Biophysical Characterization of the First Four Metal-Binding Domains of Human Wilson Disease Protein

Alia V.H. Hinz

Follow this and additional works at: https://scholarworks.wmich.edu/dissertations

Part of the Chemistry Commons, and the Medicinal and Pharmaceutical Chemistry Commons

Recommended Citation

https://scholarworks.wmich.edu/dissertations/280
BIOPHYSICAL CHARACTERIZATION OF THE FIRST FOUR METAL-BINDING DOMAINS OF HUMAN WILSON DISEASE PROTEIN

by

Alia V.H. Hinz

A dissertation submitted to the Graduate College in partial fulfillment of the requirements for the degree of Doctor of Philosophy Department of Chemistry Western Michigan University June 2014

Doctoral Committee:

David Huffman, Ph.D., Chair Blair Szymczyna, Ph.D Gellert Mezei, Ph.D Douglas Coulter, Ph.D
Wilson disease protein (WLNP) is a P$_{1b}$-type ATPase crucial for maintaining copper homeostasis in humans. Mutations in this protein result in the autosomal recessive disorder Wilson disease, a condition characterized by copper accumulation in the liver and brain. WLNP provides copper for incorporation into cuproproteins and exports excess copper into the bile for excretion. There are six metal-binding domains (MBDs) in WLNP, found within the first 650 amino acids of this 1,465 amino acid protein. Though each MBD has a different amino acid sequence, all MBDs possess a similar ferredoxin fold with a conserved hydrophobic core and a MXCXXC metal-binding motif.

The manner in which the six MBDs communicate with each other and how they affect other cytosolic-facing domains of WLNP is not understood. There is a long linker between the fourth and fifth MBDs that provides spatial separation between the first four and the last two MBDs. To better understand how the first four MBDs function, I pursued a detailed biophysical characterization of these domains. Strikingly, when MBD4 is expressed by itself, it is highly resistant to both chemical and thermal denaturation: 50% of its structure is retained in 5.9 M guanidine hydrochloride (GuHCl) and it has a melting temperature of 78°C. In contrast, when MBDs1-3 are expressed
together as a single protein, 50% of its structure is retained at 2.3 M GuHCl and the melting temperature is 58°C. Furthermore, the unusual stability of MBD4 is preserved when it is expressed in a protein construct that contains all four MBDs (MBDs1-4). In MBDs1-4, denaturation by GuHCl produced a double sigmoidal curve in which the second sigmoid correlated with that of MBD4 while the first one correlated with MBDs1-3. MBD4 also influenced the thermal denaturation of MBDs1-4, albeit in a more complicated manner. MBDs1-4 did not display a thermal unfolding transition and instead underwent a structural rearrangement that resulted in soluble aggregation.

Though the extreme conditions that MBD4 can withstand are never experienced in vivo, the enhanced stability of this domain may play a role in its ability to serve as a primary target of copper acquisition.
I am deeply indebted to many people who made it possible for me to complete this degree.

I thank my advisor, Dr. David Huffman for giving me the opportunity to do research in his lab, a place where I learned a great deal and sharpened my skills. We also had many useful conversations that helped me with my project, for which I am most appreciative. I also thank the rest of my committee: Dr. Blair Szymczyna, Dr. Gellert Mezei, and Dr. Douglas Coulter for their time and guidance. Dr. Szymczyna’s patience in explaining NMR and showing me how to process data was invaluable. Additionally, my lab colleagues were most helpful and I appreciate their insights. The friendship and support of the other graduate students in the department was also most beneficial, and I thank them. I am also grateful to both Western Michigan University and the Department of Chemistry for accepting me as a student and providing me with both funding and a place to learn and grow. The faculty and staff in the Department of Chemistry were also supportive and helpful.

My family was instrumental in my success, and I am profoundly grateful for the unwavering support and encouragement of my parents, brother, and sister-in-law. I dearly thank my fiancé, whose patience and encouragement made it possible for me to finish my degree with my sanity relatively intact.
To the many friends whose names go unmentioned here, please know that you are not forgotten. Your presence was most appreciated, and I am very grateful for your support.

Alia V. H. Hinz
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ...................................................................................... ii  
LIST OF TABLES ................................................................................................ viii  
LIST OF FIGURES ................................................................................................ x  
LIST OF ABBREVIATIONS ................................................................................. xiv  

CHAPTER  
1. COPPER HOMEOSTASIS ........................................................................ 1  
1.1 Eukaryotic Copper Requirements ...................................................... 1  
1.2 The Need for Copper Homeostasis .................................................... 2  
1.3 Copper Movement in Eukaryotes ...................................................... 4  
1.4 Wilson Disease Protein ...................................................................... 8  
1.4.1 Structural Features .................................................................... 8  
1.4.2 Copper Acquisition by the Metal-Binding Domains ................ 18  
1.4.3 Regulation ................................................................................. 25  
1.5 Research Objectives ....................................................................... 26  
References ....................................................................................................... 28  
2. MATERIALS AND EXPERIMENTAL METHODS ............................... 36  
2.1 Introduction ........................................................................................ 36  
2.2 Reagents and Instrumentation ............................................................ 36  
2.2.1 Chemical and Biological Reagents ........................................... 36
CHAPTER

2.2.2 Instrumentation ................................................................................................ 39

2.3 Summary of Plasmids and Protein Constructs.................................................. 40

2.3.1 Plasmids Containing Portions of the Wilson Disease Protein ..................... 40

2.3.2 HAH1 Plasmid ............................................................................................... 43

2.4 Construction of Plasmids, Protein Expression, and Protein Purification.................. 43

2.4.1 Preparation of the pET32-WLN1-2, pET32-WLN1-3, pET32-WLN1-4, and pET32-WLN1-4 G333R Plasmids ................................................................. 44

2.4.2 Preparation of the pET32-WLN4 Plasmid..................................................... 49

2.4.3 Transformation of Bacterial Cells for Plasmid Propagation............................ 52

2.4.4 Protein Expression ....................................................................................... 53

2.4.4.1 Transformation of Rosetta™ 2 (DE3) Cells for Protein Expression............. 54

2.4.4.2 Non-labeled Protein Expression ............................................................... 55

2.4.4.3 $^{15}$N and $^{13}$C-$^{15}$N Labeled Protein Expression..................................... 56

2.4.5 Protein Extraction and Purification............................................................... 59

2.4.5.1 Extraction for the WLN Constructs ......................................................... 59

2.4.5.2 Purification of the WLN Constructs ......................................................... 60

2.4.5.3 Metallation of the WLN Constructs ......................................................... 63

2.4.5.4 Extraction and Purification of HAH1 ...................................................... 64

2.5 Biophysical Characterization ............................................................................ 65

2.5.1 High Resolution Gel Filtration on the WLN Constructs ............................... 66
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5.2</td>
<td>Gel Filtration Analysis of the Association State of HAH1</td>
<td>68</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Light Scattering</td>
<td>69</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Circular Dichroism</td>
<td>73</td>
</tr>
<tr>
<td>2.5.5</td>
<td>Differential Scanning Calorimetry</td>
<td>81</td>
</tr>
<tr>
<td>2.5.6</td>
<td>Nuclear Magnetic Resonance Spectroscopy</td>
<td>83</td>
</tr>
<tr>
<td>2.6</td>
<td>Summary</td>
<td>86</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>88</td>
</tr>
<tr>
<td>3.</td>
<td>PLASMID PREPARATION, PROTEIN EXPRESSION, AND PURIFICATION OF THE WLN CONSTRUCTS</td>
<td>91</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>91</td>
</tr>
<tr>
<td>3.2</td>
<td>Plasmid Preparation</td>
<td>91</td>
</tr>
<tr>
<td>3.3</td>
<td>Protein Expression</td>
<td>96</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Expression of Non-labeled Proteins</td>
<td>97</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Expression of $^{15}$N and $^{13}$C,$^{15}$N Labeled Proteins</td>
<td>98</td>
</tr>
<tr>
<td>3.4</td>
<td>Protein Extraction and Purification</td>
<td>102</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Protein Extraction</td>
<td>102</td>
</tr>
<tr>
<td>3.4.2</td>
<td>Protein Purification</td>
<td>102</td>
</tr>
<tr>
<td>3.5</td>
<td>Summary</td>
<td>108</td>
</tr>
<tr>
<td>References</td>
<td></td>
<td>110</td>
</tr>
<tr>
<td>4.</td>
<td>THE BIOPHYSICAL CHARACTERIZATION OF THE WLN CONSTRUCTS</td>
<td>111</td>
</tr>
<tr>
<td>4.1</td>
<td>High Resolution Gel Filtration</td>
<td>111</td>
</tr>
</tbody>
</table>
# Table of Contents—Continued

## CHAPTER

4.2 Light Scattering Analysis ................................................................. 117
4.3 Circular Dichroism ................................................................. 124
  4.3.1 Guanidine Hydrochloride Denaturation ......................... 124
  4.3.2 Thermal Denaturation ......................................................... 130
4.4 Differential Scanning Calorimetry ........................................... 143
4.5 NMR ..................................................................................... 147
  4.5.1 $^{1}$H-$^{15}$N HSQCs of WLN1-2 and WLN1-4 ............. 147
  4.5.2 $^{1}$H-$^{15}$N HSQC of WLN4 at Different Temperatures ... 149
4.6 Summary ................................................................................ 154
4.7 Future Work .......................................................................... 161

References .................................................................................... 163

5. THE EFFECTS OF PH AND HG(II) BINDING ON THE
   ASSOCIATION STATE OF HAH1 ........................................ 166
  5.1 Introduction ........................................................................... 166
  5.2 Analytical Gel Filtration ......................................................... 169
  5.3 Discussion ............................................................................ 173
  5.4 Conclusion .......................................................................... 176

References .................................................................................... 177

APPENDIX ..................................................................................... 179
# LIST OF TABLES

1.1 Some copper containing enzymes in humans and their functions .......... 1  
1.2 Human metallochaperones and copper pumps ........................................ 2  
2.1 Chemical reagents used ........................................................................ 37  
2.2 Biological reagents and kits used in this research .................................. 38  
2.3 Instruments used .................................................................................... 39  
2.4 Columns used for the purification of the WLN constructs ...................... 40  
2.5 Plasmid names and the metal-binding domains they contained .............. 41  
2.6 The mutagenic PCR primers used to prepare the plasmids discussed above .............................................................................................................. 48  
2.7 PCR reaction contents for site directed mutagenesis ............................ 48  
2.8 PCR thermocycler conditions ................................................................ 49  
2.9 The PCR primers used to amplify the fourth metal-binding domain. ....... 50  
2.10 PCR reaction contents for preparing the WLN4 plasmid ........................ 50  
2.11 PCR thermocycler conditions ................................................................. 51  
2.12 Recipe for M9 minimal media ................................................................. 57  
2.13 The protein standards used to construct the calibration curves for HRGF .... 68  
2.14 Definition of the variables used in Equation 2.3 .................................... 72  
2.15 Summary of NMR experiments done with WLN1-2, WLN1-4, and WLN4 85  
4.1 Details of the calibration curve that was used to analyze WLN1-2............ 112  
4.2 Details of the calibration curve that was used to analyze WLN1-3 and WLN1-4 .............................................................................................................. 113
4.3 Summary of the biophysical parameters of the constructs that were determined by light scattering ........................................................................ 122
4.4 The denaturation of the Wilson disease protein constructs by guanidine hydrochloride ................................................................................................ 129
4.5 Thermal unfolding of the apo-WLN constructs ........................................................ 132
4.6 Biophysical parameters of the metal-binding domains. ........................................ 157
5.1 Table of protein standards and HAH1 elution volumes at pH 7.5 .................. 170
5.2 Table of protein standards and HAH1 elution volumes at pH 8.5 ................. 170
5.3 Table of protein standards and HAH1 elution volumes at pH 9.4 ................. 170
5.4 Summary of the masses of apo- and Hg(II)-HAH1 at various pHs ............... 173
A.1 $^1$H, $^{15}$N, and $^{13}$C resonance assignments for apoWLN4 at 37°C .......... 181
LIST OF FIGURES

1.1 Intracellular copper transport in eukaryotes .......................................................... 7
1.2 Schematic of the Wilson disease protein and its acquisition of copper ............ 11
1.3 Post-Albers scheme of the catalytic cycle of the Wilson disease protein, adapted from Banci et al. (50) ................................................................. 12
1.4 The NMR solution structure of the third metal-binding domain of Wilson disease protein, modified from PDB 2ROP .................................................... 14
1.5 A cartoon model of CopA (apo form) from *L. pneumophila*, modified from structure 3RFU from the PDB (62) .......................................................... 15
1.6 Conservation of residues among CopA proteins (62) ......................................... 17
1.7 Potential interaction between WLN4 and HAH1 (81) ....................................... 24
1.8 Diagram of a potential mechanism of copper transfer from HAH1 to WLN4 ........................................................................................................... 24
2.1 The amino acid sequence of WLN1-2 .............................................................. 41
2.2 The amino acid sequence of WLN1-3 .............................................................. 41
2.3 The amino acid sequence of WLN1-4 and WLN1-4 G333R .......................... 42
2.4 The amino acid sequence of WLN1-6 .............................................................. 42
2.5 The amino acid sequence of WLN4 ................................................................. 43
2.6 The amino acid sequence of HAH1 ................................................................. 43
2.7 The vector map of the pET-32Xa/LIC plasmid (EMD Millipore, Technical Bulletin 186VM) ................................................................. 45
2.8 Diagram of the WLN protein constructs, which were expressed as fusion proteins with thioredoxin (TRX) ................................................................. 45
2.9 The series of starter cultures grown to express $^{15}$N and $^{13}$C, and $^{15}$N labeled proteins ................................................................. 58
List of Figures—Continued

2.10 The purification scheme used to purify the WLN constructs ...................... 60

2.11 Diagram of a generic dipeptide with each atom labeled ......................... 86

3.1 Agarose gel (0.8%) of the pET32-WLN1-2, pET32-WLN1-3, and pET32-WLN1-4 G333R plasmids that were digested with BamHI (30 U) and BglII (20 U) ......................................................................................................................................................................................... 93

3.2 Agarose gel (1%) of the pET32-WLN4 plasmid digested with BamHI (30 U) and BglII (20 U) ......................................................................................................................................................................................... 96

3.3 Temperature induction study of WLN1-4 grown in minimal media containing unlabeled precursors ................................................................................................................................. 99

3.4 Time course induction study at 22°C for the expression of WLN1-4 in minimal media that contained 13C-glucose and 15N-ammonium chloride .......................................................... 100

3.5 Time course induction study at 22°C for the expression of WLN1-4 in LB media ........................................................................................................................................................................... 101

3.6 The purification scheme used to purify the WLN constructs ...................... 102

3.7 A representative chromatogram of the purification of the fusion proteins from the cell lysates using the HisPrep column ............................................................................................................................................. 103

3.8 A representative chromatogram of the purification of the TEV digest of the fusion proteins on the HisPrep column ............................................................................................................................................. 104

3.9 An overlay of the chromatograms for the constructs purified using the Superdex 75 pg column ............................................................................................................................................. 105

3.10 A chromatogram of WLN1-6 that was purified using a Superdex 200 pg column ............................................................................................................................................. 106

3.11 SDS-PAGE Gel of various WLN constructs ............................................. 107

3.12 SDS-PAGE Gel of WLN4 ........................................................................ 108

4.1 Calibration curve for the data presented in Table 4.1 ................................. 112

4.2 Calibration curve for the data presented in Table 4.2 ................................. 113

4.3 HRGF interaction study between WLN1-3 and WLN4 .............................. 116
List of Figures—Continued

4.4 Elution profiles of WLN1-3, WLN1-4, and WLN4 on the Superdex 75 HRF column ................................................................. 117

4.5 Molar mass distribution of WLN1-2 (red trace) injected onto a Superdex 75 HRF 10/300 column ..................................................... 119

4.6 Molar mass distribution of WLN1-3 (red trace) and WLN1-4 (green trace) that were injected onto a Superdex 200 HRF 10/300 column ..................................................... 120

4.7 The autocorrelation function for WLN1-2 ................................................................. 121

4.8 The autocorrelation function for WLN1-3 ................................................................. 121

4.9 The autocorrelation function for WLN1-4 ................................................................. 122

4.10 Guanidine hydrochloride denaturation of WLN1-2, WLN1-3, WLN1-4, and WLN4, respectively ..................................................... 126

4.11 Guanidine hydrochloride denaturation of the mutant WLN1-4 G333R ...... 127

4.12 Thermal unfolding curves for the WLN constructs ............................................. 131

4.13 The thermal unfolding of WLN4 (forward) and the renaturation of the protein as it was cooled (reverse) ................................................................. 133

4.14 CD spectra of WLN4 that was collected prior to the application of the thermal gradient (green line) and after the application of the reverse gradient, i.e. cooling it (orange line) ................................................................. 133

4.15 Scanning wavelength spectra of WLN1-4 were collected before and after the application of the thermal gradient ................................................................. 135

4.16 A chromatogram of WLN1-4 before and after the application of the thermal gradient ................................................................. 136

4.17 The unfolding of WLN1-4 in the presence of various concentrations of sodium chloride ........................................................................ 138

4.18 WLN1-4 with 150 mM sodium chloride scanning wavelength spectra before and after thermal denaturation ................................................................. 138

4.19 The effect of copper(I) loading on the CD spectra of WLN1-4. ..................... 140
List of Figures—Continued

4.20 Titration of WLN4 with copper ................................................................. 141
4.21 The thermogram showing the thermal denaturation of apoWLN4 .......... 145
4.22 The thermogram showing the thermal denaturation of copper-loaded WLN4 ................................................................. 146
4.23 The thermogram of apo-WLN1-4 ............................................................... 147
4.24 $^1$H-$^{15}$N HSQC overlay of apoWLN1-2 (shown in red) and apoWLN1-4 (shown in blue). ................................................................. 148
4.25 $^1$H-$^{15}$N HSQC of apoWLN4, recorded at 25°C (red) and 50°C (black) .... 151
4.26 $^1$H-$^{15}$N HSQC of apoWLN4 with peak assignment .............................. 152
4.27 An enlarged view of a portion of Figure 4.26 ............................................. 153
4.28 Sequence alignment of the 6 MBDs ......................................................... 162
5.1 Chromatograms of apo-HAH1 at various pHs (11) ................................. 171
5.2 Chromatograms of Hg(II)-HAH1 at various pHs (11) ............................. 171
5.3 $^{199}$Hg NMR of Hg(II)-HAH1 at various pHs (11) ................................. 174
5.4 $^{199m}$Hg PAC results for HAH ................................................................. 175
A.1 Thermal unfolding of WLN1-6 ................................................................. 179
A.2 CD spectra of HAH1 at pH 9.4 ................................................................. 180
A.3 CD spectra of HAH1 at pH 8.5 ................................................................. 180
A.4 Project approval certification ................................................................. 183
A.5 Permission from John Wiley for use of Figures 5.1 – 5.4 ................. 184
A.6 Permission from Nature Publishing Group for use of Figure 1.4 ........ 185
A.7 Permission from Springer for use of Figure 1.7 ................................. 186
LIST OF ABBREVIATIONS

\( \theta_{MRW} \) ........................................................................................................... Mean Residue Molar Ellipticity

ATP ................................................................................................................ Adenosine Triphosphate

\( \beta \)-ME .................................................................................................... \( \beta \)-mercaptoethanol

bp......................................................................................................................... Base Pairs

BCA ............................................................................................................... Bicinchoninic Acid

BCS ........................................................................................................ Bathocuproine Disulfonate

BSA ........................................................................................................ Bovine Serum Albumin

CD ........................................................................................................ Circular Dichroism

CHES ............................................................... N-Cyclohexyl-2-aminoethanesulfonic Acid

DLS .......................................................................................... Dynamic Light Scattering

DSC ................................................................................. Differential Scanning Calorimetry

DSS ............................................................Sodium 2,2-dimethyl-2-silapentane-5-sulfonate

DTT ........................................................................................................ Dithiothreitol

E. coli ..................................................................................................... Escherichia coli

EDTA .......................................................... Ethylenediaminetetraacetic acid, Disodium Salt

Eq.................................................................................................................. Equivalents

FPLC ................................................................................................ Fast Protein Liquid Chromatography

GuHCl ...................................................................................................... Guandine Hydrochloride

HPLC ........................................................................................................ High Pressure Liquid Chromatography

xiv
List of Abbreviations—Continued

HRGF ................................................................. High Resolution Gel Filtration
HSQC ................................................................. Heteronuclear Single Quantum Coherence
IgG ................................................................. Immunoglobulin G
IPTG ................................................................. Isopropyl β-D-1-thiogalactopyranoside
Kb ................................................................................. Kilobase
kD ................................................................................. Kilodalton
K_D ................................................................. Dissociation Constant
LB ................................................................. Luria-Bertani
LS ................................................................................. Light Scattering
MBD ................................................................. Metal-binding Domain
MWCO ................................................................. Molecular Weight Cut Off
Ni ................................................................................. Nickel
NMR ................................................................. Nuclear Magnetic Resonance
PAC ................................................................. Perturbed Angular Correlation
PCR ................................................................. Polymerase Chain Reaction
pg ................................................................................. Prep Grade
R_g ................................................................. Radius of gyration
R_H ................................................................. Hydrodynamic Radius
ROS ................................................................. Reactive Oxygen Species
SDS-PAGE ................................................ Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
SLS ................................................................. Static Light Scattering
SOC ................................................................................ Super Optimal Broth with Glucose
TCEP ..................................................................................... tris(2-carboxyethyl)phosphine
TEV ........................................................................................................Tobacco Etch Virus
TGN ..................................................................................................... trans-Golgi Network
Tris ................................................................. Tris(hydroxymethyl)aminomethane
U ................................................................................................................. Units
WLNP ................................................................. Wilson Disease Protein
CHAPTER 1
COPPER HOMEOSTASIS

1.1 Eukaryotic Copper Requirements

Copper is an essential trace metal that is required by all eukaryotes. It is used as a catalyst and structural cofactor in proteins that have a wide range of functions, including superoxide detoxification, iron metabolism, energy production, and neurotransmitter production (Table 1.1) (1–4). Another important class of copper-containing protein transport copper around the cell to specific destinations. These proteins are known as metallochaperones, examples of which are shown in Table 1.2.

Copper is acquired through diet and is primarily found in liver, shellfish, mushrooms, nuts, and seeds. An adult human contains about 110 mg of copper, making copper one of the most abundant inorganic trace metals in the body, after iron and zinc (5, 6).

Table 1.1: Some copper containing enzymes in humans and their functions.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ceruloplasmin</td>
<td>Iron metabolism; oxidizes Fe(II) to Fe(III), which is transported by transferrin</td>
<td>(2)</td>
</tr>
<tr>
<td>Cytochrome c oxidase</td>
<td>Final enzyme in the mitochondrial electron transport chain</td>
<td>(7)</td>
</tr>
<tr>
<td>Cu/Zn superoxide dismutase (SOD)</td>
<td>Antioxidant protection; catalyzes the dismutation of superoxide to hydrogen peroxide and dioxygen</td>
<td>(8)</td>
</tr>
<tr>
<td>Dopamine β-hydrolase</td>
<td>Oxygenates dopamine to produce norepinephrine</td>
<td>(9)</td>
</tr>
<tr>
<td>Hephaestin</td>
<td>Iron metabolism; helps dietary iron cross the small intestine into the blood</td>
<td>(2)</td>
</tr>
<tr>
<td>Lysl oxidase</td>
<td>Connective tissue formation; crosslinks collagen and elastin</td>
<td>(10)</td>
</tr>
<tr>
<td>Tyrosinase</td>
<td>Involved in melanin synthesis</td>
<td>(11)</td>
</tr>
</tbody>
</table>
Table 1.2: Human metallochaperones and copper pumps.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Function</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>CCS (copper chaperone for SOD)</td>
<td>Metallochaperone; transports Cu(I) to SOD</td>
<td>(12)</td>
</tr>
<tr>
<td>COX17</td>
<td>Metallochaperone; transports Cu(I) to Sco1 and COX11 for insertion into mitochondrial cytochrome c oxidase</td>
<td>(13)</td>
</tr>
<tr>
<td>HAH1 (ATOX1)</td>
<td>Metallochaperone; transports Cu(I) to either Menkes protein or the Wilson disease protein</td>
<td>(14)</td>
</tr>
<tr>
<td>Menkes protein</td>
<td>ATPase pump; embedded in the membrane of most non-liver cells; helps to control the absorption of dietary copper; also pumps Cu(I) to the interior of the trans-Golgi network for incorporation into cuproproteins; helps to remove excess Cu(I)</td>
<td>(15)</td>
</tr>
<tr>
<td>Wilson disease protein</td>
<td>ATPase pump; embedded in the membrane of the trans-Golgi network in hepatocytes; pumps Cu(I) from the cytosol to the interior of the trans-Golgi network for incorporation into cuproproteins; helps to remove excess Cu(I)</td>
<td>(16)</td>
</tr>
</tbody>
</table>

1.2 The Need for Copper Homeostasis

While copper is necessary for vital processes and for normal growth and development, excess or unligated copper is toxic. Consequently, copper levels must be carefully regulated to ensure that adequate copper is available to meet biological demand while avoiding the hazards that copper poses.

The danger that copper presents comes from the same feature that makes copper a valuable enzymatic cofactor – its ability to participate in redox reactions. Unligated copper can produce reactive oxygen species (ROS) which have been linked to cellular aging, protein and lipid oxidation, DNA/RNA damage, and cancer (17–19). A
particularly potent ROS, the hydroxyl radical, is created when Cu(I) catalyzes the Fenton reaction, shown below in Equation 1.1 (17).

\[
\begin{align*}
\text{H}_2\text{O}_2 & \xrightarrow{\text{Fenton Reaction}} 2\text{OH}^- + \cdot\text{OH} \\
\text{Cu(I)} & \quad \text{Cu(II)}
\end{align*}
\]  

[Equation 1.1]

Another way that copper can damage cells is by displacing other metals from their binding sites in proteins. Cu(II) is near the top of the Irving-Williams series and can displace weaker binding ions such as iron (20) and zinc (21) from proteins (22).

Unligated copper poses the greatest threat; consequently, cells have developed strategies to minimize this. These strategies include the use of metallochaperones to bind copper as soon as it enters the cell and ferry it to its destination as well as meticulously regulating intracellular copper levels to prevent copper excess. Additionally, glutathione binds copper and offers another form of cellular protection (23, 24). As a result of this careful orchestration, the unligated copper concentration is quite low; Rae et al. have shown that there is less than 1 atom of free copper in the yeast cells (25), and this value is expected to be similar in human cells.

Copper’s importance as a cofactor in enzymes stems from its ability to cycle between two states – Cu(I) and Cu(II). The redox potential of both of these forms of copper are biologically accessible, with the specific redox potential being determined by the nature of the microenvironment created by the protein at the metal-binding site, the identity of the specific ligands that are supplied by the protein, and the environment the protein is in. Cu(I) can be readily oxidized during redox reactions and is typically ligated by softer ligands such as the sulfur atoms in cysteine and methionine. Cu(II), on the
other hand, can be readily reduced during redox reactions; this form of copper prefers the oxygen atoms in the side chains of aspartate or glutamate or the imidazole nitrogen of histidine. Cu(I) and Cu(II) have different geometries, with Cu(I) favoring trigonal planar and Cu(II) favoring square planar or square pyramidal geometry. Enzymes that have copper cofactors must be able to adapt to the different geometries favored by the cuprous and cupric ions that are formed during the various stages of the redox reaction (26).

1.3 Copper Movement in Eukaryotes

Copper movement is carefully orchestrated to ensure that the demand for copper is met while preventing excess and unligated copper from accumulating in the cell. This encompasses the use of copper metallochaperones to bind copper upon its entering the cell and transferring it to specific destinations as well as the removal of excess copper from the cell before the cell’s chelation capacity is reached (27, 28).

Absorption of dietary copper occurs in the small intestine (29). Copper enters the enterocytes through Ctr1, an integral membrane protein that is 190 amino acids long. It contains several distinct regions, including three transmembrane domains (with a MX$_3$M motif present in the second transmembrane region), a cluster of cysteines and histidines in the intracellular C-terminal domain and a methionine rich, extracellular N-terminal domain (30). Experiments using x-ray crystallography, cryo-electron microscopy and 2D crystallography have revealed that Ctr1 is functionally active as a homotrimer, with a cone-shaped pore at the center of the trimer through which Cu(I) traverses the membrane and enters the cell (31, 32). Ctr1 does not transfer Cu(II). As extracellular copper is usually present as Cu(II), it requires reduction prior to being transported by Ctr1.
recently identified family of reductases known as Steap proteins reduce cupric ions at the brush border of enterocytes, enabling the entry of Cu(I) into the cell via Ctr1 (33).

Upon entry into the cell, copper is acquired by various metallochaperones that transport the copper to specific cellular destinations. The exact mechanism by which the metallochaperone itself acquires the copper is not presently known. Protein-protein interactions between the metallochaperone and Ctr1 may facilitate the transfer of copper, or perhaps a small ligand like glutathione acts as an intermediary between Ctr1 and the metallochaperone, as Maryon et al. have suggested (34). At present, three metallochaperones have been identified in humans. One metallochaperone, CCS, transports copper to SOD (28). Another metallochaperone, COX17, is located in the mitochondria and transfers copper to Sco1 and COX11; Sco1 donates copper to the CuA site in the cytochrome c oxidase (CCO) complex while COX11 donates copper to the CuB site (28). How copper is transported to the mitochondria has not yet been elucidated. A third metallochaperone is HAH1, which transports copper to Wilson disease protein (WLNP) and Menkes protein (27, 35). Wilson disease protein is expressed in the liver, brain, kidney, and mammary glands while Menkes protein is expressed in all other tissue (36). A diagram of copper movement in hepatocytes is shown in Figure 1.1.

