

University of Iowa Iowa Research Online

Theses and Dissertations

2008

Incorporation of histidine-rich metal-binding sites onto small protein scaffolds: implications for imaging, therapeutics, and catalysis

Samantha Lynn Soebbing University of Iowa

Copyright 2008 Samantha Lynn Soebbing

This dissertation is available at Iowa Research Online: http://ir.uiowa.edu/etd/37

Recommended Citation

Soebbing, Samantha Lynn. "Incorporation of histidine-rich metal-binding sites onto small protein scaffolds: implications for imaging, therapeutics, and catalysis." PhD (Doctor of Philosophy) thesis, University of Iowa, 2008. http://ir.uiowa.edu/etd/37.

Follow this and additional works at: http://ir.uiowa.edu/etd

Part of the <u>Chemistry Commons</u>

INCORPORATION OF HISTIDINE-RICH METAL-BINDING SITES ONTO SMALL PROTEIN SCAFFOLDS: IMPLICATIONS FOR IMAGING, THERAPEUTICS, AND CATALYSIS

by Samantha Lynn Soebbing

An Abstract

Of a thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Chemistry in the Graduate College of The University of Iowa

May 2008

Thesis Supervisor: Associate Professor Sonya J. Franklin

ABSTRACT

Many histidine-rich sites in proteins bind transition metal ions such as Zn^{2+} , natively. Such sites can encourage proper protein folding or allow access to enzymatic capabilities such as hydrolysis. Ru(II) and Tc(I) also bind to aromatic amines providing access to unique chemistries not observed in biology. Ru(II) complexes have shown efficacy in fighting cancer and catalysis, while ^{99m}Tc complexes are used in radioimaging. To incorporate such metal-ions' activities into proteins, several mutants have been designed to bind Zn^{2+} , Ru^{2+} , and Tc^{1+} by introducing three histidines onto their surfaces. The first design, Z0, utilized a chimeric approach by substituting a turn in *engrailed* homeodomain for the superimposable Zn^{2+} -binding loop of astacin. In the second design, 3HT-C, three histidine residues were incorporated into the N-terminus of the Trp-cage. The final scaffold, ubiquitin, was used to make two mutants: 3HIU with a 3-histidine containing loop inserted between residues 9 and 10, and 3HPU with 3histidine point mutations near residues 35-38. Z0 proved unstable due to incorporation of a hydrophobic patch onto its surface and was not able to be isolated in sufficient quantities for study. However, the other proteins were stable and soluble. Zn^{2+} -binding by 3HT-C was investigated by intrinsic tryptophan fluorescence quenching, circular dichroism, and RP-HPLC. Binding by 3HIU and 3HPU was studied by CD. All designed proteins bind to Zn^{2+} with K_d values in the micromolar range. 3HIU and 3HPU were further studied for their ability to bind $Ru(tacn)^{2+}$ complexes. While addition of Rucomplexes caused oligomerization to various extents depending upon reaction conditions, homogeneous Ru-protein monomers were purified by a combination of size-exclusion, cation-exchange and immobilized-metal affinity chromatographies. Ru-binding was confirmed by ESI-MS, and structural integrity was investigated by CD. Results indicate that Ru(tacn)²⁺ complexes can be bound to surface binding sites in proteins without disruption of structure, opening the door for the study of catalysis in a protein context.

 99m Tc(CO)₃⁺-binding studies were performed with 3HIU by RP-HPLC. This protein binds Tc⁺ and resists substitution by free L-histidine, suggesting that peptidic Tc-binding tags could be designed with this approach to readily incorporate the radionuclide into any protein expression system.

Abstract Approved:

Thesis Supervisor

Title and Department

Date

INCORPORATION OF HISTIDINE-RICH METAL-BINDING SITES ONTO SMALL PROTEIN SCAFFOLDS: IMPLICATIONS FOR IMAGING, THERAPEUTICS AND CATALYSIS

by Samantha Lynn Soebbing

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Chemistry in the Graduate College of The University of Iowa

May 2008

Thesis Supervisor: Associate Professor Sonya J. Franklin

Copyright by

SAMANTHA LYNN SOEBBING

2008

All Rights Reserved

Graduate College The University of Iowa Iowa City, Iowa

CERTIFICATE OF APPROVAL

PH.D. THESIS

This is to certify that the Ph.D. thesis of

Samantha Lynn Soebbing

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy degree in Chemistry at the May 2008 graduation.

Thesis Committee:

Sonya J. Franklin, Thesis Supervisor

Amnon Kohen

Claudio Margulis

Jan-Uwe Rohde

Timothy Tewson

To my family



JORGE CHAM OTHE STANFORD DAILY

THANKS TO MIGUEL ...

Jorge Cham Piled Higher and Deeper Comics

ACKNOWLEDGMENTS

First, I would like to thank my thesis advisor, Sonya J. Franklin, for all her guidance and support. She has been the world's most fantastic cheer-leader/"Mom" and is always available to shed some light on confusing results. I love her reminder that, "science is what you do...it's not who you are." I owe much gratitude to Dr. Jan-Uwe Rohde for his help with Ru-compound synthesis, to Dr. Timothy Tewson for helping with the Tc-99m project, and to Dr. Claudio Margulis for setting up my PLOP calculations. These professors have allowed me to really take ownership of these projects and make them my own. Finally, I would like to thank Dr. Amnon Kohen, for being "Dad" when "Mom" couldn't be around. He also let me take a lot of responsibility when I taught for him which allowed me to really enjoy my teaching experiences. Thanks to all these professors for serving on my committee.

I must also thank my lab-mates. I remember many a time when "us girls" would find some sort of way to get into mischief. I will never forget interpretive thesis-readings with Brooke, purple *vs.* pink tape wars with Kinesha, and teaching Jennifer that "random play" translates a little differently in English. Even though not one of the Franklin-girls, Joel did his best to "fit in". Apart from many a chemistry lesson, he coerced me to blow off my work and go to Bo James that night when Don and I first got together. You don't realize how much you need your lab-mates' companionship until they are gone. Luckily for me, I was adopted by a new posse, and they have put up with all my ups and downs in the past couple years. Poor Suhash, Asif and Sai have taken part in many a game of "Slug-Ugg" on the way to get coffee when they didn't even realize they were playing. Suhash has been my partner-in-crime since we arrived at The University of Iowa and usually knows the right thing to say (although crying girls scare him). By the way, I'm still waiting for my T-shirt.

iv

On a more serious note, there are other people within the university who deserve acknowledgement. Sharon has made sure that all my paperwork was up to date so that I didn't run into trouble with the university. Tim and Andy from chemistry stores have gotten me out of many a jam (usually during their lunch hours) and still don't hold it against me. Frank is always ready to lend a hand and a smile in exchange for hugs. Dan Ferraro spent almost a year teaching me "tricks of the trade" for molecular biology work. Debra Ferraro provided wt ubiquitin DNA so that I could build new protein mutants and even let me shadow her while purifying these proteins. Rams from the biochemistry department first suggested using ubiquitin as a scaffold, and that suggestion probably saved my sanity. Outside of The University of Iowa community, I must also thank two other people. Sarah Stewart, my high-school microbiology teacher, first inspired to me pursue research as a career, and Michelle Fry, my undergraduate research advisor, first taught me the difference between experimentation in instructional laboratories (where things should always work) and actual research (where several tries are typically needed).

Everyone thanks their families, but these people really have supported me throughout my graduate school career. Mom has been chatting with St. Jude and the Blessed Virgin ever since I came to Iowa, and Dad keeps me motivated by promising to get me a puppy when I graduate. My sibs, Kristin and Dan, have been fantastic sources of distraction. The newest addition to my family, my husband, has been my backbone these last 6 years. Don watched me go through my comps and STILL married me. Even though his job and my school have kept us apart these past few years, I have never been without his support. He has been more help getting this thesis together (albeit lastminute) than anyone could probably imagine. I'll pick up my thesis-mess soon. Thanks, Babe.

v

ABSTRACT

Many histidine-rich sites in proteins bind transition metal ions such as Zn^{2+} , natively. Such sites can encourage proper protein folding or allow access to enzymatic capabilities such as hydrolysis. Ru(II) and Tc(I) also bind to aromatic amines providing access to unique chemistries not observed in biology. Ru(II) complexes have shown efficacy in fighting cancer and catalysis, while ^{99m}Tc complexes are used in radioimaging. To incorporate such metal-ions' activities into proteins, several mutants have been designed to bind Zn^{2+} , Ru^{2+} , and Tc^{1+} by introducing three histidines onto their surfaces. The first design, Z0, utilized a chimeric approach by substituting a turn in *engrailed* homeodomain for the superimposable Zn^{2+} -binding loop of astacin. In the second design, 3HT-C, three histidine residues were incorporated into the N-terminus of the Trp-cage. The final scaffold, ubiquitin, was used to make two mutants: 3HIU with a 3-histidine containing loop inserted between residues 9 and 10, and 3HPU with 3histidine point mutations near residues 35-38. Z0 proved unstable due to incorporation of a hydrophobic patch onto its surface and was not able to be isolated in sufficient quantities for study. However, the other proteins were stable and soluble. Zn^{2+} -binding by 3HT-C was investigated by intrinsic tryptophan fluorescence quenching, circular dichroism, and RP-HPLC. Binding by 3HIU and 3HPU was studied by CD. All designed proteins bind to Zn^{2+} with K_d values in the micromolar range. 3HIU and 3HPU were further studied for their ability to bind $Ru(tacn)^{2+}$ complexes. While addition of Rucomplexes caused oligomerization to various extents depending upon reaction conditions, homogeneous Ru-protein monomers were purified by a combination of size-exclusion, cation-exchange and immobilized-metal affinity chromatographies. Ru-binding was confirmed by ESI-MS, and structural integrity was investigated by CD. Results indicate that Ru(tacn)²⁺ complexes can be bound to surface binding sites in proteins without disruption of structure, opening the door for the study of catalysis in a protein context.

vi

 99m Tc(CO)₃⁺-binding studies were performed with 3HIU by RP-HPLC. This protein binds Tc⁺ and resists substitution by free L-histidine, suggesting that peptidic Tc-binding tags could be designed with this approach to readily incorporate the radionuclide into any protein expression system.

