Characterization of the N-Terminal Domains and Disease-Causing Mutations of the Human Wilson Protein

Joshua Mutambuki Muia
Western Michigan University

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

Part of the Chemistry Commons

Recommended Citation
https://scholarworks.wmich.edu/dissertations/609
CHARACTERIZATION OF THE N-TERMINAL DOMAINS AND DISEASE-CAUSING MUTATIONS OF THE HUMAN WILSON PROTEIN

by

Joshua Mutambuki Muia

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment of the
requirements for the
Doctor of Philosophy
Department of Chemistry
Advisor: David Huffman, Ph.D.

Western Michigan University
Kalamazoo, Michigan
May 2010
The Wilson protein (ATP7B) is a copper transporting ATPase that is involved in copper trafficking and homeostasis. Unlike the other known P-type ATPases, it possesses six homologous metal binding domains at the N-terminal end. Several mutations in the gene coding for this protein lead to Wilson disease, a hepatic disorder characterized by impaired excretion of copper in the bile, and accumulation of copper in body organs such as the liver, brain, kidney, and eye cornea. Characterization of various regions of expressed and purified ATP7B has been hampered by its low stability, aggregation and degradation.

In this research, novel methods were developed to express and purify stable N-terminal copper binding domains (MBDs), a 577 amino acid construct. These domains were characterized by Gel Filtration, Light Scattering and Circular Dichroism. The results indicate that the domains are polydisperse and have a larger hydrodynamic radius relative to a globular protein of the same molecular mass.

Y532H, a novel disease-causing mutation in domains 5 and 6 of the Wilson protein was also expressed. WD5-6(Y532H) could not be purified by established protocols. Therefore, it was purified using two different methods: (1) insoluble recombinant inclusion bodies were extracted and a procedure was developed to purify the refolded fusion protein and (2) a lower temperature procedure was developed to express
soluble *non-fusion* recombinant protein. The WD5-6(Y532H) protein was characterized by Gel Filtration, Light Scattering, Circular Dichroism and Nuclear Magnetic Resonance. The results reveal that the Y532H mutation does not affect the overall structural organization of the native WD5-6 and the conformational stability of the protein. However, the NMR data of the WD5-6(Y532H) recovered by refolding reveals structural perturbations in the area around the mutation.

This work represents advances in the understanding of the function of the N-terminal region of the Wilson protein (ATP7B). The purification methods and strategies, key findings, and new developments presented in this work will be helpful not only to the individuals working in this field, but also to the scientific community at large.
ACKNOWLEDGMENTS

I am highly indebted to many individuals who have contributed immensely to my success in this work. First and foremost, I would like to thank my advisor, Professor David Huffman, for his support, guidance, and inspiration in the course of this project. He made my experience at Western Michigan University memorable. I appreciate all his commitment to and passion for my success. I would also like to thank the other members of my thesis committee—Professors James Kiddle, Ekkhard Sinn, Wendy D Ransom-Hodgkins, and Dave Reinhold—for their kind assistance and the advice they offered towards this work. I appreciate their patience and the time they spent reviewing my manuscript and attending committee meetings.

It would have been difficult to complete my graduate studies without the help and support of the dedicated faculty and staff of the chemistry department at Western Michigan University. Professor John Miller, coordinator for graduate advising, extended a great deal of assistance. Dr. Sherine Obare's input and ideas with regard to career development were very helpful. Staff members Pam McCartney, Robin Lenkart, Tom Norton, Sean Shaw and Raymond Sung, Ph.D were awesome. I am grateful for moments shared and friendships created with past and present members of the lab. Past members include Joel Lwande, Jennifer Bunce, David Achila, Patrick Ochieng, Kate Schultz, Peter Sande, and Paul Oshule. Present members include: Brian Zeider, Nastaran Salehi, Alia Hinz, Neil Blok, and Wilson Okumu. Visiting students Katrina Diaz, Alex
Acknowledgments—continued

Jubril, Sarah Dickinson, and Sunny Lee always injected fresh life into the lab just by being around us.

I am very fortunate to have met Dr. John Onyari at the University of Nairobi, Kenya, who introduced me to the world of science. He ultimately invigorated my dreams. I’m grateful to our scientific collaborator in the Bertini group at the Center for Magnetic Resonance, Florence, Italy. I was privileged to visit this dedicated group during the fall of 2007. Being far away from our home country, our stay in Kalamazoo was greatly enriched by the friendship of Dr. Huffman’s family, his wife, Ruthann, and their children, Randy and Rachel. I thank the pastors and members of Kalamazoo Valley Family Church for spiritual support and nourishment. Occasional get-togethers with the Kenyan community in Michigan kept homesickness at bay.

Lastly, I would like to thank my family for their love, encouragement, and patience. In particular, I am thankful to my wife, Jacinta, with whom I spent many late nights working through thesis manuscript drafts. I regret the sudden demise of my loving dad, William Kimondiu, over this period. He was a monumental figure in my life and taught me about hard work, honesty, and trust. The encouragement and support received from my mum and other family members is highly appreciated and will always be remembered.

Joshua Mutambuki Muia
TABLE OF CONTENTS

ACKNOWLEDGEMENTS........................................................................................................ii

LIST OF TABLES...................................................................................................................viii

LIST OF FIGURES...................................................................................................................x

LIST OF ABBREVIATIONS.....................................................................................................xv

CHAPTER

1 WILSON PROTEIN AND COPPER HOMEOSTASIS..............................................................1

1.1 Introduction......................................................................................................................1

1.2 Wilson and Menkes ATPases: Homologous Copper ATPases........................................2

1.2.1 Structural Properties of Wilson and Menkes Protein....................................................2

1.2.2 Copper Permeases and Metallochaperones.................................................................13

1.2.3 Serum Ceruloplasmin..................................................................................................17

1.2.4 Functional Properties of Wilson and Menkes Protein....................................................18

1.3 Wilson Disease...............................................................................................................28

1.3.1 The History of Wilson Disease....................................................................................28

1.3.2 Clinical Aspects and Diagnosis of Wilson Disease.......................................................29

1.3.3 Treatment of the Wilson Disease..................................................................................32

1.4 Summary.........................................................................................................................35

References................................................................................................................................37

2 MATERIALS AND EXPERIMENTAL METHODS.............................................................59

2.1 Introduction......................................................................................................................59

2.2 Chemicals and Reagents.................................................................................................59
<table>
<thead>
<tr>
<th>CHAPTER</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3 DNA Plasmids and PCR Primers............................................62</td>
</tr>
<tr>
<td>2.4 Instrumentation........................................................................66</td>
</tr>
<tr>
<td>2.5 Procedures..............................................................................68</td>
</tr>
<tr>
<td>2.5.1 Cloning, Expression, and Purification of the N-Terminal Domains</td>
</tr>
<tr>
<td>(WD 1-6)......................................................................................68</td>
</tr>
<tr>
<td>2.5.2 Cloning and Purification of WD 5-6, ATOX1 and WD 5-6(Y532H)...78</td>
</tr>
<tr>
<td>2.5.3 Tobacco Etch Virus (TEV) Protease Purification........................85</td>
</tr>
<tr>
<td>2.6 Other Proteins of Importance to this Work................................89</td>
</tr>
<tr>
<td>2.6.1 Cloning of COMMD1..................................................................89</td>
</tr>
<tr>
<td>2.6.2 Cloning of G591D, a Wilson Disease-Causing Mutation...............90</td>
</tr>
<tr>
<td>2.7 Biophysical Characterization and Analysis..................................91</td>
</tr>
<tr>
<td>2.7.1 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis.........91</td>
</tr>
<tr>
<td>2.7.2 Gel Filtration.........................................................................92</td>
</tr>
<tr>
<td>2.7.3 Light Scattering......................................................................93</td>
</tr>
<tr>
<td>2.7.4 Circular Dichroism...................................................................95</td>
</tr>
<tr>
<td>2.7.5 Nuclear Magnetic Resonance Spectroscopy...............................99</td>
</tr>
<tr>
<td>References....................................................................................102</td>
</tr>
<tr>
<td>3 CHARACTERIZATION OF THE N-TERMINAL DOMAINS OF THE</td>
</tr>
<tr>
<td>HUMAN WILSON PROTEIN.......................................................................106</td>
</tr>
<tr>
<td>3.1 Introduction..............................................................................106</td>
</tr>
<tr>
<td>3.2 Cloning of pET32Xa/wDI-6 and pET41Ek/wDI-6 Plasmids...............107</td>
</tr>
<tr>
<td>3.3 PCR Reactions and Results......................................................107</td>
</tr>
<tr>
<td>CHAPTER</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>4.3 Purification of WD5-6(Y532H) Mutant Protein</td>
</tr>
<tr>
<td>4.3.1 General Considerations and Results</td>
</tr>
<tr>
<td>4.4 Biophysical Characterization of WD5-6(Y532H) Protein</td>
</tr>
<tr>
<td>4.4.1 Gel Filtration and Light Scattering</td>
</tr>
<tr>
<td>4.4.2 Discussion</td>
</tr>
<tr>
<td>4.4.3 Circular Dichroism Spectroscopy</td>
</tr>
<tr>
<td>4.5 NMR of WD5-6(Y532H) from Refolding and Factor Xa Cleavage Method</td>
</tr>
<tr>
<td>4.5.1 The Structural Heterogeneity of WD5-6(Y532H) Mutant Protein</td>
</tr>
<tr>
<td>4.5.2 Inter-Domain Motion of WD5-6(Y532H)</td>
</tr>
<tr>
<td>4.5.3 Copper Transfer Studies of WD5-6(Y532H) with ATOX1 (HAH1)</td>
</tr>
<tr>
<td>4.6 Summary</td>
</tr>
<tr>
<td>References</td>
</tr>
<tr>
<td>5 CONCLUSION AND FUTURE DIRECTION</td>
</tr>
<tr>
<td>5.1 Significance of the Studies</td>
</tr>
<tr>
<td>5.2 Characterization of N-terminal Domains of the Human Wilson Protein</td>
</tr>
<tr>
<td>5.3 Conclusion and Future Work</td>
</tr>
<tr>
<td>APPENDIX</td>
</tr>
<tr>
<td>Project Approval Certification</td>
</tr>
</tbody>
</table>
LIST OF TABLES

1.1 The P-type ATPases subfamily ................................................................. 4
1.2 Reagents and chemicals ................................................................. 59
1.3 Enzymes, process kits, and purification aids ....................................................... 61
1.4 Notation for proteins studied ................................................................. 63
1.5 Primary primers .............................................................................. 64
1.6 Secondary primers ........................................................................ 65
1.7 Plasmid vectors ............................................................................. 66
1.8 Columns for protein purification and analysis (GE Life Sciences) ...................... 67
1.9 The PCR reaction of WD 1-6 gene for subsequent cloning into pET41EKLIC vector .................................................................................. 108
1.10 The PCR thermocycler settings of WD 1-6 gene for subsequent cloning into pET41EKLIC vector ................................................................ 108
1.11 The PCR reaction of WD1-6 gene for subsequent cloning into pET32XaLIC vector .................................................................................. 109
1.12 The PCR thermocycler settings of WD 1-6 gene for subsequent cloning into pET32XaLIC vector ................................................................ 109
1.13 The WD1-6 gene insert treatment for incorporation into pET41EKLIC vector .................................................................................. 113
1.14 The WD 1-6 gene insert treatment for pET32XaLIC vector cloning ...................... 113
1.15 LIC-compatible WD1-6 insert annealing into LIC vectors .................................. 114
1.16 Double digestion of pET32Xa/WD1-6 plasmid ........................................... 114
1.17 Cycling parameters of QuickChange II XL site-directed mutagenesis method ..................................................................................... 120
1.18 The synthesis of the pET32XaTEVII/WD1-6 plasmid from pET32XaTEVII/WD 1-6 template ........................................................................ 121
List of Figures—Continued

3.11 Notation for the expressed WD 1-6 fusion protein..............................126
3.12 The molecular weights of protein standards used to plot calibration curve........136
3.13 Molar mass of WD 1-6 by static light scattering.......................................142
3.14 Hydrodynamic radius of WD 1-6.................................................................142
3.15 The data set used to plot the linear extrapolation curve.................................145
3.16 Conformational stability of WD 1-6............................................................147
4.1 The PCR mutagenesis reaction of WD5-6(Y532H)........................................156
4.2 The cycling parameters of WD5-6(Y532H) mutagenesis................................156
4.3 The WD5-6(Y532H) gene amplification reaction with LIC-qualified primers...............................159
4.4 The PCR reaction conditions for WD5-6(Y532H) gene amplification with LIC primers.................................................................159
4.5 Calculated molecular weights (MW) of WD5-6 and WD5-6(Y532H) determined by SLS.................................................................174
4.6 Hydrodynamic radius of WD5-6 and WD5-6(Y532H) protein monomers........175
4.7 Helicity loss during denaturation.................................................................180
4.8 Y points along the WD5-6 unfolding curve (Figure 4.21) used to plot the linear extrapolation curve (Figure 4.22).........................................................182
4.9 Y points along the WD5-6(Y532H) unfolding curve (Figure 4.21) used to plot the linear extrapolation curve (Figure 4.22).........................................................183
4.10 Parameters characterizing the GnHCl denaturation of Wilson protein N-terminal domains.................................................................184
LIST OF FIGURES

1.1 The crystal structure of sarcoplasmic reticulum Ca\textsuperscript{2+}-ATPase at 2.6 Å resolution showing the nucleotide binding domain, phosphorylation domain, transmembrane domain and the actuator domain (PDB 1SUA) Toyoshima et al. [40]........................................................................................................6

1.2 The atomic structure of H\textsuperscript{+}-ATPase dimer from Arabidopsis thaliana plasma membrane without its auto-inhibitory C-terminus determined by X-ray crystallography at 3.6 Å........................................................................................................6

1.3 The illustration of the structural organization of Wilson protein/Menkes protein.................................................................................................................................................................11

1.4 The solution structure of nucleotide binding domain (PDB 2ARF). Gly1 101, Glul064, Gly 1099, and His 1069 are involved in ATP binding........................................................................................................12

1.5 The structure of human copper transporter hCTR1 determined by cryo-electron microscopy depicts a cone-shaped pore that forms an entry for copper ions into the cytosol from extracellular matrix.................................................................14

1.6 Solution structure of Cu(I)-ATX1. The ATX1 is a copper chaperone for Ccc2-ATPase in yeast..................................................................................................................................................16

1.7 The transformation of a superoxide radical anion (• 0~2) to the hyperactive hydroxyl radical (• OH) via a copper catalyzed reaction........................................................................................................20

1.8 The intracellular copper trafficking........................................................................................................21

1.9 Kayser-Fleischer ring in the cornea Descemet’s membrane......................................................................30

1.10 The copper chelating agents..............................................................................................................34

2.1 Purification columns for AKTA FPLC system (GE Life Sciences).....................................................67

2.2 Bioflo®310 fermentor/bioreactor (New Brunswick Scientific)............................................................68

2.3 The amino acid sequence of WD1-6......................................................................................................70

2.4 Schematic of steps from gene amplification to plasmid DNA sequencing........................................72

2.5 Ligation independent cloning strategy into pET41EkLIC vector.......................................................73
List of Figures—Continued

2.6 The expression and purification summary of WD1-6 protein using pET32Xa/WD1-6 plasmid ................................................................. 74

2.7 The TEV cleavage site's insertion into the pET32Xa/WD1-6 plasmid ............... 79

2.8 Recovery of WD5-6(Y532H) protein using two different routes ..................... 85

2.9 Time-course induction of TEV protease with a molecular weight of 29 kDa indicated by the arrow ........................................................................ 88

2.10 The TEV protease amino acid sequence ......................................................... 88

3.1 A 0.8% agarose gel picture of the amplified WD1-6 gene ............................. 110

3.2 pET32XaLIC vector cloning region (Novagen, Inc.) ................................. 112

3.3 pET41 EklIC vector cloning region (Novagen, Inc.) ................................. 112

3.4 The restriction enzyme mapping of pET32Xa/WD1-6 visualized in a 0.8% agarose gel ................................................................................. 115

3.5 Thioredoxin tagged WD1-6 protein and the location of the protease cleavage sites (A) Factor Xa (B) TEV (Tobacco Etch Virus) ............................................. 119

3.6 Double digest (restriction enzyme mapping) of plasmid constructs using BamHI and Bgl II, ......................................................................................... 122

3.7 The truncated TRX\textsubscript{Xa}WD1-6 protein (-35 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WD1-6 ................................................. 124

3.8 The truncated TRX\textsubscript{Xa}WD1-6 protein (-20 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WD1-6 ................................................. 124

3.9 The truncated TRX\textsubscript{Xa}WD1-6 protein (-40 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WD1-6 ................................................. 125

3.10 Expression of GST\textsubscript{Ek}WD1-6 protein (-92 kDa) in Rosetta(DE3) cells transformed with pET41 Ekl/WD1-6 ...................................................... 125

3.11 The 10% Tricine SDS-PAGE gel showing TRX\textsubscript{Xa}WD1-6 induction after mutation repairs ............................................................... 126
List of Figures—Continued

3.12 10% Tricine SDS-PAGE gel showing the TRX\textsubscript{x,TEvWD1-6} and TRX\textsubscript{JEvWD1-6} time-course induction........................................127

3.13 TRX\textsubscript{xaTEvWDl-6} protein elution from a 20 mL HisPrep column..............129

3.14 Size exclusion chromatography of TRX\textsubscript{xaTEvWDl-6}........................................130

3.15 Time-course cleavage of thioredoxin from TRX\textsubscript{xaTEvWDl-6} with TEV protease (1fxg TEV/100 jig fusion protein)..................................................131

3.16 Size exclusion chromatography of WD 1-6..........................................................132

3.17 Gel filtration calibration curve.................................................................137

3.18 The molar mass distribution of WD 1-6 determined by SLS.................................139

3.19 The autocorrelation function of WD1-6......................................................140

3.20 A plot of hydrodynamic radius against elution time in min (gel filtration) is shown for WD1-6.................................................................141

3.21 Unfolding of WD1-6 under increasing ionic strength of guanidine hydrochloride.................................................................143

3.22 WD 1-6 unfolding monitored at 222 nm with increasing molar concentration of GnHCl.................................................................144

3.23 WD 1-6 denaturation curve analysis using the linear extrapolation method (LEM).................................................................146

4.1 The NMR structure of WD5-6..................................................................155

4.2 The WD5-6(Y532H) PCR product.................................................................157

4.3 The amplified WD5-6(Y532H) gene for insertion into pET32XaLIC vector........160

4.4 Restriction mapping of pET32Xa/WD5-6(Y532H) construct.............................160

4.5 Expression and purification of non-fusion WD5-6(Y532H)...............................163

4.6 The purification of non-fusion WD5-6(Y532H) from optimized conditions........164

4.7 The refolded TRX\textsubscript{xaWD5-6(Y532H)} purification.......................................164
List of Figures—Continued

4.8 Thioredoxin tag cleavage from TRXxWD5-6(Y532H).................................165
4.9 Gel filtration of WD5-6(Y532H) recovered under optimized conditions..........167
4.10 Gel filtration of WD5-6(Y532H) obtained by refolding/Factor Xa cleavage method.............................................................................................................168
4.11 The MALDI-TOF of '*C, *N labeled WD5-6(Y532H), -16, 896 Daltons, obtained after refolding and thioredoxin cleavage; Department of Pharmaceutical Sciences, University of Florence, Italy........................................169
4.12 High resolution gel filtration of WD5-6(Y532H)........................................172
4.13 High resolution gel filtration of WD5-6.....................................................172
4.14 Molar mass distribution of WD5-6 and WD5-6(Y532H) by static light scattering..........................................................................................................................173
4.15 Autocorrelation function of WD5-6 monomer peak.....................................174
4.16 Autocorrelation function of WD5-6(Y532H) monomer peak..........................175
4.17 The R_h distribution of WD5-6 and WD5-6(Y532H) proteins..........................176
4.18 The unfolding of 29 j^M WD5-6(Y532H) under increasing molar concentration of GnHCl..............................................................179
4.19 The unfolding of 14.5 (iM WD5-6 under increasing molar concentration of GnHCl...............................................................................179
4.20 The Circular dichroism of 29^iM WD5-6 and WD5-6(Y532H) in the absence of denaturant.................................................................180
4.21 The loss of helicity during denaturation........................................................180
4.22 Ellipticity of WD5-6(Y532H) and WD5-6 monitored at 222 nm......................181
4.23 The Gibbs free energy of unfolding estimated from a liner extrapolation plot for WD5-6(Y532H) and WD5-6.........................................................184
4.24 The heteronuclear single quantum correction (HSQC) chemical shift of WD5-6(Y532H) overlaid with WD5-6.........................................................188
4.25 Chemical shift variation between wild type and Y48H-WD5-6(Y532H)...........188
List of Figures—Continued

4.26 Combined chemical shift of WD5-6(Y532H)........................................189

4.27 Sequential plot of $^{15}$N relaxation parameters for WD5-6 (black) and WD5-6(Y532H in WNDP) (red).................................................................191

4.28 The H-H bonding proximity between Y/H48 (Y/H532 in WNDP) and Y75 (Y559 in WNDP)..............................................................................192

4.29 (Left) sequential plot of CSV between apo and Cu(I)-WD5-6(Y532H) (Right) relaxation parameters of WD5-6(Y532H) at the conclusion of titration.................................................................192
LIST OF ABBREVIATIONS

A- ............................................................... Actuator domain
ADH ........................................................... Alcohol Dehydrogenase
ATOX1 .................................................... Human Anti-oxidant 1 Copper Chaperone Protein
ATP7A ...................................................... Human ATPase, Cu++ transporting, alpha polypeptide
ATP7B ...................................................... Human ATPase, Cu++ transporting, beta polypeptide
ATX1 ........................................................ ATX1 antioxidant protein 1 homolog (yeast)
BAL ............................................................... Dimercaprol
BSA ............................................................. Bovine Serum Albumin
CAD ............................................................. Carbonic Anhydrase
Cyt C ............................................................. Cytochrome C
DLS .............................................................. Dynamic Light Scattering
DMSO .......................................................... Dimethylsulfoxide
D-PCA .......................................................... D-Penicillamine
CDPP .............................................................[c/s-diamminedichloroplatinum (II)]
EDTA ............................................................ ethylenediaminetetraacetic acid
K-F ................................................................. Kayser-Fleischer rings
GnHCl .......................................................... Guanidine hydrochloride
HAH1 ........................................................... Copper chaperone, also known as ATOX1
HSQC ............................................................ Heteronuclear single quantum correlation
IMS .............................................................. Inter Membrane Space
IPTG ............................................................. Isopropyl p-D-l-thiogalactopyranoside
List of Abbreviations—Continued

LB ........................................................................................................ Luria Bertani Broth
LEM ................................................................................................. Linear Extrapolation Method
MBD ...................................................................................................... Metal Binding Domain
MBS ....................................................................................................... Metal Binding Site
MES ...................................................................................................... 2-(iV-morpholino)ethanesulfonic acid
MNKP .................................................................................................. Menkes Protein
N- ........................................................................................................... Nucleotide binding domain
OD ........................................................................................................ Optical Density
P- ........................................................................................................... Phosphorylation domain
PDA ....................................................................................................... Photodiode Array
QELS ................................................................................................. Quasi Elastic Light Scattering
RI .......................................................................................................... Refractive Index
SEC .................................................................................................... Size Exclusion Chromatography
SLS ...................................................................................................... Static Light Scattering
SOC ..................................................................................................... Super Optimal Broth
TGN, ..................................................................................................... trans-Golgi
TM ....................................................................................................... Transmembrane
WD ...................................................................................................... Wilson disease
WD4 ....................................................................................................... N-terminal domain 4 of Wilson protein
WD5 ....................................................................................................... N-terminal domain 5 of Wilson protein
WD5-6 .................................................................................................. N-terminal domains 5 and 6 of Wilson protein
WD6 ....................................................................................................... N-terminal domain 6 of Wilson protein
WD5-6(Y532H) ..................................................................................... Y532H mutant protein
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>WNDP</td>
<td>Wilson Protein</td>
</tr>
</tbody>
</table>
1.1 Introduction

Wilson protein is a copper transporting P-type ATPase. The importance of copper metal has been known for centuries although its role in biological systems did not receive much attention until a few decades ago. It is an essential cofactor for many enzymes, but, through its potential to drive oxidation/reduction reactions and form free radicals, excess copper can lead to cell toxicity and death [1]. As a result, bacterial and animal cells have developed mechanisms to control this delicate balance of intracellular copper concentration [2]. P-type ATPases are ATP-dependent, transmembrane pumps which belong to one of the protein families involved in copper uptake and regulation. Several distinct P-type ATPase subgroups designated Pi-Pv are known with type II and type IB being dominant [3]. The PiB type is not only involved in copper trafficking, but also in transport of other transition metal ions such as Zn$^{2+}$, Cd$^{2+}$, Co$^{2+}$, and Pb$^{2+}$ [4].

Scientists have identified and characterized two important PiB copper transporting ATPases in humans, ATP7A and ATP7B. Defects in ATP7A and ATP7B genes, give rise to Menkes syndrome and Wilson disease, respectively [5-8]. If not treated, these disorders can lead to liver cirrhosis, neuropathy, coma, and eventually death. To function effectively, these copper ATPases are believed to interact with other proteins such as copper chaperones, metallothioneins, and cellular target proteins [9]. The mechanisms by which these proteins coordinate copper homeostasis are not well understood.
In the following sections, I will examine recent advances in structural and functional studies of ATP7B (Wilson protein) and ATP7A (Menkes protein). Although this work is largely based on the human Wilson protein, it would not be complete without discussing other homologous proteins that are involved in copper homeostasis in plants, bacteria, and yeast. Later, I will discuss the structural and functional importance of copper chaperones and related proteins that interact with the two copper ATPases. Copper trafficking and metabolism in relation to copper ATPases, chaperones, cupro-enzymes, storage and excretion will be also examined in detail. Finally, I will address the pathogenesis, genetics, diagnostic and therapy of Wilson's disease and associated disorders. A brief summary of the reviewed work constitutes the last section, which paves way into the subject of this work — the characterization of the N-terminal domains and disease-causing mutations of the human Wilson protein.

1.2 Wilson and Menkes ATPases: Homologous Copper ATPases

1.2.1 Structural Properties of Wilson and Menkes Protein

The discovery in early 1990's of genes encoding human copper transporting ATPases [5-7] spurred unprecedented efforts by scientists to study their biological significance and function. It is also important to note that pioneering work on the ion ATPase pump was laid down almost fifty years ago, and it is worthy to take a few steps back and survey the founding work of the P-type ATPases. Nobel Laureate Jens Skou was the first scientist to discover the membrane potential generated by various cations in leg nerve homogenates from crabs in 1957 [10]. He suggested that the membrane...
potential was generated by a K\textsuperscript{+}-stimulated ATPase, which later was identified as the Na\textsuperscript{+}/K\textsuperscript{+}-ATPase which utilizes ATP hydrolysis as a source of energy to pump these ions across the axonal membrane [11]. Several other related ions pumps have been characterized since this novel discovery. They include sarcoplasmic-recticulum (SR) Ca\textsuperscript{2+}-ATPase [12] which help to control skeletal muscle contraction, the gastric H\textsuperscript{+}/K\textsuperscript{+}-ATPase that controls stomach acidity, and the H\textsuperscript{+}-ATPase that generates membrane potential in fungal and plant cells [13].

The P-type ATPase superfamily genes have been classified into ten subfamilies as shown in Table 1.1, of which six are associated with plants [14]. These six are: proton pumps (PHIA subfamily), heavy metal pumps (PIB subfamily), ER-type Ca\textsuperscript{2+}-ATPase (PHA family), auto-inhibited Ca\textsuperscript{2+}-ATPase (PnB subfamily), putative amino-phospholipid ATPase (Piv subfamily), and P-type ATPase type 5 (P\textsubscript{v}) [15]. P-type ATPases clusters, PHA (KdpB protein), P\textsubscript{NC} (Na\textsuperscript{+}/K\textsuperscript{+}-ATPases), P\textsubscript{ND} (Na\textsuperscript{+}-ATPases) and PmB (Mg\textsuperscript{2+}-ATPases) are not found in plants but only in the other eukaryotic family members. An in depth discussion of structures and functions of these sub-families is beyond the scope of this thesis. I will describe the structural organization of the best studied members of the group to enable better understanding of my work. The P-type ATPases under consideration can be divided into three main categories according to their function: (1) cell signaling, Ca\textsuperscript{2+}-ATPase; (2) electrochemical potential generating, H\textsuperscript{+}-ATPase in plants and Na\textsuperscript{+}/K\textsuperscript{+}-ATPases in animals; and (3) the essential micronutrient transporting ATPases, the Cu\textsuperscript{2+}, Zn\textsuperscript{2+}, and Co\textsuperscript{2+}-ATPases.

Generally, P-type ATPases are multi-domain membrane proteins with molecular masses in the range of 50-150 kDa. Sequence comparisons with alignment tools reveal a
characteristic pattern of conserved residues, particularly the DKTGTLT sequence motif in which D is a transiently phosphorylated aspartate. All are integral membrane proteins possessing 10 hydrophobic membrane-spanning helices (TM1-TM10), though some have only six or eight of these helices. They possess highly conserved cytoplasmic domains that link helices TM2, TM3, TM4, and TM5 [10]. P-type ATPases are structurally divided into four principal functional units according to their role and position. These units are the phosphorylation domain (P-domain), the nucleotide domain (N-domain), the actuator (A-domain) and the membrane domain (TM-domain).

Table 1.1 The P-type ATPases subfamily.

<table>
<thead>
<tr>
<th>SUBFAMILY</th>
<th>EXAMPLE</th>
<th>REFERENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>PLA</td>
<td>KdpB proteins</td>
</tr>
<tr>
<td>2</td>
<td>PLB</td>
<td>Heavy metal pumps</td>
</tr>
<tr>
<td>3</td>
<td>PLIA</td>
<td>ER-type Ca(^{2+})-ATPase</td>
</tr>
<tr>
<td>4</td>
<td>PLIB</td>
<td>Auto-inhibited Ca(^{2+})-ATPase</td>
</tr>
<tr>
<td>5</td>
<td>PLIC</td>
<td>Na(^+)/K(^+)-ATPase</td>
</tr>
<tr>
<td>6</td>
<td>PLID</td>
<td>Na(^+)-ATPases</td>
</tr>
<tr>
<td>7</td>
<td>PLIA</td>
<td>H(^+)-ATPase</td>
</tr>
<tr>
<td>8</td>
<td>PLIB</td>
<td>Mg(^{2+})-ATPases</td>
</tr>
<tr>
<td>9</td>
<td>Piv</td>
<td>Putative amino-phospholipid ATPase</td>
</tr>
<tr>
<td>10</td>
<td>Pv</td>
<td>P-type ATPase type 5</td>
</tr>
</tbody>
</table>

The Ca\(^{2+}\)-ATPase of skeletal muscle sarcoplasmic reticulum (SERCA, type 1a) is an ATP-powered Ca\(^{2+}\)-pump which is well characterized [25]. It is an integral membrane protein with molecular mass of 110 kDa, which is found in both prokaryotes and
eukaryotes. The SERCA gene family contains many SERCA pump isoforms generated by alternative RNA splicing [26-30]. Expression of these isoforms is tissue dependent. More than 20 crystal structures of SERCA Ca\(^{2+}\)-ATPase exist [31], and commonalities among these structures (Figure 1.1) include the presence of 10 transmembranes helices (TM1-TM10), three cytoplasmic domains, an A-domain (actuator domain), a P-domain (phosphorylation domain), and an N-domain (nucleotide domain) where ATP binds during catalysis [32-34]. Mutagenesis of residues near the center of transmembrane domain spanning TM4, TM5, TM6, and TM8 disable Ca\(^{2+}\) transport; therefore, it is suspected to be the calcium binding site [35]. In the SERCAla structure, the A-domain is connected to the TM1-TM3 helices with unusually long linkers, regulating Ca\(^{2+}\) binding and release by acting as an actuator for the transmembrane gating mechanism [36]. The A-domain contains a TGES sequence motif which has a role in aspartylphosphate processing. The SERCAla also possesses two unique transmembrane Ca\(^{2+}\)-binding sites which manifest cooperativity [37]; the two sites, denoted I and II, are located side by side, but Ca\(^{2+}\) binding is sequential [38].