Wilson disease protein and a related protein, Menkes protein, are located in the membrane of the trans-Golgi network where secreted cuproproteins acquire copper prior to being exported. Under basal conditions, these P_{IB}-type ATPases accept copper from HAH1 and pump it across the membrane into the interior of the trans-Golgi network by hydrolyzing ATP; this is discussed in greater detail in Section 1.4. In hepatocytes,
copper is incorporated into the multi-copper ferroxidase ceruloplasmin which is then secreted into the bloodstream. Despite containing over 95% of the copper that is found in the plasma, however, ceruloplasmin does not appear to play in pivotal role in copper metabolism; mice lacking ceruloplasmin do not display abnormal copper absorption or excretion (37). Under conditions of copper excess, Wilson disease protein and Menkes protein translocate to cytosolic vesicles and migrate to the plasma membrane, where they fuse with it and transport copper from the cell (36, 38). In hepatocytes, the excess copper is transferred to the bile and excreted. In enterocytes, excess copper is translocated across the basolateral membrane and enters the bloodstream for distribution. Most of the copper ends up in the liver, as it is the major site of copper homeostasis (39).
Figure 1.1: Intracellular copper transport in eukaryotes. Copper enters the cell as Cu(I) through Ctr1, an integral membrane protein. Once in the cell, copper is immediately acquired by either a small molecular weight ligand or a metallochaperone which then ferries the copper to a specific cellular destination. (I) The transportation of copper to the mitochondria has not been elucidated. Once copper enters the mitochondria, however, it is acquired by the metallochaperone Cox17. Cox17 delivers copper to Sco1 and Cox11, which transfers copper the CuA site in Cox1 or the CuB site in Cox2 in the CCO complex. (II) HAH1 transports copper to the Wilson disease protein or Menkes protein that is embedded in the trans-Golgi network (TGN). Under basal copper conditions, these ATPases pump copper across the membrane into the interior of the trans-Golgi network (TGN). In response to excess copper, Wilson disease protein and Menkes protein translocate into vesicles and migrate to the plasma membrane, where they exports copper from the cell. In enterocytes, copper is released into the blood for delivery to other tissues. In hepatocytes, the copper goes into the bile duct for excretion. Copper excretion is impaired in Wilson disease, resulting in the accumulation of copper. (III) The metallochaperone CCS transports copper to superoxide dismutase (SOD).
Defects in Menkes protein and Wilson disease protein result in severe diseases known as Menkes disease and Wilson disease, respectively. In Menkes disease, copper is not exported from the small intestine, resulting in intestinal copper accumulation and copper deficiency in all other tissues; this disease is characterized by neurodegeneration, growth failure, kinky hair, and early death (40, 41). In contrast, Wilson disease results when copper removal is impaired, resulting in the accumulation of excess copper in the liver and brain; this disease is characterized by liver cirrhosis as well as psychological and neurological damage (42). At present, there is no cure for Menkes disease and treatment is merely supportive; the disease is often fatal in childhood (43). No cure presently exists for Wilson disease, though the condition can be managed with copper chelating drugs, dietary restrictions, and liver transplants (44). The location of the mutation in the Wilson disease protein determines the severity of the disease, with mutations in highly conserved regions resulting in a more severe form of the disease (45).

1.4 Wilson Disease Protein

1.4.1 Structural Features

Menkes protein and Wilson disease protein are both heavy metal pumps belonging to the family of P_{\text{in}}-type ATPases. These ATPases are evolutionary related among all three domains of life and contain five primary structural domains: the transmembrane domain (TM-domain), the actuator domain (A-domain) and a large, cytosolic N-terminal metal-binding domain; two other domains, the phosphorylation domain (P-domain) and the nucleotide binding domain (N-domain), are collectively referred to as the ATP-binding domain. A cartoon representation of WLNP is shown in
Figure 1.2. The mature Wilson disease protein contains 1,465 residues and has a mass of 157 kD while Menkes protein contains 1,500 amino acids and has a mass of 163 kD.

The TM-domain consists of eight helices that function to move copper across the membrane of the trans-Golgi network. There are several conserved residues in these helices, including a CPC tripeptide motif in the sixth helix, a YN dipeptide motif in the seventh helix, and a methionine and serine in the eighth helix; the necessity of some of the conserved residues have been established by site-directed mutagenesis (46). The cysteines in the CPC motif are believed to ligate copper as it moves through membrane, as altering these residues in CopA, the Wilson disease homologue found in *E. coli*, abolishes copper transport (47).

The P-domain and the N-domain are together referred to as the ATP-binding domain. The ATP-binding domain has a dual purpose – to bind and hydrolyze ATP to provide energy to move copper across the membrane as well as to phosphorylate itself as part of the catalytic cycle. An aspartate residue that is present in the conserved DKTGT motif is phosphorylated during the course of the catalytic transfer cycle, which is shown in Figure 1.3. The A-domain, located between the fourth and fifth transmembrane helices, dephosphorylates the aspartate residue in the DKTG motif to restore WLNP to its initial state. A conserved glutamate, found within a TGE tripeptide sequence, assists with the dephosphorylation.

Wilson disease protein pumps copper across the membrane of the trans-Golgi network according to the Post-Albers scheme (48, 49), as shown in Figure 1.3. There are four states that occur during the transport cycle – E1, E1P, E2P, and E2. E1 has a high copper affinity while E2 has a low copper affinity; phosphorylation is denoted by P.
Binding of copper to the metal-binding domains triggers a rearrangement that makes it possible for ATP to bind to the N-domain; the copper then moves to the CPC motif in the sixth transmembrane helix (E1) (50). The hydrolysis of ATP yields energy and a phosphate group that is then used to phosphorylate the conserved aspartate residue in the DKTG motif in the P-domain; the copper is then moved to the metal-binding site located inside the membrane; this is the E1P state. The E2P state is characterized by the release of copper into the interior of the trans-Golgi network. To conclude the cycle, the aspartate is dephosphorylated by the conserved TGE motif in the A-domain, and the protein is in the E2 state and ready to begin the cycle anew. The A-, N-, and P- domains experience large relative movements during the catalytic cycle that correlate with the movement of the transmembrane helices (50).

The foundation for understanding the catalytic transport cycle of Wilson disease protein comes from studies that have been done on SERCA1a, a PIIA-type ATPase that transports Ca(II) into the lumen of the sarcoplasmic reticulum by using energy from ATP hydrolysis (51–53). All P-type ATPases are believed to operate in the same manner as SERCA1a (54).
Figure 1.2: Schematic of the Wilson disease protein and its acquisition of copper. A, P, and N denote the actuator domain, the phosphorylation domain, and the nucleotide binding domain, respectively. The transmembrane regions are denoted by “TM” while the metal-binding domains are denoted by MBD followed by a number. The numbers between the MBDs indicate the length of the linker between the domains. The MBDs acquire Cu(I) from HAH1 and transport it to the inside of the trans-Golgi network under basal conditions. Under conditions of copper excess, WLNP relocates to cytosolic vesicles and migrates to the cell membrane to export copper from the cell.
The N terminal metal-binding domain is comprised of smaller, individual metal-binding domains, each capable of binding one atom of Cu(I) in a distorted linear motif (55, 56). The number of tandem repeats of the metal-binding domains (MBDs) varies among species. The yeast homologue, known as Ccc2, has two metal-binding domains, enabling it to bind only one atom of Cu(I) per domain (57). Human Wilson disease protein and Menkes protein each possess six metal-binding domain repeats. In rats and mice, the Wilson disease protein has five metal-binding domain repeats, with the fourth being omitted based on sequence homology to the human Wilson disease protein (58). *Drosophila melanogaster* (fruit fly) contains four metal-binding domains (59) while *Caenorhabditis elegans* contains three (60). Though the individual metal-binding
domains all have the same three dimensional structure, a $\beta\alpha\beta\alpha\beta$ arrangement known as a ferredoxin fold, the primary structures of the domains are different (58). The NMR solution structure of the third metal-binding domain of Wilson disease protein is shown in Figure 1.4 as an example of a ferredoxin fold (61).

A crystal structure does not exist for the intact Wilson disease protein. However, in 2011 the crystal structure of the apo form of CopA, a homolog of Wilson disease protein, from *Legionella pneumophila* was determined at a resolution of 3.2 Å (62); a ribbon diagram of the structure is shown in Figure 1.5. CopA contains one metal-binding domain, but its position could not be modeled due to its mobility. Some unassigned electron density permitted researchers to assign the region to a general area, however. That location is one of the three places suggested by cryo-electron microscopy to be the location of the metal-binding domain for CopA from * Archaeoglobus fulgidus* (63), corroborating the *L. pneumophila* model (62).
Figure 1.4: The NMR solution structure of the third metal-binding domain of Wilson disease protein, modified from PDB 2ROP. The red $\alpha$-helices and yellow $\beta$-sheets of the ferredoxin fold are shown. The two side chains of the two metal-binding cysteines, located in the solvent exposed loop between the first $\beta$-sheet and the first $\alpha$-helix, are shown as orange sticks (61).
Figure 1.5: A cartoon model of CopA (apo form) from *L. pneumophila*, modified from structure 3RFU from the PDB (62). The actuator, nucleotide binding domain, and phosphorylation domains are shown in yellow, cyan, and green, respectively. The metal-binding domain is shown as a transparent lavender sphere. Six of the transmembrane helices are shown in orange while the two helices that form the platform are shown in red. Copper movement is shown, with the blue boxes denoting the transmembrane binding sites. In the larger rectangle, the side chains of the cysteines and proline in the conserved CPC motif are shown in yellow and purple, respectively.
The crystal structure of CopA revealed that two of the helices appear to form a cytosolic platform, shown as red helices in the figure above. This platform may facilitate the transfer of copper from the metal-binding domains to the copper binding site within the membrane; upon crossing the membrane, the copper is released into the lumen. One of the helices, termed MB, is kinked and forms the putative platform. It is proposed that HAH1 can deliver copper directly to the platform and bypass the metal-binding domains (62); if this is true, then the metal-binding domains may have more of a regulatory role, as others have suggested (63–65).

Comparing the sequences of CopA proteins from various species revealed that the most highly conserved residues were located in the core of the protein, including in the putative platform area (62); this is shown in Figure 1.6. Using this structure as a template and mapping the features of WLNP onto it, Gourdon et al. observed that mutations in the highly conserved core region give rise to more serious forms of Wilson disease. Several mutations occur near the phosphorylation site in the P-domain. The most common mutation, H1089Q, occurs in the ATP-binding pocket of the N-domain while the most common mutation in Asian populations, R778L, occurs near the platform entry site; several other mutations are also located near the platform, underscoring the importance of this region (45).
Figure 1.6: Conservation of residues among CopA proteins (62). The blue mesh shows the believed location of the metal-binding domain. The most conserved regions are located near the platform, indicated by the arrow. Using this structure as a template to construct a model of WLNP revealed that the most damaging mutations are located in the heavily conserved regions, particularly the platform area. Used by permission of Nature Publishing Group.

The presence of multiple metal-binding domains in the N-terminal region has prompted considerable interest, particularly with understanding why humans have six. Linkers of various lengths connect the metal-binding domains, providing an inherent separation between the metal-binding domains that may be significant. The lengths of the linker regions in human Wilson disease protein are shown in Figure 1.2. There is speculation that the domains are not all functionally equivalent and some may have more of a regulatory role (63–65).
Attempts to recombinantly express and purify large quantities of full length Wilson disease protein or Menkes protein have not been successful to date, owing to the difficulties inherent with expressing large transmembrane proteins. Consequently, researchers have instead focused on expressing and characterizing individual domains and combinations of domains, such as the MBD4 (66), MBD3-4 (61), or the actuator domain (50) of Wilson disease protein and MBD3 (67) or the actuator domain (50) of Menkes protein. By studying the structure and function of its constitutive domains, researchers hope to understand how the intact proteins function.

1.4.2 Copper Acquisition by the Metal-Binding Domains

One aspect of WLNP that has garnered a lot of focus is the MBDs, particularly with regard to delineating the mechanism by which copper is transferred between the MBDs and from HAH1 to the MBDs. By studying the copper binding affinities of the MBDs and HAH1, X-ray crystal structures of HAH1 complexed with various metals, and titrations of the MBDs with either copper loaded HAH1 or other copper loaded MBDs, researchers have learned about copper transfer and the driving forces behind copper movement. HAH1 is comprised of 68 amino acids that are arranged in the same \( \beta_\alpha \beta_\alpha \beta \) motif possessed by each metal-binding domain in WLNP. The similarity in structure between the chaperone and the target protein is a noticeable feature of proteins involved in copper homeostasis. The copper binding CXXC motif in HAH1 is located in the solvent accessible loop between the first \( \beta \)-sheet and \( \alpha \)-helix, the same location as in WLNP (68). HAH1 acquires Cu(I) after import into the cell and transports it to WLNP (16).
Determining the copper affinities of HAH1 and the MBDs has proven challenging, with different techniques and different ligands yielding different absolute values that vary over several orders of magnitude (69–71). One reason for the discrepancy is the limitations of the method chosen. The best way to obtain accurate values is to perform competition experiments in the presence of the protein using small ligands with known formation constants. This is non-trivial, because the small ligand needs to have an affinity for copper close to that of the protein and thereby buffer the ‘free’ copper in solution. Recent developments have led to the synthesis of new ligands that meet these criteria (72). Using the Cu(I) chelators BCS and BCA, Xiao et al. reported Cu(I) apparent dissociation constants (reported as logK_D’) for HAH1 and WLN5-6 to be -17.4 and -17.6, respectively (71). Both of these proteins bind copper with high affinity, yet are also able to transfer copper efficiently. Future experiments are planned to attain greater accuracy and understanding of copper binding to these types of sites (71–73).

Copper transfer from HAH1 to constructs of either single or multiple MBDs has been monitored using NMR. In 2006, Achila et al. examined the acquisition of copper by the two MBDs that are closest to the transmembrane region – WLN5-6 (66). When apoWLN5-6 was titrated with Cu(I)-HAH1, neither complex formation between the two proteins nor copper exchange was observed. In contrast, when apoWLN5-6 was titrated with Cu(I)-WLN4, partial copper transfer did occur, first to MBD6 and then to MBD5, though no complex formation was detected. The same study also found that titrating apoWLN4 with Cu(I)-HAH1 led to both copper transfer and the detection of an adduct that was in fast exchange on the NMR timescale. Additionally, the researchers found that
titrating apoWLN2 with Cu(I)-HAH1 also resulted in the formation of an NMR-observable adduct, though it was formed in smaller amounts than the adduct that was formed with WLN4. Furthermore, apoWLN5-6 behaved as a unit in solution.

WLN3-4 was investigated by Banci et al., who found that the MBDs rotated independently of each other in solution (61). Titration of apoWLN3-4 with Cu(I)-HAH1 led to the formation of an NMR-observable complex with MBD4 that was in fast exchange; no complex formation was detected for MBD3, however, and MBD3 did not acquire copper.

While the studies with constructs of double MBDs yielded important evidence as to how copper transfer occurred, it remained to be answered whether or not the behavior of these truncated constructs accurately reflected what occurs in the intact metal-binding domain of WLNP. The above results were consistent with other studies that showed that interaction (74).

In 2009 Banci et al. reported on the interaction of WLN1-6 with Cu(I)-HAH1 (75). In contrast to previous studies, this one reported that each of the six MBDs could acquire copper from Cu(I)-HAH1. Additionally, MBDs 1, 2, and 4 formed NMR-observable complexes, as evidenced by shifts in the backbone amides chemical shifts; the complex formation between Cu(I)-HAH1 and MBDs 2 and 4 were in agreement with what was previously reported (61, 66). The authors confirmed that MBDs 3, 5, and 6 acquired copper directly from HAH1 and not from intramolecular copper transfer from other MBDs by mutating the cysteines in the CXXC motifs to alanine in domains 1, 2, and 4. Titrating this mutated construct, which was unable to bind copper in the mutated domains, with Cu(I)-HAH1 still resulted in the transfer of copper to MBDs 3, 5, and 6.
As no complex formation was observed, this indicated that the exchange was slow on the NMR timescale. Unfortunately, the NMR solution structure for WLN1-6 could not be determined; the construct contained 633 residues, had limited solubility, and degraded significantly over 48-72 hours, thereby complicating efforts at determining a structure.

A study done on the structural and dynamic properties of apo and Cu(I)WLN4-6 yielded results that were not in complete agreement with the above studies (76). $^{15}$N relaxation experiments performed on both apo and Cu(I)WLN4-6 indicated that there was no large change in the dynamic properties of the protein upon copper binding. The domains tumbled independently of each other, with even MBDs 5 and 6 showing slight flexibility. This contrasted with the more rigid dimer that these two domains possessed when they were in a construct comprised only of MBDs 5 and 6 (i.e. WLN5-6) (66).

Additionally, the study on WLN4-6 found that apoWLN4-6 interacted with apoHAH1 even in the absence of copper or in low amounts of copper, with domain 4 showing the most chemical shifts; the authors suggested this meant that the transfer of copper from HAH1 to WLN4 is driven by favorable protein-protein interactions that occur even in the absence of copper. This supposition is supported by the fact that WLN4 is an acidic domain with a theoretical pI of 3.81 (preceding and succeeding linkers not included) while HAH1 is a basic protein with a theoretical pI of 7.5. Favorable interactions between an acidic face of WLN4 with a basic face of HAH1 could help to align the copper binding sites and facilitate copper transfer; a potential model of this is shown in Figure 1.7. From these results, Fatemi et al. suggest the copper transfer is driven by the motional freedom of the MBDs that is afforded by the linkers as well as favorable protein-protein interactions between both HAH1 and the MBDs and between the MBDs.
Additionally, the studies done on WLN4-6 were done at 35°C while the other studies were done at 25°C, which may account for some of the extra movement seen in WLN4-6 that was not present in other constructs.

The above studies show that it is important to consider the role of individual MBDs within the context of the intact N-terminal region. The studies all agree that MBD4 is the most likely of all domains to form an NMR-observable complex with Cu(I)-HAH1, however.

The mechanism of copper transfer from HAH1 to WLN4 was the subject of lab colleague Brian Zeider’s dissertation project (77). Part of his work focused on trapping the proposed 3-coordinate intermediate that forms between Cu(I)HAH1-WLN4 during copper transfer and delineating whether all cysteines in the CXXC motifs played a role in ligating copper. To achieve this, each of the four metal-binding cysteines in HAH1 and WLN4 were mutated one at a time to alanine, i.e. C12A HAH1, C15A HAH1, C15A WLN4 and C18A WLN4. Each mutated protein was enriched in $^{13}$C and $^{15}$N, then titrated with the unlabeled, non-mutant partner and copper, in a 1:1:1 ratio. Through the analysis of chemical shift variations and tumbling times, it was determined that both metal-binding cysteines in WLN4 are required to form an NMR-observable complex while only C12 from HAH1 was required (77). The observation that the C15 in HAH1 is not observed in a complex with WLN4 by NMR correlates with a computational study done on the proposed mechanism of copper transfer from HAH1 to WLN4 (78). This study reported that one of the most energetically favorable 3-coordinate intermediates that could form would include both cysteines in WLN4 and C12 in HAH1; the other energetically favorable intermediate is the one that initially forms when C15 from WLN4
both cysteines in HAH1 formed the initial 3-coordinate intermediate. After overcoming the 39.7 kJ /mol activation barrier, the intermediate with both cysteines in WLN4 and C12 in HAH1 is formed. Then C12 is released from the complex, finalizing the transfer of copper to WLN4, as depicted in Figure 1.8 below. This potential model of copper transfer is consistent with the intermediate complex that is found in yeast homologs, where C15 of Atx1 and C13 and C16 of Ccc2a provide the ligands to Cu(I) (79).

Intramolecular copper transfer from one metal-binding domain to another has also been documented between single domain constructs in vitro, with Cu(I)-WLN1 being able to transfer copper to apoWLN4 and vice versa (80). This strengthens the argument that copper can be moved from one MBD to another prior to being moved across the membrane.
Figure 1.7: Potential interaction between WLN4 and HAH1 (81). WLN4 has an acidic face that can interact with a basic face on HAH1. This interaction would bring the metal-binding cysteines (yellow spheres) of WLN4 and HAH1 into close proximity, thereby facilitating the transfer of copper. Used with permission of Springer.

Figure 1.8: Diagram of a potential mechanism of copper transfer from HAH1 to WLN4. While other potential 3-coordinate intermediates are possible, NMR studies have shown that the intermediate shown above dominates (77).
1.4.3 Regulation

The activity of Wilson disease protein is regulated by two means – copper levels and phosphorylation. These two factors determine the protein’s localization, i.e. whether it is in the trans-Golgi network or in vesicles.

Under basal or low copper levels, Wilson disease protein is localized to the trans-Golgi network where it pumps copper across the membrane and into the lumen for incorporation into ceruloplasmin and other cuproproteins. When copper levels exceed biological demand and threaten to overwhelm the chelation capacity of the cell, Wilson disease protein translocates into vesicles and migrates to the plasma membrane; once there, the vesicles fuse with the plasma membrane and export copper out of the cell (30).

In the liver, excess copper is exported into bile and excreted from the body. Mutations in the Wilson disease protein can prevent proper trafficking and interfere with the ability to excrete copper, resulting in the accumulation of copper in the body (82, 83). Since the liver is the primary site of copper metabolism, most of the copper accumulates there.

Phosphorylation of specific serine residues also plays a role in the regulation and localization of Wilson disease protein (84–86). Wilson disease protein exhibits a basal level of phosphorylation under basal copper levels. An increase in copper levels triggers a kinase mediated phosphorylation of other serine residues, resulting in the Wilson disease protein migrating from the trans-Golgi network to vesicles. Four serine residues become phosphorylated in response to elevated copper levels. Two of these serines (Ser478 and Ser481) are located in the linker region between metal-binding domains 3 and 4, another (Ser1121) is located in the nucleotide-binding domain and one (Ser1453) is located at the C-terminal tail (85). Protein kinase D is essential for phosphorylating these residues (86). Additionally, mutating the cysteines in the CXXC motif of metal-
binding domain 6 prevents the protein kinase D mediated phosphorylation of Wilson disease protein, thereby preventing its ability to translocate into vesicles (86). Consequently, this metal-binding domain plays a pivotal role in the proper functioning of Wilson disease protein.

1.5 Research Objectives

The role of the first four metal-binding domains of Wilson disease protein has not been clearly elucidated. Metal-binding domains 1, 2, and 4 have been shown to form NMR observable complexes with the copper metallochaperone HAH1, which suggests that one of the roles for these domains is to acquire copper. On the other hand, other experiments have pointed to a more regulatory role for the first four metal-binding domains. Deletion of the first four metal-binding domains does not alter the catalytic phosphorylation of WLNP, but it does increase the sensitivity of WLNP to copper. Lower copper levels are able to activate the truncated WLNP (64). Activation by lower levels of copper suggests that the first four metal-binding domains may affect the access of copper to metal-binding sites that are located outside of the N-terminal region, such as the platform area. Additionally, catalytic phosphorylation of Wilson disease protein was observed even when the first five metal-binding domains are deleted (87).

Several studies have shown that WLNP can function with only the sixth metal-binding domain (64, 88). Since the linker between the fifth and sixth metal-binding domains is only 8 residues long and NMR data on a construct of WLN5-6 showed that they tumbled as a single unit, it is likely that the fifth and sixth metal-binding domains in WLNP function together. Additionally, the long linker between the fourth and fifth metal-binding domains provides a large spatial separation that may enable the first four
metal-binding domains to function independently from the fifth and six metal-binding domains. Whether or not the first four metal-binding domains function independently from each other, like beads on a string, or cooperatively has not yet been elucidated and was the subject of this dissertation.

Taking all of this into consideration led to the identification of research objectives for this dissertation. The goal was to understand the structure and stability of the first four metal-binding domains of human Wilson disease protein. The structure was characterized using gel filtration chromatography, laser light scattering, and NMR spectroscopy. Stability was analyzed by CD in conjunction with thermal and GuHCl denaturation. The results of these experiments were expected to lead to a greater understanding of the role of the first four metal-binding domains of WLNP. In turn, this would lead to a greater understanding of WLNP function and the reason certain domains are favored sites for copper acquisition. The study would also lead to a greater appreciation of how mutations impair the functioning of WLNP.

A small secondary project investigated the effect of pH and Hg(II) binding on the association state of the copper metallochaperone HAH1. Gel filtration chromatography was used to characterize the association state. This project was part of a larger project that was done with Dr. Pecoraro and his group from the University of Michigan. Hg(II) was used as a mimic for Cu(I) because it is amenable to several spectroscopic techniques that Cu(I) is not. This study was undertaken with the hope of observing changes in the coordination state of Hg(II) as pH changed. Such changes may reflect the changes in coordination that Cu(I) undergoes when it is being transferred from HAH1 to WLNP and may provide a model of copper transfer.
References


CHAPTER 2
MATERIALS AND EXPERIMENTAL METHODS

2.1 Introduction

This chapter details the chemical and biological supplies used for conducting this research, as well as the techniques used for the preparation of the plasmids, and the expression, purification, and characterization of the resulting proteins.

2.2 Reagents and Instrumentation

The specialty chemical and biological reagents used are itemized in the tables below. Commonly used reagents, such as buffers and salts, were of molecular biology grade or better. Information on the instrumentation, including model number and applications, are also provided.

2.2.1 Chemical and Biological Reagents

The reagents shown below were used in this research. Separate tables have been provided for chemical and biological reagents.
Table 2.1: Chemical reagents used. The purity of the reagents is given, when available.

<table>
<thead>
<tr>
<th>Chemical reagent</th>
<th>Part Number</th>
<th>Manufacturer</th>
<th>Purity</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{15}$N-Ammonium Chloride</td>
<td>NLM-467-1</td>
<td>Cambridge Isotopes</td>
<td>99%</td>
</tr>
<tr>
<td>Copper Standard for ICP-AES</td>
<td>SC194-100</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Deuterium Oxide</td>
<td>453366</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>DSS</td>
<td>DLM-32</td>
<td>Cambridge Isotopes</td>
<td></td>
</tr>
<tr>
<td>d$_{12}$-EDTA</td>
<td>DLM-414</td>
<td>Cambridge Isotopes</td>
<td>&gt;99%</td>
</tr>
<tr>
<td>$^{13}$C-Glucose</td>
<td>CLM-1396</td>
<td>Cambridge Isotopes</td>
<td>&gt;99%</td>
</tr>
<tr>
<td>Guanidine Hydrochloride</td>
<td>G3272</td>
<td>Sigma-Aldrich</td>
<td>≥99%</td>
</tr>
<tr>
<td>Indium Standard for ICP-AES</td>
<td>IAA-049</td>
<td>Ultra Scientific</td>
<td></td>
</tr>
<tr>
<td>MEM Vitamins</td>
<td>25-020-CI</td>
<td>Mediatech</td>
<td></td>
</tr>
<tr>
<td>Nitric Acid (Trace Metal Grade)</td>
<td>A509212</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Tetrakis(acetonitrile)copper(I) Hexafluorophosphate</td>
<td>346276</td>
<td>Sigma-Aldrich</td>
<td>&gt;99%</td>
</tr>
</tbody>
</table>
Table 2.2: Biological reagents and kits used in this research.

<table>
<thead>
<tr>
<th>Biological Reagent / Kit</th>
<th>Part Number</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCA Protein Assay Kit</td>
<td>23225</td>
<td>Thermo Scientific</td>
</tr>
<tr>
<td>Benzonase</td>
<td>E1014</td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>Bio-Rad Protein Assay Reagent</td>
<td>500-0006</td>
<td>Bio-Rad</td>
</tr>
<tr>
<td>BL21(DE3) Competent Cells</td>
<td>69450</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>BugBuster® Reagent</td>
<td>70584-4</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>DNA Ladder (100 bp)</td>
<td>N3231S</td>
<td>NEB</td>
</tr>
<tr>
<td>DNA Ladder (1 kb)</td>
<td>N3232S</td>
<td>NEB</td>
</tr>
<tr>
<td>Gel Filtration Markers Kit for Protein Molecular Weights 6.5-66 kD</td>
<td>MWGF70</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>KOD Hotstart Polymerase Kit</td>
<td>71086-3</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>Low Molecular Weight Gel Filtration Calibration Kit</td>
<td>17-0442-01</td>
<td>Amersham Biosciences</td>
</tr>
<tr>
<td>MinElute PCR Purification Kit</td>
<td>28004</td>
<td>QIAGEN</td>
</tr>
<tr>
<td>Novablue Competent Cells</td>
<td>69825-4</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>pET32Xa/LIC Vector Kit</td>
<td>70072</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>QiaPrep® Spin Miniprep Kit</td>
<td>27106</td>
<td>QIAGEN</td>
</tr>
<tr>
<td>QuikChange II Site-Directed Mutagenesis Kit</td>
<td>200521</td>
<td>Agilent Technologies</td>
</tr>
<tr>
<td>Rainbow Protein Markers (Full Range)</td>
<td>RPN800E</td>
<td>GE Healthcare</td>
</tr>
<tr>
<td>Rosetta™(DE3) Competent Cells</td>
<td>70954-3</td>
<td>EMD Millipore</td>
</tr>
<tr>
<td>XL-10 Gold Competent Cells</td>
<td>200517-4</td>
<td>Agilent Technologies</td>
</tr>
</tbody>
</table>
2.2.2 Instrumentation

The instruments summarized in Table 2.3 were utilized during the course of this research. The columns used with the ÄKTA FPLC are shown in Table 2.4.

Table 2.3: Instruments used.

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Manufacturer</th>
<th>Model / Part Number</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>ÄKTA FPLC</td>
<td>GE Lifesciences</td>
<td>Part Number: 18-1900-26</td>
<td>Protein Purification</td>
</tr>
<tr>
<td>CD Spectropolarimeter</td>
<td>Jasco</td>
<td>Model Number: J-815</td>
<td>Circular Dichroism Experiments</td>
</tr>
<tr>
<td>Minicycler Thermocycler</td>
<td>MJ Research</td>
<td>Model Number: PTC-150</td>
<td>PCR</td>
</tr>
<tr>
<td>NMR (located at the University of Notre Dame, South Bend, IN)</td>
<td>Bruker</td>
<td>AVANCE II, 800 MHz</td>
<td>NMR Experiments</td>
</tr>
<tr>
<td>NMR (located at the University of Chicago, Chicago, IL)</td>
<td>Varian</td>
<td>Unity Inova 600 MHz</td>
<td>1H-15N HSQC Experiment of WLN4 at 80ºC</td>
</tr>
<tr>
<td>Refractometer</td>
<td>Bausch and Lomb</td>
<td>Abbe 3L</td>
<td>Determining the Concentration of GuHCl Solution</td>
</tr>
<tr>
<td>Shaker / Incubator</td>
<td>New Brunswick Scientific</td>
<td>Unknown</td>
<td>Bacterial Culture Growth</td>
</tr>
<tr>
<td>UV-VIS Spectrophotometer</td>
<td>Beckman</td>
<td>Model Number: DU-7400</td>
<td>Plasmid and Protein Quantitation</td>
</tr>
</tbody>
</table>
Table 2.4: Columns used for the purification of the WLN constructs. All part numbers are from GE Healthcare.