TABLE OF CONTENTS

LIST OF FIGURES	xi
LIST OF ABBREVIATIONS	.xvi
CHAPTER 1: INTRODUCTION	1
 1.1 Overview 1.2 Background	1 3 4 5 6 7 8 9 9 10 10
1.3 Thesis Goals	11
CHAPTER 2: A ZINC-BINDING CHIMERIC PROTEIN, Z0	17
 2.1 Introduction 2.2 Z0 Chimera Design 2.3 Attempts to Make Z0 2.3.1 Materials and Methods 2.3.1.1 Cloning of Z0 2.3.1.1.1 Cloning for the IMPACT System 2.3.1.1.2 Cloning for the Thermovector System 2.3.1.2 Z0 Expression and Purification 2.3.1.2.1 Z0 Purification using the IMPACT System 2.3.1.2.2 Z0 Purification using the Thermovector System 2.3.1.2.3 Z0 Purification using an Un-tagged System 2.3.2 Results and Discussion of Z0 Purification 2.3.2.1 Purification of Z0 Using the IMPACT System 2.3.2.2 Purification of Z0 Using the Thermovector System 2.3.2.3 Purification of Z0 Using the Thermovector System 2.3.2.4 Conclusions 	20 21 24 24 24 25 25 26 27 29 30 31 31 31
CHAPTER 3: METAL-BINDING TRP-CAGE PEPTIDES, 3HT-C AND 3HT-CB	50
 3.1 Introduction 3.2 Materials and Methods	50 51 51 52 52

	3.2.3.1 Fluorescence Studies	53
	3.2.3.2 CD Studies	53
	3.2.3.3 HPLC Studies	54
	3.3 Results and Discussion	54
	3.3.1 Folding and Structure of Apo-Trp-Cage Mutants	54
	3.3.2 Zn ²⁺ -binding by Modified Trp-Cages	55
	3.3.2.1 Fluorescence Studies	55
	3.3.2.2 CD Studies	57
	3.3.2.3 HPLC Studies	58
	3.4 Conclusions	59
1	Notes	65
CHAPTER 4	4: METAL-BINDING UBIQUITIN MUTANTS, 3HPU AND 3HIU	66
Ζ	4.1 Introduction	66
2	4.2 Materials and Methods	67
	4.2.1 Ubiquitin Mutant Design	67
	4.2.2 Plasmid DNA Mutations	67
	4.2.2.1 Cloning of 3HIU	68
	4.2.2.2 Cloning of 3HPU	68
	4.2.3 3HIU and 3HPU Expression and Purification	69
	4.2.4 SDS-PAGE Analysis	71
	4.2.5 ESI-MS Analysis	71
	4.2.6 Folding and Structure Preservation of Apo-3HPU	71
	4.2.7 Ubiquitin Mutant Stability	72
	4.2.7 Zn ²⁺⁻ binding by Ubiquitin Mutants	72
2	4.3 Results and Discussion	73
	4.3.1 Expression and Purification of Ubiquitin Mutants	73
	4.3.2 Stability of Ubiquitin Mutants	74
	4.3.3 Zn ²⁺ -Binding by Ubiquitin Mutants	74
2	4.4 Conclusions	76
1	Notes	83
CHAPTER 4	5 REACTIONS OF RUTHENIUM COMPLEXES WITH	
	BIOUITIN MUTANTS	84
·		
4	5.1 Introduction	84
4	5.2 Materials and Methods	85
-	5.2.1 Synthesis of Ruthenium Complexes	85
	5.2.1.1 Synthesis of $Ru(dmso)_4Cl_2 - "RuA"$	85
	5.2.1.2 Synthesis of [Ru(tacn)(dmso) ₂ Cl]Cl – "RuB"	86
	5.2.1.3 Preparation of $[Ru(tacn)(dmso)_2(OTf)]^{n+}$ – "RuC"	86
	5.2.2 Reaction of Ubiquitin Mutants with Ruthenium Compounds	87
	5.2.2.1 Method 1, Reactions in air at 37 °C	87
	5.2.2.2 Method 2, Reactions under $N_{2(9)}$ at 75 °C	88
	5.2.3 SDS-PAGE Analysis	90
	5.2.4 HPLC Analysis	90
	5.2.5 MALDI-TOF-MS Analysis	90
	5.2.6 ESI-MS Analysis	91
	5.2.7 CD Analysis.	91
	5.3 Results and Discussion	91
	5.3.1 Reactions of Ubq Mutants with Ru-Compounds, Method 1,	
	Reactions in air at 37 °C	91

5.3.2 Reactions of Ubq Mutants with Ru-Compounds, Method 2,	
Reactions under $N_{2(g)}$ at 75 °C	94
5.3.2.1 Ru-protein Purification	94
5.3.2.2 Ru-Protein Characterization	95
5.4 Conclusions	96
Notes	108
CHAPTER 6: REACTIONS OF TECHNETIUM-99M COMPLEXES WITH	
UBIQUITIN MUTANTS	110
6.1 Introduction	110
6.2 Materials and Methods	111
6.2.1 Synthesis of Tc ¹ (CO) ₂ (H ₂ O) ₂	111
$6.2.2$ Reactions of $Tc^{1}(CO)_{2}(H_{2}O)_{2}$ with Proteins	111
6.2.2 Helections of Te (CO) ₃ (11 ₂ 0) ₃ with Froteins	111
6.3 Results and Discussion	112
6.3.1 Synthesis of Tc ¹ (CO) ₂ (H ₂ O) ₂	112
6.3.2 Reactions of $Tc^{I}(CO)_{2}(H_{2}O)_{2}$ with Proteins	112
6.3.2 Reactions of Te (CO)3(Tr2O)3 with Trotenis	113
6.2.2.2 Kinotics Assays	115
6.4 Conclusions	116
0.4 Conclusions	110
Notes	123
	104
CHAPTER /: SUMMARY AND FUTURE DIRECTIONS	124
	104
7.1 Zn-Binding by Protein Designs	124
7.1.1 Engrailed Homeodomain/Astacin Chimera: Z0	124
7.1.2 Trp-cage Mutants: 3HT-C and 3HT-Cb	124
7.1.3 Ubiquitin Mutants: 3HIU and 3HPU	125
7.2 Ruthenium-Binding by Ubiquitin Mutants	126
7.3 Technetium-Binding by Ubiquitin Mutants	127
BIBLIOGRAPHY	129

LIST OF FIGURES

Figure 1.1	Methodology used by the Franklin Research group to design chimeric proteins. An EF hand of calmodulin is shown at left in purple. <i>Engrailed</i> homeodomain is shown at right in green. Overlays were created with Swiss PDB viewer
Figure 1.2	Evolution of designs in the Franklin group. All proteins are chimeras with the <i>engrailed</i> homeodomain scaffold. Proteins shown in the top row bind to lanthanides <i>via</i> the Ca ²⁺ -binding loop of calmodulin. In the lower left, a peptide designed to bind Cu ²⁺ as a chimera with the prion protein octarepeat sequence is shown. Chapter 2 of this thesis will describe attempts to make Z0, whose structure is shown at lower right. La-P3W and La-C2 are shown as NMR structures while the rest of the chimeras are shown as models13
Figure 1.3	Two views of <i>engrailed</i> homeodomain (1HDD). At top, the protein is shown bound to DNA, and at the bottom, it is shown alone. Helices 1, 2, and 3 are labeled in both structures
Figure 1.4	Two views of the Trp–cage peptide (1L2Y). At top, a ribbon diagram with labeled secondary structure elements is shown. The caged Trp25 residue is shown in pink along with other caging residues. Tyr22 is shown in purple, Gly30 is shown in dark blue, and Pro31, Pro36, Pro37, Pro38 are shown in mint. At bottom, a space-filling diagram in the same orientation is shown with the caged Trp residue shown in pink
Figure 1.5	Ribbon diagrams of wt ubiquitin (1UBQ) and Gly ₈ insertion mutants. The top picture shows wt ubiquitin with the 9-10 loop colored green and the 35-36 loop colored red. The bottom two pictures are overlays of wt Ubq (blue) and Gly ₈ -insertion mutants (green). Structural results due to insertion at the 9-10 loop are shown at left, and results due to insertion at the 35-36 loop are shown at right
Figure 2.1	Proposed ribbon structure of Z0 (center) as constructed from <i>engrailed</i> homeodomain (left) and astacin (right). The DNA- binding motif is shown in green while the Zn^{2+} -binding site is shown in yellow. The zinc ion is shown as a grey sphere bound to three histidine residues also shown in yellow
Figure 2.2	Ribbon diagrams showing deviation of astacin's β -turn from meeting helix-3 of the homeodomain if a residue is added or subtracted from the N-terminus of the substituted sequence. The middle figure shows the correct alignment of 2° structures while the top and bottom figures show resulting models due to subtraction (top) or addition (bottom) to the N-terminus of the exchanged sequence. The homeodomain motif is shown in green while astacin's Zn ²⁺ -binding site is shown in yellow. The zinc ion is shown as a grey sphere bound to three histidine residues also shown in yellow. Helix-1 of <i>engrailed</i> has been omitted for clarity37

Figure 2.3	Two ribbon diagram views showing the overlay of <i>engrailed</i> homeodomain (1ENH) and the Zn^{2+} -binding site of astacin (1AST) when sequence Z0 is used as shown at the right. Conserved residues from the homeodomain motif are shown in green while those that are substituted by asticin are shown in grey. Astacin's Zn^{2+} -binding site is shown in yellow along with its zinc-binding histidines. Helices are numbered $1 - 3$, and the zinc ion is shown as a grey sphere. Protein sequences are shown in one-letter amino acid abbreviations
Figure 2.4	PLOP results from calculations with Z0. Residues from the homeodomain motif are shown in green while astacin's Zn^{2+} -binding site is shown in yellow or pink. Yellow residues were constrained while those in pink (shown in input structure only) were not. Zn^{2+} -binding histidines are shown in blue (shown in input and first two outputs only).
Figure 2.5	Representations of Z0 and <i>engrailed</i> homeodomain comparing a structural view with that of one showing polarity assignments. In the top-left, structural view, homeodomain residues are shown in green while astacin's Zn^{2+} -binding site is shown in yellow. In the middle two views, Z0 residues are labeled by type where basic residues are blue, acidic are red, polar are yellow, and non-polar are grey. Helices are numbered $1-3$, and the zinc ion is shown as a grey sphere. In the bottom-right, polarity-labeled view, <i>engrailed</i> homeodomain is shown for comparison
Figure 2.6	SDS-PAGE analysis of Z0-CBD expression using the IMPACT system. The arrow indicates where Z0-CBD should appear in gel lanes according to the molecular weight standards shown in the middle lane
Figure 2.7	SDS-PAGE analysis of Z0-ftr expression using the Thermovector system. The arrow indicates where Z0-ftr appears in gel lanes (in agreement with molecular weight standards shown in the right-most lane). The first five lanes show soluble lysate heated to 50°C in the presence of 5% PEG 4000, 200 mM NaCl over time. The amount of time which the sample was heated increases from left to right
Figure 2.8	SDS-PAGE analysis of Z0-ftr purification using the Thermovector system. The arrows indicate where Z0-ftr, <i>ftr</i> , His-TEV, and Z0 appear in gel lanes (in agreement with molecular weight standards shown in the left-most lane). Lanes 1-7 show progression of Z0-ftr purification using IMAC while lane 8 shows protein species present after the reaction of Z0-ftr with His-TEV
Figure 2.9	SDS-PAGE analysis of Z0 purification using the Thermovector system and a second round of IMAC. The arrows indicate where Z0-ftr, <i>ftr</i> , His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the left-most lane. Column washes are shown in the next three lanes. Proteins eluted with imidazole are shown in the next lanes where imidazole concentration increases from left to right