Unlike animals, plants and fungi possess a proton pumping ATPase (H\(^{+}\)-ATPase) which couples ATP hydrolysis to proton transport out of the cell forming an electrochemical gradient across the plasma membrane. In animals, the membrane potential is established in large part by the Na\(^{+}\)/K\(^{+}\)-ATPase. The plant H\(^{+}\)-ATPase is an approximately 100 kDa monomer which can oligomerize to form dimeric and hexameric complexes [14]. A low resolution structure (8 Å) of the *Arabidopsis thaliana* plasma membrane H\(^{+}\)-ATPase isoform 2 (AHA2) determined by electron cryo-microscopy revealed a dimeric protein complex [39].
Figure 1.1 The crystal structure of sarcoplasmic reticulum Ca\(^{2+}\)-ATPase at 2.6 Å resolution showing the nucleotide binding domain, phosphorylation domain, transmembrane domain and the actuator domain (PBD 1SUA) Toyoshima et al. [40].

Figure 1.2 The atomic structure of H\(^{+}\)-ATPase dimer from Arabidopsis thaliana plasma membrane without its auto-inhibitory C-terminus determined by X-ray crystallography at 3.6 Å. The four main domains, transmembrane domain, nucleotide binding domain, phosphorylation domain, and the actuator domain are shown (PBD 3B8C) Pedersen et al. [41].
The \( \text{H}^+ \)-ATPase has N- and C-terminal segments which protrude into the cytoplasm and ten transmembrane segments (TM1-TM10) in addition to A-, P-, N-, and R- domains. Like \( \text{Ca}^{2+} \)-ATPase, the P-domain contains the catalytic phosphorylated aspartate within the canonical sequence DKTGTLT. The R-domain acts as an auto-inhibitory domain and forms the C-terminal part of the enzyme. The atomic structure of \( \text{H}^+ \)-ATPase without its auto-inhibitory C-terminus determined by X-ray crystallography at 3.6 Å [41] revealed four main domains: a transmembrane domain, a nucleotide binding domain, a phosphorylation domain, and an actuator domain (Figure 1.2).

The \( \text{Na}^+ / \text{K}^+ \)-ATPase was the first P-type pump to be described [11], and like the other P-type pumps, it is a membrane-bound ion pump. It generates a \( \text{Na}^+ \) and \( \text{K}^+ \) ion concentration gradient across the plasma membrane in animals by concurrent pumping of \( \text{Na}^+ \) out of the cytoplasm and \( \text{K}^+ \) into the cytosol [42, 43]. Two high resolution crystal structures of the sodium-potassium pump were recently reported [44, 45]. The pig renal \( \text{Na}^+ / \text{K}^+ \)-ATPase crystal structure at 3.5 Å resolution reveals a heteromultimeric P-type ATPase consisting of a- and \( \beta \)-subunits. The binding sites for \( \text{Na}^+ \), \( \text{K}^+ \), and ATP are located in the a-subunit that shows homology to a single subunit P-type ATPases such as \( \text{Ca}^{2+} \)-ATPase. The a-subunit possesses the three cytoplasmic domains—the P-domain, the N-domain, and A-domain in addition to transmembrane domains. The p-subunit is found in symport P-type ATPases. It routes the a-subunit to the plasma membrane and mediates \( \text{K}^+ \) ion entry [46]. The third subunit (\( \gamma \)-subunit in kidney outer medulla) associates with the aP-complex and regulates the pumping activity. This protein belongs to the FXYD family, a cluster of proteins expressed in tissues that specialize in fluid or solute transport [47]. The shark rectal gland \( \text{Na}^+ / \text{K}^+ \)-ATPase [45] visualized at 2.4 Å
resolution depicts features similar to the porcine renal pump; the extracellular parts of the P-subunit and FXYD protein are well characterized in the shark rectal gland pump. The Wilson protein (ATP7B) and Menkes protein (ATP7A) found in humans are members of the heavy metal transporting P-type ATPases (PIB subfamily) and they are commonly referred to as the copper transporting ATPases (Cu-ATPases) [3]. ATP7A and ATP7B are related to copper transporting ATPases found in other organisms such as the Ccc2p pump in yeast [48, 49], the CopA copper ATPase in Enterococcus hirae [50], OsHMA9 in Oryza sativa rice [51], and CUA-1 in C. elegans [52]. Unlike the other P-type ATPases that have been described, the heavy metal P-type ATPases were discovered relatively recently (1989) in studies involving cadmium resistant bacteria [53]. Structurally, Cu-ATPases are yet to be fully characterized and efforts from biophysical and theoretical chemists continue.

The Wilson protein gene ATP7B is mapped to chromosome 13ql4.3 and encodes a protein 1465 amino acid residues long [5-7]. Its complete exon-intron structure was determined [54] prior to discovery of the gene encoding ATP7A [55, 56]. The ATP7A gene is mapped to chromosome Xq13.3, and encodes a 160 kDa protein of 1500 amino acid residues [57-59]. Both genes are fairly large with ATP7A spanning approximately 150 kilobase pairs of genomic DNA and containing 23 exons. The start codon, ATG is located in the second exon. The average size of each ATP7A exon varies from 77 to 726 bp. ATP7B gene structure is very similar to ATP7A; it is made up of 21 exons ranging in size from 77 to 1234 bp; the start codon is located in exon number one. All splice sites starting from exon five in ATP7A and exon three in ATP7B occur in identical nucleotide positions except for the boundary between exon seventeen and eighteen in ATP7A (exon
fifteen and sixteen in ATP7B). These two genes are distinct from each other when observed in the 5' region that encode for amino metal binding domains. The first four domains of the six metal binding domains (MBDs) in ATP7B are encoded in one large exon (exon two), while the sequence encoding similar domains in ATP7A is spread over three exons [60].

The copper transporting ATPases ATP7A and ATP7B share general features with non-heavy metal P-type ATPases except the former have additional N-terminal domains. The five distinct domains that are revealed in ATP7A and ATP7B molecular organization are illustrated in Figure 1.3: (1) the transmembrane domains consisting of eight transmembrane domains (TM); (2) the phosphorylation domain (P-domain); (3) the nucleotide binding domain (N-domain); (4) the actuator domain (A-domain); and (5) the cytosolic N-terminal metal binding domains. The complete structure of these Cu-ATPases is unknown, but many structures of isolated domains are available [61]. The transmembrane domains, responsible for copper transport across the membrane, contains conserved residues Asn and Tyr in seventh TM [TM7], Met and Ser in eighth TM [TM8], and the CPC and YN sequence motifs in TM6 and TM7, respectively. The functional importance of these conserved residues has been confirmed by mutational analysis [62-64]. Mutations of the conserved Ser in TM 8 of the WNDP have been found in some Wilson disease patients [65].

The large cytoplasmic portion between transmembrane domain TM6 and TM7 is referred to as the adenosine triphosphate binding domain (ATP-BD) and is made up of the nucleotide binding (N) and the phosphorylation (P) domains, and contains an invariant DKTGT sequence motif that is transiently phosphorylated at the aspartate residue during
the ATP hydrolysis cycle [66]. Putative nucleotide and magnesium binding motifs, TGDN and GDGXND, are also located within the N and P domains. The actuator (A) domain lies between transmembrane domains TM4 and TM5 and possesses the invariant sequence motif, TGE. The NMR structure of the nucleotide binding domain of WNDP has been reported [67]; it is 17 kDa, and can bind ATP and analogs. The structure shows a central anti-parallel six-stranded P-sheet flanked by two a-helical hairpins (Figure 1.4). The structure is closely packed except for a flexible 29 amino acid loop located at the junction of the p-sheet and one of the a-helical motifs. The N-domain has little sequence homology (<25%) with the SERCA 1 N-domain, the Ca\(^{2+}\)-ATPase, but their shared basic structure is evident. Nucleotide binding residues (His1069, Glul064, Glyl099, and Gly1101) are present in both domains. Mutation of His1069 to glutamine is the most common Wilson disease mutation; it probably disrupts ATP binding [68]. While the cytosolic N-terminal domains of MNKP and WNDP protein possess six metal binding sites, similar Cu-ATPases in prokaryote and archea have one or two. Non-heavy metal transporting ATPases possess none.

The N-terminus of WNDP is 650 amino acids long and possesses six distinct metal binding domains, each bearing a conserved repetitive sequence motif GMT/HCxxC. The metal binding domains (MBDs) are roughly 60-70 residues in size [69-71]. In addition to the six domains, the N-terminus of WNDP has a 63 amino acid (cytosolic tail) extension located in the first domain which was found to be necessary for trafficking in polarized hepatocytes [72]. Several high resolution structures of some of these domains have been reported, both in their apo and metal bound forms [73-77]. Metal domains share a common structural characteristic, a pappap ferredoxin-like fold, a
**KEY:**

N — Nucleotide binding domain  
P — Phosphorylation domain  
A — The Actuator domain  
TM — Transmembrane domains

**SEQUENCE MOTIFS:**

CPC/YN/SEHPL — Sequence motifs characteristic of Pm-type ATPases  
TGDN — Nucleotide binding motif  
GDG — Magnesium metal binding motif  
CXXC — Copper metal binding motif

**Figure 1.3** The illustration of the structural organization of Wilson protein/Menkes protein. The transmembrane domains are embedded in plasma membrane and form a transduction channel that allows the passage of copper ions into the vesicle.

11
Figure 1.4 The solution structure of nucleotide binding domain (PDB 2ARF). Gly1101, Glu1064, Gly 1099, and His1069 are involved in ATP binding. Mutation of His 1069 causes Wilson disease, Dmitriev et al. [67].

feature also found in the human copper chaperone ATOX1 that interacts with the metal binding domains [78]. The inter-domain linkers among the metal binding domains vary in length. The longest is between the Wilson metal binding domain 4 (WD4) and Wilson metal binding domain 5 (WD5) which is 57 amino acids long, and the shortest, only 8 amino acids long is between domain WD5 and WD6. The solution structure of apo WD5-6 reveals that this protein fragment is organized into two domains and tumbles as a unit, and the CxxC metal binding motifs faces away from each other [73]. In general, all reported structures show homology and similarities in features for both MNKP and WNDP but present different surface properties [79].
1.2.2 Copper Permeases and Metallochaperones

Copper transporting ATPases in eukaryotic cells interact or associate with other proteins and organelles in the cell to orchestrate copper trafficking. These proteins include copper permeases that facilitate copper ion entry into cells and metallochaperones that deliver the acquired copper either to the Cu-transporting ATPases or to cuproenzymes.

1.2.2.1 Copper Permease: The Human Copper Transporter (hCTR1)

The hCTR1 is a 23 kDa, 190 amino acid residue long integral membrane protein. The structure, illustrated in Figure 1.5, was solved by cyro-electron microscopy at a resolution of 7 Å [80]. The CTR proteins are involved in cellular copper uptake and have been found in many different organisms [81]. Like the other characterized CTR-proteins such as yeast CTR1 [82], human CTR1 is structurally organized into four main domains: (1) an extracellular N-terminal domain [83]; (2) an intracellular loop connecting the first and second putative membrane spanning helices; (3) a membrane embedded domain that consists of 3 transmembrane helices; and (4) an intracellular C-terminal domain. hCTR1 is trimeric and contains a cone-shaped pore (~8 Å at the extracellular end, and ~22 Å at the intracellular side) along the central 3-fold axis which is believed to be the copper ion passage channel. Functionally important methionine residues in the MxxxM motif [84-86] are found in the transmembrane helix (TM2) and are strategically positioned near the extracellular end of the pore suggesting a role in copper uptake. The TM2 and TM3 residues provide the lining of the pore near the intracellular end while TM2 also provides membrane or hydrophobic interactions at the extracellular end.
Figure 1.5 The structure of human copper transporter hCTR1 determined by cryo-electron microscopy depicts a cone-shaped pore that forms an entry for copper ions into the cytosol from extracellular matrix. The MxxxxM motif in the transmembrane TM2 is near the extracellular end of the channel for copper uptake, De Feo et al. [80].
1.2.2.2 The Human Copper Chaperone ATOX1

The yeast ATX1 metallochaperone was first discovered as an anti-oxidant (ATX1) gene when overexpressed in oxidatively distressed yeast cells [87]. Later, its main function was found to be shuttling of copper to the Cu-ATPase Ccc2 [49, 88]. The human homologue of yeast ATX1, ATOX1 (also known as HAH1), is a 68 amino acid protein that shares domain homology with the N-terminal metal binding domains of WNDP and MNKP. ATOX1 exhibits a ferrodoxin-like fold and has a single conserved MTCGGC copper binding motif [70, 79]. ATOX1 binds Cu(I) and delivers it to N-terminal domains of the MNKP and WNDP which are capable of binding up to six copper ions or one copper(I) ion per MT/HCxxC motif [89, 90]. The crystal structures of Cu(I), Hg(II), and Cd(II) bound human ATOX1 have been published [91]. The NMR structure for apo and Cu(I) bound ATOX1 has also been reported [92]. All protein structures show that ATOX1 exhibits a PaPPaP fold (Figure 1.6 for yeast ATX1). The protein is homodimeric, with two monomers linked together by coordinated metal ions; Cd(II) is tetrahedral, Hg(II) assumes distorted tetrahedral geometry and Cu(I) is depicted as a pseudo tetrahedral because the fourth Cu-S bond, which is 2.4 Å is not within the range of primary bonding range of Cu(I)-S, Cu(I)-N, and Cu(I)-O complexes [93, 94].

1.2.2.3 Copper Chaperone for Superoxidase Dismutase (CCS) and COX-17

The CCS copper chaperone donates Cu within the inner mitochondrial space (IMS) to Cu,Zn-SOD, an enzyme that requires copper and zinc as co-factors [95]. Cytosolic CCS does not deliver copper into mitochondria. Instead, it is imported as a polypeptide into IMS (Inter Membrane Space) via Mia40/ERV1 pathway before oxidative refolding where
it is involved in activation of Cu,Zn-SOD by metallation using a labile copper pool within the mitochondrial matrix compartment. The labile copper pool consists of a low mass ligand complex (CuL). Binding of the copper to the ligand in the cytosol triggers translocation of the anionic complex into the mitochondria through membrane channels. There is speculation that a transporter exists that moves the CuL complex into the matrix and is also capable of shuttling it back to the IMS for subsequent use by COX-17 and SCO1 [97].

The CCS structure is organized into three functionally distinct domains: I, II, and III. Knowledge of the human CCS structure is largely derived from sequence alignments with yeast CCS, although its domain II structure has been determined [98].

Figure 1.6 Solution structure of Cu(I)-ATX1. The ATX1 is a copper chaperone for Ccc2-ATPase in yeast. The copper ion is shown as a brown sphere (PBD 1FD8) Arnesano et al. [96],
The domain I which includes the N-terminal end bears homology to human ATOX1 and includes a MTCQSC copper binding site. Domain III forms the C-terminal region, the smallest of the three domains, is approximately 30 amino acids in length and was shown to activate Cu,Zn-SOD in vivo [99]. Domain II is homologous to the target protein, Cu,Zn-SOD, and possesses a zinc binding site. COX-17 [100] is an 8.0 kDa protein that contains 6 conserved cysteine residues in both yeast and humans and functions as a copper chaperone in the inner mitochondrial space for Scol and COX-11, inner mitochondrial membrane anchored proteins. Scol and COX-11 ultimately transfer copper to COX-2 and COX-1, respectively [101]. COX-17 is also imported into IMS from cytosol in a similar fashion to CCS. In humans, Scol works with a second Sco protein, Sco2 to transfer copper into COX-2 to form the binuclear CUA site, while COX-11 copper transfer into COX-1 leads to assembly of the CUB site. Incorporation of copper in yeast cytochrome c oxidase is supported by the presence of SCOl and its homologue, SC02, and they are suspected to activate cytochrome c oxidase [102, 103].

1.2.3 Serum Ceruloplasmin

Ceruloplasmin, also called iron ferroxidase, is a blue a2-globulin oxidase that is mainly synthesized in hepatocytes and released into the circulation after copper insertion and glycosylation as holoceruplasmin [104-106]. Although copper is required for ceruloplasmin to function, its specific roles in copper trafficking and metabolism are relatively unknown [107]. Failure to incorporate copper during synthesis leads to production of apoceruloplasmin [104], which has a very short half-life of 5 h [108] compared to the holo form's half-life of 5.5 days [109]. In Wilson Disease, the defective
ATP7B protein fails to incorporate copper into apoceruloplasmin. Six copper ions are tightly bound to newly secreted ceruloplasmin. Three of these are found in mononuclear centers where they capture electrons from reductants, while the other three form the trinuclear cluster where oxygen is reduced to water [110, 111]. Ceruloplasmin intertwines iron and copper metabolism in that oxidation of Fe$^{2+}$ to Fe$^{3+}$ is catalyzed by ceruloplasmin. Fe$^{2+}$ is the source of reducing equivalents for the reduction of oxygen at the trinuclear cluster [112].

### 1.2.4 Functional Properties of Wilson and Menkes Protein

The functions of human Cu-ATPases have been largely delineated by studying the function and properties of related copper P-type ATPases in a range of organisms. As was discussed in preceding sections, the structural information of human Cu-ATPases in their entirety is currently unknown, and thus, understanding of the role and importance of each conserved amino acid residue in these proteins is not given. This has led to efforts to characterize and elucidate structure-function relationships of these Cu-ATPases. The review and discussion below links the existing pool on knowledge of the functional properties of Cu-ATPases and the roles they play in copper homeostasis and serves as an opening to the characterization of the N-terminal domains of the Wilson protein.

#### 1.2.4.1 Importance and Chemistry of Copper

Copper (Cu), is an essential trace element found in various species belonging to both plant and animal kingdoms [1]. It possess unique redox properties that have been exploited by a number of enzymatic and structural proteins [113]. Copper displays four
oxidation states: metallic Cu, cuprous ion Cu(I), cupric Cu(II), and a rare trivalent Cu ion Cu(III). Free Cu(I) does not exist in the aqueous environment of animal cells and any Cu(I) present is oxidized quickly by oxidizing agents unless it is protected by complexes formed within the cell. Copper is distributed by copper chaperones to cellular targets that require copper for their enzymatic activity and is occasionally stored by metallothioneins, a small molecular weight polypeptide. Copper drives many enzymatic reactions as an electron transfer intermediate in cellular respiration [114], free radical defense [115-117], cellular iron metabolism [118], and pigmentation [119, 120]. Copper also functions in the structural stability and architectural organization of many proteins such as the copper-zinc superoxide dismutase (Cu,Zn-SOD) [121]. In accordance with hard-soft acid-base theory, Cu(I) is a soft acid and prefers sulfur Met/Cys as proteinaceous coordination ligands while Cu(II) (borderline) preferentially coordinates with oxygen in Glu/Asp or the imidazole nitrogen in His.

While copper is essential for the well-being of growing cells and their development, excess copper is toxic and lethal. It is therefore important that the amount of copper in a cell at any particular moment does not exceed or fall below its demand. The capability of copper in redox (Fenton reaction) chemistry (Figure 1.7) to catalyze transformation of a superoxide radical anion •O'1 to the hyperactive hydroxyl radical •OH via a metal catalyzed reaction presents the potential damage that excess copper can cause [122]. These radicals oxidize and damage plasma membrane lipids, free lipids, and blood cell membranes.
\[ \cdot O_2^- + Cu^{2+} \rightarrow 0_2 + Cu^+ \]

\[ Cu^+ + H_2O_2 \rightarrow Cu^{2+} + 0H^- + \cdot O_2 \]

\[ \cdot O_2^- + H_20_2 \rightarrow 0_2 + 0H^- + \cdot O_2 \]

**Figure 1.7** The transformation of a superoxide radical anion (\( \cdot O^\cdot \)) to the hyperactive hydroxyl radical (\( \cdot OH \)) via a copper catalyzed reaction.

### 1.2.4.2 Copper Uptake, Distribution and Trafficking

Copper is acquired solely through diet in humans or nutrient medium in yeast. In the small intestines, the copper ion comes into contact with hCTR1 which translocates copper across the microvillar membrane. Divalent metal transporter 1 (DMT1) is also thought to be involved in this process [123]. The mechanism by which the copper ion is differentiated by hCTR1 from other metal ions is not known; it may be selected by size and charge. CTR1 is a high affinity membrane-associated copper importer protein conserved from yeast to humans [124, 125]. In yeast, in which CTR protein function is widely studied, the extracellular copper is delivered to one of three copper importers: yCTR1, yCTR2, and yCTR3 [126-128]. It is first reduced to Cu(I) by plasma membrane reductases before translocation by yCTR1 or yCTR3 [127, 129].

The conserved "Mets" motif (MxxxM) (see Section 1.2.2.1) in the transmembrane domain 2 of hCTR1 acquires Cu(I) for the ultimate facilitated import into the cytosol [84]. Unlike P-type ATPases, CTR proteins do not require ATP. The mechanism is thought to be driven by copper concentration differences across membranes, and by extracellular K^+, as seen in yeast [130]. The two human CTR1 transporters, hCTR1 and hCTR2, were discovered in complementation assays in yeast for
the double mutant yctr1 and yctr3 phenotypes, which were unable to transport copper into cells [125, 131]. In addition to the "Mets" motif, these two copper transporters contain a conserved, C-terminal motif HCH which was shown to be important in protein trimerization for a functional hCTR1 and interaction with the ATOX1 chaperone [132, 133]. The function of hCTR2, previously unknown due to its inability to complement a defective double mutant yctr1 and yctr3 phenotype in copper starved yeast cells [125], it was recently shown to express in late endosomes and lysosomes where it facilitated

![Figure 1.8](image)

**Figure 1.8** The intracellular copper trafficking. Copper is delivered in cytosol in the reduced form before transported to Wilson/Menkes protein (ATP7A/ATP7B) by ATOX1 chaperone. The chaperones that deliver copper to Mitochondria and Nucleus are not known, though copper ligand gets into mitochondrial through membrane channels.
cellular copper uptake [134]. CTR1 proteins are involved in permeation and export of the anti-cancer agent cisplatin in yeast and human cancer cells. Because of this phenomenon, copper permeases have been implicated in resistance to platinum based anti-cancer agents by cancer cells [135-137].

Once copper reaches the cytosol, its delivery and distribution to specific targets in the cell is undertaken by high affinity copper binding chaperones (Figure 1.8). These chaperones protect the cell from copper toxicity by binding free copper ions, act as copper scavengers, and inhibit non-specific copper binding by other cellular proteins [123]. All copper chaperones except COX-17 utilize a conserved MxCxxC motif to bind Cu(I) in a digonal geometry through the cysteine residues or trigonally by additionally utilizing glutathione or histidine/cysteine as a third ligand [138]. COX-17 binds three Cu(I) atoms arranged in a poly-copper cluster using the CCxC motif as exhibited by X-ray absorption spectroscopy [139, 140]. Yeast chaperone ATX1 delivers copper to yeast ATPase Ccc2 in the trans-Golgi network (TGN) [49], which subsequently translocates copper into vesicles for ultimate insertion into ferroxidase Fet3, a multicopper containing protein [141]. In humans, ATX1 homolog, ATOX1 delivers copper to Cu-transporting ATPases, Menkes protein (ATP7A) and Wilson protein (ATP7B) [142-144]. These Cu-ATPases differ in their trafficking routes, ATP7A delivers copper to cuproenzymes in the secretory pathway while ATP7B supplies copper to ceruloplasmin, a serum ferroxidase that binds more than 95% of copper in blood plasma [107].

Among cuproenzymes replenished with copper by ATP7A is hephaestin, a multicopper ferroxidase homologous to ceruloplasmin that oxidizes Fe$^2+$ into Fe$^{3+}$ [145]. The copper chaperone for superoxide dismutase, CCS, transports copper to Cu,Zn-SOD,
an antioxidant enzyme [146, 147]. The structural similarity between domain II of CCS and SOD1 enables direct interaction between these two proteins for copper insertion [148]. The COX-17 also relies on the CCxC motif to bind copper within the mitochondrial intermembrane space to deliver it to COX-11 and Sco proteins which in turn supply copper to the assemblage of cytochrome c oxidase complex [149, 150]. ScoI and ScoII are unique in that they bind either a single Cu(I) or Cu(II) atom [151, 152]. Under elevated copper concentrations in transfected HEK293 cells, CTR1 is readily endocytosed by a mechanism that depends on the MxxxM motif in TM2 and is quickly degraded in the lysosomes [153]. This process demonstrates how copper uptake by CTR1 may be controlled.

1.2.4.3 Intracellular Copper Trafficking and Regulation within Cells

The human copper transporting ATPases, ATP7A and ATP7B, play central roles in copper movement and regulation within the cell, and possess distinct functional properties [154]. This can be deduced from their expression, distribution, and how they are stimulated by copper concentration. ATP7A is largely expressed in intestinal epithelium and other main tissues except the liver, while ATP7B is chiefly expressed in the liver but also in the kidney and placenta, and minimal quantities are found in lung, heart, and brain tissue [5, 7, 57, 155]. In non-polarized cells and under physiological copper concentrations, ATP7A and ATP7B are localized in the trans-Golgi network (TGN) [156-160]. Elevated levels of copper stimulate ATP7A trafficking to the cell periphery and ATP7B re-localization to cytoplasmic vesicles to enable efflux of excess copper [156, 161].
Polarized epithelial cells possess a basolateral surface that orients toward the extracellular fluid and an apical surface that forms the interface with the luminal surface [162]. In vitro studies show that in polarized, under normal copper levels, ATP7A continuously shuttles between the TGN and plasma membrane while under elevated levels it pumps copper into vesicles and across the plasma membrane. The copper loaded vesicles move toward the basolateral membrane where they fuse with the membrane and release copper through exocytosis. ATP7A retreats to the TGN as the copper levels decrease [163, 164]. In like manner, the Cu-ATPase ATP7B is localized to the TGN under basal copper level and sequesters copper within vesicles at elevated copper levels [165]. These copper-loaded vesicles relocate to the apical canalicular membrane for copper excretion into the bile as ATP7B reverts to the TGN [166].

It should be noted that this trafficking process forms the regulatory basis of cell copper needs and is not unique to all polarized cells found in organisms. Other mechanisms that may influence and modulate copper trafficking and regulation by Cu-ATPases include specific trafficking by some cells [167], phosphorylation [168], protein-protein interactions [9], mutations [169], copper binding affinities [170, 171], various Cu-ATPase motifs and domains, signaling sequences [72, 172, 173] and trafficking proteins such as syntaxin 6 [72].

1.2.4.4 The Significance of Metal Binding Domains of the Wilson protein (MBDs)

The focus and basis of the work under investigation is the structural and functional properties of the N-terminal metal binding domains (MBDs) of WNDP. The spatial organization of MBDs has been described in Section 1.2.1, and the current section
focuses on their biochemical functions. There is some agreement in the published literature on the functions of the N-terminus domains but a consensus is not yet evident as new data continues to challenge the existing knowledge. The N-terminal domains are the first point of interaction between copper chaperone ATOX1 and WNDP as revealed in yeast and mammalian two hybrid systems [142, 174]. The oppositely charged surfaces of the interacting proteins facilitate ATOX1 docking and Cu(II) transfer to MBDs through a presumed three-coordinate intermediate [79]. When characterized as single isolated domains, all six MBDs appear similar in structure but surface characteristics and domain-domain interactions may give rise to different functions. The ATOX1 interacts with all six MBDs [73, 142-144, 170]. Though ATOX1 delivers copper specifically to metal binding site 2 (MBS 2) and MBS 4 [73, 142], the mechanism is not clear. In excess, ATOX1 is capable of retaliating all six MBS [142, 175]. Deletions of MBS 1-5 or MBS 6 but not MBS 1-4 were found to disrupt copper-induced trafficking of ATP7B at elevated copper levels [176]. MBS 5 and MBS 6 are in close proximity to the transduction channel; therefore, they may have a role in passing copper ions into the channel.

Inter-domain copper transfer does occur, but its physiological role is not clearly understood. From a thermodynamic standpoint, Cu(I) transfer between domains is facile [171]. A Cu(I) binding study with the N-terminus shows that each domain can accept and release Cu(I) from ATOX1; therefore, their role might be modulated by other proteins [170]. A recent NMR study by the Bertini group on Cu-ATOX1 interactions with N-terminal metal binding domains (residues 1-633) of WNDP reveals a Cu(I) bridged adduct formed between MBDs 1, 2 and 4 but not MBDs 3, 5 and 6. [175]. MBDs
3, 5 and 6 and metallated ATOX1 had a copper exchange that occurred slowly on the NMR time scale relative to MBDs 1, 2 and 4. In turn, the Cu-ATOX1/MBD 4 titration led to a higher steady state concentration of the adduct compared to Cu-ATOX1 with MMBD 1 and 2 despite having similar affinity for Cu(I) [170]. It is clear that MBD 4 plays an important role in copper transfer but how, where, and the cascade of events surrounding this observation will have to wait for more information to become available. It is possible to speculate that copper transfer in vivo may be organized in the hierarchy seen in vitro studies and may involve other molecular interactions.

The N-terminal domains influence copper-induced trafficking [177-179]. Although earlier studies had shown MBS 5 and 6, and the CPC motif were crucial for copper trafficking [176, 177, 180], Cater et al. indicated that mutations disrupting the catalytic cycle of ATP7B posed a greater risk [181]. Intracellular ATP7B trafficking requires the presence of copper and acyl-phosphate formation, but not metal ion coordination on metal binding sites, though the latter is required for efficient efflux of copper. The study of N-terminal MBDs function and properties as a single unit has not proceeded without difficulties.

The expression and purification of the WD 1-6 protein that is stable over a lengthy period of time at normal temperatures has not been very successful in other laboratories [89, 170, 175]. Under normal experimental conditions, this protein undergoes rapid degradation within 2-3 days [175]. Developing methods that would overcome this constraint, a subject of this thesis, will impact greatly not only on further characterization of the N-terminal metal binding domains but also the entire Cu-ATPase (ATP7B).
1.2.4.5 Additional Functions of Copper Transporting P-type ATPases

There is emerging evidence that copper transporting ATPases and chaperones are involved in other biological active pathways not directly related to copper homeostasis. The development of resistance after initial therapy by tumor cells to the anti-cancer agent cisplatin [cis-diamminedichroloplatinum (II) or CDPP] is partly mediated by copper transporting ATPases and chaperones as shown in vitro studies [182, 183]. The MNKP and WNDP have been found to sequester CDPP into vesicles and confer resistance to elevated levels of DPP [184-187]. These proteins are expressed in a variety of carcinoma cells [188-190]. A fluorescent CDPP has been found to co-localize with tagged ATP7B [191]. These studies may suggest that CDPP interaction with Cu-ATPases in vivo may lead to reduced efficacy to the drug therapy. Mechanistically, this interaction is not well understood, but recent work by Lutsenko's group sheds some light into process [183]. ATP7B does not confer resistance to CDPP directly; however, CDPP binding to ATP7B modulates phosphorylation activity and copper efflux is accelerated. Copper deficient cells have shown resistance to CDPP [186]. Even more intriguing is the possibility that the human copper chaperone ATOX1 can bind a stripped CDPP in the Cu metal binding site [192] without the chloride ligands. Binding of DPP by Cu(I) transporters is likely to lead to unhealthy competition between the targeting pathway and copper homeostasis.