<table>
<thead>
<tr>
<th>Column</th>
<th>Part Number</th>
<th>Resin</th>
<th>Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>HisPrep FF 16/10</td>
<td>28-9365-51</td>
<td>Ni Sepharose 6 Fast Flow</td>
<td>Purification of Fusion Proteins and Subsequent Digests</td>
</tr>
<tr>
<td>HiLoad 26/60 Superdex 75 pg</td>
<td>17-1070-01</td>
<td>Superdex 75 Prep Grade</td>
<td>Gel Filtration After HisPrep</td>
</tr>
<tr>
<td>HiLoad 16/60 Superdex 200 pg</td>
<td>17-1069-01</td>
<td>Superdex 200 Prep Grade</td>
<td>Gel Filtration After HisPrep</td>
</tr>
<tr>
<td>HiPrep 26/10 Desalting Column</td>
<td>17-5087-01</td>
<td>Sephadex G-25 Superfine</td>
<td>Buffer Exchange</td>
</tr>
<tr>
<td>Superdex 75 HR 10/300</td>
<td>17-1047-01</td>
<td>Superdex 75, 13 μm</td>
<td>Determination of Apparent Molecular Weight</td>
</tr>
</tbody>
</table>

2.3 Summary of Plasmids and Protein Constructs

This section contains a summary of all of the plasmids and protein constructs used in this research.

2.3.1 Plasmids Containing Portions of the Wilson Disease Protein

Multiple plasmids were used in this research, though all were based on the pET-32Xa/LIC vector. Table 2.5 shows the name of each plasmid and the amino acids in the human Wilson Disease protein (WLNP) for which it codes; the residue numbering is based on the Copper transporting ATPase 2 isoform A [homo sapiens], NCBI accession number NP_000044. The protein constructs are referred to as “WLN” followed by the metal-binding domain they contain. For example, WLN1-2 is a construct comprised of the first two metal-binding domains.
### Table 2.5: Plasmid names and the metal-binding domains they contained.

<table>
<thead>
<tr>
<th>Plasmid Name</th>
<th>Residues of WLNP</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET32-WLN1-2</td>
<td>57-213</td>
<td>First two metal-binding domains</td>
</tr>
<tr>
<td>pET32-WLN1-3</td>
<td>57-327</td>
<td>First three metal-binding domains</td>
</tr>
<tr>
<td>pET32-WLN1-4</td>
<td>57-429</td>
<td>First four metal-binding domains</td>
</tr>
<tr>
<td>pET32-WLN1-6</td>
<td>57-633</td>
<td>All six metal-binding domains</td>
</tr>
<tr>
<td>pET32-WLN4</td>
<td>356-429</td>
<td>Fourth metal-binding domain</td>
</tr>
<tr>
<td>pET32-WLN1-4 G333R</td>
<td>57-429</td>
<td>Identical to pET32-WLN1-4, except for a glycine to arginine mutation at residue 333 (located between the third and fourth metal-binding domains)</td>
</tr>
</tbody>
</table>

### WLN1-2 Amino Acid Sequence

GVATSTVRLGMTCQSCVKSIEDRISNLKGIISMKVSLEQGSATVKYVPSVVCL
QQVCHQIGDMGFEASIAEGKAAWSPSRLPAQEAVVKLRVEMTCQSCVSSIE
GKVRKLGQVVRKVSLSNQEAVITYQPYILIQPDLRHDHVNDMGFEEAIKSKV

**MBD1, MBD2, Linker region**

157 amino acids
16,958 Daltons

**Figure 2.1**: The amino acid sequence of WLN1-2. The different metal-binding domains (MBDs) are indicated by different colored fonts, as detailed above.

### WLN1-3 Amino Acid Sequence

GVATSTVRLGMTCQSCVKSIEDRISNLKGIISMKVSLEQGSATVKYVPSVVCL
QQVCHQIGDMGFEASIAEGKAAWSPSRLPAQEAVVKLRVEMTCQSCVSSIE
GKVRKLGQVVRKVSLSNQEAVITYQPYILIQPDLRHDHVNDMGFEEAIKSKV
APLSLGPIDIERLQSTNPKRPLSSANQNFNSETLGHQGSHVVTLLQLRIDGMHCKSCVLNIEENIQLLGVQSIQVSLNKTAQVKYDPSCTSPVALQRAIEALPPGNNFKVSL

**MBD1, MBD2, MBD3, Linker regions**

272 amino acids
29,382 Daltons

**Figure 2.2**: The amino acid sequence of WLN1-3. The different metal-binding domains (MBDs) are indicated by different colored fonts, as detailed above.
**WLN1-4 and WLN1-4 G333R Amino Acid Sequences**

GVATSTVRLGMCQSCVKSIEIDRSLNKGSSMKVSLQGATVVKYPVMVCL
QQVCHQIGDMGEASLGAASWPSRLPAQEAVKLRVEGMCQSCVSSIE
GVKVRKQGVRVVKVSLQAVQNVAPQVQPLDRLHDVNDMGEAAIKSKV
APLSLGPIDIERLQSTNPRLSSANQNFFNSETLGHQGSHVTQLRQIDGMHC
KSCVLDNIEENIQLLGVQSIQVSLNKTQVKYDPSCTSPVALQRALIEALPPGNF
KVSLPDGAEGSGTDHRSSHSSHSGSPPRNQVQGTCSTTLIAIAAGMTCASCVHSIE
GMISQLEGVQQISVSLAEVTLYNPAVISPEELRAAIEDMGEASVVSE

**MBD1, MBD2, MBD3, MBD4, Linker regions**

374 amino acids

39,783 Daltons (WLN1-4) 38,868 Daltons (WLN1-4 G333R)

**Figure 2.3:** The amino acid sequence of WLN1-4 and WLN1-4 G333R. The underlined glycine residue (G) was mutated to arginine (R) in the construct WLN1-4 G333R. The different metal-binding domains (MBDs) are indicated by different colored fonts, as detailed above.

**WLN1-6 Amino Acid Sequence**

GVATSTVRLGMCQSCVKSIEIDRSLNKGSSMKVSLQGATVVKYPVMVCL
QQVCHQIGDMGEASLGAASWPSRLPAQEAVKLRVEGMCQSCVSSIE
GVKVRKQGVRVVKVSLQAVQNVAPQVQPLDRLHDVNDMGEAAIKSKV
APLSLGPIDIERLQSTNPRLSSANQNFFNSETLGHQGSHVTQLRQIDGMHC
KSCVLDNIEENIQLLGVQSIQVSLNKTQVKYDPSCTSPVALQRALIEALPPGNF
KVSLPDGAEGSGTDHRSSHSSHSGSPPRNQVQGTCSTTLIAIAAGMTCASCVHSIE
GMISQLEGVQQISVSLAEVTLYNPAVISPEELRAAIEDMGEASVVSE

**MBD1, MBD2, MBD3, MBD4, MBD5, MBD6, Linker regions**

578 amino acids

61,505 Daltons

**Figure 2.4:** The amino acid sequence of WLN1-6. The different metal-binding domains (MBDs) are indicated by different colored fonts, as detailed above.
Figure 2.5: The amino acid sequence of WLN4. The sole metal-binding domain is indicated in orange. Neither the preceding nor succeeding linker regions were included in this construct.

2.3.2 HAH1 Plasmid

The DNA sequence that coded for the full length human HAH1 protein (68 amino acids) was inserted into the pET-11d vector using a series of overlapping primers by Dr. David Huffman. Relative to the native HAH1 sequence, the pET-11d construct contained an extra alanine residue at the beginning (MAP); the sequence is shown in Figure 2.6 below. The pET-11d vector contained no fusion or affinity tags.

Figure 2.6: The amino acid sequence of HAH1.

2.4 Construction of Plasmids, Protein Expression, and Protein Purification

This section discusses how the plasmids were constructed, the conditions under which the proteins were expressed, and the purification schemes used to purify the WLN constructs and HAH1.
2.4.1 Preparation of the pET32-WLN1-2, pET32-WLN1-3, pET32-WLN1-4, and pET32-WLN1-4 G333R Plasmids

The plasmids shown in Table 2.5 all originated from a plasmid that contained all six metal-binding domains (WLN1-6), subcloned into the pET-32Xa/LIC vector by Dr. Joshua Muia at Western Michigan University (1). The pET-32Xa/LIC vector contains a thioredoxin tag to enhance solubility and expression of the desired protein; it also contains a His$_6$ tag (6 histidines in a row) to assist with purification. The vector map for this plasmid is shown in Figure 2.7. A cartoon of the resulting fusion proteins, in which thioredoxin (TRX) was expressed at the amino terminus and a WLN construct at the carboxy terminal, is shown in Figure 2.8. A TEV site was inserted into the WLN1-6 plasmid after the fusion tags (thioredoxin and His$_6$ tag) and just upstream of the first metal-binding domain, for removal of the fusion tags following protein purification.
Figure 2.7: The vector map of the pET-32Xa/LIC plasmid (EMD Millipore, Technical Bulletin 186VM).

Figure 2.8: Diagram of the WLN protein constructs, which were expressed as fusion proteins with thioredoxin (TRX). The His\textsubscript{6} tag is shown in red font, and the TEV cleavage site is shown in purple font. TEV protease cleaves between the glutamine and glycine residues.
For all of the plasmids in Table 2.5 except WLN4, site directed mutagenesis was used to insert a stop codon at the appropriate place in the pET32-WLN1-6 plasmid to produce the desired protein construct (i.e. a stop codon at the end of metal-binding domain 2 produced the WLN1-2 plasmid). Site directed mutagenesis was performed by using the QuikChange II Site-Directed Mutagenesis Kit (Agilent Technologies). The primers used to incorporate the mutations were designed based on the parameters outlined in the manufacturer’s instruction guide (Agilent Technologies, Publication Number 200521-12) included with the mutagenesis kit. These parameters included having a primer length of 25-45 bases that terminated in one or more C or G bases, incorporating the mutation in the middle of the primer (with 10-15 bases of correct sequence on either side), a GC content ≥ 40%, and a melting temperature $T_m \geq 78^\circ C$ ($T_m = 81.5 +0.41(\%GC) - 675/N - \%mismatch$, where N is the primer length and the $\%GC$ and $\%mismatch$ are whole numbers). The primers, summarized in Table 2.6 below, were synthesized by Integrated DNA Technologies, Inc. (Coralville, Iowa). Prior to use, the lyophilized primers were resuspended in autoclaved, deionized water (18.2 MΩ•cm) at a concentration of 200 μM; working solutions of 10 μM were then prepared using the same sterile water. The primers were stored at -20°C pending usage.

The pET32-WLN1-4 plasmid was prepared by Dr. Joshua Muia and Sun Hwa Lee at Western Michigan University by using the same procedure as outlined in the above paragraph (1). The primers used to prepare this plasmid are shown in Table 2.6.

Insertion of the stop codon was achieved using mutagenic primers and PCR. The contents of each PCR reaction are shown in Table 2.7. The PCR was performed in a thermocycler using the conditions shown in Table 2.8. Mineral oil (30 μL) was overlaid
on top of each reaction after the addition of the polymerase. After PCR was completed, DpnI (10 U) was added to the PCR tube, and it was incubated at 37°C for one hour to degrade the template plasmid. Subsequently, an aliquot of the PCR product was run on a 0.8% agarose gel in order to verify the presence of a plasmid with the expected size of 7.8 kilobases (kb). Staining of the DNA was achieved by incorporating 0.5 μg/mL of ethidium bromide into the agarose gel before it was poured. Digesting the plasmid with BglII (20 U) and BamHI (30 U) yielded two fragments of the proper approximate size when run on an agarose gel: the pET-32Xa/LIC vector (5.8 kb) and the WLN1-6 gene insert (~1.78 kb). Note that all of the plasmids discussed in this section contain the WLN1-6 gene insert; the different constructs were achieved by incorporating a stop codon at the appropriate place in the DNA sequence. Consequently, the sizes of both fragments produced from the double digest were the same for all of these plasmids.
Table 2.6: The mutagenic PCR primers used to prepare the plasmids discussed above. The text in bold denotes the stop codon that was introduced. The underlined text indicates the mutation responsible for the conversion of glycine to arginine at position 333. The melting temperatures (Tm) listed were calculated according to the formula discussed in the preceding text.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Length</th>
<th>Tm (°C)</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN1-2 Stop</td>
<td>36</td>
<td>81.4</td>
<td>GAA GCT GCC ATC AAG AGC <strong>TAA</strong></td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
<td>GTG GCT CCC TTA AGC</td>
</tr>
<tr>
<td>WLN1-2 Stop</td>
<td>36</td>
<td>81.4</td>
<td>GCT TAA GGG AGC CAC <strong>TTA</strong> GCT CTT</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
<td></td>
<td>GAT GGC AGC TCT</td>
</tr>
<tr>
<td>WLN1-3 Stop</td>
<td>45</td>
<td>93.7</td>
<td>CCA CCT GGG AAT TTT AAA GTT TCT</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
<td>CTT <strong>TGA</strong> GAT GGA GCC GAA GGG</td>
</tr>
<tr>
<td>WLN1-3 Stop</td>
<td>45</td>
<td>93.7</td>
<td>CCC TTC GGC TCC ATC <strong>TCA</strong> AAG</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
<td></td>
<td>AGA AAC TTT AAA ATT CCC AGG TGG</td>
</tr>
<tr>
<td>WLN1-4 Stop</td>
<td>37</td>
<td>78.1</td>
<td>GCT TCA GTC GTT TCT GAA <strong>TGA</strong> TGT</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
<td>TCT ACT AAC CCT C</td>
</tr>
<tr>
<td>WLN1-4 Stop</td>
<td>37</td>
<td>78.1</td>
<td>GAG GGT TAG TAG AAC <strong>ATC</strong> ATT</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
<td></td>
<td>CAG AAA CGA CTG AAG C</td>
</tr>
<tr>
<td>WLN1-4 G333R</td>
<td>33</td>
<td>81.4</td>
<td>GAT GGA GCC GAA <strong>AGG</strong> AGT GGG</td>
</tr>
<tr>
<td>Forward</td>
<td></td>
<td></td>
<td>ACA GAT CAC AGG</td>
</tr>
<tr>
<td>WLN1-4 G333R</td>
<td>33</td>
<td>83.5</td>
<td>CCT GTG ATC TGT CCC ACT CCT** TTA**</td>
</tr>
<tr>
<td>Reverse</td>
<td></td>
<td></td>
<td>TTC GGC TCC ATC</td>
</tr>
</tbody>
</table>

Table 2.7: PCR reaction contents for site directed mutagenesis.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Reaction Buffer</td>
<td>5</td>
</tr>
<tr>
<td>pET32-WLN1-6 (10 ng/µL)</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer (10 µM)</td>
<td>1</td>
</tr>
<tr>
<td>Reverse Primer (10 µM)</td>
<td>1</td>
</tr>
<tr>
<td>dNTP Mix (2 mM)</td>
<td>1</td>
</tr>
<tr>
<td>Quik Change Solution</td>
<td>3</td>
</tr>
<tr>
<td>Sterile Water</td>
<td>38</td>
</tr>
<tr>
<td>Pfu Ultra HF DNA Polymerase (2.5 U/µL)</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2.8: PCR thermocycler conditions. The fourth segment was not part of the PCR cycle, but was included in the program to keep the reaction cool if the PCR product was not immediately used after completion of the program.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Number of Cycles</th>
<th>Temperature (ºC)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>3 minutes</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>95</td>
<td>50 seconds</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td></td>
<td>50 seconds</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>68</td>
<td>7 minutes 50 seconds</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>68</td>
<td>7 minutes 50 seconds</td>
</tr>
</tbody>
</table>

2.4.2 Preparation of the pET32-WLN4 Plasmid

The pET32-WLN4 plasmid was constructed using PCR to amplify the fourth metal-binding domain from the pET32-WLN1-6 plasmid. Ligation independent cloning was then used to insert the fourth metal-binding domain into the pET-32Xa/LIC vector (2, 3). KOD Hotstart Polymerase and the WLN4 primers listed in Table 2.9 were used for the amplification; these primers also incorporated a TEV restriction site into the plasmid to facilitate cleavage of the fusion protein. The primers were designed to incorporate the requisite 5’ sequences for the insertion of the PCR insert into the pET32 vector (EMD Millipore, Technical Bulletin 184-1).

The contents of the PCR reaction tube and the thermocycler conditions are shown below in Table 2.10 and Table 2.11, respectively. Mineral oil (30 µL) was overlaid on top of each reaction after the addition of the polymerase. Following PCR, the product was purified using the MinElute PCR Purification Kit from QIAGEN. An aliquot of the purified product was then run on a 1.2% agarose gel to determine if the PCR was successful. Visualization of the DNA revealed a band between 200 & 300 bases; the expected size of the PCR product was 284 bases. The PCR product was purified using...
MinElute PCR Purification Kit from QIAGEN. Quantitation of the PCR product was achieved by running an aliquot on a 1.2% agarose gel and comparing the intensity of the band to that of the 100 bp standard ladder.

**Table 2.9:** The PCR primers used to amplify the fourth metal-binding domain. The text in bold denotes the vector specific sequences required for ligation independent cloning while the underlined text denotes the TEV protease restriction site. The melting temperatures listed were supplied by the Integrated DNA Technologies, Inc.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Length</th>
<th>Tm (°C)</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN-4 Forward</td>
<td>45</td>
<td>69.1</td>
<td>GGT ATT GAG GGT CGC GAG AAC CTT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TAT TTC CAG GGÇ ACA TGC AGT</td>
</tr>
<tr>
<td>WLN-4 Reverse</td>
<td>51</td>
<td>67.9</td>
<td>AGA GGA GAG TTA GAG CCT CAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TCA GAA ACG ACT GAA GCC TCA AAT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CCC ATG</td>
</tr>
</tbody>
</table>

**Table 2.10:** PCR reaction contents for preparing the WLN4 plasmid.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10x Reaction Buffer</td>
<td>5</td>
</tr>
<tr>
<td>pET-32Xa/LIC-TEV-WLN1-6 (10 ng/µL)</td>
<td>1</td>
</tr>
<tr>
<td>Forward Primer (10 µM)</td>
<td>1.5</td>
</tr>
<tr>
<td>Reverse Primer (10 µM)</td>
<td>1.5</td>
</tr>
<tr>
<td>dNTP Mix (2 mM)</td>
<td>5</td>
</tr>
<tr>
<td>Magnesium Sulfate (25 mM)</td>
<td>3</td>
</tr>
<tr>
<td>Sterile Water</td>
<td>32</td>
</tr>
<tr>
<td>KOD Hotstart DNA Polymerase (1 U/µL)</td>
<td>1</td>
</tr>
</tbody>
</table>
Table 2.11: PCR thermocycler conditions. The fourth segment was not part of the PCR cycle, but was included in the program to keep the reaction cool if the PCR product was not immediately used after completion of the program.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Number of Cycles</th>
<th>Temperature (ºC)</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>2 minutes</td>
</tr>
<tr>
<td>2</td>
<td>30</td>
<td>95</td>
<td>20 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>58</td>
<td>30 seconds</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>70</td>
<td>30 seconds</td>
</tr>
<tr>
<td></td>
<td></td>
<td>4</td>
<td>24 hours</td>
</tr>
</tbody>
</table>

The purified PCR product was then inserted into the pET-32Xa/LIC vector using the pET-32Xa/LIC vector kit from EMD Millipore. This vector kit utilized ligation independent cloning (LIC) to directionally insert genes into the vector without performing restriction enzyme digests and ligation reactions (2, 3). Instead, the PCR products are designed to incorporate 5’ sequences that are complementary to the DNA sequence of the vector. T4 DNA polymerase was used to create the requisite vector compatible overhangs in the PCR product; the vector and the PCR product were then annealed, and the plasmid was used to transform Novabluë® cells (EMD Millipore).

After PCR, 10 μL (0.2 picomoles) of purified PCR product was combined with 2 μL of 10x T4 DNA polymerase buffer, 2 μL of dGTP (25 mM), 1 μL of DTT (100 mM), 4.6 μL of nuclease free water, and 0.4 μL of T4 DNA Polymerase (2.5 U/ μL). This mixture was incubated at 22°C for 30 minutes in the thermocycler followed by a 5 minute incubation at 75°C to inactivate the T4 DNA polymerase. Annealing of the vector and the PCR product was performed by mixing 3 μL of the above described T4 DNA Polymerase treated PCR product with 1 μL of pET-32Xa/LIC vector and incubating the mixture for 5 minutes at 22°C. Subsequently, 1.25 μL of EDTA (25 mM) was added, followed by another 5 minute incubation at 22°C.
2.4.3 Transformation of Bacterial Cells for Plasmid Propagation

The plasmids described in Sections 2.4.1 and 2.4.2 were propagated in *E. coli*. Two different strains of cells were used: XL-10 Gold Ultracompetent Cells (Agilent Technologies) were used for the plasmids that were produced using site-directed mutagenesis while Novablue® competent cells (EMD Biosciences) were used for the pET-32-WLN4 plasmid. The transformation procedures used for the two different cell lines differed slightly.

The XL10-Gold Ultracompetent Cells were transformed as follows. An aliquot of cells (5 μL) was thawed on ice for 5 minutes before 0.5 μL β-ME was added. The cells were then incubated on ice for 10 minutes, with gentle swirling occurring every 2 minutes. Afterwards, 3 μL of the DpnI treated DNA was added, and the cells were incubated on ice for 30 minutes. The cells were next heat shocked at 42°C for 10 seconds to facilitate plasmid entry into the cells. Following a 2 minute recovery on ice, 100 μL of SOC media (20 g/L tryptone, 5 g/L yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl₂, 10 mM MgSO₄, 10 mM glucose, pH7.0) that had been warmed to 37°C was added. The cells were then incubated in a 37°C shaker (250 rpm) for 1 hour. Afterwards, all of the cells were plated on a pre-warmed LB-agar plate (10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract, 15 g/L agar) that contained 100 μg/mL carbenicillin and incubated upside down for 16-18 hours at 37°C in a gravity convection oven.

Novablue® competent cells were transformed according to the following procedure. An aliquot of cells (20 μL) was thawed on ice for 5 minutes before 2 μL of the annealed plasmid mixture was added. The cells were incubated on ice for 5 minutes, then heat shocked at 42°C for 15 seconds. Following a 2 minute recovery on ice, 80 μL of SOC media (pre-warmed to 37°C) was added. The cells were grown in a shaker
maintained at 37°C and 250 rpm for 1 hour. All of the cells were then plated onto a pre-
warmed LB-agar plate that contained 100 μg/mL carbenicillin. The plate was incubated
in a gravity convection oven for 16-18 hours.

Further propagation of the plasmids was achieved by inoculating multiple 5 mL
cultures of LB media (10 g/L tryptone, 10 g/L NaCl, 5 g/L yeast extract) containing
either 100 μg/mL carbenicillin or 100 μg/mL ampicillin with cells from a single colony
from the plates. The cultures were incubated in a 37°C shaker (250 rpm) for 8-10 hours,
and then centrifuged at 13,000 rpm for 5 minutes. The supernatant was discarded, and
the pellet was either immediately subjected to plasmid purification or stored at -20°C
until processed.

The plasmids were purified using the QiaPrep® Spin Miniprep Kit from
QIAGEN, according to the manufacturer’s instructions included with the kit (QiaPrep®
plasmids were quantitated by measuring the absorbance at 260 nm, using a Beckman DU-
7400 UV-Vis Spectrometer. Aliquots of the plasmids were then sent to Retrogen (San
Diego, CA) for sequencing. All plasmids were stored at -20°C.

2.4.4 Protein Expression

Two different media were used to grow the bacteria. Specific media was chosen
based on the particular experiment for which the protein would be used. Unlabeled
proteins were required for most experiments, and these proteins were produced in LB
media. Proteins being studied with NMR required enrichment with either 15N, or 13C and
15N; M9 minimal media supplemented with the appropriate precursors was used to
prepare these proteins.
2.4.4.1 Transformation of Rosetta™ 2 (DE3) Cells for Protein Expression

Regardless of the type of liquid media used to grow the cultures and express the protein, a transformation was required to introduce the appropriate plasmid into the bacteria. This transformation procedure was the same, regardless of the type of media used subsequently.

Rosetta™ 2(DE3) cells were used to express the proteins encoded by the plasmids described in Sections 2.32 and 2.4.1-2.4.2. This BL21 derivative was selected because it is supplemented with tRNAs for 7 codons (AGA, AGG, AUA, CUA, GGA, CCC, and CGG) that are infrequently used by *E. coli* but commonly found in eukaryotes (EMD Millipore, Technical Bulletin TB009); the WLN1-4 and WLN1-4 G333R constructs contained all 7 rare codons.

Transformation of Rosetta™ 2(DE3) cells was performed according to the manufacturer’s procedure that was supplied with the cells (EMD Millipore, Technical Bulletin TB009); a summary of this procedure follows. A 5 μL aliquot of cells was removed from -80°C and thawed on ice for 5 minutes. Between 10-20 ng of plasmid (in a volume of 1-2 μL) was added to the cells, followed by gentle swirling every 2 minutes during a 10 minute incubation on ice. The cells were then heat shocked for 15 seconds at 42°C followed by a 2 minute recovery on ice. After the addition of 80μL of SOC media (pre-warmed to 37 °C), the cells were placed in a 37°C shaker (250 rpm) for 1 hour. All of the cells were then plated on an LB-agar plate that contained 100 μg/mL carbenicillin or 100 μg/mL ampicillin and 34 μg/mL chloramphenicol. The plate was then incubated in a 37°C oven for 16-20 hours.

Both carbenicillin/ampicillin and chloramphenicol were used to maintain the correct genotype; the plasmid conferred resistance to carbenicillin/ampicillin and the
Rosetta™ 2(DE3) cells conferred resistance to chloramphenicol. Carbenicillin is a more stable analogue of ampicillin that is less resistant to degradation by β-lactamase and is also more stable as the pH decreases. Sometimes cells will lose the plasmid containing the protein to be expressed (the plasmid also contains the gene conferring antibiotic resistance) if all of the antibiotic degrades in the culture during growth. Plasmid loss was not observed here, and both antibiotics were used interchangeably for cultures grown in LB media. For producing the isotopically enriched proteins, only carbenicillin and not ampicillin was used in the cultures. Additionally, chloramphenicol was not used for expressions done in minimal media; this was done to reduce the stress on the bacteria.

### 2.4.4.2 Non-labeled Protein Expression

Non-isotopically labeled proteins were used for all experiments not involving NMR. These proteins include WLN1-2, WLN1-3, WLN1-4, WLN1-6, WLN4, WLN1-4 G333R, and HAH1. These proteins were expressed by Rosetta™ 2 (DE3) cells that were grown in LB media containing 100 μg/mL carbenicillin or 100 μg/mL ampicillin and 34 μg/mL chloramphenicol.

Starter cultures were prepared by inoculating cells from a single colony into 5 mL of LB media. The cultures were placed in 37°C shaker (250 rpm) for 8-10 hours before being used to inoculate a 50 mL LB culture, which was then placed in the 37°C shaker (250 rpm) and grown for 8-10 hours.

Protein expression was done in a 2.8 L flask that contained 500 mL of media; this ratio of media volume to flask volume provided adequate aeration for the bacteria. The media was inoculated with an aliquot from the 50 mL culture and incubated in the 37°C shaker (250 rpm) until reaching an OD₆₀₀ of 0.6-0.8. Protein expression was initiated by
the adding IPTG to a final concentration of 1 mM. The cells were induced for either 4-5 hours at 37ºC or 16-18 hours at room temperature (~22ºC), depending on the availability of the shakers and time constraints. Since the thermostat on the shaker was unable to reliably maintain a temperature < 26ºC, the room temperature inductions were done by keeping the lid of the shaker open and exposing the shaker to the lab atmosphere. The temperature of the lab remained fairly constant at ~22 ºC during the course of the inductions.

Following induction, the cells were harvested by centrifugation at 5,000 rpm for 15 minutes at 4ºC. The cell pellets were either immediately lysed to extract the proteins or stored at -20ºC pending purification.

2.4.4.3 15N and 13C,15N Labeled Protein Expression

13C,15N labeled versions of WLN1-2 and WLN4 were prepared for multidimensional NMR experiments; 15N labeled WLN4 was also prepared as well. In order for the appropriate isotopic labels to be incorporated into the proteins, M9 minimal media was supplemented with 15N-ammonium chloride to incorporate 15N and 13C-glucose to incorporate 13C. The recipe for this media is shown in Table 2.12.
Table 2.12: Recipe for M9 minimal media. The final volume was adjusted to one liter using deionized water (18.2 MΩ•cm).

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Amount / Liter of media</th>
</tr>
</thead>
<tbody>
<tr>
<td>M9 Stock Solution (5x)</td>
<td>200 mL</td>
</tr>
<tr>
<td>MgSO₄ (1 M)</td>
<td>1 mL</td>
</tr>
<tr>
<td>CaCl₂ (1 M)</td>
<td>0.3 mL</td>
</tr>
<tr>
<td>MEM Vitamin Mix (100x)</td>
<td>10 mL</td>
</tr>
<tr>
<td>¹³C-glucose (or unlabeled glucose)</td>
<td>2g</td>
</tr>
<tr>
<td>¹⁵N-ammonium chloride</td>
<td>1g</td>
</tr>
</tbody>
</table>

The M9 stock solution (5x) was prepared by dissolving 15g KH₂PO₄, 42.5g Na₂HPO₄ · 2H₂O, and 2.5g NaCl in a total volume of 1 L of deionized water (18.2 MΩ•cm). This solution was autoclaved and then stored at room temperature. All other reagents were sterile filtered using a 0.2 μm polyethersulfone membrane (Pall Corporation, Part Number 4602) prior to being added to the media. Deionized water (18.2 MΩ•cm) was autoclaved and used as a diluent for the minimal media. To ensure adequate aeration, 500 mL of media was used in 2.8 L flasks.

A new transformation was performed each time labeled protein was prepared; the transformation procedure is described in Section 2.4.4.1. Using bacteria to express isotopically labeled protein required several intermediate cultures to acclimate the bacteria to the harsher conditions imposed by the minimal media and ¹³C-glucose and ¹⁵N-ammonium chloride, as bacteria directly transferred to this media from an LB-agar plate would be unlikely to grow. A series of cultures was prepared, as shown in the Figure 2.9 below. All of the cultures in this figure were grown in a shaker that was maintained at 37°C and 250 rpm; carbenicillin (100 μg/mL) was used in all cultures.
**Figure 2.9:** The series of starter cultures grown to express $^{15}$N and $^{13}$C, and $^{15}$N labeled proteins. If only $^{15}$N labeling was required, then unlabeled glucose was used in lieu of the $^{13}$C-glucose.