Figure 2.10	SDS-PAGE analysis of Z0 purification using the Thermovector system and cation exchange. The arrows indicate where Z0-ftr, <i>ftr</i> , His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the right-most lane. Proteins eluted with NaCl are shown in the next lanes where [NaCl] increases from left to right
Figure 2.11	SDS-PAGE analysis of Z0 purification using the Thermovector system and size exclusion. The arrows indicate where Z0-ftr, <i>ftr</i> , His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the right-most lane. Proteins were eluted from the column over time as shown in lanes from left to right
Figure 2.12	SDS-PAGE analysis of Z0 purification using an untagged system. The arrow indicates where Z0 appears in gel lanes. Purified Z0 can be seen in the right-most lane. All samples are labeled according to how they were purified, and lanes 1 and 8 show identical samples47
Figure 3.1	Proposed peptide structure for Trp-cage mutants. The sequences of mutants can be compared to that of "wt" Trp-cage, TC5b. Peptide sequences are shown is single-letter amino acid abbreviations
Figure 3.2	Titrations of Trp-cage mutants monitored by dependence of fluorescence emission at 350 nm upon addition of Zn^{2+} . Top: titration of 3HT-C. Bottom: titration of 3HT-Cb. Raw data is shown as closed circles and non-linear least squares fittings are shown as dashed lines. Error bars for the Zn^{2+} titration of 3HT-C were generated from calculation of the standard deviation of data from two separate trials. No error bars were generated for the Zn^{2+} titration of 3HT-Cb since only one trial was performed. Insets show representative fluorescence spectra
Figure 3.3	Representative titration of 3HT-Cb monitored by dependence of molar residue ellipticity at 222 nm upon addition of Zn^{2+} . Raw data is shown as closed circles, and the non-linear least squares fitting is shown as a dashed line. Error bars were generated by calculating the instrument noise at 222 nm. Inset shows CD spectra
Figure 3.4	HPLC analysis of absorbance at 280 nm of Zn^{2+} -binding by 3HT-C. Apo-peptide is shown in red while Zn^{2+} -tritrated peptide (10 equivalents of metal ion) is shown in blue
Figure 4.1	Proposed structures of 3HIU (top) and 3HPU (bottom). Residues conserved from wt ubiquitin are shown in green while mutated residues are shown in yellow. The sequences at the right are shown using single-letter amino acid abbreviations
Figure 4.2	Representative SDS-PAGE analysis of 3HPU purification. The arrow indicates where 3HPU appears in gel lanes. Purified 3HPU can be seen in the right-most lane. All samples are labeled according to how they were purified. 3HIU purifies similarly79

Figure 4.3	Deconvoluted ESI-MS data collected for 3HIU (top) and 3HPU (bottom). Arrows indicate positions of apo-protein peaks and their apparent masses
Figure 4.4	CD spectra of ubiquitin mutant melting studies. Top: melting of 3HIU. Bottom: melting of 3HPU
Figure 4.5	Representative titrations of Ubq mutants monitored by dependence of molar residue ellipticity at 207 nm upon addition of Zn^{2+} . Top: titration of 3HIU. Bottom: titration of 3HPU. Raw data is shown as closed circles, and non-linear least squares fittings are shown as dashed lines. Error bars were generated by calculating the instrument noise at 207 nm. Insets show CD spectra
Figure 5.1	UV-Vis Spectra of Ru-Ubq mutant reactions. Panel A shows ΔA at 450 or 460 nm during reaction progression, and Panel B shows the final spectra. Red curves are reactions with 3HIU and blue curves are reactions with 3HPU. Circles are reactions with RuB and squares are reactions with RuC. Panel C shows representative spectra for reactions of 3HIU with RuA, RuB, and RuC from left to right. Reactions with 3HPU produce similar spectra
Figure 5.2	SDS-PAGE Analysis of Ru-Ubq mutant reactions. Samples in the gel on the left were not completely denatured while samples in the gel on the right were heated extensively. Differences in gel color are due to different staining techniques used: silver on the left and Coomassie on the right. The arrow indicates where Ubq mutant monomers appear in gel lanes
Figure 5.3	Mobility analysis of species present after reactions of 3HIU and 3HPU with Ru ^{II} (tacn) complexes <i>vs.</i> protein standards101
Figure 5.4	HPLC A ₂₁₄ chromatograms for reactions of RuB and RuC with Ubq mutants. Reactions with 3HIU are shown at top while those with 3HPU are shown at the bottom. Apo-protein in shown in black while reactions with RuB are shown in red, and reactions with RuC are shown in blue
Figure 5.5	CD Spectra of unpurified Ru + Ubq mutant reactions performed in air at 37 °C for 34 days
Figure 5.6	UV-Vis Spectra of Ru-Ubq mutant reactions. Panel A shows ΔA during reaction progression with ΔA_{320} on the left and ΔA_{460} on the right. Panel B shows the final spectra of all reactions. In the upper two panels, $3HIU + RuB$ is shown in red, $3HIU + RuC$ in gold, and $3HPU + RuC$ in blue. Panel C shows raw data for reactions of Ru compounds with Ubq mutants
Figure 5.7	Representative SDS-PAGE analysis of RuB-3HIU purification. The arrow indicates where 3HIU monomers appear in gel lanes. RuC-3HIU and RuC-3HPU purified similarly

Figure 5.8	ESI-MS analysis of purified Ru-protein products. 3HIU + RuB is shown at top, 3HIU + RuC in the middle and 3HPU + RuC at the bottom. Arrows indicate positions of Ru(tacn)-Ubq mutant species and their apparent masses.	106
Figure 5.9	CD Spectra of purified Ru + Ubq mutant reactions performed at 75 °C for 6 hours in the presence of limited oxygen	107
Figure 6.1	Radioactivity HPLC chromatograms using Method I showing production of $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ from TcO_{4}^{-} generator eluate. The complexation product with L-histidine is also shown.	118
Figure 6.2	HPLC Analysis of ^{99m} Tc complexation products with 3HIU, 3HPU, and 3HT-Cb. All reactions with proteins are shown as blue curves with data points. Data points for the reaction with 3HIU are shown at squares, with 3HPU as diamonds and with 3HT-Cb as circles	119
Figure 6.3	Initial L-His competition assays with designed proteins. Tc-3HIU is shown at the top, Tc-3HPU is shown in the middle, and Tc-3HT-Cb is shown at the bottom.	120
Figure 6.4	L-His competition assays with 3HIU (top) and wt Ubq (bottom)	121
Figure 6.5	Timed binding assays for $Tc^{I}(CO)_{3}$ + 3HIU reactions (top), + wt Ubq (middle) and + L-His (bottom)	122

LIST OF ABBREVIATIONS

1°	primary
2°	secondary
2D	two dimensional
3°	tertiary
А	adenosine
Ala or A	Alanine
Amp	ampicillin
Apo-	co-factor-(i.e. metal ion) free protein
Asn or N	Asparagine
Arg or R	Arginine
Asp or D	Aspartate
b.p.	DNA base pairs
С	cytidine
CBD	chitin-binding domain
CD	circular dichroism
CHCA	α -cyano-4-hydroxycinnamic acid
CIAP	calf intestinal alkaline phosphatase
COSY	correlation spectroscopy
CV	column volume(s)
Cys or C	Cysteine
Da	Dalton (equivalent to g/mol)
ddi	doubly distilled deionized

DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DQF	double quantum filtered
DSS	2,2-dimethyl-2-silapentane-5-sulfomate
DTT	1,4-dithiothreitol
EDTA	ethylenediaminetetraacetic acid
ESI-MS	electrospray ionization-mass spectrometry
fac-	facial coordination
FPLC TM	Fast Protein Liquid Chromatography (GE Healthcare)
g	standard acceleration due to gravity at surface of Earth
G	guanosine
Gln or Q	Glutamine
Glu or E	Glutamate
Gly or G	Glycine
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
His or H	Histidine
Holo-	protein with bound co-factor (i.e. metal ion)
HPLC	high-performance liquid chromatography
IDT	Integrated DNA Technologies
Ile or I	Isoleucine
IMAC	immobilized metal ion affinity chromatography
IPTG	isopropyl-β-D-thiogalactopyranoside
K _d	dissociation constant $(^{1}/K_{a})$

LB	Luria-Bertani
Leu or L	Leucine
Lys or K	Lysine
MALDI-TOF	matrix-assisted desorption/ionization time-of-flight
MES	2-(N-morpholino)ethanesulfonic acid
Met or M	Methionine
MOPS	3-(N-Morpholino)propanesulfonic acid
NMR	nuclear magnetic resonance
NMWL	nominal molecular weight limit
NOESY	nuclear Overhauser effect spectroscopy
NTA	nitrilotriacetic acid
OD ₆₀₀	optical density at 600 nm
OTf	trifluoromethanesulfonate anion (triflate)
PAGE	polyacrylamide gel electrophoresis
PCR	polymerase chain reaction
PDB	protein data bank
PEG	polyethylene glycol
Phe or F	Phenylalanine
PLOP	Protein Local Optimization Program
PMSF	phenylmethylsulfonyl fluoride
Pro or P	Proline
psi	pounds per square inch
SDS	sodium dodecyl sulfate

Ser or S	Serine
Т	thymidine
T _m	melting point (or temperature at which a protein denatures)
tacn	triazocyclononane
TEAP	triethylammonium phosphate
TEV	tobacco etch virus protease
TFA	trifluoroacetic acid
TFE	2,2,2-trifluoroethanol
Thr or T	Threonine
TOCSY	total correlation spectrosopy
Tris	trishydroxymethylaminomethane
Trp or W	Tryptophan
Ubq	ubiquitin
UV	ultraviolet light
Val or V	Valine
Vis	visible light
WATERGATE	water suppression by tailored excitation
wt	wild type

CHAPTER 1: INTRODUCTION

1.1 Overview

While it may seem that this thesis is a compilation of many unrelated experiments, it is the author's intention to show how each of these projects led to the next. Previous work in the Franklin research group (The University of Iowa, Dept. of Chemistry) dealt with *de novo* designs of chimeric proteins where super-imposable structural motifs from unrelated proteins could be exchanged to build a new protein capable of performing the functions of each parent. Building a chimera was where this research may have begun, but it led into research of metal-binding proteins capable of new activities. By building surface-accessible metal-binding sites capable of binding reactive or radioactive ions, proteins with applications in imaging, therapeutics and catalysis could be made.

