Probing the Unfolding and the Stability of the Last Four Metal Binding Domains of Wilson Disease Protein Using Circular Dichroism and Novel Spectroscopic Techniques

Ibtesam Yaseen Alja'afreh

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

Part of the Analytical Chemistry Commons, Medicinal-Pharmaceutical Chemistry Commons, and the Organic Chemistry Commons

Recommended Citation


https://scholarworks.wmich.edu/dissertations/366
PROBING THE UNFOLDING AND THE STABILITY OF THE LAST FOUR METAL BINDING DOMAINS OF WILSON DISEASE PROTEIN USING CIRCULAR DICHROISM AND NOVEL SPECTROSCOPIC TECHNIQUES

by

Ibtesam Yaseen Alja’afreh

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

Doctoral Committee:

David Huffman, Ph.D., Chair
Ramakrishna Guda, Ph.D.
Sherine Obare, Ph.D.
Charles Ide, Ph.D.
Wilson disease protein is a copper-transporting P$_{1B}$ type ATPase. It has large N-terminal copper binding domain which is composed of six homologous sub-domains. Each of these six domains is ~72 residues and connected to one another by linking regions of various lengths. They all possess similar ferrodoxin fold, and metal-binding motif, MXCXXC.

The need of having six metal binding domains and the manner in which they are communicating with each other is not well understood. To better understand how the last four metal binding domains function, I pursued a detailed biophysical characterization of these domains. Using molecular biology I have produced domains 3 through 6 (WLN3-6), domains 4 through 6 (WLN4-6) and domains 5 through 6 (WLN5-6). Unfolding of these constructs was performed chemically using guanidine hydrochloride (GuHCl) and monitored by circular dichroism (CD). WLN5-6 unfolds in two-state model where half of its structure is retained at 3.2 M GuHCl. However, denaturation of WLN3-6 and WLN4-6 produced a double sigmoidal curve in which the first sigmoidal transition almost matches the unfolding of WLN5-6 while the second transition matches the unfolding of WLN4. The thermal unfolding illustrated high stability for all constructs.
The unfolding of a mutant of WLN5-6 where tyrosine 48 mutated to histidine, Y532H in the whole Wilson protein, was studied by CD and femtosecond time-resolved fluorescence spectroscopy. CD showed no significant differences in the thermal and chemical unfolding but the tyrosine fluorescence was reduced because the mutation disturbs the hydrophobic core of the protein.

To study the unfolding of the protein efficiently, I introduced two-photon absorption (2PA) cross-sections as novel fluorescence technique to monitor the unfolding of proteins and imaging the change in its local electric field. Green Fluorescent Protein (GFP) was used to test our hypothesis and the method was applied in WLN5-6 after binding it to an extrinsic CPM dye. Two constructs were produced, with CPM at positions 51 and 6, WLN5-6C51_CPM and WLN5-6C6_CPM. In both constructs the anisotropy and the 2PA cross-sections increased with protein unfolding and matches the CD results, this result suggests the validity of this method to monitor protein unfolding.
Copyright by
Ibtesam Yaseen Alja’afreh
2014
ACKNOWLEDGMENTS

Foremost, I would like to express my sincere gratitude to my advisor Dr. David Huffman for his endless support, guidance, patience, and advice in the course of this project. I admire and would love to acquire his intrinsic qualities of patience, objective judgment and decision making.

In addition, I would like to express my deepest gratitude to Dr. Ramakrishna Guda, committee member and collaborator in most of my projects. Without his constant guidance, motivation, enthusiasm, and immense knowledge this work would not have been possible.

I would also like to thank the other members of my dissertation committee, Dr. Sherine Obare and Dr. Charles Ide, for their availability in my committee meetings and also for their encouragements, and insightful comments.

I also convey my gratitude to the dedicated faculty and staff at Western Michigan University. Dr. Michael Barcelona, Dr. Donald Screiber, Dr. Andre Venter, Dr. Ekkhard Sinn, Dr. Steve Bertman, Dr. Gary Blackmer, and Dr. Blair Szymczyna. Staff members: Pam McCartney, Robin Lenkart, and Sean Bashaw.

In my daily work I have been blessed with a friendly and cheerful group of fellow students. I am grateful for the past and present members of both Dr. Huffman’s and Dr. Guda’s research group who have been invaluable for practical help, discussions, and constructive criticism. Past members include: Dr. Joshua Muia, Dr. Alia Hinz, and Ms.
Acknowledgments—Continued

Nastaran Marzijarani. Present members include: Dr. Jameel Hasan, Dr. Edwin Mghanga, Mr. Saad Alotaibi, Mr. Abu Bakar Abu Hagar, Mr. Javan Kisaka and Mr. Mohammad Hatshan.

Last, but not least, I would like to dedicate this dissertation to my parents for all their love and encouragement, to my brothers and sister who have loved me and supported me in all the endeavors.

In conclusion, I recognize that my study would not have been possible without the financial support from Tafila Technical University, which provided me with a scholarship for four years of my study, the Graduate College at Western Michigan University for research and travel grants, and the Department of Chemistry for teaching assistantships and recognition awards.

Ibtesam Yaseen Alja’afreh
# TABLE OF CONTENTS

ACKNOWLEDGMENTS ........................................................................................................ ii

LIST OF TABLES .................................................................................................................. xi

LIST OF FIGURES ................................................................................................................ xiii

LIST OF ABBREVIATIONS .................................................................................................. xx

CHAPTER

1. COPPER CHAPERONES AND TRANSPORTERS ............................................................... 1

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

1.2 Function and transporting of copper ........................................................................ 3

1.3 High affinity copper uptake proteins ........................................................................ 5

1.4 Copper transporter protein 1 (CTR1) ....................................................................... 7

1.5 Copper metallochaperones ....................................................................................... 10

1.5.1 COX17 ................................................................................................................ 11

1.5.2 CCS .................................................................................................................... 13

1.5.3 ATX1-type copper chaperones .......................................................................... 16

1.5.4 Putative chaperones .......................................................................................... 18

1.6 ATOX1 (HAH1) ...................................................................................................... 19

1.7 Copper transporting ATPases ................................................................................... 20

1.7.1 Menkes disease protein .................................................................................... 21

1.7.2 Wilson disease protein ..................................................................................... 22

1.8 Copper transfer from HAH1 chaperone to Cu-ATPase ........................................... 25

1.9 References ................................................................................................................ 29
Table of Contents-Continued

CHAPTER

2. MATERIALS AND EXPERIMENTAL METHODS .................................................. 40

2.1 Introduction ................................................................................................. 40

2.2 Kits, chemicals and reagents ........................................................................ 40

2.2.1 Protein notations and primers ................................................................. 42

2.3 Instrumentation .......................................................................................... 45

2.4 Procedures .................................................................................................. 48

2.4.1 Construction of WLN3-6 plasmid .......................................................... 48

2.4.2 Transformation, expression and purification of WLN3-6 ....................... 51

2.4.3 DNA plasmid, Cloning and Purification of WLN4-6 ............................... 53

2.4.4 DNA plasmid, Cloning and Purification of WLN5-6 ............................... 56

2.4.5 Transformation, expression and purification of WLN5-6 ....................... 57

2.4.6 Tobacco etch virus (TEV) protease expression and purification ............. 59

2.4.7 $^1$H, $^{15}$N NMR HSQC for WLN3-6, WLN5-6 ...................................... 60

2.5 Biophysical properties of proteins ............................................................... 63

2.5.1 Dynamic light scattering ....................................................................... 63

2.5.2 Fluorescence ............................................................................................ 67

2.5.3 Fluorescence anisotropy ......................................................................... 71

2.5.4 Measurements of fluorescence anisotropies .......................................... 74

2.5.5 Two-photon absorption ........................................................................... 76

2.5.6 Ultrafast fluorescence spectroscopy via upconversion ............................. 77
<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>SECTION</th>
<th>CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2.5.7</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td></td>
<td>2.5.8</td>
<td>Circular dichroism for biological molecules</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>References</td>
</tr>
<tr>
<td></td>
<td>3.</td>
<td>BIOPHYSICAL PROPERTIES OF WLN3-6, WLN4-6 AND WLN5-6</td>
</tr>
<tr>
<td></td>
<td>3.1</td>
<td>Introduction</td>
</tr>
<tr>
<td></td>
<td>3.2</td>
<td>Biophysical characterization and analysis of WLN5-6</td>
</tr>
<tr>
<td></td>
<td>3.3</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td></td>
<td>3.4</td>
<td>Biophysical characterization and analysis of WLN4-6</td>
</tr>
<tr>
<td></td>
<td>3.4.1</td>
<td>Chemical and thermal denaturation of WLN4-6</td>
</tr>
<tr>
<td></td>
<td>3.4.2</td>
<td>Isoelectric point (pI) of WLN4-6</td>
</tr>
<tr>
<td></td>
<td>3.4.3</td>
<td>Dynamic light scattering</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
<td>Biophysical characterization and analysis of WLN3-6</td>
</tr>
<tr>
<td></td>
<td>3.5.1</td>
<td>Mass spectroscopy (MS)</td>
</tr>
<tr>
<td></td>
<td>3.5.2</td>
<td>Dynamic light scattering</td>
</tr>
<tr>
<td></td>
<td>3.5.3</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td></td>
<td>3.5.4</td>
<td>Isoelectric point</td>
</tr>
<tr>
<td></td>
<td>3.6</td>
<td>High resolution gel filtration</td>
</tr>
<tr>
<td></td>
<td>3.7</td>
<td>Secondary structure of WLN5-6, WLN4-6 and WLN3-6</td>
</tr>
<tr>
<td></td>
<td>3.8</td>
<td>Summary and conclusion</td>
</tr>
<tr>
<td></td>
<td>3.9</td>
<td>References</td>
</tr>
</tbody>
</table>
# Table of Contents-Continued

## CHAPTER

4. **WILSON DISEASE CAUSING MUTANT Y532H** .................................................. 128

4.1 Introduction ........................................................................................................ 128

4.2 Experimental methods ....................................................................................... 131

4.2.1 Expression of Y532H ..................................................................................... 131

4.2.2 Circular dichroism ......................................................................................... 132

4.2.3 Fluorescence spectroscopy ............................................................................. 133

4.2.4 Quantum yield ............................................................................................... 133

4.3 Results and discussion ...................................................................................... 134

4.3.1 Circular dichroism ......................................................................................... 134

4.3.2 Fluorescence measurements ......................................................................... 139

4.3.3 Quantum yield ............................................................................................... 142

4.3.4 Ultrafast time-resolved fluorescence spectroscopy ..................................... 144

4.4 Comparison of Y532H to other mutants V536A, and L492S ......................... 145

4.4.1 Mechanism of ultrafast fluorescence quenching in disease causing mutations .............................................................................................................. 150

4.5 Summary and conclusion ................................................................................. 151

4.6 References .......................................................................................................... 152

5. **GREEN FLUORESCENT PROTEIN UNFOLDING** ........................................... 154

5.1 Introduction ......................................................................................................... 154

5.2 Experimental methods ....................................................................................... 158

5.2.1 Plasmid generation ......................................................................................... 158
### Table of Contents-Continued

#### CHAPTER

5.2.2 Expression and protein purification .................................................. 160
5.2.3 Unfolding of GFP .............................................................................. 164
5.2.4 Fluorescence spectroscopy ................................................................. 164
5.3 Results and discussion .......................................................................... 165
5.3.1 Mass spectrometry ............................................................................. 165
5.3.2 Steady state optical properties ............................................................. 168
5.3.3 Circular dichroism .............................................................................. 171
5.3.4 Fluorescence lifetime .......................................................................... 172
5.3.5 Fluorescence upconversion ................................................................. 174
5.3.6 Two-photon absorption ....................................................................... 177
5.3.7 Two-photon absorption cross-section .................................................. 179
5.4 Summary ............................................................................................... 181
5.5 References ............................................................................................. 182

6. MONITORING THE UNFOLDING OF THE N-TERMINAL DOMAINS 5-6 OF WILSON DISEASE PROTEIN USING EXTRINSIC FLUOROPHORE ................................................................. 185

6.1 Introduction ........................................................................................... 185
6.2 Experimental methods ............................................................................ 189
6.2.1 Construction of plasmid pET24dWLN5-6C51 ..................................... 189
6.2.2 Expression and purification of WLN5-6C51 protein .......................... 193
6.2.3 Binding WLN5-6C51 to CPM dye ...................................................... 194
6.2.4 Cloning, expression and purification of WLN5-6Cys6 protein ......... 195
Table of Contents-Continued

CHAPTER
6.2.5 Circular dichroism ............................................................... 197
6.2.6 Fluorescence spectroscopy ..................................................... 198
6.2.7 Quantum yield ................................................................. 198
6.3 Results and discussion ........................................................... 198
6.3.1 Measuring of quantum yield of WLN5-6C51_CPM .................. 198
6.3.2 Mass spectrometry ............................................................ 200
6.3.3 Circular dichroism ............................................................. 203
6.3.4 Steady-state measurements .................................................. 209
6.3.5 Two-photon absorption cross-section .................................... 215
6.3.6 Time-correlated single photon counting ................................ 217
6.3.7 Fluorescence upconversion .................................................. 218
6.4 Summary ........................................................................... 220
6.5 References ......................................................................... 224

7. SUMMARY AND FUTURE OUTLOOK ........................................... 226
7.1 Summary of dissertation work .................................................. 226
7.2 Significance ....................................................................... 229
7.3 Future outlook .................................................................. 230
7.4 References ......................................................................... 230

APPENDICES
A. Amino acid sequences of WLN5-6, WLN4-6, and WLN3-6 ................ 232
B. Fluorescence lifetime decays and residuals of GFP at various concentrations of GuHCl ................................................................. 234
APPENDICES

C. Lifetime components of GFP at various concentrations of GuHCl.............. 236

D. Project approval certification................................................................. 238
## LIST OF TABLES

1.1: Copper-dependent proteins and their functions ............................................. 3

2.1: Chemicals and reagents ..................................................................................... 41

2.2: Enzymes, biological reagents and process kits .................................................. 42

2.3: PCR primers ........................................................................................................ 43

2.4: Notations of proteins .......................................................................................... 44

2.5: Plasmids and their vectors properties ................................................................. 44

2.6: Instrumentations and their manufacturer and applications .................................. 45

2.7: Columns used with AKTA FPLC to purify proteins ........................................... 46

2.8: Solutions and buffers .......................................................................................... 47

2.9: KOD hot start DNA polymerase protocol .......................................................... 48

2.10: T4 DNA polymerase protocol ........................................................................... 50

3.1: Free energies and m values of unfolding of WLN4-6 ........................................ 101

3.2: Free energies and m values of denaturation of WLN3-6 by GuHCl .................. 114

3.3: Elution volumes of the standards and the proteins ............................................ 118

3.4: Molecular weight of proteins by high resolution gel filtration and the calculated one by amino acid composition ........................................................................ 120

3.5: Percentage of helicity for WLN5-6, WLN4-6, and WLN3-6 ............................ 121

3.6: Physical properties of WLN5-6, WLN4-6 and WLN3-6 .................................... 124

4.1: Free energies of unfolding WLN5-6 and Y532H obtained from chemical and thermal unfolding ........................................................................................................... 139

4.2: Quantum yields of WLN5-6 and Y532H ............................................................. 142

4.3: Free energy, and percentage of helicity calculated from CD measurements at 222 nm .................................................................................................................. 149
List of Tables- Continued

4.4: Lifetimes of WLN5-6, mutants, and free tyrosine in 50 mM phosphate buffer pH 7.0, at room temperature. ........................................................................................................149

5.1: KOD hot start DNA polymerase reaction...........................................................................159

5.2: PCR parameters for amplifying GFP gene ........................................................................159

5.3: T4 DNA polymerase reaction for GFP .............................................................................160

6.1: Sample preparation for mutations in WLN5-6C51 and WLN5-6C6................................192

6.2: Thermocycle parameters for mutagenesis. ......................................................................193

6.3: Quantum yield measurements of WLN5-6C51_CPM at 0, 0.5 and 4 M Guanidine Hydrochloride. ........................................................................................................199
LIST OF FIGURES

1.1: Transport mechanism. The three subunits are shown, with residues of interest depicted in asteric and the copper ion in yellow. ...............................................................9

1.2: Copper transport pathway ..................................................................................................................11

1.3: Model for membrane-dependent copper distribution to CCS and SOD1 .................................15

1.4: The major functional domains of Wilson protein ..............................................................................25

1.5: The proposed mechanism of copper transfer between HAH1 (left, green) and a domain of the Wilson or Menkes protein (right, blue). .................................................................27

2.1: 0.8% Agarose gel showing PCR product of WLN3-6 amplification ............................................49

2.2: Purification of PCR product of WLN3-6 gene amplification: (A) gel photo after excised WLN3-6 gene. (B) gel after purification of the gene .................................................................50

2.3: Digestion of WLN3-6 plasmid ........................................................................................................51

2.4: 12.5% SDS-PAGE gel shows steps of WLN3-6 expression and purification ............................53

2.5: 0.8% Agarose gel shows purified PCR product of amplifying WLN4-6 gene ............................54

2.6: Digestion of WLN4-6 plasmid ........................................................................................................54

2.7: 12.5% SDS-PAGE showing purification steps of WLN4-6 ............................................................55

2.8: SDS-PAGE gel shows pure WLN4-6 protein .................................................................................55

2.9: TEV purification steps. A, cell extract. B, flowthrough. C, eluted protein ..................................60

2.10: $^1$H,$^{15}$N-HSQC results for WLN5-6 and WLN3-6 .....................................................................62

2.11: Schematic representation of speckle pattern. Applications of DLS to Proteins ....................66

2.12: Fluorescence process .......................................................................................................................68

2.13: Schematic diagram of TCSPC setup ..............................................................................................70
List of Figures-Continued

2.14: L and T format to measure fluorescence anisotropies ....................................................... 74
2.15: Schematic diagram of upconversion ...................................................................................... 79
2.16: Scheme for fluorescence upconversion measurements ........................................................... 80
2.17: CD for proteins of different secondary structure .................................................................. 84
3.1: Cartoon structure of WLN3-4, WLN5-6, and WLN2 ............................................................... 91
3.2: ESI-IT/MS data showing the molecular weight of WLN5-6 determined by MagTran software .......................................................................................................................... 92
3.3: Molar ellipticity for the unfolded WLN5-6 using GuHCl .......................................................... 93
3.4: Fraction denatured of WLN5-6 as GuHCl is added ................................................................. 95
3.5: ΔG of WLN5-6 at different [GuHCl], at 293 K ........................................................................ 96
3.6: Thermal unfolding of WLN5-6 ................................................................................................ 96
3.7: Analysis of thermal unfolding for WLN5-6. ........................................................................... 97
3.8: ESI-IT/MS data showing the molecular weight of WLN4-6 determined by MagTran software .......................................................................................................................... 99
3.9: Unfolding of WLN3-6 by GuHCl ............................................................................................ 100
3.10: Thermal unfolding of WLN4-6 .............................................................................................. 102
3.11: Analysis of thermal unfolding of WLN4-6 ............................................................................ 103
3.12: Phast gel 3-9 and isoelectric focusing of WLN4-6 ................................................................. 105
3.13: Autocorrelation fit of the hydrodynamic radius of WLN4-6 ................................................ 106
3.14: ESI-IT/MS data showing the molecular weight of WLN 3-6 determined by MagTran software .......................................................................................................................... 108
3.15: Mass distribution of WLN3-6 protein versus volume. The solid lines (line) indicate(s) the trace from refractometer while "dots" are weight average MW ........... 110
3.16: Hydrodynamic radius (Rh) determined from “on line” dynamic light scattering of WLN3-6 ............................................................................................................................... 111
List of Figures-Continued

3.17: Autocorrelation fit of hydrodynamic radius of WLN3-6.................................................................111
3.18: Unfolding of WLN3-6 by [GuHCl]..................................................................................................113
3.19: Thermal unfolding of WLN3-6........................................................................................................115
3.20: Analysis of thermal unfolding of WLN3-6.......................................................................................115
3.21: Phast gel 3-9 and isoelectric focusing of WLN3-6..........................................................................116
3.22: Standards calibration curve of high resolution gel filtration........................................................118
3.23: Calibration curve of the standard proteins Stoke’s radius ............................................................119
3.24: CD scan of WLN5-6, WLN4-6, and WLN3-6 ..................................................................................121
3.25: Unfolding of WLN5-6, WLN4-6, and WLN3-6 by GuHCl................................................................122
3.26: Thermal unfolding of WLN5-6, WLN4-6, and WLN3-6..................................................................123
4.1: Cartoon structure of WLN5-6 (PDB:2EW9), and WLN5-6 Y532H (Tyr48 replaced by His)............131
4.2: CD scan for Y532H ..............................................................................................................................135
4.3: Thermal unfolding of Y532H monitored at 222 nm..........................................................................136
4.4: CD scan for folded and unfolded sample of Y532H .......................................................................136
4.5: Unfolding of Y532H by GuHCl monitored at 222 nm ....................................................................137
4.6: Analysis of chemical unfolding of Y532H to calculate free energy in water ............................138
4.7: Thermal unfolding analysis to obtain free energy .......................................................................138
4.8: Emission scans of Tyrosine in Y532H at different [GuHCl]...............................................................139
4.9: Fluorescence counts at 310 nm for Y532H at different [GuHCl].......................................................140
4.10: Emission scans for Tyrosine in WLN5-6 (wild type).....................................................................141
4.11: Fluorescence counts at 310 nm for WLN5-6..................................................................................141
List of Figures-Continued

4.12: Absorbance scan for quantum yield measurements of WLN5-6 and Y532H ..............143
4.13: Fluorescence scans of WLN5-6 and Y532H for quantum yield measurements ..........143
4.14: Fluorescence decay of WLN5-6, Y532H, and Tyrosine ........................................145
4.15: Cartoon structure of WLN5-6 (PDB:2EW9) showing the position of V536 and L492 ..146
4.16: Unfolding of WLN5-6, Y532H, L492S, and V536A by GuHCl monitored at 222nm ...147
4.17: CD scan of WLN5-6, Y532H, V536A, and L492S ....................................................148
4.18: Fluorescence decay of Tyrosine, WLN5-6, and the mutants (Y532H, L492S, and V536A)..................................................................................................................150
4.19: Cartoon structure of WLN5-6 showing the hydrogen bonding between Y48 (Y532), Y75 (Y559), and Y110 (Y594).....................................................................................151

5.1: Formation of GFP chromophore..................................................................................155
5.2: Cartoon structure of GFP, (PDB: 1GFL)......................................................................156
5.3: 15 % SDS-PAGE showing GFP after purification.......................................................164
5.4: ESI-IT/MS data showing the molecular weight of GFP determined by MagTran software.................................................................................................................................167
5.5: Absorbance scan and emission scan of GFP, Excitation wavelength is 470 nm.........168
5.6: Absorbance scan of GFP at different [GuHCl] and the inset represents the absorbance for each sample at 480 nm and 380 nm .................................................................169
5.7: Emission scan of GFP at different [GuHCl], the inset showing the fluorescence counts for each sample at 507 nm ....................................................................................170
5.8: Quantum efficiency of GFP..........................................................................................171
5.9: CD scan for GFP..........................................................................................................171
5.10: CD scan of GFP at different [GuHCl], inset shows the molar ellipticity at 215 nm.....172
5.11: Lifetime of GFP chromophore at various GuHCl concentrations .........................174
5.12: Fluorescence decay traces as a function of increasing Guanidine hydrochloride. Also, shown is the average lifetime with increasing GuHCl following the unfolding transitions ................................................................. 176

5.13: Anisotropy values of GFP at various [GuHCl] .................................................................................. 176

5.14: Log-log graph for GFP showing a slope of 2. .................................................................................. 178

5.15: Fluorescence emission spectra of GFP at 800 nm excitation, inset shows the fluorescence counts at 507 nm at various [GuHCl] ................................................................. 178

5.16: Two-photon absorption scan of GFP at different [GuHCl], inset shows the relative 2PA cross-section ............................................................................................................... 179

5.17: Change in local electric field in GFP as it unfolds ........................................................................... 180

6.1: CPM dye structure and covalent bond formation between maleimide group and thiol group. ........................................................................................................................... 187

6.2: Cysteine positions in both constructs of WLN5-6 ............................................................................ 188

6.3: Constructions of WLN5-6C51 ........................................................................................................ 190

6.4: Scheme represents the mutations performed for WLN5-6C6 ........................................................................ 196

6.5: ESI-IT/MS data showing the molecular weight of WLN5-6C51 determined by MagTran software .................................................................................................................. 200

6.6: ESI-IT/MS data showing the molecular weight of WLN5-6C51_CPM determined by MagTran software .................................................................................................................. 201

6.7: ESI-IT/MS data showing the molecular weight of WLN5-6C6 determined by MagTran software .................................................................................................................. 202

6.8: ESI-IT/MS data showing the molecular weight of WLN5-6C6_CPM determined by MagTran software .................................................................................................................. 203

6.9: CD signal for the unfolding of WLN5-6C51 by GuHCl ........................................................................ 204
List of Figures-Continued

6.10: Free energy of WLN5-6C51 by linear extrapolation method.................................205
6.11: CD signal of the unfolding of WLN5-6C51_CPM..................................................205
6.12: Free energy of WLN5-6C51_CPM.....................................................................206
6.13: CD signal of the unfolding of WLN5-6C6 by GuHCl............................................207
6.14: Free energy of WLN5-6C6..................................................................................207
6.15: CD signal of the unfolding of WLN5-6C6_CPM by GuHCl.................................208
6.16: Free energy of WLN5-6C6_CPM.....................................................................208
6.17: Absorption scan of WLN5-651_CPM at various [GuHCl]..................................210
6.18: Emission scan of WLN5-6C51_CPM at various [GuHCl]....................................210
6.19: log-log plot for two-photon absorption measurements......................................211
6.20: Emission scan of two-photon absorption for WLN5-6C51_CPM at various [GuHCl]...212
6.22: Neutralized emission maximum counts for WLN5-6C51_CPM at 480 nm at different [GuHCl]........................................................................................................212
6.22: Absorbtion scan for WLN5-6C6_CPM at various [GuHCl], inset shows $\lambda_{\text{max}}$ changes while protein unfolds. ................................................................................213
6.23: Emission scan of WLN5-6C6_CPM at various [GuHCl], inset shows Fl. counts at 478 nm with respect to [GuHCl].................................................................214
6.24: Emission scan of two-photon absorption for WLN5-6C6_CPM at various [GuHCl].....214
6.25: Neutralized emission maximum counts for WLN5-6C6_CPM at 480 nm as a function of [GuHCl].................................................................................................215
6.26: Two-photon absorption cross-section for the control, WLN5-6C51_CPM and WLN5-6C6_CPM........................................................................................................216
List of Figures-Continued

6.27: Change in lifetime with unfolding by GuHCl for the control, WLN5-6C51_CPM and WLN5-6C6_CPM. .......................................................... 218

6.28: Change in Anisotropy with unfolding by GuHCl for the control, WLN5-6C51_CPM and WLN5-6C6_CPM. .......................................................... 220

6.29: The direction of the dipole moment in CPM in WLN5-6C51_CPM and WLN5-6C6_CPM with respect to the electric field in WLN5-6 mutant. ............................................. 222
## LIST OF ABBREVIATIONS

- **2PA**  
  Two-Photon Absorption

- **ATP7A**  
  Copper-transporting ATPase 1 or Menkes Disease-Associated Protein

- **ATP7B**  
  Copper-transporting ATPase 2 or Wilson Disease-Associated Protein

- **BCA**  
  Bicinchoninic Acid Assay

- **CCS**  
  Copper Chaperone for Superoxide Dismutase

- **CD**  
  Circular Dichroism

- **COX17**  
  Cytochrome c Oxidase Copper Chaperone

- **CPM**  
  7-Diethylamino-3-(4’-Maleimidylphenyl)-4-Methoxycoumarin

- **Ctr1**  
  Copper Transporter 1

- **DLS**  
  Dynamic Light Scattering

- **DTT**  
  Dithiothreitol

- **EDTA**  
  Ethylenediaminetetraacetic Acid

- **GFP**  
  Green Fluorescent Protein

- **GuHCl**  
  Guanidine Hydrochloride

- **HAH1**  
  Human ATX1 Homologue

- **hCTR1**  
  Human Copper Transporter 1

- **HSA**  
  Human Serum Albumin

- **HSQC**  
  Heteronuclear Single Quantum Coherence

- **IPTG**  
  Isopropyl β-D-1-Thiogalactopyranoside

- **PAGE**  
  Polyacrylamide Gel Electrophoresis

- **PCR**  
  Polymerase Chain Reaction
List of Abbreviations - Continued

PDB .........................................................................................................................Protein Data Bank
SDS ..............................................................................................................................Sodium Dodecyl Sulfate
SOD ............................................................................................................................ Superoxide Dismutase
TCSPC .......................................................................................................................... Time Correlated Single Photon Counting
TEV .............................................................................................................................. Tobacco Etch Virus
TGN .............................................................................................................................. Trans-Golgi Network
TRX .............................................................................................................................. Thioredoxin
WLN ............................................................................................................................. N-terminal domains of Wilson Protein
CHAPTER I
COPPER CHAPERONES AND TRANSPORTERS

1.1 Introduction

Copper plays a vital role in cellular biochemistry and body health. It is needed for energy production, protection against oxidative tissue damage, production of adrenalin, collagen and elastin. Also, it is essential for blood clotting, iron metabolism and antioxidant protection because it acts as an electron donor or acceptor as a component of metalloenzymes\textsuperscript{1–5}.

Dietary copper is likely to be in the Cu\textsuperscript{2+} form. It is absorbed as Cu\textsuperscript{+} in the small intestine and, to a limited extent, in the stomach. So, it needs to be reduced before uptake at the apical membrane of the enterocyte. The major transporter of copper across the apical membrane of the intestinal cell is Ctr1 (copper transporter 1)\textsuperscript{6}. Upon copper entry into the cell, it binds to cytosolic copper chaperones to escort copper to its target protein. The three known Copper chaperones are CCS, COX17 and ATOX1 antioxidant protein 1(also known as HAH1)\textsuperscript{7,8}. The chaperone respective protein targets are Cu,Zn superoxide dismutase (SOD1), cytochrome c oxidase in the mitochondria and ATP7A and ATP7B copper transporting ATPases in the secretory pathway\textsuperscript{9}. Copper transporters exist to translocate Cu(I) from the cytoplasm, but their localization is restricted to the trans-Golgi network and internal vesicles. ATPases (ATP7A and ATP7B) traffic from the
trans Golgi network (TGN) to cytoplasmic vesicles in response to high cellular Cu concentrations where they efflux Cu(I) \(^{10}\).

The ATP7A protein in intestinal enterocytes undergoes copper-induced trafficking. In the response to copper perfusion of the intestine, ATP7A trafficked rapidly to basolateral vesicles. This response of ATP7A facilitates copper delivery into the circulation\(^{11}\). Mutations or deletions in the gene encoding ATP7A are associated with Menkes disease. Menkes syndrome is an X-linked recessive disorder that causes severe systemic copper deficiency and is usually fatal before an individual is 3 years of age\(^{12}\).

ATP7B is found primarily in the liver, with smaller amounts in the kidneys, brain and placenta. It is found in the membrane of Golgi apparatus. In hepatocytes, an increase in copper levels results in trafficking of ATP7B from the TGN to pericanalicular vesicles, and then ATP7B recycles back to the TGN when copper levels are restored\(^{9}\). ATP7B also facilitates the transfer of copper to the major plasma copper transport protein, ceruloplasmin\(^{9,13}\). Inactivation of ATP7B leads to Wilson disease. Wilson disease (also known as hepatolenticular degeneration) was first described in 1912 by S. A. Kinnier Wilson as “progressive lenticular degeneration,” a familial, lethal neurological disease accompanied by chronic liver disease leading to cirrhosis\(^{14}\). It is caused by absent or reduced function of ATP7B protein which leads to decreased hepatocellular excretion of copper into bile. This results in hepatic copper accumulation and injury. Eventually, copper is released into the bloodstream and deposited in other organs, notably the brain, kidneys, and cornea. Failure to incorporate copper into ceruloplasmin is an additional consequence of the loss of functional ATP7B protein\(^{15}\).
Wilson's disease occurs in 1 to 4 per 100,000 people. Copper takes at least five to six years to reach damaging levels in the liver and often 15 to 20 years or more to cause neurological damage. If untreated, Wilson's disease is fatal\(^\text{15}\).

### 1.2 Function and transporting of copper

Copper plays a vital role in cellular biochemistry and body health. It is needed for energy production, and protection against oxidative tissue damage. It is essential for production of adrenalin, collagen and elastin. It is required for growth, cardiovascular integrity, neuroendocrine function, and lung elasticity. Nevertheless, it is essential for blood clotting, iron metabolism and antioxidant protection because it acts as an electron donor or acceptor as a component of metalloenzymes\(^1^4\), Table 1.1.

**Table 1.1: Copper-dependent proteins and their functions\(^\text{16}\)**

<table>
<thead>
<tr>
<th>Example</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azurin</td>
<td>Electron transfer in respiration</td>
</tr>
<tr>
<td>Ceruloplasmin</td>
<td>Iron and copper transport</td>
</tr>
<tr>
<td>Cu, Zn Superoxide dismutase</td>
<td>Superoxide dismutation</td>
</tr>
<tr>
<td>Cytochrome c oxidase</td>
<td>Electron transfer in respiration, Reduction of O(_2) to H(_2)O</td>
</tr>
<tr>
<td>Dopamine β-hydroxylase</td>
<td>Catecholamine production</td>
</tr>
<tr>
<td>Hemocyanin</td>
<td>O(_2) Transport</td>
</tr>
<tr>
<td>N(_2)O reductase</td>
<td>Reduction of N(_2)O to N(_2)</td>
</tr>
<tr>
<td>Plastocyanin</td>
<td>Electron transfer in photosynthesis</td>
</tr>
<tr>
<td>Tyrosinase</td>
<td>Pigment formation</td>
</tr>
</tbody>
</table>

Copper is present in body tissue in trace amounts, but the highest amount is found in the liver, followed by the brain, kidney and heart. An adult body contains around 100 mg of copper\(^3\). It usually bound to protein complexes; 95% of blood plasma copper is bound to ceruloplasmin, while the rest is bound loosely to other blood proteins such as
albumin, transcuprein, and amino acids. However, for a given ceruloplasmin concentration in healthy people, the total copper concentration is varying by 50 µg/dL (8 µmol/L)\(^{17}\).

It is recommended to have adequate copper intake, which ranges between 0.34 mg/day for 1-3 years old to 0.90 mg/day for adults\(^{18}\). Beef liver is considered the richest dietary source of copper. Some beans, nuts, and legumes are particularly high in copper, such as cashews, sunflower seeds, and cooked lentils. Dark chocolate and enriched cereals and rice would also help add copper to diet.