The entire 30 mL culture was then used to inoculate 500 mL of M9 media that contained the appropriately labeled precursors. These large scale cultures were grown in a shaker at 37°C until the OD$_{600}$ reached between 0.6-0.8. Protein expression was initiated with the addition of IPTG (1 mM, final concentration), and the temperature was reduced to room temperature (~22°C) by keeping the lid of the shaker open. After 16 hours, the cells were pelleted by centrifugation at 5,000 rpm for 15 minutes at 4°C. If the proteins were not immediately extracted from the pellets, then the pellets were stored at -20°C.
2.4.5 Protein Extraction and Purification

All of the proteins expressed in Section 2.4.4 were purified using the same purification steps. An ÄKTA FPLC was used in conjunction with the first four columns shown in Table 2.4. Extraction of the proteins occurred immediately prior to purification.

2.4.5.1 Extraction of the WLN Constructs

Prior to purifying the proteins, they first had to be extracted from the cell pellets. If the pellets were frozen, they were thawed at room temperature on the bench for 15-25 minutes before proceeding. Bugbuster® (EMD Millipore) was used to chemically disrupt the cell membranes and release the soluble protein from the cells; 5 mL of this reagent was used for every 1g of cell pellet, as specified by the manufacturer’s instructions (EMD Millipore, Technical Bulletin 245). Benzonase was used to reduce the viscosity of the Bugbuster®/protein solution and increase protein extraction by degrading DNA and RNA; 1 μL of benzonase was used for every 5 mL of Bugbuster®. The exact concentration of the benzonase purchased from Sigma varied, though it was guaranteed to be ≥ 250 U/μL; typically, 5-8 μL of benzonase were used per 1 L of cell culture. After resuspending the cell pellet in the Bugbuster® and benzonase, the solution was incubated on a table top shaker (50 rpm) at room temperature for 30 minutes. Afterward, the cellular debris was pelleted by centrifugation at 15,000 rpm at 4°C for 15 minutes. The supernatant was decanted into a clean tube and centrifuged again under the same conditions to remove any residual particulates. The supernatant was then decanted into a clean tube and ready for purification.
2.4.5.2 Purification of the WLN Constructs

As all of the Wilson disease protein constructs used here were cloned into the pET-32Xa/LIC vector, they were all expressed as fusion tags containing thioredoxin and a His$_6$ tag. Consequently, the purification steps were the same for all proteins, with the only exception being that a Superdex 200 pg (preparative grade) column was used to purify WLN1-6 instead of the Superdex 75 pg column used for the other, smaller proteins. The purification scheme is summarized below in Figure 2.10. Protein elution was monitored at 254 nm for all columns.

Figure 2.10: The purification scheme used to purify the WLN constructs. Shown in parentheses is the sample that was loaded onto the column; this sample came from the preceding purification step.
Before loading the proteins onto the HisPrep column, the column was first cleaned with 5 column volumes (CV, 1 CV = 20 mL) of Buffer B (50 mM Tris, pH 8.0, 500 mM NaCl, 500 mM imidazole) followed by an equilibration with 5 CVs of Buffer A (50 mM Tris, pH 8.0, 500 mM NaCl, 30 mM imidazole). The proteins were then loaded onto the column, and the column was washed with 10 CVs of Buffer A to remove unbound proteins that did not contain the His6 tag. Elution was achieved using a gradient of 0-100% Buffer B over 5 CVs. Fractions containing the fusion protein were pooled following analysis by SDS-PAGE to confirm the presence of the protein.

After the initial HisPrep purification, the proteins were exchanged into TEV digestion buffer (50 mM Tris, pH 8.0, 0.5 mM EDTA) using a Desalting column, which was pre-equilibrated with 4 CVs (1 CV = 60 mL) of buffer. Elution was performed isocratically over 1.2 CVs using the same buffer. The fractions containing the protein were pooled. Protein quantitation was performed with the BCA assay.

Cleavage of the fusion protein into thioredoxin and the various WLN constructs was achieved using TEV protease in the presence of 1 mM DTT. A 1:80 mg/mg ratio of TEV:fusion protein was used. The digestion proceeded for 14-18 hours at room temperature, with >90% cleavage occurring after this time, as judged by visual inspection of a SDS-PAGE gel. Performing the digestion at 37°C was tried to determine if the digestion time could be shortened, but a significant amount of fusion protein (estimated at >50%) remained after 8 hours when an aliquot was run on an SDS-PAGE gel.

Following digestion, the WLN constructs were purified from the thioredoxin tag and the TEV protease using the HisPrep column. The conditions for column cleaning, equilibration, and elution were the same as those described above, except that only 5 CVs
of Buffer A were used to wash the column after loading the digest. Since the His$_6$ tag was attached only to the thioredoxin tag at this point and not the WLN construct, only the former would bind to the HisPrep column. Consequently, the WLN constructs would elute during the 5 CV wash with Buffer A. TEV protease also contained a His$_6$ tag, enabling it to bind to the HisPrep column. A SDS-PAGE gel was run of the fractions suspected of containing the WLN constructs, and the appropriate fractions were pooled.

To further purify the WLN constructs, gel filtration chromatography was used. For WLN1-2, WLN1-3, WLN1-4, WLN1-4 G333R, and WLN4, the Superdex 75 pg column was used. This column resolves proteins between 3-70 kD (GE Healthcare, Instructions 71-7127-00-AF). WLN1-6 required the use of a column capable of resolving larger proteins, specifically a Superdex 200 pg column (capable of resolving proteins between 10-600 kD). The Superdex columns were equilibrated with 2 CVs (1 CV =320 mL for the Superdex 75 pg column and 120 mL for the Superdex 200 pg column) of gel filtration buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 10 mM DTT, 10 mM EDTA). Elution was done isocratically with gel filtration buffer over 1.2 CVs. SDS-PAGE was used to determine which peaks contained the WLN constructs as well as the relative purity. After pooling the appropriate peaks, the sample was stored at -80°C for at least 3 days. Storing the proteins in the presence of 10 mM EDTA was determined to significantly decrease the degradation observed with several of the WLN constructs, particularly WLN1-2. The exact mechanism by which the EDTA inhibited the protein degradation was not determined.

The final step in the purification process was exchanging the proteins into phosphate buffer (50 mM, pH 7-7.5) using the Desalting column, which was equilibrated...
with 4 CVs (1 CV = 60 mL) of buffer. The protein was eluted isocratically over 1.2 CVs. The fractions containing the protein were pooled, the volume reduced to 2-4 mL, and the protein quantitated using either the BCA assay (Pierce, Rockford, IL) with BSA standards or the Bradford Assay (Bio-rad, Hercules, CA) with IgG standards (4, 5). The protein was aliquoted into 0.5 mL portions and frozen at -80°C pending use.

At several stages in the purification process, the volume of the protein needed to be reduced. This was done by using an Amicon Ultrafiltration Cell, equipped with either a 3 kD or 10 kD MWCO membrane composed of regenerated cellulose (EMD Millipore, Part Numbers PLBC04310 and PLGC04310, respectively). Note that the filters used for the Amicon Cell are treated with glycerol. While this compound would not have adversely affected most of the experiments done in this study, it would pose a significant problem with the samples used for NMR, because it contains NMR observable protons. Consequently, a more stringent cleaning procedure was used for membranes that would be exposed to proteins being analyzed by NMR. The glycerol was removed by running 50 mL of 1 M NaCl through the membrane, followed by 150 mL of sterile, deionized water. For non-NMR destined samples, the membranes were rinsed with 100 mL of sterile, deionized water.

\[\text{2.4.5.3 Metallation of the WLN Constructs}\]

Copper loading of the WLN constructs was done \textit{in vitro} in an anaerobic chamber (Vac Atmosphere Company) following purification. The constructs were diluted to 1-8 \(\mu\text{M}\) using degassed 50 mM sodium phosphate buffer (pH 7.5) / 150 mM NaCl in a 50 mL Amicon ultrafiltration device that was equipped with either a 3 or 10 kD MWCO membrane. DTT was added to a final concentration of 1 mM, and the constructs were
reduced (with stirring) for 30 minutes. A slightly substoichiometric amount of copper(I), in the form of tetrakis(acetonitrile)copper(I) hexafluorophosphate, was slowly added to the constructs dropwise over the course of 10-15 minutes. A substoichiometric amount of copper was used because protein precipitation was observed when stoichiometric or excess copper was added to the multi-domain constructs. After the addition of the copper, the protein was allowed to continue stirring for 30 minutes to allow the copper to bind. Unbound copper and DTT was then removed either by purifying the protein on a PD-10 desalting column (part number 17-0851-01, GE Healthcare) or by using the Amicon device to reduce the volume and then adding fresh buffer and reducing the volume again (total of three times).

Copper loading was quantitatively determined by using inductively coupled plasma atomic emission spectroscopy (ICP-AES). A Perkin Elmer Optima 2100DV ICP-AES instrument was used for this analysis. A series of copper standard solutions were prepared and used to construct a calibration curve. Indium was added as an internal standard to both the copper standard solutions and the protein solutions. The analysis was done by graduate students in the laboratory of Dr. Carla Koretsky at Western Michigan University.

2.4.5.4 Extraction and Purification of HAH1

HAH1 was extracted from the cells using the freeze-thaw method to disrupt the cell membrane and release the cellular contents. Pellets were frozen in liquid nitrogen for five minutes, then thawed in cool water for 15 minutes; this cycle was repeated two more times (for a total of three freezing and three thawing cycles). Following the end of the final thaw, 20 mL of extraction buffer (20 mM MES, pH 5.5, 1 mM EDTA, 5 mM DTT)
was added per liter of culture. The pellet was resuspended in the buffer by placing it on a platform shaker for one hour, followed by centrifugation at 15,000 rpm for 15 minutes.

Following centrifugation, the supernatant was loaded onto two ion exchange columns that were set up in tandem: a DEAE sepharose column (GE Healthcare) followed by a CM sepharose column (GE Healthcare). The columns had been equilibrated with 10 CVs (1 CV = 60 mL for each column) of binding buffer (20 mM MES, pH 6). The order of the columns and the pH of the binding buffer were selected to take advantage of HAH1’s isoelectric point, which is ~7.5. In the binding buffer, HAH1 had a slight positive charge and bound to the CM column. Many of the other proteins in the supernatant were either negatively charged at pH 6 and bound to the DEAE column (the first column) or were hydrophobic and not retained by either column. After the supernatant was loaded, 10 CVs of binding buffer were run through the columns to remove unbound proteins. The DEAE column was then removed, and HAH1 was eluted from the CM column using a linear gradient of 0-100% elution buffer (20 mM MES, pH 6, 1 M NaCl) over 8 CVs. Aliquots of the peaks were run on an SDS-PAGE gel, and the fractions containing HAH1 were pooled. Further purification was achieved by using a Superdex 75 26/60 column with 50 mM sodium phosphate, pH 7.5, 150 mM NaCl buffer. The relevant fractions were combined, concentrated, and aliquoted into 0.5 mL portions that were frozen at -80°C pending use.

2.5 Biophysical Characterization

Several different techniques were used to characterize the WLN constructs, including high resolution gel filtration, laser light scattering, circular dichroism, differential scanning calorimetry, and nuclear magnetic resonance spectroscopy.
2.5.1 High Resolution Gel Filtration on the WLN Constructs

High resolution gel filtration (HRGF) was used to determine the apparent molecular weight of the WLN constructs. In gel filtration chromatography, analytes are injected onto a column and eluted isocratically. In contrast to other modes of chromatographic separation, there is no chemical interaction between the column and the analyte in gel filtration chromatography. Instead, separation is achieved on the basis of the analyte’s ability to the physically access the porous beads and diffuse inside them. Smaller and more globular molecules are better able to access the pores than larger molecules and therefore have a greater distance to traverse before reaching the end of the column. Consequently, smaller molecules elute at greater volumes (i.e. they take more time to travel the greater distance) than larger molecules, which are precluded from entering the pores by their size. Globular proteins elute at volumes consistent with their actual molecular weight while proteins that are elongated (i.e. more ellipsoid) elute at volumes consistent with higher molecular weight proteins (6).

Protein standards from calibration kits were injected onto a Superdex 75 HRGF 10/300 column, which optimally separates proteins between 3-70 kD. A calibration curve was constructed by plotting the partition coefficient ($K_{av}$) against the logarithm of the molecular weight. $K_{av}$ was calculated with Equation 2.1 below, where $V_c$ is the geometric column volume (23.5 mL), $V_e$ is the elution volume, and $V_0$ is the void volume (GE Healthcare, Technical Bulletin 18-102218-AK).

$$K_{av} = \frac{V_e - V_0}{V_c - V_0}$$ [2.1]
\( V_0 \) is the volume at which an analyte will elute if it is completely excluded from entering the space inside the porous beads because it is too large; blue dextran, a polymer of 2,000 kD, was used for this purpose. After plotting the points, a linear trend line was fitted to the curve using Microsoft Excel 2007. The equation of the trend line was then solved for the molecular weight of the protein (the inverse log of \( x \)), since \( K_{av} (y) \) was known.

The lyophilized proteins from the calibration kits were resuspended in gel filtration buffer (50 mM sodium phosphate, pH 7.0, 150 mM NaCl, 10 mM DTT) to concentrations between 10-20 mg/mL; details of the standards are shown in the table below. Different combinations of the standards were used depending on their availability at the time of analysis; four standards were used for each analysis. Between 0.1-0.4 mg of each protein standard and WLN construct (in a volume of 100 μL) were injected onto the column after it had been equilibrated with 2 CVs of buffer. Each protein was eluted isocratically with 1.2 CVs of buffer. The WLN constructs were diluted with gel filtration buffer prior to injecting onto the column, and fresh DTT was added (10 mM final concentration) to ensure that the proteins were reduced. At least two injections of each protein standard and WLN construct were performed, and the average elution volume was used to calculate \( K_{av} \). The elution volume was taken from the peak apex, which was found using the Unicorn software (version 3.21.02) that was provided with the ÄKTA FPLC.
Table 2.13: The protein standards used to construct the calibration curves for HRGF.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Molecular Weight (Daltons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>6,500</td>
</tr>
<tr>
<td>Cytochrome c</td>
<td>12,400</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>13,700</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>29,000</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43,000</td>
</tr>
<tr>
<td>Albumin</td>
<td>66,000</td>
</tr>
</tbody>
</table>

HRGF can also be used to monitor the formation of complexes between proteins (7–9). This technique was used to investigate whether an interaction was occurring between WLN1-3 and WLN4, as results from circular dichroism experiments on WLN1-4 suggested that WLN1-3 and WLN4 were interacting. If an interaction was occurring between the constructs, the presence of a species with a mass approximately equal to that of WLN1-4 might be observable.

Equimolar amounts (100 μM) of both constructs were incubated in 500 μL of 50 mM sodium phosphate buffer, pH 7.5. No DTT was added (in case the interaction was mediated by a disulfide bond) and the sample was left at room temperature for 24 hours. Aliquots (100 μL) were injected onto the Superdex 75 HRGF column after 1, 5 and 24 hours. For comparative purposes, WLN1-3, WLN1-4, and WLN4 were also run on the column to verify their elution volumes. The same gel filtration buffer (without DTT, however), equilibration and elution volumes, and flow rate that were described above were used.

2.5.2 Gel Filtration Analysis of the Association State of HAH1

Gel filtration chromatography was used to investigate the influence of pH and Hg(II) binding on the association state of HAH1. A Superdex 75 HiLoad 16/600 column
was used in conjunction with one of three different buffers: pH 7.5 and 8.5 buffers contained 100 mM sodium phosphate, 200 mM sodium chloride, 1 mM TCEP, and the pH 9.4 buffer contained 100 mM CHES, 200 mM sodium chloride, 1 mM TCEP. TCEP was included in the buffer in order to prevent the oxidation of cysteines at elevated pH, as oxidation could result in the formation of HAH1 dimers. The gel filtration studies were conducted at 4°C (the FPLC was in a refrigerator), necessitating that the buffers be prepared (at room temperature) such that they would have the appropriate pH at 4°C. Prior to each injection (0.5 mL), the column was equilibrated with 1 CV (~120 mL) of buffer using a flow rate of 1 mL/min. Samples were eluted isocratically with 1.1 CV of buffer and a 1 mL/min flow rate. The elution volume of each protein was considered to be the peak apex, which was found using the Unicorn Software.

A mixture of protein standards (aprotinin, ribonuclease A, carbonic anhydrase, and ovalbumin, see Table 2.13 for details) was analyzed with each of the three buffers. The mixture contained between 0.75-1 mg of each protein in a volume of 0.5 mL. Duplicate injections were made for both the protein standards HAH1, and the average elution volume was used for the calculations. At each pH, a calibration curve was constructed from the standards using the procedure discussed in the previous section; this curve was used to calculate the apparent molecular weights of apo- and Hg(II)HAH1 at each pH.

2.5.3 Light Scattering

Light scattering (LS) is an analytical technique that can be used to determine several important biophysical characteristics of a protein, including the hydrodynamic radius ($R_h$), the radius of gyration ($R_g$), the molecular weight, and the shape. This non-
invasive technique requires small amounts of protein (~100-300 μg, depending on the molecular weight of the protein) and yields the oligomerization state of the sample (10).

There are two types of light scattering techniques: static and dynamic. Both techniques take advantage of the fact that molecules in solution scatter light in a manner that is dependent on their size and shape; the scattering intensity is also influenced by the Brownian movement of the particles (11). The primary difference between these techniques is the time scale. Dynamic light scattering (DLS) operates on the microsecond timescale and measures the rapid fluctuations in the intensity of scattered light that occurs in a small volume of solution; the $R_H$ and translational diffusion coefficient are two values that can be extracted from this technique. In contrast, static light scattering (SLS) operates on the second timescale, and the intensity of the scattered light over this time is averaged (i.e. the system being studied is static, not dynamic); consequently, the transient fluctuations in scattered light that are observed in DLS are absent in SLS. The molecular weight of the molecule, $R_g$, and the second virial coefficient ($A_2$) can be calculated from SLS data.

DLS takes advantage of the fact that rapid fluctuations in the intensity of scattered light reflect the diffusion rate of the particles, i.e. the Brownian motion (11). This diffusion rate, known as the translational diffusion coefficient ($D$), is a variable in the autocorrelation function; $D$ is calculated from the raw data through curve fitting. The autocorrelation function describes how the signal intensity changes with time and indicates the likelihood that the same signal intensity will be encountered at some later time. Since this likelihood decreases with time, the autocorrelation curve takes the form of an exponential decay curve. Once $D$ is known, the Stokes-Einstein Equation
(Equation 2.2) can be solved for $R_H$. In this equation, $k$ is the Boltzmann constant, $T$ is the temperature, $D$ is the translational diffusion coefficient, and $\eta$ is the solvent viscosity.

$$D = \frac{kT}{6\pi\eta R_H} \quad [2.2]$$

The foundation of SLS is that the amount of light scattered by a sample is directly proportional to the product of the sample concentration and the sample’s weight-average molar mass (11). Zimm’s formalism of the Rayleigh-Debye-Gans model for the scattering of light by a dilute polymer solution, shown below in Equation 2.3, established this correlation (11, 12). In this model, $K*c/R(\theta)$ is plotted against $\sin^2(\theta/2)$ and the data is fit to an equation in the form of $\sin^2(\theta/2)$. The molar mass is given by the y-intercept while the radius of gyration is given by the slope. The radius of gyration is the root mean square distance to the particle’s center of gravity. If the concentration of the sample as it elutes from the gel filtration column can be independently determined, then Equation 2.3 can be solved by ASTRA software (Wyatt Technology, Santa Barbara, CA) using the Deybe model. A refractive index detector is frequently used to measure the concentration.

$$\frac{K*c}{R(\theta)} = \frac{1}{MW \cdot P(\theta)} + 2A_2c \quad [2.3]$$
Table 2.14: Definition of the variables used in Equation 2.3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K^*$</td>
<td>An optical parameter that is equal to $4\pi^2 n^4 (dn/dc)^2/(I_0^4 N_A)$</td>
</tr>
<tr>
<td>$c$</td>
<td>The sample concentration (g/mL)</td>
</tr>
<tr>
<td>$R(\theta)$</td>
<td>The excess intensity of scattered light at angle $\theta$</td>
</tr>
<tr>
<td>$MW$</td>
<td>The weight-average molar mass</td>
</tr>
<tr>
<td>$P(\theta)$</td>
<td>The angular dependence of the scattered light</td>
</tr>
<tr>
<td>$A_2$</td>
<td>The second virial coefficient (measures sample/solvent interaction)</td>
</tr>
<tr>
<td>$N_A$</td>
<td>Avogadro’s number</td>
</tr>
<tr>
<td>$I_0$</td>
<td>Wavelength of scattered light in a vacuum (cm)</td>
</tr>
</tbody>
</table>

The amount of light that is scattered is dependent on the sample’s polarizability: the more polarizable the sample, the greater the intensity of the scattered light. This is accounted for by $dn/dc$, where $dn$ is the change in the refractive index and $dc$ is the change in concentration.

DLS and SLS analyses on WLN1-2, WLN1-3, and WLN1-4 were done at the Keck Foundation Biotechnology Resource Facility at Yale University by facility personnel. The samples were injected onto a Superdex 75 HRGF column (WLN1-2) or a Superdex 200 HRGF (WLN1-3 and WLN1-4) that was connected to an Agilent 1200 HPLC (Agilent Technologies, Wilmington, DE). The column effluent was monitored by three detectors: an OPTILAB rEX Refractive Index Detector (Wyatt Technology, Santa Barbara, CA), a DAWN-HELEOS Light Scattering Detector, and a Waters 996 Photodiode Array Detector (Waters Corporation, Milford, MA). The buffer was 50 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA and 1 mM DTT; a flow rate of 0.75 mL/min was used for the Superdex 75 column while a flow rate of 0.5 mL/min was used for the Superdex 200 column. Chemstation software (Agilent Technologies, Wilmington,
DE) controlled the HPLC and collected the UV-Vis data from the PDA while ASTRA software (Wyatt Technology, Santa Barbara, CA) controlled the other two detectors.

All samples submitted for analysis were between 4.8-5 mg/mL in 50 mM sodium phosphate buffer, pH 7.5. Prior to injection onto the column, the samples were filtered through a 0.22 μm Durapore membrane (UFC30GVNB, EMD Millipore) to remove particulates. The filtrates were then diluted to a concentration of 2.5-3 mg/mL using the gel filtration buffer, and 0.35 mL of each sample was injected.

The weight average molar mass for each construct was calculated across the entire peak and was determined using the ASTRA software (version 5.4.3.10) and the Debye model (10). Bovine serum albumin and trypsin inhibitor were used to normalize the intensity of the scattered light 90° to the sample cell. The hydrodynamic radius of each construct was calculated using ASTRA software in conjunction with the LS data taken in 2 second slices.

2.5.4 Circular Dichroism

The phenomenon of circular dichroism (CD) is observed when optically active molecules differentially absorb left- and right-handed circularly polarized light. In proteins, the peptide bond is a chromophore and absorbs circularly polarized light in the far UV region (190-260 nm). Maximum absorption of light occurs at specific wavelengths and is determined by the geometry of the peptide bond. Consequently, different secondary structural elements absorb circularly polarized light at unique wavelengths. Peptide bonds in α-helices have negative peaks at 222 and 208 nm and a positive peak at 193 nm while peptide bonds in β-sheets have a negative peak at 218 nm and a positive peak at 195 nm; bonds found in random coils have a negative peak at 195
nm (13–15). The negative peaks at 222 nm in α-helices and 218 nm in β-sheets are due to the n→π* transition of the peptide bond while the negative peak at 208 nm in α-helices and the positive peak at 195 in β-sheets are attributed to the π→π* transition (13–15).

CD experiments were performed using a Jasco J-815 Spectropolarimeter. The temperature was regulated by a Peltier Thermostated Cell Holder (Jasco, PTC-423) that was cooled with a water pump. Prior to use, the instrument was purged with nitrogen (high purity grade) for 15 minutes at 15 psi to removed oxygen, as the high voltage produced when running the instrument generates ozone that damages the optics. The same flow rate of nitrogen was maintained while the instrument was in use.

Two types of CD experiments were performed: one at fixed temperature and scanning wavelength, and the other at scanning temperature and fixed wavelength. The former was used to monitor the denaturation of the WLN constructs by guanidine hydrochloride (GuHCl), and the latter was to monitor the thermal denaturation of the constructs. The different experiments utilized different cuvettes and data acquisition parameters.

GuHCl unfolding for all constructs was performed on samples that contained 10-20 μM protein, 5 mM phosphate buffer (pH 7.5), and concentrations of GuHCl ranging from 0--7M. A series of samples was prepared of each protein that contained the same amount of protein and buffer, but varying amounts of GuHCl. The maximum concentration of GuHCl that could be attained depended on the concentration of both the protein and GuHCl stock solutions as well as the requisite final protein concentration. A CD signal of at least ~30 millidegrees (mdeg) at 222 nm was desired for each sample that contained no GuHCl (i.e. 0 M GuHCl), and the concentration of protein required to
generate this signal varied among the constructs, with smaller constructs requiring higher concentrations. Samples were equilibrated at room temperature for 2 hours before the CD spectra were recorded.

Spectra were collected from 190-260 nm, with a scanning rate of 100 nm/min, a data pitch of 0.1 nm, a bandwidth of 1 nm, and a digital integration time of 1 second. Three scans were accumulated during each run, with the final spectrum being an average of the three scans. The samples were placed in a 1 mm rectangular quartz cuvette, which was rinsed three times with filtered deionized water between samples. The temperature was held constant at 25°C.

The 8 M GuHCl stock solution was prepared by weighing the appropriate amount of GuHCl into a 50 mL conical tube and diluting to the requisite volume with deionized water. Since GuHCl is hygroscopic, however, a more accurate means of determining the concentration was required. This was achieved by measuring the refractive index of the GuHCl solution and water on a Bausch and Lomb Abbe 3L refractometer. Equation 2.4 was then used to calculate the exact molarity of the solution, where \( \Delta N \) is the difference in the refractive index between the GuHCl solution and water (16).

\[
\text{Concentration (Molarity)} = 57.147\Delta N + 38.68\Delta N^2 - 91.60\Delta N^3
\] [2.4]

The raw CD signal (mdeg) was converted to units of mean residue molar ellipticity ([\( \theta_{MRW} \)], a unit commonly used for reporting the ellipticity of proteins. This unit specifies the ellipticity for protein residues instead of entire proteins and is useful for comparing proteins of different sizes. Equation 2.5 was used to convert mdeg to mean residue ellipticity; \( \theta \) is the CD signal in mdeg, \( C_{MR} \) is the mean residue concentration.
(CMR = molar protein concentration * number of residues), and \( l \) is the path length of the cuvette in centimeters. The unit for \([\theta]_{\text{MRW}}\) is deg * cm\(^2\) * dmol\(^{-1}\).

\[
[\theta]_{\text{MRW}} = \frac{100 \times \theta}{C_{\text{MR}} \times l}
\]  

[2.5]

Several thermodynamic parameters were calculated from the GuHCl unfolding curves, including the free energy of unfolding (\(\Delta G_{H_2O}^\circ\)), the \(m\) value, and the concentration midpoint (\(C_m\)). \(\Delta G_{H_2O}^\circ\) is the free energy of unfolding extrapolated back to zero denaturant (details given below); a positive value indicates that the folded protein is more stable than the unfolded one. The \(m\) value reflects the amount of the protein’s surface area that is exposed to the solvent upon denaturation and shows how the free energy of unfolding is dependent upon the concentration of the denaturant (17). The \(C_m\) is the denaturant concentration at which 50% of the protein is folded and 50% unfolded. These parameters were all determined by fitting the curve of denaturant concentration versus the CD signal at 222 nm to either a two- or three-state equation.

In a two-state unfolding mechanism, a protein exists in either the native (N) state or the denatured state (D). The transition from the native to the denatured state is governed by an equilibrium constant, \(K_{N \leftrightarrow D}\), as shown in Equation 2.6 below. At low concentrations of denaturant, the native state is favored. As the denaturant concentration increases, the forces that stabilize the protein’s structure are disrupted and the protein begins to unfold. At the highest concentration of GuHCl (~7 M), the protein is assumed to be completely unfolded. The observed CD signal at any concentration of GuHCl results from a linear combination of the amount of protein that is in the fully native state and the fully denatured state; any intermediates that form during denaturation exist.
transiently and accumulate to negligible levels. A two-state unfolding mechanism is indicated by a sigmoidal curve when the denaturant concentration versus the CD signal at 222 nm is plotted.

\[ \text{K}_{N \rightarrow D} \]

Curves that displayed an apparent two-state unfolding mechanism were analyzed as follows. For any concentration of guanidine hydrochloride [GuHCl], the free energy of unfolding associated with the transition of the protein from its native to a denatured state was assumed to vary according to Equation 2.7 (18). In this equation, \( \Delta G^\circ_{H_2O} \) is the free energy of unfolding extrapolated back to zero denaturant and \( m \) is the slope of a plot of \( \Delta G \) versus [GuHCl].

\[ \Delta G = \Delta G^\circ_{H_2O} - m[\text{GuHCl}] \]  \[2.7\]

Since the observed CD signal \( S_{\text{obs}} \) results from the linear combination of the fraction of the protein that in its native form \( (f_N) \) and the fraction that is denatured \( (f_D) \), then Equation 2.8 results,

\[ S_{\text{obs}} = S_N f_N + S_D f_D \]  \[2.8\]
where $S_N$ is the signal of the native protein (i.e. at 0 M GuHCl) and $S_D$ is the signal of the fully denatured protein (the highest [GuHCl] that was used in each experiment, ~7 M). As all of the protein is either in the native or denatured form, $f_N + f_D = 1$. Taking into account that $\Delta G = -RT \ln(f_N/f_D)$, Equation 2.9 can be written (19); $R$ is the gas constant and $T$ is the temperature.

$$S_{obs} = \frac{S_N + S_D e^{-(\Delta G_{H_2O}-m[GuHCl])/RT}}{1+e^{-(\Delta G_{H_2O}-m[GuHCl])/RT}} \tag{2.9}$$

Proteins that displayed a sigmoidal curve when the [GuHCl] was plotted against the CD signal were considered to unfold via an apparent two-state model. The data were fit to the above equation using IGOR Pro 6 (Wavemetrics, Portland, OR), permitting the determination of $\Delta G^c_{H_2O}$, and the $m$ value. Fitting the data to IGOR’s built in sigmoid function was used for determining $C_m$.