Initially, the research goal was to design a protein capable of bringing a metal ion capable of performing hydrolysis into close proximity of DNA to act as an artificial endonuclease. A Zn^{2+} and DNA-binding chimeric protein (Z0) was designed by incorporating a Zn^{2+} -binding site from astacin into *engrailed* homeodomain. While working with this protein, a new goal emerged. Histidine residues can bind Zn^{2+} ions as well as a host of other metal ions. Why only cleave DNA when such a chimera could be used to make a library of protein devices capable of different activities with DNA when bound to different metal ions? Metal-substitution is a classic approach in bioinorganic chemistry to study metalloproteins. Use of alternate metal ions allows for incorporation of spectroscopic handles and exploration of new chemistries performed in protein contexts. However, very little research has been performed with substitutions by the more inert, second- and third-row transition metals despite their rich spectroscopic and chemical properties. This deficiency is predominantly due to the difficulty in substituting such metals into proteins. Such metal-binding sites must be surface-accessible and pre-

organized in order to facilitate binding by these inert metals and inhibit protein misfolding caused by formation of incorrectly metal-bound, yet kinetically-stable complexes.

The first inspiration for this new direction was cisplatin and new rutheniumcontaining chemotherapeutics. If a Ru^{II} or Ru^{III} ion could be brought into close proximity of DNA, irreversible cross links could be made. Rather than excising unwanted DNA codes, they could be rendered inactive by blocking them with a kinetically inert metalprotein complex. Coincidently, the Zn-binding site in Z0 contained three histidine residues that also have high affinity for Ru^{II} and Ru^{III} ions. It was postulated that such 3-His metal-binding sites would also be capable of binding facially-coordinated ^{99m}Tc⁺ ions. Binding a radioactive metal to Z0 could allow delivery of a radio-imaging moiety to specific DNA targets. One protein could have several different activities and applications depending upon the metal ion bound.

Ultimately, Z0 proved unstable and insoluble. The homeodomain motif was not as robust a scaffold as initially thought. Rather than try to optimize such a system, new scaffolds were chosen. One scaffold was a small peptide called the Trp-cage. This peptide was approximately a third the size of Z0. A mutant of this small biomolecule was chosen as a model metal-binder and named 3HT-C. The other scaffold chosen was ubiquitin; it is only slightly larger than Z0 and extremely stable. Initial studies on the effect of loop mutations upon structure had already been performed in the Robertson Laboratory (The University of Iowa, Dept. of Biochemistry). From this work, two mutants (3HIU and 3HPU) were designed. All mutants introduced three histidine residues intended to bind to ions in a facial manner.

After demonstrating Zn^{2+} binding by these mutants, work began in order to show their Ru^{II} and Tc^{I-} binding abilities as well. Unlike with Zn^{2+} , these ions cannot be added to proteins as stable, water-soluble, chloride salts. Therefore, inorganic synthesis was used to make facially-coordinated Ru^{II} (tacn) and $Tc^{I}(CO)_{3}$ complexes.

2

While the original intention of this work had been to build a chimera able to bring Zn²⁺ or kinetically-inert, second-row transition metals into close proximity of DNA, this goal could not be met due to chimera instability. Instead, new mutants were developed to serve as proof-of-principle that Ru^{II} and Tc^I-binding sites could be built into proteins. Knowledge gained from this work could aid *de novo* design of metalloproteins with applications in medical imaging, therapeutics and catalysis.

1.2 Background

Roles for biological macromolecules are continuously discovered and their mechanisms of action elucidated. One such class of bio-macromolecules, proteins, has long been studied in order to determine their various physiological functions. They convey information and perform chemistries based not only upon their chemical formulas, but upon how their sequences dictate the overall structure of such molecules. This level of complexity affords proteins the ability to perform functions with high specificity unattainable by small molecules. Studying proteins in their native contexts has greatly advanced protein research, but drawing conclusions from observations made when performing experiments amidst a biological milieu is extremely difficult. One area of research aims to learn about protein structure-function relationships by designing new sequences and studying the structural and/or functional results. This area of enzyme design research has attracted many researchers because of the biochemical, pharmacological, and catalysis applications.¹⁻⁴

1.2.1 De Novo Peptide and Protein Design

DeGrado *et.al.* define *de novo* design (in its purest form) as studies entailing, "the construction of a protein, intended to fold into a precisely defined 3-dimensional structure, with a sequence that is not directly related to that of any natural protein," but this term has been used to describe many design strategies.⁵ The aim is to test knowledge of relationships between proteins' primary, secondary, and tertiary structures by

designing protein sequences to fold into pre-determined topologies.^{6,7} While much has been learned from protein engineering, no universal theories for structure prediction from primary sequences has emerged.

As of 2005, 31830 protein structures had been deposited in the protein data bank. Of those structures, only approximately 1000 unique folds were found.⁸ Nature recycles scaffolds and makes small changes to tune for various specificities and reactivities. This fact has not gone unnoticed. While some researchers design proteins completely from first principles, others re-design known proteins.

1.2.1.1 DNA-Binding Proteins

As noted, one scaffold used extensively in the Franklin laboratory is a DNAbinding protein. A common theme in DNA-binding proteins is the interaction of an α helix with the major groove. The common DNA-binding motifs, homeodomains, zinc fingers, basic helix-loop-helix motifs and leucine zippers find their DNA targets *via* such interactions.⁹⁻¹¹ The most common DNA-binding motif, the zinc finger, binds DNA *via* an α -helix and has been studied extensively. Through engineering various mutants and studying their DNA-binding specificities, a recognition table has been developed.¹² Approaching design of DNA-binding proteins from another direction, Schepartz, *et. al.* (Yale University, Dept. of Molecular, Cellular and Developmental Biology), have transformed a non-DNA-binding peptide, the avian pancreatic polypeptide, into a DNAbinding motif.¹³ By grafting base-recognizing amino acids from helix-3 of *engrailed* homeodomain onto this scaffold, they were able to make a mutant able to bind DNA with high affinity and selectivity.

Some DNA-binding proteins perform hydrolysis reactions upon DNA in order to change its coded information. Such code changes are necessary for destruction of foreign DNA and excision or repair of detrimental mutations.¹⁴ Hydrolytic cleavage is the preferred mechanism used in biology because, unlike oxidation reactions, hydrolysis does

not produce potentially toxic radicals. Many of these enzymes use non-redox active metal ions such as Zn^{2+} , Ca^{2+} , and Mg^{2+} to perform their chemistry.^{15,16}

1.2.1.2 Metalloproteins

Due to the remarkable breadth of chemistry available to metalloproteins, approximately a third of all natural proteins need metals to function.¹⁷ Roles for metal ions in proteins include such actions as nucleophilic catalysts, (*i.e.* Zn²⁺ in carbonic anhydrase), electron transfer agents ($\text{Fe}^{2+/3+}$ in rebredoxin), and structure stabilizers (Zn^{2+} in zinc fingers).¹⁸ Yi Lu (The University of Illinois, Urbana-Champaign, Dept. of Chemistry) has done extensive work redesigning copper-binding proteins to determine relationships between type I and type II sites.^{8,19-21} Further research with copper sites in the Dennison laboratory (Newcastle University, Institute for Cell and Molecular Biosciences) indicates the exchanged copper-binding loops maintain their native conformations and chemical properties when placed in a new scaffold.²² Recent work in the Pecararo Laboratory (University of Michigan, Dept. of Chemistry) with helix-bundle proteins has attempted to determine whether pre-organization of metal binding sites is necessary or if metal ions are capable of nucleating structural elements.¹⁷ In his designs, it was found that a highly-stable binding site was not necessary for Hg^{2+} -binding. Pecararo also used helix-bundle motifs to study the effect of surrounding-non-binding residues upon ligand sets. The number of metal-binding ligands could be controlled by changing the size of surrounding residues and thereby affecting sterics in the binding pocket.²³ All the proteins prepared in this thesis are Zn²⁺, Ru^{II}, and/or Tc^I-binding proteins.

1.2.1.2.1 Zinc-Binding Protiens

After iron, zinc is the second-most abundant transition metal found in biology.²⁴ There are a few hundred Zn^{2+} -enzymes and more than 1000 Zn^{2+} -binding domains in the mammalian proteome.²⁵ Zinc ions are used ubiquitously in nature and generalizations

about Zn-binding sites have been drawn. While there are exceptions to every rule, typically, structural Zn^{2+} -binding sites have four ligands, and catalytic ones have only three (with the fourth site occupied by water).²⁶ Structural sites tend to have cysteine (thiol) ligands while catalytic sites contain histidine (imidazole), aspartate and glutamate (carboxylate) residues.

Zn-fingers are DNA-binding proteins that use their Cys_2His_2 -bound Zn^{2+} ions to stabilize their structure. Sugiura has extensively studied Zn-finger binding sites and designed mutants able to catalytically cleave DNA.^{27,28} In some designs, the cysteine residues were mutated to histidine residues to afford a reactive complex. In other designs, an "open" site was introduced by removing a ligand from the Zn^{2+} -binding coordination sphere. Conversely, the catalytic site of carbonic anhydrase can be inactivated by a mutation introducing a fourth Zn^{2+} -ligand, and thereby closing the site.²⁹

1.2.1.2.2 Ruthenium-Binding Proteins

When added to biological systems, ruthenium ions bind tightly to plasma proteins (such as albumin and transferrin) with a preference for surface histidines and thiols.^{30,31} Transferrin is able to bind two equivalents of the Ru-pharmaceutical, KP1019, while albumin binds five equivalents. Furthermore, in pre-clinical studies, 80-90% of the administered pharmaceutical was found bound to albumin while a much smaller amount was found bound to transferrin.

Ghadiri has performed much research with *de novo* design of Ru-binding helix bundle proteins. One of the earliest designs introduced 2-histidine residues on the same face of an α -helix. NMR evidence showed that added Ru^{II} ions preferentially bonded to these imidazole groups (even in the presence of competing glutamate and lysine residues). Metal-binding greatly stabilized the peptide and increased the melting temperature by 25 °C.³² Later work in his lab demonstrated metal-ion induced selfassembly of proteins.³³⁻³⁵ In some designs, 5-carboxy-2,2'-bipyridyl functional groups were coupled to the N-termini of α -helices. Upon binding Ru^{II}, triple-helix bundles were formed. Interestingly, in its apo-form, this protein forms dimers. In other designs, pyridyl functional groups were attached to helices. Instead of triple-helix bundles, Ru^{II}-binding caused formation of 4-helix bundles. In all designs, incorporation of Ru^{II} greatly stabilized the overall tertiary structure.

In *de novo* design of Ru-binding proteins using cytochrome C as a scaffold, a pair of histidine residues (separated by nineteen amino acids in sequence, but near in tertiary structure) was introduced onto the protein's surface.³⁶ By binding a Ru^{II}(bpy)₂ complex, the protein structure was greatly stabilized and the melting temperature raised by 23 °C. In contrast, when the Ru^{II} complex was bound to only one introduced histidine residue, no stabilization was observed. Cross-linking of specific protein residues by Rucomplexes provides stabilization of proteins much like formation of disulfide bridges found naturally in proteins.