Absorption of copper occurs in its ionic form in the small intestine and, to a limited extent, in the stomach. However, intestinal absorption of copper limited by other metal ions like zinc and cademium, and sulfide ions\(^{19,20}\). The presence of some amino acids and ascorbic acid in the intestine also lower the absorption of copper\(^{21,22}\). Absorbed copper carried in the blood by different transport proteins such as ceruloplasmin, albumin and transcuprein and/or histidine\(^{23}\). Once it is in the blood, it is distributed in two stages; in the first stage, most taken up by liver and kidney; in the second, it re-emerges as buried in the ceruloplasmin. Later, after bound to ceruloplasmin, it is taken by other tissues\(^{24}\). The fact that copper concentration in a mammalian liver is around 5 mg/g wet weight and in the kidney is 7-12 mg/g wet weight is considered the highest in the mammalian organs\(^5\). This fact may reflect that they are the first tissues where dietary copper is deposited and that both organs synthesize and secrete ceruloplasmin\(^{24}\).

Basically, ceruloplasmin-copper is the source of copper for body tissues. Ceruloplasmin exchanges its copper at the surface of cell membrane. A transmembrane protein called CTR1(copper transporter 1) is the primary copper transporter to the
intracellular compartments. Dietary copper is likely to be in copper(II) form; therefore, it has to be reduced before it transports through CTR1 pore, which achieved by dietary reductants like ascorbate and cupric reductases or some ferric reductases.\(^{25}\)

Upon copper (I) entry into the cell, it may be channeled to a copper storage protein known as metallothionein or binds to cytosolic copper chaperones to escort copper to its target protein. The three known Copper chaperones are CCS, COX17 and ATOX1 antioxidant protein I(also known as HAH1).\(^{7,8}\) The chaperone respective protein targets are Cu,Zn superoxide dismutase (SOD1), cytochrome c oxidase in the mitochondria, and ATP7A and ATP7B copper transporting ATPases in the secretory pathway.\(^{9}\) ATPases traffic from the trans Golgi network (TGN) to cytoplasmic vesicles in response to high cellular Copper concentrations where they efflux copper (I) ions.\(^{26}\)

1.3 **High affinity copper uptake proteins**

Ceruloplasmin is the major copper-carrying protein in the blood. Its gene in the human called CP. It is located on the long q arm in chromosome 3 between positions 23 and 25 from base pair 148,880,196 to base pair 148,939,831.\(^{27}\)

Ceruloplasmin, also called ferroxidase, functions to move iron from tissue into the blood and to prepare iron for incorporation into a transferrin, which is a molecule that transports the iron to red blood cells.\(^{28}\)

Ceruloplasmin is made in liver and brain. The form that is made in liver called serum ceruloplasmin, and it moves iron from most of the body except the brain because it is not able to enter the brain. Another form ceruloplasmin is made in brain, nervous system cells called glia, and called glycosylphosphatidylinositol (GPI)-anchored form;
this form is important for transporting iron in the brain and releasing it from brain tissue\textsuperscript{28}.

Ceruloplasmin carries more than 95\% of the total copper in healthy human plasma. It contains 6 atoms of copper in its structure. It possesses a copper-dependent activity which allows it to function in oxidation of ferrous ion into ferric ion that transferring can carry\textsuperscript{4}.

The molecular weight of the human ceruloplasmin is reported to be 151 KDa\textsuperscript{29}. Its structure has been solved by X-ray structural study in 1996\textsuperscript{30}. It is comprised of six cuprodexin type domains. Domains one and six are arranged to have three copper ions in their interface. These ions form a trinuclear cluster. Domains 2, 4, and 6 have also mononuclear sites of copper ions. In these sites, each copper ion coordinated to two histidine, one cysteine and weakly bonded fourth ligand of methionine residue, except that for the one in domain 2 which replaced by leucine. These arrangements of all six copper atoms are same as in ascorbate oxidase.

The albumins are water-soluble proteins; the most common one is serum albumin. Human serum albumin (HSA) is the most prominent protein in plasma, it concentration in blood is about 7 M, it is synthesized in liver and exported as a single non-glycosylated chain. HSA has an extraordinary ligand binding capacity specially for the compounds which are not soluble in aqueous media such as unesterified fatty acids. It usually binds up to two moles of unesterified fatty acids in normal conditions, but can accommodate up to six moles under certain disease conditions\textsuperscript{31}.

Albumin is the major zinc transporter in plasma, typically binds about 80\% of all plasma zinc. It also binds heme that is released in the bloodstream in order to transfer it to
human hemopexin. Furthermore, HSA has a good binding capacity for water, hormones, bilirubin, drugs and ions such as, calcium, sodium, potassium and copper. It has enzymatic properties and function to regulate the colloidal osmotic pressure of blood. HSA is composed of three homologous domains. Each domain has two separate subdomains connected by random coil.

1.4 Copper transporter protein 1 (CTR1)

CTR1 is conserved from overall structures from yeast to human and known to have high affinity to copper (I). Its gene was first discovered as an important gene for iron transport in *Saccharomyces Cerevisiae*. Mammalian CTR1 is a symmetrical homotrimer, where each monomer has three transmembrane domains, extracellular N-terminal domain, and short intracellular c-terminal domain. All nine transmembrane domains form a pore in the lipid bilayer that is a putative for importing of copper(I). This protein also imports Ag(I), which is considered as possible competitor for copper(I) transport because both have similar outer-electron properties. Interestingly, this protein is also able to transport cisplatin which is a platinum anticancer drug.

The human CTR1(hCTR1) gene is located on chromosome 9, more specifically 9q31/32. It is considerably smaller than in yeast. It consists of 190 amino acids. To transport copper ions, hCTR1 does not need ATP hydrolysis or ion gradient, which suggest that it is not an ion pump or a secondary transport. Structural data, functional motion, cooperative dynamics, and available mutagenesis, all those strategies provide new insight toward hCTR1 transportation mechanism, Figure 1.1. Copper ions are transported one at a time, and the TM region of hCTR alternates
between four conformations\textsuperscript{40}, Figure 1.1. In the first step ions do not bind to TM domain, the Met150 and Met154 triads regulate the passing of single ion every each cycle by controlling gating and selectivity. In the second step, activation happen due to pH shift or ion binding and this involves conformational change at the cytoplasmic end and a rotational movement at the extracellular end at the same time. At this point, copper ions bind to Met154 and interactions between protonated His139 and Gly84 occur to stabilize the change in the conformation. In the third step, another conformational change happens in the extracellular end and makes the copper ion to be passed on to the Met150 triad while Met154 closes the pore entrance. In the last step, copper ion passes freely through the polar pore. After that, it is passed through the C-terminus and from there to the copper chaperones via unrevealed mechanism.
Figure 1.1: Transport mechanism. The three subunits are shown, with residues of interest depicted in asteric and the copper ion in yellow.

A new low-resolution model suggests that the size of the apparent “pore” could be modulated by small movements of TM2 or changes in the rotamer conformation of the Met side chains. Also, hCTR1 Phe153 and Tyr147 appeared to facilitate packing of TM2 within the entrance to the pore. Therefore, substitution of Trp for Tyr147 in hCTR1 destabilized the protein. Also, more conservative replacement by Phe or reduction of the side chain size to Alanine yields transporters that transport copper at about half $V_{\text{max}}$ of...
wild type hCTR\textsuperscript{41}. TM1 and TM3 are packed very closely, and Gly residues of the GlyxxxGly motif found on TM3 are critical for forming a functional and structurally intact transporter. TM3 is more sensitive to steric bulk in yeast CTR3 than in hCTR1\textsuperscript{42}. Interestingly, the larger degree of steric constraint in TM3 of yCTR3 is mirrored by what was observed for TM1. Specifically, the TM1 component of the putative packing interface with TM3 was relatively insensitive to Trp in hCTR1 but not in yCTR3. Moreover, the Trp scan indicated that residues A70 and G71 for hCTR1 and A45, I49 and G50 for yCTR3 at the extracellular end of TM1 were sensitive to the introduction of steric bulk. In addition, the results of the comparative Trp scan leave little doubt that the ‘‘chemistry’’ of copper uptake occurs along the central threefold axis of the trimer where TM2, with some contributions from TM1 and TM3, create a copper-permeable pore through the membrane\textsuperscript{41}.

### 1.5 Copper metallochaperones

Metallochaperones is a term to describe any family of proteins that move metal ions to specific metalloenzyme or cofactor via protein-protein interaction. Copper chaperones are cytosolic peptides having CXXC metal binding motif in their N-terminal region. They were found in bacteria, yeast, plants, and mammalian cells. ATX1, a copper chaperone isolated from \textit{S. cerevisiae}, and copZ, a chaperone in \textit{Enterococcus hirae}, both show ferredoxin-like folds βαβαβ with CXXC motif in the folded chains\textsuperscript{43}. Copper (I) is bound to two cysteine sulfur groups that form a linear bidentate ligand. So far, there are three well-known human copper chaperones: CCS, COX17 and ATOX1, Figure 1.2.
Figure 1.2: Copper transport pathway

1.5.1 COX17

Cox17 is the cytochrome c oxidase copper chaperone, which first identified in yeast, it is also conserved in other eukaryotes. Yeast Cox17 is a 69-residue protein localized to both the cytosol and mitochondrial intermembrane space. COX17 homologue protein also reported in humans. Its structure from Saccharomyces cerevisiae was determined and showed unstructured N-terminal region followed by two helices and several unstructured C-terminal residues. It contains six conserved cysteines of which three are essential in copper coordination. These cysteines are: Cys23, Cys24, and Cys26, in a CCXC motif. COX17 binds three or four copper (I) ions in a poly nuclear cluster.
arrangement as indicated by X-ray absorption spectroscopy and luminescence data\textsuperscript{48,49}. The protein oligomerized into a dimer or a tetramer. Loaded Copper COX17 is primarily dimeric in the cytosol, while the one localized within the intermitochondrial membrane is predominantly tetrameric\textsuperscript{48}. Oligomerization may be important for delivering multiple copper ions.

Two other proteins COX11 and SCO1 have been implicated in copper ion delivery and insertion into cytochrome c oxidase through in vitro experiments. These proteins are anchored to the mitochondrial inner membrane through a transmembrane \(\alpha\)-helix and a large soluble copper binding domain facing in the intermembrane space where copper transfer occurs\textsuperscript{45,50}. COX17 interacts with SCO1 and COX11, as it possess number of conserved positively charged residues, which are complimentary with the negative residues on the surface of SCO1 and COX1. Therefore COX17 is a specific copper donor to both proteins\textsuperscript{51}. Copper metallation of these proteins is an intermediate step in the transfer of copper to the Cu\textsubscript{A} site in COX2 by SCO1 and to the Cu\textsubscript{B} site in COX1 by Cox11\textsuperscript{52}.

The NMR solution structure of the partially oxidized human Cox17 shows a completely unstructured N-terminal tail followed by two disulfide bonds, Cys\textsubscript{25}–Cys\textsubscript{54} and Cys\textsubscript{35}–Cys\textsubscript{44}, which stabilizes a structure consists of a coiled coil helix-coiled coil-helix domain. The redox properties of the disulfide bonds maintain the assumption that COX17 upon oxidization by Mia40 becomes partially structured and trapped in the intermembrane space. In humans, the copper(I) ion is coordinated by the sulfurs of Cys\textsubscript{22} and Cys\textsubscript{23}, which is the first example of CC binding motif in copper proteins\textsuperscript{53}. 

12
Eukaryotic cytochrome c oxidase (CcO) forms the terminal enzyme of the electron transport chain of cells. It is located within the mitochondrial inner membrane with a 37Å portion sticking out into the intermembrane space and a 32Å portion extending into the matrix$^{54,55}$.

Mammalian CcO assembly is composed of 13 subunits, and it is dependent on the insertion of several cofactors necessary for function, including two hemes and several metal ions: three copper ions, zinc, magnesium, and sodium$^{56}$. Both hemes are cofactors in subunit 1 (COX1), one of which interacts with a mononuclear copper site (Cu$_B$) forming a heterobimetallic active site (heme a$_3$-Cu$_B$); whereas, subunit 2 (COX2) contains two copper ions in a binuclear center (Cu$_A$)$^{54}$.

1.5.2 CCS

CCS is a chaperone for copper-zinc superoxide dismutase (SOD1). Its gene was first identified as a gene involved in the lysine biosynthesis in yeast, named LYS7$^{57}$. Human CCS is located in chromosome 11, more specifically 11q13. It is expressed in a wide number of organs. It is a 70 kDa intracellular homodimer which, so far, is the largest known metallochaperone$^{57}$. It folds into three functional domains; whereas, other chaperones ATOX1 and COX17 represent a single domain protein. These domains are: the N-terminal domain (domain I), central domain (domain II), and a short C-terminal domain (domain III). The N-terminal domain is very homologous to ATX1, including its copper binding site MxCXXC. This domain is proposed to work only under extreme limiting conditions of copper supply$^{58}$. The second domain is about 50% identical to
SOD1 that even a single mutation targeting aspartate in the fourth potential copper ligand D200H is sufficient to turn CCS into a SOD-like molecule with superoxide scavenging activity\textsuperscript{59}.

The major difference between domain II and SOD1 is that the former lacks two loop-regions corresponding to the SOD1 zinc subloop and electrostatic channel loop, which form the SOD1 metal binding sites and active site channel. As a result, there is no metal-binding cavity in domain II.

The C-terminal domain is highly conserved among CCS from various species. It consists of about 30 amino acid residue, including an invariant CXC motif that can bind copper. It is crucial for the activation of SOD1 in vitro. The crystal structure of this domain revealed that it is a disordered domain, but it is predicted to lie in the vicinity of domain I\textsuperscript{60}. However, some models suggested that domain III and II work together to insert copper into the active site of the target protein SOD1\textsuperscript{58,61}. Based on the affinities of domains I and III to copper(I), it has been proposed that domain I plays a role in the acquisition of copper ions; whereas, domain III act to deliver them to SOD1\textsuperscript{62}. Recently, Unger et. al.\textsuperscript{63} described surprising findings that cellular membranes play an important role in initial copper acquisition by CCS. Their data show that SOD1 can engage bilayers, and even CCS:SOD1 complex retains some affinity for the membrane. This finding suggests that copper distribution to CCS, cytosolic SOD1, and potentially other
Chaperones occur in the context of a membrane scaffold, Figure 1.3

Figure 1.3: Model for membrane-dependent copper distribution to CCS and SOD1

CCS forms a heterodimer with SOD1, and it is believed that domain II has a role in this interaction, as it resembles SOD1. This interaction is important to transfer copper to the target protein. A study in mammalian cells shows that CCS is also essential to activate mammalian SOD, but the absence of any of them is not lethal in the mouse. The fact that CCS is abundant in erythrocytes suggests it could be a potential marker of copper status. Mammalians which suffer from copper deficiency have been discovered to have a higher concentration of CCS protein.

The cytosols of eukaryotic cells contain a superoxide dismutase, which has been remarkably resistant to evolutionary change. The human SOD1 was first characterized in 1969. It is a homodimeric molecule, which contains a copper- and zinc- binding site in each subunit. The most prominent structural feature of the subunit of this enzyme is a
cylinder made of eight-stranded Greek key β-barrel fold. It catalyzes the conversion of superoxide anion to molecular oxygen and hydrogen peroxide.

1.5.3 ATX1-type copper chaperones

ATX1 is the first copper chaperone discovered. It was first known as a gene expressing a protein with antioxidant activity in *Saccharomyces Cerevisiae* cells that lacked SOD1 activity. The protein was designated as antioxidant 1 (ATX1). The antioxidant activity was then revealed in vitro to be a result of stoichiometric consumption of superoxide by Cu-ATX1 complex, which is the reason for the inactivation of the protein. The physiological activity of ATX1 is to shuttle copper to Cu-ATPase which, in turn, pumps copper ions delivered by ATX1 into the Golgi complex for later insertion into copper enzymes.

ATX1 family chaperones are highly and evolutionary conserved. After the yeast ATX1 discovery, functional homologous chaperones were identified in a variety of eukaryotes, including plants, mice, rats, dogs, and humans. In prokaryotes copper chaperones were sufficiently studied, such as CopZ from Enterococcus hirae and Bacillus, which deliver copper to the Cu-ATPases CopA and CopB. All ATX1-like proteins consist of approximately 70 amino acids and have a distinguishing MXCXXC motif close to the amino terminus.

Studying of ATX1-type copper chaperones, especially bacterial CopZ and yeast ATX1, established the concept of copper metallochaperones. They are both small and similar in amino acid sequences to the N-terminal regions of the copper transporter ATPases proteins, CopA in bacteria, Menkes and Wilson in humans, and Ccc2 in yeast.
All of these proteins have the same metal binding motif associated with the first loop of \( \beta\alpha\beta\alpha\beta\) (ferredoxin) structural fold. The single \( \beta\alpha\beta\alpha\beta\)-fold of Atx1 or CopZ is replicated once (CopA), or twice (Ccc2), or six (Menkes and Wilson) times in the cytosolic regions of the ATPases\(^7^7\).

Coimmunoprecipitations and Cu dependent two-hybrid interactions proved that ATX1 forms associations with the N-terminal region of Ccc2, and that the human homolog Atox1(sometimes called HAH1) forms associations with the Menkes or Wilson ATPases. Moreover, copper transfer between pairs of recombinant proteins was observed in vitro\(^7^8\).

A sequence of ligand exchange reactions was proposed through the solution structure of apo and copper (I) forms of ATX1, as well as for the apo and Cu (I) forms of domain a of Ccc2 (Ccc2a), in addition to the X-ray crystal structure for the homodimer of Cu(I)-HAH1, among others.

Models of docked heterodimers were obtained from the homodimer structure\(^7^9\) and examined by structural studies. Crystallographic and NMR studies reveal that the loop regions in ATX1 and ATOX1 that connect the four \( \beta\) strands and the two \( \alpha\) helices are exposed to the solvent along with the binding sites\(^4^3,7^7\). Whereas, iron sulfur clusters with similar ferredoxins and the catalytic copper sites in oxidases and oxygenases are buried in the protein molecule\(^4^3,8^0\). Therefore, the chaperone metal binding site is well suited to copper delivery rather than to a catalytic or electron transfer function.
1.5.4 Putative chaperones

There are chaperones that interact with specific cuproenzymes and are believed to transfer metal. MURR1 is a protein that is associated with Wilson protein. Its gene encodes a 190-amino acid open reading frame of unknown function or of known motif or homology that is highly conserved in vertebrate species. The deletion of MURR1 gene in animals causes hepatic copper overload. Therefore, MURR1 is required for the delivery of copper to the bile from Wilson protein.

Another possible copper chaperone is metallothionein (MT). It is a family of cysteine-rich protein that has low molecular weight, 500 to 14000 Da. In yeast, they are also known as copper metallothionein (CUP). The two widely expressed isoforms of MT, MT-1 and MT-2, have more diverse functions and may play a role in intracellular copper transfer and storage. In the perinatal liver, copper storage is associated with MT. When copper is limited in liver cells, biliary excretion of copper and holoceruloplasmin synthesis is low suggesting that MT has a role under these conditions by acting as a copper reserve. Copper can induce MT synthesis, but zinc is more likely the physiological inducer.

The third candidate as copper chaperone is APP, a membrane protein which contains a copper-binding site. It has been suggested that APP serves as a barrier to copper import in the brain; this suggestion came after studies in mice. These studies showed that in APP-null mice, the brain copper levels elevated compared to wild type mice, while brain zinc and iron levels were unaffected. Also, on transgenic mice that overexpress APP, the brain copper level reduced.
1.6 ATOX1 (HAH1)

The human ATX1 homologue (ATOX1 or HAH1) gene is located in 5q32 and was first expressed in 1996 by Klomp et al.\textsuperscript{88}. The protein has 68 amino acid residues and a mass of about 7.5 kDa. Its primary role is to shuttle copper to two types of P1B-type ATPases, Menkes (ATP7A) and Wilson (ATP7B)\textsuperscript{89,90}. Therefore, it is required for proper biliary excretion of excess, as well as delivery of copper for holoceruloplasmin\textsuperscript{91}. Moreover, other properties may be ascribed to Atox1. It may function to protect neuron cells from oxidative stress as overexpression of ATOX1 can increase neuronal viability under stress conditions, such as serum oxidation and deficiency\textsuperscript{92}. Also, it has been reported that it is involved in cisplatin resistance. Cisplatin is known as anticancer drug, and it directly binds to Atox1 in the copper binding motif\textsuperscript{93}.

Cells missing Atox1 have impaired movement of ATP7A in response to copper, providing a mechanistic explanation for copper retention and elevation\textsuperscript{94}. ATP7A is normally located in the transgolgi membrane, but when cellular copper concentration rises it moves to the plasma membrane. Atox1-null mice show signs of hypopigmentation. The skin distortion is due to the role of ATOX1 in copper delivery to tyrosinase. Skin distortion is also likely due to reduced levels of lysyl oxidase, the cuproenzyme that has a role in elastic and collagen cross-linking\textsuperscript{91}.

Recently, Kohno et. al. signify a novel role of ATOX1 in prompting vascular smooth muscle cell (VSMC) migration in mice, which specifies its role in neointimal formation after vascular injury, as Atox1 colocalizes with ATP7A in the neointimal VSMCs after wire injury in vivo. Therefore, Atox1 is a potential therapeutic target for VSMC migration and inflammation-related vascular diseases\textsuperscript{95}.
1.7 Copper transporting ATPases

Copper transporting ATPases (Cu-ATPases) are members of a large family of proteins that use energy from ATP hydrolysis to compel membrane transport of ions. Heavy metals in cells are typically regulated by P_{1B}-type ATPases. Cu ATPase expressed in yeast participates in copper insertion into Fet3, a copper dependent oxidase involved in iron uptake and oxidation.

The N-termini of P-type ATPases contain domains that are similar to ATX1 and CopZ. These domains occur in one copy as in bacterial CopA, or two copies as in yeast Ccc2, through to six copies in the human ATP7A and ATP7B ATPases. These ATPase domains have a metal binding motif MXCXXC associated with the first loop of βαββαβ, a ferrodoxin-like fold.

Most of knowledge about the structure and function of P-type ATPases originates from SERCA1, a sarco(endo)plasmic reticulum Ca^{2+}-ATPase from adult rabbits muscle. Human copper ATPases, ATP7A and ATP7B, receive copper ions from HAH1 chaperone then pump them into the lumen of the golgi network for following insertion into the ceruloplasmin.

ATP7A and ATP7B are 50–60% identical. Both are phosphorylated at Ser residue(s). Copper has a role in the regulation of the extent of a kinase-mediated phosphorylation process. Inherited mutations in both ATP7A and ATP7B are responsible for specific disorders of copper metabolism, Menkes and Wilson diseases, respectively.
1.7.1 Menkes disease protein

The gene encodes ATP7A in human is located in chromosome X in the long arm q at position 21.1. The official name of the gene is ATPase, Cu\(^{2+}\) transporting, alpha polypeptide. This gene provides instructions for making a protein that is important in regulating copper levels in the human body. Several mutations in the gene that reduce the production of ATP7A protein are responsible for a condition called occipital horn syndrome. In this syndrome, the protein impairs the absorption of copper and affects its normal distribution which, in turns, reduces the activity of copper-containing enzymes. Therefore, symptoms related to these enzymes appear, such as loose or sagging skin, coarse hair, and calcium deposition at the base of the skull and lose joints.

Mutations in the gene that prevent or delete the production of functional ATP7A protein will lead to Menkes syndrome. Many of these mutations involve deletion of part of the gene, insertion of an additional DNA blocks, or change single nucleotide. The abnormal protein will become unable to shuttle back and forth from its location, golgi apparatus, and stay on the membrane. Eventually, copper deficiency will reduce the activity of certain copper-containing enzymes and copper levels will be lower in tissues, like the brain. On the other hand, copper will accumulate in other tissues, like the kidney and the small intestine. As a result, structure and function of skin, bone, hair, blood vessels, and nervous system will be affected\(^{103}\).

The ATP7A protein is considered a large protein; it consists of 1500 residues with total mass of about 163 kDa. The apparent mass for this protein, when it is analyzed on denatured SDS gels, is higher 175-180 kDa, because of glycosylation\(^{99}\).
The N-terminal region of ATP7A consists of six sub domains. The structures of all of them have been determined individually by NMR and crystallography in both holo and apo forms (Protein Data Bank accession numbers 1KVJ, 1S6U, 2GA7, 2AW0, 1Y3J, and 1YJV).

1.7.2 Wilson disease protein

ATP7B is a gene located on the long q arm of chromosome 13 at position 14.3 that encode Wilson protein (Also called ATPase 2) in humans. Mutations that change one of the amino acids in this gene alter the dimensional structure or stability of the protein, which eventually cause Wilson disease. Introducing a stop codon, and deleting or inserting of small segments of DNA within the gene for making ATPase 2 protein, will result in producing low protein or even no protein produced. These types of mutations result in severe symptoms than a single mutation. As a result of all types of mutations, removal of excess copper from the body is impaired, and therefore, copper accumulates to toxic levels in the liver and the brain.

ATP7B as any P_{1B}-type of ATPases has certain structural features: the transmembrane domain, the ATP-binding domains, which includes two domains: nucleotide-binding domain (N-domain) and phosphorylation domain (P-domain), and the actuator domain (A-domain). The most important distinct feature is the presence of a large N-terminal region comprised of six copper-binding domains, which are connected by linkers of various lengths. It consists of 1465 residues and has a mass of about 157 kDa, but it has a mobility of a 165 kDa protein. In contrast to ATP7A, this protein is not glycosylated.
The membrane portion of ATP7B has 8 transmembrane segments (TMS) with the N- and C-termini of the protein both oriented toward the cytosol. The TMS contain the intra-membrane copper-binding sites, Figure 1.4, to which copper is delivered from the cytosol and from which it is subsequently exported into the lumen of intracellular compartments. All TMS in ATP7B are necessary for the Cu-ATPase folding or function. Furthermore, flexible connections allowing conformational changes are essential for the Cu-ATPase transport activity. Mutagenesis in the TMS suggest that CPC motif in TMS 6, NY motif in the TMS 7, and MxxS motif in TMS 8 are likely to contribute to copper coordination during transport.

Topologically, TM5–TM8 are closely equivalent to TM3–TM6 of the Ca²⁺-ATPase (SERCA1), H⁺-ATPase, or sodium and potassium-ATPase. It is believed that TM4–TM8, along with the cytosolic part, form part of a conserved functional core found in all P-type ATPases.

The ATP-binding domain of Cu-ATPases is located between TM6 and TM7, as all P-type ATPases, it contains conserved sequence motif D-K-T-G-[LIVM]-[TIS] in the P-domain. During ATP hydrolysis, the invariant aspartate residue accepts γ-phosphate from ATP bound in the N-domain and becomes reversibly phosphorylated with the formation of an acylphosphate as intermediate. During cation transport, intrinsic phosphatase activity subsequently dephosphorylates the aspartate residue, as part of the catalytic cycle.

The actuator domain Cu-ATPases is formed by a cytosolic loop, located between TM4 and TM5, contains a highly conserved T/SGE motif that inserts in the cleft between the N- and P-domains during hydrolysis of ATP. Studies in SERCA revealed the role of
the invariant glutamate residue is to activate water molecule during dephosphorylation. Other than that, the entire actuator domain rotates and regulates the release of the ion to the luminal site \(^{108-110}\).

The C-terminal region is approximately 90 residues, and interesting sequence motif of tri-leucines LLL in ATP7B has been observed preceded by a cluster of negative residues DDDGD. Mutation of two of the leucines to alanines results in trapping ATP7B in the vesicles, and not in plasma membrane. It could be the hydrophobic LLL and the negative cluster work together to retain ATP7B in vesicles or to regulate the rate of protein trafficking between intracellular compartments \(^{111}\).

One of the important cytosolic portions of Cu-ATPases is the N terminal region. The N-terminal copper binding domain is composed of 6 homologous sub-domains (WLN). Each of these six domains has ~72 residues and is connected to one another by linking regions of various lengths: 11 residues between WLN1 and WLN2, 42 residues between WLN2 and WLN3, 30 residues between WLN3 and WLN4, 57 residues between WLN4 and WLN5, and 8 residues between WLN5 and WLN6. Triple resonance NMR experiments show that each domain has the \(\beta\alpha\beta\alpha\beta\) fold and houses a copper-binding site, GMXCXXC, in which both invariant cysteines of the CXXC motif coordinate Cu(I) \(^{112}\). Residues 58-130 comprise domain 1, 143-215 domain 2, 256-327 domain 3, 357-428 domain 4, 429-485 the flexible linker, 486-556 domain 5, and 565-632 domain 6. The sub-domains 5 and 6 are connected by a short linker, and their metal-binding sites are spatially far apart \(^{113}\).

ATP7B is known to interact with other proteins: the copper chaperone HAH1, COMMD1, glutaredoxin1, and the p62 subunit of dynactin \(^{114,115}\). In all cases, the N-
terminal segment of ATP7B has been identified as the site of interaction. It has also been shown that this region interacts in a copper-dependent manner with the cytoplasmic nucleotide-binding/phosphorylation domain of ATP7B, with copper decreasing the interaction. An intact and functional N-terminal region is also required for the copper-induced phosphorylation of ATP7B which is, in turn, required for vesicular trafficking.

Figure 1.4: The major functional domains of Wilson protein

The transmembrane portion of Wilson protein is composed of 8 transmembrane segments (dark red) that form intramembrane copper-binding site(s). The ATP-binding domain is composed of the P-domain (purple) and the N-domain (orange) and together with the A-domain (turquoise) is responsible for enzymatic cycle (ATP binding, hydrolysis, phosphorylation, and dephosphorylation). The N-terminus domain has six metal (Cu)-binding subdomains (WLN 1-6, pink).

1.8 Copper transfer from HAH1 chaperone to Cu-ATPase

Human copper chaperone, HAH1, delivers copper ions to CuATPases, ATP7A and ATP7B, which pump the ions into the lumen of the Golgi complex for later insertion into
the ceruloplasmin. This copper transport occurs, both in the normal copper metabolism and in the high influx of copper, under copper metabolism disturbance.

In vivo, HAH1 transfers copper to all six domains in the truncated protein\(^{117,118}\). Although, it forms a detectable Cu(I)-mediated complex only with metal binding domain1 and 4 of ATP7A\(^{119,120}\) and WLN1, WLN2, and WLN4 of ATP7B\(^{118}\). This study implies that the rest of domains, WLN5 and WLN6, which are necessary for metal transport across the membrane, may take delivery of copper from other WLNs. In fact, copper transfer has been shown between WLN1 and WLN4, and WLN4 and WLN5\(^{113,121}\). On the other hand, mouse ATP7B lacks functional WLN4, which makes hypothesis regarding copper transferring from WLN4 to other WLNs complicated\(^{122}\).

Copper transfer between HAH1 and the N-terminal domains requires particular interactions that result in the formation of copper-WLN adducts\(^{123–125}\). The facts that HAH1 has the same overall fold as the individual copper binding subdomains, and the presence of complementary charges at the surface of HAH1 and some subdomains, suggest that HAH1 form adducts with the subdomains and transfer copper via ligand exchange\(^{79,126,127}\).

The crystallographic studies reveal a structure of two HAH1 monomers bridged by a copper. This model could be such an intermediate in copper transfer\(^{79}\). Some studies carried out on yeast ATX1 indicate that ATX1 binding site interact directly to P-type ATPases. The interaction happens between the copper-bridged ATX1dimer and the P-type ATPase using one cysteine of the later. Then a series of two and three copper (I) centers are formed leading to the transfer of copper to ATPase\(^{26,67,124}\). A few Years later, Wernimont et. al. suggested a mechanism for copper exchange between HAH1 and
human Cu-ATPases. In the first step, holo-HAH1 with two coordination copper (I) center interacts with an apo metal binding domain of the ATPases, ATP7A or ATP7B. Then the apo domain binding site donates a cysteine to form a third primary bond to copper. One of the bonds between the HAH1 cysteine is broken and rapidly forms a bond with the cysteine on the ATP7A or ATP7B. Eventually, the HAH1 copper ion transfers to the ATP7A or ATP7B and HAH1 dissociates from the complex, Figure 1.5.

Figure 1.5: The proposed mechanism of copper transfer between HAH1 (left, green) and a domain of the Wilson or Menkes protein (right, blue).

Walker et. al., in a Cys-labeling studies for a full N-terminal domain, suggested that Cu-ATOX1 preferentially delivers copper (I) to WLN2. They also demonstrated a specific role of WLN2 in the first step of HAH1 delivery of copper to Wilson protein, due to high binding affinity between HAH1 and WLN2. They also believe that binding of
copper to WLN2 works as a switch which allows the access of the HAH1 to other metal binding domains rather than a specific entrance to copper. On the other hand, NMR experiments showed that ATOX1 can form a Cu-dependent adduct with not only WLN2, but also with WLN4, when presented as individual domains. WLN4 was also preferred by Cu-ATOX1 for copper delivery in a construct of WLN3-4 by forming an ATOX1-Cu-WLN4 adduct. Other NMR experiments involved a construct with all six WLN domains, suggest WLN1, WLN2, and WLN4 formed adducts with Cu-ATOX1, whereas WLN3, WLN5, and WLN6 did not form a detectable adduct but became metalated by ATOX1.

Similar NMR experiments performed on the construct of WLN3-4, were previously performed on WLN5-6 construct. These experiments revealed that Atox1 did not interact with these domains, but, instead, Cu-loaded WLN4 could interact with WLN6 in the WLN5-6 construct and form a Cu-dependent heterocomplex. However, in Cu-chelator competition experiments using a construct of all metal binding domain in Wilson protein, the affinity for Cu-Atox1 of the individual Cu site appeared similar, and therefore, delivery could occur to all domains.
1.9 References


(12) Greenough, M.; Pase, L.; Voskoboinik, I.; Petris, M. J.; O’Brien, A. W.; Camakaris, J. Signals Regulating Trafficking of Menkes (MNK; ATP7A) Copper-


(52) Horn, D.; Barrientos, A. Mitochondrial Copper Metabolism and Delivery to Cytochrome c Oxidase. *IUBMB Life* **2008**, *60*, 421–429.


CHAPTER II
MATERIALS AND EXPERIMENTAL METHODS

2.1 Introduction

This chapter will explain and list the reagents and chemicals, enzymes and biological kits, instrumentation, notations of proteins, solution compositions, and procedures. Also, this chapter will explain the protein expression, purification, and gel electrophoresis for the proteins investigated in this research which are WLN5-6, WLN4-6, and WLN3-6. Moreover, this chapter will present an overview about the physical properties of proteins, such as fluorescence and circular dichroism.