In a three-state unfolding pathway, a stable intermediate (I) is formed as the protein transitions from its native (N) to a denatured state (D). As Equation 2.10 below shows, the transitions from $N \rightarrow I$ and $I \rightarrow D$ each have their own equilibrium constants, $K_{N\rightarrow I}$ and $K_{I\rightarrow D}$, respectively. Both transitions also have their own $\Delta G$ and $m$ values, $\Delta G_{N\rightarrow I}$ and $m_{N\rightarrow I}$ and $\Delta G_{I\rightarrow D}$ and $m_{I\rightarrow D}$, respectively. A plot of denaturant concentration vs. CD signal for a protein that unfolds via a stable intermediate has a double sigmoidal curve. Proteins displaying such a curve were considered to unfold via an apparent 3-state model and were analyzed as discussed below.

$$N \xleftrightarrow{K_{N\rightarrow I}} I \xleftrightarrow{K_{I\rightarrow D}} D \tag{2.10}$$
The accumulation of an intermediate requires a more complex equation to fit the curve. The equation, shown on the following page as Equation 2.11, is similar to Equation 2.9 but contains parameters to account for the presence of the intermediate. $S_N$, $S_D$, $R$, and $T$ have the same definitions as they did in the two-state model. $\Delta G_{N\rightarrow I}$ and $m_{N\rightarrow I}$ and $\Delta G_{I\rightarrow D}$ and $m_{I\rightarrow D}$ were calculated using the above equation. The concentration midpoints for the $N\rightarrow I$ and $I\rightarrow D$ transitions were found by using Equation 2.12, where $x_{half1}$ and $x_{half2}$ were the first and second transitions, respectively. The other parameters in this equation did not correspond to any values of interest. Both equations were fit to the curves using IGOR Pro 6.
\[ S_{obs} = S_N + S_I \exp\{-(\Delta G_{N \rightarrow I} - m_{N \rightarrow I}[\text{GuHCl}])/RT\} + S_D \exp\{-(\Delta G_{I \rightarrow D} - m_{I \rightarrow D}[\text{GuHCl}])/RT\} \]
\[ \frac{1 + \exp\{-(\Delta G_{N \rightarrow I} - m_{N \rightarrow I}[\text{GuHCl}])/RT\} + \exp\{-(\Delta G_{I \rightarrow D} - m_{I \rightarrow D}[\text{GuHCl}])/RT\} \exp\{-(\Delta G_{I \rightarrow D} - m_{I \rightarrow D}[\text{GuHCl}])/RT\}}{1 + \exp\{-(\Delta G_{N \rightarrow I} - m_{N \rightarrow I}[\text{GuHCl}])/RT\} + \exp\{-(\Delta G_{I \rightarrow D} - m_{I \rightarrow D}[\text{GuHCl}])/RT\} + \exp\{-(\Delta G_{I \rightarrow D} - m_{I \rightarrow D}[\text{GuHCl}])/RT\}} \]

[2.11]

Variable Result = base + (max1)(1 + \exp\{-(x - \text{half1})/\text{rate1}\})
Result = (max2)/(1 + \exp\{-(x - \text{half2})/\text{rate2}\})
\[ f(x) = \text{result} \]

[2.12]
Thermal unfolding studies were done on the WLN constructs as well. In these experiments, the temperature varied while the wavelength remained constant. Unfolding was monitored from 25-97°C, with a gradient of 2°C/min. Upon reaching 97°C, that temperature was held for 3 minutes before the sample was cooled according to the same gradient and the cooling spectra collected. For both the forward and reverse gradients, the band width was 1 nm and the data integration time was 1 second. A small stir bar was added to the cuvette, and the sample was stirred continuously during the experiment. The CD signal was monitored at 222 nm, as α-helices absorb maximally at this wavelength, and each metal-binding domain is known to possess two α-helices. The samples contained 3-7 μM protein and 5 mM phosphate buffer. Higher concentrations of the smaller constructs were required to attain the minimally desired signal of ~30 mdeg. A 1 cm rectangular quartz cuvette was used. No buffer blank was used. The melting temperatures (Tm) were calculated using the Spectra Manager software (version 2.08.01) that was supplied with the instrument.

2.5.5 Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) was used to determine the melting temperature (Tm) of apo WLN1-4 and apo and Cu(I)-WLN4. This analysis was performed by personnel at the Biophysics Core at the University of Colorado-Denver Anschutz Medical Campus. These proteins were analyzed by DSC because a thermal unfolding transition was not observed by CD for WLN1-4 and Cu(I)WLN4. ApoWLN4 was analyzed as well for both verification of the Tm determined by CD and as a reference for comparison of the DSC determined Tm for Cu(I)-WLN4.
In DSC, the heat capacity \( (C_p) \) of a sample and a reference cell is measured as a function of increasing temperature. If the sample undergoes a phase transition (i.e. melts or unfolds), then its heat capacity will change relative to the heat capacity observed in the reference cell, which contains the buffer blank. This change in heat capacity is indicated by a peak in the plot of temperature versus \( C_p \); the area of the peak is equal to the enthalpy change associated with the phase transition.

A MicroCal VP-DSC (GE Healthcare) instrument was used for the analysis. A temperature gradient of 1°C/min, from 20-110°C, was applied. The instrument was controlled using the VPViewer software supplied by the manufacturer. Data analysis was performed using the Microcal Origin software package.

Apo and Cu(I)-WLN4 were each run once. The sample cell was loaded with 0.5 mL of protein (both proteins were 1 mg/mL in 50 mM sodium phosphate buffer, pH 7.5) while the reference cell was loaded with 0.5 mL of 50 mM sodium phosphate buffer, pH 7.5.

Multiple concentrations of apoWLN1-4 were analyzed by DSC: 0.1 mg/mL, 0.25 mg/mL, 0.5 mg/mL, and 1 mg/mL. As with the WLN4 samples, 0.5 mL of WLN1-4 in 50 mM sodium phosphate buffer, pH 7.5, was loaded into the sample cell while 0.5 mL of 50 mM sodium phosphate buffer, pH 7.5 was loaded into the reference cell. A 1 mg/mL solution was used initially, but the protein precipitated during the course of the thermal gradient. The diluted solutions were then tried in an attempt to find a concentration that produced a clearly discernible signal without precipitating. Each of the different concentrations was run once, from 20-110°C with a temperature gradient of 1°C/min.
Prior to submitting the samples for DSC analysis, the copper content of the proteins was determined using ICP-AES (see Section 2.4.5.3 for details). The proteins were quantitated using the BCA assay.

2.5.6 Nuclear Magnetic Resonance Spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy is a powerful technique that can be used to evaluate the structure and dynamic properties of proteins in solution. Single proteins as well as protein-protein interactions can be studied with this non-invasive technique (20–22). Proteins being studied by NMR must be enriched in NMR active isotopes, i.e. $^{13}\text{C}$ and/or $^{15}\text{N}$; large proteins $> 40\text{kD}$ are also frequently enriched with deuterium ($^2\text{H}$) as well. This enrichment is achieved by expressing proteins in minimal media supplemented with the appropriately labeled feedstock. The procedures used for expressing the labeled proteins studied in this research as discussed in Section 2.4.4.3.

All of the NMR data was collected using a Bruker AVANCE II 800 MHz NMR that was equipped with a cryoprobe. This NMR, located at the Lizzadro Magnetic Resonance Research Center at the University of Notre Dame (South Bend, IN), was operated with assistance from facility personnel.

All proteins were analyzed in buffers comprised of 50 mM sodium phosphate, pH 7.5, 10 mM $d_{12}$-EDTA, and 5% D$_2$O. The protein concentration varied according to the particular WLN construct, as shown in Table 2.15 below. While high protein concentrations are desirable, the proteins had a propensity to aggregate over time at elevated concentrations; consequently, lower concentrations had to be used in order to preserve the sample integrity. As a buffer with high conductivity negatively influences the signal to noise ratio when using a cryoprobe, sodium chloride was not used in the
samples to stabilize higher protein concentrations (23). Two (2D) and three (3D) dimensional experiments were performed on WLN4 and WLN1-4. These experiments were run in succession, and samples frequently had to be exchanged for new ones after two to three days due to protein aggregation. A summary of the experiments that were run on the various WLN constructs are shown in Table 2.15.

The non-copper loaded forms of all proteins were analyzed, with the copper content being determined by ICP-AES as discussed in the previous section. The protein concentration was determined for WLN1-2 and WLN1-4 using the BCA assay while WLN4 was quantitated using the Bio-rad Assay. Thin wall NMR tube rated for 1000 MHz (part number 542-PP-8, Wilmad-Lab Glass) were used in all experiments. For experiments performed at temperatures other than 25°C, an external standard of 10 μM sodium 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) was used to reference the chemical shifts. The acquisition parameters varied according to the specific experiment being run.
Table 2.15: Summary of NMR experiments done with WLN1-2, WLN1-4, and WLN4.

<table>
<thead>
<tr>
<th>Protein Construct</th>
<th>Experiment</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{13}$C, $^{15}$N WLN4 (100 μM)</td>
<td>$^1$H-$^{15}$N HSQC</td>
<td>Collected at 25°C, 37°C, and 50°C. Correlates the amide proton and the amide nitrogen of ( i )</td>
</tr>
<tr>
<td></td>
<td>HNCACB</td>
<td>Collected at 37°C. Correlates N and HN of ( i ) with C( \beta ) and C( \alpha ) of ( i ) and ( i-1 )</td>
</tr>
<tr>
<td></td>
<td>CBCA(CO)NH</td>
<td>Collected at 37°C. Correlates N and HN of ( i ) with C( \beta ) and C( \alpha ) of ( i-1 )</td>
</tr>
<tr>
<td></td>
<td>HNCA</td>
<td>Collected at 37°C. Correlates N and HN of ( i ) with C( \alpha ) of ( i ) and ( i-1 )</td>
</tr>
<tr>
<td>$^{13}$C, $^{15}$N WLN1-2 (200 μM)</td>
<td>$^1$H-$^{15}$N HSQC</td>
<td>Collected at 25°C. Correlates the amide proton and the amide nitrogen of ( i )</td>
</tr>
<tr>
<td>$^{13}$C, $^{15}$N WLN1-4 (300 μM)</td>
<td>$^1$H-$^{15}$N HSQC</td>
<td>Collected at 25°C. Correlates the amide proton and the amide nitrogen of ( i )</td>
</tr>
</tbody>
</table>

The backbone resonance assignment of WLN4 was achieved by walking through the C\( \alpha \) chemical shifts and linking stretches of amino acids that were sequentially adjacent. Three experiments were run to collect the requisite data to perform the sequential walking: a $^1$H-$^{15}$N HSQC, which showed all correlations between N(\( i \)) and HN(\( i \)), a HNCACB, which showed all correlations between N(\( i \)), H-N(\( i \)) with C\( \alpha \) and C\( \beta \) of \( i \) and \( i-1 \) and a CBCA(CO)NH, which showed all correlations between N(\( i \)), H-N(\( i \)) with C\( \alpha \) and C\( \beta \) of \( i-1 \) (24, 25). (Figure 2.11 below shows the structure of a generic dipeptide, with the name of each particular atom indicated.)

Since the CBCA(CO)NH showed only the C\( \alpha \) and C\( \beta \) from \( i-1 \) while the HNCACB showed the C\( \alpha \) and C\( \beta \) from both \( i \) and \( i-1 \), the C\( \alpha \) and C\( \beta \) associated with a specific NH could be discerned. This made it possible to group atoms into spin systems,
which could then be linked to other spin systems through sequential walking. These short sequences were then mapped onto the full sequence of WLN4 using CARA, a computer program developed by Rochus Keller and Dr. Kurt Wüthrich at The Swiss Federal Institute of Technology in Zurich, Switzerland (www.cara.nmr.ch). NMR spectra were processed with the NMRPipe Program with the assistance of Dr. Blair Szymczyna.

![Diagram of a generic dipeptide with each atom labeled. Various NMR experiments detect specific atoms of either the i residue and/or the i-1 (preceding) residue in a protein.](image)

**Figure 2.111** Diagram of a generic dipeptide with each atom labeled. Various NMR experiments detect specific atoms of either the i residue and/or the i-1 (preceding) residue in a protein.

### 2.6 Summary

This chapter discussed the materials, reagents, and instruments used in the course of this research. The procedures used to prepare the pET32-WLN1-2, pET32-WLN1-3, pET32-WLN1-4 G333R, and pET32-WLN4 plasmids were detailed. Also discussed were the expression of the labeled and non-labeled protein constructs as well as their subsequent purification. Finally, the biophysical techniques used to characterize the
protein constructs were described. These techniques included high resolution gel filtration, laser light scattering, circular dichroism, differential scanning calorimetry, and nuclear magnetic spectroscopy.
References


CHAPTER 3
PLASMID PREPARATION, PROTEIN EXPRESSION, AND PURIFICATION OF THE WLN CONSTRUCTS

3.1 Introduction

Protein constructs comprised of single or multiple metal-binding domains of human Wilson disease protein have been amenable to expression in *E. coli* (1–7). While the single domain constructs are often readily expressed without the use of solubility tags, many of the multiple metal-binding domain constructs are better expressed as fusions with proteins (solubility tags) such as thioredoxin, maltose binding protein, or glutathione S-transferase. The benefits of expressing the metal-binding domains as fusion proteins instead of by themselves include an increased protein yield and better solubility (i.e. they are less likely to be found in inclusion bodies). Affinity tags such as a His$_6$ tag (i.e. six histidines in a row) are also frequently incorporated into these fusions to provide a convenient means of purification.

This chapter discusses the preparation of the plasmids containing the WLN constructs as well as the expression and purification of the proteins.

3.2 Plasmid Preparation

The Quick Change II XL Site-Directed Mutagenesis Kit was used to create the pET32-WLN1-2, pET32-WLN1-3, and pET32-WLN1-4 G333R plasmids. Details of the primers, PCR parameters, and reagent information can be found in Sections 2.2. and 2.4.1-2.4.2. The pET32-WLN1-2 and pET32-WLN1-3 plasmids were created by using the mutagenesis kit to insert a stop codon in the appropriate location in the pET32WLN1-6 plasmid, which codes for all six metal-binding domains. For example, the pET32-
WLN1-2 plasmid was produced by inserting a stop codon after the end of the second metal-binding domain in the pET32-WLN1-6 plasmid while the pET32-WLN1-3 plasmid was produced by inserting a stop codon after the end of the third metal-binding domain. The pET32-WLN1-4 G333R plasmid was created by using the same site directed mutagenesis kit to introduce a guanine to adenine mutation in the pET32-WLN1-4 plasmid; this missense mutation resulted in a glycine to arginine substitution at position 333, which is located in the linker region between the third and fourth metal-binding domains. Following mutagenesis, the plasmids were used transformed into XL-10 Gold cells (Agilent Technologies); the transformation procedure can be found in Sections 2.4.3. The plasmids were propagated by using cells from single colonies to inoculate 5 mL of liquid LB media that contained 100 μg/mL carbenicillin; multiple 5 mL cultures were prepared. The cells were pelleted after 8-10 hours of growth (in a shaker maintained at 37°C, 250 rpm) by centrifugation at 15,000 rpm in a tabletop centrifuge. After purification using Qiagen’s QiaPrep® Spin Miniprep Kit, the plasmids were digested with BglII (20 U) and BamHI (30 U) and run on a 0.8% agarose gel to verify the presence of the expected fragments. As shown in Figure 3.1 below, all three of the plasmids contained the expected fragment sizes of 5800 bp (pET-32Xa/LIC vector) and 1,731 bp (the WLN1-6 insert). Note that all of the plasmids constructed here contain the same 1,731 bp insert, although the plasmids express different proteins. This is because the plasmids were constructed by inserting a stop codon at the appropriate position instead of cloning only the desired metal-binding domains into the pET-32Xa/LIC vector. After verification of the correct fragment sizes by agarose gel electrophoresis, the plasmids were quantitated by measuring their absorbance at 260 nm and sent for
sequencing at Retrogen (San Diego, CA). The plasmids with the correct sequence were stored at -20°C and used for future experiments.

**Figure 3.1:** Agarose gel (0.8%) of the pET32-WLN1-2, pET32-WLN1-3, and pET32-WLN1-4 G333R plasmids that were digested with BamHI (30 U) and BglII (20 U). The lane contents were as follows: 1) 1 kb markers (New England Biolabs), 2) pET32-WLN1-2, 3) pET32-WLN1-3, 4) pET32-WLN1-4 G333R. The ~1.7 kb fragment in lanes 2-4 is the WLN1-6 insert, the ~5.8 kb fragment is the pET-32Xa vector, and the band at ~7.5 kb is undigested plasmid.

The pET32-WLN1-6 plasmid used as the template DNA was graciously provided by a fellow lab member, Dr. Joshua Muia. The pET32-WLN1-4 plasmid used as the template DNA to prepare the pET32-WLN1-4 G333R mutant was prepared by Dr. Joshua Muia and Sun Hwa Lee. Since the pET32-WLN1-6 plasmid was engineered with a TEV cleavage site (ENYLFQG), all the plasmids prepared using this plasmid as the
template DNA contained this site as well; these included all of the plasmids discussed above.

These plasmids expressed the WLN constructs as fusion proteins with thioredoxin (TRX) located at the N-terminal of fusion, as shown in Figure 2.8. The fusion proteins also contained a His\textsubscript{6} tag located between the thioredoxin tag and the metal-binding domains.

The pET32-WLN4 plasmid was created by using KOD Hotstart DNA Polymerase to amplify the fourth metal-binding domain in the pET32-WLN1-6 plasmid. The primers were designed to incorporate a TEV protease restriction site into the plasmid sequence. Though the pET-32Xa/LIC vector possesses a Factor Xa cleavage site (IEGR), using Factor Xa to cleave fusion proteins has proven problematic for other members of this lab. Factor Xa is expensive, and in order to get a high level of cleavage, many units of enzyme were required. Additionally, this enzyme apparently produces non-specific cleavage in the fusion proteins, as evidenced by the presence of protein bands with unexpected sizes on SDS-PAGE gels (data not shown). In contrast, TEV protease is a much cheaper protease (it can be produced in-house) that has not shown any evidence of non-specific cleavage. Consequently, the decision was made to incorporate the TEV restriction site into the pET32-WLN4 plasmid and avoid using the Factor Xa cleavage site.

Following PCR, the PCR product was then cloned into the pET-32Xa/LIC vector using ligation independent cloning. Details on the primers, PCR parameters, reagents, and ligation into the pET-32XA/LIC vector can be found in Sections 2.2. and 2.4.1-2.4.2, respectively. After ligation, the plasmid was used transformed into XL-10 Gold cells.
(Agilent Technologies), as discussed in Section 2.4.3. The plasmid was propagated by using cells from single colonies to inoculate 5 mL of liquid LB media that contained 100 μg/mL carbenicillin or ampicillin. Multiple colonies were screened by inoculating multiple 5 mL aliquots of LB, each with cells from a single colony. After 8-10 hours of growth at 37°C in a shaker (250 rpm), the cells were pelleted by centrifugation at 15,000 rpm in a tabletop centrifuge. Subsequently, the plasmids were purified using QIAGEN’s QiaPrep® Spin Miniprep Kit. An aliquot of the plasmids were digested with BglII (20 U) and BamHI (30 U) and run on a 1.0% agarose gel to verify the presence of the expected fragments, as shown in Figure 3.2 below. The size of the pET-32XaLIC vector was 5800 base pairs while the size of the WLN4 insert is 284 base pairs. Quantitation of the plasmid was done by measuring the absorbance at 260 nm. Multiple plasmids were submitted to Retrogen (San Diego, CA) for DNA sequencing. Plasmids with the correct sequence were kept and stored at -20°C.

As with the plasmids constructed above, the pET32-WLN4 plasmid expressed WLN4 as a fusion protein with thioredoxin located at the N-terminal, followed by a His\textsubscript{6} tag. A diagram of the fusion protein is shown in Figure 2.8.
Figure 3.2: Agarose gel (1%) of the pET32-WLN4 plasmid digested with BamHI (30 U) and BglII (20 U). The lane contents were as follows: 1) 100 bp markers (New England Biolabs), 2) pET32-WLN4 digest 3) 1 kb marker (New England Biolabs).

3.3 Protein Expression

All protein expression was done in Rosetta™ 2(DE3) cells. This BL21 derivative was selected because it contained tRNAs for 7 codons that are uncommonly used by *E. coli* but frequently used by eukaryotes (EMD Millipore, Technical Bulletin TB009). The pET32-WLN1-4 and pET32-WLN1-4 G333R plasmids contained all of these codons while the plasmids of the smaller WLN constructs had fewer.

The transformation procedure used for inserting the plasmids into Rosetta™ 2(DE3) cells was discussed in Section 2.4.4.1. Briefly, 5 μL of cells were transformed using 10-20 ng of plasmid DNA. Following a one hour incubation with SOC media in a
37°C shaker, the cells were plated onto an LB-agar plated that contained either 100 μg/mL carbenicillin or ampicillin and 34 μg/mL chloramphenicol. The plates were then incubated in a 37°C oven for 16-18 hours, after which time 50-75 colonies were typically produced.

3.3.1 Expression of Non-labeled Proteins

LB media was used for the production of non-labeled proteins, which was used in all experiments except those involving NMR. The procedure used for growing and expressing the proteins is detailed in Section 2.4.4.2. In summary, a 5 mL starter culture was grown using cells from a single colony on the LB-agar plate. This culture was used to inoculate 50 mL of LB; the 50 mL culture was then used to inoculate 500 mL of LB. After the OD₆₀₀ of the 500 mL culture reached between 0.6-0.8, protein expression was induced by the addition of IPTG (1 mM final concentration). To maintain adequate aeration, only 500 mL of LB was contained in a 2.8 L flask. Total culture volumes of either 1 or 2 liters were grown simultaneously. The cultures were all grown in a shaker maintained at 37°C and 250 rpm. Protein expression was done at either 37°C for 4 hours or room temperature (~22°C) for 16 hours. As the thermostat on the shaker could not reliably sustain temperatures < 25°C, the lid on the shaker was kept open during these inductions to maintain the temperature. The temperature of the lab remained constant during the course of the inductions. Anecdotal evidence collected during this research project suggested that longer induction times at lower temperatures produced more recombinant protein. Intensive studies were not done to confirm this, however. For production of non-labeled proteins, the two temperatures were used interchangeably, depending of the availability of the shaker and time constraints.
Following induction, the cultures were centrifuged at 5,000 rpm for 15 minutes at 4°C to produce a cell pellet. If protein purification did not begin immediately, then the cell pellet was frozen at -20°C. A 1 L culture typically produced ~5-7 g of wet pellet.

3.3.2 Expression of $^{15}$N and $^{13}$C, $^{15}$N Labeled Proteins

The procedure used to express single labeled ($^{15}$N) and double-labeled ($^{13}$C, $^{15}$N) proteins was discussed in detail in Sections 2.4.4.3. The cultures were all grown at 37°C while protein expression occurred at room temperature (~22°C) for 14-16 hours. A 1 L culture of singly/doubly labeled protein typically produced ~5-6 g of wet pellet.

An initial induction study was performed at two temperatures, 22°C and 37°C, in order to determine the optimal temperature for protein expression in minimal media. For this experiment, unlabeled glucose and ammonium chloride was used in the M9 minimal media. WLN1-4 was expressed in the induction studies. Cultures were grown to an OD$_{600}$ between 0.6-0.8 at 37°C and induced with IPTG (1 mM, final concentration). Aliquots (1 mL) were taken at the time points indicated in Figure 3.3. The amount of protein loaded onto the gels was normalized for all samples based on the OD$_{600}$ recorded for each aliquot.
Figure 3.3: Temperature induction study of WLN1-4 grown in minimal media containing unlabeled precursors. The lane contents were as follows: 1) Amersham Full Range Rainbow Markers, 2) 22°C, 0 hr, 3) 22°C, 4.5 hr, 4) 22°C, 5.5 hr, 5) 22°C, 6.5 hr, 6) 22°C, 7.5 hr, 7) 37°C, 4.5 hr, 8) 37°C, 5.5 hr, 9) 37°C, 6.5 hr, 10) 37°C, 7.5 hr, where the time listed is the number of hours after induction that the sample was collected. No 0 hr sample for the 37°C induction was run on the gel due to an insufficient number of lanes; previous induction studies indicated that there was no induction at 0 hr (data not shown). The TRX-WLN1-4 fusion protein (58 kD) is shown within the red oval. A 12.5% acrylamide gel was used; staining was done using 0.1% Coomassie Blue R-250.

As indicated in the gel above, the induction performed at 22°C yielded higher levels of TRX-WLN1-4 than the induction done at 37°C. Consequently, 22°C was selected as the induction temperature for expressing the label proteins.

Further induction tests were run using WLN1-4 in order to determine the optimal induction time for protein expression in minimal media that contained labeled precursors. In order to gauge the effect of the labeled supplements on the induction time, two different media were used: M9 with $^{13}$C-glucose and $^{15}$N ammonium chloride and LB media (used as a positive control, since this relatively nutrient rich media was conducive to optimal protein expression). Culture growth, induction, and sampling were done the same way as described for the temperature study above. Again, the amount of protein
The gels run on the induction test samples are shown below in **Figures 3.4-3.5**.

**Figure 3.4:** Time course induction study at 22°C for the expression of WLN1-4 in minimal media that contained $^{13}$C-glucose and $^{15}$N-ammonium chloride. The lane contents were as follows: 1) Amersham Full Range Rainbow Markers, 2) 0 hr, 3) 10 hr, 4) 14 hr, 5) 21 hr, 6) 24 hr, 7) 37 hr, and 8) 44 hr, where the time listed is the number of hours after induction that the sample was collected. The TRX-WLN1-4 fusion protein (58 kD) is shown within the red rectangle and appeared to run at a slightly higher molecular weight than usual. This anomaly was attributed to a problem with the gel, as the standards also ran oddly. A 12.5% acrylamide gel was used; staining was done using 0.1% Coomassie Blue R-250.
Figure 3.5: Time course induction study at 22°C for the expression of WLN1-4 in LB media. The lane contents were as follows: 1) Amersham Full Range Protein Standards, 2) 0 hr, 3) 4 hr, 4) 8 hr, 5) 20 hr, 6) 24 hr, where the time listed is the number of hours after induction that the sample was collected. The band highlighted with the red oval is TRX-WLN1-4; the mass of this fusion protein is 58 kD. A 12.5% acrylamide gel was used; staining was done using 0.1% Coomassie Blue R-250.

In the M9 media supplemented with labeled precursors, maximal protein expression is reached at 14 hours and appeared to remain constant for up to 44 hours, which suggested that the protein was not being degraded during the extended induction times. Predictably, the LB culture reached maximum protein expression the soonest, at 8 hours, and the protein levels remained constant for up to 24 hours. Based on the results of these studies, induction of the singly/doubly proteins was performed at 22°C for between 14-16 hours.
3.4 Protein Extraction and Purification

3.4.1 Protein Extraction

The same procedure was used to extract the proteins from the cell pellets, regardless of the specific construct or labeling. This procedure, described in Section 2.4.5.1, was performed immediately prior to protein purification.

3.4.2 Protein Purification

The same purification scheme, described in Section 2.4.5.2 was used to purify all constructs, regardless of the specific construct or labeling. The same columns were used with all constructs, with the exception of the one used for gel filtration; WLN1-6 used a Superdex 200 pg column while the other constructs used a Superdex 75 pg column.

Figure 3.6: The purification scheme used to purify the WLN constructs. Shown in parentheses is the sample that was loaded onto the column; this sample came from the preceding purification step.
During the initial purification step, the cell lysate was loaded onto the HisPrep column. Since only proteins containing a His$_6$ tag should have been bound to this column, there was only one peak produced on the chromatogram. The apex of this peak typically occurred at ~25% Buffer B for all constructs. A representative chromatogram is shown in Figure 3.7. All of the fractions in the peak were combined and exchanged into TEV digestion buffer for cleavage of the fusion protein with TEV protease.

![Chromatogram](image)

**Figure 3.7:** A representative chromatogram of the purification of the fusion proteins from the cell lysates using the HisPrep column. The first peak contained proteins that did not bind to the column and were removed during the wash with Buffer A. The second peak was the fusion protein that eluted during the gradient.

A 1:80 TEV protease:fusion protein ratio produced >95% cleavage after 16 hours at room temperature, as estimated from SDS-PAGE gels (data not shown). No promiscuity was observed with this enzyme. The TEV protease recognition site is EQNLYFQG, with the protease cleaving between the residues in bold (8).

As a result, all the WLN constructs that began with the first metal-binding domain contained an amino terminal glycine that was not part of the native protein sequence.
The WLN4 construct also began with glycine, but first amino acid in the metal-binding domain is a glycine (the preceding linker was not included in the construct). Separation of the WLN constructs from the thioredoxin tag and the TEV protease was achieved by using the HisPrep column. Since the WLN constructs lacked a His6 tag, they eluted during the wash with Buffer A. Both the TEV protease and the thioredoxin tag possessed a His6 tag and were bound to the column when the digest was loaded; they were subsequently eluted in a single peak during the gradient. A representative chromatogram is shown in Figure 3.8 below.

![Figure 3.8](image.png)

**Figure 3.8:** A representative chromatogram of the purification of the TEV digest of the fusion proteins on the HisPrep column. The first peak contained the WLN construct, which eluted during the wash with Buffer A. The second peak, containing the thioredoxin tag and the TEV protease, eluted during the gradient.

Gel filtration chromatography was used to further purify the WLN constructs after the initial purification of the digest on the HisPrep column. The column used depended on the construct being purified: a Superdex 200 pg was used for WLN1-6 while a Superdex 75 pg was used for the other constructs. The proteins were eluted isocratically.
from the columns using gel filtration buffer (50 mM Tris, pH 8.0, 150 mM NaCl, 10 mM EDTA, 10 mM DTT). The volume at which each construct eluted was determined by its apparent molecular weight, with the larger constructs being eluted at smaller volumes. Typically, the chromatogram for each construct contained a major peak (the WLN construct) and one or more smaller peaks (impurities). An overlay of the constructs purified using the Superdex 75 pg column is shown in Figure 3.9 below. A chromatogram of WLN1-6 that was purified using the Superdex 200 pg column is shown in Figure 3.10.

Figure 3.9: An overlay of the chromatograms for the constructs purified using the Superdex 75 pg column. The elution volume of each construct (taken from the peak apex) was: 238 mL for WLN4, 216 mL for WLN1-2, 184 mL for WLN1-3, and 171 mL for WLN1-4. The UV absorption at 254 nm was monitored and is shown on the y-axis as milli-Absorbance Units (mAU).
Figure 3.10: A chromatogram of WLN1-6 that was purified using a Superdex 200 pg column. This construct eluted at 73 mL. The noisy baseline was due to lamp problems on the FPLC.