The Cu-ATPases also interact with other proteins of clinical significance. COMMD1 (formerly known as MURR1), is a protein product coded by a defective gene in Bedlington terrier dogs suffering from liver copper toxicosis [193, 194]. This protein has been proposed to be involved in regulation of copper trafficking and other cellular mechanisms [9, 195], and it has been shown to interact with the N-terminus of WNDP.
but not MNKP [196]. The dynactin subunit p62 interacts with the N-terminus of WNDP, specifically to MBDs 4-6. This interaction is copper-dependent, and may be responsible for copper mediation through microtubules [197].

1.3 Wilson Disease

Wilson Disease (WD) is an inherited autosomal recessive disorder of copper homeostasis that leads to accumulation of copper in key organs, such as the liver, brain and kidneys [198]. The main cause of WD is the defective ATP7B that fails to deliver copper to the bile excretory system. Many mutations [65, 199-203] in the ATP7B gene alter its ability to effectively function; a database for these mutations has been created and is maintained by the Cox laboratory at the University of Alberta, Canada (http://www.wilsondisease.med.ualberta.ca/database.asp). If left untreated, WD can lead to permanent organ damage, neurological, and psychological problems, eventually, death if a liver transplant is not performed [204]. The prevalence of Wilson disease is reportedly 1 in 30,000 births; this number might be an underestimate because many cases remain undiagnosed or misdiagnosed [205-207].

1.3.1 The History of Wilson Disease

The historical events of Wilson disease were compiled by Walshe, who initiated the initial treatment of Wilson disease (WD) with D-penicillamine in 1955 [208]. The American-born neurologist Samuel Alexander Kinnier Wilson was the first physician to describe the progressive lenticular degeneration disorder that later came to be known as Wilson disease (WD) [209]. His original description, which was published in a single
issue of the journal Brain in March 1912 (213 pages in all) described WD as a rare, familial, progressive, and invariable fatal disease, accompanied by neurological deficit and liver cirrhosis. Preceding Wilson's work was Kayser who described pigmented corneal rings (now known as Kayser-Fleischer rings or K-F) in patients suffering from "pseudosclerosis" [210, 211].

A correlation between K-F rings and Wilson disease was not apparent because it is not found in all patients. In 1913, Rumpel identified excess copper deposits in the liver of WD patients, while Vogt in 1929, Haurowitz in 1930, and later Glazebrook [212], reported copper overloads in the liver or brain of the dying patients. Earlier in 1921, Hall had postulated that WD exhibited autosomal recessive inheritance. Cumings [213] initiated the modern understanding of WD in 1948 by publishing a comparative study of copper and iron content in brain, the liver, and hepatodegeneration. He was first to propose the use of Dimercaprol for treatment of the disease. The 1950-1960's saw discovery of ceruloplasmin deficiency in Wilson patients [214, 215] and the emergence of treatment therapies such as Dimercaprol (BAL) [216], D-penicillamine [217, 218], and zinc salts [219]. An advance in WD research in the 1970's and 1980's culminated in the identification of the ATP7B gene in 1993. This has ushered WD into the molecular genetics era [5, 7, 8].

1.3.2 Clinical Aspects and Diagnosis of Wilson Disease

The onset of the WD ranges from early age to adulthood. Children as young as 2 years old and adults over 70 have developed the disease [220]. There are four main aspects of WD clinical presentations: hepatic, hemolytic, neurological and psychiatric ailments.
Hepatic and hemolytic symptoms are manifested by chronic hepatitis, acute hepatitis, cirrhosis, jaundice and liver failure [221]. They can be also accompanied by high serum transaminase levels, anemia secondary to hemolysis [220], and Kayser-Fleischer rings (K-F) (Figure 1.9) have been found in roughly 55% of patients at this stage [222]. The K-F rings are caused by copper deposition on Descemet’s membrane in the cornea. Neurological features which tend to present later in life include hypokinetic speech, dystonia, tremor, lack of motor incoordination, and inability to perform basic functions [223]. Nearly all neurological patients show K-F rings and may exhibit Parkinson disease-like behavior as copper accumulates in the basal ganglia [224]. Behavioral changes, hallucinations, depression, and anxiety are among many of the psychiatric symptoms. Other nonclassical features that occur in WD patients are hypoparathyroidism [225], infertility [226], renal abnormalities [227], and skeletal muscle diseases [228].

![Copper metal deposits](image)

**Figure** 1.9 Kayser-Fleischer ring in the cornea Descemet’s membrane. A copper deposit is the golden brown ring indicated by arrow in at the edge of the iris (220).
The diagnosis of WD is not based on any on single procedure but an accumulation of evidence from clinical observations and laboratory tests. Normally, cases showing known symptoms are easily diagnosed; the results can be confirmed by laboratory tests such as the measurement of 24 h urinary copper excretion, hepatic copper concentration, serum 'free' copper concentration, and ceruloplasmin concentration. Since these tests can also indicate other liver diseases, a single method diagnosis should not be relied upon as a positive test for WD [229]. The presence of WD can be indicated by urinary copper concentration greater than 100 (µg/24 h (>1.6 (imol/24 h), liver copper content more than 250 µg/g dry weight (normal < 50 ng/g dry weight), serum free copper concentration more than 25 ng/dL (normal <15 µg/dL) [220], and D-penicillamine urine challenge test (urinary excretion of copper > 1600 µg per day following D-penicillamine administration) [230]. The determination of ceruloplasmin concentration is also important, but it has limited usage for diagnosis.

Despite the fact that about 20% of WD patients have normal ceruloplasmin concentration (200-400 mg/L), most patients, heterozygous WD carriers, and some normal individuals have ceruloplasmin concentrations of less than 200 mg/L. Therefore, the enzymatic assay method (oxidase activity) of determining ceruloplasmin levels is preferred for Wilson disease diagnosis [231, 232], Other parameters used for diagnosis is the presence of K-F rings described earlier. WD patients with hepatic symptoms and neuropsychiatric disorders are likely to exhibit K-F rings [233]. Another tool that is gaining momentum in WD diagnosis is genetic testing of first degree relatives of a patient [234]. By mutational analysis, it is becoming easier both to confirm the presence of a defective gene and to determine the severity of the mutation on the malfunctioning
WNDP. Even though not common, newborn screening for WD is being encouraged to identify affected individuals early. This is done by measurement of ceruloplasmin in Guthrie dried-blood spots or urine samples from newborn infants [235-237].

1.3.3 Treatment of the Wilson Disease

The treatment of WD is intended to decrease the copper concentration in the body either through accelerated copper excretion or impaired copper absorption from diet. Currently, this is accomplished by use of copper chelating drugs (Figure 1.10) and diet management. The discovery of WD in the early twentieth century was not followed by any form of therapy until 1948. Cumings proposed and Mandlebrote and co-workers subsequently used BAL for treatment of WD [216, 238]. Incidentally, BAL, a chelating agent, was developed as antidote agent to the arsenical war gas lewisite [239, 240]. BAL was a promising candidate for management of WD as it was able to remove copper from body rapidly. BAL usefulness was overshadowed by its toxicity and the requirement of painful intramuscular injections; a substitute was needed [208]. Now, WD is managed by several therapies and drugs. D-penicillamine (D-PCA) is an orally administered chelating drug [217]; it forms copper-pencillamine complexes that are excreted in urine. D-PCA possesses a free sulfhydryl group that is capable of chelating copper. It works fairly well for patients with hepatic symptoms [220]. Its side effects are numerous and include Vitamin B6 deficiency, hypersensitivity, nephrotoxicity, hematuria, bone marrow thinning, and Goodpasture's syndrome [223].

Trientine (triethylene tetramine dihydrochloride) is another copper chelating agent that works in the same fashion as D-PCA though it has fewer side effects. It is the first
line of defense in WD therapy, but not as effective as D-PCA. Unlike D-PCA, trientine lacks a sulfhydryl group, is less toxic, stimulates tissue copper efflux, and patients with neurological symptoms tend to worsen when this is used as the primary treatment [241]. Trientine has a polyamine-like structure and chelates copper to form a stable complex using four nitrogens in a planar ring [223]. Zinc has been used to manage WD for almost six decades [242]. Its mode of action and application is distinct from Trientine and D-PCA. Zinc promotes enterocyte metallothionein synthesis. Metallotheinein has a higher affinity for copper than zinc, and this induction promotes copper excretion [243]. Metallotheinein bound copper is not absorbed but excreted into fecal matter as cells undergo shedding. Gastric irritation and immune suppression are just a few side effects associated with zinc therapy [244, 245].

Ammonium tetrathiomolydate (TM) is another copper removal agent which is no longer used to treat WD. Its mode of action involves two mechanisms: First, it complexes with copper in intestinal lumen, hindering absorption. Second, it inhibits copper uptake by forming a copper complex with albumin [246]. Recently, TM was found to form a complex with copper chaperone ATOX1 homolog, ATX1; therefore, it could inhibit copper trafficking proteins through metal cluster formation [247]. TM was preferred for patients with neurological signs and symptoms, but due to limited clinical trial data, this drug is not commercially available in the United States or Europe. Patients receiving TM may suffer from bone marrow depression and hepatotoxicity side effects [248, 249]. WD is also managed by lifestyle change. Foods rich in copper such as liver, chocolate, nuts and mushrooms should be avoided. Antioxidants, mainly vitamin E, have been found low in WD patients so supplements are encouraged [220]. Liver transplantation is the last
option for WD patients with acute liver failure who are non-responsive to medical therapy. A prognostic scoring system for WD has been developed to assess those who might be at higher risk and require liver transplant [250].

Figure 1.10 The copper chelating agents. They possess sulfhydryl and polyamines functional groups that bind to copper. Ammonium tetrathiomolybdate (TM) is no longer used for WD treatment due to its toxicity.
1.4 Summary

Copper is an essential element required in the body for many enzymatic activities and maintenance of structural integrity of proteins. Maintaining cellular copper concentration within physiological limits is critical, excess copper in the body leads to production of superoxide radicals that damage cell membranes, whereas copper deficit is associated with pigmentation and growth disorders. The role and importance of the copper transporting ATPase ATP7B (Wilson protein) in copper homeostasis was discussed in detail. WNDP is structurally and functionally related to Menkes protein (MNKP), a homologous protein that shares a complementary function with WNDP. The Cu-ATPases belong to the heavy metal transporting P-type ATPase (PiB) subfamily and are distinguished by their possession of N-terminal metal binding domains. Structurally, Cu-ATPases are organized into five functional domains, the transmembrane (TM-), phosphorylation (P-), nucleotide binding (N-), the actuator (A-), and the N-terminal metal binding (MBD) domains. To effect copper trafficking, Cu-ATPases acquire copper from the metallochaperone ATOX1 and deliver it either to secretory vesicles or cuproenzymes. Copper transporting ATPases and the human copper transporter (hCTR1) are also found to be involved in resistance to the anticancer drug cisplatin by tumor cells.

The N-terminal domains play an important role in copper transfer because they form the first point of contact between Cu-ATPases and the metallochaperone ATOX1. Deletion of N-terminal domains has been shown to abrogate the enzymatic activity of Cu-ATPases. Cu-ATPase activity is in turn regulated by cellular copper levels. Several mutations in the ATP7B gene cause WD, a disorder characterized by accumulation of copper in key organ tissues such as the liver and kidney. WD is managed by
administration of copper chelating agents, BAL, D-penicillamine and trientine. Zinc salts
which promote in vivo copper binding by metallothioneins are also used along with
addition to limiting copper intake through diet. Liver transplant is usually the last option
for patients who are non-responsive to treatment. Clearly, understanding the function of
both native and mutant copper metabolism proteins at the molecular level is a worthy
pursuit highly germane to human health.
References


45


Cumings, J. N. "The copper and iron content of brain and liver in the normal and in hepatolenticular degeneration" (1948) *Brain* 71, 410-415.


CHAPTER 2

MATERIALS AND EXPERIMENTAL METHODS

2.1 Introduction

This chapter covers the materials and methodology that were used in this study. A description of the following major subtopics is given: materials, instrumentation, procedures, DNA plasmids and PCR primers, and biophysical characterization of the proteins.

2.2 Chemicals and Reagents

The chemicals and reagents are listed in Table 2.1 along with the purity level where applicable, and the supplier's name and location. In general, all chemical and reagents met molecular grade standards. Milli-Q water (18.2 MQ/cm) purified using US Filter PURELAB Plus UV/UF (Millipore) was used in all experiments. The water for cell culture and processes prone to microbial contamination were autoclaved for 30 min at 121°C. Enzymes, process kits, and purification aids are listed separately in Table 2.2.

Table 2.1 Reagents and chemicals.

<table>
<thead>
<tr>
<th>Chemical/Reagent</th>
<th>Catalog</th>
<th>Percent Purity</th>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acrylamide</td>
<td>164855000</td>
<td></td>
<td>Acros Organics</td>
<td>Morris Plain, NJ</td>
</tr>
<tr>
<td>Agarose</td>
<td>15510-019</td>
<td></td>
<td>Invitrogen</td>
<td>Carlsbad, CA</td>
</tr>
<tr>
<td>Ammonium Chloride</td>
<td>A661-500</td>
<td>99.8%</td>
<td>Fisher Scientific</td>
<td>Fair Lawns, NJ</td>
</tr>
</tbody>
</table>
Table 2.1 — Continued

<table>
<thead>
<tr>
<th>Chemical/Reagent</th>
<th>Catalog</th>
<th>Percent Purity</th>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ammonium Persulfate</td>
<td>A3678-25G</td>
<td>&gt;98%</td>
<td>Sigma-Aldrich</td>
<td>St. Louis, MO</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>A-9518</td>
<td>92.5-100.5%</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Antifoam A</td>
<td>A-5757</td>
<td></td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Bis-Acrylamide</td>
<td>BP171-25</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Boric Acid</td>
<td>BP 168-1</td>
<td>&gt;99%</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Bromophenol Blue</td>
<td>B-6131</td>
<td></td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Carbenicillin</td>
<td>762024</td>
<td></td>
<td>Bio-World</td>
<td>Dublin, OH</td>
</tr>
<tr>
<td>Chloramphenicol</td>
<td>220551</td>
<td>&gt;99.6%</td>
<td>EMB Biosciences</td>
<td>La Jolla, CA</td>
</tr>
<tr>
<td>Coomasie Blue R-250</td>
<td>20278</td>
<td></td>
<td>Thermo Scientific</td>
<td>Rockford, IL</td>
</tr>
<tr>
<td>Coomasie Plus</td>
<td>23236</td>
<td></td>
<td>Pierce</td>
<td>Rockford, IL</td>
</tr>
<tr>
<td>Dextrose</td>
<td>BP350-500</td>
<td>99.8%</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>DTT</td>
<td>1758-9030</td>
<td></td>
<td>Inalco Spa</td>
<td>Milan, Italy</td>
</tr>
<tr>
<td>EDTA</td>
<td>BP119-500</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Glutathione, reduced</td>
<td>G425-500</td>
<td>&gt; 99 %</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Glycerol</td>
<td>G33-1</td>
<td>99.9%</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Guanidine HCl</td>
<td>50933</td>
<td>&gt; 99.5%</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>HEPES</td>
<td>H7006-500G</td>
<td>99.5%</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Imidazole</td>
<td>56750</td>
<td>99.5%</td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Iodoacetamide</td>
<td>653k5320</td>
<td></td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>IPTG</td>
<td>1758-1400</td>
<td></td>
<td>Inalco Spa</td>
<td></td>
</tr>
<tr>
<td>Magnesium Chloride</td>
<td>BP214-500</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>MEM Vitamins</td>
<td>75-020-C1</td>
<td></td>
<td>MediaTech</td>
<td>Herndon, VA</td>
</tr>
<tr>
<td>MES</td>
<td>BP300-100</td>
<td>98%</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Nickel (II) Sulfate</td>
<td>211085000</td>
<td>99%</td>
<td>Acros Organics</td>
<td></td>
</tr>
<tr>
<td>Pefabloc SC PLUS</td>
<td>1187601001</td>
<td></td>
<td>Roche Diagnostics</td>
<td>Indianapolis, IN</td>
</tr>
<tr>
<td>Sodium Dodecyl sulfate (SDS)</td>
<td>1535-017</td>
<td>&gt;99.5%</td>
<td>GIBCOBRL</td>
<td>Grand Island, NY</td>
</tr>
<tr>
<td>Thiamine HCl</td>
<td>BP892-100</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
</tbody>
</table>
**Table 2.1**—Continued

<table>
<thead>
<tr>
<th>Chemical/Reagent</th>
<th>Catalog</th>
<th>Percent Purity</th>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tricine</td>
<td>22561 1</td>
<td>&gt;98.5%</td>
<td>USB Corp.</td>
<td>Cleveland, OH</td>
</tr>
<tr>
<td>Tricine</td>
<td>807420</td>
<td></td>
<td>MP Biochemicals</td>
<td>Solon, OH</td>
</tr>
<tr>
<td>Tris Base</td>
<td>BP152-5</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Tryptone</td>
<td>211713</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Tryptone</td>
<td></td>
<td></td>
<td>MoBio Labs</td>
<td>Carlsbad, CA</td>
</tr>
<tr>
<td>Tryptone</td>
<td>T7293-2509</td>
<td></td>
<td>Sigma-Aldrich</td>
<td></td>
</tr>
<tr>
<td>Urea</td>
<td>BP 169-500</td>
<td>99%</td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Vitamin H</td>
<td>B0463</td>
<td></td>
<td>Tokyo Kasei Kogyo</td>
<td>Tokyo, Japan</td>
</tr>
<tr>
<td>Xylene Cyanol FF</td>
<td>BP565-10</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>Yeast Extract</td>
<td>212720</td>
<td></td>
<td>Fisher Scientific</td>
<td></td>
</tr>
<tr>
<td>(13C) Glucose</td>
<td>CLM-1396-1</td>
<td>99%</td>
<td>Cambridge Isotope</td>
<td>Andover, MA</td>
</tr>
<tr>
<td>(15N) Ammonium Chloride</td>
<td>NLM-467-1</td>
<td>99%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2.2** Enzymes, process kits, and purification aids.

<table>
<thead>
<tr>
<th>Enzymes/Kits</th>
<th>Catalog</th>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 kb DNA Ladder</td>
<td>N3232S</td>
<td>NEB</td>
<td>Cambridge, MA</td>
</tr>
<tr>
<td>100 bp DNA Ladder</td>
<td>N3231S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>BamHI</td>
<td>B0136S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>Benzonase</td>
<td>70746</td>
<td>Novagen</td>
<td>San Diego, CA</td>
</tr>
<tr>
<td>Bglll</td>
<td>R0144S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>BugBuster® Protein Extraction Reagent</td>
<td>70584-4</td>
<td>Novagen</td>
<td></td>
</tr>
<tr>
<td>Deep Vent (exo-) DNA Polymerase</td>
<td>M0259S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>Deep Vent DNA Polymerase</td>
<td>M0258S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>Dialysis Cassettes</td>
<td>66030</td>
<td>Pierce</td>
<td>Rockford, IL</td>
</tr>
</tbody>
</table>
Table 2.1 —Continued

<table>
<thead>
<tr>
<th>Enzymes/Kits</th>
<th>Catalog</th>
<th>Supplier</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>DH5a Cells</td>
<td></td>
<td>Dr. Huffman’s glycerol stocks</td>
<td></td>
</tr>
<tr>
<td>Factor Xa</td>
<td></td>
<td>Enzyme Research</td>
<td>South Bend, IN</td>
</tr>
<tr>
<td>Fast-Link DNA Ligation Kit</td>
<td>LK11025</td>
<td>Epicentre</td>
<td>Madison, WI</td>
</tr>
<tr>
<td>HotStar HF Start DNA Polymerase</td>
<td>202602</td>
<td>Qiagen</td>
<td>Valencia, CA</td>
</tr>
<tr>
<td>KOD Hot Start DNA Polymerase</td>
<td>71086-3</td>
<td>Novagen</td>
<td></td>
</tr>
<tr>
<td>MinElute PCR Purification Kit</td>
<td>28004</td>
<td>Qiagen</td>
<td></td>
</tr>
<tr>
<td>Nco I</td>
<td>R0193</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>Ndel</td>
<td>R0111S</td>
<td>NEB</td>
<td></td>
</tr>
<tr>
<td>pET32Xa/LIC Kit</td>
<td>70072</td>
<td>Novagen</td>
<td></td>
</tr>
<tr>
<td>pET41Ek/LIC Kit</td>
<td>71071</td>
<td>Novagen</td>
<td></td>
</tr>
<tr>
<td>QIAfilter Plasmid Midi kit</td>
<td>12243</td>
<td>Qiagen</td>
<td></td>
</tr>
<tr>
<td>QIAprep Spin Miniprep Kit</td>
<td>27106</td>
<td>Qiagen</td>
<td></td>
</tr>
<tr>
<td>QIAquick Gel Extraction Kit</td>
<td>28706</td>
<td>Qiagen</td>
<td></td>
</tr>
<tr>
<td>Rainbow Protein Marker</td>
<td>RPN800E</td>
<td>GE</td>
<td>Piscataway, NJ</td>
</tr>
<tr>
<td>Rosetta(DE3) Cells</td>
<td>70954</td>
<td>Novagen</td>
<td></td>
</tr>
<tr>
<td>Site-Directed Mutagenesis Kit</td>
<td>200521-5</td>
<td>Stratagene Cedar Creek, TX</td>
<td></td>
</tr>
<tr>
<td>TEV Protease</td>
<td>12575-015</td>
<td>Invitrogen</td>
<td></td>
</tr>
</tbody>
</table>

2.3 DNA Plasmids and PCR Primers

The Wilson protein cDNA encoding the entire ATP7B was kindly provided by Professor Jonathan Gitlin of Washington University at St. Louis (now at Vanderbilt University). The Wilson cDNA was transformed into DH5a cells and prepared glycerol stocks were stored at -80°C. The DNA Strider 1.3 software was used to design the gene inserts encoding the protein's domains of interest. PCR mutagenic primers were designed and analyzed using DNA Strider 1.3 software or an online based oligo-analyzer at
The primers were ordered from Integrated DNA Technologies, Inc. (Coralville, Iowa). The notation used for Wilson's protein constructs and plasmids vectors is shown in Tables 2.3 and 2.6. The primary primers (Table 2.4) were expressly designed either to amplify the genes of interest or to incorporate mutations, whereas secondary primers (Table 2.5) were additional primers for which need arose in the course of this work. All lyophilized primers were resuspended in double distilled nuclease-free water to make a 200 pM stock solution. This stock solution was further diluted to make 10 μM polymerase chain reaction (PCR) working aliquots and stored at -20°C.

Table 2.3 Notation for proteins studied.

<table>
<thead>
<tr>
<th>Notation</th>
<th>Residues in ATP7B</th>
<th>Name</th>
<th>bp in gene</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD1-6</td>
<td>57-633</td>
<td>N-terminal Domains 1-6</td>
<td>1731</td>
</tr>
<tr>
<td>WD5-6</td>
<td>485-633</td>
<td>N-terminal Domains 5-6</td>
<td>447</td>
</tr>
<tr>
<td>WD5-6(Y532H)</td>
<td>485-633</td>
<td>Y532H mutation of WD5-6</td>
<td>447</td>
</tr>
<tr>
<td>WD5-6(G591D)</td>
<td>485-633</td>
<td>G591 mutation of WD5-6</td>
<td>447</td>
</tr>
<tr>
<td>Other protein</td>
<td></td>
<td>COMMD1</td>
<td>450</td>
</tr>
<tr>
<td>COMMD1</td>
<td></td>
<td>COMMD1</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.4 Primary primers.

<table>
<thead>
<tr>
<th>Primer</th>
<th>bp</th>
<th>% GC</th>
<th>TM °C</th>
<th>Sequence 5’ to 3’</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMMD1 BamHI Rev</td>
<td>37</td>
<td>56.8</td>
<td>68.2</td>
<td>GTC AGT GGG GAG GGA TCC TGG ACA CCT TGG TTT CAA C</td>
</tr>
<tr>
<td>COMMD1 Ncol Fwd</td>
<td>21</td>
<td>81.0</td>
<td>71.2</td>
<td>GCC CGC GAA CCA TGG CGG CGG</td>
</tr>
<tr>
<td>Ek/LICWD5Fwd</td>
<td>39</td>
<td>53.8</td>
<td>67.7</td>
<td>GAC GAC GAC AAG ATG GCA CCG CAG AAG TTC TTA CAT</td>
</tr>
<tr>
<td>Ek/LICWD6Rev</td>
<td>40</td>
<td>60.0</td>
<td>70.4</td>
<td>GAG GAG AAG CCC GGT CTA CTG GGC CAG GGA AGC ATG AAA G</td>
</tr>
<tr>
<td>GC_KD-Y532H Rev</td>
<td>34</td>
<td>55.9</td>
<td>66.4</td>
<td>CTG GAT GAC CTC TGG GTC ATG CTT GAT CTC TGC C</td>
</tr>
<tr>
<td>GC TEV PI</td>
<td>43</td>
<td>58.1</td>
<td>70.0</td>
<td>CCA CGC GAC CCT CAA TAA GGT TCT CAC CAC CGG TAC CCA GAT C</td>
</tr>
<tr>
<td>GC_TEV PII</td>
<td>39</td>
<td>58.9</td>
<td>70.4</td>
<td>GCT GGT GGC CAC GGC ACC GAA ATA AAG GTT CTC ACC ACC</td>
</tr>
<tr>
<td>GC TEV PIII</td>
<td>39</td>
<td>56.4</td>
<td>69.3</td>
<td>CTG TGC TGG TGG CCA CGC CCT GGA AAT AAA GGT TCT CAC</td>
</tr>
<tr>
<td>KD-Y532H Fwd</td>
<td>34</td>
<td>55.9</td>
<td>66.4</td>
<td>GCC AGA GAT CAA GCA TGA CCC AGA GGT CAT CCA G</td>
</tr>
<tr>
<td>TEV PI</td>
<td>43</td>
<td>58.1</td>
<td>70.1</td>
<td>GAT CTG GGT ACC GGT GGT GAG AAC CTT ATT GAG GGT CGC GTGG</td>
</tr>
<tr>
<td>TEV PII</td>
<td>39</td>
<td>58.9</td>
<td>70.4</td>
<td>GGT GGT GAG ACC CT TAT TAT TTC GGT CGC GTG GCC ACC AGC</td>
</tr>
<tr>
<td>TEV PIII</td>
<td>39</td>
<td>56.4</td>
<td>69.3</td>
<td>GTG AGA ACC TT ATT TCC AGG CCG TGG CCA CCA GCA CAG</td>
</tr>
<tr>
<td>WD1 EkLIC Forward</td>
<td>45</td>
<td>60</td>
<td>71.9</td>
<td>GACGACGACAAGATGTGGCCACCCAGACACAGTCGGATCTTGGGC</td>
</tr>
<tr>
<td>WD1 XaLIC Forward</td>
<td>45</td>
<td>62.2</td>
<td>72.9</td>
<td>GGTATTTGAGGCTCGTGCGCCACCCAGACACGATCAGGTCTTTGGGC</td>
</tr>
<tr>
<td>WD1 XaTEVLIC Forward</td>
<td>36</td>
<td>52.8</td>
<td>66.2</td>
<td>GGT ATT GAG GGT CGC GAG AAC CT TAT TAT TTC CAG GGC</td>
</tr>
<tr>
<td>WD6 EkLIC reverse</td>
<td>49</td>
<td>57.1</td>
<td>72.0</td>
<td>GAGGAAGAACCGGGTTACTGGGGCCAGGGAAAGCATGAAAGCCAAATTCCCC</td>
</tr>
<tr>
<td>WD6 XaLIC Reverse</td>
<td>51</td>
<td>52.9</td>
<td>70.8</td>
<td>AGAGGGAGAGTTAGAGCCTCCTGGGGCCAGGGAAAGCATGAAAGCCAAATTCCCC</td>
</tr>
<tr>
<td>WD5 XaLIC Fwd</td>
<td>42</td>
<td>54.8</td>
<td>69.1</td>
<td>GGT ATT GAG GGT CGC ATG GCA CCG CAG AAG TGC TTC TTA CAG</td>
</tr>
<tr>
<td>WD6 XaLIC Rev</td>
<td>42</td>
<td>54.9</td>
<td>69.1</td>
<td>AGA GGA GAG TTA GAG CCC TAC TGG GCC AGG GAA GCA TGA AAG</td>
</tr>
</tbody>
</table>
Table 2.5 Secondary primers.

<table>
<thead>
<tr>
<th>Primer</th>
<th>bp</th>
<th>% GC</th>
<th>TM °C</th>
<th>Sequence 5' to 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET32XaLICWD1-6 Fix Deletion</td>
<td>27</td>
<td>62.9</td>
<td>66.5</td>
<td>CTG AGA CCT TGG GCC ACC AAG GAA GCC</td>
</tr>
<tr>
<td>GC - pET32XaLICWD1-6 Fix Deletion</td>
<td>27</td>
<td>62.9</td>
<td>66.5</td>
<td>GGC TTC CTT GTG GCC CCA AGG TCT CAG</td>
</tr>
<tr>
<td>Fix GtoE at base 519</td>
<td>37</td>
<td>48.6</td>
<td>64.3</td>
<td>GGT CCG GAA ACT GCA AGA AGT GAG AGT CAA AGT C</td>
</tr>
<tr>
<td>GC - Fix GtoE at base 519</td>
<td>37</td>
<td>48.6</td>
<td>64.3</td>
<td>GAC TTT GAC TCT CAC TAC TTC TTG CAG TTT CCG GAC C</td>
</tr>
<tr>
<td>Fix Insertion WD 1-6</td>
<td>33</td>
<td>60.6</td>
<td>68.3</td>
<td>GTC ATG GAG GAC TAC GCA GGC TCC GAT GGC AAC</td>
</tr>
<tr>
<td>GC - Fix Insertion WD 1-6</td>
<td>33</td>
<td>60.6</td>
<td>68.3</td>
<td>GTC ATG GAG GAC TAC GCA GGC TCC GAT GGC AAC</td>
</tr>
<tr>
<td>Fix AtoT at 395 (6709)</td>
<td>38</td>
<td>52.6</td>
<td>67.5</td>
<td>GCA AAT ATC GGT GTC TTT GCC GGA AGG GAC TGC AAC AG</td>
</tr>
<tr>
<td>Fix 2 MIS at 468,9 (6930)</td>
<td>37</td>
<td>64.8</td>
<td>72.3</td>
<td>CAC TGG GAG CCT CCC TGC AAA CCA TGC CCC GGA CAT C</td>
</tr>
<tr>
<td>Fix GtoD at 496 (7013)</td>
<td>35</td>
<td>48.5</td>
<td>64.3</td>
<td>GCT TCT TAC AGA TCA AAG GCA TTA CTG CAT CC</td>
</tr>
<tr>
<td>GC - Fix AtoT at 395 (6709)</td>
<td>38</td>
<td>52.6</td>
<td>67.5</td>
<td>CTG TCG TCC CTT CGG CCA AAG ACA CCG TTA TTT GC</td>
</tr>
<tr>
<td>GC - Fix 2 MIS at 468,9 (6930)</td>
<td>37</td>
<td>64.8</td>
<td>72.3</td>
<td>GAT GTC CGG GCC ATG GTT TGC AGG GAG CCT CCC AGT G</td>
</tr>
<tr>
<td>GC - Fix GtoD at 496 (7013)</td>
<td>35</td>
<td>48.5</td>
<td>64.3</td>
<td>GGA TGC ACA GGT CAT GCC TTT GAT CTG TAA GAA GC</td>
</tr>
<tr>
<td>Fix D to G (M1) at base 6626</td>
<td>37</td>
<td>56.7</td>
<td>69.1</td>
<td>CTC TGA TTG CCA TTG CCG GCC TGA CCT GTG CAT CCT G</td>
</tr>
<tr>
<td>Fix G to E at base 6035</td>
<td>37</td>
<td>51.3</td>
<td>65.4</td>
<td>GGT CCG GAA ACT GCA AGG AGT AGT CAA AGT C</td>
</tr>
<tr>
<td>GC - Fix D to G (M1) at base 6</td>
<td>37</td>
<td>56.7</td>
<td>69.1</td>
<td>CAG GAT GCA CAG GTC CGG CCA GAT GCA ATC AGA G</td>
</tr>
<tr>
<td>GC - Fix G to E at base 6035</td>
<td>37</td>
<td>51.3</td>
<td>65.4</td>
<td>GAC TTT GAC TCT CAC TAC TCC TTG CAG TTT CCG GAC C</td>
</tr>
<tr>
<td>FIX D to G (DEC 1506)</td>
<td>37</td>
<td>54.0</td>
<td>67.4</td>
<td>CTC TGA TTG CCA TTG CCG GTA TCA CCT GTG CAT CCT G</td>
</tr>
<tr>
<td>FIX G to E (DEC 1506)</td>
<td>37</td>
<td>54.0</td>
<td>66.9</td>
<td>GGT CCG GAA ACT GCA AGG AGT GGT GAG AGT CAA AGT C</td>
</tr>
<tr>
<td>GC - FIX D to G (DEC 1506)</td>
<td>37</td>
<td>54.0</td>
<td>67.4</td>
<td>CAG GAT GCA CAG GTC ATA CCG GCA ATG GCA ATC AGA G</td>
</tr>
<tr>
<td>GC - FIX G to E (DEC 1506)</td>
<td>37</td>
<td>54.0</td>
<td>66.9</td>
<td>GAC TTT GAC TCT CAC TCC TTG CAG TTT CCG GAC C</td>
</tr>
<tr>
<td>Fix G to R (AA#339)</td>
<td>39</td>
<td>53.8</td>
<td>67.0</td>
<td>GGA ATG AGA ACT GGA AGA CCT GTG ATC TGT CCC ACT CCC</td>
</tr>
<tr>
<td>GC - Fix G to R (AA#339)</td>
<td>39</td>
<td>53.8</td>
<td>67.0</td>
<td>GGA ATG AGA ACT GGA AGA CCT GTG ATC TGT CCC ACT CCC</td>
</tr>
</tbody>
</table>
Table 2.6 Plasmid vectors.

