atoms assigned by NMR for cleaved Cd-wMT-2.

	H_N	N	Nδ	Nε	Hα	Hβ	Hγ	Hδ	Hε	Cα
2 Ser	-	-	-	-	-	-	-	-	-	-
3 Met	-	-	-	-	-	-	-	-	-	59.18
4 Ala	9.22	135.77	-	-	4.56	1.04	-	-	-	47.72
5 Asp	8.92	127.75	-	-	4.16	3.04,2.76	-	-	-	57.08
6 Ala	-	-	-	-	4.11	0.95	-	-	-	51.24
7 Phe	7.86	111.38	-	-	4.54	2.85,3.10	-	7.18	7.26	54.39
8 Asn	8.11	119.68	112.06	-	4.75	2.20,2.36	-	7.17,6.53	-	49.35
9 Thr	8.75	109.59	-	-	4.04	-	-	-	-	58.39
10 Gln	8.09	119.93	-	112.08	4.17	2.50	2.30,2.24	-	6.72,7.14	55.15
11 Cys	7.95	114.69	-	-	4.29	3.36,2.93	-	-	-	55.23
12 Cys	7.48	112.35	-	-	4.80	3.07,3.12	-	-	-	54.01
13 Gly	7.63	113.38	-	-	3.81	-	-	-	-	44.30
14 Asn	6.99	115.41	-	-	5.08	2.89,2.80	-	-	-	48.47
15 Lys	-	-	-	-	-	-	-	-	-	56.17
16 Thr	7.93	118.73	-	-	4.39	4.01	1.18	-	-	59.42
17 Cys	9.07	127.58	-	-	3.90	2.94,3.05	-	-	-	59.28
18 Pro	-	-	-	-	4.33	1.98,2.24	-	3.47,3.88	-	60.21
19 Arg	8.22	118.67	-	-	4.80	1.60,1.37	1.26	-	-	52.07
20 Glu	8.41	123.02	-	-	4.35	1.70,1.81	2.03	-	-	52.33
21 Gly	-	-	-	-	-	-	-	-	-	-
22 Ser	7.81	-	-	-	4.20	3.96,3.70	-	-	-	56.51
23 Thr	7.30	107.48	-	-	4.48	4.06	0.99	-	-	56.28
24 Cys	8.40	124.82	-	-	4.82	2.52,3.30	-	-	-	57.33
25 Ala	8.30	-	-	-	4.40	-	-	-	-	-
26 Cys	7.84	-	-	-	4.53	3.60,3.83	-	-	-	54.14
27 Ser	7.05	124.22	-	-	4.01	2.96,2.84	-	-	-	58.32
28 Lys	7.41	119.78	-	-	4.26	1.50	-	-	-	53.12
29 Cys	7.26	123.63	-	-	4.50	3.14,3.50	-	-	-	56.27
30 Arg	9.12	131.95	-	-	4.38	1.75,1.90	1.61	-	-	53.42
31 Cys	8.51	126.09	-	-	4.10	2.87,3.00	-	-	-	57.59
32 Pro	-	-	-	-	4.35	-	-	3.79	-	59.68
33 Lys	8.61	121.15	-	-	3.91	1.77	-	-	-	55.22
34 Asp	8.12	114.82	-	-	4.19	2.77	-	-	-	50.43
35 Asp	7.90	120.34	-	-	4.70	2.48,2.19	-	-	-	50.12
36 Cys	8.13	123.21	-	-	4.50	2.46,2.73	-	-	-	56.62
37 Ala	8.77	132.70	-	-	4.52	1.33	-	-	-	48.68
38 Pro	-	-	-	-	3.96	2.09,2.20	1.68	3.51	-	62.14
39 Asn	8.03	115.22	-	-	4.54	-	-	-	-	51.03
40 Cys	7.61	121.28	-	-	3.96	3.66,2.78	-	-	-	59.46
41 Lys	8.46	128.21	-	-	4.37	2.00	1.39	-	-	50.99
42 Lys	9.42	124.55	-	-	4.24	1.81	1.27	1.63	-	53.56
43 Leu	8.77	116.86	-	-	3.69	1.53,2.47	1.63	0.98,0.87	-	53.11
44 Cys	8.48	122.81	-	-	4.41	2.93,3.08	-	-	-	55.25