Another use of Ru^{II/III}-bound proteins is in the study of electron transfer in proteins. In one such study, ruthenium was tethered *via* a histidine residue to a helix dimer/bundle to model electron transfer.³⁷ In another system, Ru^{II/III}(NH₃)₅ was appended to various histidines introduced onto the surface of myoglobin. Correlations were drawn showing the dependence of electron transfer efficiency between the heme group and Ru-ion upon distance.³⁸

1.2.1.2.3 Technetium-Binding Peptides

For use in medical imaging, Tc-chelates must meet several criteria: (1) have extremely high binding affinities so that the Tc-ions remain bound even at very low concentrations, (2) provide homogeneous products in high yields, (3) stabilize the metalion under *in vivo* conditions, and (4) form an inactive complex that does not affect physiological functions.³⁹ Currently, work to design Tc-binding biomolecules for pharmaceutical applications has been limited to studies with antibodies and small peptides. One such Tc¹(CO)₃-antibody complex (with Tc-binding performed by a 6-His tag) even went into clinical trials.⁴⁰ In another study, a bombesin analogue was coupled to a Tc-chelate complex.⁴¹ The chelate selected was the unnatural amino acid, N^{α}-histidine, capable of tridentate binding to metal ions. A mono-dentate histidine analogue was also studied, but binding was not stable to substitution by free L-histidine. Despite high receptor affinity and rapid plasma clearance, this design had limited utility because it was degraded under physiological conditions. To specifically target cellular compartments, cell-penetrating peptides (typically < 30 amino acids) have been coupled to Tc-chelates.⁴² Such complexes were successfully internalized into HeLa cells.

1.2.1.3 Previous Work in Chimeric-Designed Proteins

As stated previously, the design approach used in the Franklin laboratory was to exchange super-imposable loops from different proteins to make a new mutant capable of performing the functions of both parents. This approach was called chimeric design or modular turn substitution. When designing initial chimeras, the following similarities were noted: (1) the turn connecting helices 2 and 3 of *engrailed* homeodomain was geometrically super-imposable to the EF hand, Ca^{2+} -binding site from calmodulin, and (2) calcium binding sites are capable of binding to lanthanide ions very tightly. If the calmodulin metal-binding loop was substituted for the turn in the homeodomain, and a lanthanide was bound instead of Ca^{2+} , a new endonuclease could be built from other proteins just as the chimera monster of mythology was built from other beasts (**Figure 1.1**). The homeodomain motif was the DNA-binding "lion's head", the loop from calmodulin was the metal-binding "goat" middle and the lanthanide was the hydrolytic "snake". Hence, the protein was designed to bring a metal-ion capable of activating a bound water molecule into close proximity to DNA in order to promote cleave of its phosphodiester bonds.

As can be seen in **Figure 1.2**, chimeric protein design in the Franklin laboratory has evolved. In the beginning, chimeras were built only from the Ca²⁺-binding loop from calmodulin substituted for the turn between helices 2 and 3 of *engrailed* homeodomain. These peptides were very small (33 amino acids) but capable of sequence selective cleavage of DNA when bound to one equivalent of lanthanide.⁴³⁻⁵⁰ Further work aimed at tuning the hydrolytic activity of the bound lanthanide by changing its coordination environment.⁵¹ In the next generation, helix 1 of the homeodomain motif was added onto the chimera, thereby almost doubling the protein size to ~60 amino acids.^{52,53} In another doubling of chimera size in generation 3, a homo-dimer of the entire metallohomeodomain was designed.⁵⁴ While the 60 amino acid monomer showed only a modest improvement of activity compared to the smaller, 33 amino acid peptides, the dimer showed incredible hydrolytic activity and sequence-selective binding.

In another branch of the evolutionary tree, transition metal-binding sites were designed into *engrailed* homeodomain. In one successful design, the octarepeat Cu^{2+} -binding loop from the prion protein was substituted for the same homeodomain turn described above to make a 28-mer peptide.⁵⁵⁻⁵⁷ While DNA-binding and cleavage assays were not performed with this chimera, it did show exceptional specificity for Cu^{2+} . This work with transition metal-binding chimeras led to the design of a Zn^{2+} -binding chimera (Z0) incorporating the binding site from astacin into *engrailed* homeodomain. This protein design will be described in chapter 2.

1.2.2 Scaffolds Used in this Research

1.2.2.1 Engrailed homeodomain

Homeodomains make attractive scaffolds for protein engineering because of their relative simplicity. They are small in size (60 amino acids), have only one element of secondary structure (α -helices), do not have any disulfide bonds, do not need any cofactors to function, and operate as monomers. They are held together by a conserved

hydrophobic core made up of Leu16, Phe20, Leu34, Lue40, Ile45, Trp48, and Phe49 located at the interface of its three helices.^{58,59} Helix 1 is oriented anti-parallel to helix 2 while helix 3 binds to the major groove of DNA (**Figure 1.3**). When compared to proteins of similar size, homeodomains are less stable. However, upon binding to DNA, their disordered termini attain structure.⁶⁰ Another source of instability is that helix-3 is highly positively charged leading to repulsion between residues. DNA-binding balances these charges.⁶⁰

1.2.2.2 Trp-Cage

The Trp-cage is a 20 residue peptide optimized from exendin-4, a 39 amino acid peptide found in Gila monster saliva (**Figure 1.4**).⁶¹ It gets its name from its unique fold in which a tryptophan residue is encapsulated by prolines. In order to stabilize this peptide, the N-terminus of exendin-4 was removed, a salt bridge introduced by making point mutations, and a Phe22Tyr mutation was performed to aid in hydrophobic packing. Residues 21 – 27 (exendin-4 numbering) are α -helical. Residues 30-33 comprise a small 3₁₀ helix, and the remaining C-terminal residues are part of a PPII helix. The original designers described Trp-cage as very stable, and 90 – 95% of the peptide is correctly folded at low temperatures.⁶² It's T_m is between 42 – 46 °C at neutral pH.^{61,62} While the folding mechanism of Trp-cage is not fully understood, most research indicates a cooperative, 2-state folding mechanisms. The peptide is most stable at 20 °C, and much of its structural elements survive at high temperatures.⁶³

1.2.2.3 Ubiquitin

Ubiquitin is an extremely stable, small protein (76 amino acids) used in signaling networks. In work performed in the Robertson laboratory (The University of Iowa, Dept. of Biochemistry), the effect of loop insertion mutations upon ubiquitin structure was investigated (**Figure 1.5**).^{64,65} They chose ubiquitin for their model because the protein is very stable and well characterized. They found that the site of the insertion mutation

had a greater impact on protein structure than the sequence of the inserted loop. Specifically, insertions between residues 9 and 10 had very little impact on the overall protein geometry while insertions between residues 35 and 36 caused subtle structure changes throughout the protein. They hypothesized that the 9-10 loop can better accommodate insertions due to its greater flexibility and localized stabilizing factors. In contrast, the 35-36 loop is more extended, stabilized by more global protein interactions, and is closer to the protein's hydrophobic core. Of note was the finding that the veryflexible Gly₈-loops (with 8 sequential glycine residues inserted) provided the most stable mutants for each insertion site.⁶⁵

1.3 Thesis Goals

In chapters 2 - 4, the design of 3-His, metal-binding proteins will be described. Since the original goal of this work was to design Zn^{2+} -proteins, their metal-binding affinity will be demonstrated by Zn^{2+} -binding studies. In chapter 5, synthesis and binding of Ru^{II} complexes will be described, and in chapter 6, synthesis and binding of a ^{99m}Tc^I-complex will be described. Almost all prior research of Ru^{II} and Tc^I-binding proteins/peptides utilizes non-natural chelates conjugated to biomolecules. This work serves as proof of principle that proteins composed completely of naturally-occurring amino acids can be designed to stably bind to non-physiological metal-ions.


Figure 1.1 Methodology used by the Franklin Research group to design chimeric proteins.⁵¹ An EF hand of calmodulin is shown at left in purple. *Engrailed* homeodomain is shown at right in green. Overlays were created with Swiss PDB viewer.



Figure 1.2 Evolution of designs in the Franklin group. All proteins are chimeras with the *engrailed* homeodomain scaffold. Proteins shown in the top row bind to lanthanides *via* the Ca²⁺-binding loop of calmodulin. In the lower left, a peptide designed to bind Cu²⁺ as a chimera with the prion protein octarepeat sequence is shown. Chapter 2 of this thesis will describe attempts to make Z0, whose structure is shown at lower right. La-P3W and La-C2 are shown as NMR structures^{47,53} while the rest of the chimeras are shown as models.^{54,55}



Figure 1.3 Two views of *engrailed* homeodomain (1HDD). At top, the protein is shown bound to DNA, and at the bottom, it is shown alone. Helices 1, 2, and 3 are labeled in both structures.



Figure 1.4 Two views of the Trp–cage peptide (1L2Y). At top, a ribbon diagram with labeled secondary structure elements is shown. The caged Trp25 residue is shown in pink along with other caging residues. Tyr22 is shown in purple, Gly30 is shown in dark blue, and Pro31, Pro36, Pro37, Pro38 are shown in mint. At bottom, a space-filling diagram in the same orientation is shown with the caged Trp residue shown in pink.



Figure 1.5 Ribbon diagrams of wt ubiquitin (1UBQ) and Gly_8 insertion mutants. The top picture shows wt ubiquitin with the 9-10 loop colored green and the 35-36 loop colored red. The bottom two pictures are overlays of wt Ubq (blue) and Gly_8 -insertion mutants (green). Structural results due to insertion at the 9-10 loop are shown at left, and results due to insertion at the 35-36 loop are shown at right.