<table>
<thead>
<tr>
<th>Vector</th>
<th>Source</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>pET24d</td>
<td>Novagen</td>
<td>Non-fusion expression</td>
</tr>
<tr>
<td>pET32Xa/LIC</td>
<td>Novagen</td>
<td>Thioredoxin tag</td>
</tr>
<tr>
<td>pET41Ek/LIC</td>
<td>Novagen</td>
<td>Glutathione S-Transferase tag</td>
</tr>
</tbody>
</table>

2.4 Instrumentation

The amplification of genes and mutagenesis was performed with a Minicycler (Model PTC-150) thermocycler, MJ Research, Inc. (Watertown, MA). Plated cells were grown at 37°C in a gravity convection incubator, Precision 4L (Thermo Fisher Scientific, Waltham, MA). The agarose electrophoresis was done with a Fisher Biotech electrophoresis system, FB-SB-7120. For the culture growth, small scale growth (less than 2 liters) was carried out in a Controlled Environment Incubator Shaker, New Brunswick Scientific. All large scale production of protein (10 L) was done either by using a BioFlo310® Fermentor/bioreactor (Figure 2.2) or Microferm Fermentor (New Brunswick Scientific).

The concentrations of protein and DNA samples were spectroscopically measured using a UV/VIS Beckman Diode Array Spectrophotometer, DU-7400. Protein concentration was determined by BCA assay (Pierce, Rockford, IL) and Bio-Rad protein assay using a Benchmarker microplate reader (Bio-Rad, Hercules, CA). The protein purification and gel filtration experiments were done with an Amersham AKTA FPLC (product code 18-1900-26) system in an Isotemp laboratory refrigerator at 4°C (Fisher Scientific). The list and nature of the columns that were used are shown in Figure 2.1 and
Several centrifugation instruments were available; for small samples, Eppendorf 5415 and Eppendorf *Minispin* (Germany) Plus centrifuges were used, while a Sorvall RC 5B plus high speed centrifuge was used for large samples.

*Figure 2.1* Purification columns for AKTA FPLC system (GE Life Sciences).

**Table 2.7** Columns for protein purification and analysis (GE Life Sciences).

<table>
<thead>
<tr>
<th>Column Name</th>
<th>Product Code</th>
<th>Matrix/Size</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAE Sepharose</td>
<td>XK 26</td>
<td>DEAE Sepharose FF</td>
<td>Cation exchange</td>
</tr>
<tr>
<td>GSTpreppFF 16/10</td>
<td>28-9365-50</td>
<td>Glutathione sepharose</td>
<td>GST purification</td>
</tr>
<tr>
<td>HisPrep FF 16/10</td>
<td>28-9365-51</td>
<td>Nickel Sepharose</td>
<td>His-tag purification</td>
</tr>
<tr>
<td>HiPrep 26/10 Desalt</td>
<td>17-5087-01</td>
<td>Superdex G-25 superfine</td>
<td>Buffer exchange</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</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</td>
</tr>
<tr>
<td>Superdex 75HR 10/300</td>
<td>17-1047-01</td>
<td>Superdex 75, 13^M</td>
<td>Gel filtration</td>
</tr>
<tr>
<td>Superdex 200 10/300 GL</td>
<td>17-5175-01</td>
<td>Superdex 200, 13 nM</td>
<td>Gel filtration</td>
</tr>
</tbody>
</table>
2.5 Procedures

2.5.1 Cloning, Expression, and Purification of the N-Terminal Domains (WD1-6)

The N-terminal metal binding domains construct with a molecular weight of 61.3 kDa, includes the amino acid residues 57-633 of Wilson protein (Figure 2.3). Due to its large size, expressing WD1-6 as a fusion protein was most desirable. This is because fusion tagged proteins tend to display increased cytoplasmic solubility and also simplify the purification process [1]. The unwanted tagged protein can be easily cleaved with an appropriate protease. To clone WD1-6, two plasmid vectors, pET41Ek/LIC and pET32Xa/LIC that encode fusion tag proteins Glutathione S-Transferase (GST) and Thioredoxin (TRX), respectively, were chosen. The "Ek" stands for the Enterokinase cleavage site, -AspAspAspAspLys- while "Xa" designates the Factor Xa cleavage site, -IleGluGlyArg-. [2]. The cleavage sites are positioned such that all vector encoded
sequences are cleaved from the target protein after digestion with protease. pET41Ek/LIC is a kanamycin resistance vector while pET32XaLIC is an ampicillin resistance vector.

The LIC in the vector notation refers to Ligation Independent Cloning (Figure 2.5), a method that facilitates the directional cloning of PCR products without use of restriction enzymes, digestion, or ligation reactions [3, 4]. LIC utilizes the 3'→5' exonuclease activity of T4 DNA polymerase to create a specific 12- or 15- base single-stranded overhangs in both the amplified DNA and cloning vector. The PCR primers are designed by incorporating appropriate 5' sequences to generate overhangs complementary to the vector's LIC cloning site. The purified PCR product is then incubated with T4 DNA polymerase to generate compatible overhangs in the presence of the 'recognition' deoxynucleotide. When mixed together, the PCR product and LIC vector anneal immediately before transformation into competent E. coli cells. Circularized plasmid in E. coli results from covalent bond formation at vector-insert junctions by endogenous DNA ligase. Purified plasmid DNA is then transformed into suitable E. coli hosts for protein expression after sequencing analysis (Protocol TBI84), Novagen [5]. A schematic of the procedure from gene amplification to plasmid DNA recovery and sequencing is presented in Figure 2.4. The WD 1-6 LIC primers were designed according to the LIC kit user protocol. The presence of the amplified gene insert was confirmed by 0.8% SDS-PAGE agarose gel; it was run alongside 100 base and 1 kilobase DNA ladders for 1.25 h at 80V. The gel was stained with a 0.5 μg/mL ethidium bromide solution for 20 min, rinsed 3X with distilled water, and DNA visualized using an Image Master® VDS (Amersham).
IVATSTVRILGMTCQSCVKSIEDRISNLKGISMKVSVLEQG SATVKYVPSV
VCLQQVCHQIGDMGFEASIAEKGAAASWPSRLPAQEAVKLRVEGMCQS
CVSSIEGKVRKLGQVRKVLSNQEAVITYQPYLIQPEDLRDHVNDMGFE
AAIKSKVAPLSLGIDLERLQSTNPKRPLSSANQNFWNSETLGHQGSHHVVT
QLRIDGMHCKSCVLNIEENIGQQLLGQVSQISVLQSVLENKTAQVKYDPSCTSPVAL
QRAIEALPPGNFKVSLPDGAEGSGTDRRSSSSHSGSPRNNQVQGTCSTTLIA
IAGMTCASCVHSIEGMIQLEGVQQISVLASEGTATVLYNPAVISPEELRAAI
EDMGFEASVVAESCSTPNLGNHSAGNSMVQTTDGTPTSLQEVAPHTGRLPA
NHAPDILAKSPQSTRAVPQKCFLQIKGMTCASCVSNIERNLQKEAGVLSV
LVALMAGKAEIKDPEVIQPLEIAQFIQDLGFEEAMEDYAGSDGNIELTIT
GMTCASCVHINESKLRGRTNGITYASVALATSKALVKFDPEIGPRDIKIEEI
G FHASLAQ  633

**KEY:**

Yellow - GMH/TCXXCV motif

Red - Copper binding site cysteines

2 - Glycine is added after thioredoxin tag cleavage with TEV protease.

Q- Tyrosine 532, Y532H mutation site

**Figure 2.3** The amino acid sequence of **WD1-6**.
The PCR product was purified with a Qiagen MiniElute PCR purification kit but the process was not sufficient to remove residual deoxyribonucleotides that interfere with T4 DNA polymerase activity. The PCR product was further purified using a gel extraction protocol (QIAquick Gel Extraction Kit). Quantification of the purified PCR product was performed either by agarose gel estimation or UV absorbance measurements, depending on the yield. The treatment of PCR product or the WD1-6 insert with T4 DNA polymerase to generate compatible overhangs was done according to the pET32XaLIC or pET41EkLIC kit instruction (Novagen). This was followed by the insert gene annealing to the vector and transformation into Novablue Gigasingles™ competent cells (Novagen). In most cases, the competent cell transformation followed a standard method; 10 to 20 nanograms of plasmid DNA were transformed for 5-10 min in pre-chilled competent cells. The cells were placed on ice for 5 min with Novablue Gigasingles™ or 30 min for other phenotypes and heat-shocked at 42°C for 15-60 sec depending on the volume of the cells in the tube. Then, they were again pre-chilled on ice for another 2 min before adding warm, sterilized Luria Bertani (LB) media or Super Optimal Growth (SOC) media [6, 7].

The transformed competent cells were either plated immediately in the case of Novablue Gigasingles™ cells or the other cell phenotypes were placed in an incubator shaking at 250 rpm set at 37°C for 1 h to recover. The cells were then plated on LB/agar media containing antibiotics, either kanamycin (30 μg/mL) or ampicillin/carbenicillin (100 μg/mL). A secondary antibiotic in the agar media was added to screen the transformed cells which carried genotypic resistance to that particular antibiotic. For example, Rosetta(DE3) E. coli cells are resistant to chlorampenicol but Novablue Gigasingles™ cells are not. The cells were incubated for 16 h at 37°C. The screening of
Gene insert amplification.

1 Ligation of gene insert into a cloning vector (ligation independent cloning).

1 Transformation of ligated plasmid into Gigasingles™ Novablue cells.

Transformed cells grown on LB/agar and appropriate antibiotic.

Colony screening: Single colonies inoculated in 5 mL of LB/antibiotic, grown at 37°C, 250 rpm for 12-16 h and cells harvested.

Plasmid DNA recovery: QIA prep Spin Miniprep kit (Qiagen).

Restriction mapping of purified plasmid DNA with BamHI and BgII.

DNA sequencing by primer extension method.

**Figure 2.4** Schematic of steps from gene amplification to plasmid DNA sequencing.
Figure 2.5 Ligation independent cloning strategy into pET41EkLIC vector.

Colonies was done by inoculating 5 mL of LB media containing the appropriate antibiotic with a single colony. The colonies were shaken for 12-16 h at 37°C and cells harvested by centrifugation. Plasmid DNA from the cell pellet was recovered using a QiaPrep Spin MiniPrep kit. It was quantified and analyzed by restriction mapping. The plasmids found to contain the WD1-6 gene insert were further sequenced by primer extension (Retrogen, Inc., San Diego CA, and also by Dr. Todd Barkman, Biological Sciences, Western Michigan University).

The WD1-6 purification summary is shown in Figure 2.6. WD1-6 protein was first expressed by transforming either pET41Ek/WD1-6 or pET32Xa/WD1-6 plasmid into Rosetta(DE3) competent cells. The transformation procedure was the same as described in the above paragraph except that DMSO was added to certain competent cells.
Transformation of pET32Xa/WD1-6 plasmid into Rosetta(DE3) expression cells.

Transformed cells grown on LB/agar and appropriate antibiotic, single colonies inoculated in 5 mL LB/antibiotic.

After a 5 h growth at 37°C, 250 rpm, 5 mL cultures scaled up to 0.1-1 L of LB/antibiotic culture, cells induced after OD600nm=0.6 with 0.1-3 mM IPTG. Cells induced at 28-30°C, 250 rpm and harvested after 4-6 h.

Protein extracted by freeze/thaw (4 rounds) and BugBuster® reagent (simultaneously). Supernatant recovered.

Protein purified by Immobilized Metal Affinity Column (HisPrep sepharose column). Protein purified further by gel filtration.

Cleavage of thioredoxin tag with Factor Xa or TEV protease.

Digested product is loaded onto HisPrep sepharose column, thioredoxin tag binds, and WD 1-6 passes through.

Concentration of WD 1-6 flowthrough and gel filtration of cleaved WD1-6.

Figure 2.6 The expression and purification summary of WD 1-6 protein using pET32Xa/WD 1-6 plasmid.
to increase plasmid DNA permeation. The plated cells were incubated overnight at 37°C; five colonies were inoculated in 5 mL LB culture containing the appropriate antibiotic and grown in a 250 rpm shaking incubator set at 37°C for 5-6 h. The 5 mL culture was either induced with a final concentration of ImM IPTG (Isopropyl P-D-l-thiogalactopyranoside) or transferred to 100-500 mL LB media for further growth. The cells grown in LB media greater than 5 mL were induced with IPTG after the optical density at ODeoonm reached 0.5-0.6. The concentration of the sugar inducer, IPTG, was varied to optimize expression. Protein induction was performed for 3-4 h and cells were harvested by centrifugation at 8,000xg at 4°C for 15 min. The cell pellet was stored at -20°C when protein extraction was not started immediately. The expressed fusion protein, TRXxWD1-6 (derived from pET32XaLIC vector) or GSTEkWD1-6 (derived from pET41EkLIC vector) possessed a six poly-histidine tag, (-HHHHHH-) that was utilized in the purification protocol. To isolate the TRXXWD1 fusion, the frozen cell pellet was thawed in a 50 mL falcon tube (for a freshly harvested sample, this was disregarded) and resuspended with 5 mL of BugBuster® protein extraction reagent per gram of pellet. The BugBuster® protein extraction reagent buffer releases soluble proteins by breaking and lysing the bacterial cell wall. The high viscosity due to DNA and RNA release was lowered by adding 20 KU/mL of Benzonase (Novagen) enzyme. Supernatant containing soluble proteins was separated from cell debris by spinning for 15 min at 16,000xg. The supernatant was decanted or pipetted into a clean 50 mL falcon tube ready to load on a purification column.

The HisPrep FF 16/10 is a nickel chelating sepharose column with high affinity for poly-histidine tagged proteins. The AKTA FPLC was connected to this column,
cleaned with 10 column volumes (CV) of water, and subsequently equilibrated with 10 CV of HisPrep binding buffer (20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, 20 mM imidazole). The supernatant containing TRX, WD1-6 and other proteins from the cell was loaded onto the column at a 1 mL/min flow rate. The size of supernatant that was loaded was always less or equal to 15 mL (3/4 CV). This step was followed by washing the column with HisPrep binding buffer until the UV absorbance returned to the range of baseline. The tagged bound protein was eluted either isocratically or by a gradient from the column with elution buffer (20 mM sodium phosphate buffer, pH 7.4, 500 mM NaCl, 500 mM imidazole). The phosphate elution buffer was later substituted with 50 mM HEPES, pH 7.5, 500 mM NaCl, 300 mM imidazole, as this was found to stabilize the pure protein. Before elution with HEPES buffer, about 3 CV of 30 mM imidazole in HEPES buffer was passed through the protein-bound column to remove residual phosphate buffer. Protein fractions were characterized by SDS-PAGE and UV. The protein fractions containing WD1-6 were pooled together and concentrated with an Amicon device with a 10,000 molecular weight cut-off membrane (Cat. 13622, Millipore). If the protein was pure, it was prepared for fusion protein cleavage; otherwise, the protein was subjected to gel filtration. To remove the TRX and His-tags from WD1-6, the purified protein was exchanged into Factor Xa or TEV buffer using a bench top PD-10 desalting column or the AKTA FPLC connected to the HiPrep 26/10 Desalt column. To determine the correct amount of Factor Xa, the purified protein was first quantified by the BCA protein assay. The buffer conditions for Factor Xa digestion were 50 mM Tris/Cl, pH 8.0, 100 mM NaCl, 5 mM CaCl₂ and for the Tobacco Etch Virus (TEV), 50 mM Tris/Cl, pH 8.0, 1 mM DTT, 0.5 mM EDTA. After the addition of Factor
Xa, cleavage mixture was incubated at room temperature for 16 h. TRX<sub>Xa</sub>WD1-6 cleavage with Factor Xa never went beyond trial experiments; therefore, quantifying the amount of Factor Xa needed was not done. The cleavage reaction was stopped with protease inhibitors, 4 mM Pefabloc SC PLUS (Roche); PMSF, PhenylMethylSulphonyl Fluoride (Sigma-Aldrich).

The cleaved protein was first exchanged into HisPrep binding buffer before loading onto a smaller HisPrep chelating column (CV = 1 or 5 mL) at a 0.2 mL/min flow rate using a peristaltic pump, and the flowthrough fractions that contained WD1-6 protein were collected. TRX or GST protein bearing the His-tag after cleavage was left bound on the column while WD1-6 protein passed through. The WD1-6 flowthrough was concentrated and purified further by gel filtration (Superdex 75 or 200 Sepharose columns). The purity was determined by SDS-PAGE and the yield quantified by BCA (bicinchoninic acid) protein assay [8] (Pierce, Rockford, IL).

The cleavage of WD1-6 protein with Factor Xa was found to be inefficient and accompanied by protein degradation. The Bertini group at the Center for Magnetic Resonance, Florence, Italy had produced a WD1-6 construct containing residues 1-633 of WNDP incorporating the protease Tobacco Etch Virus (TEV) cleavage site [9]. Gateway technology (Invitrogen) was utilized to incorporate TEV cleavage site, -GluAsnLeuTryPheGlnGly- [10-12]. In this thesis, a method is described to incorporate the TEV cleavage site directly into pET32Xa/WD1-6 as illustrated in Figure 2.7. This was achieved by substituting all the amino acids of the Factor Xa protease site and another 3 amino acids upstream of the site with TEV cleavage site in a three step mutagenesis reaction involving three sets of primers: TEV PI, TEV PII, TEV PIII and
their complementary primers (Table 2.3). The resulting plasmid construct, named pET32TEV/WD1-6 was sequenced to confirm the incorporation of TEV cleavage site. It was not known how the introduction of new amino acids (TEV cleavage site) would affect the protein expression and stability. To maintain the pET32XaLIC vector sequence, a new WD1XaTEVLIC forward primer (Table 2.3) beginning with TEV cleavage site instead of Domain 1 of the Wilson protein was designed. Using pET32TEV/WD1-6 plasmid as the DNA template, the WD1-6 gene insert from the TEV cleavage site was amplified with KOD HiFidelity Hot Start DNA polymerase. The insert was then ligated into the pET32XaLIC Vector by LIC. This resulting plasmid was named pET32XaTEV/WDI-6 and contained both Factor Xa and TEV cleavage sites (Figure 2.7). The plasmid was sequenced and transformed, and the protein expressed and purified as described in Figure 2.6. Purified WD1-6 was stored at -80°C either with protease inhibitors added or stabilizing agents such as glycerol. Various modifications and changes in procedures are also discussed under results and discussion in Chapter 3.

2.5.2 Cloning and Purification of WD 5-6, ATOX1 and WD 5-6(Y532H)

The structural and functional properties of the isolated WD5-6 of human Wilson protein had been studied by members of our lab [13]. The gene inserts encoding WD5-6 and ATOX1 proteins had been previously sub-cloned by the restriction enzyme cleavage and ligation method into pET24d+ and pET11d (Novagen), respectively. Purification of WD5-6 and ATOX1 followed established protocols [13]. Mutagenic primers, KD-Y532H Fwd and GC KD-Y532H Rev (Table 2.3), coding for substitution of tyrosine with histidine were used to incorporate the Y532H mutation into the native WD5-6 gene.
**Figure 2.7** The TEV cleavage site's insertion into the pET32Xa/WD1-6 plasmid.
Site-directed mutagenesis (Stratagene) was performed using the manufacturer's instructions. Subsequent expression studies under typical conditions resulted in high level expression of WD5-6(Y532H) mutant protein as inclusion bodies (insoluble aggregate). Therefore, the WD5-6(Y532H) mutant was cloned into a fusion vector. The WD5-6 gene with Y532H mutation was amplified with mutagenic primers for subsequent cloning into pET32Xa/LIC vector using LIC (Section 2.5.1). The gene insert coding for WD5-6 and WD5-6(Y532H) mutant contains 148 amino acids, residues 485-633 of WNDP. The sub cloned plasmids were transformed into *E. coli* Rosetta(DE3) cells (Novagen) for protein expression.

### 2.5.2.1 Protein Expression

The cells were grown in the presence of kanamycin (30 jxg/mL) for the pET24d/WD5-6 plasmid and ampicillin/carbenicillin (100 jg/mL) for the pET32Xa/WD5-6(Y532H) and pET1ld/ATOX1 plasmids. Protein expression was induced with 0.5-1 mM IPTG (final concentration). For WD5-6 protein isolation, the cells were harvested after 4 h of induction. The protein was extracted into 20 mM MES/Na, 0.1 mM EDTA, 1 mM DTT, pH 6.0 buffer by four rounds of freezing and thawing in liquid nitrogen. The supernatant was applied to a DEAE-Sepharose (Diethylaminoethyl cellulose) cation exchange column. The non-binding proteins were stripped from the column with 20 mM MES/Na, pH 6.0, 0.1 mM EDTA, and 10 mM DTT (low salt buffer). Elution was performed by a NaCl gradient; WD5-6 eluted at -0.16 M NaCl. The protein containing fractions were identified by SDS-PAGE and concentrated by an Amicon device fitted with 3,000 molecular cut-off membrane (Cat.
PLBC02510, Millipore) to 2-3 mL. The concentrate was then loaded on a HiLoad 26/60 Superdex 75 column (Table 2.7) and eluted with 20 mM MES/Na, pH 6.0, and 150 mM NaCl, and 10 mM DTT.

2.5.2.2 Inclusion Body Extraction

For isolation of WD5-6(Y532H) mutant protein (Figure 2.8), TRX<sub>Xa</sub>WD5-6(Y532H) fusion protein was extracted from the cell pellet using the 5-10 mL BugBuster® protein extraction reagent per gram of cell pellet. Using this method, very little soluble protein was recovered due to inclusion body formation. To increase the yield, the inclusion bodies were processed according to BugBuster® protein extraction reagent user protocol for inclusion bodies (Novagen). The pellet obtained from the initial extraction was resuspended again in the same amount of BugBuster® reagent used in that first initial step and resuspended to form a uniform mixture. Recombinant lysozyme (Novagen) or chicken lysozyme (Sigma-Aldrich) was added to a final concentration of 1KU/mL and incubated for 5 min at room temperature. Then, two vol of BugBuster® reagent (1 vol equals the amount of BugBuster® reagent added in the initial protein extraction step) was added. The resuspension was mixed by mild vortexing for 1 min, and inclusion bodies were recovered by centrifugation at 5,000xg for 15 min at 4°C. The inclusion bodies were resuspended in 2 vol BugBuster® reagent, mixed by vortexing and inclusion bodies were again obtained by centrifugation. This step was repeated twice. The inclusion bodies were finally resuspended in 1 vol of BugBuster® reagent then centrifuged at 16,000xg for 15 min at 4°C.
2.5.2.3 Mutant Protein Refolding Procedure

The inclusion bodies so obtained were solubilized in a 50 mL polycarbonate tube (Nalgene) with 6 M GnHCl in HisPrep binding buffer, gently mixed in an ice bath, and a supernatant was obtained by centrifugation at 16,000xg for 30 min at 4°C. The supernatant containing denatured fusion protein was slowly loaded onto HisPrep FF 16/10 column. The protein was refolded in situ by washing off GnHCl with 10 CV of HisPrep binding buffer (20 mM Phosphate, pH 7.4, 500 mM NaCl, and 20 mM imidazole). The elution was performed isocratically with 5 CV of 500 mM imidazole in HisPrep binding buffer.

2.5.2.4 Thioredoxin Fusion Protein Cleavage

The protein fractions were verified by SDS-PAGE, concentrated, and exchanged into Factor Xa cleavage buffer (50 mM Tris/Cl, pH 8.0, 100 mM NaCl, 5mM CaCl\(_2\)) using a PD-10 column (GE). The protein was cleaved at 37°C for 16 h by adding 0.5 U/mg of Bovine Factor Xa. Time-course cleavage products were analyzed by SDS-PAGE. A significant percentage of the fusion protein precipitated out during and after buffer exchange, requiring centrifugation to remove the insoluble debris. After completion of digestion, Factor Xa activity was quenched by addition of a serine protease inhibitor (4 mM Pefabloc SC PLUS), or, alternately, the protease was captured by Xarrest agarose matrix (Novagen).

To separate thioredoxin from mutant WD5-6(Y532H), the cleavage product was passed through a 5 mL nickel chelating HisPrep column using a small peristaltic pump (Pump P-1, GE). The flowthrough containing cleaved WD5-6(Y532H) was
concentrated to about 2-3 mL before loading onto the HiLoad 26/60 Superdex 75 column. The protein was eluted with 20 mM MES/Na, pH 6.0, 150 mM NaCl, 10 mM DTT, at a flow rate of 2.6 mL/min and a 5.2 mL fraction size. Fractions that contained pure protein, as determined by SDS-PAGE analysis, were combined, concentrated to a total volume of 3 mL, and protease inhibitors added before storage.

2.5.2.5 Modifications of WD5-6(Y532H) Protein Expression and Purification

The purified refolded WD5-6(Y532H) mutant was found to be unstable and readily degraded to a product of two protein fragments within a week or so. To overcome this challenge, the purification procedure for WD5-6(Y532H) was modified in the following way (Figure 2.8). The pET24d/WD5-6(Y532H) plasmid containing a non-fusion WD5-6(Y532H) mutant gene was transformed into Rosetta(DE3) cells according to established protocols. Freshly grown colonies from prepared kanamycin LB/agar plates were inoculated into 5 mL LB and grown for 4-5 h at 250 rpm at 37°C. Each of the 5 mL aliquots were transferred to a 100 mL LB containing 30 ng/mL of kanamycin and grown for one hour before transferring to 1 L of LB. The 1 L cultures were grown until the cell optical density at OD600nm reached 0.5-0.6. Two 1-L flasks placed in an incubator at 37°C were induced with different amounts of IPTG, 1 mM and 3 mM. After 30 min, the flask induced with 3 mM IPTG was placed in a water bath at 28-30°C, and shaken at 200 rpm while the other flask (with 1mM IPTG) remained in the shaking incubator with the temperature lowered to 30°C. The cells were harvested after 5 h and pellets stored at -80°C.
The extraction of WD5-6(Y532H) protein was initiated by slow thawing of the frozen cell pellet. Then, five rounds of freezing and thawing in liquid nitrogen were performed. A freeze-thaw buffer (20 mM MES/Na, pH 6.0, 1 mM EDTA, 10 mM DTT) was added in the third round. The pellet was resuspended in freeze-thaw buffer (10 mL/g pellet), then benzonase (25 U/mL) and Pefabloc SC PLUS protease inhibitor (1 mM) were added. The resuspension was incubated on an orbital shaker (Bellco Biotech, Vineland, NJ) for 15 min at room temperature and 80 rpm while waiting for the viscosity to decrease. The cell resuspension was then placed into ice cold water and sonicated (F50 Sonicator Dismembrator, Fisher Scientific) for 10 sec, 10 times at 30 sec time interval (pulse 12). The supernatant (called S) was obtained by centrifugation at 16,000xg, and 4°C for 30 min.

The DEAE Sepharose column (CV = 60 mL) was prepared before use by stripping unwanted protein with 4 CV of 1M NaCl, rinsing with 5 CV of water, and equilibrating with 5 CV of low salt buffer (20 mM MES/Na, pH 6.0, 0.1 mM EDTA, 10 mM DTT). The supernatant was directly applied to the column at 1 mL/min and all unbound protein was eluted with low salt buffer. The WD5-6(Y532H) protein was eluted by a 1M NaCl gradient. The fractions containing WD5-6(Y532H) were pooled and concentrated. The protein was further purified by size exclusion chromatography using the HiLoad 16/60 Superdex 200 column. The Superdex column was cleaned with filtered milli-Q water and equilibrated with elution buffer, 50 mM HEPES, pH 7.5, 200 mM NaCl, and 1% glycerol. The protein was eluted at 2 mL/min for the first 20 mL of elution, and then the flow rate was reduced to 1 mL/min for the rest of the elution. Eluted protein was recovered in 2 mL fractions.
The fractions containing the pure WD5-6(Y532H) were identified by both SDS-PAGE analysis and UV absorbance analysis. The protein was concentrated with an Amicon device equipped with a 3,000 molecular weight cut-off filter. The concentrate was stored at -80°C.

Figure 2.8 Recovery of WD5-6(Y532H) protein using two different routes.

2.5.3 Tobacco Etch Virus (TEV) Protease Purification

TEV protease cleaves protein with high specificity and efficiency, targeting a 7 amino acid cleavage site (Section 2.5.1). The purification of TEV protease was done entirely in our laboratory, except for a small quantity obtained from Invitrogen for
standardization purposes. The TEV-encoding plasmid was kindly supplied by Dr. Tim Cross at Florida State University. The TEV plasmid was transformed into BL21(DE3) *E. coli* cells (Novagen) according to previously described protocol. The cells were grown on LB/agar containing 100 µg/mL carbenicillin. Three to five colonies were inoculated in 5 mL of LB culture containing 100 µg/mL carbenicillin and grown in a 250 rpm shaking incubator at 37°C for 5 h. One of the culture tubes was transferred to 100 mL LB media, grown for one hour, transferred to 1000 mL of LB with antibiotic added, and grown until the *OD600nm* reached 0.5-0.6. The TEV induction was initiated by adding 1mM IPTG final concentration and the media was transferred to a water bath shaker at 25°C. The induction proceeded for 4-6 h (*Figure 2.9*). Cells were harvested by centrifugation and resuspended in 20 mM Tris/Cl, pH 7.9, 500 mM NaCl, 5 mM imidazole buffer, before storage at -20°C.