2.2 Kits, chemicals and reagents

All chemicals and reagents used met molecular grade standards. Water used in all experiments was Milli-Q water with conductivity of 18.2 MΩ/cm purified using US Filter PURELAB plus UV/UF (Millipore). The water and media for cell culture were autoclaved for 30 minutes at 121°C. List of all chemicals used along with the purity and the supplier are in Table 2.1.
<table>
<thead>
<tr>
<th>Chemical/Reagents</th>
<th>Catalog number</th>
<th>Purity</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>$^{13}$C Glucose</td>
<td>CLM-1396-1</td>
<td>99 %</td>
<td>Cambridge Isotope</td>
</tr>
<tr>
<td>$^{15}$N Ammonium chloride</td>
<td>NLM-467-1</td>
<td>99 %</td>
<td>Cambridge Isotope</td>
</tr>
<tr>
<td>Acrylamide</td>
<td>A0268205</td>
<td>&gt; 99%</td>
<td>Acros Organics</td>
</tr>
<tr>
<td>Agarose, ultrapure</td>
<td>16500-100</td>
<td>Ultra pure</td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Amonium persulfate</td>
<td>BP179-100</td>
<td>Electrophoresis grade</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Ampicillin</td>
<td>A-9518</td>
<td>&gt;92.5 %</td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>Bis-Acrylamide</td>
<td>BP171-25</td>
<td>Electrophoresis grade</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Boric acid</td>
<td>BP168-1</td>
<td>≥99 %</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Bromopheno blue</td>
<td>B-6131</td>
<td></td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>Coomassie brilliant blue R-250</td>
<td>20278</td>
<td></td>
<td>Thermo scientific</td>
</tr>
<tr>
<td>CPM</td>
<td>765860</td>
<td></td>
<td>Invitrogen</td>
</tr>
<tr>
<td>Dibasic potassium phosphate</td>
<td>39076</td>
<td>99.5 %</td>
<td>J.T.Baker</td>
</tr>
<tr>
<td>Dibasic sodium phosphate</td>
<td>S-5136</td>
<td></td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>Dithiothritol (DTT)</td>
<td>1758-9030</td>
<td></td>
<td>Inalco Sas</td>
</tr>
<tr>
<td>EDTA</td>
<td>H45476</td>
<td>99.7 %</td>
<td>Mallinckrodt chemicals</td>
</tr>
<tr>
<td>Glycerol</td>
<td></td>
<td>99.9 %</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Guanidine Hydrochloride</td>
<td>G3272-500G</td>
<td>&gt; 99.5 %</td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>Glutathione, reduced</td>
<td></td>
<td>&gt; 99 %</td>
<td>Sigma-Aldrich</td>
</tr>
<tr>
<td>HEPES</td>
<td>BDH4518-1KGP</td>
<td>≥ 99.9 %</td>
<td>VWR</td>
</tr>
<tr>
<td>Imidazole</td>
<td>A10221</td>
<td>99 %</td>
<td>Alfa Aesar</td>
</tr>
<tr>
<td>IPTG</td>
<td>1758-1400</td>
<td></td>
<td>Inalco Sas</td>
</tr>
<tr>
<td>Kanamycin sulphate</td>
<td>420311</td>
<td></td>
<td>Calbiochem</td>
</tr>
<tr>
<td>Magnesium chloride</td>
<td>BP214-500</td>
<td>≥ 99 %</td>
<td>Fischer Scientific</td>
</tr>
<tr>
<td>MEM Vitamins</td>
<td>11120-052</td>
<td></td>
<td>Invitrogen</td>
</tr>
<tr>
<td>MES</td>
<td>BDH4544</td>
<td>99 %</td>
<td>VWR</td>
</tr>
<tr>
<td>Methanol</td>
<td></td>
<td>≥ 99.8 %</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>Monobasic potassium phosphate</td>
<td>4008-01</td>
<td>100.3%</td>
<td>J.T.Baker</td>
</tr>
<tr>
<td>Monobasic sodium phosphate</td>
<td>BP330-1</td>
<td>98 %</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Potassium chloride</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sodium chloride</td>
<td>BDH0286-500G</td>
<td>99 %</td>
<td>VWR</td>
</tr>
<tr>
<td>Sodium dodecyl sulfate (SDS)</td>
<td>15525-017</td>
<td>&gt; 99.5%</td>
<td>Gibcobrl</td>
</tr>
<tr>
<td>Sodium hydroxide</td>
<td>S612-3</td>
<td></td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>TEMED</td>
<td>BP150-100</td>
<td>≥ 97 %</td>
<td>Fisher Bioreagents</td>
</tr>
<tr>
<td>Trizma base</td>
<td>T6066-1KG</td>
<td>≥ 99.9 %</td>
<td>Sigma Aldrich</td>
</tr>
<tr>
<td>Tryptone</td>
<td>12111-1</td>
<td></td>
<td>Mo Bio labs</td>
</tr>
<tr>
<td>Xylene cyanol FF</td>
<td>BP565-10</td>
<td>Biotech grade</td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Yease Extract</td>
<td>DF0127071</td>
<td></td>
<td>Fisher Scientific</td>
</tr>
<tr>
<td>Agar</td>
<td>J637-5</td>
<td></td>
<td>Midsci</td>
</tr>
<tr>
<td>Trichloroacetic acid</td>
<td>BDH0310-500G</td>
<td>99 %</td>
<td>VWR</td>
</tr>
<tr>
<td>Dextrose</td>
<td>BP350-1</td>
<td>≥ 99.8 %</td>
<td>Fisher Scientific</td>
</tr>
</tbody>
</table>
Table 2.2: Enzymes, biological reagents and process kits

<table>
<thead>
<tr>
<th>Biomolecules/Kits</th>
<th>Catalog #</th>
<th>Supplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Kb DNA Ladder</td>
<td>N3232S</td>
<td>NEB</td>
</tr>
<tr>
<td>BamHI</td>
<td>N3231S</td>
<td>NEB</td>
</tr>
<tr>
<td>BglII</td>
<td>R0144S</td>
<td>NEB</td>
</tr>
<tr>
<td>Benzonase</td>
<td>70746</td>
<td>Novagen</td>
</tr>
<tr>
<td>BugBuster® Protein Extraction</td>
<td>70584-4</td>
<td>Novagen</td>
</tr>
<tr>
<td>KOD Hot Start DNA Polymerase</td>
<td>71086-3</td>
<td>Novagen</td>
</tr>
<tr>
<td>MinElute PCR Purification Kit</td>
<td>28004</td>
<td>Qiagen</td>
</tr>
<tr>
<td>pET32Xa/LIC Kit</td>
<td>70072</td>
<td>Novagen</td>
</tr>
<tr>
<td>QIAprep Spin Miniprep Kit</td>
<td>27106</td>
<td>Qiagen</td>
</tr>
<tr>
<td>QIAquick Gel Extraction Kit</td>
<td>28706</td>
<td>Qiagen</td>
</tr>
<tr>
<td>Rainbow Protein Marker</td>
<td>RPN800E</td>
<td>GE</td>
</tr>
<tr>
<td>Quick Change II Site-Directed</td>
<td>200521-5</td>
<td>Stratagene</td>
</tr>
<tr>
<td>Mutagenesis Kit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bio-Rad Protein Assay Reagent</td>
<td>500-0006</td>
<td>Bio-Rad</td>
</tr>
<tr>
<td>BCA™ Protein Assay Kit</td>
<td>23225</td>
<td>Thermo Scientific</td>
</tr>
<tr>
<td>Isoelectric Focusing Calibration</td>
<td>17-0442-01</td>
<td>Amersham Bioscience</td>
</tr>
<tr>
<td>Kit Broad PI(pH3-10)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosetta DE3 Competent Cells</td>
<td>70954-3</td>
<td>Novagen</td>
</tr>
<tr>
<td>NovaBlue</td>
<td>71227-3</td>
<td>Novagen</td>
</tr>
<tr>
<td>Gigasingles™ Competent Cells</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BL21DE3 Competent Cells</td>
<td>69450</td>
<td>Novagen</td>
</tr>
<tr>
<td>LMW Gel filtration Calibration</td>
<td>17-0442-01</td>
<td>Amersham Bioscience</td>
</tr>
<tr>
<td>Kit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HMW Gel Filtration Calibration</td>
<td></td>
<td>Amersham Bioscience</td>
</tr>
<tr>
<td>Kit</td>
<td></td>
<td></td>
</tr>
<tr>
<td>XL-10 Gold Competent Cells</td>
<td>200517-4</td>
<td>Stratagene</td>
</tr>
<tr>
<td>ATP7B gene</td>
<td>Generous donation from Dr. Jonathan Gitlin, University of Washington</td>
<td></td>
</tr>
<tr>
<td>TEV gene</td>
<td>Kind gift from Dr. Timothy A. Cross, National High Magnetic Field Laboratory, Florida State University</td>
<td></td>
</tr>
</tbody>
</table>

2.2.1 Protein notations and primers

Forward and backward primers were designed for amplification of the gene in interest in order to clone it into the preferred vector, and primers used for mutations are listed in Table 2.3 with their number of base pair (bp), GC content percentage, melting point $T_m$, and their sequences. The primers were purchased from, Integrated DNA Technologies, Inc. (Coralville, Iowa).
The name, notation, and the residues in ATP7B corresponding to each protein are listed in Table 2.4.

Table 2.3: PCR primers

<table>
<thead>
<tr>
<th>Primer</th>
<th>bp</th>
<th>% GC</th>
<th>% Tm</th>
<th>Sequence 5'-3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN3-6 (Forward)</td>
<td>54</td>
<td>56</td>
<td>75.2</td>
<td>5'- GGT ATT GAG GGT CGC GAG AAC CTT TAT TTC CAG GGC GTG GTC ACC CTC CAA CTG -3'</td>
</tr>
<tr>
<td>WLN3-6 (Reverse)</td>
<td>51</td>
<td>53</td>
<td>73.4</td>
<td>5'- AGA GGA GAG TTA GAG CCT CAC TGG GCC AGG GAA GCA TGA AAG CCA ATT TCC -3'</td>
</tr>
<tr>
<td>WLN4-6 (Forward)</td>
<td>45</td>
<td>51</td>
<td>70.9</td>
<td>5'- GGT ATT GAG GGT CGC GAG AAC CTT TAT TTC CAG GGC ACA TGC AGT -3'</td>
</tr>
<tr>
<td>WLN4-6 (Reverse)</td>
<td>51</td>
<td>53</td>
<td>73.4</td>
<td>5'- AGA GGA GAG TTA GAG CCT CAC TGG GCC AGG GAA GCA TGA AAG CCA ATT TCC -3'</td>
</tr>
<tr>
<td>WLN5-6 (C499S,C502S)</td>
<td>41</td>
<td>46</td>
<td>67.5</td>
<td>Forward: 5'- GAT CAA AGG CAT GAC CAG TGC ATC CAG TGT GTC TAA CAT AG -3' Reverse: 5'- CTA TGT TAG ACA CAC TGG ATG CAC TGG TCA TGC CTT TGA TC-3'</td>
</tr>
<tr>
<td>WLN5-6 (C575S,C578S)</td>
<td>38</td>
<td>61</td>
<td>72</td>
<td>Forward: 5'- CAG GGA TGA CCA GCC CGT CCA GCG TCC ACA ACA TAG AG -3' Reverse: 5'- CTC TAT GTT GTG GAC GCT GGA CGC GCT GGT CAT CCC TG -3'</td>
</tr>
<tr>
<td>WLN5-6 (C490A)</td>
<td>39</td>
<td>54</td>
<td>69.7</td>
<td>Forward: 5'- GCA CCG CAG AAG GCC TTC TTA GAT CAC AAA AGC ATG ACC -3' Reverse: 5'- GGT CAT GCC TTT GAT CTG TAA GAA GGC CCT TCT CGG TGC -3'</td>
</tr>
<tr>
<td>WLN5-6 (E535C)</td>
<td>39</td>
<td>56</td>
<td>70.8</td>
<td>Forward: 5'- GAG ATC AAG TAT GAC CCA TGC GTC ATC CAG CCC CTC GAG -3' Reverse: 5'- CTC GAG GGG CTG GAT GAC GCA TGG GTC ATA CTT GAT CTC -3'</td>
</tr>
<tr>
<td>GFP (Forward)</td>
<td>50</td>
<td>54</td>
<td>73.6</td>
<td>5'- GGT ATT GAG GGT CGC GAG AAC CTT TAT TTC CAG GGC ATG GTG AGC AAG GG -3'</td>
</tr>
<tr>
<td>GFP (Reverse)</td>
<td>49</td>
<td>53</td>
<td>72.9</td>
<td>5'- AGA GGA GAG TTA GAG CCT CAC TTG TAC AGC TCA TCC ATG CCG TGG GTG A -3'</td>
</tr>
</tbody>
</table>

* Residue number in ATP7B  
*# Residue number in WLN5-6
### Table 2.4: Notations of proteins

<table>
<thead>
<tr>
<th>Notation</th>
<th>Residues in ATP7B</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN3-6</td>
<td>256-633</td>
<td>N-terminal metal binding domains three through six</td>
</tr>
<tr>
<td>TRX_TEV_WLN3-6</td>
<td>256-633</td>
<td>N-terminal metal binding domains three through six with fusion thioredoxin tag and TEV protease digestion site</td>
</tr>
<tr>
<td>WLN4-6</td>
<td>355-633</td>
<td>N-terminal metal binding domains four through six</td>
</tr>
<tr>
<td>TRX_TEV_WLN4-6</td>
<td>355-633</td>
<td>N-terminal metal binding domains four through six with fusion thioredoxin tag and TEV protease digestion site</td>
</tr>
<tr>
<td>WLN5-6</td>
<td>485-633</td>
<td>N-terminal metal binding domains five through six</td>
</tr>
<tr>
<td>Y532H</td>
<td>485-633</td>
<td>Y532H mutation in WLN5-6</td>
</tr>
<tr>
<td>V536A</td>
<td>485-633</td>
<td>V536A mutation in WLN5-6</td>
</tr>
<tr>
<td>L492S</td>
<td>485-633</td>
<td>L492S mutation in WLN5-6</td>
</tr>
<tr>
<td>WLN5-6C51</td>
<td>485-633</td>
<td>Six mutations in WLN5-6 C490A, C499S, C502S, E535C, C575S, C578S.</td>
</tr>
<tr>
<td>WLN5-6C51_CPM</td>
<td>485-633</td>
<td>WLN5-6Cys51 bound to CPM at C535</td>
</tr>
<tr>
<td>WLN5-6C6</td>
<td>485-633</td>
<td>Four mutations in WLN5-6, C499S, C502S, C575S, C578S</td>
</tr>
<tr>
<td>WLN5-6C6_CPM</td>
<td>485-633</td>
<td>WLN5-6Cys6 bound to CPM in position</td>
</tr>
<tr>
<td>GFP</td>
<td></td>
<td>Green Fluorescent Protein</td>
</tr>
<tr>
<td>TEV</td>
<td></td>
<td>Tobacco etch Virus Protease</td>
</tr>
</tbody>
</table>

### Table 2.5: Plasmids and their vectors properties

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>Vector</th>
<th>Size of plasmid</th>
<th>Size of inserted gene</th>
<th>Tag</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>pWLN3-6</td>
<td>pET 32 Xa/ LIC</td>
<td>5926</td>
<td>1134 + 21 for TEV site</td>
<td>Thioredoxin</td>
<td>Ampicillin/carbenicillin</td>
</tr>
<tr>
<td>pWLN4-6</td>
<td>pET 32 Xa/ LIC</td>
<td>5926</td>
<td>831 + 21 for TEV site</td>
<td>Thioredoxin</td>
<td>Ampicillin/carbenicillin</td>
</tr>
<tr>
<td>pWLN5-6</td>
<td>pET 24 d+</td>
<td>5307</td>
<td>453</td>
<td>No tag</td>
<td>Kanamycin</td>
</tr>
<tr>
<td>pGFP</td>
<td>pET 32 Xa/ LIC</td>
<td>5926</td>
<td>714 + 21 for TEV site</td>
<td>Thioredoxin</td>
<td>Ampicillin/carbenicillin</td>
</tr>
</tbody>
</table>
2.3 Instrumentation

Instruments which used in the research, manufacturers, and their applications are listed in Table 2.6. Columns for FPLC are listed in Table 2.7.

Table 2.6: Instrumentations and their manufacturer and applications

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Model #</th>
<th>Manufacturer</th>
<th>Application</th>
</tr>
</thead>
<tbody>
<tr>
<td>AKTA FPLC</td>
<td>18-1900-26</td>
<td>GE Lifescience</td>
<td>Protein purification</td>
</tr>
<tr>
<td>UV/VIS spectrophotometer</td>
<td>DU-7400</td>
<td>Beckman</td>
<td>Nucleic acid and protein analyzer</td>
</tr>
<tr>
<td>Jasco CD spectropolarimeter</td>
<td>J815</td>
<td>Jasco</td>
<td>CD measurements</td>
</tr>
<tr>
<td>Sorval centrifuge</td>
<td>RC 5B plus</td>
<td>Thermofisher scientific</td>
<td>Concentrating protein and cell culture spinning</td>
</tr>
<tr>
<td>Minicycler</td>
<td>PTC 150</td>
<td>MJ research</td>
<td>PCR</td>
</tr>
<tr>
<td>Controlled environment incubator shaker</td>
<td>New Brunswick scientific</td>
<td>Growing cultures</td>
<td></td>
</tr>
<tr>
<td>Minispin plus centrifuge</td>
<td>5415</td>
<td>Eppendorf</td>
<td>Plasmid purification</td>
</tr>
<tr>
<td>Fluorimeter</td>
<td>FL 900</td>
<td>Edinburgh</td>
<td>Two photon fluorescence and lifetime measurements</td>
</tr>
<tr>
<td>UV-VIS scanning spectrophotometer</td>
<td>UV2101PC</td>
<td>Shimadzu</td>
<td>UV scan for proteins</td>
</tr>
<tr>
<td>Fluorescence spectrophotometer</td>
<td>F-2500</td>
<td>Hitachi</td>
<td>Fluorescence scan</td>
</tr>
<tr>
<td>Glove box</td>
<td></td>
<td>Vacuum atmosphere company</td>
<td>Anaerobic environment</td>
</tr>
<tr>
<td>Fluorescence up-conversion system</td>
<td>FOG 100</td>
<td>CDP corp</td>
<td>Up-conversion measurements</td>
</tr>
<tr>
<td>Phastsystem separation and control unit</td>
<td>280578</td>
<td>Pharmacia</td>
<td>Isoelectric focusing</td>
</tr>
<tr>
<td>pH benchtop meter</td>
<td>Orion 3 star pH benchtop</td>
<td>Thermo electron corporation</td>
<td>Adjusting pH of buffers</td>
</tr>
<tr>
<td>Analytical balance</td>
<td>PG603-SDR</td>
<td>Mettler Toledo</td>
<td>Weighing &gt; 0.01 g</td>
</tr>
<tr>
<td>Analytical balance</td>
<td>AT400</td>
<td>Mettler Toledo</td>
<td>Weighing &gt; 0.1 mg</td>
</tr>
<tr>
<td>DynaPro dynamic light scattering</td>
<td></td>
<td>Wyatt Technology Corporation</td>
<td>DLS of proteins</td>
</tr>
<tr>
<td>Thermal imaging system</td>
<td>FTI-500</td>
<td>Pharmacia Biotech</td>
<td>View DNA in Agarose gel</td>
</tr>
</tbody>
</table>
Table 2.7: Columns used with AKTA FPLC to purify proteins

<table>
<thead>
<tr>
<th>Column</th>
<th>Function</th>
<th>Resin</th>
<th>Details</th>
<th>Binding buffer</th>
<th>Eluting buffer</th>
<th>Protein purified/usage</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEAE sepharose</td>
<td>Anion exchange</td>
<td>Diethylaminoethyl sepharose</td>
<td>FF XK26</td>
<td>20 mM MES, 10 mM DTT, 1 mM EDTA, pH=6.0-6.3</td>
<td>20 mM MES, 10 mM DTT, 1 mM EDTA, 1M NaCl, pH= 6.0-6.3</td>
<td>WLN5-6 Y532H WLN5-6Cys51 WLN5-6Cys6</td>
</tr>
<tr>
<td>HisPrep</td>
<td>His-tag affinity</td>
<td>Sepharose</td>
<td>FF16/10</td>
<td>50 mM HEPES, 30 mM Imidazole, 500 mM NaCl, pH =7.5</td>
<td>50 mM HEPES, 300-500 mM Imidazole, 500 mM NaCl, pH =7.5</td>
<td>TEV WLN3-6 WLN4-6 GFP</td>
</tr>
<tr>
<td>Desalting column</td>
<td>Buffer exchange</td>
<td>Superdex G-25 superfine</td>
<td>HiPrep 26/60</td>
<td>Desired buffer</td>
<td></td>
<td>All investigated proteins</td>
</tr>
<tr>
<td>Superdex 75 pg</td>
<td>Gel Filtration</td>
<td>Superdex 75 prep grade</td>
<td>HiLoad 26/60</td>
<td>50 mM HEPES, 200 mM NaCl, 5-10 mM DTT, pH = 7.5</td>
<td></td>
<td>All investigated proteins</td>
</tr>
<tr>
<td>Superdex 200 HR</td>
<td>High resolution gel filtration</td>
<td>Crosslinked agarose and dextrane</td>
<td>10/30</td>
<td>50 mM phosphate buffer, 150 mM NaCl, pH = 7.0</td>
<td></td>
<td>All investigated proteins</td>
</tr>
</tbody>
</table>

The solutions and buffers that needed were prepared as recommended and their compositions are listed in Table 2.8.
### Table 2.8: Solutions and buffers

<table>
<thead>
<tr>
<th>Solution</th>
<th>Composition</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA sample buffer (6X)</td>
<td>0.25% Bromophenol blue, 0.25% xylene cyanol FF, 30% glycerol, distilled water</td>
</tr>
<tr>
<td>LB broth</td>
<td>10 g Bacto-tryptone, 10 g NaCl, 5 g Bacto-yeast extract, 1 L milliQ water.</td>
</tr>
<tr>
<td>TBE buffer (5X)</td>
<td>54 g Tris base (FW=121 g/mol), 27.5 g Boric acid (FW=61.83 g/mol), 20 ml EDTA (0.5 M, pH=8), up to 1 L milliQ water, pH=8.0.</td>
</tr>
<tr>
<td>SDS sample buffer (6X)</td>
<td>7 ml Tris (0.5M, pH=6.8), 3 ml Glycerol, 1 g SDS, 0.93 g DTT, 1.2 mg Bromophenol Blue.</td>
</tr>
<tr>
<td>SDS cathodic buffer (5X)</td>
<td>0.5 M Tris base, 0.5M Tricine, 0.5 % SDS, pH = 8.25</td>
</tr>
<tr>
<td>SDS anodic buffer (5X)</td>
<td>1M Tris, pH= 8.9</td>
</tr>
<tr>
<td>SDS staining buffer</td>
<td>0.2% Coomassie blue R-250 dye, 50% Methanol, 7% Glacial acetic acid, 43 % distilled water.</td>
</tr>
<tr>
<td>Gel drying solution</td>
<td>350 ml Ethanol, 40 ml, Ethylene glycol, 50 ml Glycerol, 600 ml distilled water.</td>
</tr>
<tr>
<td>Agarose gel</td>
<td>0.8 % Agarose, 40 ml TBE buffer (1X).</td>
</tr>
<tr>
<td>SDS separating gel (12.5%)</td>
<td>8.3 ml Acrylamide/Bis-acrylamide (30 % / 0.8%), 5 ml Tris (1.5M, pH8.8), 0.2 ml Ammonium persulfate (10 %), 0.2 ml SDS (10 %), 0.008 ml TEMED, 6.26 ml distilled water</td>
</tr>
<tr>
<td>SDS stacking gel (5%)</td>
<td>1.7 ml Acrylamide/Bis-acrylamide (30 % / 0.8%), 1.25 ml Tris (1 M, pH6.8), 0.1 ml Ammonium persulfate (10 %), 0.1 ml SDS (10 %), 0.01 ml TEMED, 6.8 ml distilled water</td>
</tr>
<tr>
<td>M9 salts (5X)</td>
<td>15 g KH₂PO₄, 42.5 g Na₂HPO₄.2H₂O, 2.5 g NaCl, up to 1 L milliQ water, pH =7.3-7.5, solution filtered through 0.22 μm then autoclaved.</td>
</tr>
<tr>
<td>Minimal media</td>
<td>1 ml MgSO₄ (1M), 0.3 ml CaCl₂ (1M), 1 ml MEM vitamin mix, 5 ml ¹⁵NH₄Cl (0.2g/ml), 10 ml glucose (0.2g/ml), 200 ml M9 salts (5X), diluted to 1 L and autoclaved</td>
</tr>
</tbody>
</table>
| IEF staining solution         | A. 0.2 % (w/v) CuSO₄ in 20 % glacial acetic acid  
B. 60% methanol  
C. 0.4 g Coomassie blue R-250, 400 ml H2O, and 600 ml methanol |
| IEF drying solution           | 5 % glycerol, and 10 % glacial acetic acid                                                            |
2.4 Procedures

2.4.1 Construction of WLN3-6 plasmid

WLN3-6 gene was amplified using pET32Xa/WLN1-6 plasmid as template, which was prepared by Dr. Joshua Muia. Amplification was performed by PCR reaction using KOD hot start DNA polymerase (Novagen) protocol, Table 2.9. Mutagenic primers were designed so that the forward primer encodes Tobacco Etch Virus (TEV) protease recognition site, Table 2.1. Both forward and reverse primers include a sequence compatible with LIC cloning into pET32Xa/LIC vector. “Xa” designates the Factor Xa cleavage site. LIC notation refers to the Ligation Independent Cloning, a method that facilitates the directional cloning of PCR products without the use of restriction enzymes, digestion, or ligation reactions.

Table 2.9: KOD hot Start DNA polymerase protocol

<table>
<thead>
<tr>
<th>Sample preparation</th>
<th>V (µL)</th>
<th>PCR parameters</th>
<th>duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reverse primer (5µM)</td>
<td>1.78</td>
<td>1- Activation (95°C)</td>
<td>3 min</td>
</tr>
<tr>
<td>Forward primer (5µM)</td>
<td>2.56</td>
<td>2- Denaturation (94°C)</td>
<td>20 sec</td>
</tr>
<tr>
<td>dNTP (2 mM)</td>
<td>5.0</td>
<td>3- Annealing (53°C)</td>
<td>30 sec</td>
</tr>
<tr>
<td>MgSO₄ (25 mM)</td>
<td>3</td>
<td>4- Extension (72°C)</td>
<td>1 min 40 sec</td>
</tr>
<tr>
<td>10X Buffer for KOD</td>
<td>5</td>
<td>Repeat steps 2-4</td>
<td>30 cycles</td>
</tr>
<tr>
<td>Template (WLN1-6)</td>
<td>1</td>
<td>Store (4°C)</td>
<td>infinity</td>
</tr>
<tr>
<td>KOD hot start DNA polyrase</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autoclaved water</td>
<td>35.66</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The PCR product corresponding to WLN3-6 gene expected to have 1130 bp. 0.8 % Agarose gel was used to confirm the results of the PCR. The results showed other products beside the one that is expected for WLN3-6, Figure 2.1. Therefore, the PCR
reaction performed again and the gel exposed to UV light at lower energy (365 nm) then the band corresponding to WLN3-6 gene was excised. The excised pieces of gel were purified using a gel extraction kit from Qiagen, then quantified using Agarose gel, Figure 2.2.

![Figure 2.1: 0.8 % Agarose gel showing PCR product of WLN3-6 amplification](image)

0.2 pmol of the pure insert was treated with T4 DNA polymerase as described in Table 2.10. The solution then was mixed and incubated at 22°C for 30 minutes. Then the enzyme deactivated by incubation at 75°C for 20 minutes. 2 µL of T4 DNA polymerase treated PCR insert was then ligated to a microliter of Xa/LIC vector at 22°C for 5 minutes before a one microliter of 25 mM MgSO₄ added and incubated for 5 minutes at 22°C. After ligation, some of the ligation product was transformed into GigaSingles NovaBlue cells. Positive growth was observed in the plate after leaving it in an oven overnight at 37°C. A single healthy colony was inoculated in 5 mL LB media containing 100 µg/mL of ampicillin. The resulted cells were centrifuged, and the plasmid was extracted and
purified using Qiagen plasmid purification kit. The plasmid was confirmed by double 
digestion using BamHI and BglIII enzymes, the pieces of the digestion are shown in 
Figure 2.3. Then a sample was sent for sequencing in Retrogen, Inc. at San Diego, CA. 
S-tag primer in addition to T7 and T7 term primers were used for sequencing. The results 
show a successful sequence that aligned perfectly to the WLN3-6 gene stored in DNA strider ® program at our laboratory.

Table 2.10: T4 DNA polymerase protocol

<table>
<thead>
<tr>
<th>Solution</th>
<th>V (µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10 X T₄ DNA polymerase buffer</td>
<td>2</td>
</tr>
<tr>
<td>25 mM dGTP</td>
<td>2</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>1</td>
</tr>
<tr>
<td>Nuclease free water</td>
<td>13.6</td>
</tr>
<tr>
<td>2.5 U/µL DNA polymerase</td>
<td>0.4</td>
</tr>
<tr>
<td>0.2 pmol of WLN3-6 insert</td>
<td>1</td>
</tr>
</tbody>
</table>

![Figure 2.2: Purification of PCR product of WLN3-6 gene amplification: (A) gel photo after excised WLN3-6 gene. (B) gel after purification of the gene](image)
2.4.2 Transformation, expression and purification of WLN3-6

WLN3–6 protein was expressed by transforming pET32TEV/WLN3–6 plasmid into Rosetta (DE3) competent cells. A single colony was inoculated in an LB culture contain ampicillin (100 µg/ml). The culture was induced with a final concentration of 1mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) when the optical density at 600 nm reached 0.6-0.8. Protein induction was performed for 3-4 h and cells were harvested by centrifugation.

The TRX_{TEV}WLN3–6 fusion protein was extracted by freezing and thawing the cell pellet three times followed by the addition of BugBuster® (Novagen) protein extraction reagent. The high viscosity due to DNA and RNA release was lowered by adding 20 KU/mL of Benzonase (Novagen). The supernatant containing soluble proteins was separated from cell debris by centrifugation. The supernatant was decanted and
loaded onto the HisPrep FF 16/10 (nickel chelating sepharose column) with high affinity for poly-histidine tagged protein. The expressed fusion protein, TRX_{TEV}WLN3–6 possessed the six poly-histidine tag (-HHHHHH-) that was utilized in the purification. The tagged bound protein was eluted by an imidazole gradient from the column. The protein was further purified by gel filtration, and then prepared for fusion protein cleavage. Fusion protein has mass of 58 kDa, lane E in Figure 2.4.

To remove the TRX and His-tags from WLN3–6, the purified protein was exchanged into TEV buffer and quantified by the BCA protein assay (Pierce). The cleavage mixture has the ratio of 1mg of TEV protease per 80 mg of the protein to be cleaved. The mixture was incubated at room temperature for 16 h. The cleaved protein was loaded onto the HisPrep column and the flowthrough fractions that contained WLN3–6 protein were collected and purified by gel filtration using Superdex 75. The purity was determined by SDS-PAGE, Figure 2.4. WLN3-6 protein appears in final purification step as band around 40 kDa. The yield quantified by BCA protein assay was almost 20 mg/ L culture.
Figure 2.4: 12.5 % SDS-PAGE gel shows steps of WLN 3-6 expression and purification


2.4.3 DNA plasmid, Cloning and Purification of WLN4-6

WLN4-6 gene was amplified by PCR reaction using KOD hot start DNA polymerase (Novagen) protocol. Mutagenic primers were designed so that the forward primer encode Tobacco Etch Virus (TEV) protease recognition site. Both forward and reverse primer including sequence compatible with LIC cloning into pET32Xa/LIC, Table 2.3. The template encoding the N-terminal portion of ATP7B was prepared by Dr. Joshua Muia\textsuperscript{1}. The PCR product was produced using annealing temperature of 58°C then purified after it was excised from the gel. 0.2 pmol of the pure PCR product was then treated with T4 DNA polymerase before it was ligated to pET32 Xa/LIC vector as was performed for WLN3-6. The protein was expressed and purified by the same methods used for WLN3-6. Results of: PCR product after purification, plasmid digestion, SDS-
PAGE for protein purification steps, and pure protein are shown in Figure 2.5, Figure 2.6, Figure 2.7, and Figure 2.8, respectively.

Figure 2.5: 0.8 % Agarose gel shows purified PCR product of amplifying WLN4-6 gene

Figure 2.6: Digestion of WLN4-6 plasmid
Figure 2.7: 12.5 % SDS-PAGE showing purification steps of WLN4-6

A, non-induced. B, after 4 hours induction. C, freeze/thaw extract. D, TRX\textsubscript{TEV}WLN4-6 after HisTrap column. E, TRX\textsubscript{TEV}WLN4-6 after gel filtration. F, digestion of protein using TEV protease.

Figure 2.8: SDS-PAGE gel shows pure WLN4-6 protein
2.4.4 DNA plasmid, Cloning and Purification of WLN5-6

The plasmid of WLN5-6 was constructed by Dr. Muia\(^1\). It was ligated into pET24d\(^+\) vector, which contains a gene resistance to the antibiotic Kanamycin. 10-50 ng of the plasmid was transformed into 100 µL of NovaBlue cells, incubated in ice for 20 min then heat shocked at 42 °C for 30 seconds then immediately incubated in ice for 2 minutes. 1000 µL of LB broth was added to the cell and incubated at 37 °C with shaking at a rate of 225 rpm. An hour later, the cells were centrifuged at 7000 rpm for 2 minutes. The pellet were then gently resuspended in about 250 µL of the LB and plated in LB/agar plate containing 30 µg/mL of Kanamycin. The plate was incubated at 37 °C overnight. One colony of the plate was inoculated into 5 mL autoclaved LB media. Kanamycin was added to a total concentration of 30 µg/mL then incubated at 37 °C overnight in a shaker at 225 rpm. Overnight, the cells were grown and became dense so they were centrifuged and resuspended using 250 µL of P1 buffer from QIAprep spin miniprep plasmid purification kit, 250 µL of P2 buffer was added and the tube was flipped several times then 350 µL of N3 buffer was added. The tube was flipped 4 to 5 times and centrifuged at 13000 rpm for 10 minutes. The supernatant was then applied to the spin column and centrifuged for one minute. The column was then washed by adding 500 µL buffer PB and centrifuged for another minute. Another wash for the column was performed by adding 750 µL PE buffer and centrifuging for 1 minute. The flow-through was discarded then the column was centrifuged for another minute. To elute the plasmid, the spin column was placed in a clean 1.5 mL microcentrifuge tube and 50 µL of EB buffer were added to the center of the spin column let stand for two minutes then centrifuged for a minute. The plasmid then stored at −20 °C.
2.4.5 Transformation, expression and purification of WLN5-6

20 ng of WLN5-6 plasmid was transformed into 5 µL of Rosetta2(DE3) cells. The cells were incubated on ice for 15 minutes prior to heat them at 42 °C for 25 seconds. The cells then were incubated again in ice for 2 minutes. 250 µL of LB was added and the tube contains the cells and LB was incubated at 37 °C for an hour with shaking at a rate of 225 rpm. An LB/Agar plates containing 30 µg/ml kanamycin was used to plate the cells. The plate was then incubated at 37 °C overnight.