Notes

- (1) Szymkowski, D. E. Curr. Opin. Drug Discov. Devel. 2005, 8, 590-600.
- (2) Pokala, N.; Handel, T. M. J. Struct. Biol. 2001, 134, 269-281.
- (3) Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. *Mol. Cell* **2002**, *10*, 895-905.
- (4) Lu, Y. Angew. Chem. Int. Ed. 2006, 45, 5588-5601.
- (5) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779-819.
- (6) Franklin, S. J.; Welch, J. T. Comm. Inorg. Chem. 2005, 26, 127-164.
- (7) Woolfson, D. N. Adv. Protein Chem. 2005, 70, 79-112.
- (8) Lu, Y. Inorg. Chem. 2006, 45, 9930-9940.
- (9) Freemont, P. S.; Lane, A. N.; Sanderson, M. R. *Biochem. J.* **1991**, *278*, 1-23.
- (10) Patikoglou, G.; Burley, S. K. Annu. Rev. Biophys. Biomol. Struct. **1997**, 26, 289-325.
- (11) Brennan, R. G.; Matthew, B. W. Trends Biochem. Sci. 1989, 14, 286-290.
- (12) Sera, T.; Uranga, C. *Biochem.* **2002**, *41*, 7074-7081.
- (13) Montclare, J. K.; Schepartz, A. J. Am. Chem. Soc. 2003, 125, 3416-3417.
- (14) Hegg, E. R.; Burstyn, J. N. Coord. Chem. Rev. **1998**, 173, 133-165.
- (15) Cowan, J. A. Chem. Rev. 1998, 98, 1067-1088.
- (16) Wilcox, D. E. Chem. Rev. 1996, 96, 2435-2458.
- (17) Ghosh, D.; Pecararo, V. L. Inorg. Chem. 2004, 43, 7902-7915.
- (18) Regan, L. Trends Biochem. Sci. **1995**, 20, 280-285.
- (19) Lu, Y.; Roe, J. A.; Bender, C. J.; Peisach, J.; Banci, L.; Bertini, I.; Gralla, E. B.; Valentine, J. S. *Inorg. Chem.* **1996**, *35*, 1692-1700.
- (20) Lu, Y.; Berry, S. M.; Pfister, T. D. Chem. Rev. 2001, 101, 3047-3080.
- (21) Lu, Y.; Gralla, E. B.; Roe, J. A.; Valentine, J. S. J. Am. Chem. Soc. **1992**, 114, 3560-3562.
- (22) Li, C.; Banfield, M. J.; Dennison, C. J. Am. Chem. Soc. 2007, 129, 709-718.
- (23) Lee, H.-W.; Matzapetakis, M.; Mitra, S.; Neil, E.; Marsh, G.; Pecararo, V. L. J. *Am. Chem. Soc.* **2004**, *126*, 9178-9179.

- (24) Lipscomb, W. N.; Strater, N. Chem. Rev. 1996, 96, 2375-2433.
- (25) Maret, W. J. Trace Elem. Med. Biol. 2005, 19, 7-12.
- (26) Berg, J. M.; Godwin, H. A. Annu. Rev. Biophys. Biomol. Struct. 1997, 26, 357-371.
- (27) Nomura, A.; Sugiura, Y. J. Am. Chem. Soc. 2004, 126, 15374-15375.
- (28) Nomura, A.; Sugiura, Y. *Inorg. Chem.* **2004**, *43*, 1708-1713.
- (29) Ippolito, J. A.; Baird, T. T.; McGee, S. A.; Christianson, D. W.; Fierke, C. A. *PNAS* **1995**, *92*, 5017-5021.
- (30) Kostova, I. Curr. Med. Chem. 2006, 13, 1085-1107.
- (31) Reedijk, I. PNAS 2003, 100, 3611-3616.
- (32) Ghadiri, M. R.; Fernholz J. Am. Chem. Soc. 1990, 112, 9633-9635.
- (33) Ghadiri, M. R.; Soares, C.; Choi, C. J. Am. Chem. Soc. 1992, 114, 4000-4002.
- (34) Ghadiri, M. R.; Case, M. A. Angew. Chem. Int. Ed. 1993, 32, 1594-1597.
- (35) Ghadiri, M. R.; Soares, C.; Choi, C. J. Am. Chem. Soc. 1992, 114, 825-831.
- (36) Muheim, A.; Todd, R. J.; Casimiro, D. R.; Gray, H. B.; Arnold, F. H. J. Am. *Chem. Soc.* **1993**, *115*, 5312-5313.
- (37) Hong, J.; Kharenko, O. A.; Ogawa, M. Y. Inorg. Chem. 2006, 45.
- (38) Axup, A.; Albin, M.; Mayo, S. L.; Crutchley, R. J.; Gray, H. B. J. Am. Chem. Soc. **1988**, *110*, 435-439.
- (39) Abram, U.; Alberto, R. J. Braz. Chem. Soc. 2006, 17, 1486-1500.
- (40) Guo, Z.; Sadler, P. J. Angew. Chem. Int. Ed. 1999, 38, 1512-1531.
- (41) Garayoa, W. G.; Ruegg, D.; Blauenstein, P.; Zwimpfer, M.; Khan, I. U.; Maes, V.; Blanc, A.; Beck-Sickinger, A. G.; Tourwe, D. A.; Schubiger, A. Nuc. Med. Biol. 2007, 34, 17-28.
- (42) Costantini, D. L.; Hu, M.; Reilly, R. M. *Cancer Biother. Radiopharm.* **2008**, *23*, 3-24.
- (43) Kim, Y.; Welch, J. T.; Lindstrom, K. M.; Franklin, S. J. J. Biol. Chem. 2001, 6, 173-181.
- (44) Welch, J. T.; Sirish, M.; Lindstrom, K. M.; Franklin, S. J. *Inorg. Chem.* **2001**, *40*, 1982-1984.
- (45) Sirish, M.; Franklin, S. J. J. Inorg. Biochem. 2002, 91, 253-258.
- (46) Kim, Y.; Franklin, S. J. *Inorg. Chim. Acta* **2002**, *341*, 107-112.

- (47) Welch, J. T.; Kearney, W. R.; Franklin, S. J. PNAS 2003, 100, 3725-3730.
- (48) Jain, S.; Welch, J. T.; Horrocks, W. D.; Franklin, S. J. *Inorg. Chem.* **2003**, *42*, 8098-8104.
- (49) Kovacic, R. T.; Welch, J. T.; Franklin, S. J. J. Am. Chem. Soc. **2003**, 125, 6656-6662.
- (50) Wong-Deyrup, S. W.; Kim, Y.; Franklin, S. J. J. Biol. Inorg. Chem. 2006, 11, 17-25.
- (51) Harris, K. L.; Lim, S.; Franklin, S. J. Inorg. Chem. 2006, 45, 10002-10012.
- (52) Lim, S.; Franklin, S. J. Prot. Sci. 2006, 15, 2159-2165.
- (53) Lim, S., The University of Iowa, 2006.
- (54) Wong-Deyrup, S. W., The University of Iowa, 2007.
- (55) Shields, S. B., The University of Iowa, 2005.
- (56) Shields, S. B.; Franklin, S. J. *Biochem.* **2004**, *43*, 16086-16091.
- (57) Shields, S. B.; Franklin, S. J. J. Inorg. Biochem. 2007, 101, 783-788.
- (58) Laughon, A. *Biochem.* **1991**, *30*, 11357-11367.
- (59) Ledneva, R. K.; Alexeevskii, A. V.; Vasil, S. A.; Sprin, S. A.; Karyagina, A. S. *Mol. Biol.* **2001**, *35*, 647-659.
- (60) Stollar, E. J.; Mayori, U.; Lovell, S. C.; Federici, L.; S.M.V., F.; Fersht, A. R.; Luisi, B. F. *J. Biol. Chem.* **2003**, *278*, 43699-43708.
- (61) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425-430.
- (62) Streicher, W. W.; Makhatadze, G. I. *Biochem.* **2007**, *46*, 2876-2880.
- (63) Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. J. Am. Chem. Soc. 2005, 127, 10943-10950.
- (64) Ferraro, D. M., Ferraro, D.J., Ramaswamy, S., Robertson, A.D. *J. Mol. Biol.* **2006**, *359*, 390-402.
- (65) Ferraro, D. M.; Hope, E. K. R., A.D. J. Mol. Biol. 2005, 352, 575-584.

CHAPTER 2: A ZINC-BINDING CHIMERIC PROTEIN, Z0

2.1 Introduction

Chimeric protein design aims to introduce new functionalities into proteins by redesign. Specifically, structurally super-imposable, super secondary protein motifs are exchanged in order to combine the activities of parent structures in one protein. Previous chimeras combined the Ca^{2+} -binding loop from calmodulin with the DNA-binding domain from *engrailed*.¹⁻¹² Lanthanide ions were used in place of Ca^{2+} to increase their reactivity toward DNA. A Zn^{2+} -binding chimera was a "next logical step" in the evolution of such designs. The ion can perform hydrolysis reactions much like lanthanides due to its similar charge density and water exchange rate.¹³ Also, like most lanthanide ions, Zn^{2+} is redox inactive, and it easily accommodates a variety of ligand geometries.¹⁴ However, unlike lanthanides, Zn^{2+} occurs naturally in biological systems and is less toxic. Therefore, the ion can be more easily used in biosystem applications. This chimera was designed to bind Zn^{2+} by combining the Zn^{2+} -binding loop from astacin with the DNA-binding domain from *engrailed*.

Homeodomains are the second-most prevalent DNA-binding motifs after zinc fingers and bind DNA with an affinity of approximately 1 nm.¹⁵ The protein makes specific contacts with 3-4 base pairs in the major groove and contacts another 2 base pairs in the adjacent minor groove.¹⁶ Even though homeodomains are relatively small proteins, they have specificity comparable to their larger cousins.¹⁷ In *engrailed* homeodomain, residues Leu47, Gln50, Asn51, and Met54 in helix 3 are most responsible for DNA specificity, along with Arg3 and Arg5 in the N-terminus.^{18,19} Gln50 has been shown to be especially important to binding specificity. When this residue was mutated, the homeodomain recognized TAATCC instead of its natural TAATTA binding site.^{20,21} Omission of the N-terminal helix (as in the first chimeras in the Franklin group) reduces DNA specificity and affinity but also leads to greater protein instability.²² The helix-

turn-helix motif was previously shown to be amenable to turn substitutions and thus was the initial target scaffold used in this work.

The complimentary helix-turn-strand Zn^{2+} -binding site from astacin, a digestive enzyme from crayfish (*Astacus astacus* L.), was selected for substitution of the turn in the homeodomain motif.²³ Astacin belongs to the metzincin family of peptidases which contain the conserved Zn^{2+} -binding sequence: His-Glu-B-X-His-X-B-Gly-B-X-His (where "B" stands for bulky, hydrophobic residues and "X" can be any amino acid) which folds into a 90° turn structurally similar to the turn in homeodomain motifs. The bulky, hydrophobic residues are critical for packing into the hydrophobic core of astacin. Astacin binds Zn^{2+} in a trigonal bipyramidal coordination mode with three histidine residues and a glutamate-bridged water molecule from the conserved sequence. Unlike in other metzincins, a tyrosine residue from elsewhere in the primary sequence also binds to the Zn^{2+} ion. One proposed role of this Zn^{2+} -binding tyrosine residue is as a "switch". When tyrosine is bound, the zinc site is closed and inactive. When tyrosine is not bound, the zinc site is able to bind a water molecule and perform hydrolysis reactions.

Z0 was originally designed to perform as a novel endonuclease. Substitution of Zn^{2+} ions for Ru^{II} and Tc^{I} complexes would have led to chimeras with possible chemotherapeutic and medical imaging applications. Roles of metal ions in natural hydrolases could also have been revealed by study of this chimera.