To purify the TEV protease bearing a poly-histidine tag (*Figure 2.10*), the frozen cells were thawed slowly. The cell pellet underwent four rounds of freezing and thawing in liquid nitrogen. Afterwards, benzonase (25 U/mL) was added to reduce viscosity. The cells were sonicated (F50 Sonicator Dismembrator, Fisher Scientific) 7 times for 10 sec at 30 sec interval (Pulse 12) while placed on ice cold water before centrifuging at 16,000xg for 15 min to recover the supernatant. The supernatant was loaded onto a nickel chelating HisPrep column equilibrated with Tris/Cl HisPrep binding buffer (20 mM Tris/Cl, pH 7.9, 500 mM NaCl, 30 mM imidazole), which was used to remove the unbound protein contaminants. The TEV protease was eluted with 20 mM Tris/Cl, pH 7.9, 500 mM NaCl, and 300 mM imidazole. The protein was concentrated to approximately 5 mL; half of this concentrate was exchanged into TEV cleavage buffer.
(50 mM Tris/Cl, pH 8.0, 0.5 mM EDTA, 1 mM DTT) before storing at -80°C in 50%
glycerol, 1 mM EDTA and 5 mM DTT. Although the intention of exchanging the buffer
was to remove imidazole, it was found not to interfere with the TEV activity at 20-30 (JM
levels. The time between TEV protease elution and storage was very critical because the
protein showed a tendency to precipitate out of solution with time.

The activity of the TEV protease was quantified using a known amount of
TRXαaTEvWDl-6 in a time-course digestion reaction since the activity of TEV varied
from batch to batch. During time-course digestion trials, samples were incubated up to 48
h with protein samples taken out periodically. The activity of purified TEV protease was
calculated as the amount in micromoles of the TRXαaTEvWDl-6 protein cleaved by one
micromole (μmole) of TEV in 16 h at room temperature (25°C) in buffer conditions, 50
mM Tris/Cl, pH 8.0, 0.5 mM EDTA, and 1mM DTT. The large scale cleavage reaction
was also set up in the same conditions. The TEV protease has an added advantage over
Factor Xa protease because of its high specificity [14]. The TEV protease, since it has a
poly-histidine tag, was removed together with the His-tagged thioredoxin from the
cleaved product, WD 1-6, with a nickel chelating HisPrep column.
Figure 2.9 Time-course induction of TEV protease with a molecular weight of 29 kDa indicated by the arrow.

<table>
<thead>
<tr>
<th>Time (h)</th>
<th>0</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38 k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31 k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24 k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17 k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 k</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 2.10 The TEV protease amino acid sequence. The poly-histidine tag residues are in positions two to seven.

<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>GHHHHHHGES LFKGRDYNP ISSSICHTLN ESDGHTTSLY GIGFGPFIIT NKHFLRRNNG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70</td>
<td>80</td>
<td>90</td>
<td>100</td>
<td>110</td>
<td>120</td>
</tr>
<tr>
<td>TLLVQSLHGV FKVDTTTLQ QHLVGRDME TIIIRMETPK DFPPFPQKLK FREPQREERI</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>130</td>
<td>140</td>
<td>150</td>
<td>160</td>
<td>170</td>
<td>180</td>
</tr>
<tr>
<td>CLVTRNFQPK SMETSSMETV SOTSCTPSS DGIFWKHIQ TKDGQCSPG VSTRDGFIVG</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>190</td>
<td>200</td>
<td>210</td>
<td>220</td>
<td>230</td>
<td>240</td>
</tr>
<tr>
<td>INHASNFTNT NNYFTSVPNK FMETELLTNQ EAQQWVSNGW LNADSVLWGG HKVFMTVVKP</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>260</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EEPFQPVKEA TQLMETNELV YSQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
2.6 Other Proteins of Importance to this Work

2.6.1 Cloning of COMMD1

The finding that COMMD1 also interacts with WNDP [15] cultivated interest in conducting further experiments between this protein and the isolated WD1-6 multidomains. The COMMD1 cDNA was acquired from ORIGENE Technologies, Rockville, MD. Mutagenic primers bearing Bam I and Nco I restriction sites, COMMD1 BamHI Fwd and COMMD1 Nco I Rev (Table 2.3), respectively, were designed with DNA Strider 1.3 software. The thermocycling conditions and the mixing of PCR components were performed according to HotStar HiFidelity PCR handbook (Qiagen). In a PCR tube, the COMMD1 gene template, primers, dNTPs mix, HotStar DNA polymerase buffer and nuclease-free water were mixed and amplified with HotStar DNA polymerase (Qiagen). The amplified DNA was visualized by agarose gel electrophoresis. The COMMD1 amplified insert was purified by gel extraction and resuspended in water.

The ligation of the COMMD1 insert into pET24d vector was performed using the Fast-Link™ DNA ligation kit (Epicenter Biotechnologies, Madison, WI). The pET24d and COMMD1 insert were first digested with restriction enzymes BamHI and Nco I in BamHI buffer and cleaned up using the gel extraction method. The reaction mixture was incubated at room temperature for 5 min before transferring to 70°C for 15 min to inactivate Fast-Link™ DNA ligase. Cooled ligated DNA was transformed into Novablue Gigasingles™ cells and colonies were screened after growth on LB/agar and kanamycin (30ng/mL). The resulting plasmid DNA from screened colonies was analyzed by restriction enzyme mapping for the insert with BamHI and Nco I. The plasmid DNA
containing the insert was sequenced at Retrogen, Inc., San Diego, CA. The pET24d/COMMD1 plasmid was then transformed into Rosetta(DE3) cells and protein expression was induced with 1-5 mM IPTG (final concentration). The expression of COMMD1 protein was not successful.

2.6.2 Cloning of G591D, a Wilson Disease-Causing Mutation

The WD5-6(G591D) mutant gene was subcloned into the pET41EkLIC vector resulting to pET41Ek/WD5-6(G591D) construct by Patrick Ochieng, a former student in our lab. Due to the similarity of this construct and the WD5-6(Y532H) mutant, a study comparing the two mutations was deemed valuable to this work. The protease enterokinase was not able to cleave the GST fusion protein efficiently, rendering the pET41Ek/WD5-6(G591D) construct undesirable. A new construct subcloned in a pET32XaLIC vector was prepared utilizing the same methodology used for domains WD1-6 and WD5-6(Y532H) plasmid vector constructs (Sections 2.5.1 and 2.5.2). The resulting plasmid was transformed into Rosetta(DE3) cells, then grown on LB/agar containing 30 |g/mL kanamycin. The resulting colonies were inoculated in 5mL medium, scaled up, and the protein induction was performed with 1 mM final IPTG concentration. The expressed protein formed inclusion bodies; the purification and processing followed the refolding procedure of WD5-6(Y532H) mutant protein (Section 2.5.2). Cleavage of thioredoxin from WD5-6(G591D) mutant protein with Factor Xa was not successful.
2.7 Biophysical Characterization and Analysis

2.7.1 Sodium Dodecyl Sulfate-Polyacrylamide Gel Electrophoresis

The SDS-PAGE method analysis relies entirely on the migratory behavior of proteins in a polyacrylamide gel based on their molecular sizes. Both 15% and 10% Tricine SDS resolving gels were prepared from a 30% Acrylamide (w/v), 0.8% Bis-acrylamide (w/v) solution. The separating and stacking gels were prepared as described in the molecular cloning laboratory manual [16]. The mixture was degassed for 10 min under vacuum to remove air bubbles before adding ammonium persulfate and TEMED (N,N,N',N'-tetramethylethylenediamine) to initiate cross-linking. The mixture was poured into a 0.75 mm thickness gel pre-cast mold, and the top overlaid with water-saturated n-butanol to form a uniform and level surface for the stacking gel overlay. The separating gel was left for 45-60 min to polymerize; the n-butanol was washed with distilled water and the unit dried with pressurized air. The stacking gel comb was inserted before pouring the degassed stacking gel solution [16] mixed with ammonium persulfate and TEMED. It took 30-45 min for the stacking gel to fully polymerize. The gels were then dismantled before the assembly of the electrophoresis apparatus (Mighty Small II, SE250/SE260, Amersham).

Protein samples for gel electrophoresis were prepared by mixing 6X gel loading buffer (300 mM Tris/Cl, 600 mM DTT, 12% SDS (w/v), 0.06% bromophenol blue (w/v), 60% glycerol (v/v)) with the appropriate protein solution and water to make a 1X solution, boiled at 92°C for 10 min for cell pellet samples and 3 min for pure proteins. The cathodic buffer (0.5M Tris/Cl, pH 8.25, 0.5% SDS) was added in the separating gel.
chamber and anodic buffer (1M Tris/Cl, pH 8.9) in the outer chamber. Samples were loaded alongside protein molecular standards; 5-15 pL was loaded per well. The samples were run at 95V for 1 hour and 15 min. The gels were incubated with staining solution (53% water (v/v), 40% methanol (v/v), 7% glacial acetic acid (v/v) and 0.25% Coomassie Brilliant Blue R-250 (w/v)) for 30 min on a table orbital shaker at 80 rpm or microwaved (Model R-215EW, SHARP, Osaka, Japan) for 50 sec. The excess dye was destained by the same solution without Coomassie blue. The gels were fixed and dried using the cellophane MiniGel drying kit (Invitrogen).

2.7.2 Gel Filtration

Gel filtration elution of a protein is dependent on the sizes of both the matrix pores and protein size [17]. Large proteins do not enter the matrix and therefore elute in the void volume, $V_0$, for they move at the same flow rate as the eluting buffer. Small proteins, having full access to the pores, elute nearly at the end of one total column volume, $V_t$. In between these two extremes, proteins of different sizes are eluted according to their elution volume, $V_e$, although elution of a sample is characterized best by the distribution coefficient, Kd, which is defined as the fraction of the stationary phase that is available for diffusion of a given molecular species. Though comparable to Kd, the partition coffiecient, $K_{av}$, is used for protein comparison and is calculated by Equation 2.1. Since $K_{av}$ is related to the size of the protein, plotting the logarithm of molecular weight against $K_{av}$ of known protein standards gives a calibration curve that can used to estimate the unknown molecular weight of a protein. Equation 2.1 works well for
proteins of similar shape and conformation like globular proteins. This method is also an excellent tool to study protein-protein dimerization.

\[ V_v - V_0 \]

The protein sample was first quantified by the BCA protein assay, and the concentration calculated. The gel filtration analysis was done with an FLPC AKTA system fitted with a Superdex 200 10/300 GL column (for WD1-6) or Superdex 75HR 10/300 for (for WD5-6 and WD5-6(Y532H)) and equilibrated with several CV of 50 mM sodium phosphate, pH 7.5, 150 mM NaCl, 10 mM DTT. A 25 pL sample of each the protein standards with a concentration of 1-5 mg/mL was injected onto column via a loop and eluted with the same buffer at 0.5 mL/min. The amount of protein used per run was approximately 25-125 p.g. The void volume, \( V_0 \), was determined with Blue dextran and the elution volume, \( V_e \), was the same as the column volume. The calibration curve was obtained by a plot of logarithm of molecular weights of the protein standards against the corresponding \( K_{av} \).

### 2.7.3 Light Scattering

Light scattering works on the principle that a high intensity, monochromatic and collimated laser light impending on solution sample will be scattered in a manner dependent on the properties of the solute. Light scattering can be used to determine the molar mass, \( M_w \); radius of gyration, \( r_g \); second virial coefficient \( A_2 \), and translational diffusion coefficient \( D \), of the solute in solution. Static light scattering (SLS) or time averaged fluctuations in the intensity of scattered light due to molecular motions are used
to calculate $M_w$, $r_g$ and $A_2$. The Rayleigh-Gans-Debye derived model equations are used to compute these parameters [18]. Dynamic Light Scattering (DLS) measure the fluctuations in the intensity of scattered light at the nanosec time scale [19]. The scattered light intensity depends on the particle size ($r_g$, radius of gyration) and the magnitude of the scattering wave vector $q$ [20], Equation 2.2, where $k$ is radiation wavelength, $\theta$ is scattering angle, and $q$ is refractive index of the solvent.

$$q = \frac{4\pi \eta \sin(\theta/2)}{\lambda}$$  \hspace{1cm} [2.2]

The most important measurement obtained from a DLS measurement is the translational diffusion coefficient, $D$ (units of $\text{cm}^2/\text{sec}$), which is related to the hydrodynamic radius $R_H$ through the Stokes-Einstein equation, Equation 2.3, where $k$ is the Boltzmann constant, $D$ the translational diffusion coefficient, $T$ the temperature and $\eta$ the solvent viscosity.

$$D = \frac{kT}{6\pi \eta R_H}$$  \hspace{1cm} [2.3]

DLS-like gel filtration is also used to determine the unknown molar mass of a protein when calibrated with protein standards of known molecular weights [21].

The Light Scattering experiments were done at the Keck Foundation Biotechnology Resource Laboratory at Yale University. The light scattering data was collected using a Superdex 200 (for WD1-6) or 75 (for WD5-6 and WD5-6(Y532H)) 10/300, HR Size Exclusion Chromatography (SEC) column (GE Healthcare), connected to an Agilent 1200 HPLC, (Agilent Technologies, Wilmington, DE) equipped with an
autosampler. The elution from SEC was monitored by a photodiode array (PDA) UV/VIS detector (Agilent Technologies), differential refractometer (RI) (optilab Wyatt Corp., Santa Barbara, CA), and a static and dynamic, multi-angle, laser light scattering (LS) detector (HELEOS II with QELS capability, Wyatt Corp., Santa Barbara, CA). The SEC-UV/LS/RI system was equilibrated with 20 mM HEPES, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM DTT buffer at the flow rate of 1 mL/min (WD1-6) or 0.5 mL/min (WD5-6 and WD5-6 (Y532H)). Two software packages were used for data collection and analysis: Chemstation software (Agilent Technologies, Wilmington, DE) controlled the HPLC operation and data collection from the multi-wavelength UV/VIS detector, while the ASTRA software (Wyatt Corp., Santa Barbara, CA) collected the data from the RI detector, the LS detectors, and recorded the UV trace at 280 nm sent to the PDA detector. The weight average molecular masses, Mw, were determined across the entire elution profile in the intervals of 1 sec from static LS measurement using ASTRA software as previously described [22]. The Hydrodynamic radii, RH, were calculated from "on-line" dynamic LS measurements taken every 2 sec. The dynamic light scattering signal was analyzed by the methods of cumulants [23].

### 2.7.4 Circular Dichroism

Circular dichroism (CD) spectroscopy utilizes the ability of the asymmetric chiral active macromolecules to absorb circularly polarized light. CD is a useful tool for characterization of secondary structure and unfolding properties of proteins and polypeptides. The spectra of proteins in the far UV produce unique structural fingerprints dependent on the geometry of the polypeptide backbone. This is because these spectra
arise from $ji^*-k^*$ transitions of the amide bonds [24]. For example, $a$-helical proteins have spectra distinct from those of all $\beta$-sheet structures. Circularly polarized light is separated into two vectors with equal strength with one rotating clockwise ($ER$) and the other counterclockwise ($EL$). As the asymmetric molecules interact with light, they absorb left and right handed circular polarized light differently. The plane of the light is therefore rotated and the summation of $ER$ and $EL$ produce a single vector that is ellipsoid giving the light a property of being elliptically polarized. The CD is thus reported in degrees ellipticity or the difference in the absorbance of the $ER$ and $EL$ by the asymmetric molecule [25].

The raw data from CD spectra in millidegrees was converted into mean residue ellipticity, $[\theta]$ (deg-cm$^2$-decimol$^{-1}$), by Equation 2.4 [26] with the parameters: $\theta$, millidegrees; $M_w$, molecular weight of the protein in Dalton; $c$, protein concentration in mg/ml; $l$, cell pathlength in cm; $N_A$, number of amino acid residues of the protein.

$$[\theta] = \frac{100 \times M_w}{c \times l \times N_A}$$ [2.4]

The protein samples for CD spectra measurements were purified by gel filtration chromatography followed by exchange into 50 mM Na phosphate, pH 7.5. The protein concentration was determined by the BCA assay (Pierce). Urea and GnHCl stock solutions of 8 M were prepared. To accurately prepare these 8 M GnHCl stock solution, the mass of the tube was predetermined or tared over weighing balance and refractive index used to validate concentration. The salt was weighed into the tube and 50 mM Na phosphate buffer added until the desired 50 mL volume was reached.
CD spectra were acquired with a JASCO J-815 CD Spectropolarimeter. Before acquisition, the 0.1 cm path-length cell was cleaned and rinsed with buffer. The protein samples for measurement were prepared in two final concentrations, 14.5 pM and 29 ^M, depending on the pretrial CD run. The protein unfolding studies were done by preparing samples of the same concentration but with GnHCl added to 1 to 6 M final concentrations. The protein samples were incubated for at least an hour to equilibrate before acquiring data. The samples were passed through a Millex®-GV 0.22 µm filter unit (SLGV004SL, Millipore) prior to CD spectral analysis. The CD measurements were carried out with a 0.1 cm path-length rectangular cell. The thickness of the cell can be obtained by measuring the transmittance of the empty cell with air as the reference. An interference pattern is produced when transmitted radiation interact with radiation reflected off from the two windows of the cell. The thickness $t$ of the cell in centimeters is obtained from the Equation 2.5, where, $t$ is the cell thickness in centimeters and $N$, the number of peaks between two wave numbers, $\sigma_1$ and $\sigma_2$.

$$t = \frac{N}{2(\sigma_1 - \sigma_2)}$$  \[2.5\]

The cell was pre-rinsed with sample solution and filled to ~300 fxL with the same sample. The baseline or the control was the buffer without the protein. Water blanks were also measured. The CD Spectropolarimeter, purged with nitrogen gas at 15 PSI per min was turned on and allowed to warm up for 15 min. The scan rate was 50 nm/min and an average of three runs was recorded for each sample. Temperature was controlled by a single position Peltier Thermostated Cell holder, JASCO model PTC-423, cooled with a small water pump.
The acquired data was computed, processed and tabulated using the instrument's RIS Spectra manager software (JASCO). The secondary structure analysis was performed with CDpro software (http://lamar.colostate.edu/~sreeram/CDPro/main.html) [27-30]. The protein unfolding analysis was based on established phenomena in which denaturants alter the equilibrium between the native (folded) and denatured (unfolded) states of protein [31-34]. The Gibbs free energy of unfolding was calculated by the Linear Extrapolation Method [35], starting with by Equation 2.6, where, AG, Gibbs free energy; R, the gas constant; ln, natural logarithm; U, unfolded protein fraction; F, folded protein fraction.

$$\text{AG} = -RT\ln \frac{v^1}{v^2}$$  \hspace{1cm} [2.6]

The relative concentrations of F and U were derived from CD spectra data, assuming the protein was fully folded in the absence of denaturant. A two-state folding model was used to determine the denatured protein fraction, Fj, using Equation 2.7, where, Yobs is the observed variable parameter, and Yn and Yj are the values of Y characteristic of the native and denatured protein [34].

$$\text{Fd} = \frac{Y_n - Y_{\text{obs}}}{Y_n - Y_b}$$  \hspace{1cm} [2.7]

The free energy of unfolding between the folded and unfolded state of the protein AG, is calculated by Equation 2.8, where, T is the absolute temperature.
The least squares analysis method was used to obtain the free energy of unfolding, \( AG^{H_2O} \), in the absence of denaturant. This value is related to free energy according to Equation 2.9, where, \( AG \), Gibbs free energy; \( m \), measure of the dependence of \( AG \) on denaturant concentration, \([D]\), denaturant concentration.

\[
\Delta G = \Delta G^{H_2O} - m[D]
\]  

2.7.5 Nuclear Magnetic Resonance Spectroscopy

The NMR technique is used for determination of protein structure and probing of protein-protein interaction dynamics. Unlike X-ray crystallography, proteins in NMR are studied in their aqueous form. NMR requires either a singly or doubly isotopically labeled protein sample. To make isotopic labeled samples, cells were grown on \(^{13}\)C glucose as a carbon source and on \(^{15}\)N ammonium chloride as a nitrogen source. The protein-encoding plasmid was transformed into Rosetta(DE3) cells, grown on LB/agar containing 100 (ig/mL carbenicillin and colonies were inoculated into the media. To express labeled protein, cells were grown in a minimal media solution.

To make a minimal media solution, 1 L of phosphate buffer (6 g Na2HPO4, 3 g KH2PO4, and 0.5 g NaCl) was prepared with ultrapure water, conductivity, 18.2 MQ/cm (US Filter PURELAB Plus UV/UF, Millipore) in a clean 2.8 L flask, the pH was adjusted to 7.2 with 0.1 M sodium hydroxide, and the buffer autoclaved. In a sterile environment
or near a flame, 2 or 4 g of (13C) glucose and 1 g of (15N) ammonium chloride were added. The following components were also added per L of the media with a syringe fitted with an Acrodisc® syringe (PN 4612, Pall Life Sciences, Ann Arbor, MI) 0.22 filter: 2.0 mL of 1 M MgSO4, 2 mL of 100 mM CaCl2, 5.0 mL of MEM Vitamin mix and 1 mL of the appropriate antibiotic (100 ng/mL ampicillin/carbencillin or 35 (ig/mL kanamycin). Metal salts supplements were occasionally added [16]. The cells were first grown in 5 mL LB for 6-8 h and then centrifuged in a sterile 50 mL tube at 4500xg at 37°C for 10 min. The supernatant was discarded and the cells were resuspended in 10 mL of the minimal media. The cells were grown in a 250 rpm shaking incubator at 37°C until the OD600nm reached between 0.6-1.0 (this took 8-10 h) and protein expression was induced with 1-3 mM IPTG final concentration. The induction time varied with temperature; for protein expressed at 37°C, time ranged from 3-5 h while for 25°C, it was 4-6 h range. The cells were harvested and protein purified in the same way described in Section 2.5.2.

The NMR protein sample was prepared under N2 in a Vac Atmospheres® Nexus chamber; the sample was first exchanged with oxygen-depleted NMR buffer (100 mM sodium phosphate, pH 7.2) using a PD-10 column and sample fraction containing protein were identified with Coomasie dye (Pierce, Rockford, IL). The protein sample was concentrated to about 600-800 μL, then D2O was added to 10% final concentration (v/v) and protease inhibitor (1 mM Pefabloc SC PLUS final concentration) added. In some cases, the protein sample was first treated with 10 mM DTT to reduce disulfide bonds. When samples were prepared in Kalamazoo, Michigan, a small amount (5-10 μL) was taken for concentration determination, and the remainder of the prepared NMR samples
was shipped to the Center for Magnetic Resonance, Florence, Italy using an ice-cold box. Samples were also prepared on site by the author during a 3 month visit to CERM in Fall of 2007. Dr. Shenlin Wang (at the time a student at CERM) and Prof. David Huffman (on sabbatical at CERM 2007-2008) assisted in the NMR spectral assignment.

For both WD5-6(Y532H) mutant and WD5-6 wild type, the NMR spectra for backbone assignments were obtained through standard triple resonance experiments [36]. The backbone assignment experiments of WD5-6(Y532H) were performed on 0.2 mM protein samples at 500 MHz (298K), Bruker® using standard pulse sequences.

Variations in the position of signals between WD5-6(Y532H) and wild type WD5-6 were quantified through combined chemical shift variation (CSV). This was calculated from the experimental 'H and 15N chemical shift variations (A5 ('H) and A5 (15N), respectively), using Equation 2.10.

\[
CSV = \sqrt{\frac{(\Delta \delta^{(1}H))^{2} + (\Delta \delta^{(15}N)_5)^2}{2}}
\]  

[2.10]

15N-RI, 15N-R2 and steady-state heteronuclear 'H-^N NOEs were measured at 600 MHz (298K). Processing of NMR data was accomplished with the XWINNMR program. 15N RI and 15N R2 parameters were determined by fitting the integration values obtained from the XWINNMR program, measured as a function of the delay within the pulse sequence, to a single-exponential decay. The heteronuclear NOE values were obtained from the ratio of the peak height for 'H-saturated and unsaturated spectra.


5. Xa/LIC Cloning Kits, User protocol TB184 (EMD Chemicals Inc.).


32. Myers, J. K., Pace, C. N., & Scholtz, J. M. "Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding" (1995) Protein Sci. 4, 2138-2148.


34. Pace, C. N. "Measuring and increasing protein stability" (1990) Trends Biotechnol. 8, 93-98.

CHAPTER 3

CHARACTERIZATION OF THE N-TERMINAL DOMAINS OF THE HUMAN WILSON PROTEIN

3.1 Introduction

The importance of the N-terminal domains in the function of the Cu-ATPase ATP7B is evident from several mutational studies [1-3]. For example, deletion of MBDs leads to copper sensitivity, and an increase in the rate of phosphorylation of the Wilson protein. MBDs 5-6, the metal binding domains closest to the membrane and transduction channel, when deleted, lead to abrogation of the Cu-ATPase activity [4]. The characterization of the complete N-terminal domains of the Wilson protein has been attempted by various groups [3, 5-8] with mixed success. It is evident from these studies that the protein's instability due to susceptibility to degradation and aggregation of the purified multi-domain construct continues to hinder further studies, particularly with longer procedures. Research which overcomes these challenges and elucidates the unknown properties not only of the N-terminal domains, but also of the Wilson protein in its entirety will be a significant contribution to this field. In this chapter, cloning, purification and characterization of the N-terminal domains will be reported and discussed. Major advancements in the study of the N-terminal domains of Wilson protein are highlighted.
3.2 Cloning of pET32Xa/WD1-6 and pET41Ek/WD1-6 Plasmids

The successful cloning of pET32Xa/WD1-6 and pET41Ek/WD1-6 plasmids involved many trial reactions under varying conditions. The WD1-6 gene is approximately 1731 bp in length when amplified with two sets of LIC primers, WD1 EkLIC Forward and WD6 EkLIC Reverse; WD1 XaLIC Foward and WD6 XaLIC Reverse (Table 2.3). The primers were designed to have a slightly higher melting temperature than the extension temperature of the PCR thermocycler. Other factors that were considered in the primer design were guanine-cytosine content, length, intra- and inter- primer interactions, and low sequence similarity to the non-target sites of the template. The choice of the DNA polymerase was found to be crucial, especially in the amplification of large pieces of DNA. The working environment was kept clean; the PCR tubes, pipette tips and milli-Q water (US Filter PURELAB Plus UV/UF, Millipore) for use were autoclaved (Vacuum Steam Sterilizer, M/C 3533 Castle, Rochester, NY).

3.3 PCR Reactions and Results

PCR reaction conditions utilized to amplify the gene encoding WD1-6 gene is shown in Tables 3.1 and 3.3 from Wilson protein cDNA. The PCR conditions are outlined in corresponding Tables 3.2 and 3.4. The pET41EkLIC vector target WD1-6 gene amplification was achieved by addition of an additive (1 U per 100 ng of DNA template), Perfect Match Enhancer (Stratagene, Cedar Creek, TX). The amplified WD1-6 gene from both reactions was visualized in a 0.8% agarose gel stained with 0.5 ng/mL of ethidium bromide. A 0.8% agarose gel picture of amplified WD1-6 gene for subsequent cloning into pET41EkLIC vector with 1 Kb DNA ladder (NEB) is shown in Figure 3.1.
Table 3.1 The PCR reaction of WD 1-6 gene for subsequent cloning into pET41 EkLIC vector.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>69.33</td>
</tr>
<tr>
<td>10X Thermopolymerase buffer (NEB)</td>
<td>10.0</td>
</tr>
<tr>
<td>dNTPs mix (2 mM)</td>
<td>10.0</td>
</tr>
<tr>
<td>Wilson protein cDNA (1.6 µg/µL)</td>
<td>6.25</td>
</tr>
<tr>
<td>WD1 EkLIC Forward (10 mM)</td>
<td>1.78</td>
</tr>
<tr>
<td>WD6 EkLIC Reverse (10 mM)</td>
<td>1.64</td>
</tr>
<tr>
<td>Deep Vent Polymerase (exo-)</td>
<td>2.0</td>
</tr>
<tr>
<td>Total</td>
<td>101.0</td>
</tr>
</tbody>
</table>

Table 3.2 The PCR thermocycler settings of WD 1-6 gene for subsequent cloning into pET41 EkLIC vector.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature (°C)</th>
<th>Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Activation cycle</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>Extension</td>
<td>72</td>
</tr>
<tr>
<td>3</td>
<td>Melting</td>
<td>95</td>
</tr>
<tr>
<td>4</td>
<td>Annealing</td>
<td>60</td>
</tr>
<tr>
<td>5</td>
<td>Repeat step 2 to 4 for 30 cycles</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Extension</td>
<td>72</td>
</tr>
<tr>
<td>7</td>
<td>End/store</td>
<td>4</td>
</tr>
</tbody>
</table>
Table 3.3 The PCR reaction of WD 1-6 gene for subsequent cloning into pET32XaLIC vector.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>33.3</td>
</tr>
<tr>
<td>10X KOD Hot Start polymerase buffer (Novagen)</td>
<td>5.0</td>
</tr>
<tr>
<td>dNTPs mix (2 mM)</td>
<td>5.0</td>
</tr>
<tr>
<td>Wilson protein gene (10.0 μg/μL)</td>
<td>1.0</td>
</tr>
<tr>
<td>WD1 XaLIC Forward (10 mM)</td>
<td>0.9</td>
</tr>
<tr>
<td>WD6 XaLIC Reverse (10 mM)</td>
<td>0.8</td>
</tr>
<tr>
<td>KOD Hot Start DNA polymerase (Novagen)</td>
<td>1.0</td>
</tr>
<tr>
<td>Total</td>
<td>25.0</td>
</tr>
</tbody>
</table>

Table 3.4 The PCR thermocycler settings of WD 1-6 gene for subsequent cloning into pET32XaLIC vector.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature (°C)</th>
<th>Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Activation cycle</td>
<td>95</td>
</tr>
<tr>
<td>2</td>
<td>Extension</td>
<td>70</td>
</tr>
<tr>
<td>3</td>
<td>Melting</td>
<td>95</td>
</tr>
<tr>
<td>4</td>
<td>Annealing</td>
<td>67</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>End/store</td>
<td>4</td>
</tr>
</tbody>
</table>

Repeat steps 2 to 4 for 25 cycles
**Figure 3.1** A 0.8% agarose gel picture of the amplified WD1-6 gene. The -1.8 Kb WD1-6 gene product, visualized here along with 1 Kb DNA ladder (NEB) was subsequently cloned into pET41EkLIC vector using the LIC method.
3.3.1 Ligation Independent Cloning (LIC)

The amplified WD1-6 gene inserts were incorporated into target vectors, pET32XaLIC and pET41EkLIC by the LIC method. This cloning technique involved three main stages: (1) PCR product purification by agarose gel extraction; (2) treatment of target insert with T4 DNA polymerase to generate compatible overhangs to vector; and (3) annealing of the insert gene to the vector. The agarose gel purification removes residual deoxynucleotides, DNA polymerases, and any contaminating DNA that may interfere with T4 DNA polymerase activity in the generation of overhangs. The cloning region of pET32XaLIC and pET41EkLIC vectors is shown in Figures 3.2 and 3.3, respectively.