Appendix 18

	H_N	N	Nδ	Nε	Hα	Hβ	Hγ	Hδ	Hε	Cα
45 Cys	8.84	123.14	-	-	4.16	3.40,3.32	-	-	-	58.35
46 Ala	6.82	121.14	-	-	3.88	1.28	-	-	-	49.84
47 Asp	7.94	117.26	-	-	4.43	2.58,2.37	-	-	-	51.37
48 Ala	8.29	124.42	-	-	4.17	1.14	-	-	-	49.19
49 Gln	8.27	121.63	-	112.11	4.20	2.05,2.19	1.89	-	6.64,7.42	53.71
50 Cys	8.09	117.81	-	-	4.96	2.63,2.95	-	-	-	57.87
51 Gly	7.88	110.87	-	-	4.29,4.19	-	-	-	-	42.76
52 Asn	8.63	122.26	111.20	-	4.80	2.84,3.16	-	6.95,6.82	-	48.52
53 Ala	-	-	-	-	-	-	-	-	-	49.52
54 Ser	7.88	118.60	-	-	4.84	-	-	-	-	54.62
55 Cys	-	-	-	-	-	-	-	-	-	-
56 Ser	-	-	-	-	4.19	4.02	-	-	-	58.69
57 Cys	8.53	123.21	-	-	3.86	3.42,2.62	-	-	-	60.09
58 Gly	7.43	106.53	-	-	3.69,3.82	-	-	-	-	43.29
59 Ala	-	-	-	-	-	-	-	-	-	51.24
60 Ala	7.45	117.43	-	-	4.28	1.36	-	-	-	48.90
61 Cys	7.39	121.28	-	-	4.15	3.18,2.88	-	-	-	59.45
62 Lys	9.80	131.96	-	-	4.33	1.97,2.02	1.47	1.58	-	53.24
63 Cys	8.59	125.82	-	-	3.77	2.70,2.97	-	-	-	60.66
64 Ala	8.17	123.92	-	-	3.91	1.24	-	-	-	49.14
65 Ala	-	-	-	-	-	-	-	-	-	-
66 Gly	7.85	111.32	-	-	4.04	-	-	-	-	43.13
67 Ser	8.49	115.59	-	-	4.54	3.92,3.67	-	-	-	55.83
68 Cys	8.21	128.03	-	-	4.63	3.13,2.42	-	-	-	55.89
69 Ala	8.41	131.49	-	-	4.32	1.27	-	-	-	48.18
70 Ser	-	-	-	-	-	-	-	-	-	-
71 Gly	-	-	-	-	-	-	-	-	-	42.47
72 Cys	7.74	122.89	-	-	3.63	3.73,2.76	-	-	-	59.54
73 Lys	8.40	127.67	-	-	4.25	2.01	1.41	-	-	51.18
74 Lys	9.25	125.58	-	-	4.14	1.61,1.53	1.32	-	-	53.40
75 Gly	8.32	108.59	-	-	3.51,3.95	-	-	-	-	44.19
76 Cys	8.17	129.57	-	-	4.26	2.99,3.15	-	-	-	56.20
77 Cys	7.84	121.42	-	-	4.06	3.14,3.17	-	-	-	58.75
78 Gly	7.37	106.76	-	-	3.89,3.45	-	-	-	-	42.43
79 Asp	7.42	125.51	-	-	4.30	2.57,2.42	-	-	-	53.27

Appendix 19

Appendix 19. Alignment of the Sea Urchin MTA sequence with Earthworm wMT-2. Alignment performed in ClustalW2 (http://www.ebi.ac.uk/Tools/msa/clustalw2/).