One fresh colony from overnight plate was inoculated for each 5 mL LB media. I usually use 4 inoculations to make sure that I will have growth in at least one culture tube. Kanamycin antibiotic were added to each LB media used for the growth in a concentration of 30 µg/mL. The tubes were incubated at 37 °C in the shaker for 6 hours then each tube transferred to a flask contains 100 mL of LB and same as before incubated at the same temperature for 2.5 hours. The LB in each flask then transferred to another one liter LB and incubated at the same temperature until the OD at 600 nm reaches between 0.6-0.8. At that value of OD, an aqueous solution of 1M of IPTG was used to add IPTG to a final concentration of 1 mM in each flask. Three hours later the cells were harvested by centrifugation at 5000 rpm at 4 °C to obtain the pellet. The pellet was then stored at – 20 °C for later purification.

WLN5-6 protein was expected to have an isoelectric point of 4.9 as calculated by DNA Strider™ 1.3f16. Therefore at pH 6.0 it will have a net negative charge for the side chains and so ion exchange chromatography can be used to purify the protein. DEAE column is an anion exchange resin and so it was used to purify the protein. For further purification of the protein gel filtration chromatography was also used.
The cells from the frozen pellet were lysed by repeated freeze and thaw cycles. The cells were frozen for one minute in liquid nitrogen and thawed under cold tap water while keeping them in a centrifuge bottle. After 5 cycles, the pellet were resuspended with the pipet pump and extraction buffer (20 mM MES/Na, 1 mM DTT, 0.1 mM EDTA, pH=6.0), 5 ml buffer were added per 1 gram pellet. 2 uL of Benzonase nuclease (2.5 KU) was also added and then the solution agitated at 4 °C for 30 minutes. Insoluble cell debris was removed from soluble protein in the supernatant by centrifugation at 15000xg for 20 min at 4 °C. The supernatant containing protein was then recovered for protein purification.

Purification of WLN5-6 was performed by using DEAE-Sepharose anion exchange column. The column was washed with elution buffer (20 mm MES/Na, 0.1 mM EDTA, 5mM EDTA, 1 M NaCl, pH =6.3) and pre-equilibrated by low salt washing buffer (20 mm MES/Na, 0.1 mM EDTA, 5mM EDTA, pH =6.3). The supernatant, 20 ml, was then loaded into the column at a flow rate of 2 ml/min. The unbound proteins were washed out of the column by passing the washing buffer through the column until the UV absorbance returned to baseline. Then the protein eluted with the elution buffer. The fraction containing WLN5-6 were identified by SDS-PAGE (15 %). The fractions were collected and concentrated by Amicon device with a 3,000 MWCO membrane. WLN5-6 was further purified by gel filtration. The concentrated sample (< 10 ml) was injected in equilibrated Hiload 26/60 Superdex 75 column at a flow rate of 2.5 ml/min. The protein was then eluted using HEPES buffer (50 mM HEPES, 200 mM NaCl, 10 mM EDTA, pH = 7.5). SDS-PAGE confirmed the purity and BCA assay used to measure protein concentration. Approximately, 30 mg of protein was produced per liter LB culture.
2.4.6 Tobacco etch virus (TEV) protease expression and purification

Tobacco etch virus (TEV) protease is the common name for the 27 kDa catalytic domain of the nuclear inclusion a (NiA) protein encoded by the tobacco etch virus. TEV protease is a site-specific protease with a canonical recognition site, Glu-Asn-Leu-Tyr-Phe-Gln-Gly. Cleavage occurs between the Gln and Gly, resulting in a Gly at the amino terminus of the protein of interest. TEV protease is an ideal choice for removal of tags from fusion proteins in the research and production of recombinant protein.

TEV protease that used to remove the tag from WLN3-6 and WLN4-6 was produced in our laboratory from a vector, kindly, supplied by Dr. Tim Cross of the National High Magnetic Field Laboratory. The protein plasmid was transformed to Rosetta 2 cells and the cells grown in LB media at 37 °C until the OD$_{600}$ reached 0.6-0.8, then IPTG was added to a final concentration of 1 mM. Four hours later, the cells were harvested at 4 °C by centrifugation. The pellet were thawed then re-suspended in 35 ml of 50 mM sodium phosphate buffer, 0.3 M NaCl, pH=8.0 and lysed by sonication at 12 pulse for 5 minutes.

The supernatant was collected by centrifugation at 15000 rpm for 20 min. the resulting supernatant was loaded onto Ni$^{2+}$ affinity column equilibrated with equilibrium buffer (50 mM phosphate buffer, 0.3 M NaCl, pH=8.0, 10 mM Imidazole). The column was then washed with the washing buffer (50 mM phosphate buffer, 0.3 M NaCl, pH=8.0, 10 mM Imidazole) and then eluted with the elution buffer (50 mM phosphate buffer, 0.3 M NaCl, pH=8.0, 250 mM Imidazole). SDS-PAGE was used to confirm the purity of the TEV, Figure 2.9. The eluted fraction was concentrated prior to load into Desalting column and then eluted with 50 mM Tris-HCl, pH 8.0, 0.5 mM EDTA. The
eluted fractions were concentrated using Amicon device with 3000 Da cut-off membrane and then BCA assay was performed, and then glycerol and DTT were added to a final concentration of 50 % and 5 mM, respectively. Protein then was stored for later use at -80°C in 500 μl increments.

Figure 2.9: TEV purification steps. A, cell extract. B, flowthrough. C, eluted protein

2.4.7 \(^1\text{H},^{15}\text{N}\) NMR HSQC for WL.N3-6, WL.N5-6

NMR \(^1\text{H},^{15}\text{N}\)-HSQC experiments of \(^{15}\text{N}\)-NMR labeled protein is a good spectrum to start with to test whether a protein is folded and assess the quality of the spectrum to
see whether it is worth recording other spectra and possibly proceeding on to other, more expensive labeling schemes.

The plasmid for proteins in interest was transformed into Rosetta 2 (DE3) cells. The following day fresh colonies were inoculated into 5 ml LB then incubated at 37°C for 10 hours. 10 µL of the media were transferred into 30 ml of minimal media containing the proper antibiotic. The media were incubated at 37°C with shaking (250 rpm) overnight. The next day the minimal media were transferred to 470 ml minimal media that also contain antibiotic. When OD_{600} reached 0.6-0.8 the cells were induced by adding IPTG to a final concentration of 1 mM and the temperature reduced to room temperature. 22 hours later, the cells harvested by centrifugation at 4 °C and 5000 rpm. The pellet was formed then saved at -20 °C for later purification. The purification of each protein was conducted as described earlier for WLN3-6 and WLN5-6 in section 2.4.2 and section 2.4.4, respectively.

Around 600 µL of 0.5 mM of protein required for obtaining $^1$H,$^{15}$N-HSQC. The proteins have been analyzed in 800 MHz NMR instrument at University of Notre Dame in Indiana. The results are shown in Figure 2.10. Spectra for all the proteins show that it is possible to proceed to more experiments.
Figure 2.10: $^1$H, $^{15}$N-HSQC results for WLN5-6 and WLN3-6
2.5 Biophysical properties of proteins

2.5.1 Dynamic light scattering

When light impinges on matter, the electric field of the light induces an oscillating polarization in the molecules. Consequently, the molecule gives a secondary source of light and scatters the light. If the source of light is monochromatic and coherent, as in laser light, then a time-dependent fluctuation in the scattered intensity is observed. This fluctuation is because of the Brownian motion, which makes the distance between the scattered molecules in the solution is continuously changing with time. Brownian motion is the random movement of particles due to collisions caused by the bombardment of the solvent molecule that surrounded the particles. As a result, a Doppler shift between the frequency of the incident light and the frequency of the scattered light is observed; this change is related to the size of the particles\(^3\). Then electrodynamics and theory of time-dependent statistical mechanics help to obtain information through the characteristics of the scattered light about molecular dynamics and the structure of the scattering molecules.

Dynamic light scattering (DLS), also called photon correlation spectroscopy, is a non-invasive technique to measure the speed of particles and molecules undergoing Brownian motion. This motion is influenced by the particle size, sample viscosity, and the temperature. Small particles at high temperature have rapid Brownian motion\(^3\). Velocity of the Brownian motion is defined by the translational diffusion coefficient (D) which, in turn, can be converted to particle size through Stokes-Einstein equation, equation (2.1).
\[
d_h = \frac{k_B T}{6 \pi \eta D}
\]  \hspace{1cm} (2.1)

Where, \(d_h\) is the hydrodynamic radius, which defined as the diameter of a hard sphere that has the same average of translational diffusion coefficient as the particle being measured, \(k_B\) is the Boltzmann’s constant, \(\eta\) is the viscosity, and \(T\) is the temperature in Kelvin\(^4\).

The translational diffusion coefficient (\(D\)) depends not only on the size of the particle, but also on any surface structure, in addition to the concentration and kind of ions in the medium. Ions and their concentrations affect the thickness of the electric double layer (Debye length, \(K^{-1}\)), which then changes the particle diffusion speed and the apparent particle’s size. Moreover, the conformation of the polymer can be affected due to the factors mentioned above, which can modify the apparent size by quite a few nanometers. The amount of scattered light is related to the particle size. The intensity of the scattering is related to the sixth power of the particle diameter. There are two terminologies used for the studying of scattering light: static light scattering and dynamic light scattering. In static, the experiment probes the time-average intensity of the scattered light, while in dynamic light scattering, the fluctuations in light intensity are probed.

2.5.1.1 Description of how DLS works

If an immobile cuvette containing sample is illuminated by a laser with a frosted glass screen to view the sample cell, a classical speckle pattern will be seen. This speckle has dark and bright blobs due to destructive and constructive phases of the scattered light, Figure 2.11. Because the particles in the system undergo Brownian motion, the speckle pattern is observed where the spot of each speckle is seen to be in continuous motion.
This is because the phase addition from the moving particles is constantly evolving and forming new patterns. The rate at which these intensity fluctuations occur will be rapid for small particles and slower for large particles. Then a device called digital auto correlator measures the degree of similarity between two signals, or one signal with itself at varying time intervals. If the signal intensity is compared with itself at certain time, then there is perfect correlation, as the signals are identical. The degree of correlation ranges between (1.00) and zero where, perfect correlation is designated by unity (1.00) and no correlation is designated by zero (0.00). However, when the signal is derived from Brownian motion, which is a random process, the correlation reduces with time. The correlation ranges from nanoseconds or microseconds. The correlator constructs a correlation function $G(\tau)$ of the scattered intensity, equation (2.2):

$$G(\tau) = <I(t).I(t+\tau)>$$  \hspace{1cm} (2.2)

Where $\tau$ is the sample time difference of the correlator.

For monodisperse particles in Brownian motion, the correlation function is an exponential decaying function of $\tau$, equation (2.3):

$$G(\tau) = A[1 + B \exp(-2\Gamma\tau)]$$  \hspace{1cm} (2.3)

Where A is the baseline of the correlation function, and B is the intercept of the correlation function.

$\Gamma = D q^2$, where $D$ is the translational diffusion coefficient, $q$ is the wave vector

$q = (4 \pi n / \lambda_o) \sin (\theta/2)$, where $n$ is the refractive index of dispersant, $\lambda_o$ is the wavelength of the laser, $\theta$ is the scattering angle.

For polydisperse samples, the equation is written as:

$$G(\tau) = A(1 + B g_1(\tau)^2)$$

where $g_1(\tau)$ is the sum of all the exponential decays contained in the correlation function.
Figure 2.11: Schematic representation of speckle pattern. Applications of DLS to proteins

Because scattering intensity is proportional to the sixth power of the molecular radius, DLS became a valuable and sensitive technique to trace protein aggregates. The presence of a small amount of aggregate will cause significant change in the mean hydrodynamic radius and in the percentage polydispersity. DLS can also be valid in thermal stability studies by the measurement of the size and the scattering intensity as a function of temperature. The heat influences the denaturation of protein and leads to enormous aggregates on top of increases in the scattering intensity.

If DLS is measured in high concentrated solutions, then the radiation may scatter many times, as the scattered centers are grouped together. This effect is called multiple scattering; it is deterministic more than a random phenomenon, therefore, the randomness of the interaction tends to be averaged out by large number of scattering events. Concentrated protein, which is charged, could increase the repulsion forces between molecules due to the reduction in the intermolecular distance. That, in turn, increases the diffusion speed, and so, decreases the apparent size. This leads to the importance of the
dynamic Debye plot. This plot is actually a measurement of diffusion coefficient at several concentrations of the sample in order to determine the diffusion coefficient at infinite dilution. Also, this plot provides information about the intermolecular interaction in a given solution. Repulsive interactions cause faster motion while attractive interactions induce slower motion.

2.5.2 Fluorescence

Fluorescence is a radiative decay from the singlet state of molecules known as chromophores. The process require three steps; excitation, excited-state lifetime, and fluorescence emission, illustrated in Figure 2.12. In excitation, the molecule absorbed a photon of energy \( (h\nu_{Ex}) \) to produce an electronic singlet excited state \( S_1 \). The excited state last for nanoseconds, typically 1-10 ns, during this time, the fluorophore interacts with its molecular environment or/and undergoes conformational change. As a result, the excited state loses some of its energy, therefore, became relaxed to a lower singlet excited-energy state. The fluorescence emission originates from the lower singlet excited state of \( S_1 \) and a photon of \( (h\nu_{Em}) \) is emitted making the fluorophore return back to its original ground state\(^5\). Under the same conditions, the fluorescence emission spectrum of fluorophore will have the same profile no matter what the excitation wavelengths were; only the intensity will change. This independency on the excited state is due to the dissipation of energy during the excited-state lifetime. The fluorescence technique is considered very sensitive because it is continually repeats the fluorescence process to generate thousands of detectable photons, unless the fluorophore is bleached.
Fluorescence measurements are broadly classified into two types: steady-state measurements and time-resolved measurements. The most common type is the steady-state measurements where a continuous beam of light is used to illuminate the sample and the intensity of the emission spectrum is observed. The second type of measurements, time-resolved, is used to measure the intensity decays or anisotropy decays. In these measurements the sample is exposed to a pulse of light which whose width is typically shorter than the decay time of the sample. This type requires a high-speed detection system to measure the intensity or anisotropy in the nanosecond timescale. Steady-state and time-resolved measurements are related, since steady-state observation is the average of the time-resolved phenomena over the complete intensity decay of the sample. This information is lost during the averaging process and so it is important to have time-resolved measurements.
Monitoring of time-resolved emission can be performed using time-correlated single-photon counting (TCSPC)\textsuperscript{6,7}. In this technique, the sample is excited with pulses from a laser or a flash lamp, then the detection system monitors the time difference between the excitation pulse and the first fluorescence photon from the sample. This measurement requires that only one photon is observed for a large number of excitation pulses. In this case, an extremely low count rate must be ensured such that the system operates in single photon counting mode. Under this condition, statistics mimic the poisson distribution and a true time-resolved emission profile is obtained. A schematic diagram of a TCSPC experiment can be represented in Figure 2.13.
An excitation pulse (optical pulse) is split into two fractions, one part is used to excite the sample and the other part is used to generate a start pulse in the start PMT or photodiode. The optical signal at the start PMT generates an electrical START pulse, which is then running through a constant fraction discriminator (CFD) to START the input of time to amplitude converter (TAC) to initialize the charging operation. The other
part of the optical pulse excites the sample and gives rise to an emission of photons. These photons are then detected by STOP PMT to generate an electrical STOP pulse. The STOP pulse pass through another CFD and a variable delay line before it routed to the TAC. On receiving the STOP signal, TAC stops its charging operation and generates an electrical output, having an amplitude proportional to the time difference ($\Delta t$) between the START and STOP pulses reaching the TAC. The TAC output pulse is then fed to the input of a multi channel analyzer (MCA) through an analogue to digital converter (ADC). The ADC generates a numerical value corresponding to the TAC output pulse and thus selects an appropriate channel of the MCA and the count is added to the channel. The above cycle (from excitation to data storage) is repeated a large number of times and as a result a histogram of the counts versus the channel number of MCA is generated. It represents the true emission decay, when the collection rate of emission photons by the STOP pulse is very low, as illustrated by the statistical treatment$^{6,7}$.

The emission decay observed has to be de-convoluted with the instrument response to get the actual lifetime. The instrument response function (IRF) is measured using a scatterer placed in place of the sample. The width of IRF function depends on the excitation source and detection systems used.

### 2.5.3 Fluorescence anisotropy

Anisotropy measurements are based on excitation of fluorophores by absorbing photons of polarized light whose electric vectors are aligned parallel to the fluorophores transition moment. This photoselective excitation results in a partially oriented population of fluorophores in partially polarized fluorescence emission. The relative
angle between the excitation and emission moments determines the maximum anisotropy.

Anisotropy ($r$) is directly related to polarizability ($p$), equations (2.4) and (2.5):

$$\begin{align*}
    r &= \frac{l_{yy} - G l_{yy}}{l_{yy} + 2G l_{yy}} , G = \frac{l_{yy}}{l_{yy} + 2G l_{yy}} \\
    p &= \frac{3 * r}{2 + r} 
\end{align*}$$

The polarized fluorescence, gradually, returns to its unpolarized fluorescence, depending on rotational diffusion and other factors. Rotational diffusion occurs during the lifetime of the excited state, and so, displaces the emission dipole of the fluorophore. The excited state lifetime is about 1-10 ns. In fluid solutions, the fluorophores rotate broadly many times during the excited lifetime, since it rotates in 50-100 ps. Therefore, the orientation of the polarized emission is randomized and the anisotropy value approaches near zero. Another factor that can decrease the anisotropy is the transfer of the excitation between fluorophores. However, the effect of rotational diffusion can be decreased if the fluorophore is bound to a macromolecule, which usually has a high rotational correlation time. In this case, assuming no other processes result in loss of anisotropy, the expected anisotropy is given by the Perrin equation $^8$ (2.6).

$$\begin{align*}
    r &= \frac{r_0}{1 + \frac{\tau}{\theta}} \\
    \theta &= \frac{\eta V}{RT} 
\end{align*}$$

$r_0$ is the anisotropy in the absence of rotational diffusion, and $\theta$ is the rotational correlation time for the diffusion process. Based on that, binding the fluorophore to a protein will slow the the probe’s rate of rotational motion. The smaller the proteins the shorter the correlation times and therefore an expected a lower anisotropy; however, large proteins can have low anisotropy if they are bound to a long-lifetime fluorophore. The
most important point here is that measurement of fluorescence anisotropy is sensitive to factors that affect the rate of rotational diffusion.

Actually, for any fluorophore randomly distributed in solution, with one-photon excitation, the value of \( r_0 \) must be within the range of -0.20 to 0.40 for single photon excitation. A special condition needs to be applied for measuring fundamental anisotropy in order to avoid rotational diffusion. The probes are usually examined in solvents that form a glass at low temperature such as propylene glycol and glycerol, and in diluted solutions. The commonly used solvent for measuring fundamental anisotropy is propylene glycol at -60 to -70°C. In these conditions, the fluorophore remains immobile during the lifetime of the excited state. Therefore, the measured anisotropy value (\( r_0 \)) provides a measure of the angle between the absorption and the emission dipoles (\( \beta \)).

The fundamental anisotropy is given by equation (2.7),

\[
r_0 = \frac{2}{5} \left( \frac{3 \cos^2 \beta - 1}{2} \right) \tag{2.7}
\]

where the angle (\( \beta \)) varies with the excitation wavelength, and consequently, the anisotropy will change. The reason for this change is due to the rotation of the absorption transition moment. As the excitation wavelength changes, the fraction of light absorbed by each transition will change too.

The anisotropy is usually independent of the emission wavelength, because the emission always occurs from the lowest singlet state. The anisotropy spectrum is a plot of anisotropy versus excitation wavelength of a fluorophore in a solution, where solvent relaxation does not occur. If solvent relaxation or emission occurs from more than one state and shows a different emission spectra, then in these cases the anisotropy can be
dependent upon emission wavelength. And so, anisotropy decreases with increasing wavelength because the average lifetime is longer for longer wavelengths.

2.5.4 Measurements of fluorescence anisotropies

There are two methods used to measure fluorescence anisotropy, the L-format and the T-format, Figure 2.14. In the L-format method, a single emission channel is used, while in the T-format, two separate detection systems are used to detect the intensities of the parallel and perpendicular components.

![Diagram of L and T format to measure fluorescence anisotropies](image)

Figure 2.14: L and T format to measure fluorescence anisotropies
In the L-format, each one of the four intensities needs to be measured separately. While in the T-format, a pair of the fluorescence intensities pairs can be measured simultaneously. Therefore, measurements in the T-format are faster than in the L-format. To measure the fluorophore anisotropy and to minimize the background, the measured intensities are corrected to the reference by subtracting the intensities of the reference sample from measured sample. The reference, usually, contains all components of the sample with the exception of the fluorophore. Also, having sample components that produce low fluorescence compared to the fluorophore minimizes the background signal. The excitation wavelength used to measure anisotropy should give the strongest emission intensity, unless at that wavelength there is greater emission of the background\textsuperscript{11}.

The total intensity is given by \((I_{VH} + 2I_{VV})\) which reflect the sum of the fluorophore intensity along the three axes; however, the intensity measured is not proportional to the total intensity or either intensities. It is actually proportional to another combination of vertical and horizontal intensities. To make the anisotropy proportional to the total intensity, the excitation polarizer is oriented in the vertical position and the emission polarizer is oriented 54.7° (the magic angle) from the vertical. At this orientation the perpendicular intensity is twice the intensity of the horizontal, since \(\cos^2 54.7° = 0.333\) and \(\sin^2 54.7° = 0.667\). Another important factor that should be considered during fluorescence anisotropy measurements is the alignment of the excitation and emission polarizers. Both should be exactly positioned in the vertical and the horizontal orientations.
2.5.5 Two-photon absorption

Two-photon absorption (2PA) is a non-linear process, where in two photons absorbed simultaneously to excite the molecule from its ground state to a higher state. The transition energy for this process is equal to the sum of the energies of the two photons absorbed. This phenomena was first described theoretically in 1931 by Maria Goppert-Mayer\textsuperscript{12}, in her doctoral dissertation. She concluded that the energy absorbed by the two photons is similar to that of a one photon that spans the combined gap. Thirty years later, and after the invention of the laser, the phenomenon was observed by Franken et. al. and experimentally verified by Kaiser and Garrett in a Europium-doped crystal, and then observed in a Cesium vapor\textsuperscript{13}.

Two-photon absorption is a rare phenomena, therefore, an intense laser beam is used to observe its effect. The rate equation for the 2PA is second-order in the photon flux and a first-order in the concentration of molecules, equation (2.8),

\[
\frac{dN_{TP}}{dt} = \frac{1}{2} \delta N_{GS} F^2
\]  

(2.8)

\(\delta\) is the two-photon absorption cross section, \(N_{TP}\) is the number of molecules per unit volume in the excited state due to 2PA, \(N_{GS}\) is the number of molecules in the ground state, \(F\) is the photon flux (number of photons per unit area and time, light intensity), and \(t\) is the time.

The great interest in two-photon absorption is due to several aspects; the first is due to the advantages of two-photon excitation instead of one-photon. Excitations with two-photons of near infrared wavelength during 2PA process lower the scattering effect. According to Rayleigh’s law of scattering, equation (2.9), doubling the excitation wavelength (\(\lambda\)) will reduce the scattering effect (\(\sigma_s\)) 16 times\textsuperscript{14–16}.
The second aspect of the 2PA is its ability to increase the penetration depth, and so, to reduce the focusing volume in materials. Also, 2PA depends on the polarization of the excitation light. This is because it is involved in two different transition moments, as well as different states\textsuperscript{17,18}.

### 2.5.6 Ultrafast fluorescence spectroscopy via upconversion

Fluorescence spectroscopy is widely used for studying the structure and function of proteins. In particular, the fluorescence lifetime of the chromophore in protein, such as tyrosine and tryptophan, is important to study protein interactions. This is due to its sensitivity to subtle environmental changes\textsuperscript{19}, and to local conformational changes\textsuperscript{20,21}, in addition to processes accompanied with translational and rotational motion, via anisotropy decay\textsuperscript{22}. The majority of fluorescence decay changes happen within a few picoseconds to nanoseconds, which requires a very short light pulse. The development of ultrafast lasers such as the Ti:sapphire laser enables studies of proteins in the subpicosecond timescale.

In 1975, Maher and Hirsch used a new nonlinear technique to mix the fluorescence excited by an ultrafast laser pulse with another laser pulse in a nonlinear crystal such as Potassium Dihydrogen Phosphate (KDP) and β-Barium Borate (BBO) to generate sum or difference frequency radiation\textsuperscript{23}. The significance of this technique is the ability to provide time resolution equivalent to the pulse width because the mixing process occurs merely for the period of the existence of the second laser pulse. As a result, delaying the gate pulses with a mechanical stage leads to an “optical boxcar approach”. This
technique, called sum frequency generation (also known as fluorescence upconversion), was first applied in biochemical research by Halliday and Topp in 1977\textsuperscript{24}.

The basic concept of the time resolution mechanism using the upconversion technique is based on a cross-correlation between the fluorescence and the probe laser pulse, Figure 2.15. At initial time ($t = 0$), the sample is excited by an ultrafast laser pulse, which could be a second or third harmonic pulse, with a frequency of $\omega_p$. When the fluorescence from the sample $\omega_F$ and the probe laser pulse arrive together at time $t = \tau$ both are cofocused into a nonlinear crystal such as BBO. The crystal usually is oriented at an angle with respect to the fluorescence and laser beams. When the probe laser pulse is present in the crystal, only at that time, the sum frequency photons are generated. The probe laser pulse works as a “light gate” and gives time resolution within the laser pulse width. Varying the delay $\tau$ of the probe laser beam is used to trace the time evolution of the fluorescence.
The sum frequency generation is a “parametric process”, where the energy is conserved and the interaction of light with matter leaves the matter unchanged\(^\text{25}\), equation (2.10).

\[
\hbar \omega_s = \hbar \omega_p + \hbar \omega_F \tag{2.10}
\]

\(\omega\) is the photon frequency. For sum frequency to occur efficiently a condition called “phase matching” must be satisfied, as given in equation (2.11)

\[
\hbar k_s \approx \hbar k_p + \hbar k_F \tag{2.11}
\]

Where, \(\hbar\) and \(k\) are the reduced Planck’s constant and the angular wavenumber, respectively. The subscripts \(s\), \(p\), and \(F\) denote the sum frequency generation, probe laser beam, and the fluorescence beam, respectively.

Figure 2.15: Schematic diagram of upconversion
The quantum efficiency for phase matched sum frequency generation is proportional to probe pulse peak power and square of crystal thickness, but inversely proportional to the focal spot \(^{25}\). BBO has a wide phase-matching wavelength, from 250 – 2000 nm, for mixing with 800 nm. Also, BBO has high damage threshold and low group velocity mismatch. The mismatch between the group velocity of the probe and fluorescence pulse may lead to temporal broadening of the generated sum frequency pulse. Therefore, BBO is preferred for UV upconversion experiments for subpicoseconds pulses\(^ {26}\). In our experiments, to excite Tyrosine, we used THG and BBO crystal, and the layout for studying Tyrosine is shown in Figure 2.16.

![Figure 2.16: Scheme for fluorescence upconversion measurements. THG is the third harmonic generator, BBO is β-Barium Borate, L1-L3 are lenses, M1-M4 are mirrors and D is the diaphragm.](image-url)
2.5.7 Circular dichroism

Circular Dichroism (CD) is an absorption spectroscopy uses circularly polarized light. CD is the difference in the absorption between the left-handed and right-handed polarized light, equation (2.10). It exhibited in molecules that contain one or more chiral chromophores.

\[ \Delta A(\lambda) = A_{LCPL} - A_{RCPL} \]  

\( \lambda \) is the wavelength and \( A \) is the absorbance. CD usually is measured over a range of wavelengths. UV-CD is mostly applied to study the structure of small organic molecules, DNA, and Proteins, while UV/VIS CD is used to study charge transfer in protein-metal complexes. The CD signal can be positive, or negative, depending on the selective absorption of circularly polarized light. Circular polarized light is produced when two linear polarized lights wave are out of phase with respect to each other by a quarter-wave. This can be achieved by a quarter-wave plate; the orientation of this plate will slow one of the linear components with respect to the other by quarter-wave, and therefore, change their refractive indices. The result will be either left circular polarized light (LCPL), or right circular polarized light (RCPL).

Chiral molecules are non superimposable mirror-image isomers. These enantiomers are identical in their photophysical properties, except for their interactions with polarized light and other chiral molecules. A solution of chiral molecules is birefringence, or an anisotropic medium in which right and left circularly polarized light propagate in different speeds. A linearly polarized wave is a superposition of two different handed circularly polarized waves, with equal amplitude and phase. When it passes through a birefringent medium, the phase relationship between the circularly
polarized wave changes, and the resultant linearly polarized wave rotates. This phenomenon is called optical rotation, and when it is measured as a function of wavelength it is called optical rotatory dispersion spectroscopy (ORD).

However, when the linear polarized light passes through an optically active sample, the amplitude of the less absorbed component will be larger than the more strongly absorbed component. Therefore, when the amplitudes are combined, an ellipse forms rather than a circular shape. The degree of ellipticity ($\theta$) is defined as the tangent ratio between the minor and major axes of the elliptic. Linearly polarized light has zero degrees of ellipticity, while fully LCPL has a positive 45° ellipticity and RCPL has a negative 45°.

Since optical rotation is associated with the change of the phase that is due to the change in the refractive index, it is observed at all wavelengths. CD is associated with the change in the absorbencies of LCPL and RCPL, and it is observed only at wavelengths where a chiral molecule absorbs light. CD and ORD can be interconverted by Kramers-Kronig transformations relation, equation (2.11):

$$\Delta A = \frac{\theta}{32.982}$$  \hspace{1cm} (2.11)

Molar ellipticity is reported in deg.cm$^2$.dmol$^{-1}$ or deg.M$^{-1}$.m$^{-1}$ and calculated as in equation (2.12):

$$[\theta]_M = 10 \frac{\theta}{c \times l}$$  \hspace{1cm} (2.12)

Where $\theta$ is in mdeg, $c$ is the concentration in molar, and $l$ is the pathlength in cm. The values of molar ellipticity can be converted to a molar extinction coefficient by equation (2.13):

$$\Delta \varepsilon = \frac{[\theta]_M}{3298.2}$$  \hspace{1cm} (2.13)
2.5.8 Circular dichroism for biological molecules

Most biological molecules are chiral, for instance proteins consist of amino acids and 19 of the 20 common amino acids are chiral. The secondary structure of proteins, DNA and RNA, is also chiral and contribute to the circular dichroism signal. The CD spectrum of a protein or a DNA molecule is highly influenced by the three dimensional structure of these macromolecules, and so, the spectrum is not a simple reflect of the individual residues or bases\textsuperscript{28,29}. Each macromolecule structure has a specific circular dichroism signature, therefore, this can be used to classify structural elements and to observe changes in their structure\textsuperscript{30}.

The secondary structure of proteins is easily determined by CD spectroscopy in the far-UV region (190-250 nm). At these wavelengths, the peptide bond absorbs light and the signal arises from the protein fold. The alpha helix has negative minima at 208 nm and 222 nm, and positive maxima at 190 nm. Beta sheets show a negative peak at 218 nm and a positive peak at 196 nm. The random coil has a positive peak at 212 nm and a negative one at around 195 nm, Figure 2.17. The CD signal determines the structure based on the average of the entire molecular, but it cannot determine which specific residues are involved in each secondary structure portion.
In the near-UV spectral region (250-350 nm), CD can be useful to determine the tertiary structure of proteins. In this region aromatic amino acids and disulfide bonds are the chromophores and their CD signals are sensitive to the overall protein tertiary structure. Tyrosine residues are responsible for the signal from 270-290 nm, phenylalanines are responsible for the signal from 250-270 nm, while tryptophan residues are responsible for the signals from 280-300 nm. Disulfide bond signals are broad and weak signals spread throughout the entire near-UV region. Therefore, the presence of
significant signals in the near-UV region is an indication of a well-defined folded protein. However, the signals in this region are weak, relative to the far-UV region\(^3\).

The CD signal for proteins are usually expressed in mean residue ellipticity \([\theta]_{MR}\), equation (2.14). This unit reports the molar ellipticity for individual protein residues to make it easy to compare different proteins which differ in molecular weight.

\[
[\theta]_{MR} = 100 \frac{\theta}{CMR \cdot l} \tag{2.14}
\]

CMR is the mean residue concentration and can be calculated in two different ways, either by multiplying the molar concentration of protein (\(C\)) by the number of residues (\(N\)) or by dividing the concentration of protein in unit of gram per liter over the average amino acid residue weight, which is 113, equation (2.15)\(^3\).

\[
CMR = \frac{C \text{ (molar)}}{N}, \quad \text{or} \quad CMR = \frac{C (\frac{\theta}{l})}{113} \tag{2.15}
\]

2.6 References


CHAPTER III

BIOPHYSICAL PROPERTIES OF WLN3-6, WLN4-6 AND WLN5-6

3.1 Introduction

The understanding of the structure and stability of Wilson disease protein will shed light into the copper transport mechanism and the effect of mutations on it. Among the three major domains of Wilson disease protein, the N-terminal domain is the largest one and plays a major role in copper acquisition and transport. The N-terminal region of Wilson protein has about 650 amino acid residues, and it has a mass of about 72 kDa. It consists of six metal binding domains; each one has a conserved GMT/HCXXCXXXIE motif except that the threonine residue in the in the third domain is replaced by histidine. Each domain has a pair of cysteines, which are two residues away from each other, that bind copper ions. Each of the metal binding domains has around 70 amino acids, and a mass of around 7 kDa.\textsuperscript{1,2}

All these domains are linked with chains of various lengths. The longest one is between domains 4 and 5, consisting of 57 residues. The linker between domains 2 and 3 is also long, but it is only 40 amino acids. Linkers between 1 and 2, and 5 and 6 are the shortest with 12 and 8 amino acids, respectively. The linker between domains 3 and 4 is 30 amino acids. The structure of the N-terminal metal binding domains (WLN1-6) has not been determined yet. NMR structures of individual metal binding domains from the highly homologous human copper-transporting ATPase ATP7A have been determined\textsuperscript{3-5}
and some Wilson protein constructs, WLN5-6\textsuperscript{1}, WLN2\textsuperscript{6} and WLN3-4\textsuperscript{7}, are also available, Figure 3.1.