The treatment of the WD1-6 insert with T4 DNA polymerase to generate compatible overhangs for both target vectors was done as shown in Tables 3.5 and 3.6. The reaction mixture was incubated for 30 min at 22°C followed by incubation for 20 min at 75°C to inactivate the T4 DNA polymerase. The treated WD1-6 insert was annealed to the vectors as shown in Table 3.7 by mixing the plasmid vector, treated WD1-6 insert, and EDTA, and then incubating for 5 min at 22°C. The ligated plasmid DNA recovered from colony screening was analyzed by restriction mapping. Both vectors possesses Bgl II and BamH I restriction sites (see Figures 3.2 and 3.3) that lie on either end of the LIC cloning site and therefore, can be used to map the presence of the insert. Table 3.8 shows the reaction protocol for the restriction enzyme mapping of pET32Xa/WD1-6, also known as double digests and shown in Figure 3.4 are the results.
**Figure 3.2** pET32XaLIC vector cloning region (Novagen, Inc.).

**Figure 3.3** pET41 EkLIC vector cloning region (Novagen, Inc.).
Table 3.5 The WD 1-6 gene insert treatment for incorporation into pET41 EkLIC vector.

<table>
<thead>
<tr>
<th>Volume (μL)</th>
<th>Volume (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD 1-6 insert (-30 ng/μL) in TE buffer</td>
<td>7.0</td>
</tr>
<tr>
<td>10X T4 DNA polymerase buffer</td>
<td>2.0</td>
</tr>
<tr>
<td>25 mM dATP</td>
<td>2.0</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>1.0</td>
</tr>
<tr>
<td>T4 DNA polymerase (2.5U/μL)</td>
<td>0.4</td>
</tr>
<tr>
<td>Water</td>
<td>8.4</td>
</tr>
<tr>
<td>Total</td>
<td>20.0</td>
</tr>
</tbody>
</table>

* TE Buffer; 50 mM Tris/Cl, 0.1M EDTA, pH 8.0.

Table 3.6 The WD 1-6 gene insert treatment for pET32XaLIC vector cloning.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD 1-6 insert (-50 ng/μL) in TE buffer</td>
<td>4.0</td>
</tr>
<tr>
<td>10X T4 DNA polymerase buffer</td>
<td>2.0</td>
</tr>
<tr>
<td>25 mM dGTP</td>
<td>2.0</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>1.0</td>
</tr>
<tr>
<td>T4 DNA polymerase (2.5U/μL)</td>
<td>0.4</td>
</tr>
<tr>
<td>Water</td>
<td>11.6</td>
</tr>
<tr>
<td>Total</td>
<td>20.0</td>
</tr>
</tbody>
</table>
Table 3.7 LIC-compatible WD 1-6 insert annealing into LIC vectors.

<table>
<thead>
<tr>
<th>Step</th>
<th>Component</th>
<th>Volume (HL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>pET41EklC or pET32XaLIC vector</td>
<td>1.0</td>
</tr>
<tr>
<td>2</td>
<td>T4 DNA polymerase treated WD 1-6 insert</td>
<td>2.0</td>
</tr>
<tr>
<td>3</td>
<td>Mixed and incubated for 5 min at 22°C</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>25 mM EDTA</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>Incubated for further 5 min at 22°C</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>Transformation into Novablue Gigasingles™ cells</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>Colony screening, plasmid purification, and restriction enzyme mapping</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.8 Double digestion of pET32Xa/WD1-6 plasmid.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (jL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water</td>
<td>13.0</td>
</tr>
<tr>
<td>10 X BamHI buffer</td>
<td>2.0</td>
</tr>
<tr>
<td>BamHI (10U/μL)(^b)</td>
<td>0.5</td>
</tr>
<tr>
<td>Bgl II (10 U/μL)(^b)</td>
<td>0.5</td>
</tr>
<tr>
<td>Total</td>
<td>20.0</td>
</tr>
</tbody>
</table>

Incubated for 2 h at 37°C followed by agarose gel electrophoresis

\(^b\) One unit (U) is defined as the amount of the enzyme required to digest 1 μg of X DNA in 1 h at 37°C in a total reaction of 50 jL.
Figure 3.4 The restriction enzyme mapping of pET32Xa/WDl-6 visualized in a 0.8% agarose gel.
3.3.2 Construction of pET32Xa/WD1-6 Plasmid

The amplification of WD1-6 gene by PCR (1731 bp) was found to be dependent on several factors, including the source of DNA polymerases, additives, history of deoxynucleotide mix, and temperature settings of the PCR thermocycler. Among several DNA polymerases used, the Deep Vent and Deep Vent (exo-) DNA polymerase (NEB, MA), HotStar HiFidelity polymerase (Qiagen), and KOD Hot Start DNA polymerase (Novagen) were able to amplify the WD1-6 gene, but not without errors. Sequenced WD1-6 inserts amplified with KOD Hot Start polymerase and Deep Vent (exo-) polymerase for insertion into the pET41 EkLIC vector revealed introduction of several missense mutations scattered throughout the sequence. The HotStar HiFidelity polymerase had been successful in amplifying a 449 bp COMMD1 gene [9]. Further tests of WD1-6 amplification for insertion into the pET41 EkLIC vector were performed with the HotStar HiFidelity polymerase. Exploration of the pET41Ek/WD1-6 construct was discontinued after the reaction for amplification failed (in Table 3.4) to work in the absence of solution Q*, an additive supplied with the polymerase kit (Qiagen). All polymerases except KOD Hot start polymerase required the presence of a PCR additive, either Perfect Match Enhancer (Stratagene) or solution Q supplied with the HotStar HiFidelity polymerase kit (Qiagen).

The pET32Xa/WD1-6 construct prepared from both HotStar HiFidelity polymerase and KOD Hot Start polymerase amplification also resulted in some mutations. Deletions, insertion, and missense mutations were introduced during amplification. Despite the presence of mutations and truncations, the protein from

* An innovative PCR additive that facilitates amplification of difficult templates by modifying the melting behavior of DNA (HotStar HiFidelity PCR Handbook 09/2005).
pET32Xa/WD1-6 expressed at higher level than that of pET41Ek/WD1-6. The mutations of pET32Xa/WD1-6 were well spaced, either far from each other or separated by one or two base pairs, and a Quickchange II XL site-directed mutagenesis kit (Stratagene) was used to correct them following the manufacturer's instruction. In total, 13 mutations were corrected, 1 deletion, 1 insertion and 11 missense mutations. The primers used for these corrections are listed in Table 2.4 in Section 2.3. For unknown reasons, the corrected pET32Xa/WD1-6 mutation-free construct contained a number of silent mutation changes that were not in the original Wilson protein cDNA obtained from Washington University at St. Louis.

Since the construct's mRNA sequence can affect protein expression, retaining the original Wilson protein cDNA sequence was considered important. Therefore, additional amplifications of the gene for WD1-6 were performed under a variety of conditions. KOD Hot Start DNA polymerase (Novagen) amplified the WD1-6 gene with high efficiency and fidelity under newly optimized conditions shown in Tables 3.3 and 3.4. This insert was cloned into the pET32XaLIC vector and adapted for subsequent work. It was found that KOD Hot Start DNA polymerase was sensitive to both temperature and the number of PCR cycles.

The amplification of the gene fragment WD1-6 using various DNA polymerases was difficult. Deep Vent, Deep Vent (exo-), and Hot Star DNA polymerases isolated from Pyrococcus strain have fairly slow elongation rates in comparison to Thermococcus kodakaraensis derived KOD Hot Start DNA polymerase [10], and this could be a factor considering that the former cannot amplify this large sequence without a PCR additive.
The PCR temperature settings, pure and fresh dNTPs, PCR additives, and a PCR thermocycler with heated lid affect the amplification fidelity.

3.4 Strategy for Incorporating TEV Cleavage Site

The protease cleavage of a fusion protein with Factor Xa can result in incomplete cleavage or cleavage at other sites. Therefore, a construct encoding Tobacco Etch Virus (TEV) cleavage site was prepared, pET32XaTEV/WD1-6 (Figure 3.5). TEV protease cleaves more efficiently than Factor Xa resulting in higher yield. The TEV restriction site can be introduced into an expression vector using the Gateway Entry vector pENTR/TEV/D-TOPO (Invitrogen) [5, 11] followed by Gateway recombination (Invitrogen). Although the Gateway system is rapidly gaining popularity, it is an expensive, fairly complex procedure, and the TEV cleavage from this system leaves four new encoded amino acid residues at the N-terminal of the purified protein. Since the native sequence was desired, a different method was used and is outlined below.

3.4.1 Method of Incorporating TEV Protease Cleavage Site

The incorporation of a TEV encoding site into pET32Xa/WD1-6 was performed using two sequential steps, site-directed mutagenesis followed by LIC. Three sets of primers and complements, TEV PI, TEV PII, and TEV PIII (see Table 2.3) were prepared to encode the TEV cleavage site and complementary overlapping sequences encoding the Factor Xa cleavage site (Figure 3.5) and the first 3 amino acid residues downstream. The primers were designed for optimum annealing to the template while
Figure 3.5 Thioredoxin tagged WD 1-6 protein and the location of the protease cleavage sites (A) Factor Xa (B) TEV (Tobacco Etch Virus).
minimizing intra- and inter-primer dimer formation [12]. The synthesis of the new plasmid encoding the TEV cleavage site, termed pET32TEV/WDI-6, was performed stepwise with 3 mutagenic reactions using a QuickChange II XL site-directed mutagenesis kit (examples are shown in Tables 3.9 and 3.10).

Modifications were made to place the TEV site immediately downstream of the Factor Xa site. The newly constructed TEV/WD1-6 gene segment in pET32TEV/WDI-6 was amplified and cloned back into pET32XaLIC vector. The WD1-6 insert containing the TEV cleavage site was amplified using WD1 XaTEVLIC Forward and WD6 XaLIC Reverse primers (Table 2.3) with pET32TEV/WD1-6 plasmid as the template. The PCR product, treated to produce LIC compatible overhangs, was ligated into pET32XaLIC to produce pET32XaTEV/WD1-6, encoding both Factor Xa and TEV cleavage sites. Figure 3.6 shows the restriction enzyme mapping of pET32XaLIC, pET32Xa/WD1-6, and pET32XaTEV/WD1-6 using BamHI and Bgl II endonucleases.

Table 3.9 Cycling parameters of QuickChange II XL site-directed mutagenesis method.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycle</th>
<th>Temperature (°C)</th>
<th>Time (Sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>12-18</td>
<td>95</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td>60/Kb</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>68</td>
<td>420</td>
</tr>
</tbody>
</table>
Table 3.10 The synthesis of the pET32XaTEVII/WDl-6 plasmid from pET32XaTEVI/WDl-6 template. This is the second mutagenesis reaction in the sequence of three.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (ML)</th>
<th>Final mass (ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Water</td>
<td>37.4</td>
<td></td>
</tr>
<tr>
<td>2 10X reaction Pfu polymerase buffer</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>3 pET32XaLICTEVI/WD 1-6 (-6.5 ng/L)</td>
<td>1.5</td>
<td>10.0</td>
</tr>
<tr>
<td>4 TEV PII Forward primer (10 μM)</td>
<td>1.04</td>
<td>125.0</td>
</tr>
<tr>
<td>5 GC TEV PII Reverse primer (10 μM)</td>
<td>1.06</td>
<td>125.0</td>
</tr>
<tr>
<td>6 DNTPs</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>7 QuikSolution (an additive)</td>
<td>3.0</td>
<td></td>
</tr>
<tr>
<td>8 PfuUltra HF DNA polymerase</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>9 Total reaction volume</td>
<td>50.0</td>
<td></td>
</tr>
<tr>
<td>10 New plasmid synthesized following parameters in Table 3.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Parental DNA digested with Dpn I restriction enzyme for 2 h at 37°C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 New plasmid transformation, colony screening, and DNA purification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Primer extension DNA sequencing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 3.6 Double digest (restriction enzyme mapping) of plasmid constructs using BamHI and Bgl II. Lane 1, pET32Xa/WDl-6; Lane 2, pET32XaTEV/WDl-6; Lane 3, pET32XaLIC vector; Lane 4, 1 Kb Ladder (NEB). The product was visualized in a 0.8% agarose gel. The size of pET32XaLIC vector is due to migratory behavior of linearized DNA.
3.4.2 Discussion

The incorporation of a TEV cleavage site in expression vectors has gained popularity, especially in the last decade [13]. The reason TEV protease is a valuable reagent is its ability to function under a variety of conditions and its specificity compared to Factor Xa, enterokinase, and thrombin proteases. Several factors need to be considered when incorporating TEV cleavage sites and these include new amino acid residues and the preservation of the native sequence. In some cases the introduction of a TEV cleavage site has decreased the solubility of target protein, limiting its usage [13].

3.5 The Expression and Purification of WD1-6 Protein

The use of several different WD1-6 expression vectors helped to optimize expression and purification of the N-terminal domains. The protein expression and purification methods were discussed in Section 2.5.1. TRXxₜ WD1-6 expressed from pET32Xa/WD1-6 containing truncations is shown in Figures 3.7, 3.8 and 3.9. Missense mutations found in the pET41Ek/WD1-6 vector led to poor yield of expressed protein, Figure 3.10. Further trials with this DNA plasmid construct were discontinued.

Correction of all mutations (Section 3.3.2) led to high yield of expressed protein (Figure 3.11). Mutations repair were confirmed by DNA sequencing (Retrogen, Inc.). The induction of TRXₓₜEvWD1-6 and TRXₜEvWD1-6 proteins (see Table 3.11 for further information) is shown in Figure 3.12. Subsequent steps in the purification and thioredoxin cleavage of TRXₓₜ WD1-6 and TRXₓₜEvWD1-6 are discussed in the following sections.
Figure 3.7 The truncated TRXxaWDl-6 protein (~35 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WDl-6. Arrow indicates protein of interest.

Figure 3.8 The truncated TRXxaWDl-6 protein (~20 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WDl-6. Arrow indicates protein of interest.
**Figure 3.9** The truncated TRX$_{Xa}$WD1-6 protein (-40 kDa) expressed in Rosetta(DE3) cells transformed with pET32Xa/WD1-6. Arrow indicates protein of interest.

**Figure 3.10** Expression of GST$_{Ek}$WD1-6 protein (-92 kDa) in Rosetta(DE3) cells transformed with pET41Ek/WD1-6. Arrow indicates protein of interest.
Table 3.11 Notation for the expressed WD1-6 fusion protein.

<table>
<thead>
<tr>
<th>Expressed protein</th>
<th>Fusion protein</th>
<th>Cleavage site present</th>
</tr>
</thead>
<tbody>
<tr>
<td>TRXx&lt;sub&gt;a&lt;/sub&gt;WD1-6</td>
<td>Thioredoxin</td>
<td>Factor Xa</td>
</tr>
<tr>
<td>TRXTEVWD1-6</td>
<td>Thioredoxin</td>
<td>TEV</td>
</tr>
<tr>
<td>TRXxaTEVWD1-6</td>
<td>Thioredoxin</td>
<td>Factor Xa/TEV</td>
</tr>
</tbody>
</table>

Figure 3.11 The 10% Tricine SDS-PAGE gel showing TRX<sub>x</sub>WD1-6 induction after mutation repairs. Lane 1 (0 h), control, Lanes 2 (1 h) and 3 (4 h), expressed truncated TRX<sub>x</sub>aWD1-6, Lane 4 (4 h), TRX<sub>x</sub>aWD1-6 after deletion mutation repair, and Lane 5 (4 h), TRX<sub>x</sub>aWD1-6 after insertion mutation repair.
Figure 3.12 A 10% Tricine SDS-PAGE gel showing the TRX_{XAT}E_{VWDL}-6 and TRX_{TEVWDL} time-course induction. Lanes 1-2, control, Lanes 2-6, TRX_{XAT}E_{VWDL}-6, Lanes 7-9, TRX_{TEVWDL}-6, and M, full range rainbow protein marker. The arrow indicates the protein of interest.

3.5.1 Immobilized Metal Affinity Chromatography (IMAC) Purification

Both TRX_{Xa}WDl-6 and TRX_{XaTEVWDL} contain a poly-histidine tag upstream of the N-terminal end of WD 1-6 and downstream of the thioredoxin fusion protein. This poly-histidine tag was used in the initial purification of the cell lysate utilizing immobilized metal-affinity chromatography (IMAC) [14, 15]; nickel ion was the chelating agent in a Sepharose 6 Fast Flow HisPrep FF 16/10 column. The cell lysate (supernatant) was applied to the column. Bound protein was washed with HisPrep binding buffer (20 mM Na Phosphate, pH 7.4, 500 mM NaCl, 20 mM imidazole) to remove contaminants and then equilibrated with 50 mM HEPES, pH 7.5, 500 mM NaCl, 30 mM imidazole. The protein elution was isocratically eluted with 50 mM HEPES, pH
7.5, 500 mM NaCl, 300 mM imidazole. The HEPES buffer was found to stabilize WD1-6 relative to phosphate buffer. The $\text{TRX}_{Xa}\text{TEVWD1-6}$ isocratically eluted protein is shown in Figure 3.13.

### 3.5.2 Size Exclusion Chromatography

Initial IMAC chromatography did not remove all impurities. The protein was further purified by size exclusion chromatography (gel filtration) by using either a 320 mL Superdex 75 26/60 column or a 120 mL Superdex 200 16/60 column. The flow rate for the initial one sixth of the elution was set slightly higher (4X faster) than the rest of the elution. The $\text{TRX}_{Xa}\text{TEVWD1-6}$ protein eluted at -65.6 mL, whereas cleaved WD1-6 eluted at -71.05 mL for the 120 mL Superdex 200 16/60 column (Figures 3.14 and 3.16). The conditions for the elution were 50 mM HEPES, pH 7.5, 200 mM NaCl, and 1% glycerol, 2 mL/min flow for the initial 20 mL, and then 0.5 mL/min for the rest of elution.

Thioredoxin fusion cleavage with TEV protease is shown in Figure 3.15 and was complete after 24 h under the conditions noted. The cleavage of thioredoxin fusion protein with Factor Xa was unsuccessful (not shown) largely due to low yield and rapid degradation of the purified $\text{TRX}_{Xa}\text{WD1-6}$ protein construct, most likely due to the identity of the amino acid immediately after the protease site.
Figure 3.13 TRXxATEvWDl-6 protein elution from a 20 mL HisPrep column. Lanes 1-8 shows fractions sampled by SDS-PAGE with corresponding elution volume at the bottom of the gel. CL in Lane 10 is the crude cell lysate (BugBuster® protein extraction reagent supernatant).
Figure 3.14 Size exclusion chromatography of TRXxaTEvWD1–6. Lanes 1–5 shows elution fractions (below) and the chromatogram profile (top) for Superdex 200 16/60 column. Solid black bars indicate protein trace.
Figure 3.15 Time-course cleavage of thioredoxin from TRX$_{X_a}$TEvWDl-6 with TEV protease (1pg TEV/100 pg fusion protein).
Figure 3.16 Size exclusion chromatography of WD 1-6. The elution and chromatogram profile for Superdex 200 16/60 column is shown on top plot. Lanes 1-5 contain sample fractions assayed and their corresponding elution volume indicated at the bottom of the gel. Solid black bars indicate the protein trace.
3.5.3 Discussion

The purification of WD1-6 of Wilson protein has been a hurdle for many research groups. The Rosenzweig group found that the solubility of the N-terminal domains (1-633 amino acids), including the cytosolic tail (molar mass -67 kDa) improved if cells were induced at 16°C for 20 h [3]. DiDonato and co-workers found that - 70% of the protein ended up in the induced cells as inclusion bodies. To recover the expressed protein, the inclusion bodies were solubilized and the denatured protein refolded [7]. The Bertini group [5] was able to purify WD1-6 (1-633 amino acids) but the method used [16] introduced four additional vector-encoded amino acid residues at the N-terminal end. The effects of these "extra" amino acids relative to the native WD1-6 have not been determined. Since the first 56 amino acids of WNDP are predicted to be unstructured, we chose not to include them in our WD1-6 construct. This probably led to the higher degree of stability we observed.

In this work, the N-terminal region of the Wilson protein lacking the first 56 amino acids, a total molar mass of 61.3 kDa, has been purified to homogeneity utilizing improved methods and optimized conditions. The yield and solubility of the expressed fusion WD1-6 protein was found to increase tremendously (3-5 times) when the culture was induced with 3 mM IPTG and at temperature between 28-30°C. The optimal time for induction of protein expression was 4-5 h. Protein extraction was initially done by freezing and thawing the cell pellet 3-5 times in liquid nitrogen before resuspension in BugBuster® protein extraction reagent. The amount of TRXx₅WD1-6 or TRXxaTEvWD1-6 protein released into solution increased with incorporation of this step. The highest yield that was achieved for TRXxaTEvWD1-6 was 55 mg fusion protein
per liter of LB culture, corresponding to 10 mg after cleavage. Although Rosenzweig and co-workers were successful in cleaving WD 1-6 fusion protein with Factor Xa, it did not work well with the construct used in this study. Instead, the purified fusion protein readily degraded and to an extent accelerated by Factor Xa.

Buffer components and storage conditions were found to be important in purification of WD 1-6. Both the fusion and purified WD 1-6 protein were more stable in Tris and HEPES buffers supplemented with 1% glycerol than in phosphate buffer. This was evaluated from the formation of precipitates at higher concentrations and degradation during purification, thioredoxin cleavage and storage. Tris and HEPES, bearing amine and hydroxy 1 groups, might be stabilizing the solvated WD 1-6 protein better than the phosphate group. In spite of precipitation and time-dependent degradation of pure WD1-6 protein, it can be concluded that this protein entity is sufficiently stable for biophysical and biochemical characterization.

3.6 The Biophysical Characterization

3.6.1 Gel Filtration Results and Discussion

High Resolution Gel Filtration Chromatography was utilized. A 25 (iL protein solution of 0.2-1.0 |ig/iL was injected into Superdex 200 10/300 GL high performance column (CV, 24 mL). Protein standards were obtained from both Sigma-Aldrich and GE life sciences. The gel filtration results are shown in Table 3.12 and in Figure 3.17. Proteins fractionate according to their hydrodynamic volume in gel filtration. Proteins with small hydrodynamic volume take longer than large proteins to elute due to a higher
equilibration rate into pores of the medium matrix (the stationary phase). This technique is useful in protein purification, determination of oligomeric state of complexes, and molecular weight determination. Using the calibration curve from known protein standards (Figure 3.17), the molecular weight of WD1-6 was determined to be -134 kDa. WD1-6 protein was expected to elute in the same range as Bovine Serum Albumin which has a molecular weight of 66 kDa. WD5-6, a 15 kDa segment of the N-terminal region, also elutes faster than expected. The difference between the estimated and actual molar mass of WD5-6 was relatively small (40%) compared to WD1-6 (103.2%).

Although the movement of proteins in a size exclusion column depends on hydrodynamic volume, proteins with the same physical hydrodynamic volume but dissimilar shapes will migrate differently [17]. The molecular mass of WD1-6 could not be determined with precision by size exclusion chromatography; therefore, further experiments using static light scattering were performed. To further probe the shape of WD1-6 in relative to proteins of similar mass, its hydrodynamic radius was determined by dynamic light scattering [18].

### 3.6.2 Light Scattering Results

Static and dynamic light scattering were at performed at Yale University's Keck Biotechnology Resource Laboratory. The molar mass and hydrodynamic radius of WD1-6 was calculated using ASTRA software. Figure 3.18 shows the molar mass distribution across the elution peaks of Alcohol Dehydrogenase (ADH), Bovine Serum Albumin (BSA) and WD1-6. The autocorrelation function or time dependence of diffusive motion, and derived translational diffusion coefficient for WD1-6 is shown in Figure 3.19.
Table 3.12 The molecular weights of protein standards used to plot calibration curve. The calculated molecular weights of the N-terminal domains, **WD 1–6** and **WD 5–6** are shown in bold face type.

<table>
<thead>
<tr>
<th>Protein</th>
<th>MW (Da)</th>
<th>$V_0$</th>
<th>$V_e$</th>
<th>$v_t$</th>
<th>$V_e-V_r$</th>
<th>$v_r-v_r$</th>
<th>$K^d$</th>
<th>log M.W</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blue dextran</td>
<td></td>
<td>8.63</td>
<td></td>
<td>24</td>
<td></td>
<td></td>
<td>15.37</td>
<td></td>
</tr>
<tr>
<td>BSA</td>
<td>66000</td>
<td>8.63</td>
<td>15.04</td>
<td>24</td>
<td>6.41</td>
<td>15.37</td>
<td>0.42</td>
<td>11.10</td>
</tr>
<tr>
<td>CAD</td>
<td>29000</td>
<td>8.63</td>
<td>17.82</td>
<td>24</td>
<td>9.19</td>
<td>15.37</td>
<td>0.60</td>
<td>10.28</td>
</tr>
<tr>
<td>CytC</td>
<td>12400</td>
<td>8.63</td>
<td>19.16</td>
<td>24</td>
<td>10.53</td>
<td>15.37</td>
<td>0.69</td>
<td>9.43</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>44000</td>
<td>8.63</td>
<td>16.31</td>
<td>24</td>
<td>7.68</td>
<td>15.37</td>
<td>0.50</td>
<td>10.69</td>
</tr>
<tr>
<td>Conalbumin</td>
<td>75000</td>
<td>8.63</td>
<td>15.18</td>
<td>24</td>
<td>6.55</td>
<td>15.37</td>
<td>0.43</td>
<td>11.23</td>
</tr>
<tr>
<td>Aldolase</td>
<td>158000</td>
<td>8.63</td>
<td>13.70</td>
<td>24</td>
<td>5.07</td>
<td>15.37</td>
<td>0.33</td>
<td>11.97</td>
</tr>
<tr>
<td>Ferritin</td>
<td>440000</td>
<td>8.63</td>
<td>11.57</td>
<td>24</td>
<td>2.94</td>
<td>15.37</td>
<td>0.19</td>
<td>12.99</td>
</tr>
<tr>
<td><strong>WD 1-6</strong></td>
<td>134194</td>
<td>8.63</td>
<td>13.97</td>
<td>24</td>
<td>5.34</td>
<td>15.37</td>
<td>0.35</td>
<td>11.81</td>
</tr>
<tr>
<td><strong>WD5-6</strong></td>
<td>23609</td>
<td>8.63</td>
<td>17.81</td>
<td>24</td>
<td>9.18</td>
<td>15.37</td>
<td>0.60</td>
<td>10.07</td>
</tr>
</tbody>
</table>

$d$ $K_{av} = (V_e-V_0)/(V_e V_0)$.
Figure 3.17 Gel filtration calibration curve. The known molecular weight protein standards are in black while the N-terminal domains of Wilson protein, WD1-6 and WD5-6 are displayed in red.
The hydrodynamic radius distribution of WD 1-6 and a BSA standard is shown in Figure 3.20.

The anomalous elution of WD1-6 in size exclusion chromatography relative to BSA can be explained by comparing their hydrodynamic radius, R_H (Figure 3.20). The BSA (molar mass of 66 kDa), with a hydrodynamic radius of 3.4 nm, elutes after WD1-6 (Figures 3.17 and 3.20). WD1-6 has a larger hydrodynamic radius, and migrates faster through the column than a compact globular protein. The static LS results also indicate the shape of WD1-6 is nonspherical and polydisperse. This is evident from the slanted molar mass and RH distribution across the elution peak of the protein. The interdomain linkers are known to be flexible such that the domains can move relative to each other. From the gel filtration data, it is possible to state that the six domains are somewhat extended.

3.6.3 Circular Dichroism

The unfolding of WD 1-6 as a function of GnHCl denaturant concentration, measured from 270 to 205 nm, is shown in Figure 3.21. The change in ellipticity at 222 nm was used to monitor unfolding of WD 1-6, as shown in Figure 3.22. The Gibbs free energy of unfolding with denaturant was calculated using a two-state folding model (Table 3.15 and Figure 3.23). The calculation of AG° energy, m, and [GnHCl]i/2 values for WD 1-6 was done using the Linear Extrapolation Method [19, 20]. Results are tabulated in Table 3.16.

The conformational stability of a protein can be expressed as the free energy change that occurs when a protein at equilibrium and under favorable conditions
Figure 3.18 The molar mass distribution of WD 1-6 determined by SLS. A plot of molar mass (g/mol) against elution time in min (gel filtration), monitored by a UV detector at 280 nm is shown. The "fragmented" lines labeled A-C are weight average MW for each slice, i.e., measured every 1 s. A: Alcohol Dehydrogenase, ADH (blue), B: WD 1-6, (red), C: Bovine Serum Albumin, BSA (green). The slanted distribution of MW across the eluting peak of WD 1-6 suggests non-homogeneity of the sample while early elution implies nonspherical shape of the protein.
Figure 3.19 The autocorrelation function of WD1-6. Time dependence of diffusive motion, also referred to as the intensity autocorrelation function, was determined by "on line" DLS measurements at an angle of 100° with a 1 sec collection time. Time resolved scatter intensity fluctuations were analyzed using ASTRA Software (Wyatt Corp., Santa Barbara, CA) implementing the cumulants method [21]. The translational diffusion coefficient derived from this data, shown in the graph, is used to calculate R_h using the Einstein-Stokes equation.

\[
Dt = (5.97 \pm 0.14) \times 10^{-7} \text{ cm}^2/\text{sec}
\]
Figure 3.20 A plot of hydrodynamic radius against elution time in min (gel filtration) is shown for WD1-6. BSA, a globular protein with known $R_h$ is used as a standard. The "square dot" lines, A-B, represents the $R_h$ distribution within the eluting peak. A: WD1-6 (Brown), B: BSA (Magenta). The slanted $R_h$ distribution is indicative of a polydisperse material in the elution peak. The $R_h$ for a monodisperse material remains relatively constant across the elution peak.
Table 3.13 Molar mass of WD1-6 by static light scattering. WD16 exhibits polydispersity; therefore, the molar mass given is a weighted average.

<table>
<thead>
<tr>
<th>Protein</th>
<th>Molar mass by SLS g/mol</th>
<th>Actual molar mass g/mol</th>
<th>Dispersity</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD 1-6</td>
<td>-61,570</td>
<td>61,378</td>
<td>Polydisperse</td>
</tr>
<tr>
<td>BSA</td>
<td>66,000</td>
<td>66,000</td>
<td>Monodisperse</td>
</tr>
</tbody>
</table>

Table 3.14 Hydrodynamic radius of WD1-6. Although WD1-6 has a similar mass to BSA, it has a larger hydrodynamic radius than BSA. Therefore, WD1-6 is nonspherical in shape and it can migrate faster than BSA in a gel filtration column. * Molar mass varies due to polydispersity of the protein.

<table>
<thead>
<tr>
<th>Sample</th>
<th>RH (nm)</th>
<th>Molar mass (g/mol)</th>
<th>Apparent shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD 1-6</td>
<td>4.09±0.09</td>
<td>-65,000*</td>
<td>Nonspherical</td>
</tr>
<tr>
<td>BSA</td>
<td>3.4</td>
<td>66,000</td>
<td>Spherical</td>
</tr>
</tbody>
</table>

reversibly changes from a folded to an unfolded state [22]. The $AG/^{20}$ (free energy change in absence of denaturant) can be obtained from urea and GnHCl denaturation curves. Most globular proteins have a conformational stability in the range of 21-63 kJ/mole, much weaker than the strength of a chemical bond [23]. The common techniques used to monitor folded and unfolded conformations of proteins are UV absorbance, fluorescence, and circular dichroism [24]. The CD of WD1-6 unfolding in the presence of denaturant has not been reported. Copper binding studies have been used
to monitor CD from a WD1-6 protein reconstituted from inclusion bodies consisting of residues 1-633 [6, 25]. The copper binding CD data reports a mean residue ellipticity of 4500 for apo WD1-6, and 5000 for copper loaded WD1-6 protein at 222 nm. Also, a different set of data reported by the same authors that included zinc binding studies gives apo WD1-6 protein a mean residue ellipticity of 5500, and 7100 for copper loaded WD1-6 at 222 nm.