The final stage of purification involved exchanging the constructs into 50 mM phosphate buffer, pH 7-7.5. This was performed using the same Desalting column that was used to exchange the fusion protein into TEV digestion buffer.

All constructs, except WLN4, were quantitated using the BCA assay; that construct was quantitated using the Bio-rad Bradford Assay. WLN4 proved unresponsive to the BCA assay, necessitating the use of an alternate assay.

For cultures grown in LB media, ~12-20 mg of purified WLN construct was obtained per liter of culture. Cultures grown in M9 media supplemented with either $^{15}$N-ammonium chloride and unlabeled glucose or $^{13}$C-glucose and $^{15}$N-ammonium chloride typically yielded ~12-15 mg of purified WLN construct per liter of culture.
Figure 3.11: SDS-PAGE Gel of various WLN constructs. The lane contents were as follows: 1) Amersham Full Range Protein Standards, 2) WLN1-6, 3) WLN1-4, 4) WLN1-3, and 5) WLN1-2. The molecular weights for the WLN constructs were 61.4 kD, 39.7 kD, 29.3 kD, and 16.9 kD, respectively. A 12.5% SDS-PAGE gel was used; staining was done with 0.1% Coomassie R-250.
Figure 3.12: SDS-PAGE Gel of WLN4. The lane contents were as follows: 1) Amersham Full Range Protein Standards, and 2) WLN4. The molecular weight of WLN 4 was 7.5 kD. A 10% NuPAGE Bis-Tris gel from Invitrogen was used; staining was done with 0.1% Coomassie R-250.

3.5 Summary

Four plasmids that contained different combinations of the N-terminal metal-binding domains were prepared: pET32-WLN1-2, pET32-WLN1-3, pET32-WLN1-4 G333R, and pET32-WLN4. The first two plasmids were prepared using site directed mutagenesis to insert a stop codon at the appropriate location in the pET32-WLN1-6 plasmid, resulting in constructs with the appropriate number of metal-binding domains. The third plasmid was also created using the same site directed mutagenesis kit to introduce a guanine to adenine mutation, which resulted in a missense mutation and the conversion of a glycine residue to an arginine residue; pET32-WLN1-4 was used as the template plasmid. The pET32-WLN4 plasmid was prepared by using PCR to amplify the
fourth metal-binding domain from the pET32-WLN1-6 plasmid and inserting it into the pET-32Xa/LIC vector; KOD Hotstart DNA Polymerase was used for the PCR amplification.

All proteins were expressed using Rosetta™ 2(DE3) cells. LB media was used to grow most of the cultures, since non-isotopically labeled protein was required for most analyses. Proteins that were singly (\(^{15}\)N ) or doubly (\(^{13}\)C, \(^{15}\)N) labeled were prepared by adding the appropriately labeled precursors to M9 minimal media.

Induction tests were performed to optimize the time and temperature for production of the labeled proteins. Based on these results, the expression of the singly/doubly labeled proteins was done at 22°C for 14-16 hours while the expression of the triply labeled protein was done for 24 hours at 22°C.

All of the protein constructs, regardless of their labeling, were purified using the same chromatographic steps: HisPrep column, Desalting column, HisPrep column, Superdex 75 pg or Superdex 200 pg, and Desalting column. The proteins were quantitated using either the BCA or Bradford assay and stored at either -20°C or -80°C until needed.
References


CHAPTER 4
THE BIOPHYSICAL CHARACTERIZATION OF THE WLN CONSTRUCTS

This chapter discusses the results of the biophysical analysis of the various WLN constructs. The techniques used to characterize the constructs included high resolution gel filtration, light scattering, circular dichroism, differential scanning calorimetry, and nuclear magnetic resonance spectroscopy.

4.1 High Resolution Gel Filtration

High resolution gel filtration (HRGF) was used to determine the apparent molecular weights of WLN1-2, WLN1-3 and WLN1-4. This technique was also used to investigate a possible interaction between WLN1-3 and WLN4. A Superdex 75 HRGF 10/300 column, which fractionates proteins between 3-70 kD, was used for these analyses. Information on the operational parameters used with the column and the experimental design of the interaction study can be found in Section 2.5.1.

WLN1-2, WLN1-3, and WLN1-4 were analyzed by HRGF to determine their apparent molecular weights. A calibration curve was constructed by injecting a set of globular protein standards onto the column and plotting the log of their molecular weights against their partition coefficients, $K_{av}$ (see Equation 2.1). Separate calibration curves were done for WLN1-2 and WLN1-3/WLN1-4 because the constructs were analyzed at different times, and the same protein standards were not available at both times. Based on the calibration curves, the apparent molecular weights of the constructs were calculated to be: 19,100 Daltons for WLN1-2, 35,500 Daltons for WLN1-3, and 46,800 Daltons for WLN1-4. These results are summarized in Table 4.1 and 4.2; the
calibration curves are shown in Figures 4.1 and 4.2.

Table 4.1: Details of the calibration curve that was used to analyze WLN1-2. The calculated molecular weight derived from the protein sequence is shown in parentheses; the experimentally calculated molecular weight obtained by using the calibration curve is shown in bold font. \( V_e \) was the average elution volume of at least two injections of the protein, and \( K_{av} \) was the partition coefficient that was calculated using Equation 2.1.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Concentration (mg/mL)</th>
<th>MW (Daltons)</th>
<th>Log MW</th>
<th>( V_e ) (mL)</th>
<th>( K_{av} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>3</td>
<td>6,500</td>
<td>3.81</td>
<td>15.39</td>
<td>0.49</td>
</tr>
<tr>
<td>Cytochrome C</td>
<td>2</td>
<td>12,400</td>
<td>4.09</td>
<td>13.81</td>
<td>0.39</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>3</td>
<td>43,000</td>
<td>4.63</td>
<td>10.81</td>
<td>0.19</td>
</tr>
<tr>
<td>Albumin</td>
<td>2.5</td>
<td>66,000</td>
<td>4.82</td>
<td>9.85</td>
<td>0.13</td>
</tr>
<tr>
<td>Blue Dextran</td>
<td>1</td>
<td>2,000,000</td>
<td>n/a</td>
<td>7.73</td>
<td>n/a</td>
</tr>
<tr>
<td>WLN1-2</td>
<td>2.2</td>
<td>19,100</td>
<td>n/a</td>
<td>12.84</td>
<td>0.32</td>
</tr>
<tr>
<td>(16,958 Daltons)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 4.1: Calibration curve for the data presented in Table 4.1.
Table 4.2: Details of the calibration curve that was used to analyze WLN1-3 and WLN1-4. The calculated molecular weight derived from the protein sequence is shown in parentheses; the experimentally calculated molecular weight obtained by using the calibration curve is shown in bold. $V_e$ was the average elution volume of at least two injections of the protein, and $K_{av}$ is the partition coefficient.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Concentration (mg/mL)</th>
<th>MW (Daltons)</th>
<th>Log MW</th>
<th>$V_e$ (mL)</th>
<th>$K_{av}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>3.0</td>
<td>6,500</td>
<td>3.81</td>
<td>15.04</td>
<td>0.47</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>3.0</td>
<td>13,700</td>
<td>4.14</td>
<td>13.19</td>
<td>0.35</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>3.0</td>
<td>29,000</td>
<td>4.46</td>
<td>11.63</td>
<td>0.25</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>4.0</td>
<td>44,000</td>
<td>4.64</td>
<td>10.52</td>
<td>0.18</td>
</tr>
<tr>
<td>BSA</td>
<td>2.5</td>
<td>66,000</td>
<td>4.82</td>
<td>9.51</td>
<td>0.12</td>
</tr>
<tr>
<td>Blue Dextran</td>
<td>0.5</td>
<td>2,000,000</td>
<td>n/a</td>
<td>7.68</td>
<td>n/a</td>
</tr>
<tr>
<td>WLN1-3 (29,382 Daltons)</td>
<td>3.2</td>
<td>35,500</td>
<td>n/a</td>
<td>11.00</td>
<td>0.21</td>
</tr>
<tr>
<td>WLN1-4 (39,783 Daltons)</td>
<td>3.2</td>
<td>46,800</td>
<td>n/a</td>
<td>10.36</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Figure 4.2: Calibration curve for the data presented in Table 4.2.
HRGF was used to characterize WLN1-6. Using a Superdex 200 HRGF 10/300 column, a lab colleague, Joshua Muia, determined that the apparent molecular weight of this construct was ~134 kD; the actual size of the construct was 61.4 kD (1). The substantial difference between the actual and apparent molecular weight were attributed to WLN1-6 having a larger hydrodynamic radius ($R_H$) than a globular protein with the same molecular weight would have (see Section 4.2 for information on the hydrodynamic radii of the constructs). As a consequence of its increased hydrodynamic radius, WLN1-6 adopted a more ellipsoid than spherical shape and migrated through the gel filtration column as a protein with a higher molecular weight would (2).

HRGF data illustrated that the addition of metal-binding domains to the Wilson disease constructs resulted in decreased elution volumes, which indicated that the constructs had higher apparent molecular weights and larger hydrodynamic radii than what would be expected for globular proteins of similar molecular weights. This trend became more pronounced as more domains were added. It is important to keep in mind that HRGF reports only on the average state of the protein ensemble and not on the conformational state of individual protein molecules. Therefore, as more metal-binding domains were added, the average state of the protein increased.

WLN4 and WLN1-4 G333R were not analyzed by HRGF because the Superdex 75 HRGF 10/300 was unavailable for use after these constructs had been created and purified.

HRGF was also used to investigate a potential interaction between WLN1-3 and WLN4, as described in Section 2.5.1. Briefly, equimolar amounts of WLN1-3 and WLN4 were incubated at room temperature for 24 hours. An aliquot was injected onto
the Superdex 75 HRGF column after 1, 5, and 24 hours. If an interaction were occurring, it would be expected that the new species would elute near 9.2 mL, the elution volume for WLN1-4.

The chromatograms in Figure 4.3 do not show the presence of a species other than WLN1-3 and WLN4. After 24 hours, a peak at 6.9 mL appeared. This peak eluted below the void volume (7.7 mL) of the column and was attributed to aggregation of WLN1-3, as the WLN1-3 peak at 10.1 mL decreased in absorbance. The elution profiles of WLN1-3, WLN1-4, and WLN4 are shown for reference in Figure 4.4. The lack of an observed interaction between WLN1-3 and WLN4 could have several causes. WLN1-3 and WLN4 may simply not interact, or interaction may require the linker region that is between metal-binding domains 3 and 4, which neither the WLN1-3 nor the WLN4 construct possessed.
Figure 4.3: HRGF interaction study between WLN1-3 and WLN4. Equal volumes of the sample containing WLN1-3 and WLN4 were injected at the indicated times. As the chromatograms show, there is no evidence of a species with an increased apparent molecular weight at any of the time points. The peak at 6.9 mL in the 24 hour sample elutes before the void volume of the column and is not considered evidence of an interaction. That peak is likely due to the aggregation of WLN1-3, since its peak area and height have decreased at 24 hours. Absorbance was monitored at 254 nm.
Figure 4.4: Elution profiles of WLN1-3, WLN1-4, and WLN4 on the Superdex 75 HRGF column. Equimolar amounts of all proteins were injected. Absorbance was monitored at 254 nm.

4.2 Light Scattering Analysis

WLN1-2, WLN1-3, and WLN1-4 were characterized by static (SLS) and dynamic (DLS) light scattering at the Keck Biotechnology Resource Laboratory at Yale University. The constructs were injected onto a high resolution gel filtration column that was connected to an OPTILAB rEX Refractive Index Detector, a DAWN-HELEOS Light Scattering Detector, and a Waters 996 Photodiode Array Detector. ASTRA software (version 5.4.3.10) was used to calculate the weight-average molar mass and the hydrodynamic radius ($R_h$) of each construct. The Debye model was used for determining the weight-average molar masses, and a $dn/dc$ value of 0.175 mL/g was used in this calculation, as this value was adequate for the analysis of the standard proteins analyzed with the WLN constructs (3, 4). Figure 4.5 shows the molar mass distribution of WLN1-2; the trace indicated that there was only one apparent form of WLN1-2 present, i.e. the
sample was monodisperse. Figure 4.6 shows the molar mass distribution for WLN1-3 and WLN1-4. While the trace for WLN1-4 indicated the presence of a monodisperse sample, the trace for WLN1-3 indicated the presence of other proteins. Subsequent analysis of another aliquot of WLN1-3 by SDS-PAGE revealed that the protein had partially degraded (data not shown). These degraded proteins appeared as a trailing edge at the back of the peak in the trace, as shown in Figure 4.6; this portion of the peak was not used for calculating the molecular weight and $R_{h}$.

The autocorrelation curves for WLN1-2, WLN1-3, and WLN1-4 are shown in Figures 4.7, 4.8, and 4.9, respectively. Once the translational diffusion coefficient, D, was determined, the hydrodynamic radius of each construct was calculated by solving the Stokes-Einstein equation. The hydrodynamic radius is the radius that a hard sphere with the same translational diffusion coefficient as the protein would have and indicates the radius of the hydrated protein.

The molecular weights and hydrodynamic radii of the constructs are summarized in Table 4.3.
Figure 4.5: Molar mass distribution of WLN1-2 (red trace) injected onto a Superdex 75 HRGF 10/300 column. Trypsin inhibitor (pink trace) and bovine serum albumin (blue trace) were run for comparative purposes; these globular proteins have molar masses of 20.0 kD and 66.4 kD, respectively. The solid lines denote the UV absorbance at 280 nm while the dots denote the weight-average molar mass for each 1 second slice.
Figure 4.6: Molar mass distribution of WLN1-3 (red trace) and WLN1-4 (green trace) that were injected onto a Superdex 200 HRGF 10/300 column. The asymmetry of the WLN1-3 peak indicated the presence of other association states and/or other proteins. Subsequent analysis by SDS-PAGE revealed that the construct had partially degraded (data not shown). Bovine serum albumin (blue trace), ovalbumin (brown trace) and carbonic anhydrase (aqua trace) were run for comparative purposes; these globular proteins have molar masses of 66.4 kD, 42.8 kD, and 29.0 kD, respectively. The solid lines denote the UV absorbance at 280 nm while the dots denote the weight-average molar mass for each 1 second slice.
Figure 4.7: The autocorrelation function for WLN1-2. The translational diffusion coefficient (D) was calculated as $1.13 \pm 0.03 \times 10^{-6}$ cm$^2$/sec by ASTRA software.

Figure 4.8: The autocorrelation function for WLN1-3. The translational diffusion coefficient (D) was calculated as $9.61 \pm 0.53 \times 10^{-7}$ cm$^2$/sec by ASTRA software.
Figure 4.9: The autocorrelation function for WLN1-4. The translational diffusion coefficient (D) was calculated as $7.11 \pm 0.24 \times 10^{-7}$ cm$^2$/sec by ASTRA software using.

Table 4.3: Summary of the biophysical parameters of the constructs that were determined by light scattering. The actual MW refers to the MW that was calculated from the protein sequence. The $R_H$ and ASTRA calculated MW were calculated from the apex of one peak. *WLN1-6 eluted as a polydisperse sample, i.e. multiple peaks were present. **Data is from the unpublished work of a lab colleague, Ibtesam Alja’afreh.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Actual MW (kD)</th>
<th>ASTRA Calculated MW (kD)</th>
<th>$R_H$ (nm)</th>
<th>Frictional Ratio</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN1-2</td>
<td>16.9</td>
<td>17.0</td>
<td>2.16 ± 0.06</td>
<td>1.24</td>
<td>This work</td>
</tr>
<tr>
<td>WLN1-3</td>
<td>29.3</td>
<td>29.8</td>
<td>2.54 ± 0.14</td>
<td>1.24</td>
<td>This work</td>
</tr>
<tr>
<td>WLN1-4</td>
<td>39.7</td>
<td>40.9</td>
<td>3.43 ± 0.11</td>
<td>1.50</td>
<td>This work</td>
</tr>
<tr>
<td>WLN4</td>
<td>7.6</td>
<td>7.6</td>
<td>1.33 ± 0.07</td>
<td>1.02</td>
<td>(5)</td>
</tr>
<tr>
<td>WLN1-6</td>
<td>61.4</td>
<td>66*</td>
<td>4.09 ± 0.09</td>
<td>1.72</td>
<td>(1)</td>
</tr>
<tr>
<td>WLN3-6</td>
<td>61.4</td>
<td>39.1</td>
<td>3.67 ± 0.12</td>
<td>1.63</td>
<td>**</td>
</tr>
</tbody>
</table>
The results summarized in Table 4.3 indicated that there was no difference between the ASTRA calculated molecular weight and the sequence predicted molecular weight for WLN1-2, WLN1-3, and WLN1-4. WLN4 was prepared and sent for light scattering analysis by a lab colleague, Dr. Wilson Okumu (5). The results, also shown in Table 4.3, indicate that the ASTRA calculated molecular weight and the actual molecular weight are the same. In contrast to the monodisperse nature of the previous constructs, WLN1-6 exhibited polydispersity. The ASTRA calculated molecular weight (of the apex of the main peak) for WLN1-6 (prepared and submitted for analysis by a lab colleague, Dr. Joshua Muia) was similar to the actual molecular weight (1). The hydrodynamic radius of WLN1-6 was calculated to 4.09 ± 0.09 nm, which is greater than the 3.4 nm hydrodynamic radius of bovine serum albumin, a globular protein with nearly the same molecular weight (1).

The frictional ratio for each construct was calculated using the Dynamics 7 software supplied with the light scattering instrument. A frictional ratio of 1.0 denotes a perfect sphere; the higher the ratio, the less spherical the protein. The cutoff of the frictional ratio that describes spherical vs non-spherical proteins is not precisely defined, but it is usually regarded to be 1.5. Based on this cut-off, the shape of WLN1-6 was determined to be non-spherical while WLN1-4 was deemed spherical, though it is at the border of the designation and displays more of a non-spherical shape than WLN1-3. As the frictional ratios in Table 4.3 indicate, the more domains that were added, the less spherical the construct became. For comparative purposes, the frictional coefficient for the purely fibrous protein fibrinogen is 2.3 (6).
4.3 Circular Dichroism

Stability of the WLN constructs was assessed by using circular dichroism to monitor denaturation as a function of guanidine hydrochloride and temperature. Analysis of the data revealed that the constructs have different degrees of stability despite having the same ferredoxin fold (βαββαβ).

4.3.1 Guanidine Hydrochloride Denaturation

Guanidine hydrochloride was used to denature the protein constructs (apo forms only) and assess their stability as described in Section 2.5.4. Briefly, a series of samples (20-40, more samples were prepared for constructs that unfolded via a three-state pathway) were prepared that contained a fixed protein concentration and a varying concentration of guanidine hydrochloride. After equilibrating the samples at room temperature, a scanning wavelength spectra (200-250 nm) was collected. The CD signal in mean residue ellipticity (θ_{MRW}) at 222 nm was plotted as a function of guanidine hydrochloride concentration; these curves are shown in Figures 4.10-4.11. The resulting sigmoidal or double sigmoidal curves were fitted to a two- or three-state model, respectively, with IGOR Pro 6; these results are shown in the Table 4.4 (7).

WLN1-2 and WLN1-3 unfolded via an apparent two-state mechanism, and the curves appeared very similar to each with a single, well-defined transition occurring between ~1.5-3.5 M guanidine hydrochloride. In WLN1-2 there were larger fluctuations in the CD signal between 0-2 M guanidine hydrochloride than there were for WLN1-3. The C_m (the concentration of guanidine hydrochloride at which the protein is 50% denatured) for both constructs was calculated to be 2.47 ± 0.03 M and 2.48 ± 0.03 M guanidine hydrochloride for WLN1-2 and WLN1-3, respectively. The free energy of
unfolding for WLN1-2 and WLN1-3 were 16.4 ± 1.4 kJ/mol and 15.2 ± 1.1 kJ/mol, respectively. The presence of metal-binding domain 3 in WLN1-3 did not significantly affect stability, as its free energy of unfolding and $C_m$ are very similar to WLN1-2.
Figure 4.10 A-D: Guanidine hydrochloride denaturation of WLN1-2, WLN1-3, WLN1-4, and WLN4, respectively. The CD signal at 222 nm is plotted.
WLN4 also unfolded via an apparent two-state model and exhibited a single unfolding transition that occurred between 5-7 M guanidine hydrochloride, with the $C_m$ calculated to be $5.71 \pm 0.02$ M guanidine hydrochloride. Prior to $\sim 5$ M guanidine hydrochloride, there were only slight fluctuations in the curve, which emphasized the extreme stability of the construct. The free energy of unfolding was calculated as $42.9 \pm 2.8$ kJ/mol.

In contrast to the constructs discussed so far, WLN1-4 had two distinct unfolding transitions and therefore unfolded via an apparent 3-state model. The first transition occurred at $2.32 \pm 0.02$ M guanidine hydrochloride, which is close to the $C_m$ value of $2.48 \pm 0.03$ M guanidine hydrochloride that was obtained for WLN1-3. The second transition of WLN1-4 occurred at $5.91 \pm 0.1$ M guanidine hydrochloride, which is similar to the $C_m$ value of $5.71 \pm 0.02$ M guanidine hydrochloride for WLN4. The free energy of the unfolding for the first transition was $16.5 \pm 1.0$ kJ/mol and $38.8 \pm 7.2$ kJ/mol for
the second transition, which are similar to the values for WLN1-3 and WLN4, respectively. Given similarity of the results of WLN-1-4 to those obtained for WLN1-3 and WLN4, WLN1-4 appears to unfold as a result of two independent events – the unfolding of the first three metal-binding domains (i.e. WLN1-3) and the fourth metal-binding domain (i.e. WLN4).

WLN1-4 G333R, which possessed a mutation in the linker region between the third and fourth metal-binding domains, unfolded similarly to WLN1-4. An apparent 3-state unfolding was observed, with the $C_m$ being $2.35 \pm 0.02$ M guanidine hydrochloride and $5.72 \pm 0.1$ M guanidine hydrochloride for the first and second transitions, respectively. The free energies of unfolding were $16.0 \pm 1.0$ kJ/mol and $34.3 \pm 6.7$ kJ/mol for the first and second transitions, respectively. Since the values for the mutant are quite similar to the values of WLN1-4, this suggested that the mutation does not greatly impact the stability. This is consistent with what Banci et al. reported for the NMR solution structure of WLN3-4 - the disordered linker region between the two metal-binding domains allowed free rotation of the domains (8). A mutation in a disordered region would not be expected to have a high impact on well-structured regions of a protein.
Table 4.4: The denaturation of the Wilson disease protein constructs by guanidine hydrochloride. $\Delta G_{\text{H}_2\text{O}}$ is the free energy of unfolding in the absence of guanidine hydrochloride. $m$ describes the amount of surface area of the protein that is exposed following denaturation. $C_m$ is the concentration of guanidine hydrochloride in which the protein is 50% denatured. For constructs that followed a three-state unfolding pathway, the data in the “Second Transition” columns is for the intermediate to denatured portion of the unfolding (I $\rightarrow$ D) while the data in the previous three columns is for the protein as it moved from the native to intermediate state (N $\rightarrow$ I). In constructs without a second transition, no intermediate accumulated to appreciable levels and the protein went directly from its native structure to a denatured form (N $\rightarrow$ D). One curve (containing 20-40 points) was analyzed for each construct and the standard deviation from the curve fitting is shown.

<table>
<thead>
<tr>
<th>Construct</th>
<th>$\Delta G_{\text{H}_2\text{O}}$ (kJ/mol)</th>
<th>$m_1$ (kJ/mol $\times$ M)</th>
<th>$C_m$ (M)</th>
<th>$\Delta G_{\text{H}_2\text{O}}$ (kJ/mol)</th>
<th>$m_1$ (kJ/mol $\times$ M)</th>
<th>$C_m$ (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN1-2</td>
<td>16.4 ± 1.4</td>
<td>6.6 ± 0.5</td>
<td>2.47 ± 0.03</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>WLN1-3</td>
<td>15.2 ± 1.1</td>
<td>6.1 ± 0.4</td>
<td>2.48 ± 0.03</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>WLN1-4</td>
<td>16.5 ± 1.0</td>
<td>7.1 ± 0.4</td>
<td>2.32 ± 0.02</td>
<td>38.8 ± 7.2</td>
<td>6.5 ± 1.3</td>
<td>5.91 ± 0.08</td>
</tr>
<tr>
<td>WLN4</td>
<td>42.9 ± 2.8</td>
<td>7.5 ± 0.5</td>
<td>5.71 ± 0.02</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>WLN1-4 G333R</td>
<td>16.0 ± 1.0</td>
<td>6.8 ± 0.4</td>
<td>2.35 ± 0.02</td>
<td>34.3 ± 6.7</td>
<td>6.0 ± 1.2</td>
<td>5.72 ± 0.10</td>
</tr>
</tbody>
</table>
The free energy of unfolding for most proteins is between 20-60 kJ/mol, which is on the order of a few hydrogen bonds (9). The multi-domain constructs of WLN1-2 and WLN1-3 were slightly lower than that. This somewhat low stability may be a consequence of the fact that these constructs are fragments of a protein and are not naturally expressed on their own. Adding the fourth metal-binding domain increased the complexity of the unfolding curve, but the first unfolding transition had a free energy of unfolding and a $C_m$ that were consistent with WLN1-3, which suggested that WLN1-2 and WLN1-3 may simply be slightly less stable intrinsically.

In contrast to WLN1-2 and WLN1-3, WLN4 displayed a much higher free energy of unfolding and $C_m$. Indeed, the $C_m$ of WLN4 is very close to the concentration of guanidine hydrochloride that is typically used to denature proteins that have been incorporated into inclusion bodies (10, 11). Given that each metal-binding domain possesses the same ferredoxin fold ($\beta\alpha\beta\alpha\beta$), it is surprising that the stability of the fourth metal-binding domain is so much greater than multi-domain constructs.

### 4.3.2 Thermal Denaturation

The thermal denaturation of apo- and some copper-loaded forms of the constructs were performed as discussed in Section 2.5.4. Briefly, 3-5 uM protein in 5 mM phosphate buffer (pH 7.5) was placed in a 1 cm quartz cuvette that was inserted into a Peltier device that was connected to a CD instrument. A thermal gradient of 2°C/min was used to heat the sample from 25°C to 97°C, and the CD signal was monitored at 222 nm; these curves are shown in Figure 4.12. The melting temperature, $T_m$, was calculated using the software supplied with the instrument; each construct was analyzed three times. The results are shown in Table 4.5 below.
Figure 4.12: Thermal unfolding curves for the WLN constructs. A representative curve for each construct is shown.
Table 4.5: Thermal unfolding of the apo-WLN constructs. The $T_m$ shown is the average of three trials, ± the standard deviation.

<table>
<thead>
<tr>
<th>Construct</th>
<th>$T_m$ (°C)</th>
<th>Reversibility of Unfolding</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN1-2</td>
<td>59.5 ± 1.7</td>
<td>No</td>
<td>Protein precipitated</td>
</tr>
<tr>
<td>WLN1-3</td>
<td>58.3 ± 0.6</td>
<td>No</td>
<td>Protein precipitated</td>
</tr>
<tr>
<td>WLN4</td>
<td>78.3 ± 0.2</td>
<td>Yes</td>
<td>Refolding curve very similar to the unfolding curve</td>
</tr>
<tr>
<td>WLN1-4</td>
<td>No unfolding transition</td>
<td>N/A</td>
<td>Protein did not unfold – the minimal wavelength of absorption changed and the protein aggregated</td>
</tr>
<tr>
<td>WLN1-4 G333R</td>
<td>No unfolding transition</td>
<td>N/A</td>
<td>Protein did not unfold – the minimal wavelength of absorption changed and the protein aggregated</td>
</tr>
</tbody>
</table>

WLN1-2 and WLN1-3 displayed similar melting temperatures and both irreversibly precipitated after being heated. WLN4 demonstrated high thermal stability and its melting temperature was 20°C higher than that of the three domain construct WLN1-3. Additionally, WLN4 did not precipitate after the thermal gradient and was able to mostly refold as the reverse gradient was applied, as shown in Figures 4.13 and 4.14.
Figure 4.13: The thermal unfolding of WLN4 (forward) and the renaturation of the protein as it was cooled (reverse).

Figure 4.14: CD spectra of WLN4 that was collected prior to the application of the thermal gradient (green line) and after the application of the reverse gradient, i.e. cooling it (orange line); both spectra were collected at 25°C. After being cooled, WLN4 was able to regain 94.4% of the CD signal at 222 nm that it had at the beginning.
Neither WLN1-4 nor its G333R mutant displayed an apparent unfolding transition. A fluctuation in the CD signal is observed at ~58°C for both proteins, which suggested a structural or association state perturbation. A scanning wavelength spectra collected before and after the thermal gradient indicated that structural rearrangement had taken place, as shown in Figure 4.15 below; though only the WLN1-4 curve is shown, WLN1-4 G333R displayed the same wavelength shift. Prior to being heated, WLN1-4 had a MRE minimum at ~222 nm, a wavelength at which α-helices absorb. After the gradient, however, the MRE at 222 nm increased while the MRE at ~209 nm, a wavelength at which β-sheets absorb, decreased. Since the thermal denaturation was monitored at 222 nm, this may explain the slight fluctuation in CD signal that occurred at around 58°C - this may have been the temperature at which the structural transition occurred. After this temperature, the CD signal for WLN1-4 decreases slightly, which is consistent with the loss of some of the α-helices. The CD signal for the mutant, WLN1-4 G333R, decreased slightly (data not shown).

To verify that the lack of an unfolding transition was in fact reflective of what would occur in the intact amino terminal of the Wilson disease protein, a construct comprised of all six metal-binding domains (WLN1-6) was expressed and purified. When WLN1-6 was heated, there was no apparent melting temperature (see Figure A1 in the Appendix). Additionally, the scanning wavelength spectra that were taken of the protein before and after heating were the same as those observed for WLN1-4 and its mutant (data not shown). Therefore, the thermal behavior of WLN1-4 was representative of the intact amino terminus of the Wilson disease protein and was not an aberration that was introduced by truncating this region.
Figure 4.15: Scanning wavelength spectra of WLN1-4 were collected before and after the application of the thermal gradient; both spectra were collected at 25°C. The above spectra show that the most negative ellipticity shifted from ~222 nm to ~211 nm following exposure to heat. The mutant WLN1-4 G333R displayed the same wavelength shift (data not shown).
Figure 4.16: A chromatogram of WLN1-4 before and after the application of the thermal gradient. The proteins were injected onto Superdex 200 HRGF column and eluted using 50 mM sodium phosphate, pH 7.5 / 150 mM sodium chloride at a flow rate of 0.5 mL/min. In the “Before” sample, WLN1-4 eluted at 19.1 mL while in the “After” sample it eluted at 10.8 mL.