2.2 Z0 Chimera Design

The proposed structure for Z0 is a combination of structures of the homeodomain from *engrailed* and the Zn^2 -binding site of astacin as shown in **Figure 2.1**. Crystal structures of *engrailed* homeodomain (PDB accession code: 1ENH) and astacin (PDB accession code: 1AST) were loaded into Swiss PDB Viewer.²⁴ Helix 2 of the homeodomain was aligned with helix 2 of astacin so that their turns into their next secondary structure elements were superimposed. Unlike in previous chimera designs

from the Franklin laboratory, Z0 substitutes parts of proteins in which elements of 2° structure need to be preserved. Because two of the histidine residues needed for binding Zn^{2+} were located in an α -helix from astacin, there were not many options for how to replace the homeodomain sequence with the Zn^{2+} -binding one. Helical residues had to be superimposed so that the C-terminal β -turn from astacin could meet helix-3 of the homeodomain. Any addition or deletion of amino acids from the N-terminal end of the substituted sequence would rotate the β -turn away from its designated C-terminus at helix-3 of the homeodomain (**Figure 2.2**).

Ultimately, the chimera sequence was created by exchanging part of helix 2 and the turn from the homeodomain (residues 31 - 44) for the similar orthogonal turn made from helix 2 and the β -turn (residues 92 - 102) in astacin. The exchanged astacin residues were His-Glu-Leu-Met-His-Ala-Ile-Gly-Phe-Tyr-His. **Figure 2.3** shows the correct alignment of residues and resulting protein sequence for Z0. The resulting chimera had 59 residues, a calculated pI of 10.5, a calculated molar mass of 7.2 kDa, and a calculated extinction coefficient of $8250 \pm 5\%$ cm⁻¹ M⁻¹ (DNAstar, Inc).

To see if the exchanged sequence could span the distance between helices 2 and 3 of *engrailed* homeodomain, molecular modeling with PLOP was used²⁵. PLOP generates protein structures from primary sequences using iterative loop-building/energy evaluating cycles. (Energy calculations were based on the OPLS-AA force field and the Surface Gerneralized Born implicit solvent model.) PLOP allows researchers to explore the capability of amino acid loops to span given distances by generating large numbers of candidate-loops (thousands) built up based on dihedral angles of given amino acids. Candidates are clustered in calculations to reduce redundancy and calculations are performed to optimize side-chains in these candidate loops. Finally, complete energy minimization calculations are performed on the loop, and lowest-energy outputs are given. In this experiment, the first seven residues (His31 – Ile37) of the exchanged sequence were constrained to be helical while the remaining four residues (Gly38 –

His41) were unconstrained. No distance constraints for metal-binding were used, so in essence, only the apo-loop was predicted in calculations. Ribbon structures were generated with Swiss PDB viewer. Results from PLOP calculations are shown in Figure 2.4. The two lowest-energy outputs (structures numbered 1 and 3) are shown above structures for the other lowest-energy outputs. Output 3 looks most like the input structure. It is not difficult to imagine different rotamers of the histidine residues being able to turn and bind to a Zn^{2+} ion. In the other lowest-energy structure (output 1), the second histidine residue is blocked by the turn connecting the helices. In the other three outputs, helix-2 is interrupted after the first histidine residue. Swiss PDB viewer calculated which residues were part of secondary structure elements by comparing torsion angles to standard values. When this helix was distorted in order to have the Cterminal turn meet helix-3, the program determined that some of the residues were no longer within α -helix parameters. Whether those residues truly lose helicity or simply do not meet the viewer's standards for what constitutes a helix is unknown. Further, the extra helix shown in structures 2, 4, and 5 is probably also an artifact of the viewer's secondary structure determinations. Regardless, all structures show that it is possible for the eleven residues from astacin to span the distance between helix-2 and 3 of the homeodomain even if the first seven residues are constrained in a helix.

After many trials trying to optimize expression and purification conditions (described below in section 2.3), Z0's structure was examined in a different way. Instead of looking at Z0 from a strictly structural view, the protein's hydrophobicity and hydrogen-bonding network were examined (**Figure 2.5**). Different representations of Z0 were made with Swiss PDB viewer. Coloring of basic, acidic, polar and non-polar residues was performed using the "coloring by type" function in the viewer. The presence of hydrogen bonds was also calculated with the viewer. Substitution of *engrailed's* helix-2 residues with those from astacin introduced a hydrophobic patch on the protein surface. This region would have been crucial for maintaining the hydrophobic

core in astacin where its metal-binding site was buried in a cleft. However, in the new, Z0 structure, these residues were designed to be on the water-accessible surface of the protein. Furthermore, substitution of Ise37 from astacin for *engrailed's* glutamate abolished a hydrogen bond to Arg15 in helix-1. Introduction of this hydrophobic patch and loss of an important hydrogen bond were probably the causes for the inability to acquire pure Z0.

2.3 Attempts to Make Z0

2.3.1 Materials and Methods

2.3.1.1 Cloning of Z0

Z0 DNA cloned into pProEx-HTC plasmid was received from Kathy Stern (Franklin Laboratory, The University of Iowa, Dept. of Chemistry). Since several purification methods were attempted, the cassette DNA was cloned into several different vectors. All DNA primers were ordered from IDT (Coralville, IA) with desalting but no further purification. Vector changes were made by PCR using *Pfu turbo* polymerase (Stratagene, Agilent Technologies, Co.) and standard PCR procedures²⁶. *E. coli* DH5 α cells (Invitrogen, Corp.) were transformed with plasmid DNA and their plasmids purified following a mini-prep kit manual (Qiagen, Inc.) in order to make DNA plasmid stocks. All DNA sequences were verified by sequencing performed by the DNA Facility (The University of Iowa). All vectors used in these studies carried T7 promoters and a gene for Amp resistance.

2.3.1.1.1 Cloning for the IMPACT System

The first vector that Z0 DNA was inserted into was pTYB1 from the IMPACT (Intein-Mediated Purification with an Affinity Chitin-Binding Tag) System (New England Biolabs, Inc.). This system attaches a chitin-binding domain to the C-terminus of the cloned protein *via* a self-cleaving intein peptide sequence. Forward and reverse

DNA primers were designed to insert Z0 DNA at the *EcoR1* and *Sap1* sites in pTYB1. The forward primer was used to engineer a "start codon" (ATG) and an *EcoR1* restriction site, while the reverse primer engineered a *Sap1* restriction site. Forward: 5'-CAA GAA TTC ATG GAT GAG AAG CGC CCA CGC (30 b.p., $T_m = 67$ °C). Reverse: GGT GGT TGC TCT TCC GCA AGT CGA CTT CTT GAT CTT GG (38 b.p., $T_m = 70$ °C). A total of 45 PCR cycles were used to amplify the cassette DNA out of the pProEx-HTC vector before digestion with *EcoR1* and *Sap1*. The pTYB1 vector was digested with the same restriction enzymes before the resulting termini were dephosphorylated with CIAP. The Z0-cassette was ligated into the pTYB1 vector using T4 ligase.

2.3.1.1.2 Cloning for the Thermovector²⁷ System

The pETM90 vector was received from Ario deMarco (European Molecular Biology Laboratory, Heidelberg, Germany). This system attaches a thermophilic protein (*ftr*) to the N-terminus of the cloned protein *via* a 6-His tag and a TEV site. Forward and reverse DNA primers were designed to insert Z0 DNA at the *Nco1* and *Xho1* sites in pETM90. The forward primer was used to engineer an *Nco1* restriction site, while the reverse primer engineered a "stop codon" (TAA) and an *Xho1* restriction site. Forward: 5'-GAT CGT CAC CAT GGA TGA GAA GCG TCC ACG (30 b.p., T_m = 65 °C). Reverse: 5'-AGT CCT CGA GTT AAG TCG ACT TCT TGA TCT TGG (33 b.p., T_m = 64 °C). A total of 30 PCR cycles were used to amplify the cassette DNA out of the pTYB1 vector (of the previous clone) before digestion with *Nco1* and *Xho1*. The pETM90 vector was digested with the same restriction enzymes before the resulting termini were dephosphorylated with CIAP. The Z0-cassette was ligated into pETM90 vector using T4 ligase.

2.3.1.1.3 Cloning for an Un-tagged System

The pET21a(+) vector was received from Sunghyuk Lim (Franklin Laboratory, The University of Iowa, Dept. of Chemistry). This system expresses the protein of interest without any protein tags to aid in purification. Forward and reverse DNA primers were designed to insert Z0 DNA at the *Nde1* and *Xho1* sites in pET21a(+). The forward primer was used to engineer a "start codon" (ATG) and an *Nde1* restriction site, while the reverse primer engineered a "stop codon" (TAA) and an *Xho1* restriction site. Forward: 5'-GAT CGT CAC ATA TGG ATG AGA AGC GTC CAC G (31 b.p., $T_m = 65$ °C). Reverse: 5'-AGT CCT CGA GTT AAG TCG ACT TCT TGA TCT TGG (33 b.p., $T_m = 64$ °C). A total of 30 PCR cycles were used to amplify the cassette DNA out of the pTYB1 vector before digestion with *Nde1* and *Xho1*. The pET21a(+) vector was digested with the same restriction enzymes before the resulting termini were dephosphorylated with CIAP. The Z0-cassette was ligated into pET21a(+) vector using T4 ligase.

2.3.1.2 Z0 Expression and Purification

Three different systems were used utilizing DNA cloned as described above. Z0 in pTYB1 was used with the IMPACT System (New England Biolabs, Inc.), Z0 in pETM90 was used with the Thermovector System²⁷, and Z0 in pET21a(+) was used in an untagged system. While not necessarily described thoroughly here, many conditions and parameters were varied in attempts to optimize expression and purification of Z0. All cultures were grown in LB + Amp media. Expression and purification were assessed by SDS-PAGE. Soluble proteins (as described in this thesis) were those soluble after cell lysis without the addition of denaturants. Insoluble proteins were those that could only be solubolized upon addition of strong denaturants such as 8 M urea.

2.3.1.2.1 Z0 Purification using the IMPACT System

Many parameters were varied in purification attempts of CBD-tagged Z0. BL21(DE3)star (Invitrogen, Corp.) or ER2566 (New England Biolabs, Inc.) competent cells were transformed with purified Z0-pTYB1 plasmid DNA before inducing cultures at various OD_{600} measurements of cell growth (0.5 – 0.8). Various concentrations of IPTG (0.2 - 0.8 mM) were used to induce protein production at temperatures ranging from $16 - 37^{\circ}$ C. Depending on the temperature used for protein expression, cells were allowed to express for 3 - 18.5 hours. Cells were lyzed using lysozyme or a French Press.