![Unfolding of WD1-6 under increasing ionic strength of guanidine hydrochloride.](image)

**Figure 3.21** Unfolding of WD1-6 under increasing ionic strength of guanidine hydrochloride.
Figure 3.22 WD 1-6 unfolding monitored at 222 nm with increasing molar concentration of GnHCl.
Table 3.15 The data set used to plot the linear extrapolation curve. The Y represents the selected points along the WD 1-6 unfolding curve at 222 nm that were used to calculate the free energy of unfolding.

<table>
<thead>
<tr>
<th>YF</th>
<th>Y</th>
<th>FF</th>
<th>K</th>
<th>FU</th>
<th>YU</th>
<th>[AG]</th>
<th>[GnHCl]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-6316.28</td>
<td>-6214.75</td>
<td>0.97857</td>
<td>0.0219</td>
<td>0.02143</td>
<td>-1578.77</td>
<td>9467.442</td>
<td>0.59</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-6000.42</td>
<td>0.93333</td>
<td>0.07143</td>
<td>0.06667</td>
<td>-1578.77</td>
<td>6538.244</td>
<td>1.23</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-5885.87</td>
<td>0.90915</td>
<td>0.09993</td>
<td>0.09085</td>
<td>-1578.77</td>
<td>5706.553</td>
<td>1.48</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-5266.91</td>
<td>0.7785</td>
<td>0.28453</td>
<td>0.2215</td>
<td>-1578.77</td>
<td>3114.144</td>
<td>2.09</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-4976.83</td>
<td>0.71727</td>
<td>0.39418</td>
<td>0.28273</td>
<td>-1578.77</td>
<td>2306.488</td>
<td>2.31</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-4407.76</td>
<td>0.59715</td>
<td>0.67463</td>
<td>0.40285</td>
<td>-1578.77</td>
<td>975.1544</td>
<td>2.61</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-3977.26</td>
<td>0.50628</td>
<td>0.97521</td>
<td>0.49372</td>
<td>-1578.77</td>
<td>62.20582</td>
<td>2.84</td>
</tr>
<tr>
<td>&quot;-6316.28</td>
<td>-3193.86</td>
<td>0.34092</td>
<td>1.93328</td>
<td>0.65908</td>
<td>-1578.77</td>
<td>-1633.26</td>
<td>3.27</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-2674.67</td>
<td>0.23132</td>
<td>3.32293</td>
<td>0.76868</td>
<td>-1578.77</td>
<td>-2975.19</td>
<td>3.62</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-2244.17</td>
<td>0.14045</td>
<td>6.11978</td>
<td>0.85955</td>
<td>-1578.77</td>
<td>-4488.18</td>
<td>4</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-2029.84</td>
<td>0.09521</td>
<td>9.50279</td>
<td>0.90479</td>
<td>-1578.77</td>
<td>-5578.47</td>
<td>4.3</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-1878.34</td>
<td>0.06323</td>
<td>14.81428</td>
<td>0.93677</td>
<td>-1578.77</td>
<td>-6678.52</td>
<td>4.6</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-1726.83</td>
<td>0.03125</td>
<td>30.99688</td>
<td>0.96875</td>
<td>-1578.77</td>
<td>-8507.7</td>
<td>5.09</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-1664.01</td>
<td>0.01799</td>
<td>54.57743</td>
<td>0.98201</td>
<td>-1578.77</td>
<td>-9909.35</td>
<td>5.6</td>
</tr>
<tr>
<td>-6316.28</td>
<td>-1651.08</td>
<td>0.01526</td>
<td>64.5152</td>
<td>0.98474</td>
<td>-1578.77</td>
<td>-10323.8</td>
<td>6</td>
</tr>
</tbody>
</table>
Figure 3.23 WD 1-6 denaturation curve analysis using the Linear Extrapolation Method (LEM). The Gibbs free energy calculated from selected points along denaturation curve is plotted against denaturant concentration [GnHCl]; the y-intercept is the free energy of unfolding in the absence of denaturant.
Table 3.16 Conformational stability of WD 1-6. The free energy of unfolding, $m^\circ [\text{GnHCl}]^1/2$ for WD 1-6 is compared with other N-terminal domains and mutant protein, WD5-6(Y532H).

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WD4$^\circ$</th>
<th>WD 5-6</th>
<th>WD 1-6</th>
<th>WD5-6(Y532H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$AGU^H^H^O^\circ$ (kJ mol$^{-1}$)</td>
<td>44.5 ± 0.4</td>
<td>7.9 ± 0.6</td>
<td>11.0 ± 0.2</td>
<td>8.0 ± 0.0</td>
</tr>
<tr>
<td>$m$ (kJ mol$^{-1}$ M$^{-1}$)</td>
<td>-8.3 ± 0.1</td>
<td>-2.8 ± 0.1</td>
<td>-3.8 ± 0.0</td>
<td>-2.9 ± 0.0</td>
</tr>
<tr>
<td>[GnHCl] $^1/2$, M</td>
<td>5.38</td>
<td>2.82</td>
<td>2.91</td>
<td>2.78</td>
</tr>
</tbody>
</table>

The mean residue ellipticity reported here for the WD1-6, residues 57-633 of Wilson protein, was -6100 at 222 nm. The copper content was not quantified before CD measurement, but purification was done in the presence of 0.5 mM EDTA. The Gibbs free energy of unfolding of WD 1-6 in the absence of denaturant is shown in Table 3.16. The value obtained using the linear extrapolation method was 11.04 kJ/mole. The $m$, the parameter which is proportional to the amount of surface area of protein in contact with solvent upon unfolding, [26, 27] was found to be -3.8 kJ/mole/M. The concentration at which the protein is half denatured, the [GnHCl]$^1/2$, was 2.9 M for WD 1-6. The thermodynamic parameters for WD 1-6 unfolding (denaturation curve) by GnHCl have not been previously reported. It is recommended that future calculations be extended to more inclusive methods like the denaturant binding model [28], Tanford's methods [29], and the method of Stainforth et al [30].

$^\circ$ From PhD Dissertation of Wilson Okumu, Western Michigan University ©2010.
3.7 Summary

The N-terminal domains of human Wilson protein (residues 57-633 of ATP7B) have been cloned, expressed, and purified utilizing both established protocols and newly developed methods. WD1-6 gene amplification was found to be dependent on the DNA polymerase used, the nature of deoxynucleotide mixture, and PCR thermocycler settings. DNA polymerases with faster elongation rates and fewer PCR cycles favored the process. Amplification of an entire plasmid vector containing the WD1-6 insert gene using site-directed mutagenesis kit was conducted with high fidelity and specificity, however. Three WD1-6 expression vectors, pET41Ek/WDl-6, pET32Xa/WDl-6, and pET32XaTEV/WDl-6 were prepared, which coded for GST and an enterokinase cleavage site, TRX and a Factor Xa cleavage site, and TRX, and Factor Xa/TEV cleavage sites. Work on pET41Ek/WDl-6 was discontinued due many missense mutations introduced by DNA polymerase errors. WD1-6 protein obtained from the pET32Xa/WDl-6 construct readily degraded during purification and the removal of the thioredoxin tag was inefficient. Modification of pET32Xa/WDl-6 construct to pET32XaTEV/WD 1-6 led to a stable protein.

Expression of WD1-6 fusion protein at lower (28-30°C) than physiological temperatures improved its cellular solubility. The protein was extracted by both freeze/thaw and BugBuster® protein extraction reagent. The protein was subsequently purified by Immobilized Metal Affinity Chromatography (IMAC) and Size Exclusion Chromatography (SEC). Tobacco Etch Virus (TEV) protease that was used to cleave the TRX tag was purified in our lab and its activity standardized. Cleaved WD1-6 was cleaned up by both IMAC and SEC before storage. The DNA integrity after amplification
was confirmed by primer extension sequencing and restriction enzyme mapping. Protein expression and purification was followed by SDS-PAGE analysis.

Biophysical characterization of WD1-6 was performed using gel filtration, light scattering and circular dichroism techniques. The protein was found to migrate faster than protein of similar mass in gel filtration column; therefore, estimation of its molar mass using this technique was too high. This phenomena was also seen during light scattering experiments, though the LS scattering data showed that the protein was polydisperse and non-spherical. This implies that WD1-6 is somewhat elongated. This feature may be important \textit{in vivo} in that the N-terminal domains are likely to require flexibility to interact with cytoplasmic domains, transmembrane domains, and copper metallochaperone. As seen in the NMR titrations of Cu-ATOX1 and WD1-6, N-terminal metal binding domains 1, 2, and 4 are able to form a complex with Cu-ATOX1, but not 3, 5, and 6 [5]. These N-terminal metal binding domains are separated with amino acid linkers of different lengths (see Section 1.2.1). Domains 5-6 and 1-2 are separated by an 8 amino acid and 11 amino acid linkers, respectively. Domain 4 is separated from domain 5 by a 57 amino acid linker, and a 30 amino acid linker from domain 3, which is then separated by a 42 amino acid linker from domain 2. Since we have found that WD1-6 is not globular in structure, we can compare these finding with reported structural studies of CopA copper transporting ATPase in bacteria.

Limited proteolysis studies and cryo-electron microscopy have been used to delinate structural and functional organization of the N-terminal metal binding domain of the CopA, a bacterial Cu-ATPase from \textit{Thermatoga maritima}. Hatori et al. found that N-, A-, and N-terminal metal binding domain undergo rearrangement when apo CopA is
metallated, and when bound to phosphate analogs [31]. The papain cleavage site at N-terminal metal binding domain-Ma helix loop is protected against papain attack when CopA is bound to copper and phosphate analogs. This work suggests that N-terminal metal binding domains' association with A-domain lead to overall dominnal reorganization, and thus, aid the protein to attain required conformations for phosphorylation, copper transfer, and regulation. This work is also supported by Stokes et al. through a cryo-electron microscopy studies of CopA [32]. The N-terminal metal binding domain was found to be localized within the cytoplasmic domains, N-, A-, and P-domains. It is hypothesized that N-terminal metal binding domain interact with N-domain in the absence of copper and prevents ATP binding. This state is likely disrupted by metallated metallochaperone docking on N-terminal metal binding domain and commencement of copper efflux through the transduction channel. We hypothesize that the N-terminal metal binding domains of human Wilson protein, WD1-6, follow a similar mechanism. We suggest that domains 3, 5 and 6 are not available for ATOX1 docking; they are likely to be associated with cytoplasmic domains in a copper free state. The long linkers separating these domains may permit domianal arrangement in which domains 1, 2, and 4 are facing away from cytoplasmic domains for possible interaction with metallated ATOX1. Binding of copper on N-terminal metal binding domains initiates the conformational changes in WNDP leading to activation of the enzyme. This arrangement may constitute the control mechanism for copper binding, translocation and the ATP hydrolysis as the energy source.
References


10. KOD Hot Start DNA polymerase, EMD Chemicals Inc., Gibbstown, NJ.

11. pENTR® Directional TOPO® Cloning kits, Invitrogen Corp., Carlsbad, CA.
   (Integrated DNA Technologies, Inc., Coralville, IW).


CHAPTER 4

CHARACTERIZATION OF Y532H MUTATION OF THE HUMAN WILSON PROTEIN

4.1 Introduction

Over 260 disease-causing mutations have been found in the gene encoding the copper transporting ATP7B (Wilson protein) [1-6]. The geographical distribution of the mutations is as diverse as the world population itself; stretching all the way from Hungary to Brazil. Because of this large number, the characterization and analysis of these mutations has remained at the genetic and clinical levels [7]. Although computer modeling is gaining momentum, no structures exist for full length Wilson protein [8]. To advance the understanding of the molecular basis for these mutations, the Y532H mutation (Figure 4.1) was selected due to (1) its location within the N-terminal domains 5 and 6, which have been extensively studied and (2) its proximity to the interfacial region between domains 5 and 6.

This chapter begins with a description of site-directed mutagenesis and cloning of the WD5-6(Y532H) mutation. The protein purification and biophysical characterization steps are then summarized as described in Section 2.4.2, followed by results and discussion. Findings from the research are also highlighted.
Domain 5

Domain 6

Figure 4.1 The NMR structure of WD5-6. The position of the Y532 amino acid residue (also called Y48 when numbered from position 485 of the human Wilson protein) is shown in red (PDB 2EW9).

4.2 Construction of pET24d/WD5-6(Y532H), pET32Xa/WD5-6(Y532H) Plasmids

4.2.1 PCR Reactions and Results

The Y532H mutation, in which the amino acid tyrosine is substituted with histidine, was incorporated into pET24d/WD5-6 by site-directed mutagenesis. The cloning of pET24d/WD5-6 vector containing the N-terminal domains 5 and 6 of Wilson protein has been described before in a paper published by members of our lab and the Bertini group at the Center for Magnetic Resonance, Florence, Italy [9]. The PCR reaction conditions for Y532H mutagenesis are shown in Tables 4.1 and 4.2. The PCR product is shown in the Figure 4.2 along with the pET24d/WD5-6 DNA template.
Table 4.1 The PCR mutagenesis reaction of WD5-6(Y532H).

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (µL)</th>
<th>Final mass (ng)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Water</td>
<td>37.25</td>
<td></td>
</tr>
<tr>
<td>2 1 OX <em>PfuUltra</em> polymerase buffer</td>
<td>5.0</td>
<td></td>
</tr>
<tr>
<td>3 pET24d/WD5-6 DNA (-7.4 ng/µL)</td>
<td>1.35</td>
<td>10.0</td>
</tr>
<tr>
<td>4 KD-Y532H Fwd primer (10</td>
<td>iM)</td>
<td>1.2</td>
</tr>
<tr>
<td>5 GC KD-Y532H Rev primer (10 pM)</td>
<td>1.2</td>
<td>125.0</td>
</tr>
<tr>
<td>6 dNTPs Mix</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>7 Quiksolution (Additive)</td>
<td></td>
<td>3.0</td>
</tr>
<tr>
<td>8 <em>PfuUltra</em> HF DNA polymerase</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>9 Total reaction volume</td>
<td></td>
<td>50.0</td>
</tr>
<tr>
<td>10 New plasmid synthesized following parameters in Table 4.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11 Parental DNA digested with <em>Dpn I</em> restriction enzyme for 2 hours at 37°C</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12 New plasmid transformation, colony screening and DNA purification</td>
<td></td>
<td></td>
</tr>
<tr>
<td>13 Primer extension DNA sequencing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2 The cycling parameters of WD5-6(Y532H) mutagenesis.

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycle</th>
<th>Temperature (°C)</th>
<th>Time (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95</td>
<td>60</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>95</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68</td>
<td>330</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>68</td>
<td>420</td>
</tr>
</tbody>
</table>
Figure 4.2 The WD5-6(Y532H) PCR product. The 5.7 Kb PCR product (Lane 3) is visualized alongside 1 Kb DNA ladder (NEB), and pET24d/WD5-6 DNA template (Lane 2) in an ethidium bromide stained 0.8% agarose gel.
After transformation in Rosetta(DE3), the protein expressed from pET24d/WD5-6(Y532H) was found to be insoluble, leading to low yield for the mutant protein. Expressing WD5-6(Y532H) as a fusion protein with thioredoxin was pursued since (1) the solubility of expressed protein was likely to increase, (2) the presence of a polyhistidine tag would aid in the purification process, and (3) if the insoluble protein persisted, refolding could be performed with the protein remaining bound on a HisPrep column. The gene encoding WD5-6(Y532H) was inserted into the pET32XaLIC vector using the LIC method discussed in Chapter 3 (Section 3.3.1). The amplification reaction conditions are shown in Tables 4.3 and 4.4. The amplified PCR product (450 bp) is shown in Figure 4.3. The LIC cloned insert was confirmed by restriction mapping with BamHI and Bgl II, and primer extension DNA sequencing (Figure 4.4).

4.2.2 Discussion

The mutagenesis of Y532H into domains WD5-6 was straightforward. The primers were designed with a high content of guanine and cytosine to ensure correct binding to the DNA template. The PCR product was incubated for 2 h with Dpn I instead of the recommended 1 h. The extended incubation not only decreased chances of false positives due to parental DNA, but also made screening for the mutant plasmid easier. Unlike the WD1-6 gene insert, the amplification of WD5-6 containing the Y532H mutation (hereafter referred as WD5-6(Y532H)) with LIC primers went perfectly. The amplification of WD5-6(Y532H) gene insert using LIC primers by PCR did not show any specificity to a particular DNA polymerase and needed no thermocycler optimization or external additives.
Table 4.3 The WD5-6(Y532H) gene amplification reaction with LIC-qualified primers.

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5X HotStar DNA polymerase buffer (Qiagen)</td>
<td>20.0</td>
</tr>
<tr>
<td>Water</td>
<td>74.15</td>
</tr>
<tr>
<td>pET24d/WD5-6(Y532H) (13.5 ng/μL)</td>
<td>1.0</td>
</tr>
<tr>
<td>WD5 XaLIC Fwd primer (10 pM)</td>
<td>1.93</td>
</tr>
<tr>
<td>WD6 XaLIC Rev primer (10 pM)</td>
<td>1.92</td>
</tr>
<tr>
<td>HotStar HiFidelity DNA polymerase (Qiagen)</td>
<td>1.0</td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 4.4 The PCR reaction conditions for WD5-6(Y532H) gene amplification with LIC primers.

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Temperature (°C)</th>
<th>Duration (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Activation cycle</td>
<td>95</td>
<td>5.0</td>
</tr>
<tr>
<td>2 Extension</td>
<td>72</td>
<td>1.0</td>
</tr>
<tr>
<td>3 Melting</td>
<td>95</td>
<td>0.25</td>
</tr>
<tr>
<td>4 Annealing</td>
<td>60</td>
<td>1.0</td>
</tr>
<tr>
<td>5 Repeat steps 2 to 4 for 30 cycles</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7 Final extension</td>
<td>72</td>
<td>10.0</td>
</tr>
<tr>
<td>6 End/store</td>
<td>4</td>
<td>Indefinite</td>
</tr>
</tbody>
</table>

f Contained deoxynucleotides (dNTPs) mix.
Figure 4.3 The amplified WD5-6(Y532H) gene for insertion into pET32XaLIC vector. Lane 1, the 450 bp PCR product is visualized alongside a 1 Kb DNA ladder, Lane 2, in a 0.8% ethidium bromide stained agarose gel. The Xa before Y532H denotes Factor Xa.

Figure 4.4 Restriction mapping of pET32Xa/WD5-6(Y532H) construct. Lane 1, pET32XaLIC vector as control, Lane 2, the pET32Xa/WD5-6(Y532H) construct, and Lane 3, 1 Kb DNA ladder (NEB). This 0.8% agarose gel was stained with 0.5 ng/mL ethidium bromide solution.
4.3 Purification of WD5-6(Y532H) Mutant Protein

4.3.1 General Considerations and Results

The purification and characterization of native WD5-6 has been described before [9]. The protein expresses very well and is easily extracted from the cell pellet by the freeze/thaw method in liquid nitrogen. When the Y532H mutation was introduced into the native WD5-6 domains, the mutant protein expressed at 37°C with 1 mM IPTG was largely insoluble. The mutant protein yield from extraction by freeze/thaw method in liquid nitrogen was very low.

The WD5-6(Y532H) mutant protein was then expressed as a fusion protein with thioredoxin containing Factor Xa cleavage site (TRXx\_a WD5-6(Y532H)). The expression of WD5-6(Y532H) as a fusion did not increase the solubility as expected. Therefore, the entire fusion protein was purified using an inclusion body protocol, followed by denaturation with 6 M GnHCl and refolding. The thioredoxin fusion protein was subsequently cleaved off with Factor Xa. The WD5-6(Y532H) prepared by the refolding method and Factor Xa cleavage was prone to precipitation and degradation. The expression and purification of WD5-6(Y532H) without a fusion protein (pET24d/WD5-6(Y532H) plasmid construct) was later optimized. The induction temperature was lowered between 28-30°C from 37°C and the concentration of IPTG was increased from 1 mM to 3 mM. The induction time was also extended from 3-4 h to 4-6 h. The purification methods and SDS-PAGE results are discussed in following sections.
4.3.1.1 **DEAE (diethylaminoethyl) Cation Exchange Chromatography**

The non-fusion WD5-6(Y532H) protein (expressed from pET24d/WD5-6(Y532H) plasmid construct) was first purified by a cation exchange sepharose column. The negatively charged WD5-6(Y532H) protein bound while positively charged proteins did not. **Figure 4.5 (A)** shows 15% Tricine SDS-PAGE gel of time-course induction of WD5-6(Y532H) at 37°C and **Figure 4.5 (B)** shows WD5-6(Y532H) protein extraction using different methods before modification of expression and purification methods. After development of the new method to express the protein at 28°C, it was extracted by the freeze/thaw method, followed by sonication and purification with DEAE Sepharose column. **Figure 4.6** shows eluted DEAE fractions of WD5-6(Y532H) obtained after conditions were optimized.

4.3.1.2 **Immobilized Metal Affinity Chromatography (IMAC) Purification**

The inclusion bodies resulting from the expression of TRX_{Xa}WD5-6(Y532H) were processed according to the description given in materials and experimental methods section (see Section 2.4.2). The denatured fusion protein was then directly loaded on the column; the bound denatured protein was washed with 20 mM phosphate, pH 7.5, 500 mM NaCl, and 20 mM imidazole, causing simultaneous refolding of the protein *in situ*. The refolded fusion protein was then eluted with 300-500 mM imidazole (**Figure 4.7**).

4.3.1.3 **Fusion Protein Cleavage of TRX_{Xa}WD5-6(Y532H)**

The refolded fusion protein recovered from inclusion bodies was relatively pure. The refolded TRX_{Xa}WD5-6(Y532H) had high precipitation rates when buffer conditions
Figure 4.5 Expression and purification of non-fusion WD5-6(Y532H). (A) Time-course induction at 37°C with 1 mM IPTG. (B) Protein extraction with different methods, Lane 1, BugBuster® reagent extract; Lane 2, cell pellet after Bugbuster® reagent extraction; Lane 3, cell pellet after freeze/thaw extraction; Lane 4, freeze/thaw extract; Lane 5, induced cell pellet as a control. Arrows indicate the protein of interest.
Figure 4.6 The purification of non-fusion WD5-6(Y532H) from optimized conditions. Arrow indicates the protein of interest. Lanes 1-8, protein elution fractions from DEAE Sepharose column. The protein eluted -0.16 M NaCl. Lane 9 is crude extract (supernatant) before DEAE purification.

Figure 4.7 The refolded TRX<sub>X</sub>WD5-6(Y532H) purification. The 10% Tricine gel shows isocratically eluted TRX<sub>X</sub>WD5-6(Y532H). Lanes 1-8, from a HisPrep FF 16/10 column charged with nickel. Molar mass of the fusion protein is 33 kDa.
were altered to Factor Xa buffer (50 mM Tris/Cl, pH 8.0, 100 mM NaCl, 5 mM CaCl₂).

Even though Factor Xa cleavage was kept under 16 h, significant amounts of protein were lost to precipitation. The cleaved thioredoxin tag was removed using a 5 mL HisPrep column pre-charged with nickel and the refolded WD5-6(Y532H) was purified further by gel filtration. TRXₓₓₓₓₓₓ WD5-6(Y532H) digestion by Factor Xa is shown in Figure 4.8 below.

Figure 4.8 Thioredoxin tag cleavage from TRXₓₓₓₓₓₓ WD5-6(Y532H). Lanes 1, 2, WD5-6 as a control; Lanes 3, 4, the cleavage of TRXₓₓₓₓₓₓ WD5-6(Y532H); A, B, and C indicate the positions of TRXₓₓₓₓₓₓ WD5-6(Y532H), TRX and WD5-6(Y532H), respectively in lanes 3 and 4 after 16 h. The protein samples were loaded in duplicate but varying amounts for clarity.
4.3.1.4 Size Exclusion Chromatography of WD5-6(Y532H) Protein

Size Exclusion Chromatography (gel filtration) was the final purification step. It enabled removal of contaminating proteins and residual protease from the digestion reactions. The protocol for this step is discussed in the material and experimental methods section (Section 2.5.2). The SDS-PAGE gel of WD5-6(Y532H) recovered from optimized growth and purification conditions is shown in Figure 4.9 with an elution peak at 94.85 mL from the 120 mL Superdex 200 16/60 column. Before gel filtration, WD5-6(Y532H) obtained by refolding and Factor Xa cleavage method was first purified with a HisPrep tag column to remove the thioredoxin tag. The elution trace for this WD5-6(Y532H) from the 320 mL superdex 75 26/60 column is shown in Figure 4.10. The molar mass of $^{15}$N, $^3$C isotopically labeled WD5-6(Y532H) for NMR studies determined by MALDI-TOF mass spectrometry is shown in Figure 4.11. The protein was 99.8% isotopically labeled; the molar mass obtained by MALDI-TOF mass spectrometry was 16,896 Daltons, whereas the actual molar mass of a fully $^{15}$N, $^{13}$C isotopically labeled WD5-6(Y532H) is 16,924 Daltons.

4.3.1.5 Caveats of WD5-6(Y532H) Purification

The solubility of the thioredoxin tagged TRX$_x$WD$_5$-6 (Y532H) protein was not better than that of the protein without the fusion tag. The protein extraction by freeze/thaw method and BugBuster® reagent extraction methods could only recover minute quantities. With protein rich inclusion bodies, protein solubilization and refolding method was an appropriate option. The fusion protein was first processed to remove extraneous protein matter following the BugBuster® protein extraction protocol and then
Figure 4.9 Gel filtration of WD5-6(Y532H) recovered under optimized conditions. Lanes 1, 2, 3 are protein fractions with contaminants; lanes 4, 5 are fractions with pure WD5-6(Y532H). The UV trace on the top is for the 120 mL Superdex 200 16/60 prep grade column and corresponding elution volume for each fraction is shown at the bottom of the gel. Arrow indicates the protein of interest.
Figure 4.10 Gel filtration of WD5-6(Y532H) obtained by refolding/Factor Xa cleavage method. Lanes 1-7 are elution fractions sampled by 10% Tricine SDS-PAGE gel from Superdex 75 26/60 column. The UV trace for elution profile is shown on top while the elution volume for each fraction is indicated at the bottom of gel. Arrow indicates the protein of interest.
Figure 4.11 The MALDI-TOF of $^{13}$C, $^{15}$N labeled WD5-6(Y532H), -16, 896 Daltons, obtained after refolding and thioredoxin cleavage; Department of Pharmaceutical Sciences, University of Florence, Italy. Actual molar mass of a fully $^{13}$C, $^{15}$N isotopically labeled WD5-6(Y532H) is 16,924 Daltons.

solubilized with 6 M GnHCl, loaded on HisPrep nickel column and refolded in situ by washing the bound protein with HisPrep binding buffer (20 mM Na Phosphate, pH 7.4, 500 mM NaCl, 20 mM imidazole). The drawbacks of inclusion body processing and refolding were time and cost. In addition, WD5-6(Y532H) obtained by the refolding/Factor Xa cleavage method was unstable especially when exchanged from high to low salt buffer. For example, protein was stable in phosphate buffer with 500 mM NaCl and 500 mM imidazole near pH 7 but precipitated when exchanged into 50 mM Tris/Cl, 100 mM NaCl.

Although the cleavage of fusion protein with Factor Xa was efficient, aggregation of cleaved WD5-6(Y532H) hindered recovery. The analysis of WD5-6(Y532H) degradation by SDS-PAGE gel revealed two protein fragments, 5 kDa and 10 kDa in size. Degradation of the protein might have been due to poorly refolded protein undergoing self autolysis or digestion of misfolded protein by residual Factor Xa.
Due to the limited storage life of the WD5-6(Y532H) obtained by refolding/Factor Xa cleavage method, the protein needed to be expressed and purified continually for experiments. This proved to be tedious and costly. The expression and purification of non-fusion WD5-6(Y532H) was modified. The induction of the WD5-6(Y532H) was done at low temperature (28-30°C), concentration of IPTG was increased up to 3 mM, and the induction time extended by 2 h. Sonicating the cell suspension after rounds of freeze/thaw in liquid nitrogen caused more WD5-6(Y532H) protein to be released. WD5-6(Y532H) recovered by this method was comparable to the native WD-6, both in yield and physical stability. It is possible that WD5-6(Y532H) expressed at 37°C (physiological temperature) was prone to misfolding forming inclusion bodies. This behavior may be the reason this mutation leads to Wilson disease. At body temperature, the ATP7B enzyme bearing this mutation may have a higher probability of misfolding.

4.4 Biophysical Characterization of WD5-6(Y532H) Protein

Except for NMR, all others experiments were performed on WD5-6(Y532H) recovered from non-fusion construct, pET24d/WD5-6(Y532H) using optimized expression and purification methods.

4.4.1 Gel Filtration and Light Scattering

The primary objective for high resolution gel filtration analysis of mutant WD5-6(Y532H) protein was to find the migratory behavior of the protein through a stationary phase. The properties evaluated included hydrodynamic radius, aggregation state, and dimer formation. In this experiment, a Superdex 75 10/300 column was used for
separation with the native WD5-6 as the control. The injection volume was 25 pL of 1-5 mg/mL protein. Figures 4.12 and 4.13 show the elution profiles for WD5-6 and WD5-6(Y532H) proteins through the size exclusion column, respectively. The elution trace for WD5-6(Y532H) obtained by refolding/Factor Xa cleavage method is not shown but gives a similar elution time.

Lights scattering measurements were also performed for both WD5-6 and WD5-6(Y532H). The molar mass distribution determined by static light scattering (SLS) is shown in Figure 4.14 while calculated molecular weights of different species in the protein are tabulated in Table 4.5. The hydrodynamic radius, R_h was also determined using dynamic static scattering (DLS). The autocorrelation function curves for WD5-6 and WD5-6(Y532H) along with the derived translational diffusion coefficient which is used to calculate R_h are shown in Figure 4.15 and Figure 4.16. The hydrodynamic radius distribution of the proteins and the constituent species is shown in Figure 4.17. The R_H, the apparent shape, and calculated molar mass of the protein monomers determined by DLS are reported in Table 4.6.

4.4.2 Discussion

The gel filtration data indicates that both the mutant and native protein domains do not possess any significant differences that can be determined by this method. This is also supported by light scattering experiments. The elution times in Figure 4.12 and Figure 4.13 are within experimental error and both chromatograms show a Gaussian function. Although the entire elution profile is not shown for gel filtration results,
Figure 4.12 High resolution gel filtration of WD5-6(Y532H). The protein elutes at -12.88 mL using a Superdex 75 10/300 GL column.

Figure 4.13 High resolution gel filtration of WD5-6. The protein elutes at -12.79 mL using a Superdex 75 10/300 GL column.
Figure 4.14 Molar mass distribution of WD5-6 and WD5-6(Y532H) by static light scattering. A plot of molar mass (g/mol) against elution time in min (gel filtration) monitored by UV at 280 nm is shown. The "non-continuous" lines labeled A-E are weight average MW for each slice, i.e., measured every 1 s. A: WD5-6(Y532H) aggregate (blue), B: WD5-6 tetramer (red), C: WD5-6 dimer (red), D: WD5-6 monomer, (red), E: WD5-6(Y532H) monomer (red overlapped with blue color). Computed values are tabulated in Table 4.5.
Table 4.5 Calculated molecular weights (MW) of WD5-6 and WD5-6(Y532H) determined by SLS. The elution profile is shown in Figure 4.15. Note that peaks' elution is simultaneously monitored by UV, RI and LS detectors.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Peak Elution at RI trace (mL)</th>
<th>MW (kDa) Calculated by ASTRA Average for the major peak</th>
<th>MW (kDa) Calculated by ASTRA Range of Mw observed</th>
<th>MW for monomer (kDa)</th>
<th>Mono disperse peak</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD5-6(Y532H)</td>
<td>11.87</td>
<td>16.1</td>
<td>15.5-16.5</td>
<td>16</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>7.9</td>
<td>350</td>
<td>280-420</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>WD5-6</td>
<td>11.89</td>
<td>16.1</td>
<td>15.5-16.5</td>
<td>16</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>10.4</td>
<td>32</td>
<td>31.5-32</td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>9.1</td>
<td>72</td>
<td>64-85</td>
<td></td>
<td>No</td>
</tr>
</tbody>
</table>

Figure 4.15 Autocorrelation function of WD5-6 monomer peak. The translational diffusion coefficient derived from this exponential graph is shown.
Translational diffusion coefficient

\[ D_t = (1.19 \pm 0.03) \times 10^{-6} \text{ cm}^2/\text{sec} \]

**Figure 4.16** Autocorrelation function of WD5-6(Y532H) monomer peak. The translational diffusion coefficient derived from this exponential graph is shown.