All the six N-terminal metal binding domains of ATP7A, also called Menkes (MNK), receive copper from the cytoplasmic metallochaperone HAH1\textsuperscript{8}. However, the six soluble domains interact differently with HAH1. Investigations using yeast two-hybrid assays\textsuperscript{9–11} and NMR\textsuperscript{12,13} show that MNK first and fourth domains form a metal-mediated adduct with HAH1 while the sixth domain is simultaneously loaded with Cu(I) without formation of the adduct. On the other hand, in case of Wilson protein (WLN), both the second and fourth domains can form a copper-dependent adducts with HAH1\textsuperscript{14}. Interestingly, WLN5-6 can receive Cu(I) from domain 4, but not from HAH1\textsuperscript{1}. On the other hand, all domains were able to receive Cu(I) from HAH1, when metal binding cysteines were mutated to serine in five of the six metal binding motifs\textsuperscript{15}.

The fact that WLN has six repeated ferrodoxin folds is really interesting, since other species, such as bacteria or yeast have fewer domains\textsuperscript{16}. Moreover, not all domains form adducts with HAH1 or are needed for copper trafficking. Some domains accept copper from other domains. All these facts raised the question about the biophysical properties and the identity of each domain. In other words, why do similar domains behave differently? And what is the importance of each domain within the whole protein? To address these questions, biophysical properties of domains within a construct of other domains need to be studied. Although NMR structure studies are available on a few domains of the WLN, the biophysical properties of these domains are not well understood. Detailed investigations can shed light on the structure and stability of these domains. In this chapter, the biophysical properties of WLN5-6, WLN4-6 and WLN3-6
are determined by using circular dichroism, dynamic light scattering, isoelectric focusing gel, and mass spectrometry.

Figure 3.1: Cartoon structure of WLN3-4, WLN5-6, and WLN2
3.2 Biophysical characterization and analysis of WLN5-6

WLN5-6 was characterized during the purification process using SDS-PAGE (15%) in order to collect the fractions that contain the protein fractions. Mass spectroscopy was used to determine the mass of the protein. ESI-MS showed the mass of the protein to be 16042 g/mol, Figure 3.2, the average mass calculated by DNA Strider® program is 16048 Da.

Figure 3.2: ESI-IT/MS data showing the molecular weight of WLN5-6 determined by MagTran software
3.3 Circular dichroism

The protein samples for CD spectra were prepared by buffer exchange into 50 mM sodium phosphate buffer, pH 7.5. The protein concentration was determined by the BCA assay (Pierce). Guanidine Hydrochloride (GuHCl) 8 M was prepared as a stock solution. The refractive index of GuHCl solution was measured to accurately determine GuHCl concentration using the following equation (3.1):

\[
C = 57.147 (\Delta N) + 38.68 (\Delta N)^2 - 91.6 (\Delta N)^3
\]  \hspace{1cm} (3.1)

where \(C\) is the concentration of GuHCl in moles per liter, and \(\Delta N = (\text{refractive index of solution} - \text{refractive index of water})^{17}\).

The protein concentration was constant in all samples, 18 \(\mu\)M, whereas the GuHCl concentration varied from 0 – 8 M. CD was acquired with a JASCO J-815 CD Spectropolarimeter. The unfolding was monitored at 222 nm, Figure 3.3, in 0.5 M increments.

![Graph](image.png)

**Figure 3.3**: Molar ellipticity for the unfolded WLN5-6 using GuHCl
The denaturation profile of WLN5-6 shows a sigmoidal and a cooperative transition from the native (N) to the denatured state (D), \( N \leftrightarrow D \). GuHCl is a salt, and will be ionized in aqueous solution to give the positive guanidinium ion and negative chloride ion. At low concentrations, these ions are presumed to neutralize and mask the positive and negative charged amino acid side chains. This affects eliminates electrostatic interactions that may stabilize or destabilize the protein. If the protein has more electrostatic repulsive than attractive forces then at low concentration of GuHCl, increased protein stability may mask electrostatic repulsion. The reverse is true though\(^\text{18}\).

At high concentrations, the binding ability of the guanidinium ion is expected to predominate despite the types of electrostatic interactions present in the protein, and therefore, it becomes a denaturant and drives the equilibrium toward the unfolded state. In all the unfolding processes, the high ionic strength of GuHCl solution is also expected to contribute in masking the electrostatic interaction. The protein stability derived from GuHCl denaturation is actually a relative estimation of the hydrophobic interactions\(^\text{18}\).

The standard free energy of unfolding \( \Delta G_D^0 \) was calculated using the linear extrapolation method. The slope (m) of the plots between \( \Delta G_D^0 \) and [GuHCl] is a measure of the ability of GuHCl and any denaturant, in general, to unfold a protein, equation (3.2). It was also observed that the m value is proportional to the number of denaturant binding sites exposed on unfolding\(^\text{19,20}\). The dependent variable was normalized to show the fractional denaturation (\( F_D \)), equation (3.3), Figure 3.4.

\[
\Delta G_D^0 = \Delta G_{water}^0 - m[GuHCl] \quad (3.2)
\]
\[ F_D = \frac{Y_S - Y_N}{Y_D - Y_N} \]  

(3.3)

\( Y_S \) is the CD signal at a specific concentration of GuHCl, \( Y_N \) is the CD signal for the native protein, \( Y_D \) is the CD signal for the denatured protein. \( Y_N \) and \( Y_D \) for WLN5-6 is the average molar ellipticity of the native and denatured states calculated from sigmoidal fit to the unfolding. \( Y_N \) is found to be \((-1.99 \times 10^6) \text{ deg.cm}^2\text{.dmol}^{-1} \) and \( Y_D \) is \((-3.02 \times 10^5) \). Since the free energy is related to the equilibrium constant via the equation (3.4), and \( K_{eq} \) is independent of protein concentration, therefore it can be calculated from the ratio of the molecules folded, equation (3.5).

\[ \Delta G_D^0 = -RT \ln K_{eq} \]  

(3.4)

\[ K_{eq} = \frac{F_D}{1 - F_D} \]  

(3.5)

Figure 3.4: Fraction denatured of WLN5-6 as GuHCl is added
Figure 3.5: $\Delta G$ of WLN5-6 at different [GuHCl], at 293 K

$\Delta G_{\text{water}}$ was calculated by extrapolating the straight line to [GuHCl] = 0, giving a value of 12.2 ± 0.7 kJ/mol. GuHCl concentration where F_D equal 0.5 is 3.3±0.7 M., Figure 3.5.

Figure 3.6: Thermal unfolding of WLN5-6
The thermal unfolding of WLN5-6, Figure 3.6, closely approaches a two-state unfolding model. The unfolding profile can be fitted to a sigmoidal fit in order to obtain the melting temperature ($T_m$), which was found to be $81.5 \pm 2.7 ^\circ C$. The two-state model was used to determine the free energy of unfolding, $\Delta G$, at 293 K, Figure 3.7. $K_{eq}$ can be used to calculate $\Delta S$ and $\Delta H$ for the transition from native to denatured state, $N \leftrightarrow D$, assuming the denaturation is cooperative (no intermediates, and that $\Delta H$ is independent of temperature over the range studied). According to the van’t Hoff equation (3.6), a plot of $\ln K_{eq}$ versus $(1/T)$ is a straight line with a slope of $(-\frac{\Delta H}{R})$ and a y-intercept of $(-\frac{\Delta S}{R})$.

$$\ln K_{eq} = -\frac{\Delta H}{RT} + \frac{\Delta S}{R} \quad (3.6)$$

$\Delta G$ then calculated at certain temperature by the following equation, equation (3.7):

$$\Delta G = \Delta H - T \Delta S \quad (3.7)$$

![Figure 3.7: Analysis of thermal unfolding for WLN5-6](image-url)
The value of $\Delta G$ at 293 K is calculated for WLN5-6 by the two state model, $12.7 \pm 3$ kJ/mol, which agree well with $\Delta G_{H2O}$ which is 12.2 kJ/mol.

### 3.4 Biophysical characterization and analysis of WLN4-6

WLN4-6 purity was characterized using SDS-PAGE (15%), Figure 3.8. Mass spectroscopy was also used to determine the mass of the protein. ESI-MS showed the mass of the protein to be 29198 g/mol, Figure 3.8, the average mass calculated by DNA Strider® program is 29192 Da.
Figure 3.8: ESI-IT/MS data showing the molecular weight of WLN4-6 determined by MagTran software
3.4.1 Chemical and thermal denaturation of WLN4-6

![Graph showing the denaturation of WLN4-6 by GuHCl](image)

**Figure 3.9: Unfolding of WLN3-6 by GuHCl**

The denaturation of WLN4-6 by GuHCl exhibits multiple phases, Figure 3.9. CD data were fit to a three-state model, equation (3.8), and the values of the free energies of unfolding extrapolated to zero denaturant and the m values are given in Table 3.1. The first transition midpoint occurs at 3.0 M GuHCl, while the second transition occurs at 5.8 M GuHCl. The first transition state matches the transition for WLN5-6 that was described earlier. Therefore, the second transition at a higher [GuHCl] is due to the unfolding of domain 4. Similar results are also observed for unfolding of WLN4 and WLN1-4.
Table 3.1: Free energies and m values of unfolding of WLN4-6

<table>
<thead>
<tr>
<th>Conversions</th>
<th>$\Delta G^{H_2O}$ (kJ/mol)</th>
<th>$m$ (kJ/mol/[GuHCl])</th>
<th>[GuHCl]$_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N → I</td>
<td>19.9 ± 2.8</td>
<td>6.8 ± 0.9</td>
<td>3.00 ± 0.58</td>
</tr>
<tr>
<td>I → U</td>
<td>35 ± 11</td>
<td>6.1 ± 2.0</td>
<td>5.8 ± 2.6</td>
</tr>
</tbody>
</table>

\[
S_{obs} = \frac{S_N + \exp\left\{-\Delta G_{N\rightarrow I}/RT\right\} + S_U \cdot \exp\left\{-\Delta G_{N\rightarrow I}/RT\right\} \cdot \exp\left\{-\Delta G_{I\rightarrow U}/RT\right\}}{1 + \exp\left\{-\Delta G_{N\rightarrow I}/RT\right\} + \exp\left\{-\Delta G_{N\rightarrow I}/RT\right\} \cdot \exp\left\{-\Delta G_{I\rightarrow U}/RT\right\}}
\]  

(3.8)

Where, $\Delta G_{N\rightarrow I} = \Delta G_{N\rightarrow I}^0 - m_{N\rightarrow I}[GuHCl]$ and $\Delta G_{I\rightarrow U} = \Delta G_{I\rightarrow U}^0 - m_{I\rightarrow U}[GuHCl]$.

$\Delta G_{N\rightarrow I}$ for the first transition from the native state to the intermediate has value of 20 kJ/mol. The second transition, $\Delta G_{I\rightarrow U}$, from the intermediate state to the denatured state is 35 kJ/mol, it’s higher by 15 kJ/mol but the m-value is almost the same. m-value is the slope of the linear fitting and represents a proportionality constant show rate of $\Delta G$ changes with increasing [GuHCl], $\delta \Delta G / \delta [GuHCl]$. The m-values are 6.8 kJ/mol and 6.1 kJ/mol [GuHCl] for first and second transitions, respectively.

Heat denaturation of WLN4-6 was performed on 3.5 µM samples in a cuvette with a 10 mm path length. The program was set to increase the temperature from 20°C to 95°C at rate of 1°C/min. The CD signal was measured at 222 nm after maintaining the temperature at each value for 10 seconds. The signal was recorded once a minute. The WLN4-6 thermal denaturation profile shows a sigmoidal shape that has a transition melting point ($T_m$) of 74.4 °C, Figure 3.10. To obtain thermodynamic properties such as $\Delta H$ and $\Delta S$, the Van’t Hoff equation was used, equation (3.6). $K_{eq}$ was calculated based on the two-state model where the fraction of molecules in the denatured state is ($F_D$) and $K_{eq}$ is determined from equation (3.3) and equation (3.5), respectively.
Plotting ln $k_{eq}$ against $1/T$ gives a linear fit that can be used to obtain the change in enthalpy ($\Delta H$) from the slope and change in entropy ($\Delta S$) from the intercept. The van’t Hoff plot of WLN4-6 is shown in Figure 3.11. The change in entropy is 0.45 KJ/mol, $\Delta H$ is 156.2 KJ/mol and $\Delta G$ is 24.51 kJ/mol. $\Delta G$ was calculated at 20°C according to equation (3.7). Free energy is almost that obtained for the first transition from unfolding using GuHCl. This means that it is corresponding to the unfolding of WLN5-6, and $\Delta G$ for WLN5-6 is half of that was obtained for WLN4-6.

Figure 3.10: Thermal unfolding of WLN4-6
The presence of water molecules is critical for protein folding and the maintenance of the folded structure. Thermal denaturation of protein involves a change in the protein structure and results in increased hydrogen bond breakage. Therefore, the free energy change of unfolding is due to the combined effects of both protein unfolding and hydration changes. The free energy of stability of a typical protein is only 40-90 kJ mol$^{-1}$ which equivalent to very few hydrogen bonds. There are both enthalpic and entropic contributions to the free energy that change with temperature, and so give rise to heat denaturation. Upon thermal unfolding the interior polar and nonpolar groups become exposed to water and so have consequential changes in the water-water interactions on $\Delta G_N^D$, $\Delta S_N^D$ and $\Delta H_N^D$.

If $\Delta G_N^D$ has a negative value then denaturation is favored. $\Delta S_N^D$ and $\Delta H_N^D$ are generally both positive during thermal denaturation. Water naturally has more structure at lower temperatures, and so the enthalpy of transfer of polar groups from the protein
interior into water is positive at low temperatures and negative at higher temperatures. This is because polar groups generate negative enthalpy change due to the ordered water molecules that surrounded them and the increasing of molecular interactions. The breakdown of the water structure is the major contributor to the positive change in the overall enthalpy at lower temperatures\textsuperscript{23}.

### 3.4.2 Isoelectric point (pI) of WLN4-6

The isoelectric point of the individual domains of the N-terminal metal binding domains of Wilson protein vary from acidic for WLN4 to basic for WLN1. Isoelectric focusing of pure WLN4-6 was determined by using a pH 3-9 gradient gel on a PhastSystem Separation unit. The pI of WLN4-6 was determined to be 4.86 from linear fitting of the broad pI protein standard migration against their pIs, in Figure 3.12. The predicted and experimental pI of WLN5-6 is 4.96. Therefore, WLN4 and the linker between domains 4 and 5 must have on average a pI similar to WLN5-6. The pI of WLN4 is 3.85 as confirmed experimentally and the linker pI is 6.21 as predicted by DNA Strider™.
3.4.3 Dynamic light scattering

The hydrodynamic radius of WLN4-6 was determined through dynamic light scattering performed in the Keck laboratory at Western Michigan University using a DynaPro Titan instrument from Wyatt Technology Corporation. A 200 µM sample of WLN4-6 in 50 mM Sodium Phosphate buffer was centrifuged at 15,000 rpm for 14 minutes, then the supernatant was filtered through a 0.22 µm filter. The instrument was set up to make 10 acquisitions for 10 seconds each. Multiple measurements were performed and the representative data were cumulatively fit, Figure 3.13.

Figure 3.12: Phast gel 3-9 and isoelectric focusing of WLN4-6
Figure 3.13: Autocorrelation fit of the hydrodynamic radius of WLN4-6

The value of hydrodynamic radius ($R_h$) obtained from the measurements was 3.53±0.4 nm. The predicted hydrodynamic radius for WLN4-6 if considered as globular protein, which has mass of 29193 Da by the Dynamics 6.7.6 software, is 2.5 nm. Therefore, the frictional ratio is 1.41. This ratio implies that WLN4-6 molecules have non-spherical shape. Spherical protein should have a frictional ratio of 1.0.

3.5 Biophysical characterization and analysis of WLN3-6

3.5.1 Mass spectroscopy (MS)

SDS-PAGE was used during the purification of WLN3-6 to determine the fractions that contain the desired protein which elute from the column and to determine the purity of the protein. WLN3-6 migrates on SDS-PAGE on apparent mass of 40 kDa.
WLN3-6 has 379 amino acids, and its mass calculated based on average amino acid usage by DNA Strider® program is 39.852 kDa. MS separates proteins based on their mass-to-charge ratio. To separate by MS, a protein is ionized in one of several ways, and then accelerated by electric or magnetic fields. In some cases, the charged protein will break apart to produce ions. The pattern of ions produced is dependent on the structure of the protein so that they may be used to determine the primary structure of the protein. Most MS instruments in use today ionize proteins in ways that minimizes protein fragmentation to allow a true mass determination\textsuperscript{24}.

WLN3-6 was prepared for MS by exchanging the buffer into water using a PD10 column (GE Healthcare). The protein concentration was determined using the BCA assay (Pierce). A sample of 8 µM protein, 50% methanol and 0.1% formic acid was injected into the spectrometer. The mass of the protein was determined using electrospray ionization-ion trap mass spectroscopy (ESI-ITMS) in Dr. Andre Venter’s laboratory at Western Michigan University. The error was within 0.01 % of the predicted mass of 39852 Da, Figure 3.14.
Figure 3.14: ESI-IT/MS data showing the molecular weight of WLN 3-6 determined by MagTran software
3.5.2 Dynamic light scattering

Dynamic Light Scattering (DLS) is used to measure hydrodynamic sizes, polydispersities and aggregation effects of protein samples. It measures the laser light that is scattered from dissolved macromolecules or suspended particles. Due to the Brownian motion of the molecules and particles in solution, fluctuations of the scattering intensity can be observed. From the correlation function, the diffusion coefficient \( D \) (units of \( \text{cm}^2/\text{sec} \)) of the molecules can be calculated by fitting the data. Finally the hydrodynamic radius \( (R_h) \) of the particles and molecules can be calculated by equation (3.9)\(^{25}\):

\[
D = \frac{kT}{6\pi\eta R_h}
\]

where \( k \) is the Boltzmann constant, \( D \) the translational diffusion coefficient, \( T \) the temperature and \( \eta \) the solvent viscosity.

DLS interfaced with gel filtration is also used to determine the unknown molar mass of a protein when calibrated with protein standards of known molecular weights. The light scattering experiment for WLN3-6 was performed at the Keck Foundation Biotechnology Resource Laboratory at Yale University. HPLC SEC Laser Light Scattering of WLN3-6 reveals that the average molecular weight is 39220 g/mol, Figure 3.15, while the molecular mass calculated by the amino acid composition is 39852 Da.
Figure 3.15: Mass distribution of WLN3-6 protein versus volume. The solid lines (line) indicate(s) the trace from refractometer while "dots" are weight average MW

The hydrodynamic radius ($R_h$) determined from “on line” dynamic light scattering measurement is $3.67 \pm 0.12$ nm, Figure 3.16. It was calculated using an autocorrelation function fitting of the dynamic light scattering data, Figure 3.17. This value corresponds to a translational diffusion coefficient, $D_t$ ($6.66 \pm 0.22 \times 10^{-7}$ cm$^2$/sec. If we assume a molecular weight of 39190 g/mol and a partial specific volume of 0.734 ml/g an equivalent sphere radius, $R_s$, can be determined as 2.25 nm. The frictional ratio can be determined by dividing $R_h$ by $R_s$, giving a value of 1.63. If WLN3-6 were a perfect sphere, it would have a frictional ratio very close to 1. Therefore, WLN3-6 molecules have a non-spherical shape.
Figure 3.16: Hydrodynamic radius (Rh) determined from “on line” dynamic light scattering of WLN3-6

Figure 3.17: Autocorrelation fit of hydrodynamic radius of WLN3-6
3.5.3 Circular dichroism

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; there is a weak but broad \( n \rightarrow \pi^* \) transition centered around 220 nm and a more intense \( \pi \rightarrow \pi^* \) transition around 190 nm\(^{26}\).

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 raw data from CD spectra in millidegrees can be converted into molar ellipticity, \([\theta]\) (deg·cm\(^2\)·dmol\(^{-1}\)), by equation (3.10) with the parameters \( \theta \), millidegrees; \( c \), protein concentration in mol/L; \( l \), cell pathlength in cm:

\[
[\theta] = \frac{\theta}{10 \times c \times l} \quad (3.10)
\]

Usually, protein molar ellipticity is converted to a normalized value that is independent of the length of the protein. Mean residue ellipticity (MRE) is used for this purpose; it is the measured molar ellipticity \([\theta]\) of the molecule divided by the number of residues in the molecule, equation (3.11).

\[
MRE = \frac{[\theta]}{\text{# of residues}} \quad (3.11)
\]

The protein samples for CD spectra were prepared by buffer exchange into 50 mM Na phosphate, pH 7.5. The protein concentration was determined by the BCA assay
Guanidine Hydrochloride (GuHCl) 8 M was prepared as a stock solution. The refractive index of GuHCl solution was measured to accurately determine GuHCl concentration using the following equation (3.12):

\[ C = 57.147 \Delta N + 38.68 (\Delta N)^2 - 91.6 (\Delta N)^3 \]  

(3.12)

where \( C \) is the concentration of GuHCl solution in moles per liter and \( \Delta N \) = refractive index of solution - refractive index of water\(^{17}\).

The unfolding of WLN 3–6 was performed chemically using GuHCl and thermally by increasing the temperature from 25°C to 95°C, Figure 3.18. CD spectra were acquired with a JASCO J-815 CD Spectropolarimeter. The unfolding was monitored at 222 nm.

![Graph showing molar ellipticity vs. GuHCl concentration with data points and fitted curve.](image)

\( \Delta G_{H2O}^{N,U} = 24.9 \pm 6.9 \text{ KJ/mol} \)

\( C_m = 4.8 \pm 1.8 \text{ M} \)

\( \Delta G_{H2O}^{N,I} = 20.4 \pm 4.0 \text{ KJ/mol} \)

\( C_m = 2.6 \pm 0.51 \text{ M} \)

Figure 3.18: Unfolding of WLN3-6 by [GuHCl]

The denaturation of WLN3-6 shows multiple phases as that of WLN4-6. The denaturation data were fit to a three-state model, equation (3.8), and the values of the
free energies of unfolding extrapolated to zero GuHCl ($\Delta G_{N\rightarrow I}$ and $\Delta G_{I\rightarrow U}$), the corresponding m values for the two steps are given in Table 3.2. The $\Delta G_0$ of $N \leftrightarrow I$ and $I \leftrightarrow N$ transitions, extrapolated to zero GuHCl provides an independent estimation of the free energies of the $N \leftrightarrow I$ and $I \leftrightarrow N$ transitions.

Table 3.2: Free energies and m values of denaturation of WLN3-6 by GuHCl

<table>
<thead>
<tr>
<th>Conversions</th>
<th>$\Delta G^{H2O}$ (kJ/mol)</th>
<th>m (kJ/mol/[GuHCl])</th>
<th>[GuHCl]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N \rightarrow I$</td>
<td>20.4 ± 4.0</td>
<td>7.9 ± 1.6</td>
<td>2.6 ± 0.5</td>
</tr>
<tr>
<td>$I \rightarrow U$</td>
<td>24.9 ± 6.9</td>
<td>5.1 ± 1.4</td>
<td>4.8 ± 1.8</td>
</tr>
</tbody>
</table>

Chemical unfolding shows that there is a transition around 2.6 M GuHCl probably corresponding to the unfolding of domains WLN5-6 and WLN3 together. The second transition occurs at 4.8 M GuHCl, perhaps due to domain 4 unfolding as noticed for WLN4-6 unfolding. However, the intermediate in the case of WLN3-6 is less stable than WLN4-6 when comparing the free energy of the second transition, $I \leftrightarrow U$, this instability could be due to the repulsion forces between the acidic residues from domain 4 and the residues from domain 3 and the linker, but the error of this determination is high.
The thermal unfolding of WLN3-6 is shown in Figure 3.19. This experiment shows a melting temperature ($T_m$) of 75 °C, which is relatively high and similar to...
WLN4. The Van’t Hoff plot of WLN3-6 is shown in Figure 3.20. The free energy (ΔG) of unfolding obtained from that plot is 21.91 kJ/mol. It is similar to that obtained for the first transition state from unfolding using GuHCl.

3.5.4 Isoelectric point

In IEF, each protein migrates through the support layer until it is "trapped" at the point where the pI of the protein is the same as the pH gradient formed in the support media. IEF was used to determine the pI of WLN3-6 protein, and it was found to be 4.96, Figure 3.21, which is same as calculated by DNA Strider program of, 4.94.

Figure 3.21: Phast gel 3-9 and isoelectric focusing of WLN3-6
3.6 High resolution gel filtration

High resolution gel filtration was performed in order to determine the molecular weight, and obtain information about the shape and the radius of WLN constructs. In this experiment, Superdex 200 HR 10/30 column (GE Healthcare), high molecular weight calibration kit (GE Healthcare) and low molecular weight markers kit (Aldrich) were used. The standard proteins were run individually and as a mixture, the mixture was prepared as the recommended concentrations of proteins by the protocol GE Healthcare, code no 28-4038-41. The void volume of the column was determined through the elution volume of Ferritin. All proteins were diluted with elution buffer (50 mM sodium phosphate, 150 mM NaCl, 10 mM DTT, pH = 7.2). Proteins were injected into a 100 μl loop, and 240 μl of buffer was used to empty the loop. The flow rate was 0.4 ml/min. The molecular weight (M_r) of proteins of interest was calculated according to a calibration curve, Figure 3.22, plotting the logarithm of the molecular weight (Log M_r) of the standards versus the partition coefficient (K_{av}), equation (3.13),

\[
K_{av} = \frac{v_e - v_0}{v_c - v_0}
\] (3.13)

where \(v_e\) is the elution volume, \(v_0\) is the void volume, and \(v_c\) is the column volume (which is 23.55 ml for Superdex 200 HR 10/30). Results are shown in Table 3.3. The Stoke’s radius was determined by plotting the square root of the logarithm of partition coefficient against the stoke’s radius of the standard proteins, Figure 3.23.
Table 3.3: Elution volumes of the standards and the proteins

<table>
<thead>
<tr>
<th>Standard Proteins</th>
<th>$M_r$</th>
<th>$R_{st}$</th>
<th>Conc (mg/ml)</th>
<th>$V_e$</th>
<th>$K_{av}$</th>
<th>Log $M_r$</th>
<th>Log $K_{av}$</th>
<th>$\sqrt{\log K_{av}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ferritin</td>
<td>440000</td>
<td>61</td>
<td>0.3</td>
<td>8.747</td>
<td>-</td>
<td>5.643</td>
<td>-</td>
<td>-0.161 0.4012</td>
</tr>
<tr>
<td>Ribonuclease A</td>
<td>13700</td>
<td>16.4</td>
<td>3</td>
<td>18.965</td>
<td>0.690</td>
<td>4.137</td>
<td>-0.354</td>
<td>0.5952</td>
</tr>
<tr>
<td>Ovalbumin</td>
<td>43000</td>
<td>30.5</td>
<td>4</td>
<td>16.278</td>
<td>0.509</td>
<td>4.633</td>
<td>-0.293</td>
<td>0.5418</td>
</tr>
<tr>
<td>Conalbumin</td>
<td>75000</td>
<td>-</td>
<td>3</td>
<td>15.294</td>
<td>0.442</td>
<td>4.875</td>
<td>-0.354</td>
<td>0.5952</td>
</tr>
<tr>
<td>Aldolase</td>
<td>158000</td>
<td>48.1</td>
<td>4</td>
<td>13.856</td>
<td>0.345</td>
<td>5.199</td>
<td>-0.462</td>
<td>0.6797</td>
</tr>
<tr>
<td>Carbonic Anhydrase</td>
<td>29000</td>
<td>-</td>
<td>2</td>
<td>17.489</td>
<td>0.591</td>
<td>4.462</td>
<td>-0.229</td>
<td>0.4783</td>
</tr>
<tr>
<td>WLN 5-6</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>16.925</td>
<td>0.552</td>
<td>4.548</td>
<td>-0.258</td>
<td>0.5076</td>
</tr>
<tr>
<td>WLN 4-6</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>15.613</td>
<td>0.464</td>
<td>4.818</td>
<td>-0.334</td>
<td>0.5776</td>
</tr>
<tr>
<td>WLN3-6</td>
<td>-</td>
<td>-</td>
<td>1</td>
<td>14.784</td>
<td>0.408</td>
<td>4.988</td>
<td>-0.39</td>
<td>0.6241</td>
</tr>
</tbody>
</table>

Figure 3.22: Standards calibration curve of high resolution gel filtration
Determination of the molecular weight by high resolution gel filtration is based into two assumptions; the proteins have spherical shape and they do not interact with the gel. The results showed that the molecular weight of the unknowns protein was almost two times the actual masses, Table 3.4. This indicates that the proteins either interact with the gel itself or contaminants in the gel, or the proteins have a nonspherical shape which makes them elute earlier than standards of the same mass. To investigate the interaction with the contaminants in the gel, the column was washed with 0.1 M NaOH solution and then the standards and the samples were run again. The results were reproducible and, if there is an interaction between the gel and the proteins, then that would cause the proteins to elute later. Based on this, an interaction of the proteins with the gel was excluded. Interestingly, when other domains of Wilson Protein (WLN 1-4, 39783 Da and WLN 5-6, 16045 Da) were used as internal standards, the mass was close to the masses calculated from amino acid composition and the radii were close to that measured by the dynamic light scattering, Table 3.4. This indicates that the proteins do not have the same spherical
shape as the standards. This result is also supported by dynamic light scattering experiment for WLN5-6, WLN 3-6 and WLN4-6, which show a frictional ratio of 1.26, 1.63 and 1.43, respectively.

Table 3.4: Molecular weight of proteins by high resolution gel filtration and the calculated one by amino acid composition

<table>
<thead>
<tr>
<th>Protein</th>
<th>$M_r$ (Da)</th>
<th>$R_{st}$ (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Amino acids composition</td>
<td>Gel Filtration using internal standards</td>
</tr>
<tr>
<td>WLN5-6</td>
<td>16045</td>
<td>16045</td>
</tr>
<tr>
<td>WLN4-6</td>
<td>29193</td>
<td>30840</td>
</tr>
<tr>
<td>WLN3-6</td>
<td>39854</td>
<td>48986</td>
</tr>
</tbody>
</table>

3.7 Secondary structure of WLN5-6, WLN4-6 and WLN3-6

The estimation of the helical content of a protein can be estimated through the CD signal at 222 nm by comparison to the maximum helicity ($X_H^\infty$) for the protein expected based on the number of residues. The MRE for maximum helicity ($X_H^\infty$) is calculated by the following equation (3.14)\(^{27}\), and the percentage helicity is calculated as in equation (3.15). CD scan for all proteins discussed is shown in Figure 3.24. The results are summarized in Table 3.5.

$$X_H^\infty = -40000 \times (1 - \left( \frac{2.5}{\# \text{Res}} \right))$$ (3.14)
\[
\%\alpha = \frac{MRE_{222}}{X_H^{\infty}} \times 100
\]

(3.15)

Table 3.5: Percentage of helicity for WLN5-6, WLN4-6, and WLN3-6

<table>
<thead>
<tr>
<th>Protein</th>
<th># Res</th>
<th>(X_H^{\infty})</th>
<th>(%\alpha)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN5-6</td>
<td>150</td>
<td>-39333.3</td>
<td>34.9 ± 2.6</td>
</tr>
<tr>
<td>WLN4-6</td>
<td>277</td>
<td>-39639</td>
<td>28.3 ± 1.2</td>
</tr>
<tr>
<td>WLN3-6</td>
<td>379</td>
<td>-39736.1</td>
<td>20.9 ± 1.8</td>
</tr>
</tbody>
</table>

Figure 3.24: CD scan of WLN5-6, WLN4-6, and WLN3-6
3.8 Summary and conclusion

Three constructs of N-terminal metal binding domains of WLN, WLN5-6, WLN4-6 and WLN3-6, were characterized using mass spectrometry and dynamic light scattering. The structure and the stability were investigated via chemical unfolding using GuHCl and by increasing the temperature and monitoring by circular dichroism. The frictional ratio increased as more domains are added to the constructs, and so, the constructs are no longer spherical, which is expected due to the presence of a long linker and the large size of the constructs. Circular dichroism measurements of the WLN5-6 showed a two state unfolding profile while WLN4-6 and WLN3-6 exhibit three state models, Figure 3.25. Thermal denaturation of all constructs shows two states with a transition midpoint at temperature of 75°C or higher, Figure 3.26.

![Figure 3.25: Unfolding of WLN5-6, WLN4-6, and WLN3-6 by GuHCl](image)

The chemical unfolding results for WLN5-6, WLN4-6 and WLN3-6 show lack of cooperativity within domains. Since, adding domains did not affect the first transition in
all constructs. However, the second transition in WLN3-6 and WLN4-6 matches the transition for WLN4 alone. This means that WLN4 maintains its integrity in the constructs that were studied in this research. In a previous study, the unfolding of WLN1-2, WLN1-3, and WLN1-4\textsuperscript{22}, a lack of cooperativity was observed and WLN4 also maintains its integrity. These results suggest that WLN1-3 work as one unit and WLN5-6 work as another unit, while WLN4 operates separately. Therefore, we assume that WLN4 acts independently and can transfer copper to other domains as well as acquire copper from HAH1.