Another possibility for the lack of an apparent melting point for WLN1-4 and its mutant is that the proteins are forming soluble aggregates after being heated, and that it was these aggregates that were resistant to thermal unfolding. This hypothesis was investigated by analyzing WLN1-4 on a Superdex 200 HRGF column after the protein had undergone the thermal gradient. As shown in Figure 4.16 above, there was a marked difference in the elution volume after the protein was heated relative to the elution volume of the protein prior to being heated. After being heated, WLN1-4 eluted near the void volume of the column (10.6 mL), which indicated that aggregation had occurred, though no precipitation was visible to the eye. This is known as soluble aggregation.

WLN1-4 does not have an apparent melting temperature but it did undergo an apparent structural rearrangement as it was heated. In an attempt to characterize the
nature of the forces that were contributing to the stability of WLN1-4, sodium chloride was used to see if the protein could be thermally unfolded. As shown in Figure 4.17, at concentrations of 75 mM or higher, WLN1-4 did show an apparent melting temperature. The $T_m$ could not be calculated, however, because the CD signal had not leveled off after the protein unfolded and before data collection ended at 98°C.

To verify that an actual unfolding transition had occurred and the decrease in the CD signal was not due to soluble aggregation, a scanning wavelength spectra was collected before and after the gradient. As Figure 4.18 shows, the shape of the curve after WLN1-4 was heated was the same as it was prior to being heated, though the MRE increased as secondary structure was lost upon heating. The wavelength minima (222 nm) were the same in both curves, unlike when WLN1-4 was unfolded in the absence of sodium chloride and the wavelength minimum shifted from 222 nm to 211 nm.
Figure 4.17: The unfolding of WLN1-4 in the presence of various concentrations of sodium chloride. The same conditions that were used in the previous thermal unfolding experiments (i.e. thermal gradient, protein and buffer concentration) were used here. The CD signal was monitored at 222 nm. No visible precipitation was observed after the gradient.

Figure 4.18: WLN1-4 with 150 mM sodium chloride scanning wavelength spectra before and after thermal denaturation.
The ability of WLN1-4 to thermally unfold in the presence of low concentrations of sodium chloride suggests that electrostatic contacts may play a role in stabilizing WLN1-4. In this construct, 9.6% of the residues are acidic (aspartic acid and glutamic acid) and 10.1% are basic (arginine, histidine, and lysine). It is possible that a salt bridge(s) could be forming between an acidic and basic residue, thereby conferring structural stability. Since neither an X-ray nor a NMR structure is yet known for WLN1-4 or WLN1-6, it is not possible to determine what residues in particular may be interacting. A study by Nilsson et al. showed that the thermal stability of WLN5-6 was decreased by 30ºC when NaCl was added, which demonstrated that electrostatic contacts are important in that construct (12). Since the structure of all of the metal-binding domains is conserved, there may also be electrostatic contacts present in WLN1-4.

Thermally denaturing copper loaded forms of the constructs proved to be challenging, and it was not possible to obtain consistent melting temperatures. Whether a single, copper-loaded protein sample was split into two and run successively on the CD, or two different samples were processed in parallel, or the protein was metallated immediately prior to unfolding, the melting temperatures varied. Additionally, WLN4 and WLN1-4 were the only constructs that could be consistently metallated; WLN1-2 and WLN1-3 tended to precipitate upon the addition of copper.

The method used to titrate copper into the proteins differed from the procedure used to anaerobically metallate the proteins that was outlined in Section 2.4.5.3. The procedure was modified as discussed below.

An aliquot of WLN1-4 was reduced using TCEP (1 mM) and placed in a 1 cm cuvette. Sub-stoichiometric equivalents of copper were titrated in, and a scanning
wavelength spectra was collected on the CD; see Figure 4.19. After 3.5 equivalents of copper were added, the thermal gradient was applied. As with the apo-form of WLN1-4, copper loaded WLN1-4 did not exhibit an apparent melting point; several runs confirmed this. Additionally, the same spectroscopic shift that apo-WLN1-4 exhibited after being heated was also present in copper loaded WLN1-4 (Figure 4.19).

WLN4 was titrated with copper in the same manner as WLN1-4 as discussed above. Copper binding led to an increase in MRE, as shown below in Figure 4.20. Thermal unfolding of copper-loaded protein was performed several times, but the results were always inconsistent. This inconsistency may be due to metal induced aggregation of WLN4.

![Figure 4.19](image)

**Figure 4.19:** The effect of copper(I) loading on the CD spectra of WLN1-4. The data was collected at 25°C using a 1 cm cuvette. The addition of copper results in a less negative CD signal. This is due to the absorption of the copper-thiolate bonds that absorb in this region and does not mean that the secondary structure of WLN1-4 is being decreased. Upon heating, Cu(I)-WLN1-4 exhibited that same shift in minima that apo-WLN1-4 did, indicating that soluble aggregation occurred.
Figure 4.20: Titration of WLN4 with copper. Data was collected in at 25°C using a 1 cm cuvette. The addition of copper results in a less negative CD signal. This is due to the absorption of the copper-thiolate bonds that absorb in this region and does not mean that the secondary structure of WLN4 is being decreased. Consistent unfolding temperatures for Cu(I)-WLN4 could not be obtained.

The addition of copper to WLN1-4 and WLN4 resulted in an increase of the CD signal, as indicated in the above figures. This change in CD signal is not due to a decrease in secondary structure, however, as NMR experiments performed on constructs of several the metal-binding domains have revealed. In the literature, the addition of copper to WLN3-4 resulted in the shifting of resonances associated with several residues, which are predominately within the loops of the metal-binding domains, but the $\alpha$-helices and $\beta$-sheets were unaffected (13). Additionally, the titration of copper into WLN5-6 revealed that the $\alpha$-helices and $\beta$-sheets were not perturbed, though there were changes in the copper binding loops (14).

The increase in CD signal observed in Figures 4.19-4.20 is likely due to the asymmetric nature of the copper binding site once copper is bound. The copper binding
site is chiral and able to absorb circularly polarized light. Since the CD signal that is reported by the spectrometer is defined as the difference between the amount of left-handed circularly polarized light minus the amount of right-handed circularly polarized light, the increase in CD signal must be due to the copper-binding site absorbing more right-handed circularly polarized light.

The increase in CD signal upon addition of copper that was observed for WLN1-4 and WLN4 is in contrast to what DiDonato et al. reported for the titration of a construct comprised of all six metal-binding domains (15). In this study, the researchers found that the addition of copper led to a more negative CD signal at wavelengths associated with secondary structure motifs (i.e. 208 and 222 nm), and the authors interpreted this to mean that the construct was becoming increasingly structured upon binding copper. The aforementioned NMR experiments (done after the study by DiDonato) indicated that the domains did not undergo significant structural rearrangement upon copper binding (13,14).

The most likely explanation for the discrepancy is that the absorption of CD light by the copper-thiolate bonds in MBDs were not taken into consideration. As NMR experiments have convincingly demonstrated that copper binding does not introduce large scale conformational changes, the changes in CD signal that were observed by DiDonato et al. must have another explanation. Another possible explanation is that the protein DiDonato et al. used for their study was primarily expressed in inclusion bodies and had to be denatured and refolded prior to use. Perhaps after the protein was refolded it did not exactly adopt the native conformation. This may have caused minor
differences in the metal-binding sites that led to observable changes in the CD spectra upon the addition of copper.

4.4 Differential Scanning Calorimetry

Differential scanning calorimetry (DSC) was performed on apo- and copper-loaded WLN4 and apo-WLN1-4 in order to determine the melting points ($T_m$) of these constructs. The analysis utilized a MicroCal VP-DSC (GE Healthcare) instrument and occurred at the Biophysics Core at the University of Colorado-Denver Anschutz Medical Campus. These constructs were analyzed by DSC because of difficulties that were encountered when the thermal denaturation was monitored by CD.

When the thermal denaturation of copper-loaded WLN4 was analyzed by CD, the melting temperature varied. Two samples that came from the same metallated sample were run on successive days under the same experimental conditions (a thermal gradient of 25-95°C, with a 2°C/min ramp rate was used). For one experiment, the melting point was calculated as 74.5 ± 0.99°C; for the other experiment, it was calculated as 66.7 ± 0.4°C. The reason for this discrepancy was not conclusively determined. A melting point of 74.5°C was consistent with the melting point of non-copper loaded WLN4 of ~78°C, which was reproducible. Consequently, DSC was used in an attempt to resolve the ambiguity.

In the case of WLN1-4, this construct did not appear to thermally unfold at temperatures up to 95°C. Since an aqueous buffer (phosphate) was used as a diluent and the cuvette did not have a screw cap lid, higher temperatures could not be used without boiling the sample and evaporating the water. This obstacle was alleviated by DSC
because the sample cells were pressurized and the particular DSC instrument used here could reach temperatures of up to 110°C.

An aliquot of WLN4 was loaded with copper according to the procedure outlined in Section 2.4.5.3. ICP-AES was used to determine the copper content of all samples prior to their submission for analysis. The copper content of the metallated WLN4 was 77.4% while the copper content of the non-metallated WLN4 was 1.4%. When analyzed by DSC, the non-copper loaded WLN4 exhibited an endothermic peak in the thermogram at 84.5°C and an exothermic peak at 107°C. The endothermic peak was expected and the temperature at the apex of this peak represented the $T_m$ of the protein; the exothermic peak was attributed to protein aggregation; as no aggregation was visibly apparent when the cuvette was examined post-run, the aggregation must have been soluble. In contrast, no unfolding transition was observed for the copper-loaded WLN4, as neither an endothermic nor an exothermic peak was produced. No precipitation was visible after the experiment, but soluble aggregates may have formed; if this were the case, then these aggregates may have been resistant to thermal denaturation. The thermograms of non-copper and copper-loaded WLN4 are shown in Figures 4.21 and 4.22, respectively.

The melting temperature for apo-WLN1-4 could not be determined. As the endotherms in Figure 4.23 show, there was concentration dependent aggregation that occurred at protein concentrations $\geq 0.5$ mg/mL. One possible source of this aggregation was the WLN1-4 was being partially denatured and the denatured proteins were aggregating; if this were the case, then higher protein concentrations would produce transitions at lower temperatures since more protein would be available to be denatured; this was what was observed for apo-WLN1-4 (GE Healthcare, Differential Scanning
Calorimetry: Theory and Practice, Technical Bulletin 28-9870-36 AA). Lower concentrations of apo-WLN1-4 (0.1 and 0.5 mg/mL) did not generate sufficient signal for a transition to be observed. Consequently, a melting temperature could not be determined for apo-WLN1-4.

**Figure 4.21:** The thermogram showing the thermal denaturation of apoWLN4; the temperature in °C is shown on the x-axis, and the heat capacity (cal/°C) is shown on the y-axis. WLN4 (1 mg/mL) was loaded in the sample cell, and 50 mM sodium phosphate, pH 7.5 was loaded into the reference cell. A thermal gradient from 20-110°C with a 1°C/min ramp rate was applied. The buffer thermogram was not subtracted from the WLN4 thermogram because to the significant baseline difference between the two baselines (buffer baseline not shown here). Consequently, the enthalpy of denaturation could not be calculated. The significant difference between the two baselines was likely due to incomplete buffer exchange of WLN4. No protein precipitation was observed following the thermal gradient. Only one run of DSC was performed.
Figure 4.22: The thermogram showing the thermal denaturation of copper-loaded WLN4; the temperature in °C is shown on the x-axis, and the heat capacity (cal/°C) is shown on the y-axis. Copper-loaded WLN4 (1 mg/mL) was loaded in the sample cell, and 50 mM sodium phosphate, pH 7.5 was loaded into the reference cell. The same thermal gradient described in the previous figure was used for this sample. As with the non-metallated WLN4 sample, there was a significant difference between the thermograms of the buffer blank and copper-loaded WLN4; this difference was attributed to incomplete buffer exchange of WLN4 following metallation. The buffer thermogram was not subtracted from the WLN4 thermogram for this reason. There was no unfolding transition observed. No protein precipitation was observed following the thermal gradient. Only one run of DSC was performed.
Figure 4.23: The thermogram of apo-WLN1-4. The temperature in ºC is shown on the x-axis, and the heat capacity (cal/ºC) is shown on the y-axis. Various concentrations of apo-WLN1-4 were loaded into the sample cell while 50 mM sodium phosphate, pH 7.5 was loaded into the reference cell. The same thermal gradient described in the previous figures was used for this sample, but no melting point could be determined. At concentrations high enough to yield a signal, the protein aggregated (≥ 0.5 mg/mL); lower concentrations produced insufficient signal for analysis.

4.5 NMR

4.5.1 \(^1\)H-\(^{15}\)N HSQCs of WLN1-2 and WLN1-4

\(^1\)H-\(^{15}\)N HSQC spectra were collected on \(^{13}\)C and \(^{15}\)N isotopically enriched samples of apoWLN1-2 and apoWLN1-4 using a Bruker AVANCE II 800 MHz NMR, as discussed in Section 2.5.6. The spectra were overlaid and are shown below in Figure 4.24. Both proteins display well dispersed resonances, which is indicative of a folded, non-aggregated protein.
Figure 4.24: $^1$H-$^{15}$N HSQC overlay of apoWLN1-2 (shown in red) and apoWLN1-4 (shown in blue). Both spectra display well dispersed resonances, indicative of a folded protein. The intra-protein overlapping peaks between 7.9-8.5 ppm are found in regions of proteins that are not in $\alpha$-helices or $\beta$-sheets, i.e. the linkers and loops. WLN1-2 contained 157 residues and WLN1-4 contained 374 residues.

The overlaid spectra showed that many of the apoWLN1-2 resonances are either on top of or very near to resonances in apoWLN1-4. This indicates that WLN1-2 has a similar structure in WLN1-4 as it does when it is by itself. The least amount of signal overlap between the two spectra occurs in the lower left quadrant, from 8.7-10 ppm on the x-axis and 123-134 ppm on the y-axis. The identity of these residues is not known because the resonances were unable to be assigned for either WLN1-2 or WLN1-4. However, resonances within this region are found in $\beta$-sheets. The shifting of resonances
indicates that the residues involved in one (or more) β-sheet(s) change positions slightly when WLN1-2 is expressed as part of WLN1-4.

4.5.2 $^1$H-$^{15}$N HSQC of WLN4 at Different Temperatures

$^1$H-$^{15}$N HSQC spectra were collected on $^{13}$C and $^{15}$N isotopically enriched samples of apoWLN4 at 25°C, 37°C, and 50°C using a Bruker AVANCE II 800 MHz NMR, as discussed in Section 2.5.5. Two 3-dimensional experiments, the HNCACB and the HN(CO)CACB, were collected at 37°C. These experiments, which correlated the N and HN of i with Cα and Cβ of i and i-1 and the Cα and Cβ of i-1, respectively, were necessary to assign the resonances in the $^1$H-$^{15}$N HSQC.

The $^1$H-$^{15}$N HSQCs from 25°C and 50°C are overlaid in Figure 4.23. At both temperatures, the resonances are well dispersed, indicating that the protein is folded and not aggregated. The thermal stability of WLN4 that was observed by CD was verified by the $^1$H-$^{15}$N HSQC at 50°C, which indicated no denaturation or aggregation. In fact, the peak resolution was better at 50°C than at 25°C; this was due to the faster tumbling time of the protein. The melting temperature of apoWLN4 that was calculated by using CD was 78.3 ± 0.2°C, and the unfolding curve (Figure 4.23) displayed little change in CD signal below 60°C. This NMR data further establishes the thermal stability of WLN4.

The $^1$H-$^{15}$N HSQC at 37°C is shown with the peak assignment in Figure 4.26; the amino acid sequence of WLN4 is shown in Figure 2.5. Of the 74 amino acids in WLN4, 9 non-proline residues were unable to be located in the HSQC (there are two prolines in the sequence: Pro50 and Pro55); the unassigned peaks were the first three amino acids (Gly1, Thr2, and Cys3) and the loop containing the metal-binding loop (Cys15, Ala16, Ser17, Cys18, Val19, and H20). The metal-binding loop, which contains Cys15 and
Cys18, is flexible, and the residues in that region could not be assigned with the 3D spectra that were collected. An enlarged view of the center of the spectra of Figure 4.26 is shown in Figure 4.27. A list of the $^1$H, $^{15}$N, and $^{13}$C resonances is in Table A1 in the Appendix.
Figure 4.25: $^1$H-$^{15}$N HSQC of apoWLN4, recorded at 25°C (red) and 50°C (black). At both temperatures show sharp, well dispersed resonances, indicative of a folded protein. At 50°C, the resolution increased as the tumbling time increased, resulting less overlapping peaks. Additionally, 50°C WLN4 exhibited no signs of aggregation or denaturation. The thermal stability observed here is consistent with the CD results indicated that the construct did not unfold until at mesophilic temperatures.
Figure 4.26: $^1$H-$^{15}$N HSQC of apoWLN4 with peak assignment.
Figure 4.27: An enlarged view of a portion of Figure 4.26.
4.6 Summary

This chapter discussed the biophysical characterization of the WLN constructs using HRGF, laser light scattering, circular dichroism, differential scanning calorimetry, and NMR.

HRGF results indicated that the constructs were globular, as they eluted at volumes consistent with their molecular weights. The hydrodynamic radii and frictional ratios were calculated from light scattering experiments and confirmed the globular nature of the constructs. As more domains were added to the constructs, the frictional ratio increased and the constructs were no longer perfectly spherical, as was expected due to the increasing size of the constructs.

The stability of the constructs was assessed by using guanidine hydrochloride and thermal denaturation in conjunction with circular dichroism. The results revealed that metal-binding domain four is highly resistant to denaturation by these methods, as it retained 50% of its structure at 5.9 M GuHCl and at 78.3ºC. WLN1-2 and WLN1-3 displayed similar properties to each other, unfolding at 2.5 M GuHCl and ~55ºC. In contrast to the other constructs, the GuHCl unfolding of WLN1-4 occurred via a three-state unfolding pathway; the first transition coincided with that of WLN1-3 and the second transition coincided with that of WLN4. Additionally, a thermal unfolding transition was not observed for WLN1-4; instead, a structural rearrangement occurred during heating which resulted in soluble aggregation. The addition of 150 mM NaCl to WLN1-4 resulted in thermal denaturation, though the melting point could not be calculated due to insufficient data points that could be collected after the unfolding transition. The ability of WLN1-4 to thermally denature in the presence of NaCl suggests
that electrostatic contacts are important in stabilizing this construct; as no structure for the construct is available, these contacts cannot be determined. When copper was titrated into the constructs, the MRE at 222 became less negative. As NMR experiments of WLN3-4 and WLN5-6 have clearly established that no loss of secondary structure occurs upon the addition of copper, there must be another explanation. One possibility is that the change in CD signal is due to the introduction of a chiral site that created when copper binds to the construct and this site absorbs circularly polarized light in the far-UV region; if more right-handed circularly polarized light were absorbed, then this would explain the decrease in the CD signal.

DSC was used to determine the melting temperatures of apo- and Cu(I)-WLN4, and apo-WLN1-4. The melting temperature of apo-WLN4 was ~84.5°C, similar to the value that was calculated by using CD (77.9°C). Cu(I)-WLN4 did not display an unfolding transition, and this may have been due to the presence of soluble aggregation; the melting temperature for metallated constructs varied wildly when CD was used, and that may have been due to soluble aggregation. No transition point was detected for apo-WLN1-4 for temperatures up to 110°C, though a small endothermic peak was visible; this peak was probably the result of soluble aggregation.

$^1$H-$^{15}$N HSQCs were collected for WLN1-2, WLN1-4, and WLN4. The spectrum of WLN1-2 overlaid well with the spectrum of WLN1-4, indicating that in the context of WLN1-4, WLN1-2 adopted largely the same structure as it did when it was expressed by itself. This could suggest that WLN1-2 is not interacting with the third and fourth metal-binding domains in the absence of copper. Decreased signal overlap in the lower left-hand region in Figure 4.24 was observed, indicating that residues involved in β-sheets
have slightly different positions in WLN1-2 and WLN1-4. \(^1\)H-\(^{15}\)N HSQCs collected at 25°C and 50°C revealed that WLN4 retained its structure at elevated temperatures and did not aggregate, which corroborated the thermal stability that CD experiments indicated. Triple resonance experiments collected at 37°C permitted the assignment of 63 of the 74 amide resonances in the \(^1\)H-\(^{15}\)N HSQC (collected at 37°C).

This research revealed the unexpected chemical and thermal stability of WLN4. One explanation of WLN4’s stability may be its hydrophobicity. As Table 4.6 below shows, WLN4 is the most hydrophobic domain – it is nearly twice as hydrophobic as the next most hydrophobic domain, WLN1. The large number hydrophobic forces likely stabilize the domain and make it resilient to GuHCl denaturation as well as thermal denaturation. Another biophysical parameter, the alipathic index, is also shown in the table below. This value indicates the relative volume of the protein that is occupied by alanine, isoleucine, leucine, and valine; larger values are correlated with higher thermostability (16). The aliphatic index value for WLN4 was in the lower half of the values of the six metal-binding domains, suggesting that it should be less thermostable. WLN3 has the highest aliphatic index, suggesting that it should have the highest melting point. In the context of WLN1-3, however, the presence of WLN3 did not result in a higher melting point than that of WLN1-2, as both WLN1-2 and WLN1-3 unfolded at nearly the same temperature (59.5°C and 58.3°C, respectively). WLN3 has a very low hydropathic value of 0.077, indicating that it is only slightly hydrophobic.

Perhaps the hydrophobic nature of WLN4 also plays a role in its thermal stability. Though WLN4 is the most hydrophobic, it is also the most acidic, with a theoretical pI of 3.93. There are 9 acidic residues, and many of these are clustered together on the surface
to form a negatively charged patch (Figure 1.7). This patch, predicted to interact with a basic patch on HAH1, helps to solubilize WLN4. The low pI greatly enhanced the solubility of WLN4, and during this research, it was frequently concentrated to 1 mM with no precipitation.

**Table 4.6:** Biophysical parameters of the metal-binding domains. These values were computed by using the ProtParam tool on the ExPASy website ([http://web.expasy.org/protparam/](http://web.expasy.org/protparam/)). The metal-binding domains did not include either the preceding or succeeding linker. The GRAVY value (Grand Average of Hydropathicity) is calculated by summing the hydropathicity values for all of the amino acids using the Kyte-Doolittle scale; positive values are hydrophobic and negative values are hydrophilic. The aliphatic index is the relative volume occupied by alanine, isoleucine, leucine, and valine residues in a protein; a higher value is correlated with greater thermostability.

<table>
<thead>
<tr>
<th>Metal-binding domain</th>
<th>Theoretical pI</th>
<th>GRAVY</th>
<th>Aliphatic Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>7.86</td>
<td>0.338</td>
<td>100.00</td>
</tr>
<tr>
<td>2</td>
<td>8.71</td>
<td>-0.169</td>
<td>98.61</td>
</tr>
<tr>
<td>3</td>
<td>6.72</td>
<td>0.135</td>
<td>113.61</td>
</tr>
<tr>
<td>4</td>
<td>3.93</td>
<td>0.579</td>
<td>105.69</td>
</tr>
<tr>
<td>5</td>
<td>4.53</td>
<td>0.236</td>
<td>105.69</td>
</tr>
<tr>
<td>6</td>
<td>5.51</td>
<td>0.251</td>
<td>110.00</td>
</tr>
<tr>
<td>1-2</td>
<td>8.26</td>
<td>96.11</td>
<td>0.046</td>
</tr>
<tr>
<td>1-3</td>
<td>8.36</td>
<td>97.39</td>
<td>-0.074</td>
</tr>
<tr>
<td>1-4</td>
<td>5.81</td>
<td>92.22</td>
<td>-0.064</td>
</tr>
</tbody>
</table>

WLN4 does display thermal stability that is reminiscent of a protein from a thermophilic organism while MBDs1-3 seem more mesophilic in origin. For example, the melting temperature of Cold Shock Protein B (CspB) was determined for two
different organisms - one was a mesophile, *Bacillus subtilis*, and the other was a thermophile, *Thermus thermophilus*. Both proteins contained 67 amino acids, but the melting temperature CspB from *T. thermophilus* was 86°C while it was 54°C from *B. subtilis* (17).

A study that analyzed the sequences of mesophilic and thermophilic proteins found that both classes of proteins have similar hydrophobicities and compactness; additionally, both have similar polar and non-polar contributions to the surface area as well as main chain and side chain hydrogen bonds (17). Thermophilic proteins had more side chain – side chain bonds and salt bridges, however, and also contained more arginines and tyrosines and fewer cysteines and serines (17).

To see if side chain–side chain bonds played a role in the thermal stability of WLN4, the solution structures of WLN3 and WLN4 (the linker region was not included) using PyMol and VMD (http://www.ks.uiuc.edu/Research/vmd/) to look for side chain - side chain bonds and salt bridges, respectively. There is only a solution structure for WLN3-4 (PDB ID 3ROP) and not for the domains 3 and 4 separately. The solution structure of WLN3-4 was used to obtain the structures of domains 3 and 4 by themselves. No striking differences between the two domains were observed. Additionally, WLN4 does not contain more arginines and tyrosines than other MBDs nor fewer cysteines and serines, so the thermal stability cannot be rationalized that way.

The metallochaperone HAH1 also contains a ferredoxin fold and has had its thermal and GuHCl stability reported in the literature. In one study, the melting temperatures of apo and Cu(I)-HAH1 have been reported as 74°C and 92°C, respectively; the unfolding of both apo- and Cu(I)-HAH1 was reversible (18). In that study, the
samples contained 20 mM Tris, pH 7.5 and 150 µM DTT and the CD signal was monitored at 220 nm. In another study the same authors reported that the thermal stability of apo- and Cu(I)-HAH1 was 68°C and 71°C, respectively; the unfolding was reversible (19). In the second study, the samples contained 20 mM Tris, pH 7.5 and 1 mM DTT; the CD signal was again monitored at 220 nm. The difference in melting temperatures may be due to the formation of an intermolecular disulfide bond at elevated temperatures; a high [DTT] would prohibit the disulfide bond from forming, resulting in a lower melting temperature. These thermal unfolding conditions were similar to those used in this dissertation project: the CD signal was monitored at 222 nm and the samples were buffered at pH 7.5 with 5 mM phosphate. For the research reported in this dissertation, DTT was not used. The GuHCl unfolding of both apo- and Cu(I)-HAH1 was reported to be two-state (19). The presence of Cu(I) increased the stability of HAH1, with the Cm for apo and Cu(I) forms calculated as 1.8 ± 0.1 M and 2.3 ± 0.1 M, respectively (20). The free energy of unfolding increased from 14.5 ± 0.6 kJ/mol to 16.7 ± 0.7 kJ/mol in the apo- and Cu(I)- forms (19).

The sequence of HAH1 (see Figure 2.6) was also analyzed by ProtParam tool on the ExPASy website (http://web.expasy.org/protparam/). The aliphatic index value was 87.54, the GRAVY value was -0.123, and the predicted pI is 6.70. If the aliphatic index is the primary basis for thermostability, then the expected melting temperature would be quite low. Additionally, HAH1 displays a negative GRAVY value, indicating that it is hydrophilic. This GRAVY value is in sharp contrast to the GRAVY value of WLN4, which is 0.576 (hydrophobic). While both apo-HAH1 and apoWLN4 have the same
ferredoxin fold, nearly the same melting temperatures and display reversible unfolding, the domains have different biophysical characteristics.

Another copper containing human protein that displays high thermal stability is copper-zinc superoxide dismutase (SOD1). When mature, SOD1 contains one atom each of copper and zinc as well disulfide bond that dimerizes that makes a homodimer; in this form, the melting point is >90°C (20). When SOD1 lacks the both metal ions as well as the disulfide bridge, it exists as a monomer and the melting temperature drops to 42°C; adding the disulfide bridge (without the metals) dimerizes SOD1 and increases the melting temperature to ~52°C (21). SOD1 does not contain a ferredoxin fold. Instead, each subunit contains an eight-stranded, antiparallel, Greek-key-β-barrel motif as well as several loops; one of the loops contains the metal-binding sites for both metal ions (22). In SOD1, the high thermal stability is linked to the binding of the copper and zinc ions, as in their absence SOD1 has a very low melting temperature. In contrast, both HAH1 and WLN4 have high melting temperatures even in the absence of copper.

Sequence alignment of the six MBDs was performed used Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalw2/) and appears in Figure 4.28. WLN4 does not appear to possess any readily apparent differences in amino acid sequence that would explain why this domain has such high stability. Discounting the highly conserved metal-binding region (residues 10-16 in the figure), there are 9 residues that are fully conserved among all six MBDs. Additionally, another 12 residues have highly similar properties (e.g. bulky, hydrophobic amino acids like valine or isoleucine appear in all MBDs at position 8) while an additional 4 residues have weakly similar properties (e.g. a small hydrophobic residue like alanine or a bulkier β-branched amino like valine appears
in position 67 in all MBDs.) Of the 72 amino acids that appear in the sequence alignment, 16 are fully conserved, 12 are strongly conserved, and 4 are weakly conserved; this amounts to 32 residues or 44.4% of the amino acids in the MBDs. The remaining 40 amino acids are not conserved among the MBDs.

The research presented in this dissertation regarding the stability of WLN4 has not appeared in the literature. While the extreme conditions that this domain can tolerate do not exist the cell, the remarkable stability that WLN4 possesses may play a role in its ability to function as a primary site of copper acquisition.

4.7 Future Work

The research presented in this dissertation revealed the remarkable chemical and thermal stability of WLN4. This surprising stability exists both when WLN4 is expressed by itself and when it is expressed as part of WLN1-4, which suggests that the stability is important for the function of WLN4 within the N-terminal metal-binding domain.