**Table 4.6** Hydrodynamic radius of WD5-6 and WD5-6(Y532H) protein monomers. Molecular weight determined by DLS and apparent shape of the monomers is also given.

<table>
<thead>
<tr>
<th>Sample</th>
<th>RH (nm)</th>
<th>MW (g/mol) DLS</th>
<th>Actual MW (g/mol)</th>
<th>Apparent shape</th>
</tr>
</thead>
<tbody>
<tr>
<td>WD5-6</td>
<td>2.10±0.06</td>
<td>16,060±20</td>
<td>16,045</td>
<td>Spherical</td>
</tr>
<tr>
<td>WD5-6(Y532H)</td>
<td>2.05±0.06</td>
<td>16,140±10</td>
<td>16,031</td>
<td>Spherical</td>
</tr>
</tbody>
</table>
Figure 4.17 The $R_H$ distribution of WD5-6 and WD5-6(Y532H) proteins. A plot of hydrodynamic radius against elution volume in mL (gel filtration) is shown. The "dotted" lines, A-D are $R_H$ distributions within the eluting peaks. A: WD5-6, tetramer (red), B: WD5-6 dimer (red), C: WD5-6 monomer (red) and, D: WD5-6(Y532H) monomer (blue). The slanted $R_H$ distribution is indication of a polydisperse material in the elution peak, whereas $R_H$ for a monodisperse material remains relatively constant across the elution peak.
dimerization to a higher-order by WD5-6 and WD5-6(Y532H) protein monomers was detected by light scattering experiments (Figures 4.14 and 4.17). A prior analysis with SDS-PAGE found that WD5-6(Y532H) dimerized more relative to the native from when exposed to the air, but this was not detected in freshly prepared samples. The presence of aggregates of higher-order in WD5-6(Y532H) at 7.9 mL peak (Figure 4.17) but not in WD5-6 could be due to reduced association stability dimers and tetramers. Unlike the N-terminal metal binding domains, native WD5-6 and mutant WD5-6(Y532H) monomers were found to be monodisperse and spherical. Thus, the calculated average mass of these monomers was independent of the averaging method (if a peak contains a mixture of species of different molar sizes, average molar mass is given). This mean the monomer peak contains only one type of the macromolecule. The \( R_h \) and molar mass determined by SLS and DLS (Tables 4.5 and 4.6) for both proteins show a negligible deviation from each other.

It seems that the mutation does not have profound effects on the overall structure and shape of the native protein, since protein migration in gel filtration is a function of the overall shape of the macromolecule. As is discussed in Section 4.3.1.3, the WD5-6(Y532H) protein obtained by refolding followed by thioredoxin cleavage had a tendency to aggregate and degrade. Because of the small amounts of protein sample injected, it was hard to estimate the fraction of protein that was been lost due to aggregation. WD5-6(Y532H) protein recovered by optimized expression and purification methods and the native WD5-6 were found to be very stable; degradation and precipitation was not observed. During light scattering measurements, it was possible to recover up to -85% of the sample that was injected into the size exclusion column. While it is evident that gel
filtration analysis and light scattering results show no defined effects due to mutation, this information might provide some insights on how slight perturbations at atomic level due to missense mutations, can be ultimately lethal without affecting physical properties of a protein.

4.4.3 Circular Dichroism Spectroscopy

4.4.3.1 Results

The CD data conversion and processing was done with RIS Manager Software (JASCO). The Origin 7.5 software (Origin Lab) and Microsoft Excel were used for data plotting and tabulation of the unfolding studies. The CD spectra for WD5-6(Y532H) and WD5-6 were acquired under the following conditions: (1) in the absence of denaturant, (2) under increasing molar concentration of GnHCl, and (3) at two different protein concentrations 14.5 and 29 p.M. The unfolding was monitored from 260-190 nm. The proteins were freshly purified and samples prepared by mixing with an appropriate amount of 8 M GnHCl and incubated for 2-6 h.

The CD spectra of WD5-6(Y532H) and WD5-6 unfolding are shown in Figures 4.18-20. To find the relationship in the unfolding pattern between the native and mutant protein, superimposition of spectra at a select denaturant concentrations is shown in Figure 4.20. The Gibbs free energy of unfolding, $AG_{ij}^{H2O}$ in the absence of denaturant was calculated using a two-state model as explained in Sections 2.7.4 and 3.6.3.A plot of Gibbs free energy of unfolding against concentration of denaturant using the data sets in Tables 4.8 and 4.9 from unfolding curve model model at 222 nm (Figure 4.22) is shown.
Figure 4.18 The unfolding of 29 jaM WD5-6(Y532H) under increasing molar concentration of GnHCl.

Figure 4.19 The unfolding of 14.5 iM WD5-6 under increasing molar concentration of GnHCl.
Figure 4.20 The Circular dichroism of 29pM WD5-6 and WD5-6(Y532H) in the absence of denaturant.

Table 4.7 Helicity loss during denaturation.

<table>
<thead>
<tr>
<th>GnHCl</th>
<th>% Helicity WD5-6</th>
<th>% Helicity WD5-6(Y532H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 M</td>
<td>23.9</td>
<td>23.4</td>
</tr>
<tr>
<td>1 M</td>
<td>23.4</td>
<td>22.4</td>
</tr>
<tr>
<td>2M</td>
<td>23.7</td>
<td>16.5</td>
</tr>
<tr>
<td>3 M</td>
<td>16.3</td>
<td>13.3</td>
</tr>
<tr>
<td>4 M</td>
<td>9.0</td>
<td>5.3</td>
</tr>
<tr>
<td>5 M</td>
<td>3.9</td>
<td>3.5</td>
</tr>
<tr>
<td>6 M</td>
<td>2.7</td>
<td>3.0</td>
</tr>
</tbody>
</table>
Figure 4.21 The loss of helicity during denaturation.

Figure 4.22 Ellipticity of WD5-6(Y532H) and WD5-6 monitored at 222 nm.
Table 4.8 Y points along the WD5-6 unfolding curve (Figure 4.21) used to plot the linear extrapolation curve (Figure 4.22). YF; fraction of folded protein at Y, FF; fully folded protein fraction, K; equilibrium constant, FU; fully unfolded protein fraction, YU, fraction of unfolded protein at Y.

<table>
<thead>
<tr>
<th>YF</th>
<th>Y</th>
<th>FF</th>
<th>K</th>
<th>FU</th>
<th>YU</th>
<th>AG(J/Mol)</th>
<th>fGnHClIi</th>
</tr>
</thead>
<tbody>
<tr>
<td>-7123</td>
<td>-2615</td>
<td>0.2945</td>
<td>2.3958</td>
<td>0.7055</td>
<td>-733.5</td>
<td>-2165</td>
<td>3.9</td>
</tr>
<tr>
<td>-7123</td>
<td>-2179</td>
<td>0.2262</td>
<td>3.4218</td>
<td>0.7739</td>
<td>-733.5</td>
<td>-3048</td>
<td>4.2</td>
</tr>
<tr>
<td>-7123</td>
<td>-1867</td>
<td>0.1774</td>
<td>4.6371</td>
<td>0.8226</td>
<td>-733.5</td>
<td>-3801</td>
<td>4.4</td>
</tr>
<tr>
<td>-7123</td>
<td>-1571</td>
<td>0.1311</td>
<td>6.6261</td>
<td>0.8689</td>
<td>-733.5</td>
<td>-4685</td>
<td>4.6</td>
</tr>
<tr>
<td>-7123</td>
<td>-1369</td>
<td>0.0995</td>
<td>9.0548</td>
<td>0.9005</td>
<td>-733.5</td>
<td>-5459</td>
<td>4.8</td>
</tr>
<tr>
<td>-7123</td>
<td>-1044</td>
<td>0.0486</td>
<td>19.592</td>
<td>0.9514</td>
<td>-733.5</td>
<td>-7371</td>
<td>5.2</td>
</tr>
<tr>
<td>-7123</td>
<td>-918.7</td>
<td>0.029</td>
<td>33.497</td>
<td>0.971</td>
<td>-733.5</td>
<td>-8700</td>
<td>5.6</td>
</tr>
<tr>
<td>-7123</td>
<td>-841.4</td>
<td>0.0169</td>
<td>58.211</td>
<td>0.9831</td>
<td>-733.5</td>
<td>-10069</td>
<td>5.8</td>
</tr>
<tr>
<td>-7123</td>
<td>-825.5</td>
<td>0.0144</td>
<td>68.458</td>
<td>0.9856</td>
<td>-733.5</td>
<td>-10471</td>
<td>6.0</td>
</tr>
<tr>
<td>-7123</td>
<td>-6215</td>
<td>0.8578</td>
<td>0.1657</td>
<td>0.1422</td>
<td>-733.5</td>
<td>4453.4</td>
<td>0.59</td>
</tr>
<tr>
<td>-7123</td>
<td>-5886</td>
<td>0.8064</td>
<td>0.2401</td>
<td>0.1936</td>
<td>-733.5</td>
<td>3534.5</td>
<td>1.5</td>
</tr>
<tr>
<td>-7123</td>
<td>-5267</td>
<td>0.7095</td>
<td>0.4094</td>
<td>0.2905</td>
<td>-733.5</td>
<td>2212.4</td>
<td>2.1</td>
</tr>
<tr>
<td>-7123</td>
<td>-4408</td>
<td>0.575</td>
<td>0.739</td>
<td>0.425</td>
<td>-733.5</td>
<td>749.32</td>
<td>2.6</td>
</tr>
<tr>
<td>-7123</td>
<td>-3194</td>
<td>0.3851</td>
<td>1.597</td>
<td>0.6149</td>
<td>-733.5</td>
<td>-1160</td>
<td>3.3</td>
</tr>
<tr>
<td>-7123</td>
<td>-2675</td>
<td>0.3038</td>
<td>2.2916</td>
<td>0.6962</td>
<td>-733.5</td>
<td>-2055</td>
<td>3.6</td>
</tr>
<tr>
<td>-7123</td>
<td>-2030</td>
<td>0.2029</td>
<td>3.9289</td>
<td>0.7971</td>
<td>-733.5</td>
<td>-3390</td>
<td>4.3</td>
</tr>
<tr>
<td>-7123</td>
<td>-1727</td>
<td>0.1555</td>
<td>5.4324</td>
<td>0.8445</td>
<td>-733.5</td>
<td>-4193</td>
<td>5.1</td>
</tr>
</tbody>
</table>
Table 4.9 Y points along the WD5-6(Y532H) unfolding curve (Figure 4.21) used to plot the linear extrapolation curve (Figure 4.22). YF; fraction of folded protein at Y, FF; fully folded protein fraction, K; equilibrium constant, FU; fully unfolded protein fraction, fraction of unfolded protein at Y.

<table>
<thead>
<tr>
<th>YF</th>
<th>Y</th>
<th>FF</th>
<th>K</th>
<th>FU</th>
<th>YU</th>
<th>AG (J/MOL)</th>
<th>[GnHCl]</th>
</tr>
</thead>
<tbody>
<tr>
<td>-9574</td>
<td>-8968</td>
<td>0.9309</td>
<td>0.0743</td>
<td>0.0692</td>
<td>-815.3</td>
<td>6441.3</td>
<td>0.59</td>
</tr>
<tr>
<td>-9574</td>
<td>-8599</td>
<td>0.8887</td>
<td>0.1253</td>
<td>0.1113</td>
<td>-815.3</td>
<td>5146.3</td>
<td>0.98</td>
</tr>
<tr>
<td>-9574</td>
<td>-8132</td>
<td>0.8354</td>
<td>0.197</td>
<td>0.1646</td>
<td>-815.3</td>
<td>4025.2</td>
<td>1.37</td>
</tr>
<tr>
<td>-9574</td>
<td>-7663</td>
<td>0.7819</td>
<td>0.279</td>
<td>0.2181</td>
<td>-815.3</td>
<td>3163.1</td>
<td>1.68</td>
</tr>
<tr>
<td>-9574</td>
<td>-6944</td>
<td>0.6998</td>
<td>0.429</td>
<td>0.3002</td>
<td>-815.3</td>
<td>2096.6</td>
<td>2.1</td>
</tr>
<tr>
<td>-9574</td>
<td>-5876</td>
<td>0.5778</td>
<td>0.7309</td>
<td>0.4223</td>
<td>-815.3</td>
<td>776.81</td>
<td>2.5</td>
</tr>
<tr>
<td>-9574</td>
<td>-5136</td>
<td>0.4934</td>
<td>1.0269</td>
<td>0.5066</td>
<td>-815.3</td>
<td>-65.75</td>
<td>-2.8</td>
</tr>
<tr>
<td>-9574</td>
<td>-4298</td>
<td>0.3976</td>
<td>1.5149</td>
<td>0.6024</td>
<td>-815.3</td>
<td>-1029</td>
<td>3.1</td>
</tr>
<tr>
<td>-9574</td>
<td>-3658</td>
<td>0.3246</td>
<td>2.0807</td>
<td>0.6754</td>
<td>-815.3</td>
<td>-1815</td>
<td>3.4</td>
</tr>
<tr>
<td>-9574</td>
<td>-2800</td>
<td>0.2266</td>
<td>3.4133</td>
<td>0.7734</td>
<td>-815.3</td>
<td>-3042</td>
<td>3.8</td>
</tr>
<tr>
<td>-9574</td>
<td>-2061</td>
<td>0.1422</td>
<td>6.032</td>
<td>0.8578</td>
<td>-815.3</td>
<td>-4452</td>
<td>4.3</td>
</tr>
<tr>
<td>-9574</td>
<td>-1615</td>
<td>0.0913</td>
<td>9.9585</td>
<td>0.9088</td>
<td>-815.3</td>
<td>-5695</td>
<td>4.8</td>
</tr>
<tr>
<td>-9574</td>
<td>-1322</td>
<td>0.0578</td>
<td>16.293</td>
<td>0.9422</td>
<td>-815.3</td>
<td>-6914</td>
<td>5.2</td>
</tr>
<tr>
<td>-9574</td>
<td>-1148</td>
<td>0.038</td>
<td>25.296</td>
<td>0.962</td>
<td>-815.3</td>
<td>-8004</td>
<td>5.6</td>
</tr>
</tbody>
</table>

183
Figure 4.23 The Gibbs free energy of unfolding estimated from a linear extrapolation plot for WD5-6(Y532H) and WD5-6.

Table 4.10 Parameters characterizing the GnHCl denaturation of Wilson protein N-terminal domains

<table>
<thead>
<tr>
<th>Parameter</th>
<th>WD4&lt;sup&gt;8&lt;/sup&gt;</th>
<th>WD 5-6</th>
<th>WD 1-6</th>
<th>WD5-6 (Y532H)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ΔG&lt;sub&gt;20&lt;/sub&gt; (kJ mol&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>44.5 ± 0.4</td>
<td>7.9 ± 0.6</td>
<td>11.0 ± 0.2</td>
<td>8.0 ± 0.0</td>
</tr>
<tr>
<td>m (kJ mol&lt;sup&gt;-1&lt;/sup&gt; M&lt;sup&gt;-1&lt;/sup&gt;)</td>
<td>-8.3 ± 0.1</td>
<td>-2.8 ± 0.1</td>
<td>-3.8 ± 0.0</td>
<td>-2.9 ± 0.0</td>
</tr>
<tr>
<td>[GnHCl] 1/2, M</td>
<td>5.38</td>
<td>2.82</td>
<td>2.91</td>
<td>2.78</td>
</tr>
</tbody>
</table>

<sup>8</sup> From the PhD Dissertation of Wilson Okumu, Western Michigan University ©2010.
Figure 4.23. The calculation of $\Delta G^{20}$ energy, m, and [GnHCl]i/2 values for WD5-6(Y532H), WD5-6, WD4, and WD1-6 based on the Linear Extrapolation Method, LEM [10, 11] are tabulated in Table 4.10.

4.4.3.2 Discussion

Circular Dichroism can be used to estimate the secondary structure of proteins [12-17]. This is common especially when the 3-dimensional structure is not available. CD can be also used to monitor degradation, denaturation and refolding of protein during purification. The wavelength most utilized for this purpose is 222 nm. The solution NMR structure of WD5-6, published by members of our lab in collaboration with Bertini group, CERM, Italy reveals that WD5-6 possesses about 28% P-sheet and 30% α-helical character [9]. Using CDpro (http://lamar.colostate.edu/~sreeram/CDPro/main.htm) and the data in Figures 4.18 and 4.19, the tertiary structure for both WD5-6(Y532H) and WD5-6 in the absence of GnHCl was predicted to be a beta/alpha.

The main objective for WD5-6 and WD5-6(Y532H) CD experiments was to determine the changes in conformational stability due to the single amino acid mutation. This phenomenon, monitored by the disappearance of secondary structure in the presence of denaturant, is shown in Figures 4.18-21, and in Table 4.7. The denaturants alter equilibrium between the native (folded) and denatured (unfolded) protein states. From the unfolding spectra, it can be inferred that the stabilities of the native and mutant are identical, particularly in the absence of denaturant (Figure 4.20). WD5-6 is very stable up to 2M GnHCl, but the WD5-6(Y532H) mutant begins unfolding immediately upon addition of 1M GnHCl. There are differences in the loss of helicity during unfolding.
(Figure 4.21 and Table 4.7). WD5-6(Y532H) loses more helicity than WD5-6 at each denaturant concentration. The Gibbs free energy of unfolding in absence of denaturant, $\Delta G_{\text{U}}$, $m$, and $[\text{GnHCl}]_{i/2}$ of various N-terminal domains of Wilson protein has been determined by the Linear Extrapolation Method (LEM), shown in Table 4.5. Due to small differences in stability (less than a 0.1 kJ/mole) between the mutant and native protein, the change in free energy of unfolding due to the mutation (AAG) is not significant. Professor Pace at Texas A&M University, who has extensively studied conformational stability and the mechanism of folding of proteins, cautions about relying on a single methodology to calculate the free energies of unfolding. He identifies deviations from a two-state mechanism, presence of intermediates, and the empirical nature of LEM as some of the factors that can interfere with precise measurements [18]. The absence of intermediates can be confirmed by differential scanning calorimetry during thermal unfolding.

A better approach to acquire more accurate measurements is to use multiple probes to monitor denaturation. They include UV difference spectroscopy, fluorescence and NMR. The m values for WD5-6(Y532H) and WD5-6 support the hypothesis that m values correlate strongly with amount of protein surface exposed to the solvent upon unfolding [19]. The WD5-6(Y532H) mutation is likely to distort the domain 5 structure, thus, increasing the surface area available for solvent binding, but this is not observed in this study. Conformational stability values of a protein determined from a variety of methods have a higher probability of being accurate. Urea and GnHCl are also known to give $\Delta G_{\text{U}}^{20}$ estimates that are in closer agreement [20], though only GnHCl is shown here.
4.5 NMR of WD5-6(Y532H) from Refolding and Factor Xa Cleavage Method

4.5.1 The Structural Heterogeneity of WD5-6(Y532H) Mutant Protein

The $^1$H-$^{15}$N HSQC spectra of WD5-6(Y532H) show well-dispersed resonances indicative of a folded protein (Figure 4.24). This was also exhibited by CD spectra indicating that the overall fold of both domains of WD5-6(Y532H) is not disturbed by mutation. The residues localized in the linker regions, including E73-G77 (E557-G561, in WNDP), show two sets of assignments in $^1$H-$^{15}$N HSQC spectrums, indicating the presence of structural heterogeneity, i.e. two possible relative orientations of the two domains of WD5-6(Y532H). Amide chemical shifts are highly similar between WD5-6 and WD5-6(Y532H) for domain 6, G80-Q149 (G564-Q633, in WNDP) as shown by chemical shift variation (CSV) data.

The most affected residues include Q4-F7 (Q488-F491 in WNDP), K47-H(Y)48 (K532-H(Y)532) and V52-I53 (V536-I537 in WNDP) (Figure 4.25). Mapping to the structure of wild type WD5-6 (Figure 4.26) shows these residues are spatially close to the mutation. One set of the double conformation of residues, E73-G77 (E557-G561 in WNDP), also show large variation in combined chemical shift, while the other set of chemical shifts are quite similar to wild type. This result indicates that one of the two conformations of this linker region is quite similar to the wild type, whereas the other one leads to a different conformation.
Figure 4.24 The heteronuclear single quantum correction (HSQC) chemical shift of WD5-6(Y532H) overlaid with WD5-6. Red is WD5-6, green is WD5-6(Y532H). Residues indicated by arrows correspond to the following in WNDP: C6, C490; 146, 1530; Y/H48, Y/H532; D47, D531.

Figure 4.25 Chemical shift variation between wild type and Y48H-WD5-6(Y532H). Double conformations of E73-G77 (E557-G561 in WNDP) are shown in green.
Figure 4.26 Combined chemical shift of WD5-6(Y532H). Residues Q4-F7 (Q488-F491 in WNDP) experiencing CSV values above 0.1 are mapped in green on the structure of apo WD5-6. Residues V52-I53 (V536-I537 in WNDP) experiencing CSV smaller than 0.1ppm but bigger than 0.05 are in pink; H48 (H532) mutation is in red; residues E73-G77 (E557-G561 in WNDP) showing double conformations are depicted in blue.

4.5.2 Inter-Domain Motion of WD5-6(Y532H)

4.5.2.1 Results and Discussion

Disease-causing point mutations could alter the global fold and/or dynamic properties of a protein, thus $^{15}$N relaxation parameters on apo WD5-6(Y532H) were analyzed for 122 backbone NH groups. As shown in Figure 4.28, residues of domain 6
had lower R2 values showing that these residues experience faster tumbling than the overall protein tumbling. Average R1 and R2 values for the folded domains are $1.7 \pm 0.1$ s$^{-1}$ and $11.2 \pm 2.1$ s$^{-1}$. While the average R1 and NOE values for the mutant protein are identical to those of wild type, the R2 values are significant lower than those of the wild type through the whole peptide sequence ($15.4 \pm 2.1$ s$^{-1}$ for wild type protein). Our previous work has shown that the two domains of wild type WD5-6 reorient in solution as a single molecule instead of two [9]. The decreasing R2 values induced by WD5-6(Y532H) can be interpreted as the presence of slower breathing motions in domain 5.

By looking at the structure of wild type WD5-6, the loss of the rigidity of the two domains may arise also from the loss of H-bond interaction between the OH group of Y48 (H532) and the backbone carbonyl of Y75 (Y559). The missing H-bond is also likely to be the reason there are multiple backbone conformations for the linker region (Figure 4.28).

Thus it can be concluded that the mutation Y48H (Y532H) disturbs the hydrophobic packing in domain 5, which is critical for a compact structure, inducing motions on the side chain groups, and breaching the rigidity between the two domains, causing inter-domain motion and structural heterogeneity not observed in wild type WD5-6.
Figure 4.27 Sequential plot of $^{15}$N relaxation parameters for WD5-6 (black) and WD5-6(Y532H in WNDP) (red). Relaxation parameters were measured at 600MHz and 298K.

4.5.3 Copper Transfer Studies of WD5-6(Y532H) with ATOX1 (HAH1)

The interaction of WD5-6(Y532H) with ATOX1 was studied by titrating the $^{15}$N-apo-(WD5-6(Y532H) with unlabelled Cu(I)-ATOX1 and following through $^1$H-$^1$N HSQC spectra. Copper transfer between the mutant protein and ATOX1 were observed. When 2 eq. Cu(I)-ATOX1 were added stepwise to apo WD5-6(Y532H), the copper forms of WD5-6 were identified. The process is slow over NMR time scales, where the intensities of apoprotein signals decreased with increasing ATOX1 concentration and, concomitantly, signals corresponding to Cu(I)-WD5-6(Y532H) appeared and increased in intensity. As shown in Figure 4.29, the residues showing the largest CSV between apo
Figure 4.28 The H-H bonding proximity between Y/H48 (Y/H532 in WNDP) and Y75 (Y559 in WNDP).

Figure 4.29 (Left) Sequential plot of CSV between apo and Cu(I)-WD5-6(Y532H), and (Right), relaxation parameters of WD5-6(Y532H) at the conclusion of titration. The peaks whose line widths are too broad to be correctly integrated are excluded from the list.
WD5-6(Y532H) and apo WD5-6(Y532H) when titrated with 2 eq. Cu(I)-ATOX1 are localized within the CXXC motif of both domains. This indicates that both domains are able to receive one copper ion from Cu(I)-ATOX1, consistent with recently published work on the interaction of ATOX1 with N-terminal domains ofWNDP [21].

4.6 Summary

Several mutations in the gene encoding ATP7B (Wilson protein) are known to cause Wilson disease, a hepatological disorder characterized by copper accumulation mainly in the liver. Among these mutations is the Y532H substitution found at the interface of the N-terminal domains 5 and 6 of the human Wilson protein. The characterization of the Y532H mutation in WD5-6 domains of Wilson protein is reported. The Y532H mutation was introduced into a WD5-6 plasmid construct, pET24d/WD5-6 by site-directed mutagenesis. The resulting plasmid construct, pET24d/WD5-6(Y532H) was subsequently used for protein expression in Rosetta(DE3) cells. Despite the good expression of the mutant protein, it formed an inclusion body, an insoluble form of the protein. Protein extraction using methods that work for the native WD5-6 was unsuccessful. The protein was then prepared as a fusion protein with thioredoxin to increase its solubility.

The fusion protein, TRXaWD5-6(Y532H), also formed inclusion bodies. The protein was recovered from inclusion bodies by refolding and Factor Xa cleavage. The stability of WD5-6(Y532H) obtained by this method was poor; it was prone to degradation. The expression and purification methods for non-fusion WD5-6(Y532H)
were then modified and optimized. This led to purification of stable WD5-6(Y532H) and with high yield.

Biophysical characterization was performed for WD5-6(Y532H) recovered by both refolding/Factor Xa cleavage and optimized purification methods. Only the NMR method was used to probe the atomic organization of WD5-6(Y532H) recovered by refolding/Factor Xa cleavage method. The results revealed that the mutation does not change the overall structure of the native protein although it may be perturb the hydrogen bonding in residues near the mutation site. The gel filtration, light scattering, and circular dichroism data also indicates that the mutation does not change the overall shape and stability of the native WD5-6 protein. A few distinct differences are observed at the molecular level, though. When native WD5-6 is expressed at 37°C, it is readily soluble in the extraction buffer. This is not the case with mutant WD5-6(Y532H); it forms inclusion bodies, which are not readily soluble. The Y532H mutation is a model for how various disease-causing mutations can be characterized. While we found no structural perturbations due to Y532H mutation, we speculate that the mutation may interfere with the protein's refolding in vivo.
References


CHAPeR 5

COuNSeLIO N AND FUTURE DIRECTION

5.1 Significance of the Studies

The main objective of these studies was to elucidate the biophysical properties of the N-terminal domains and disease-causing mutations of the human Wilson protein. As was discussed in previous chapters, the N-terminal domains are unique in that they are found in the heavy metal transporting P-type ATPases. Copper trafficking in cells is orchestrated by copper permeases, copper chaperones, and copper transporting ATPases. Several studies involving isolated domains of the Wilson protein and the related homologous protein Menkes protein has provided initial understanding of the functions and structure of the Wilson protein copper pump. Study of the N-terminal domains of Wilson protein as a whole, however, has been limited by poor availability of stable purified protein at experimental ambient temperatures. The purified N-terminal region readily degrades and is prone to precipitation.

A stable protein is very important, especially in studies that are time-dependent and require the structural integrity to be maintained for a lengthy period of time. Wilson protein structure in its entirety is not yet available since structural information is hard to obtain without a good quality protein sample. Within the APTPB gene exists several mutations that are known to cause Wilson disease. Due to the magnitude and number of these mutations across the gene, the molecular bases of these mutations have not been explored and investigations have been confined to genetic and clinical levels. Therefore, there is need to (1) overcome the challenges that hinder availability of purified stable
protein and, (2) find ways to stabilize the protein in vitro and (2) design a general model that can be used to characterize the disease-causing mutations at large.

5.2 Characterization of N-Terminal Domains of the Human Wilson Protein

We have presented and described the approaches that were applied to overcome the challenges and constraints that prevented the characterization of the N-terminal domains and disease-causing mutations. We showed that expression and purification of the N-terminal domains is dependent on many factors. The amplification of N-terminal domains gene might show DNA polymerase specificity and the protein expression can be dramatically increased by careful optimization of growing culture conditions such as the temperature, inducer, and lengthy of induction. Expression of protein with fusion tags is a widespread technique, but the removal of the fusion tag protein using encoded protease can turn out to be inefficient or fail completely. This phenomena was seen with expressed TRX<sub>xa</sub>/WD1-6 protein, adding another hurdle to the work. We subsequently developed and applied a simple method to substitute and incorporate any protease cleavage site within an existing plasmid vector construct without performing trial-and-error screening of proteolytic sites. It is also possible to probe the stability contributions to a protein of an amino acid residue located at the N-terminal end following the same method.

The purification of the N-terminal domains was also investigated. New knowledge was added on the systemic purification not only of the N-terminal domains, but also of similar recombinant proteins. We showed that large proteins can be released from the cell by sequential extraction process, for example, using a combination of the freeze/thaw
method and BugBuster® reagent extraction. Y532H, a mutation known to cause Wilson disease, was characterized in the attempt to design a model for characterization of the multifold mutations in the ATP7B gene. The purification protocol followed similar approaches to that described in preceding paragraphs. In this work, a native versus mutant approach was adapted. We were able to follow slight deviations between the native and mutant protein. With this approach, shortcomings were easily identified and appropriate action taken. To this end, we report the characterization of Y532H mutation as a model for the characterization of any disease-causing mutation of the Wilson protein.

5.3 Conclusion and Future Work

The improvements in expression and purification of the N-terminal domains of both mutant proteins ushered in the biophysical studies. The findings and major insights were described in detail in chapter 3 and 4. In conclusion, we highlight that (1) novel methods were developed for expression and purification of recombinant proteins, (2) a stable construct of N-terminal domains protein has been successfully purified for biophysical studies and (3) mutations may largely hamper polypeptides refolding in vivo thereby leading to protein misfolding. Further work is recommended in the following areas: (1) the expression and purification of the entire protein pump for possible atomic structure determination will aid in understanding in depth functions of the pump and (2) conformational stability studies of the proteins need to be extended to other methods so that the data can be inclusive. Many disease-causing mutations in the Wilson protein remain uncharacterized; therefore, it would be interesting to see the effect of each and every individual mutation known in the Wilson protein.
Appendix

Project Approval Certification
Recombinant DNA Biosafety Committee

Project Approval Certification

For rDNA Biosafety Committee Use Only

Project Title: Characterization of MetaUoproteins
Principal Investigator: David Huffman
RDBC Project Number: 09-DHd

Date Received by the rDNA Biosafety Committee: 2/27/2009

Reviewed by the rDNA Biosafety Committee

[I] Approved

[ ] Approval not required

2/27/2009
Chair of rDNA Biosafety Committee Signature Date

Revised 5/02 WMU RDBC
All other copies obsolete

201