![Graph showing thermal unfolding](image)

**Figure 3.26: Thermal unfolding of WLN5-6, WLN4-6, and WLN3-6**

The pI for all constructs is around 5, which mean they are acidic and negatively charged at physiological pH. They all have elongated shape. The values of the biophysical properties studied for WLN5-6, WLN4-6, and WLN3-6 are summarized in Table 3.6.
Table 3.6: Physical properties of WLN5-6, WLN4-6 and WLN3-6

<table>
<thead>
<tr>
<th>Property</th>
<th>WLN5-6</th>
<th>WLN4-6</th>
<th>WLN3-6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mass spectroscopy</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mass</td>
<td>16042</td>
<td>29198</td>
<td>39848</td>
</tr>
<tr>
<td>Chemical unfolding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>[GuHCl]_{N→U} (M)</td>
<td>3.27 ± 0.25</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[GuHCl]_{N→I} (M)</td>
<td>-</td>
<td>3.0 ± 0.58</td>
<td>2.6 ± 0.51</td>
</tr>
<tr>
<td>[GuHCl]_{I→U} (M)</td>
<td>-</td>
<td>5.78 ± 2.6</td>
<td>4.8 ± 1.8</td>
</tr>
<tr>
<td>ΔG^{H_2O}_{N→U} (kJ/mol)</td>
<td>12.2 ± 0.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ΔG^{H_2O}_{N→I} (kJ/mol)</td>
<td></td>
<td>19.9 ± 2.8</td>
<td>20.4 ± 0.4</td>
</tr>
<tr>
<td>ΔG^{H_2O}_{I→U} (kJ/mol)</td>
<td></td>
<td>35.2 ± 11.3</td>
<td>24.9 ± 6.9</td>
</tr>
<tr>
<td>m_{N→U} (kJ/mol/M)</td>
<td>3.7 ± 0.19</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>m_{N→I} (kJ/mol/M)</td>
<td></td>
<td>6.8 ± 0.9</td>
<td>7.9 ± 1.6</td>
</tr>
<tr>
<td>m_{I→U} (kJ/mol/M)</td>
<td></td>
<td>6.1 ± 2.0</td>
<td>5.1 ± 1.4</td>
</tr>
<tr>
<td>IEF</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pI</td>
<td>4.90^{28}</td>
<td>4.86</td>
<td>4.96</td>
</tr>
<tr>
<td>Thermal Unfolding</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T_m</td>
<td>81.5 ± 2.5</td>
<td>74.4 ± 0.4</td>
<td>75.2 ± 0.3</td>
</tr>
<tr>
<td>ΔS</td>
<td>0.167 ± 0.004</td>
<td>0.449 ± 0.007</td>
<td>0.397 ± 0.008</td>
</tr>
<tr>
<td>ΔH</td>
<td>59 ± 1.6</td>
<td>156.2 ± 2.5</td>
<td>138.3 ± 2.87</td>
</tr>
<tr>
<td>ΔG @ 293 K</td>
<td>12.7 ± 3.0</td>
<td>24.6 ± 3.3</td>
<td>22 ± 3.8</td>
</tr>
<tr>
<td>DLS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>R_h (Å)</td>
<td>21.0 ± 0.6^{25}</td>
<td>35.25 ± 0.4</td>
<td>36.7 ± 0.12</td>
</tr>
<tr>
<td>Frictional ratio</td>
<td>1</td>
<td>1.41</td>
<td>1.63</td>
</tr>
<tr>
<td>Shape</td>
<td>Spherical</td>
<td>Non-spherical</td>
<td>Non-spherical</td>
</tr>
<tr>
<td>Percentage α helix (%)</td>
<td>34.9 ± 2.6</td>
<td>28.3 ± 1.2</td>
<td>20.9 ± 1.8</td>
</tr>
</tbody>
</table>
3.9 References


(3) Banci, L.; Bertini, I.; Del Conte, R.; D’Onofrio, M.; Rosato, A. Solution Structure and Backbone Dynamics of the Cu(I) and Apo Forms of the Second Metal-Binding Domain of the Menkes Protein ATP7A. *Biochemistry* 2004, 43, 3396–3403.


CHAPTER IV

WILSON DISEASE CAUSING MUTANT Y532H

4.1 Introduction

In the previous chapter, investigations focused on understanding the structure and stability of different domains of the Wilson disease protein. It is equally important to study how the disease causing mutations influence stability. More than 400 mutations are found in the ATP7B gene of Wilson disease patients. Around 28 of disease-causing missense mutations are within the six metal-binding domains. Interestingly, 18 are in domains 5 and 6. Several mutations in the metal binding domains cause the disease as they affect the interaction between ATP7B and metallochaperones HAH1 and COMMD1. For example; G85V, L492S, G591D, and A604P result in increased binding of ATP7B to COMMD1. A486S and Y532H mutations also have increased interaction slightly but it was not reproducible. On the other hand, a decrease in the interaction with HAH1 was observed. HAH1 is actually part of the copper metallation pathway and intracellular copper efflux.

The fact that Wilson protein has six similar metal binding domains has been the subject of many investigations. Based on solution structures of individual metal binding domains, they all have a similar ferrodoxin-like fold like that of the HAH1; however, genetic and biochemical data imply that the six domains are not functionally equivalent, and that they function to finetune and regulate the copper transport activity. As part of
our efforts to understand the role of the six metal binding domains in the N-terminus of Wilson protein, our laboratory studied interaction between a construct of WLN5-6 and the ATOX1 chaperone and two other domains, domains 2 and 4. Based on that study, a pathway was proposed for copper delivery and it was shown that ATOX1 forms a protein-protein adduct with domains 2 and 4 as observed via NMR. Mutants in domains 5 and 6, specifically L492S and Y532H, have shown a change in the level of interaction between the ATP7B and ATOX1. We assume that missense mutations in the N-terminus of Wilson protein destabilize protein folding structure and disrupt the network of atomic contacts between domains. So, it is important to understand the structural stability of the mutants.

Studying protein folding or structure is extremely complicated in eukaryotic cells. Therefore, protein folding has been widely studied in pure proteins in vitro using mainly biophysical techniques. These techniques include fluorescence resonance energy transfer or intrinsic fluorescence of tryptophan and tyrosine residues. Most investigations normally probe the tryptophan fluorescence. Unfortunately, WLN5-6 does not possess any tryptophan but it possesses three tyrosine residues, which are fluorescent but not as strongly fluorescent as tryptophan. In most studies, steady-state fluorescence of tryptophan or tyrosine is used which provides an ensemble average of interactions around the chromophore. On the other hand, time-resolved fluorescence measurements can provide specific information about the hydrogen bonding or other ligand-ligand interactions that play a crucial role in quenching the excited state lifetimes. Lifetimes with femtosecond resolution can interrogate the very fast interactions that are not observed with nanosecond time resolution. Femtosecond
time-resolved studies with intrinsic tryptophan fluorescence are difficult with high laser energy of UV excitation (i.e. 266 nm) which degrade the tryptophan easily by multiphoton ionization\textsuperscript{14}. On the other hand, WLN5-6 has three tyrosine residues and multiple tyrosine residues of a protein can serve as localized probes of structural changes during thermal and chemical unfolding through ultrafast time-resolved fluorescence spectroscopy. Our femtosecond time-resolved fluorescence measurements on tyrosine show little degradation during 266 nm laser irradiation. This system is unique to be probed with femtosecond measurements.

Two disease causing-mutants, V536A and L492S, were studied by another laboratory member\textsuperscript{15}. These mutants were incorporated in WLN5-6 construct. The mutants are less stable than the wild type based on the $\Delta G$ unfolding values. However, tyrosines in V536A exhibit a faster decay and sub picosecond lifetimes relative to wild type or other mutants. In this study, another important disease-causing mutant, Y532H, is investigated, Figure 4.1, and its unfolding was monitored by circular dichroism. Also, we applied tyrosine emission and time-resolved fluorescence spectroscopy to understand the change in the tertiary structure of this mutant.
Figure 4.1: Cartoon structure of WLN5-6 (PDB:2EW9), and WLN5-6 Y532H (Tyr48 replaced by His)

4.2 Experimental methods

4.2.1 Expression of Y532H

Site-Directed Mutagenesis was used to produce the three mutant’s plasmids of WLN5-6, V536A, L492S and Y532H\textsuperscript{15,16}. All primers for mutagenesis were designed
based on the QuickChange® II XL Site-Directed Mutagenesis kit manual. They were purchased from, Integrated DNA Technologies, Inc. (Coralville, Iowa). The sequences of the plasmids were confirmed by DNA sequencing, Retrogen Inc. (San Diego, CA). WLN5-6 and its mutant, Y532H, were transformed into *Escherichia Coli* strain Rosetta2 (DE3) cells. A single colony was grown in LB containing Kanamycin antibiotic (30 μg/L). When the OD$_{600}$ reached 0.6-0.8, cells were induced with 1 mM IPTG at 37 °C with shaking at 250 rpm for 2-3 h. Then cells were harvested by centrifugation. Cell pellets were extracted by the freeze-thaw method. Proteins were extracted with 20 mM MES buffer pH 6.3 containing 1 mM EDTA and 5 mM DTT. For purification, each protein was passed through a DEAE-Sepharose anion exchange column then gel filtration (Hiload 26/60 Superdex 75). SDS-PAGE confirmed the purity of the protein. Protein concentration was measured by the BCA assay.

### 4.2.2 Circular dichroism

Protein samples were unfolded using Guanidine Hydrochloride (GuHCl) at different concentrations ranging between 0-8 M. Each sample contains 5 mM Sodium Phosphate buffer and were equilibrated at room temperature for 2-3 hours. CD measurements in the UV (190-300 nm) region were recorded on a JASCO J-815. A quartz cuvette of path length 1 mm was used. Each spectrum was the average of three scans. Protein concentrations were in the range of 12 - 30 µM. After subtracting the appropriate blanks, molar ellipticities were calculated using equation (4.1):

$$[\theta] = \frac{\theta_{obs}}{10 \times c \times l} \quad (4.1)$$
$\theta_{\text{obs}}$ is the observed ellipticity in millidegrees, "c" is concentration of protein (mole per liter) and "l" is the path length in centimeters. Thermal unfolding were performed using 2-3 micromolar of the protein in a 1 cm path length cuvette, temperature were increased one degree per minute from 20˚C to 95˚C.

4.2.3 Fluorescence spectroscopy

Optical absorption measurements were carried using Shimadzu UV 2101 PC absorption spectrometer. Fluorescence measurements were carried out by using a F900 Edinburgh Spectrofluorimeter and a Hitachi F-2500 Spectrofluorimeter with a rectangular cell of 4 mm path length. Time-resolved fluorescence measurements were carried out using femtosecond fluorescence up-conversion using the third harmonic of Ti: Sapphire laser as an excitation source. Emission was recorded at 340 nm, the Exalite dye was used to optimize the fluorescence upconversion setup with a 266 nm excitation. Protein concentration for all the steady-state fluorescence measurements was 2 µM. The excitation wavelength of 270 nm was used while the emission spectra were recorded from 280 to 450 nm with 5- and 10-nm slit widths for excitation and emission, respectively.

4.2.4 Quantum yield

The fluorescence quantum yield ($\Phi_F$) is the ratio of photons absorbed to photons emitted through fluorescence. In other words, the quantum yield gives the probability of the excited state being deactivated by fluorescence rather than by another, non-radiative mechanism. The most reliable method for recording $\Phi_F$ is the comparative or relative
method of Williams et. al\textsuperscript{17}, wherein standard samples with known $\Phi_F$ values are used. Essentially, solutions of the standard and test samples with identical absorbance at the same excitation wavelength can be assumed to be absorbing the same number of photons. Hence, a simple ratio of the integrated fluorescence intensities of the two solutions (recorded under identical conditions) will yield the ratio of the quantum yield values. Since $\Phi_F$ for the standard sample is known, then the $\Phi_F$ for the test sample is easy to be calculated by equation (4.2):

$$\phi_F = \left(\phi_F\right)_S \frac{\int J(\nu)\,d\nu \,(J_a)_S \eta^2}{\int J_s(\nu)\,d\nu \,J_a \eta^2_s}$$  \hspace{1cm} (4.2)$$

$\int J(\nu)\,d\nu$ and $\int J_s(\nu)\,d\nu$ are the area under the fluorescence emission curve for the sample and the standard, respectively. $(J_a)_S$ and $J_a$ are the absorbance of the standard and sample, respectively. $\eta$ and $\eta_s$ are the refractive index of the solvent for the sample and the standard respectively. Here, the quantum yields were measured with respect to free tyrosine fluorescence.

4.3 Results and discussion

4.3.1 Circular dichroism

Mutations of Tyrosine 48 in WLN5-6 to Histidine (Y532H) were studied using circular dichroism (190-300 nm), and samples were unfolded with GuHCl and heat. The far UV CD scan can monitor $\alpha$-helix structure, Figure 4.2, since a canonical $\alpha$-helical protein like myoglobin has minima at 222 nm and 208 nm. The molar residue ellipticity expected for a protein of 150 amino acids is -39333.3 deg·cm$^2$·dmol$^{-1}$ based on the signal at 222 nm and equation (3.14). The molar residue ellipticity for Y532H at 222 is -9471.6.
therefore, the percentage of α-helix is 23.5±1 %. The wild type has α-helix percentage of 34.9±2.6 %. This decrease indicates disruption in the α-helix region in the mutant due to the substitution of a histidine residue instead of a tyrosine in position 48. The side chain of tyrosine, specifically the hydroxyl group, was involved in a network of contacts that maintains the helical structure at domain 57.16.

![Figure 4.2: CD scan for Y532H](image)

The CD signal was also monitored at 222 nm for the same protein after gradually increasing the temperature to 90°C. As shown in Figure 4.3, the signal changes little before 50°C then it starts to increase and levels off at 90°C. A sigmoidal fit for the data shows a T_m at around 79±2.4°C. The T_m value is slightly lower than the value obtained for the native protein, 81.5 ± 2.6°C, Figure 3.6. After incubating the protein at 90°C for 5 minutes, a CD scan was performed and compared to the scan for the same sample before unfolding, Figure 4.4. The signal at 222 nm increased noticeably, and the signal at 208 nm decreased and slightly shifted to lower wavelength. This could be due to increasing
amount of random coil structure. However, the protein was able to recover some of its original structure after cooling down to 20°C.

**Figure 4.3:** Thermal unfolding of Y532H monitored at 222 nm

**Figure 4.4:** CD scan for folded and unfolded sample of Y532H
Unfolding of Y532H using GuHCl is shown in Figure 4.5. The protein starts to unfold at a lower concentration of GuHCl, 1 M and then completely unfolds at around 5 M. [GuHCl] at the midpoint (C_m), is 2.8 ± 0.1 M. The free energy of unfolding was obtained using the linear extrapolation method, Figure 4.6. ΔG_H2O is 10.6 ±1.55 kJ/mol, and the m-value is 3.34 ± 0.37 kJ/mol/M. The m-value reflects the dependence of free energy on the concentration of denaturant and the area of the hydrophobic residues that are accessible to the solvent. Y532H has slightly higher values of m, ΔG_H2O, and C_m, compared to wild type WLN5-6.

Y532H has a higher free energy of thermal unfolding than WLN5-6, Figure 4.7, of 13.5±3.5 kJ/mol. If we consider the error of the free energy values, then the values for both thermal and chemical unfolding are not far from each other, energy values are summarized in Table 4.1.

![Molar Ellipticity vs [GuHCl] for Y532H](image)

**Figure 4.5: Unfolding of Y532H by GuHCl monitored at 222 nm**
Figure 4.6: Analysis of chemical unfolding of Y532H to calculate free energy in water

Figure 4.7: Thermal unfolding analysis to obtain free energy
Table 4.1: Free energies of unfolding WLN5-6 and Y532H obtained from chemical and thermal unfolding

<table>
<thead>
<tr>
<th></th>
<th>WLN5-6</th>
<th>Y532H</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_m$</td>
<td>3.30 ± 0.7</td>
<td>3.2 ± 0.6</td>
</tr>
<tr>
<td>$\Delta G_{H_2O}$</td>
<td>12.2 ± 0.7</td>
<td>10.6 ± 1.5</td>
</tr>
<tr>
<td>$T_m$</td>
<td>81.5 ± 2.7</td>
<td>79 ± 2.4</td>
</tr>
<tr>
<td>$\Delta G$</td>
<td>12.7 ± 3</td>
<td>13.5 ± 3.5</td>
</tr>
</tbody>
</table>

4.3.2 Fluorescence measurements

Tyrosine fluorescence of Y532H was monitored through an emission scan ($\lambda_{ex} = 275$ nm). 20 µM Y532H at different concentrations of GuHCl ranging between 0 – 7 M were prepared and equilibrated at room temperature for 2 hours. A CD scan for the samples was obtained from 300 to 190 nm, Figure 4.2. The Tyr fluorescence was monitored often at 270 nm and the emission from 280 – 450 nm was recorded, Figure 4.8.

![Emission scans of Tyrosine in Y532H at different [GuHCl]](image-url)

Figure 4.8: Emission scans of Tyrosine in Y532H at different [GuHCl]
The tyrosine emission at 310 nm increases while the protein unfolds, Figure 4.9. However, when the protein started to unfold, at 2.5 M GuHCl, the fluorescence suddenly decreases, levels off, then increases again up to 7 M GuHCl. In the folded state the fluorescence of the Tyrosine is quenched, presumably due to hydrogen bonding of the tyrosyl hydroxyl group.
Figure 4.10: Emission scans for Tyrosine in WLN5-6 (wild type)

Figure 4.11: Fluorescence counts at 310 nm for WLN5-6

In the wild type, shown in Figure 4.10 and Figure 4.11, the fluorescence increases wherein the Y532H fluorescence rate decreases as the protein unfolded.
4.3.3 Quantum yield

Tyrosine has a quantum yield of 0.14 and it was used as a standard to obtain the quantum yield for WLN5-6 and Y532H. Diluted samples of Tyrosine, WLN5-6, and Y532H, in 5 mM sodium phosphate buffer (pH 7.5), were prepared so the absorbance at 275 nm is equal to 0.1, Figure 4.12. The emission scans were recorded from 285-550 nm for all three samples, Figure 4.13. The value of the refractive index for all samples was assumed to equal the value for the buffer, 1.333. Y532H exhibits more fluorescence than the wild type. The fact that Y532H has a higher quantum yield (1.5 fold), Table 4.2, than the wild type suggests that Y532, in the wild type, acts as a self quencher for the protein, and once it replaced with Histidine, the quenching effect was removed. Probably the tyrosine residue in position 532 of WLN5-6 has the most important role among all three tyrosine residues for non-radiative processes and so its quenching is more significant.

<table>
<thead>
<tr>
<th></th>
<th>Area</th>
<th>$J\left(A_{275}\right)$</th>
<th>QY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tyr</td>
<td>101894.6</td>
<td>0.097</td>
<td>0.14</td>
</tr>
<tr>
<td>WLN5-6</td>
<td>23987.72</td>
<td>0.100</td>
<td>0.032</td>
</tr>
<tr>
<td>Y532H</td>
<td>36542.64</td>
<td>0.098</td>
<td>0.050</td>
</tr>
</tbody>
</table>
Figure 4.12: Absorbance scan for quantum yield measurements of WLN5-6 and Y532H

Figure 4.13: Fluorescence scans of WLN5-6 and Y532H for quantum yield measurements
4.3.4 Ultrafast time-resolved fluorescence spectroscopy

Time-resolved fluorescence decays were obtained using femtosecond fluorescence up-conversion after excitation at 266 nm and the emission monitored at 300 nm. One important factor to note in this study is that the protein is stable and all kinetics were reproducible even after long exposure to the laser. This is one of the first studies to use fluorescence upconversion to monitor the tyrosine fluorescence in a protein. This paves the way to use ultrafast time-resolved fluorescence spectroscopy to probe specific amino acid interactions in proteins. The fluorescence decay of the protein is faster than for the free tyrosine, and Y532H is faster than WLN5-6, Figure 4.14. The lifetime for free tyrosine is 1.4 ps, Y532H is less than 0.2 ps and for WLN5-6 it is 0.35 ps, Table 4.4. Since typically tyrosine residues in the proteins are buried and in a constrained environment, their decay will be slower. However, the tyrosines in Y532H have a fast decay and this could be due exposure of the two tyrosines to the solvent after the third tyrosine (Y532) have been removed.
Figure 4.14: Fluorescence decay of WLN5-6, Y532H, and Tyrosine

4.4 Comparison of Y532H to other mutants V536A, and L492S

Several other mutants of Wilson disease were studied by Marzijaran\textsuperscript{15}, and these mutants were incorporated in WLN5-6 using site directed mutagenesis. These mutants are V536A and L492S (V52A and L8S in WLN5-6), Figure 4.15. They were unfolded using GuHCl, Figure 4.16, and the $\Delta G_{H_2O}$ for all these mutants and the wild type were obtained as described in section 3.3 and are summarized in Table 4.3.
Figure 4.15: Cartoon structure of WLN5-6 (PDB:2EW9) showing the position of V536 and L492
Figure 4.16: Unfolding of WLN5-6, Y532H, L492S, and V536A by GuHCl monitored at 222nm

V536A has a transition midpoint of 3.0 M GuHCl which is less than the wild type of 3.3 M. The $\Delta G_{H2O}$ and m values show that V536A secondary structure is less stable than wild type. The mutation is on the surface of the protein and Alanine is less hydrophobic residue than Valine, so the mutation affects the Van der Waals interactions. The position of the mutation is in the loop between $\beta3$ and $\alpha2$ of domain 5, Table 4.3. Y532H has almost the same unfolding profile as the wild type. The $\Delta G_{H2O}$ value is lower than the wild type but higher than L492S and V536A, perhaps because the position of the mutation in Y532H is in interface of domains 5 and 6, and not really in the hydrophobic core.

L492S has a lower stability based on $\Delta G_{H2O}$ value. It was predicted because of the replacement of the hydrophobic leucine residue with highly hydrophilic residue serine. This mutation definitely alters the stable interaction between the hydrophobic residues.
and so lowers the helicity of the mutant as compared to the wild type. L492S has the lowest helicity among the mutants, Figure 4.17.

![Figure 4.17: CD scan of WLN5-6, Y532H, V536A, and L492S](image)

The CD data shows similar unfolding behavior for L492S and Y532H, compared to V536A. However, all the mutations have an effect on the structure and on the conformational stability of the protein. The mutations lowered the free energy value of unfolding and the transition midpoint which means the mutants are less stable than the wild type. Dynamic light scattering experiments for V536A and Y532H show almost the same hydrodynamic radii as WLN5-6 (data not shown)\textsuperscript{15,16}. We expect that the mutations do not alter the tertiary structure, but do affect the ionic and hydrogen bonds interaction in the mutants, and lead to a molten globule structure.
Table 4.3: Free energy, m values, and $C_m$ calculated from CD measurements at 222 nm

<table>
<thead>
<tr>
<th>Sample</th>
<th>$\Delta G_{H2O}$ (kJ/mol)</th>
<th>$m$ (kJ/mol/M)</th>
<th>$C_m$ (M)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN5-6</td>
<td>12.2 ± 0.7</td>
<td>3.7 ± 0.2</td>
<td>3.3 ± 0.3</td>
</tr>
<tr>
<td>WLN5-6 V536A</td>
<td>8.6 ± 0.7</td>
<td>3.4 ± 0.2</td>
<td>2.5 ± 0.3</td>
</tr>
<tr>
<td>WLN5-6 L492S</td>
<td>7.1 ± 0.50</td>
<td>2.8 ± 0.2</td>
<td>2.6 ± 0.2</td>
</tr>
<tr>
<td>WLN5-6 Y532H</td>
<td>10.6 ± 1.5</td>
<td>3.34 ± 0.4</td>
<td>3.2 ± 0.6</td>
</tr>
</tbody>
</table>

Quantum yield and time-resolved fluorescence spectroscopy measurements were performed for all the mutants. In contrast to Y532H, the L492S quantum yield is lower than wild type. Therefore, this mutation lets Y532 have a highly quenching effect on the fluorescence of other tyrosines and increases the energy transfer to Y532. In the case of the tyrosine lifetime, tyrosines in WLN5-6 and mutant V536A exhibit a difference in the magnitude of the picosecond tyrosine lifetime, Figure 4.18 and Table 4.4. This is probably due to the fact that the movement and rotation of tyrosine residues are limited and they are buried in WLN5-6. This means that the fluorescence energy is being utilized by the mutant in different processes. Interestingly, the decay for L492S is similar to the wild type, since Y532 is still present and the tyrosine interactions are not altered.

Table 4.4: Lifetimes of WLN5-6, mutants, and free tyrosine in 50 mM phosphate buffer pH 7.0, at room temperature

<table>
<thead>
<tr>
<th>Sample</th>
<th>Lifetime (population %)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tyrosine</td>
<td>1.4 ps (46.1 %), &gt;10 ps (53.9 %)</td>
</tr>
<tr>
<td>WLN5-6</td>
<td>350 fs (76.7 %), &gt;10 ps (23.3 %)</td>
</tr>
<tr>
<td>WLN5-6 V536A</td>
<td>360 fs (97.7 %), 1.5 ps (1.9 %), &gt;10 ps (0.4 %)</td>
</tr>
<tr>
<td>WLN5-6 L492S</td>
<td>350 fs (77.5 %), &gt;10 ps (22.5 %)</td>
</tr>
<tr>
<td>WLN5-6 Y532H</td>
<td>&lt;200 fs (95%), 1.5 ps (1%), &gt;10 ps (4%)</td>
</tr>
</tbody>
</table>
This quenching is not due to exposure by the solvent, otherwise the lifetime will be similar to the free tyrosine. Tyrosine residue lifetimes can be quenched via hydrogen exchange and non-radiative pathways. Since the tyrosine residues in L492S mutant have almost the same lifetime as the wild type and Y532H, then it means that tyrosine Y532 alone is not responsible for the emission of the protein. However, it has a role in the quenching in the V536A mutant. Mutation of Valine to Alanine perturbs the structure and allows more interaction and energy transfer to Y532.

![Fluorescence decay of Tyrosine, WLN5-6, and the mutants (Y532H, L492S, and V536A)](image)

**Figure 4.18:** Fluorescence decay of Tyrosine, WLN5-6, and the mutants (Y532H, L492S, and V536A)

### 4.4.1 Mechanism of ultrafast fluorescence quenching in disease causing mutations

In the case of wild-type protein, the three tyrosine residues form a unique triad with interactions between them, Figure 4.19. This leads to energy transfer between them
and less non-radiative relaxation. It has to be mentioned that the hydrogen bonding interactions between the tyrosines in WLN5-6 are also responsible for some amount of quenching of the fluorescence that was reflected in their faster lifetimes compared to that of free tyrosine. V536A influences the interactions around Y532, leading to ultrafast excited state decay. This triad of tyrosines is very important for the stability of the protein and ultrafast fluorescence measurements allowed us to begin understand their important role.

![Figure 4.19: Cartoon structure of WLN5-6 showing the hydrogen bonding between Y48 (Y532), Y75 (Y559), and Y110 (Y594)]](image)

4.5 Summary and conclusion

The mutation of Y48 in WLN5-6 (Y532H) has no significant effect in the stability of the protein as there is no huge difference in the free energy of unfolding. By investigating the NMR structure of WLN5-6 (PDB:2EW9), we noticed that all the
tyrosines in WLN5-6 (in positions 48, 75, and 110) are connected together through hydrogen bonds directly or indirectly, Figure 4.19. Y48 is bonded to the backbone/peptide bond Y75 and Y48 has 51 NOEs (Nuclear Overhauser Effects) based on the NMR structure of WLN5-6. On the other hand, Y75 has a hydrogen bond with N81 that connects with K123, which also has a H-bond with Y110. Y75 and Y110 are responsible for the fluorescence of the protein while Y48 is quenching the fluorescence, when replaced by Histidine. Mutant V536A and Y532H affected the rates of the fluorescence decay which means both perturb the influence of Y48. Present results have conclusively demonstrated the importance of a tyrosine triad in WLN5-6. We have provided a new way to monitor specific interactions in proteins with ultrafast time-resolved spectroscopy.

4.6 References


CHAPTER V
GREEN FLUORESCENT PROTEIN UNFOLDING

5.1 Introduction

In this work, protein unfolding is being used as a tool to estimate and compare the stability of the N-terminal metal binding domains in Wilson’s protein. We used CD (Chapter 3), and intrinsic Tyrosine fluorescence (Chapter 4) as tools to monitor the unfolding. In this chapter, we propose a new tool to monitor protein unfolding, based on the change in the two-photon absorption (2PA) cross-section during unfolding. To test our hypothesis, we used protein with a well-known structure and a chromophore buried in the core of the protein. Therefore, we chose green fluorescent protein to test our model.

Green fluorescent protein was discovered in 1962 by Osamu Shimomura\textsuperscript{1} accidentally, when he was isolating a blue luminescent protein from the jellyfish \textit{Aequorea victoria} that he named Aequorin. The jellyfish glowed green instead of blue, due to the absorbance of the protein’s blue light by a second jellyfish protein called later green fluorescent protein (GFP). In 1979 Shimomura showed that GFP contained a special chromophore that absorbs and emits light\textsuperscript{2}. The chromophore is a p-hydroxybenzylideneimidazolinone formed from residues 65–67, which are Ser-Tyr-Gly in the native protein. The formation mechanism is divided into three steps, Figure 5.1; the first step is the folding of the GFP protein. The second step is rapid cyclization and dehydration. The last step, which is the rate determining step, is the oxidation by oxygen
of the Tyrosine side chain\textsuperscript{3,4}. This formation process is auto-catalytic; no enzyme or cofactor is required. However, it is thermo sensitive above 30\textdegree C. Holo GFP is thermostable.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{gfp_chromophore.png}
\caption{Formation of GFP chromophore}
\end{figure}

The structure of GFP consists of eleven beta strands that form a beta-barrel and an \(\alpha\)-helix that runs through the center of the barrel. The chromophore is also located in the center of the \(\beta\)-barrel. The structure classification of the GFP is \(\beta\)-can\textsuperscript{5,6}, Figure 5.2. Folding of the protein into a \(\beta\)-can structure protects the chromophore from solvent and solvent-quenching effects. This also, prevents other proteins from accessing the chromophore region and interferring with spontanuous chromophore formation. The GFP structure highlights the importance of the entire protein in the function of the chromophore, as any deletion in the amino acid sequence leads to diruption of chromophore formation.
Figure 5.2: Cartoon structure of GFP, (PDB: 1GFL)

GFP consists of 238 amino acids and its wild type excitation peaks are at 395 nm and 475 nm with extinction coefficients of roughly 30,000 M$^{-1}$ cm$^{-1}$ and 7,000 M$^{-1}$ cm$^{-1}$, respectively. Excitation at 395 nm leads to a decrease over time of the 395 nm excitation peak and a reciprocal increase in the 475 nm excitation band. This may be due to the photoisomerization effect with irradiation of GFP by UV light. GFP has an emission peak at 508 nm.

GFP is the first known example of the Förster cycle within a protein core in which the fluorophore in the GFP protein absorbs a photon at 395 nm which makes the tyrosine residue (Tyr66) more acidic and then loses a proton. The produced anionic fluorophore will absorb a photon at 475 nm and finally emit a photon at 508 nm and accept a proton to return to its original state. The physical studies of native GFP reveals that it is very resistant to denaturation requiring 6 M guanidine hydrochloride at 90°C or a pH of <4.0 or >12.0. The biological role of GFP is to transduce the blue
chemiluminescence of the protein aequorin into green fluorescent light by energy transfer. Since fusion of GFP to any protein does not alter the function or location of that protein, the gene has become a useful tool for making chimeric proteins of GFP linked to other proteins where it functions as a fluorescent protein tag\textsuperscript{9,10}. Another important character of GFP is that it tolerates N- and C-terminal fusion to a broad variety of proteins. It has been expressed in many species: bacteria\textsuperscript{12}, yeast\textsuperscript{13}, slime mold\textsuperscript{11}, plants\textsuperscript{14}, drosophila\textsuperscript{15}, zebrafish\textsuperscript{16}, and in mammalian cells\textsuperscript{17}. GFP works as a noninvasive fluorescent marker in living cells which allows wide range of applications where it may function as a cell lineage tracer\textsuperscript{18}, reporter of gene expression\textsuperscript{12}, or as a measure of protein-protein interactions\textsuperscript{19}.

Folding of the protein into the β-can structure protects the chromophore from being solvent accessible and so eliminates solvent-quenching effects. GFP chromophore fluorescence is quenched upon protein denaturation, presumably by exposing the chromophore to the aqueous solvent environment\textsuperscript{20}. Therefore, the green fluorescence is used as an insightful probe of the folding of the protein. GFP needs to fold efficiently in order to function and inefficient folding is known to limit its use in some applications\textsuperscript{21}.

Much research has been conducted on the folding of this protein, either in vitro or in vivo. A model in which GFP folds through several intermediates states has been proposed through the acid-denatured state\textsuperscript{22}. Some studies monitored the chromophore fluorescence to measure the folding and unfolding kinetics while others use NMR to analyse structural dynamics\textsuperscript{23–25}. Static and kinetic FTIR, UV/vis and fluorescence spectroscopy were also used for pressure-induced refolding and unfolding monitoring\textsuperscript{26}. Single-molecule force spectroscopy was used to drive GFP molecule from the native state
to the unfolded state and quantitative analysis of force distributions along with the lifetimes has led to a detailed picture of a complex mechanical unfolding. In a recent study, fluorescent proteins have shown to enhance 2PA cross-section, which was attributed to the local electric field around the chromophore. When a protein unfolds, the electric field around the chromophore changes as the chromophore experiences a new environment. So, we would like to use this change to monitor the unfolding of proteins. In this investigation, we have carried out measurements to test this hypothesis and develop a new technique to monitor the proteins unfolding. With the development of this technique, a way to image the folded and unfolded forms of proteins is quite feasible. In addition to 2PA cross-sections, we will use ultrafast fluorescence lifetimes to follow the unfolding of GFP.

5.2 Experimental methods

5.2.1 Plasmid generation

GFP gene, from *Aequorea coerulescens*, was amplified using KOD Hot Start DNA Polymerase (Novagen ®) By using a pAcGFP1-N1 vector (clontech) as a template and appropriate primers (with TEV site in the forward primer) which are compatible with pET 32 Xa/LIC vector, Table 2.3. The amplification reaction and the PCR parameters are shown in Table 5.1 and Table 5.2.
Table 5.1: KOD hot start DNA polymerase reaction

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- Template cDNA, pAcGFP1-N1 (10 ng/μl in TE buffer)</td>
<td>1 μL</td>
</tr>
<tr>
<td>2- Forward primer 1 (10 μM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>3- Reverse primer 1 (10 μM)</td>
<td>1 μL</td>
</tr>
<tr>
<td>4- Buffer for KOD (10 X)</td>
<td>5 μL</td>
</tr>
<tr>
<td>5- dNTPs (2 mM)</td>
<td>5 μL</td>
</tr>
<tr>
<td>6- MgSO4 (25 mM)</td>
<td>3 μL</td>
</tr>
<tr>
<td>7- PCR grade water</td>
<td>33 μL</td>
</tr>
<tr>
<td>8- KOD hot Start DNA polymerase (1U/μL)</td>
<td>1 μL</td>
</tr>
</tbody>
</table>

Table 5.2: PCR parameters for amplifying GFP gene

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1- 95 °C</td>
<td>2 minutes</td>
</tr>
<tr>
<td>2- 95 °C</td>
<td>20 seconds</td>
</tr>
<tr>
<td>3- 70 °C</td>
<td>24 seconds</td>
</tr>
<tr>
<td>4- Steps 2-3</td>
<td>30 cycles</td>
</tr>
<tr>
<td>5- 4 °C</td>
<td>infinity</td>
</tr>
</tbody>
</table>

0.8 % Agarose gel was used to verify the amplification of the desired gene. The gene is 723 bp. The product was purified using Qiagen PCR purification kit and quantified using the intensity of the bands through exposure to UV light in an ethidium bromide stained gel. 0.2 pmol of purified insert was treated with T4 DNA polymerase to generate compatible overhangs. The following component was assembled in a small PCR tube, Table 5.3.
### Table 5.3: T4 DNA polymerase reaction for GFP

<table>
<thead>
<tr>
<th>Component</th>
<th>Volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purified PCR product</td>
<td>1 μl</td>
</tr>
<tr>
<td>10X T4 DNA polymerase buffer</td>
<td>1 μl</td>
</tr>
<tr>
<td>25 mM dGTP</td>
<td>1 μl</td>
</tr>
<tr>
<td>100 mM DTT</td>
<td>0.5 μl</td>
</tr>
<tr>
<td>Nuclease-free water</td>
<td>6.3 μl</td>
</tr>
<tr>
<td>2.5 U/μl T4 DNA polymerase</td>
<td>0.2 μl</td>
</tr>
</tbody>
</table>

The reaction was incubated at 22°C for 30 minutes. Then the enzyme was deactivated by incubation at 75°C for 20 minutes. The reaction tube was then chilled on ice and 0.5 μl of vector was added to 2 μl of the treated insert, and incubated at 22°C for 5 minutes prior to the addition of 0.5 μl of EDTA solution (25 mM). The produced plasmid was transformed into Novablue Giga Singles cells. The plasmid was purified and digested with BamH1 and BglII to assure the ligation of the insert to the vector. Three plasmids were sent for sequencing to Retrogen Company, CA. using S-tag, T7 and T7 term primers. The 6700 bp (approximately) plasmid is called pET32\textsubscript{TEV}LIC-GFP.