To continue investigations in this area, there are several avenues of research that could be considered. One avenue would be to examine the effect of copper binding on the stability of WLN4. Another avenue would be to confirm the chemical and thermal stability of the copper metallochaperone HAH1, since previous studies gave conflicting data (18,19). As WLN4 is one of the preferred sites of copper acquisition from HAH1, it would be interesting to see if HAH1 possesses the same stability. A third avenue of research would be to determine either a NMR solution structure or X-ray crystal structure for WLN1-4. A complete structural description of WLN1-4 provide insight into how the metal-binding domains are arranged and would aid in understanding their role in WLNP.
Figure 4.28: Sequence alignment of the 6 MBDs. The sequences were aligned using Clustal Omega. Amino acids are color coded as follows: red – small and hydrophobic (AVFPMILW), blue – acidic (DE), magenta – basic (RK), Green – hydroxyl, sulfhydryl, amino, and glycine (STYHCNGQ). The symbols below the line denote consensus, with “*” denoting a fully conserved residue, “:” denoting conservation between residues with very similar properties, and “.” denoting conservation between groups with weakly similar properties. A dash represents a gap in the sequence. The same sequences that were analyzed for Table 4.6 were aligned here.
References


CHAPTER 5
THE EFFECTS OF PH AND HG(II) BINDING ON THE ASSOCIATION STATE OF HAH1

This chapter discusses the HRGF study that was done in order to determine the effect of pH and Hg(II) binding on the association state of the copper metallochaperone HAH1. The study was performed as part of a collaboration with Professor Vincent Pecoraro and his group at the University of Michigan and included $^{199}$Hg NMR and $^{199m}$Hg PAC experiments. The goal of the study was to better characterize the coordination environment of Hg(II) in HAH1 as a function of pH. Seeing the coordination environment at both physiologically relevant and more extreme pHs would enable a better understanding of the coordination states that are available to Hg(II), a model for Cu(I), and may suggest a mechanism by which copper can be transferred from HAH1 to other proteins (or may suggest intermediates that are formed when Cu(I) is transferred from HAH1 to WLN or MNK).

5.1 Introduction

HAH1 is a small, single domain metallochaperone that possesses a ferredoxin fold and can ligate a single atom of Cu(I) via two cysteines that are located in a solvent exposed loop (1, 2). This protein, also known as Atox1, delivers cytosolic copper to the trans-Golgi network for incorporation into cuproproteins such as ceruloplasmin. HAH1 delivers the copper to a membrane-embedded ATPase that pumps the copper across the membrane into the interior of the trans-Golgi network. There are two ATPases that accept the copper, each being expressed in different locations: in liver cells, the ATPase
is Wilson disease protein while the ATPase in other tissues is Menkes protein (also known as ATP7A).

Experimental evidence has shown that the transfer of copper from HAH1 to both the Wilson disease protein and the Menkes protein involves the formation of a heterodimer in which the copper is ligated by both HAH1 and one of the metal-binding domains of the target protein (2–4). The nature of the intermediate that is formed during copper transfer is not well characterized. The intermediate could be a two-, three-, or four-coordinate structure, as each metal-binding domain and HAH1 possesses two cysteines in the metal-binding loop. Computation analysis suggests that a three-coordinate intermediate is most energetically favored, with two three-coordinate intermediates occurring during copper transfer (5). In this scheme, two ligands initially come from HAH1 and one from the metal-binding domain; in the second of the intermediates, two ligands come from the metal-binding domain and one comes from HAH1.

The NMR solution structures of apo- and Cu(I)-HAH1 have been solved and the structures indicate that the copper is ligated by two cysteines from a single molecule of HAH1 and that no homodimers were detected (1). It seems probable that a higher order intermediate would be formed, however, as the copper is transferred.

Crystal structures for HAH1 complexed with Cd(II), Hg(II) or Cu(I) have been determined (2). In these crystals, HAH1 formed a three-coordinate homodimeric complex in the presence of Hg(II) and a 4-coordinate homodimeric complex in the presence of Cd(II). The Cu(I)-HAH1 crystals were also homodimeric as well, with the Cu(I) found in either a 3- or 4-coordinate state; the resolution could not unambiguously
determine whether or not a fourth ligand was present, though if it were it would be close to the limit of known copper-sulfur bond lengths. The presence of the higher coordinated copper center in the crystals provides evidence that such a coordination state is possible. Such a state may only be an artifact of the crystal packing, however, and not necessarily indicative of a physiologically relevant state. The quest for experimental evidence of a higher coordinated intermediate state was the impetus for this study.

To study the coordination of metals in proteins in a simplified system, Pecoraro and his group have designed a metallopeptide that contains a metal-binding site in the interior of an α-helical coiled coil (6). In the presence of Hg(II), the metallopeptides self-assemble into dimers and trimers. Each metallopeptide contributes one metal-binding site to coordinate the Hg(II) atom in the hydrophobic interior of the coil. Physiological pH favors a two-coordinate dimer while a pH of 8.5 favors the formation of a 3-coordinantate trimer. Variations of the peptide have been used to study the coordination states and binding geometries of Cd(II) and As(III), as well as to further study Hg(II) (7, 8).

Both Cu(I) and Hg(II) are soft acids and have similar preferences for coordination geometries. Since Hg(II) will bind to HAH1 and there is an isotope of Hg(II), $^{199}$Hg that is amenable to techniques such as NMR and perturbed angular correlation (PAC), Hg(II) was used as a model for Cu(I) to investigate the possible coordinate environments that could exist for Cu(I). Using Hg(II) as a model was further validated by the fact the yeast Atx(I) can transfer Hg(II) to the Ccc2, the yeast homologue of Wilson disease protein; this suggests that Hg(II) and Cu(I) are transported by similar mechanisms and may form similar intermediates during copper transfer (9).
5.2 Analytical Gel Filtration

Analytical gel filtration experiments were done to determine how pH and Hg(II) changed the association state of HAH1, which was expressed and purified as discussed in Section 2.4.5.4. The procedure outlined below was developed by Dr. Brian Zeider during the course of his PhD dissertation (10).

A Superdex 75 16/600 column was equilibrated with 1 CV (120 mL) of buffer using a flow rate of 1.0 mL/min. Three buffers were used: 100 mM sodium phosphate (pH 7.5), 100 mM sodium phosphate (pH 8.5), and 100 mM CHES (pH 9.4); each buffer also contained 200 mM NaCl and 1 mM TCEP (to prevent the oxidation of cysteines at elevated pHs); DTT was not used because it is a ligand for Hg(II). As the gel filtration experiments were run at 4°C, the buffers were prepared at room temperature such that they would have the desired pH at 4°C. Hg(II)-HAH1 samples were freshly prepared by adding 0.5 equivalents of Hg(II) (from HgCl₂) to HAH1 that was in the desired buffer. Duplicate injections of apo- and Hg(II)-HAH1 as well as the protein standards were analyzed for each buffer. The elution volumes, taken at the peak apex, were determined by using the Unicorn Software that was supplied with the ÄKTA FPLC and are shown in Tables 5.1-5.3 below.
### Table 5.1: Table of protein standards and HAH1 elution volumes at pH 7.5.

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (Daltons)</th>
<th>Log MW</th>
<th>$V_e$</th>
<th>$K_{av}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>6,500</td>
<td>3.81</td>
<td>91.43</td>
<td>0.62</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>13,700</td>
<td>4.14</td>
<td>78.94</td>
<td>0.45</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>29,000</td>
<td>4.46</td>
<td>65.56</td>
<td>0.27</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43,000</td>
<td>4.63</td>
<td>57.64</td>
<td>0.16</td>
</tr>
<tr>
<td>Apo-HAH1</td>
<td>9,500*</td>
<td>N/A</td>
<td>85.30</td>
<td>0.53</td>
</tr>
<tr>
<td>Hg(II)-HAH1</td>
<td>9,700*</td>
<td>N/A</td>
<td>84.65</td>
<td>0.53</td>
</tr>
</tbody>
</table>

*The MW that was calculated from the standard curve. The actual MW for apo- and Hg(II)-HAH1 are 7,341 and 7,351 Daltons, respectively.

### Table 5.2: Table of protein standards and HAH1 elution volumes at pH 8.5.

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (Daltons)</th>
<th>Log MW</th>
<th>$V_e$</th>
<th>$K_{av}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>6,500</td>
<td>3.81</td>
<td>91.48</td>
<td>0.62</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>13,700</td>
<td>4.14</td>
<td>78.79</td>
<td>0.45</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>29,000</td>
<td>4.46</td>
<td>65.72</td>
<td>0.27</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43,000</td>
<td>4.63</td>
<td>57.52</td>
<td>0.16</td>
</tr>
<tr>
<td>Apo-HAH1</td>
<td>9,500*</td>
<td>N/A</td>
<td>84.94</td>
<td>0.53</td>
</tr>
<tr>
<td>Hg(II)-HAH1</td>
<td>10,500*</td>
<td>N/A</td>
<td>83.15</td>
<td>0.51</td>
</tr>
</tbody>
</table>

*The MW that was calculated from the standard curve. The actual MW for apo- and Hg(II)-HAH1 are 7,341 and 7,351 Daltons, respectively.

### Table 5.3: Table of protein standards and HAH1 elution volumes at pH 9.4.

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (Daltons)</th>
<th>Log MW</th>
<th>$V_e$</th>
<th>$K_{av}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aprotinin</td>
<td>6,500</td>
<td>3.81</td>
<td>90.29</td>
<td>0.60</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>13,700</td>
<td>4.14</td>
<td>78.32</td>
<td>0.44</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>29,000</td>
<td>4.46</td>
<td>67.34</td>
<td>0.29</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43,000</td>
<td>4.63</td>
<td>59.39</td>
<td>0.19</td>
</tr>
<tr>
<td>Apo-HAH1</td>
<td>8,400*</td>
<td>N/A</td>
<td>84.94</td>
<td>0.53</td>
</tr>
<tr>
<td>Hg(II)-HAH1</td>
<td>11,500*</td>
<td>N/A</td>
<td>83.15</td>
<td>0.51</td>
</tr>
</tbody>
</table>

*The MW that was calculated from the standard curve. The actual MW for apo- and Hg(II)-HAH1 are 7,341 and 7,351 Daltons, respectively.
Figure 5.1: Chromatograms of apo-HAH1 at various pHs (11). Equal amounts of each sample were injected onto a Superdex 75 16/600 column and eluted using an isocratic gradient. Apo-HAH1 was a monomer at pHs 7.5 and 8.5, though the mass of the protein decreased slightly at pH 9.4; this decrease may be due to the protein adopting a molten globule structure. Figure used by permission of John Wiley and Sons.

Figure 5.2: Chromatograms of Hg(II)-HAH1 at various pHs (11). Equal amounts of each sample (the same amounts as used in Figure 5.1) were injected onto a Superdex 75 16/600 column and eluted using an isocratic gradient. As the pH increases, the apparent mass of Hg(II)-HAH1 increases, as evidenced by smaller elution volumes. The shoulder at ~85 mL at pH 9.4 was likely apo-HAH1, as that was the elution volume for the protein at this pH. Figure used by permission of John Wiley and Sons.
As the chromatograms in **Figure 5.1** show, apo-HAH1 existed as a monomer at pHs 7.5 and 8.5. At pH 9.4 there was a slight decrease in the apparent molecular weight and an increase in A$_{254}$. The decrease in molecular weight may be due to HAH1 adopting a molten globule structure at this pH. CD confirmed that there were no significant alterations to the secondary structure in the far-UV region (see A2 in the Appendix). This was consistent with the presence of a molten globule, which possesses the correct secondary structure but lacks the correct tertiary structure and appropriate side chain packing of the properly folded protein. The increase in A$_{254}$ may be explained by the partial deprotonation of the two tyrosine residues (12). The pKa for the side chain of the free amino acid in solution is 10.4, and at a pH of 9.4, the side chain would be 9% deprotonated, assuming that the pKas were not significantly altered by an extreme microenvironment in HAH1.

In the presence of Hg(II), HAH1 began to dimerize as the pH increased; this is shown in **Figure 5.2**. This behavior is in contrast to what occurred for apo-HAH1, which remained monomeric at each pH tested. The apparent molecular weights are summarized in **Table 5.4**. While equal amounts of protein was injected for all of the apo- and Hg(II) runs, the signal intensity increased in the Hg(II) samples due to ligand-to-metal charge transfer.
Table 5.4: Summary of the masses of apo- and Hg(II)-HAH1 at various pHs. As the pH increases, HAH1 begins to dimerize in the presence of Hg(II).

<table>
<thead>
<tr>
<th>pH</th>
<th>Apo-HAH1</th>
<th>Hg(II)-HAH1</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.5</td>
<td>9,500 Da</td>
<td>9,700 Da</td>
</tr>
<tr>
<td>8.5</td>
<td>9,500 Da</td>
<td>10,500 Da</td>
</tr>
<tr>
<td>9.4</td>
<td>8,400 Da</td>
<td>11,500 Da</td>
</tr>
<tr>
<td>Actual MW</td>
<td>7,341 Da</td>
<td>7,541 Da</td>
</tr>
</tbody>
</table>

While these gel filtration studies provided evidence for a dimeric association state for HAH1 in the presence of Hg(II), information regarding the nature of the coordination state could not be deduced. To investigate the coordination states, $^{199}\text{Hg}$ NMR and $^{199m}\text{Hg}$ PAC experiments were performed.

5.3 Discussion

Analytical gel filtration results showed that in the presence of Hg(II), HAH1 began to dimerize as the pH increased. Investigation of the coordination states of the Hg(II) at the various pHs was carried out using $^{199}\text{Hg}$ NMR and $^{199m}\text{Hg}$ PAC (11).

The $^{199}\text{Hg}$ NMR spectra shown in Figure 5.3 has a single signal at -819 ppm for the pH 7.5 sample (10, 11). This is consistent with the chemical shift of -821 ppm that O’Halloran’s group observed for a 2-coordinate Hg(II) thiolate species of Cu(I)-Atx1, the yeast homologue of HAH1 (13). At pH 8.5, two signals appear – one at -806 ppm and the other at -348 ppm. The former signal was still consistent with a HgS$_2$ species. The signal at -348 ppm could be two possible structures - either a T-shaped 3-coordinate HgS$_3$ species or a 4-coordinate HgS$_4$ species (13–15). At pH 9.4, the signal at ~800 ppm disappeared, leaving only the signal at -348 nm. From the NMR experiments it was concluded that at physiological pH a 2-coordinate HgS$_2$ species was present; as the pH increased, the concentration of this species decreased and a species that was either 3- or
4-coordinate appeared. Since NMR could not be used to distinguish between the 3- and 4-coordinate structures, $^{199}$mHg PAC was used.

**Figure 5.3:** $^{199}$Hg NMR of Hg(II)-HAH1 at various pHs (11). The signal at -819 ppm is consistent with a 2-coordinate Hg(II) thiolate species while the signal at -348 ppm could be attributed to either a T-shaped 3-coordinate HgS$_3$ species or a 4-coordinate HgS$_4$ species (11). Figure used by permission of John Wiley and Sons.

The $^{199}$mHg PAC data is shown below in **Figure 5.4**. At physiological pH (solid line), the peak labeled with the down arrow dominated the spectra and was consistent with HgS$_2$ (16). As the pH was increased to 8.5 (dashed line), the amplitude of that peak decreased by one third and a new peak, labeled with the up arrow, began to emerge; this new peak was consistent with a distorted HgS$_4$ structure (17). Increasing the pH to 9.4 further reduced the amplitude of the signal for the HgS$_2$ species while the amplitude for the distorted HgS$_4$ increased. The signal complexity increased at pH 9.4, which complicated the analysis since there were several satisfactory models that could fit the data; one of the models that fit best was a T-shaped HgS$_3$ species. It was concluded that
at pH 9.4 an equilibrium of several species was present, including HgS$_2$, a distorted HgS$_4$ structure, and possibly a T-shaped HgS$_3$ species.

The $^{199}$Hg NMR data indicated only a single peak in the spectra, which corresponded to either a 3- or 4-coordinate species while the $^{199m}$Hg PAC data was more complex and indicated the presence of multiple species. The authors pointed out that this may be due to the exchange rate of the species relative to the time scale of the technique. NMR operates on the millisecond time scale, and an exchange rate that is faster than this will produce only one peak in the spectra. In contrast, PAC operates on a faster time scale: microseconds-nanoseconds (µs-ns). Since there was a mixture of higher coordinate species present at pH 9.4 in the PAC experiments but not the NMR experiment, the exchange rate for the speciation must be on the µs-ns timescale.

![Figure 5.4](image-url)

**Figure 5.4:** $^{199m}$Hg PAC results for HAH. The solid line is at pH 7.5, the dashed line is at pH 8.5, and the dotted line is at pH 9.4. The down arrow denotes a species consistent with HgS$_2$ while the up arrow denotes species that are consistent with a distorted HgS$_4$ structure (11). Figure used by permission of John Wiley and Sons.
5.4 Conclusion

These experiments gave insight into the association and coordinate states Hg(II)-HAH1. Analytical gel filtration experiments showed that Hg(II)-HAH1 began to dimerize as the pH increased while apo-HAH1 remained monomeric. $^{199}$Hg NMR revealed the presence of a 2-coordinate HgS$_2$ species at physiological pH, a combination of 2-coordinate and 3- and/or 4-coordinate species at pH 8.5, and only a 3- and/or 4-coordinate species at pH 9.4. $^{199m}$Hg data confirmed the presence of a 2-coordinate HgS$_2$ species at pH 7.5 while at pH 8.5 there was a 2-coordinate HgS$_2$ species and a distorted HgS$_4$ species. The signal at pH 9.4 was more complex and could not be unambiguously deconvoluted. What was apparent, however, was that multiple species were present. The models that best fit the spectra are a T-shaped HgS$_3$ species and a distorted HgS$_4$ species. The presence of multiple species in the PAC data and only a single signal in the NMR spectra provided information regarding the exchange rate of the complex. The exchange rate had to be faster than the NMR timescale (ms) since only one peak was seen in that spectra, but slower than the PAC timescale ($\mu$s-ns) since multiple peaks were observed. Overall, the experiments showed that higher coordination states could be achieved for mercury-thiolate compounds, some of which may mimic the intermediates that are formed during copper transfer from HAH1 to the metal-binding domains of Wilson disease proteins and Menkes protein.
References


Figure A.1: Thermal unfolding of WLN1-6. The data is plotted at 222 nm. As with WLN1-4, no discernible unfolding transition is observed for WLN1-6. Additionally, the slight decrease in CD signal is likely due to the same structural rearrangement was observed with WLN1-4 that resulted in the formation of soluble aggregates. Gel filtration experiments were not done with WLN1-6 to confirm this, however.
**Figure A.2:** CD spectra of HAH1 at pH 9.4. Data was collected at 4°C in a 1 mm cuvette. HAH1 maintained its secondary structure after 24 hours at pH 9.4. Adding Hg(II) did not immediately produce significant changes in the region of 200-250 nm, though after 24 hours the signal became less negative in this area.

**Figure A.3:** CD spectra of HAH1 at pH 8.5. Data was collected at 4°C in a 1 mm cuvette.
Table A.1: $^1$H, $^{15}$N, and $^{13}$C resonance assignments for apoWLN4 at 37°C.

<table>
<thead>
<tr>
<th>Residue</th>
<th>HN</th>
<th>N</th>
<th>CA</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>G1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S4</td>
<td>8.657</td>
<td>120.172</td>
<td>57.536</td>
<td>65.510</td>
</tr>
<tr>
<td>T5</td>
<td>8.438</td>
<td>116.047</td>
<td>61.366</td>
<td>71.061</td>
</tr>
<tr>
<td>T6</td>
<td>9.260</td>
<td>122.985</td>
<td>60.076</td>
<td>69.849</td>
</tr>
<tr>
<td>L7</td>
<td>8.172</td>
<td>129.923</td>
<td>53.587</td>
<td>45.143</td>
</tr>
<tr>
<td>I8</td>
<td>9.174</td>
<td>125.235</td>
<td>59.255</td>
<td>41.742</td>
</tr>
<tr>
<td>A9</td>
<td>9.377</td>
<td>131.610</td>
<td>51.672</td>
<td>19.225</td>
</tr>
<tr>
<td>I10</td>
<td>8.509</td>
<td>122.610</td>
<td>60.663</td>
<td>41.039</td>
</tr>
<tr>
<td>A11</td>
<td>9.291</td>
<td>131.610</td>
<td>50.968</td>
<td>20.945</td>
</tr>
<tr>
<td>G12</td>
<td>8.517</td>
<td>107.487</td>
<td>45.261</td>
<td></td>
</tr>
<tr>
<td>M13</td>
<td>8.767</td>
<td>119.985</td>
<td>56.988</td>
<td>34.315</td>
</tr>
<tr>
<td>T14</td>
<td>9.158</td>
<td>113.988</td>
<td>62.079</td>
<td>71.407</td>
</tr>
<tr>
<td>C15</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A16</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S17</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C18</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>V19</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>H20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S21</td>
<td>8.282</td>
<td>116.422</td>
<td>61.836</td>
<td>62.657</td>
</tr>
<tr>
<td>I22</td>
<td>8.047</td>
<td>122.797</td>
<td>66.058</td>
<td>38.693</td>
</tr>
<tr>
<td>E23</td>
<td>8.650</td>
<td>117.36</td>
<td>60.467</td>
<td>29.428</td>
</tr>
<tr>
<td>G24</td>
<td>8.368</td>
<td>106.109</td>
<td>47.059</td>
<td></td>
</tr>
<tr>
<td>M25</td>
<td>7.687</td>
<td>119.61</td>
<td>58.161</td>
<td>32.752</td>
</tr>
<tr>
<td>I26</td>
<td>8.462</td>
<td>118.86</td>
<td>61.523</td>
<td>35.097</td>
</tr>
<tr>
<td>S27</td>
<td>7.899</td>
<td>110.797</td>
<td>61.523</td>
<td>63.165</td>
</tr>
<tr>
<td>Q28</td>
<td>7.015</td>
<td>116.422</td>
<td>55.542</td>
<td>29.624</td>
</tr>
<tr>
<td>L29</td>
<td>7.554</td>
<td>121.297</td>
<td>54.877</td>
<td>41.352</td>
</tr>
<tr>
<td>E30</td>
<td>8.579</td>
<td>125.235</td>
<td>58.004</td>
<td>28.998</td>
</tr>
<tr>
<td>G31</td>
<td>8.595</td>
<td>109.859</td>
<td>45.105</td>
<td></td>
</tr>
<tr>
<td>V32</td>
<td>7.742</td>
<td>121.110</td>
<td>64.181</td>
<td>31.735</td>
</tr>
<tr>
<td>Q33</td>
<td>9.205</td>
<td>128.423</td>
<td>56.597</td>
<td>29.936</td>
</tr>
<tr>
<td>Q34</td>
<td>7.891</td>
<td>116.422</td>
<td>55.777</td>
<td>32.869</td>
</tr>
<tr>
<td>I35</td>
<td>8.790</td>
<td>121.11</td>
<td>58.395</td>
<td>40.257</td>
</tr>
<tr>
<td>S36</td>
<td>8.319</td>
<td>120.41</td>
<td>56.128</td>
<td>64.963</td>
</tr>
<tr>
<td>V37</td>
<td>9.894</td>
<td>129.548</td>
<td>61.288</td>
<td>32.360</td>
</tr>
</tbody>
</table>
Table A.1—Continued

<table>
<thead>
<tr>
<th>Residue</th>
<th>HN</th>
<th>N</th>
<th>CA</th>
<th>CB</th>
</tr>
</thead>
<tbody>
<tr>
<td>S38</td>
<td>8.877</td>
<td>120.922</td>
<td>54.955</td>
<td>63.477</td>
</tr>
<tr>
<td>L39</td>
<td>8.915</td>
<td>131.223</td>
<td>57.848</td>
<td>41.899</td>
</tr>
<tr>
<td>A40</td>
<td>8.392</td>
<td>119.435</td>
<td>54.916</td>
<td>18.561</td>
</tr>
<tr>
<td>E41</td>
<td>7.711</td>
<td>112.297</td>
<td>56.128</td>
<td>30.562</td>
</tr>
<tr>
<td>G42</td>
<td>8.024</td>
<td>109.672</td>
<td>46.825</td>
<td></td>
</tr>
<tr>
<td>T43</td>
<td>7.468</td>
<td>106.859</td>
<td>58.943</td>
<td>73.563</td>
</tr>
<tr>
<td>A44</td>
<td>9.674</td>
<td>122.235</td>
<td>50.186</td>
<td>22.118</td>
</tr>
<tr>
<td>T45</td>
<td>8.822</td>
<td>119.985</td>
<td>62.383</td>
<td>69.576</td>
</tr>
<tr>
<td>V46</td>
<td>9.158</td>
<td>127.673</td>
<td>60.898</td>
<td>33.846</td>
</tr>
<tr>
<td>L47</td>
<td>9.072</td>
<td>130.485</td>
<td>54.095</td>
<td>43.384</td>
</tr>
<tr>
<td>Y48</td>
<td>9.362</td>
<td>123.735</td>
<td>55.229</td>
<td>41.547</td>
</tr>
<tr>
<td>N49</td>
<td>8.939</td>
<td>119.047</td>
<td>49.013</td>
<td>39.162</td>
</tr>
<tr>
<td>P50</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A51</td>
<td>7.961</td>
<td>119.985</td>
<td>53.157</td>
<td>18.834</td>
</tr>
<tr>
<td>V52</td>
<td>7.922</td>
<td>115.86</td>
<td>62.618</td>
<td>35.332</td>
</tr>
<tr>
<td>I53</td>
<td>8.321</td>
<td>121.297</td>
<td>60.311</td>
<td>42.329</td>
</tr>
<tr>
<td>S54</td>
<td>7.554</td>
<td>116.61</td>
<td>55.190</td>
<td>65.979</td>
</tr>
<tr>
<td>P55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E56</td>
<td>8.298</td>
<td>115.297</td>
<td>58.865</td>
<td>28.451</td>
</tr>
<tr>
<td>E57</td>
<td>7.594</td>
<td>119.422</td>
<td>59.332</td>
<td>30.250</td>
</tr>
<tr>
<td>L58</td>
<td>7.578</td>
<td>121.11</td>
<td>57.927</td>
<td>42.212</td>
</tr>
<tr>
<td>R59</td>
<td>8.431</td>
<td>119.235</td>
<td>59.881</td>
<td>29.311</td>
</tr>
<tr>
<td>A60</td>
<td>8.337</td>
<td>120.36</td>
<td>54.955</td>
<td>18.053</td>
</tr>
<tr>
<td>A61</td>
<td>7.703</td>
<td>120.36</td>
<td>54.902</td>
<td>17.842</td>
</tr>
<tr>
<td>I62</td>
<td>7.570</td>
<td>118.86</td>
<td>65.589</td>
<td>38.146</td>
</tr>
<tr>
<td>E63</td>
<td>8.384</td>
<td>122.047</td>
<td>58.982</td>
<td>29.272</td>
</tr>
<tr>
<td>D64</td>
<td>8.431</td>
<td>120.735</td>
<td>56.793</td>
<td>40.843</td>
</tr>
<tr>
<td>M65</td>
<td>7.617</td>
<td>116.235</td>
<td>55.190</td>
<td>65.979</td>
</tr>
<tr>
<td>G66</td>
<td>8.133</td>
<td>105.359</td>
<td>44.714</td>
<td></td>
</tr>
<tr>
<td>F67</td>
<td>6.905</td>
<td>118.485</td>
<td>56.832</td>
<td>41.273</td>
</tr>
<tr>
<td>E68</td>
<td>8.024</td>
<td>119.985</td>
<td>56.011</td>
<td>30.718</td>
</tr>
<tr>
<td>A69</td>
<td>9.041</td>
<td>127.673</td>
<td>50.655</td>
<td>23.134</td>
</tr>
<tr>
<td>S70</td>
<td>8.759</td>
<td>114.735</td>
<td>56.91</td>
<td>65.745</td>
</tr>
<tr>
<td>V71</td>
<td>9.111</td>
<td>126.173</td>
<td>64.259</td>
<td>31.969</td>
</tr>
<tr>
<td>V72</td>
<td>8.947</td>
<td>128.985</td>
<td>63.165</td>
<td></td>
</tr>
<tr>
<td>S73</td>
<td>8.141</td>
<td>117.922</td>
<td>57.966</td>
<td>64.611</td>
</tr>
<tr>
<td>E74</td>
<td>8.235</td>
<td>128.61</td>
<td>31.422</td>
<td></td>
</tr>
</tbody>
</table>
Figure A.4: Project approval certification.
Figure A.5: Permission from John Wiley for use of Figures 5.1 – 5.4.
Figure A.6: Permission from Nature Publishing Group for use of Figure 1.4.
**SPRINGER LICENSE**
**TERMS AND CONDITIONS**

May 19, 2014

This is a License Agreement between Alia Hinz ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

*All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.*

<table>
<thead>
<tr>
<th>License Number</th>
<th>3392530936750</th>
</tr>
</thead>
<tbody>
<tr>
<td>License date</td>
<td>May 19, 2014</td>
</tr>
<tr>
<td>Licensed content publisher</td>
<td>Springer</td>
</tr>
<tr>
<td>Licensed content publication</td>
<td>Springer eBook</td>
</tr>
<tr>
<td>Licensed content title</td>
<td>Biological Copper Transport</td>
</tr>
<tr>
<td>Licensed content author</td>
<td>Prof. David L. Huffman</td>
</tr>
<tr>
<td>Licensed content date</td>
<td>Jan 1, 2013</td>
</tr>
<tr>
<td>Type of Use</td>
<td>Thesis/Dissertation</td>
</tr>
<tr>
<td>Portion</td>
<td>Figures</td>
</tr>
<tr>
<td>Author of this Springer article</td>
<td>Yes and you are a contributor of the new work</td>
</tr>
<tr>
<td>Order reference number</td>
<td>Figure 1.7</td>
</tr>
<tr>
<td>Original figure numbers</td>
<td>Figure 3</td>
</tr>
<tr>
<td>Title of your thesis / dissertation</td>
<td>Structural Characterization of the First Four Metal-binding Domains of Human Wilson Disease Protein</td>
</tr>
<tr>
<td>Expected completion date</td>
<td>May 2014</td>
</tr>
<tr>
<td>Estimated size(pages)</td>
<td>200</td>
</tr>
<tr>
<td>Total</td>
<td>0.00 USD</td>
</tr>
</tbody>
</table>

**Terms and Conditions**

**Introduction**

The publisher for this copyrighted material is Springer Science + Business Media. By clicking "accept" in connection with completing this licensing transaction, you agree that the following terms and conditions apply to this transaction (along with the Billing and Payment terms and conditions established by Copyright Clearance Center, Inc. ("CCC"), at the time that you opened your Rightlink account and that are available at any time at http://myaccount.copyright.com).

**Limited License**

With reference to your request to reprint in your thesis material on which Springer Science and Business Media control the copyright, permission is granted, free of charge, for the use