#### 5.2.2 Expression and protein purification

pET32\textsubscript{TEV}LIC-GFP was transformed into of BL21(DE3) cells. A single colony was removed from a previously transformed agar plate using an inoculation loop. The loop was inserted and stirred to release the cells into a 50 mL culture tube containing 5 mL LB media and 100 mg/L ampicillin. The tube was then placed in the shaker to
incubate for 4 hours at 37°C at a speed of 220-230 rpm. The culture tube then was transferred to a 500 mL flask containing 100 mL LB media and 100 mg/ L ampicillin. The flask was incubated in the shaker at the same temperature and speed previously used (vide supra). Two hours later, the LB media in the flask was transferred to one liter LB media containing 100 mg/L ampicillin. The flask was placed in the shaker at 37°C at a speed of 220-230 rpm. The optical density (OD$_{600}$) was checked at regular intervals to evaluate cell growth using a Beckman/Coulter DU 730 Life Science UV/Vis Spectrophotometer. When the OD$_{600}$ reached approximately 0.6-0.8, cells were induced with isopropyl β-D-1-thiogalactopyranoside (IPTG) to 1 mM. Immediately after adding IPTG, the temperature of the shaker was lowered to 28-30°C. Four hours later, cells were collected using a Sorvall RC-5B superspeed centrifuge (5,000 rpm) for 20 minutes and stored at -20°C for later purification.

The expressed protein contains 6x-His-tag so it was purified using nickel-sepharose affinity chromatography and gel filtration chromatography. Five grams of cell pellet was frozen in liquid nitrogen for two minutes and then thawed under cold water for 30 minutes. The freeze/thaw cycle was repeated 5 times. Then the pellet was re-suspended in 25 mL BugBuster® (Novagen) extraction reagent. A few microliters of (25 KU) benzonase were added to the extract to decrease cell lysate viscosity by digesting chromosomal DNA. The solution was centrifuged at 15,000 rpm for 20 minutes (SS-34 rotor). The supernatant was loaded into His Trap FF 16/10 column in order to start protein binding. The column was attached to an ATKA FPLC system. The column was washed with 140 mL (7 CV) equilibrium buffer contains 50 mM Potassium Phosphate, 500 mM Sodium Chloride, and 30 mM Imidazole at pH of 7.5. The protein has a visible
green color which makes it easy to determine the binding of the protein to the column. The protein was eluted with buffer containing 50 mM potassium phosphate, 500 mM Sodium Chloride, and 300 mM Imidazole, pH 7.5. The protein was further purified using gel filtration chromatography. The eluted protein was concentrated by an Amicon ultrafiltration device with a 3000 Da MWCO membrane at 4°C to reduce the volume of pooled fractions to 6 mL.

Concentrated solution was injected into a HiLoad 16/26 Superdex 75 prepgrade column using the FPLC system. Prior to injection, the column was equilibrated with 640 mL (2 CV) of 50 mM MOPS (3-(N-morpholino)propanesulfonic acid), 200 mM sodium chloride, 10 mM DTT, pH 7.2. The protein was eluted as a sharp peak at 165 mL after injection; the total CV was 320 mL. Protein containing fractions were concentrated to 7 mL using an Amicon device (vide supra). The purified GFP fusion solution was injected onto a desalting column to change the buffer into the buffer for TEV protease proteolysis. The TEV protease buffer contains 50 mM Tris-HCl, 0.5 mM EDTA and 1mM DTT, pH 8.0 (or. The protein concentration was determined using the BCA assay. TEV protease was added to the protein solution in the TEV protease buffer as a ratio of 1:80, TEV protease:protein. After addition of TEV protease the solution was incubated at room temperature overnight.

Purification of the digested GFP-fusion protein was completed using HisTrap, and Superdex 75 gel filtration fitted to FPLC (Amersham Biosciences, Piscataway, NJ). To separate the GFP from the His-tag, the digested GFP solution was loaded into a clean HisTrap column and then GFP was eluted by using the equilibration buffer that contains 50 mM Potassium Phosphate, 500 mM Sodium Chloride, and 30 mM Imidazole, pH 7.5.
Fractions containing GFP were collected, concentrated, and further purified by gel filtration with Superdex 75 column as before. The protein eluted at around 200 mL after injection in a CV of 320 ml. The fractions containing the protein were collected, concentrated, and desalted against 50 mM Potassium Phosphate buffer, pH 7.5. The protein was quantified using the BCA assay. By the end of the purification, 10 mg of the protein was obtained, then stored in phosphate buffer at -20°C. The extinction coefficient of the protein in water at 280 nm is equal to 23505 M⁻¹cm⁻¹ and this value was used for assaying the protein. The yield of the protein is 2 mg per gram pellet, and approximately 10 mg per one liter of culture. The protein purity of GFP was identified by SDS-PAGE, Figure 5.3 and mass spectrometry, Figure 5.4. The induction for protein expression was at 28-30°C. The yield could be potentially increased if the induction temperature was at room temperature instead of 28-30°C. However, the formation of the chromophore depends on the temperature of the induction.
5.2.3 Unfolding of GFP

A solution consists of 16 μM GFP protein mixed with varying amounts of GuHCl range from 0 M (control) to 3.0 M. All samples were equilibrated for 48 hours at 4°C prior to the spectroscopic measurements. Steady state fluorescence and circular dichroism were used to follow the denaturation of the GFP. Data points were acquired at room temperature using the same protein concentration in an appropriate GuHCl concentration.

5.2.4 Fluorescence spectroscopy

Optical absorption measurements were performed on a Shimadzo UV 2101 PC spectrometer. Fluorescence measurements were carried out by using F900 Edinburgh
Spectrofluorimeter and a Hitachi F-2500 Spectrofluorimeter with a rectangular cell of 4 mm path length. Time-resolved fluorescence and fluorescence anisotropy of unfolded protein samples and the control samples were performed using femtosecond fluorescence up-conversion. Polarization of the excitation beam for the anisotropy measurements was controlled using a Berek compensator and the sample is continuously rotated in a cell with 1 mm thickness. The polarizer was alternated between parallel and perpendicular orientations to the excitation polarization, with a dwell time of 30 s in each position to correct for excitation intensity drifts. The emerging light was imaged onto a microchannel plate photomultiplier detector (MCP-PM; Hamamatsu R1712U-03). Fluorescence decays were collected using standard TCSPC electronics in 1K channels at 0.0675 ns/ch. Instrument response function (IRF) was measured using Raman scattering from water. The excitation average power was around 0.9 ± 0.005 mW. 2PA cross-sections (δ) measurements were performed using a method employed by Rebane et al\textsuperscript{30}. A Coumarin 485 solution in methanol was used as reference. A neutral density filter was used to change the power as the power dependence is measured.

5.3 Results and discussion

5.3.1 Mass spectrometry

The GFP protein produced is a mutant that differs from the native protein by having a Valine residue in the second position at the N-terminus. Also, it has an extra Glycine residue due to TEV cleavage at position -1. It consists of 240 amino acids, and its sequence is shown below:
The average mass of the protein calculated by DNA\textsuperscript{©} Strider Software is 26931 Da. When chromophore forms, after dehydration and oxidation, the protein loses a mass of 18 Da. So, the expected mass for GFP is 26912 Da. The mass spectrum shown in Figure 5.4 shows a major component of a mass of 26904 Da.

**GFP sequence**

GMVSKGAEFL TGIVPILIEL NGDVNGHKFS VSGEGEGDAT YGKLTLLKFC
TTGKLVPWP TLVTTL\textbf{SYG} VQCFSDYPDHM KQHDFFKSAM PEGYIQERTI
FFEDDGNYKS RAEVKFEGDT LVRRIELTGT DFKEDGNILG NKMEXTYNNAH
NVYIMTDKAK NGIKVNFIR HNIEDGSVQI ADHYQQNTPN GDGVLLPDPN
HYLSTQSALS KDPNEKRDHM IYFGFVTAAA ITHGMDELYK

\textbf{SYG}: Amino acids which form the chromophore
Figure 5.4: ESI-IT/MS data showing the molecular weight of GFP determined by MagTran software
5.3.2 Steady state optical properties

The absorbance and the emission scans of a pure diluted sample (20 μM) of GFP are shown in Figure 5.5. The absorption spectrum shows two maxima, at 280 nm and at 475 nm. The absorbance at 280 nm is due to the aromatic amino acids. GFP has 12 tyrosines, 13 phenylalanines, and one tryptophan residue. The absorbance at 475 is due to the chromophore at its anionic form. The emission due to chromophore is observed at 505 nm when the protein excited at 470 nm. Note that the biophysical parameters of this protein from *Aequorea coerulescens* are not identical to the original GFP isolated from *Aequorea victoria*

Figure 5.5: Absorbance scan and emission scan of GFP, Excitation wavelength is 470 nm
When the protein unfolds in guanidine hydrochloride, another absorbance peak appears at 380 nm while the absorbance at 475 nm decreases until the protein totally unfolds in 3.0 M GuHCl. This new peak is due to the neutral form of the chromophore, so the chromophore converts from its anionic form to its neutral form due to exposure to the solvent, Figure 5.6. However, the absorption band at 380 nm grows in is slower than the decrease of the 475 nm band, which suggests that the non-fluorescent state is not formed directly from the fluorescent state\textsuperscript{20}. 1.6 M GuHCl is the concentration where 50% of the protein is unfolded and so half of the chromophore is on its neutral form and half is on its anionic form. At that concentration of GuHCl, absorbance value at 475 nm is 1.5 fold more than the absorbance at 380 nm.

![Figure 5.6: Absorbance scan of GFP at different [GuHCl] and the inset represents the absorbance for each sample at 480 nm and 380 nm](image)

The neutral form of the chromophore does not fluoresce, and this is observed when the fluorescence decreases as the protein is denatured; in other words, percentage of
the neutral form of the chromophore decreases, Figure 5.7. The fact that denatured GFP is not fluorescent and its absorption spectrum is significantly different from native GFP implies that noncovalent interactions of the chromophore with its local environment have a great influence on the spectral characteristics, and that fluorescence is mediated by amino acids close to the chromophore in the tertiary structure of GFP.

![Figure 5.7: Emission scan of GFP at different [GuHCl], the inset showing the fluorescence counts for each sample at 507 nm](image)

The quantum efficiency of the chromophore does not change significantly until the protein totally unfolds, at 2.4 M GuHCl, Figure 5.8. The quantum efficiency was calculated according to the following equation (5.1):

\[
\text{Relative Quantum efficiency} = \frac{F_s}{A_s} * \frac{A_0}{F_0}
\]  

(5.1)

F_s and A_s are the fluorescence and the absorption of a sample at certain concentration of GuHCl, and F_0 and A_0 are the fluorescence and the absorption of protein without GuHCl.
5.3.3 Circular dichroism

GFP consists of 11 β-sheets (PDB ID: 1GFL), therefore its CD scan shows a minima at 215 nm as expected, Figure 5.9.
A CD scan was also obtained for each GFP sample prepared at various concentrations of GuHCl. The GFP concentration was maintained at 32 µM for all samples, Figure 5.10.

![CD Scan of GFP at different [GuHCl], inset shows the molar ellipticity at 215 nm](image)

**Figure 5.10:** CD scan of GFP at different [GuHCl], inset shows the molar ellipticity at 215 nm

### 5.3.4 Fluorescence lifetime

Time-correlated single photon counting (TCSPC) method was used to measure the fluorescence lifetime\(^1\). In this method the sample was excited at 370 nm using the laser source (section 2.5.2). The sample was excited periodically and the fluorescence decay profile was reconstructed from multiple single photon events collected over many cycles. The emission scan was recorded at 510 nm. The instrument response fluorescence (IRF) was obtained using aluminum foil to insure total scattering of the light. The fluorescence decay was added to the IRF and the deconvolution decay fitted using single, bi-, or tri-exponential fitting, as needed, according to equation (5.2).
The fluorescence lifetime was fitted to a single exponential for samples of 0 M GuHCl through 1.2 m GuHCl. At higher GuHCl concentration, the fitting was multiexponential and the average lifetime was calculated according to equation (5.3)

\[
\langle \tau_{av} \rangle = \frac{a_0 + a_1 \tau_1 + a_2 \tau_2 + a_3 \tau_3}{a_0 + a_1 + a_2 + a_3} \quad (5.3)
\]

Where \( \tau_1, \tau_2, \text{ and } \tau_3 \) are two decay lifetimes while \( a_0, a_1, a_2 \text{ and } a_3 \) are corresponding amplitudes.

The fluorescence lifetime of GFP is measured as 3.3 ± 0.1 ns, Figure 5.11. During denaturation, the fluorescence lifetime did not change significantly at lower concentration of GuHCl or at higher concentration of GuHCl. However, the fluorescence intensity decreases at higher concentrations. A constant fluorescence lifetime with decreasing fluorescence intensity implies that the number of fluorescent protein molecules decreases during denaturation but not the characteristic of the fluorescent state. Between 1 M and 1.6 M GuHCl, the lifetime decreases drastically due to the exposure of the chromophore to the solvent. Conformational changes of the amino acids far from the chromophore did not affect the fluorescence intensity since they do not make the chromophore solvent accessible.
5.3.5 Fluorescence upconversion

Time-resolved fluorescence and fluorescence anisotropy of GFP were studied using femtosecond fluorescence upconversion. The upconversion system used in our experiments was obtained from the CDP system, FOG100, in the laboratory of Dr. Ramakrishna Guda at Western Michigan University. The system uses frequency doubled light from a mode-locked broad band Ti-sapphire laser (Spectra Physics, Tsunami, 710-920 nm). Samples of about 1.0 absorbance unit are continuously rotated in a rotating cell which is 1 mm thick to avoid the degradation of the sample. Horizontally polarized fluorescence emitted from the sample was upconverted in a nonlinear crystal borate using a beam pump at 800 nm, which first passed through a variable delay line, Figure 2.16.

Time-resolved fluorescence anisotropy $r(t)$ measurement was carried out with parallel and perpendicular excitations whose polarization is changed with a Berek compensator. Fluorescence anisotropy is calculated from traces obtained after parallel
(I\textsubscript{par}) and perpendicular (I\textsubscript{per}) excitation with the following equation, based on the ratio of difference between parallel and perpendicular polarized emissions over the magic-angle fluorescence, equations (5.4) and (5.5).

\[
 r(t) = \frac{I\textsubscript{par}(t) - G I\textsubscript{per}(t)}{I\textsubscript{par}(t) + 2GI\textsubscript{per}(t)} \quad (5.4)
\]

\[
 Magic \ angle = \frac{I\textsubscript{par}(t) + G I\textsubscript{per}(t)}{3} \quad (5.5)
\]

The G factor accounts for the differences in sensitivities for the detection of emission in perpendicular and parallel polarized configuration.

The fluorescence emission decay indicates the depolarization of the chromophore. The main reason for the depolarization is molecular rotation or energy transfer to another molecule with a different orientation\textsuperscript{32}. The chromophore absorbs the light that is parallel to its absorption dipole. When polarized light is used to excite a sample, only a subset of molecules will be excited. Brownian motion causes this subset to become disordered (partially polarized). With sufficient time the fluorescence will arise from randomly oriented molecules. Monitoring both parallel and perpendicular planes of polarization enables us to follow this path from order to disorder that caused by Brownian motion\textsuperscript{32}.

Fluorescence anisotropy is useful in obtaining information regarding molecular size and mobility of the chromophore, because the molecular motion depends on molecular confinement, and the size of the molecule. An exponential decay of fluorescence intensity at 510 nm is observed in the investigated concentration range of GuHCl. As shown in Figure 5.12, the fluorescence intensity decreases faster at higher concentrations of GuHCl. This decrease is due to the change in local environment when the protein is unfolded. Although the chromophore is covalently bonded to GFP and
located at the center of the barrel, unfolding by 1.4 M GuHCl significantly changes the local environment around the chromophore and makes it accessible to the solvent. At higher concentrations of GuHCl, the chromophore becomes fully exposed to solvent and the anisotropy decreases drastically and becomes independent of GuHCl concentration.

Figure 5.12: Fluorescence decay traces as a function of increasing Guanidine hydrochloride. Also, shown is the average lifetime with increasing GuHCl following the unfolding transitions

Figure 5.13: Anisotropy values of GFP at various [GuHCl]
The anisotropy value calculated for the folded GFP was about 0.40, Figure 5.13, and did not significantly change during unfolding, as expected, since the chromophore is bound to the backbone and no energy transfer to, or from other ligands is possible.

5.3.6 Two-photon absorption

The power dependence of the fluorescence detected at 850 nm excitation was determined from a log-log plot of the fluorescence signal at 510 nm versus incident peak power, shown in Figure 5.14. The induced fluorescence obeyed power-squared intensity dependence as indicated by the measured slope of 1.98, thereby confirming the existence of two-photon absorption. GFP excitation at 800 nm yielded the fluorescence spectra displayed in Figure 5.15. The characteristic GFP bands had a peak at 507 nm, and were almost indistinguishable from those observed for one photon fluorescence at 400 nm excitation. The 2PA at 800 nm yielded more fluorescence relative to excitation at 400 nm in the range of 0 – 1.2 M GuHCl, then decreased to reach a value lower than 1 at 0 M GuHCl at 3.0 M, Figure 5.15.
Figure 5.14: Log-log graph for GFP showing a slope of 2

Figure 5.15: Fluorescence emission spectra of GFP at 800 nm excitation, inset shows the fluorescence counts at 507 nm at various [GuHCl]
5.3.7 Two-photon absorption cross-section

Two-photon absorption cross-sections ($\delta$) measurements were performed using a method employed by Rebane et al$^{30}$, equation (5.6).

$$\delta(s) = \frac{I_{2P}(s) \cdot I_{1P}(folded)}{I_{1P}(s) \cdot I_{2P}(folded)} \cdot \delta(folded)$$

(5.6)

Where, $I_{2P}(s)$, and $I_{2P}(folded)$ are the fluorescence intensities of two-photon absorption of the sample at a certain [GuHCl] in the folded state. $I_{1P}(s)$ and $I_{1P}(folded)$ are the fluorescence intensities of one-photon absorption at a certain [GuHCl] and in the folded state.

![Graph showing 2PA cross-section of GFP at different [GuHCl]].

Figure 5.16: Two-photon absorption scan of GFP at different [GuHCl], inset shows the relative 2PA cross-section

The 2PA cross section of samples at 0, 1.0, 1.6 and 2.6 M GuHCl were measured as the wavelength was scanned from 710 – 880 nm. The sample at 2.6 M GuHCl showed almost a constant value over all wavelengths measured. For samples 0, 1.0, and 1.6 M GuHCl, the value increased as the wavelength increased. The samples of 0 M and 1.0 M exhibit a lower 2PA cross-section than 1.6 M, Figure 5.16. Since the electric field is
related to the 2PA cross section ($\delta$) the relative change of the electric field should be the same, as shown in Figure 5.17. The 2PA cross section ($\delta$) is directly related to the change in the dipole moment from the ground state to the excited state ($\mu_{ge}$). And the latter is related to the change in the local electric field ($\vec{E}$), equation (5.7):

$$\delta \propto (\Delta \mu_{g\rightarrow e})^2 \quad \Delta \vec{\mu}_{g\rightarrow e} = \Delta \vec{\mu}_{g\rightarrow e} \pm \frac{1}{2} \vec{E} \quad (5.7)$$

In the initial state of unfolding the dipole moment of the GFP chromophore is parallel to the local electric field in the beta barrel. When GuHCl is added to the protein sample, the electrostatic interaction from its ions enhances the electric field around the chromophore. After the protein completely unfolds, the chromophore no longer feels the electric field and so it drops dramatically.

![Figure 5.17: Change in local electric field in GFP as it unfolds](image)

Figure 5.17: Change in local electric field in GFP as it unfolds
5.4 Summary

GFP was cloned into the pET-32 Xa/LIC, and expressed as a fusion protein, then purified using His Tag affinity column and gel filtration. The protein is a monomer with a mass of 27 kDa. Unfolding of GFP by GuHCl was monitored by CD. 50% was unfolded at around 1.6 M GuHCl. The anionic form of the GFP chromophore absorbs light at 480 nm and emits light at around 508 nm. During the unfolding, the neutral form of the chromophore increases which was noted by its absorbance at 380 nm. As the protein unfolds the chromophore fluorescence is quenched due to solvent accessibility. The fluorescence lifetime and the fluorescence intensity both decrease while unfolding in a profile similar to that obtained by the change in the CD signal.

The 2PA cross-section increased to reach its maximum at around 1.6 M of GuHCl, and then dropped to a value lower than the initial value. This is due to the change in the local electric field around the chromophore and to the change in the chromophore state. Initially, the chromophore is in its anionic form and its dipole moment seems to slightly align with the local electric field. In the initial stages of unfolding, the anionic form of the chromophore experiences additional electrostatic interactions that amplify the electric field around it. However, when the protein is unfolded completely, the chromophore becomes neutral and the protein loses its beta barrel structure and no influence of electric field is seen in the 2PA cross-sections. These results have shown that the 2PA cross-sections of chromophores can be used as markers to study proteins unfolding. The ratio of two-photon to one-photon fluorescence can also be used to image the folded, unfolded or mis-folded forms of proteins. In addition, ultrafast fluorescence measurements were able to monitor the unfolding of GFP quite effectively.
5.5 References


(26) Herberhold, H.; Marchal, S.; Lange, R.; Scheyhing, C. H.; Vogel, R. F.; Winter, R. Characterization of the Pressure-Induced Intermediate and Unfolded State of Red-


6.1 Introduction

Novel techniques of two-photon absorption (2PA) and ultrafast fluorescence to monitor protein folding/unfolding were demonstrated with a case study of GFP in chapter 5. With this background, we have utilized the same techniques to study the folding and unfolding of WLN5-6. It is well established in the literature that fluorescence spectroscopy is an important tool for tackling protein folding because of its sensitivity and applicability\(^1\). Natural intrinsic fluorescence from the tryptophan, and to lesser extent from tyrosine residues, is used to monitor the tertiary structure of proteins in solution and can provide information on conformational changes of proteins during unfolding. Also, we have shown with GFP, that it is possible to use the linear and nonlinear optical properties of chromophores to study the unfolding (Chapter 5). However, for WLN5-6, there is no intrinsic chromophore that can be accessed with the fundamental Ti:Sapphire laser. An alternative approach is to use extrinsic fluorescence where a fluorescent dye is added to the protein. Extrinsic dyes can be covalently bound to the protein through the thiol group of the cysteine or through the amine group of the \(\alpha\)-amino terminus or the \(\varepsilon\)-amine group of lysine\(^2\)\(^-\)\(^4\). However, there is a risk that added dye may affect the conformational stability, aggregation tendency and interfere with the interaction of other ligands with the protein\(^5\). Therefore, fluorescence from a protein using a protein’s
intrinsic fluorescence is the preferred approach to monitor conformational changes. Fluorescence signals can provide information regarding conformational changes in proteins due to changes in the polarity around the fluorophor probe, the emission spectra can shift to shorter or longer wavelength, fluorescence anisotropy \(^6\), emission maximum, fluorescence lifetime, fluorescence anisotropy, and rotational correlation time \(^7\).

In this work, it is shown that two-photon absorption is a valid method to monitor protein unfolding (Chapter 5), and GFP unfolding is investigated by the two-photon absorption cross-section method. In this chapter, I investigate the unfolding of WLN5-6 construct by a new method, in addition to the time-resolved fluorescence techniques, such as anisotropy and fluorescence decay. The unfolding of WLN5-6 was monitored through extrinsic fluorescence rather than intrinsic fluorescence because it has no tryptophan residue in its sequence. It has three tyrosine residues and five phenylalanine residues, which are weak fluorophores compared to tryptophan and need a high excitation energy which may cause them to degrade.

The probe used is a thiol-reactive reagent called, 7-diethylamino-3-(4'-maleimidylphenyl)-4-methylcoumarin (CPM). CPM contains a maleimide group that forms a covalent bond with the cysteine residue. Figure 6.1 shows the structure of CPM and the covalent bond formed.
Two constructs of WLN5-6 were produced via a series of site-directed mutagenesis experiments in order to obtain two protein constructs that possesses only one cysteine in each ferrodoxin fold, at positions 6 and 51, Figure 6.2. Position 6 is located in the β1 sheet of domain 5 while position 51 located in the loop between β3 and α2 of domain 5. CPM then was bound to the protein, and the resultant dye-protein system was
unfolded using GuHCl and the unfolding was monitored by CPM fluorescence. All other cysteines at positions 15, 18, 91, 94 were mutated to serine.

Figure 6.2: Cysteine positions in both constructs of WLN5-6
6.2 Experimental methods

6.2.1 Construction of plasmid pET24dWLN5-6C51

The gene encoding WLN5-6 was amplified and cloned into pET24d<sup>9</sup> producing pET24dWLN5-6 by Dr. David Achila, a previous laboratory member. It was used to produce both mutants with only one cysteine residue in the sequence, one mutant has the cysteine at position 51 (WLN5-6C51) and the second has a cysteine at position 6 (WLN5-6C6). The resultant plasmids are called pET24dWLN5-6C51 and pET24dWLN5-6C6, respectively. pET24dWLN5-6C51 was constructed, from pET24dWLN5-6, by performing 6 mutations targeting 5 cysteine residues and one glutamate residue in order to have only one cysteine residue at position 51. A mutant of pET24dWLN5-6 with 7 mutations, which targeted the five cysteins and glutamate at positions 51 and 127, was available in the laboratory, as a starting point<sup>10</sup>. The cysteines were mutated to serines and selected glutamate was mutated to cysteine. This plasmid was used as template to produce pET24dWLN5-6C51 by performing two mutations, C127E and S6A, as shown in the scheme in Figure 6.3.
Figure 6.3: Constructions of WLN5-6C51

Mutations were performed using QuikChange® II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA). All required primers (forward and reverse) were designed as recommended (Table 2.3). First, cysteine at position 127 was altered to glutamate, (C127E). The sample was prepared as described in Table 6.1, then 1 µL of pfuUltra HF DNA polymerase (2.5 U/µL) was added. The sample was overlaid with 30 µL mineral oil and placed in PCR thermocycler machine. The thermocycle parameters, shown in Table 6.2, were carried out. After segment 3 was finished, about 2.5 hours, 1 µL of Dpn I (10 U/µL) was added to the solution below the mineral oil layer and was mixed well by pumping it up and down with a micropipette. The resultant sample was incubated for 1
hour at 37°C and then chilled on ice, prior to transformation into XL 10-Gold competent cells. The competent cells were thawed on ice for 5 minutes, then 2 µL of β-Mecaptoethanol mix was added to the cells for higher efficiency of transformation. After incubating the cells for 10 minutes, 2 µL of Dpn I-treated DNA was added to the cells. The cells were incubated in ice for another 30 minutes. Then the cells were heat shocked by incubating the tube in 42°C for 30 seconds and then immediately incubated on ice for 2 minutes. 500 µL of pre-heated SOC medium was added to the tube and then incubated at 37°C for 1 hour with shaking at 250 rpm. 250 µL of the transformed cells were plated in the LB/Agar plate containing 30 µg/mL of Kanamycin. The plate was incubated at 37°C for 18 hours. After incubation, about 100 colonies were grew on the plate. A single colony was inoculated in a 5 mL LB autoclaved medium (contains 30 µg/mL Kanamycin). Then the tube was incubated at 37°C with shaking for 16 hours. The cells then were harvested by centrifugation and the plasmid was purified using QIAGEN® QIAprep spin Miniprep Plasmid Purification protocol.

The cell pellet was suspended in 250 µL of buffer P1, then 250 µL of buffer P2 was added and the tube was inverted for 4-6 times. 350 µL of buffer N3 was added and the tube was inverted for 4-6 times. The solution then was centrifuged at 13.5 rpm for 10 minutes. Supernatant was decanted into the spin column then centrifuged for 1 minute. The flow through was poured away and 750 µL of buffer PE was added to the spin column, then centrifuged for 1 minute. Again, the flow through was poured away and then the spin column was centrifuged for another 1 minute. The spin column was placed in a clean eppendorf tube and 30 µL of EB buffer was added to the column. Then it was let stand for 2 minutes before being centrifuged for 1 minute and the flow through that
contained the pure plasmid was collected in the eppendorf tube and stored at – 20℃. The plasmid was analyzed on 0.8 % agarose gel. A 1000 ng sample was sent to Retrogen, Inc. (San Diego, CA) for sequencing using T7 and T7-term primers. After confirming the success of the mutation, the resulting plasmid was used as a template for the incorporation of the second mutation, S6A. In order to incorporate the second mutation, S6A, a solution was assembled as described in Table 6.1, and PCR was performed as, Table 6.2. The mutation and the native plasmid purification were performed as stated previously.

Table 6.1: Sample preparation for mutations in WLN5-6C51 and WLN5-6C6

<table>
<thead>
<tr>
<th>Construct</th>
<th>Mutation</th>
<th>10 X Reaction buffer (μL)</th>
<th>DNA Template 10-50 ng (μL)</th>
<th>Forward primer 125 ng (μL)</th>
<th>Reverse primer 125 ng (μL)</th>
<th>dNTP mix (μL)</th>
<th>Quick Solution (μL)</th>
<th>MilliQ H2O (μL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WLN56C51</td>
<td>C127E</td>
<td>5</td>
<td>3</td>
<td>1.087</td>
<td>1.11</td>
<td>1</td>
<td>3</td>
<td>35.8</td>
</tr>
<tr>
<td></td>
<td>S6A</td>
<td>5</td>
<td>4</td>
<td>1.036</td>
<td>1.047</td>
<td>1</td>
<td>3</td>
<td>34.9</td>
</tr>
<tr>
<td>WLN5-6C6</td>
<td>C15,18S</td>
<td>5</td>
<td>2</td>
<td>1.0</td>
<td>1.0</td>
<td>1</td>
<td>3</td>
<td>37</td>
</tr>
</tbody>
</table>

192
Table 6.2: Thermocycle parameters for mutagenesis

<table>
<thead>
<tr>
<th>Segment</th>
<th>Cycles</th>
<th>Temperature</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>95°C</td>
<td>1 min</td>
</tr>
<tr>
<td>2</td>
<td>18</td>
<td>95°C</td>
<td>50 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>60°C</td>
<td>50 sec</td>
</tr>
<tr>
<td></td>
<td></td>
<td>68°C</td>
<td>6.5 min</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>68°C</td>
<td>7 min</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4°C</td>
<td>20 min</td>
</tr>
</tbody>
</table>

In the mutant, pET24dWLN5-6C51, cysteine 6 was mutated to encode alanine rather than serine, because when mutated to serine the protein folded improperly\(^\text{10}\). The protein sequence of WLN5-6C51 is listed below, and the incorporated amino acids are shown in yellow.

WLN5-6C51 sequence

MAPQK AFLQIKGMSASSVSNIERNLQKEAGVLSVLVALMAGKAEIKYDP
CVIQPLEIAQFIQDLGFEAAVMEDYAGSDGNIELTTGMTASSVHNIES
KLRTNGITYASVALATSKALVKFDPEIIGPRDIIIEEIEIGFHASLAQ

6.2.2 Expression and purification of WLN5-6C51 protein

pET24dWLN5-6C51 was transformed into Rosetta 2 (DE3), and then grown on LB/agar plate containing 30 μg/mL kanamycin. A single colony was inoculated in a 5 mL LB medium, scaled up to 1 L, and the protein expression was induced with 1 mM IPTG. After 3 hours of induction, protein was harvested by centrifugation and extracted from
the pellet using the freeze-thaw method. The extraction buffer consists of 20 mM Mes, pH 6.0, 0.1 mM EDTA, and 5mM DTT. The pI of the protein is around 5, according to IEF and DNA strider® prediction, and so, at pH 6 it is expected to possess an overall negative charge. WLN5-6C51 purification was achieved by passing the supernatant of the extraction step through a DEAE-Sepharose anion exchange column. The negatively-charged proteins will bind to the column, and the positively-charged proteins will pass through. Bound protein was eluted using the extraction buffer, except that 1M NaCl was added to it. The fractions that contain the protein were identified using SDS-PAGE, then collected and concentrated. The protein was further purified by gel filtration using Superdex 75 column (GE Healthcare). The buffer contained 50 mM HEPES, 200 mM NaCl, 5 mM DTT, pH 7.5. The protein eluted at around 200 ml. The fractions containing the protein were collected and concentrated. The buffer was exchanged to 50 mM sodium phosphate buffer, pH 7.5, using a desalting column. The protein was concentrated using a 3000 MWCO membrane in an Amicon device and stored at –20°C.

6.2.3 Binding WLN5-6C51 to CPM dye

WLN5-6C51 was reacted with CPM using a general procedure suitable for binding most thiol-reactive probes. In this process, 50-100 µM of WLN5-6C51 was dissolved in 50 mM sodium phosphate buffer at pH 7.5 at room temperature. In pH range of 7.0-7.5, the thiol group of the cysteine is sufficiently nucleophilic, while the amines groups are protonated and unreactive. The CPM dye was dissolved in DMSO just immediately prior to use, to about 1-10 mM. The dye is not stable when exposed to light, therefore aluminum foil was used to wrap it and prevent exposure to the light. 10-20
moles of dye were added per mole of protein. The dye was added dropwise with gentle stirring. The reaction was allowed to proceed for 2 hours at room temperature, then for 17 hours at 4°C. The reaction mixture was protected from light all the time. Upon completion of the reaction, excess reduced glutathione solution was added to the reaction mixture and stirred for 30 minutes in order to consume excess CPM dye. The mixture was concentrated using an Amicon ultrafiltration device with a 3000 MWCO membrane to a volume of 2.5 mL. This reduced volume was passed through a PD-10 column (GE Healthcare) to separate the CPM-modified protein from free CPM dye and exchange the buffer into 50 mM sodium phosphate buffer. The protein then was concentrated again and quantified using the BCA assay. Almost 75% of the protein was recovered. The labeled protein mass spectrum was obtained to verify the binding of the dye.

6.2.4 Cloning, expression and purification of WLN5-6Cys6 protein

Four mutations were performed on pET24dWLN5-6 plasmid to prepare a protein with only one cysteine residue at position 6. At this position the cysteine is expected to be partially buried between domain 5 and 6. A plasmid of WLN5-6 containing two mutations was available in the laboratory\textsuperscript{10}, and was used as a template to produce the above mentioned mutants. The two mutations changed the two cysteines at positions 91 and 94 to serine, Figure 6.4. To change C15 and C18 to serine, one site directed mutagenesis reaction with QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene) was performed. Both mutations were completed using one set of primers, Table 2.3. The final plasmid construct had 4 mutations, cysteine 15, 18, 91, and 94 changed to serine, and therefore the plasmid contains only one cysteine at position 6.
Figure 6.4: Scheme represents the mutations performed for WLN5-6C6

The plasmid was sequenced to verify mutation incorporation. The resultant plasmid was transformed into Rosetta2(DE3), and plated on the LB/agar containing 30 μg/mL kanamycin. A single colony was inoculated into 5 mL LB, scaled up to 1 L, and the protein expression was induced with 1 mM IPTG final concentration. After 3 hours, the protein was harvested by centrifugation and extracted from the pellet using a freeze-thaw method. Protein purification was performed by DEAE-Sepharose anion exchange and gel filtration by Superdex 75 column.
The protein has 149 amino acid residues. The protein sequence is shown below. The protein has an average mass of 15984 Da, verified by Electrospray ionization-ion trap (ESI-IT) mass spectrometry, performed in Professor Andre Venter’s laboratory at Western Michigan University. This protein was then bound to CPM dye as described for WLN5-6C51, Section 6.2.3.

<table>
<thead>
<tr>
<th>WLN5-6C6 sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAPQKCFQIKGMT$ASSVSNIERNLQKEAGVLSVLVALMAGKAEIKYDP</td>
</tr>
<tr>
<td>EVIQPLEIAQFIQDLGFEAAVMEDYAGSDGNIELTITGMT$ASSVHNIES</td>
</tr>
<tr>
<td>KLTRTNGITYASVALATSKALVKFDPEIIGPRDIKIIEEIGFHASLAQ</td>
</tr>
</tbody>
</table>

6.2.5 Circular dichroism

The mutant proteins, WLN5-6C51 and WLN5-6C6, and their dye bound forms were unfolded using GuHCl. The unfolding was monitored by circular dichroism (CD). Protein solutions of 20-30 μM were prepared with various GuHCl concentrations, ranging from 0 to 8 M, then equilibrated at room temperature for at least 2 hours prior to CD measurements. A CD scan from 300 nm to 190 nm was performed; each scan was the replicated and averaged 3 times. The secondary structure was monitored at 222 nm using a J-815 CD Spectropolarimeter with a rectangular cell of 1 mm path length.

Guanidine hydrochloride (GuHCl) stock solution of 8 M was used for preparing various GuHCl concentrations. Its concentration was accurately determined from refractive index measurements described elsewhere\textsuperscript{11}. 
6.2.6 Fluorescence spectroscopy

Steady-state measurements, time-resolved fluorescence and fluorescence anisotropy were performed as described in Section 5.2.4. Two-photon absorption cross-section (δ) measurements were performed using a method employed by Rebane et al.\textsuperscript{12}.

6.2.7 Quantum yield

The emission spectra of the WLN5-6C51 at 0, 0.5, and 4 M GuHCl as well as C485 were measured upon excitation at 370 nm. For quantum yield analysis, the absorbance of all samples was fixed at around 0.1. C485 in methanol was used as the standard. The refractive index for the dye in solution was considered the same as that of water.

6.3 Results and discussion

6.3.1 Measuring of quantum yield of WLN5-6C51_CPM

The fluorescence quantum yield (Φ\textsubscript{F}) is the ratio of photons absorbed to photons emitted through fluorescence. In other words the quantum yield gives the probability of the excited state being deactivated by fluorescence rather than by another non-radiative mechanism. The most reliable method for recording Φ\textsubscript{F} is the comparative or relative method of Williams \textit{et. al.}\textsuperscript{13} which involves the use of standard samples with known Φ\textsubscript{F} values. Essentially, solutions of the standard and test samples with identical absorbance at the same excitation wavelength can be assumed to be absorbing the same number of photons. Hence, a simple ratio of the integrated fluorescence intensities of the two
solutions (recorded under identical conditions) will yield the ratio of the quantum yield values. Since $\Phi_{F}$ for the standard sample is known then the $\Phi_{F}$ for the test sample is easy to be calculated.

The $\Phi_{F}$ for three samples of WLN5-6C51_CPM was measured. These samples are at 0, 0.5 and 4 M GuHCl. The comparative dilute method was used and the following equation, equation (6.1), was applied.

$$\Phi_{F} = (\Phi_{F})_{S} \frac{\int J(\nu) d\nu (J_{a})_{s} \eta_{s}^{2}}{\int J_{s}(\nu) d\nu J_{a} \eta_{a}^{2}}$$  \hspace{1cm} (6.1)

Where $\int J(\nu) d\nu$ and $\int J_{s}(\nu) d\nu$ were the area under the fluorescence emission curve for the sample and the standard, respectively. $(J_{a})_{s}$ and $J_{a}$ were the absorbance of the standard and sample, respectively. $\eta$ and $\eta_{s}$ are the refractive index of the solvent for the sample and the standard, respectively. The standard that has been used for the quantum yield measurements is C485 dye and the results are summarized in Table 6.3.

**Table 6.3: Quantum yield measurements of WLN5-6C51_CPM at 0, 0.5 and 4 M guanidine hydrochloride**

<table>
<thead>
<tr>
<th>Sample ID</th>
<th>O.D (370 nm)</th>
<th>Area</th>
<th>RI ((\eta))</th>
<th>$\eta^{2}$</th>
<th>Area (\frac{Sam}{Std})</th>
<th>O.D (\frac{Std}{Sam})</th>
<th>$\eta^{2}$ (\frac{Sam}{Std})</th>
<th>(Q.Y)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C485</td>
<td>0.096</td>
<td>4.0E+08</td>
<td>1.328</td>
<td>1.7646</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.160</td>
</tr>
<tr>
<td>0 M</td>
<td>0.089</td>
<td>1.4E+09</td>
<td>1.330</td>
<td>1.7689</td>
<td>3.47036</td>
<td>1.078</td>
<td>1.0024</td>
<td>0.600</td>
</tr>
<tr>
<td>0.5 M</td>
<td>0.099</td>
<td>1.7E+09</td>
<td>1.340</td>
<td>1.7956</td>
<td>4.09509</td>
<td>0.97</td>
<td>1.0175</td>
<td>0.646</td>
</tr>
<tr>
<td>4 M</td>
<td>0.1</td>
<td>1.5E+09</td>
<td>1.390</td>
<td>1.9321</td>
<td>3.83492</td>
<td>0.96</td>
<td>1.0949</td>
<td>0.645</td>
</tr>
</tbody>
</table>
6.3.2 Mass spectrometry

Mass spectrometry was used to ensure the incorporation of the mutations and to show absence of post translational modifications for both proteins, WLN5-6C51 and WLN5-6C6. The average mass expected for WLN5-6C51 based on its sequence is about 15927 Da. The mass spectrum, Figure 6.5, shows a peak with a mass of 15926 Da which corresponds to the protein. However, another peak shows a mass of 15795 Da. This mass corresponds to WLN5-6C51 without the initial methionine. It is estimated that about 60% of mature proteins lose their first methionine.\textsuperscript{14}

Figure 6.5: ESI-IT/MS data showing the molecular weight of WLN5-6C51 determined by MagTran software
The CPM dye has a mass of 402 g/mol. When it bound to the protein, the protein mass became 16320 g/mol as shown in Figure 6.6.

![ESI-IT/MS data showing the molecular weight of WLN5-6C51_CPM determined by MagTran software](image)

Figure 6.6: ESI-IT/MS data showing the molecular weight of WLN5-6C51_CPM determined by MagTran software

Similar observation was obtained for WLN5-6C6. It is calculated average mass is 15985 Da. The mass spectrum shown in Figure 6.7 indicates a peak at 15985 Da, peak A,
corresponds to the protein while the peak at 15853 Da indicates the protein when the methionine was lost.

Figure 6.7: ESI-IT/MS data showing the molecular weight of WLN5-6C6 determined by MagTran software

The mass spectrometry shows that both proteins, with and without methionine, bound to CPM dye and the mass spectrum is shown in Figure 6.8
6.3.3 Circular dichroism

6.3.3.1 WLN5-6C51 and WLN5-6C51_CPM

The unfolding of WLN5-6C51 protein was monitored by CD. The unfolding profile shows a two state model, and at 2.7 M GuHCl half of the protein molecules are unfolded, Figure 6.9. The least-squares linear extrapolation regression was used to determine at the thermodynamics parameters of the protein, Figure 6.10. The results show a free unfolding energy of $9.67 \pm 0.46$ kJ/mol. There are no significant differences between the unfolding of the wild type and this mutant. Therefore, the 6 mutations that were incorporated to WLN5-6C51 (C15S, C18S, C91S, C94S, C6A, and E51C) did not change the folding properties. This result was different when C6 was mutated to serine$^{10}$. This result may indicate that C6 interacts with hydrophobic residues. So, when mutated
to a hydrophobic residue, such as alanine, the interaction was not affected. Therefore, WLN5-6C51 maintains a native-like structure. Serine, which has an –OH moiety, is more polar and so disrupts the native hydrophobic interactions that stabilize the protein.

Labeling WLN5-6C51 with CPM did not affect the folding of the mutant. It actually stabilized it slightly. WLN5-6C51_CPM maintains 50% of its folded structure at 3.1 M GuHCl compared to 2.7 M prior to labeling, Figure 6.11. The literature indicates that labeling could stabilize the protein if it binds to a hydrophobic region\textsuperscript{15-17}. The free energies for both the mutant and the labeled protein are almost the same when the range of the error is considered. The free energy for the mutant is 9.67 ± 0.46 kJ/mol, while after labeling it is 9.3 ± 0.52 kJ/mol, Figure 6.12.

![Figure 6.9: CD signal for the unfolding of WLN5-6C51 by GuHCl](image)
Figure 6.10: Free energy of WLN5-6C51 by linear extrapolation method

Figure 6.11: CD signal of the unfolding of WLN5-6C51_CPM
6.3.3.2 WLN5-6C6 and WLN5-6C6_CPM

The unfolding of WLN5-6C6 was monitored by CD, Figure 6.13. In this mutant, the four cysteines in the metal-binding sites of WLN5-6 were changed to serine. Obviously, the mutations did not affect the folded state of the protein or the unfolding profile, probably, because the cysteines were located in the loops and not in the core of the protein. Therefore, they are not interacting with other residues. Also, the free energy calculation shows higher free energy comparing to the other mutant, WLN5-6C51, Figure 6.14. The free energy is 15.36 ± 0.65 kJ/mol. The mutant also remains stable and folded even after labeling it with CPM as shown in Figure 6.15. Its free energy was determined to be 15.14 ± 0.49 kJ/mol, Figure 6.16.
Figure 6.13: CD signal of the unfolding of WLN5-6C6 by GuHCl

Figure 6.14: Free energy of WLN5-6C6
Figure 6.15: CD signal of the unfolding of WLN5-6C6_CPM by GuHCl

Figure 6.16: Free energy of WLN5-6C6_CPM
6.3.4 Steady-state measurements

6.3.4.1 WLN5-6C51_CPM

The steady state absorption and fluorescence measurements were obtained using the parameters Section 6.2.6. A quartz cuvette with a width of 4 mm was used for the measurements. A free CPM dye absorbs light with excitation maxima at 384 nm and emission maxima at 469 in methanol solvent\(^{18}\). However, the fluorescence of the dye is sensitive to the environment. In this chapter, the change in the emission intensities and in the emission maxima of the CPM attached to the protein were used to follow the unfolding of the proteins. The steady-state measurements of the dye attached to the protein were performed on the same samples that were prepared for the CD measurements. So, all the samples contained the same concentration of the protein, 20-30 \(\mu\text{M}\), while GuHCl were varied from 0 – 7.5 M, in 5 mM phosphate buffer.

The absorption spectrum and the fluorescence spectrum of the dye for WLN5-6C51_CPM are shown in Figure 6.17 and Figure 6.18, respectively. The absorption maxima increased slightly while the protein was unfolding. This change was due to the change in the environment around the CPM probe while unfolding. The CPM dye became more exposed to the solvent, therefore its absorbance increased and shifted toward longer wavelength, probably due to an increase in the refractive index of the solution when GuHCl concentration increased. The fluorescence was measured immediately after the absorption was measured. It also showed the same trend for the absorption. The emission maxima (apex) increased from 468 nm to the folded protein to 480 nm for the unfolded protein, Figure 6.18, which reflects an increase in the quantum
efficiency for the excited dye molecules when they are not constrained by protein structure.

Figure 6.17: Absorption scan of WLN5-651_CPM at various [GuHCl]

Figure 6.18: Emission scan of WLN5-6C51_CPM at various [GuHCl]
The dye was also excited using laser light at 800 nm to measure the two-photon absorption fluorescence. The logarithm of the fluorescence intensity was plotted against the logarithm of the power to ensure that the excitation was due to the absorption of two photons. The slope of the plot had value of about 2, as shown in Figure 6.19. Similar to one-photon fluorescence, the fluorescence from two-photon absorption increases with the unfolding of the protein, Figure 6.20 and Figure 6.21.

![Log-log plot for two-photon absorption measurements](image-url)

**Figure 6.19:** log-log plot for two-photon absorption measurements
Figure 6.20: Emission scan of two-photon absorption for WLN5-6C51_CPM at various [GuHCl]

Figure 6.21: Neutralized emission maximum counts for WLN5-6C51_CPM at 480 nm at different [GuHCl]
6.3.4.2 WLN5-6C6_CPM

The absorbance of WLN5-6C6_CPM was also measured in the same way as for the WLN5-6C51_CPM. Similar results were obtained: the absorbance maxima increased and shifted to a longer wavelength, Figure 6.22. The fluorescence changed in the same way, but the folded protein has longer maxima. The fluorescence maximum for the folded protein was 473 nm and started to increase with the addition of GuHCl until it reached 478 nm, Figure 6.23. Two-photon absorption was also performed on the samples using a laser to excite the samples at 800 nm, Figure 6.24. When the enhancements of two-photon and one-photon fluorescence were compared, two-photon fluorescence was more than one photon, Figure 6.25.

![Absorption Scan](image)

*Figure 6.22: Absorption scan for WLN5-6C6_CPM at various [GuHCl], inset shows λ_max changes while protein unfolds*
Figure 6.23: Emission scan of WLN5-6C6_CPM at various [GuHCl], inset shows Fl. counts at 478 nm with respect to [GuHCl]

Figure 6.24: Emission scan of two-photon absorption for WLN5-6C6_CPM at various [GuHCl]
6.3.5 Two-photon absorption cross-section

A control was prepared by binding CPM dye to a free cysteine residue in order to compare the behavior of the CPM dye bound to the protein, with respect to the dye not bound to the protein. The dye was dissolved in DMSO then an excess of cysteine was added to the dye solution and stirred for two hours at room temperature and in a dark environment. The product, Cys-CPM, was used as a control for the spectroscopic measurements of protein systems that bound to CPM.

Different solutions of GuHCl solutions were prepared (from 0 M - 8 M), and the dye concentration in all GuHCl solutions were constant. All the experiments that were performed for the labeled proteins were applied for the control under the same conditions.
The two-photon absorption (2PA) cross-section for the control, WLN5-6C51_CPM, and WLN5-6C6_CPM at different GuHCl concentrations were shown in Figure 6.26. The two photon absorption cross-section (δ) was measured as the ratio between the intensities of the two-photon fluorescence (I_{2P}) and the intensities one-photon fluorescence (I_{1P}) as in the following equation (6.2):

\[ \delta_{(s)} = \frac{I_{2P \ (s)}}{I_{1P \ (s)}} \times \frac{I_{1P \ (folded)}}{I_{2P \ (folded)}} \times \delta_{(folded)} \]  

(6.2)

2PA cross-sections increased as the protein unfolded; this change was due to the change of the dipole moment of the dye which was produced because of the change in the electric field. 2PA increasing means that the dye dipole moment was working against the electric field of the protein. Therefore, 2PA cross-section can be used to image the local electric field in the protein through the dipole moment measurements of the fluorophore.

![Figure 6.26](image.png)

Figure 6.26: Two-photon absorption cross-section for the control, WLN5-6C51_CPM and WLN5-6C6_CPM
6.3.6 Time-correlated single photon counting

The fluorescence lifetime of the CPM probe was measured through time-correlated single photon counting (TCSPC). The samples were excited with 370 nm pulsed laserlight. The emission decay was observed and the signal de-convoluted using the instrument response to obtain the actual lifetime. The instrument response function (IRF) was obtained from measuring the emission decay for a scattered sample; usually it was either a turbid solution or a piece of aluminum foil placed in the sample holder. The acquired emission decay curves were fitted to either single or multi-exponential decay equation, based on the fluorescence decay profile, equation (6.3):

\[ I(t) = \sum_{i=0}^{n} a_i e^{-t/\tau_i} \]  

(6.3)

The change in lifetimes for the control, WLN5-6C51_CPM, and WLN5-6C6_CPM is shown in Figure 6.27. The lifetime for the folded state of both constructs was relatively higher than the unfolded form of the proteins. However, WLN5-6C51_CPM has a higher lifetime than WLN5-6C6_CPM, 3.3 ns and 2.8 ns, respectively, while the control has a lifetime of 1.7 ns. This is due to the change in the local environment around the CPM probe. When the probe was in a very constrained environment with few ionic interactions, or hydrogen bonding, and with no accessibility to the solvent, its fluorescence lifetime was higher. Therefore, the control has lower lifetime than the CPM bound to the proteins. However, WLN5-6C51_CPM appeared to be in the most constrained environment. In the case of the control, the fluorescence lifetime was continuously increasing with the addition of GuHCl. That is due to change in the pH and an increase in the electric constant. When CPM dye was bound to the
protein, its fluorescence lifetime decreases as the protein starts to unfold despite high GuHCl concentration. The effect of GuHCl did not overcome the effect of having the probe in a constrained environment until the protein completely unfolded. Then the bound CPM behavior seemed to be similar to the control, Figure 6.27.

\[ r = \frac{I_{VV} - G \cdot I_{VH}}{I_{VV} + 2 \cdot G \cdot I_{VH}} \] (6.4)

6.3.7 Fluorescence upconversion

Fluorescence upconversion was used to study time-resolved polarized fluorescence of the CPM dye at femtosecond timescales. The sample was constantly rotated in a 1 mm cylindrical cell. Long, intermediate, and short scans for both parallel and perpendicular decays were measured in order to calculate anisotropy based on equation (6.4):
The anisotropy measurements for the control, WLN5-6C51_CPM, and WLN5-6C6_CPM are shown in Figure 6.28. The anisotropy values for the CPM bound to the protein were lower than the control. The anisotropy increased in all cases, but in the case of the protein, the change was similar to the CD unfolding profile. In case of the control, which is a CPM dye bound to a cysteine, the anisotropy is in direct relation with the increasing of [GuHCl]. Therefore, at higher [GuHCl] the CPM is changing its orientation quickly. Similar observations were noticed when the dye was bound to the proteins, WLN5-6C51_CPM and WLN5-6C6_CPM, but the rate of the increase suddenly became higher when the protein started to unfold. Unfolding of the protein will release the CPM from being constrained to become free to orient itself and so the anisotropy increases. The value of the anisotropy for WLN5-6C6_CPM is 0.30 which is higher by almost 10% more than the value for WLN5-6C51_CPM, 0.27. This result shows that the CPM orientation changes more when it is at position 51 rather than position 6. The cysteine at position 51 is more exposed to the solvent since it is in the loop between β3 and α2; however, position 6 is enveloped in the protein core.
Figure 6.28: Change in Anisotropy with unfolding by GuHCl for the control, WLN5-6C51_CPM and WLN5-6C6_CPM

6.4 Summary

All the methods that have been used to monitor the unfolding showed a transition at around 2.5 M GuHCl which was confirmed by the circular dichroism as a transition for unfolding the protein. The maximum emission wavelength was shifted toward the longer wavelength while increasing the denaturant concentration. During the unfolding, the local environment around the CPM dye was changed and the dye became more exposed to the solvent which stabilizes the excited state of the dye and lowers its energy. As a result, the emission shifted toward longer wavelength while unfolding the protein. On the other hand, the control fluorescence did not shift. The change of the anisotropy indicates a change in the dipole moment orientation of the CPM dye while the protein unfolds. This change is attributed to the change of the electric field around the dye. The change in the anisotropy was quick compared to the changes based on other methods. The dye
anisotropy changed from 0.27 in the folded protein to 0.34 for the unfolded state. The control shows a slight increase in the anisotropy due to the increasing concentration of denaturant. The two photon absorption cross-section ($\delta$) is measured as the ratio between the intensities of the two-photon fluorescence ($I_{2p}$) and the intensities of one-photon fluorescence ($I_{1p}$) as in the equation (6.2). 2PA cross section also increases as the protein unfolds. This change is due to the change of the dipole moment of the dye which was produced because of the change in the electric field, equation (6.5).

$$\Delta \mu_{eg}^0 = \Delta \mu_{eg} + 0.5 (\alpha_1 - \alpha_0) E \cos \theta$$  \hspace{1cm} (6.5)

$\mu_{eg}^0$ and $\mu_{eg}$ are the dipole moment at zero and non zero electric field. $\alpha_1$ and $\alpha_0$ are the polarizabilities of excited and ground state, respectively. $\theta$ is the angle between the dipole moment and the electric field.

The increasing in 2PA means that the dye dipole moment is working against the electric field of the protein and that they are antiparallel to each other as shown in Figure 6.29. The 2PA cross section can be used to image the local electric field in the protein through the dipole moment measurements of the fluorophore.
Figure 6.29: The direction of the dipole moment in CPM in WLN5-6C51_CPM and WLN5-6C6_CPM with respect to the electric field in WLN5-6 mutant
The lifetime of the CPM dye is very dependent on its environment. So if it is in a rigid environment as in the case of the folded protein, the excited state lifetime will be longer than if it is more flexible. The lifetime of CPM in the folded state was 3.3 ns. While in the unfolded state of the protein it decreased to be 2.7 ns. However, the CPM dye lifetime in the control showed a continuous increase in the lifetime due to the stability of the excited state as the denaturant concentration increases.
6.5 References


(10) Marzijarani, N. Biophysical Characterization of the N-Terminal Metal Binding Domains 4-6, 5-6 and Two Disease-Causing Mutations of the Human Wilson Protein.; Western Michigan University: Kalamazoo, MI. M.S. Thesis., 2010., Western Michigan University, 2010, p. 195.


CHAPTER VII
SUMMARY AND FUTURE OUTLOOK

7.1 Summary of dissertation work

The need of having six metal-binding domains in Wilson protein — and the manner in which they are communicating with each other — is not well understood. To better understand how the last four metal-binding domains function, a detailed biophysical characterization of these domains was pursued. Molecular biology techniques were used to produce domains 3 through 6 (WLN3-6), domains 4 through 6 (WLN4-6), and mutants of domains 5 through 6 (WLN5-6). Unfolding of these constructs was performed chemically using guanidine hydrochloride (GuHCl), and this process was monitored by circular dichroism (CD). WLN5-6 unfolds with a two-state model, and half of its α-helical structure is retained at 3.2 M GuHCl. However, denaturation of WLN3-6 and WLN4-6 produces a double sigmoidal curve. In which, the first sigmoidal transition confers with the unfolding parameters of WLN5-6 while the second transition values mirror the unfolding of WLN4 alone. These experiments agree with those reported in the doctoral dissertation of Alia Hinz (Ph. D., Chemistry Department, Western Michigan University, 2014) for WLN1-4 and WLN4 alone. The thermal-unfolding midpoints, which were monitored by CD, are higher than most globular proteins by about 20 degrees, and these experiments show that both WLN3-6 and WLN4-6 are stable constructs. These observations suggest that the WLN1-3 and WLN5-6 form a unit
structure with a certain function, while WLN4 serve as a linker with and perhaps have certain function that is different from the other two constructs (units).

On the other hand, a Wilson disease-causing mutant of WLN5-6, Y48H, which is Y532H in the whole protein, was studied by both CD and femtosecond time-resolved fluorescence spectroscopy. CD data show no significant differences in the thermal and chemical unfolding relative to native WLN5-6. However, in fluorescence up-conversion experiments, the mutant relaxed significantly faster than native WLN5-6 (~10 fold) in picosecond time regime. This is attributed to the triad interaction of the three tyrosines present in the interfacial region between WLN5 and WLN6. This study showed that fluorescence up-conversion experiments probe the local protein environment near the tyrosine residues, and monitor the interfacial region and communication between domains 5 and 6 of WLN5-6.

To study the unfolding of the protein efficiently, two-photon absorption (2PA) cross-sections was used as a novel fluorescence technique to monitor the unfolding of proteins and imaging the change in its local electric field. A derivative of Green Fluorescent Protein (GFP) was used to test our hypothesis. GFP was successfully cloned into pET32 Xa LIC vector, and expressed as a fusion protein, then purified using His Tag affinity column and gel filtration technique. The unfolding of GFP by GuHCl was monitored by CD. 50 % was unfolded at around 1.6 M GuHCl. The anionic form of GFP chromophore absorbs light at 480 nm and emits light at around 508 nm. During the unfolding, the neutral form of the chromophore percentage increased which was noticed by its absorbance at 380 nm. As the protein unfolds, the chromophore fluorescence quenched due to the solvent accessibility. The fluorescence lifetime and the fluorescence
intensity both were decreasing while unfolding in a profile similar to the profile obtained by CD, except that the CD signal was increasing. The 2PA cross-sections were increasing to reach its maximum at around 1.6 M of GuHCl, and then it dropped to be lower than the initial value. This is due to the change in the local electric field around the chromophore and to the change in the chromophore form. Initially, the chromophore is in its anionic form, its dipole moment seems to align with the local electric field, and the cross-section was higher. At the initial stages of unfolding, the anionic form of chromophore experiences additional electrostatic interactions that amplify the electric field around it. However, when the protein is unfolded completely, the chromophore becomes neutral, the protein loses its beta-barrel structure, and no influence of electric fields is seen on the 2PA cross-sections. Present results have conclusively shown that the 2PA cross-sections of chromophores can be used as good markers to study the unfolding of proteins. Also, the ratio of two-photon to one-photon fluorescence can be used to image the folded, unfolded or mis-folded forms of proteins. In addition, ultrafast fluorescence measurements were able to monitor the unfolding of GFP quite effectively.

Lastly, we applied the 2PA cross-section method on two mutants of WLN5-6 after binding it to an extrinsic CPM dye. Two constructs were produced, with CPM at positions 51 and 6, WLN5-6C51_CPM and WLN5-6C6_CPM. In both constructs the anisotropy and the 2PA cross-sections increased with protein unfolding and matched the CD results. The change in the anisotropy was sharp compared to the changes based on the other methods. 2PA cross section also increases as the protein unfolds. This change is due to the change of the dipole moment of the dye which produced because of the change in the electric field. Its increasing means that the dye dipole moment is working against the
electric field in an anti-parallel orientation in the folded form. When the protein unfolds, the influence of electric field is lost and the 2PA cross-section has increased. This result suggests the validity of these novel ultrafast fluorescence methods; fluorescence lifetimes, anisotropy and 2PA cross-section can be used to monitor protein unfolding.

### 7.2 Significance

The major contribution of my research work is in revealing of the stability of WLN4 in WLN3-6 and WLN4-6 constructs of the N-terminal metal binding domain in Wilson protein where we observed that WLN4 remains its integrity in almost all constructs containing it. With this study, we were able to gain further evidence that WLN domains 1-3 and 5-6 work cooperatively as units and domain 4 connecting them. A second contribution of our research is in successfully applying ultrafast fluorescence spectroscopic techniques to probe the structural changes around tyrosines in the interfacial region between WLN5 and WLN6 in WLN5-6 and its mutants. Our results, for the first time have shown that femtosecond fluorescence spectroscopy in the UV region can be used to probe the local environment around any protein which contains tyrosine residues but not tryptophan. The results have shown the importance of having three tyrosines arranged spatially close to one another, explaining the mechanic behind fast fluorescence quenching in the disease-causing mutant Y532H.

Another important contribution of this research is developing a novel technique to monitor local electric fields in the form of 2PA cross-sections of chromophores. The influence of environment on the 2PA cross sections of chromophores was utilized to monitor local electric fields in proteins and thereby protein folding/unfolding transitions.
7.3 Future outlook

The knowledge that was gained regarding the lack of co-operativity in WLN5-6 and the fact that WLN4 maintains its integrity in all constructs are interesting results that are important to fully understand the functionality of the N-terminal metal binding domains. We believe, after this research, that the six metal binding domains interacting as two units, WLN1-3 and WLN5-6, with WLN4 acting independently. These results give a better understanding of the mechanism whereby copper acquired by the N-terminal domains.

The most important contribution of my research is the use of 2PA to monitor the unfolding of the protein. This method will have applications in monitoring changes in the local environment at a specific position. This method can also be utilized to image protein-ligand interactions, protein aggregation, and drug delivery.

7.4 References

Appendix A

Amino acid sequences of WLN5-6, WLN4-6, and WLN3-6
**WLN5-6 Sequence**

MAPQKCFQI KGMTCASCVS NIERNLQKEA GVLSVLVALM AGKAEIKYDP EVIQPLEIAQ
FIQDLGFEAA VMEDYAGSDG NIELTITGTM CASCVHNIES KLTRTNGITY ASVALATSKA
LVKFDPEIIG PRDIKIIIEE IGFHASLAQ

**WLN4-6 Sequence**

GTCSTTLIAI AGMTCASCVH SIEGMIQLE GVQQISVSLA EGTATVLYNP AVISPEELRA
AIEDMGFEAS VVSESCSTNP LGNHSAGSNM VQTDDGTPTS LOEVAPHTGR LPANHAPDIL
AKSPQSTRAV APQKCFQIK GMTCASCVSN IERNLQKEAG VLSVLVALMA GKAEIKYDPE
VIQPLEIAQF IQLDLGFEAAV MEDITGSMG IELTITGTM TASCVHNIES KLTRTNGITYA
SVALATSKAL VKFDPEIIGP RDIKIIIEE IGFHASLAQ

**WLN3-6 Sequence**

GVVTQLRVID GMHCKSCVLN IEENGQLIG VQSIQVSLEN KTAQVYDPS CTSPVALQRA
IEALPPGNFK VSLPDAEGS GTHRSSSH SPGSPPRNQV QGTCSTTLIA IAGMTCASCV
HSIEGIMISL EGVQQISVSL AEGATATVLYN PAVISPEELR AAIEDMGFEA SVVSESCSTN
PLGNSAGSN MVQTDGTPT SLQEVAPHTG RLPANHAPDI LAKSPQSTRA VAPQKCFQI
KGMTCASCVS NIERNLQKEA GVLSVLVALM AGKAEIKYDP EVIQPLEIAQ FIQDLGFEAA
VMEDYAGSDG NIELTITGTM CASCVHNIES KLTRTNGITY ASVALATSKA LVKFDPEIIG
PRDIKIIIEE IGFHASLAQ
Appendix B

Fluorescence lifetime decay and residuals of GFP at various concentrations of GuHCl
Appendix C

Lifetime components of GFP at various concentrations of GuHCl
<table>
<thead>
<tr>
<th>GuHCl (M)</th>
<th>$\tau_1$</th>
<th>$a_1$</th>
<th>$\tau_2$</th>
<th>$a_2$</th>
<th>$\tau_3$</th>
<th>$a_3$</th>
<th>Time (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>3.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.3841</td>
</tr>
<tr>
<td>0.2</td>
<td>3.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.407</td>
</tr>
<tr>
<td>0.4</td>
<td>3.4</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.3563</td>
</tr>
<tr>
<td>0.6</td>
<td>3.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.34</td>
</tr>
<tr>
<td>0.8</td>
<td>3.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3.2712</td>
</tr>
<tr>
<td>1</td>
<td>3.2</td>
<td>0.94</td>
<td>0.02</td>
<td>0.06</td>
<td>0</td>
<td>0</td>
<td>2.97915</td>
</tr>
<tr>
<td>1.2</td>
<td>3.1</td>
<td>0.85</td>
<td>0.02</td>
<td>0.15</td>
<td>0</td>
<td>0</td>
<td>2.62463</td>
</tr>
<tr>
<td>1.4</td>
<td>2.9</td>
<td>0.28</td>
<td>0.276</td>
<td>0.518</td>
<td>0.0135</td>
<td>0.206</td>
<td>0.80914</td>
</tr>
<tr>
<td>1.6</td>
<td>2.8</td>
<td>0.05</td>
<td>7.5E-4</td>
<td>0.704</td>
<td>0.0118</td>
<td>0.25</td>
<td>0.13259</td>
</tr>
<tr>
<td>1.8</td>
<td>2.8</td>
<td>0.04</td>
<td>8E-4</td>
<td>0.716</td>
<td>0.0075</td>
<td>0.241</td>
<td>0.12171</td>
</tr>
<tr>
<td>2</td>
<td>2.8</td>
<td>0.03</td>
<td>8E-4</td>
<td>0.726</td>
<td>0.0065</td>
<td>0.246</td>
<td>0.07919</td>
</tr>
<tr>
<td>2.2</td>
<td>2.8</td>
<td>0.01</td>
<td>6E-4</td>
<td>0.745</td>
<td>0.0085</td>
<td>0.255</td>
<td>0.03013</td>
</tr>
<tr>
<td>2.4</td>
<td>2.3</td>
<td>0.01</td>
<td>6E-4</td>
<td>0.713</td>
<td>0.0058</td>
<td>0.287</td>
<td>0.02491</td>
</tr>
<tr>
<td>2.6</td>
<td>2.0</td>
<td>0.01</td>
<td>7.18E-4</td>
<td>0.735</td>
<td>0.0052</td>
<td>0.265</td>
<td>0.02181</td>
</tr>
</tbody>
</table>

$\tau$: is the lifetime  
$a$: is the corresponding population
Appendix D

Project approval certification
Institutional Biosafety Committee

Project Approval Certification

For Institutional Biosafety Committee Use Only

Project Title: Characterization of Metalloproteins

Principal Investigator: David Huffman

IBC Project Number: 14DHd

Date Received by the Institutional Biosafety Committee: November 29, 2013

☑ Reviewed by the Institutional Biosafety Committee

☐ Approved
☐ Approval not required

1 [Signature]
Chair of Institutional Biosafety Committee Signature

12/13/2013
Date