Only one attempt was made to purify CBD-tagged Z0. Lysate was loaded onto a 10 mL chitin column (New England Biolabs, Inc.) equilibrated with 20 mM HEPES, 50 mM KCl, pH 7. The column was washed with another 10 CV of the same buffer, 5 CV of 20 mM HEPES, 1 M KCl, pH 8, and then another 2.5 CV mL of the first buffer again. To cleave Z0 off the column, 2.5 CV of 20 mM HEPES, 50 mM KCl, 50 mM DTT, pH 8.5 buffer was added and allowed to incubate with the protein/column resin at 4 °C for two days. Protein was eluted, and the column was washed with more 20 mM HEPES, 50 mM KCl, pH 7 buffer. Protein was concentrated using two different methods. In the first, the protein solution was concentrated using a 5 kDa NMWL spin filter (Millipore, Corp). In the other method, protein was dialyzed against a concentrated solution of PEG-8000.

2.3.1.2.2 Z0 Purification using the Thermovector System

BL21(DE3)star competent cells (Invitrogen, Corp.) were transformed with purified Z0-pETM90 plasmid DNA before inducing cultures at an OD_{600} of 0.5 with 0.4 mM IPTG. Cells were allowed to express at 28 °C for 3 hours before harvesting the cells by centrifugation (6000 x g for 10 minutes at 4 °C). Cells were lyzed using a French Press at 1000 psi. Both the *ftr*-thermophilic tag and the 6-His-tag were exploited for protein purification.

First attempts to purify Z0 using the thermovector system exploited the thermostable *ftr* protein tag. In order to stabilize *ftr* at high temperatures, PEG 4000 or PEG 400 were added at concentrations ranging from 0 - 10%. NaCl was also added at various concentrations (0 - 300 mM) to try to keep heated proteins in solution. Tagged

protein was heated at 50 or 70 °C for 0 - 40 minutes. Cleavage of Z0 from *ftr* was never attempted since yields of fusion protein after heating were so low.

Subsequent attempts to purify Z0 in this system exploited the 6-His Tag joining *ftr* to Z0. All attempts used IMAC to first purify *ftr*-tagged Z0. The IMAC resin employed was Ni-NTA Superflow (Qiagen, Inc.). Cleared cell lysate was applied to Ni-NTA resin and allowed to incubate with gentle shaking at room temperature for at least 30 minutes before allowing the resin to settle. Protein was eluted using an imidazole gradient applied by FPLC (ÄKTA, GE Life Sciences Co., formerly Amersham-Pharmacia Biotech) where buffer A was 50 mM Tris, 300 mM NaCl, 20 mM imidazole, pH 8 and buffer B was 50 mM Tris, 300 mM NaCl, 250 mM imidazole, pH 8. Protein was detected by absorbance at 280 nm. FPLC fractions containing Z0-*ftr* were pooled and dialyzed (against 50 mM Tris, 100 mM NaCl, pH 8) into conditions amenable to His-TEV reactions with substrate. His-TEV (expressed and purified in house) was added to Z0-*ftr* at a ratio of 1:20 (by molarity) and allowed to incubate at 37 °C for 2 hours.

Once *ftr*-tagged Z0 was purified and allowed to react with 6-His-TEV, Z0 needed to be purified from the tag and protease. In one strategy, a second round of IMAC was used to attempt purification of Z0 from 6-His-TEV and the *ftr* tag. The Z0-*ftr* + His-TEV reaction mixture was re-applied to Ni-NTA as described above. After incubation with resin, proteins were eluted using the same buffer systems and gradient. In the other strategy, other chromatography methods (or combination of methods) were used to purify Z0 from *ftr* and 6-His-TEV. One such trial utilized cation exchange with Heparin Fast-Flow Resin (GE Life Sciences, Co.) The Z0-*ftr* + His-TEV reaction mixture was applied to heparin resin and proteins were eluted using a salt gradient applied by FPLC where buffer A was 50 mM Tris, 100 mM NaCl, pH 8, and buffer B was 50 mM Tris, 1 M NaCl, pH 8. Protein was detected by absorbance at 280 nm. In another trial, 1% Triton X-100 and 1 M urea were added to put the Z0-*ftr* + His-TEV reaction mixture into mildly denaturing conditions. This solution was applied to an anion exchange resin called

DEAE sepharose (Sigma-Aldrich, Co.). Protein was eluted with 50 mM Tris, 1 M NaCl, pH 8 buffer. Z0-containing fractions were dialyzed against buffer containing only 100 mM NaCl, and 1 M urea was added to the dialysate. This solution was applied to a size exclusion column (Sephadex G-50F, GE Life Sciences, Co.) and eluted with 50 mM Tris, 100 mM NaCl, pH 8 buffer. In a final trial, the Z0-*ftr* + His-TEV reaction mixture was concentrated and directly applied to a size-exclusion column (pre-packed Superdex, GE Life Sciences, Co.). Proteins were eluted with 50 mM Tris, 100 mM NaCl, pH 8 buffer.

If any Z0 could be purified, buffer conditions needed to be changed so that metal ions could be added (pH \leq 7, [NaCl] ~150 mM, no denaturant or surfactants). Buffer change attempts were performed by dialysis. In some attempts, buffer conditions were changed drastically in one step while in others, conditions were changed gradually using several different dialysis buffers in sequence.

2.3.1.2.3 Z0 Purification using an Un-tagged System

Many parameters were varied when optimizing expression of Z0. Some expression conditions produced soluble protein while others produced Z0 expressed in insoluble inclusion bodies. BL21(DE3)star competent cells (Invitrogen, Corp.) were transformed with purified Z0-pET21a(+) plasmid DNA. Cells were induced with 0.8 mM IPTG at OD_{600} measurements ranging from 0.4 – 0.7. Protein was expressed at 28 or 37 °C. Cells were harvested by centrifugation before attempting lysis using either a French Press only or with pre-treatment with lysozyme.

Z0 purification was attempted under native, mildly denaturing and completely denaturing conditions utilizing many chromatography techniques including IMAC, cation exchange, and size exclusion. Refolding of protein was attempted under various concentrations and pH conditions with additives such as urea, ZnCl₂ and L-arginine. In the most successful attempts, a purification strategy adapted from Sunghyuk Lim's purification of C2 was used.⁶

Briefly, cells were resuspended in 50 mM Tris, 100 mM NaCl, 10 mM EDTA, pH 8 buffer and incubated with lysozyme. This lysate was centrifuged at 20 000 x g for 30 minutes at 4 °C. The pellet was washed 50 mM Tris, 100 mM NaCl, pH 8 buffer containing DNaseI and the solid fraction was harvested by centrifugation. More buffer containing DNaseI was added before lyzing with a French Press. This lysate was centrifuged at 20 000 x g for 45 minutes at 4 °C before adding (NH₄)₂SO₄ to the soluble fraction to 35% saturation to precipitate Z0. Proteins in both the insoluble lysate and (NH₄)₂SO₄ precipitate were extracted using buffered urea (100 mM Tris, 8 M urea, pH 8). 1% TritonX-100 was added to diluted, extracted, denatured proteins, and the buffer was very slowly changed to 20 mM Tris, 100 mM NaCl, 1 M urea, 2 mM EDTA, pH 8 by dialysis. Dialyzed protein was applied to heparin resin and eluted with a salt gradient applied by FPLC where buffer A was 20 mM Tris, 100 mM NaCl, pH 8 and buffer B was 20 mM Tris, 2 M NaCl, pH 8. Z0-containing fractions were pooled and concentrated using a stirred cell fitted with a YM-3 membrane (Millipore, Corp.). This strategy produced 95% pure Z0, but in a buffer system that was not useful for further studies (pH \sim 8 and in 1 M NaCl). Further purification was attempted by size exclusion using Sephadex G-50F or Superdex. Buffer changes to lower pH and/or [NaCl] and exchange Tris buffer to non-metal-binding ones (*i.e.* MOPS or MES) were also attempted.

2.3.2 Results and Discussion of Z0 Purification

While many parameters and conditions were varied in attempts to purify Z0, only a few of the more successful attempts are discussed here.

2.3.2.1 Purification of Z0 Using the IMPACT System

Z0-pTYB1 plasmid DNA was successfully cloned and used to transform different strains of bacteria (BL21(DE3)star or ER2566). When protein expression was induced, yields were extremely low (**Figure 2.6**). Since proteins can be expressed as soluble cytoplasmic proteins or in insoluble inclusion bodies, both cell fractions were analyzed.

Z0-CBD (62 kDa) did not appear in either the soluble or insoluble fractions from cell lysates. Since the IMPACT system was the first one used in purification attempts, it was assumed that the low-temperature recommended for protein expression (16 $^{\circ}$ C) was responsible for low yields.

Previous work with this system (Ramaswamy Laboratory, The University of Iowa, Dept. of Biochemistry) indicated that expression of CBD-tagged proteins at higher temperatures produced proteins in the insoluble lysate. Proteins expressed in this fraction are typically misfolded, denatured, or have undergone some level of proteolysis. There is no guarantee that such proteins can refolded into their native state. Further, it is very difficult to get a homogeneous sample of protein. Expression at high temperature may produce more protein, but the yield of useful protein may be very low.

Even though expression was low, purification of Z0 using a chitin column was attempted in case (1) Z0-CBD was expressed but not detected by SDS-PAGE amidst other cellular proteins or (2) Z0 refolded on the column before elution. Still, no Z0 was detected. Since protein refolding was not a desirable method for obtaining pure Z0, and expression at low temperatures gave extremely low yields, it was decided that IMPACT was not a good system for making Z0.

2.3.2.2 Purification of Z0 Using the Thermovector System

Z0-pETM90 plasmid DNA was successfully cloned and used to transform BL21(DE3)star cells. Soluble fusion-protein expression was very high. When exploiting the thermostability of the *ftr* tag, there was no selectivity for the Z0-ftr (42 kDa) fusion protein (**Figure 2.7**). Upon heating (in the presence of various recommended stabilizing additives), all proteins precipitated out of solution to similar extents. Since so much optimization would be required to use this method, it was abandoned.

Using IMAC, Z0-ftr was much more easily purified (**Figure 2.8**). There were still some contaminating proteins, but more than 90% of them were removed using IMAC.

Also shown in this figure are the species present after treatment of Z0-ftr with His-TEV, a protease recognizing a cleavage site between Z0 and its *ftr* tag. Shown in lane 8 are unreacted Z0-ftr (42 kDa), the *ftr* tag (35 kDa), His-TEV (29 kDa), and freed Z0 (7 kDa). In the particular trial shown, much Z0-ftr was lost to the column flow-through fractions because too small a column was used. In other trials, this loss was minimized. This trial does show the level of Z0-ftr purity achieved using IMAC, though. However, Z0-ftr was not the goal of this purification. Z0 still needed to be purified from the other contaminating proteins (unreacted Z0-ftr, *ftr*, and His-TEV).

When a second round of IMAC was used, Z0 eluted in the column wash fractions despite having a designed metal-binding site (**Figure 2.9**). It was hoped that Z0 would be retarded by the column media as typically observed for metal-binding proteins. Not only did Z0 elute in the wash fractions, but it also co-eluted with other, contaminating proteins. This Z0 was also very dilute. These results indicated that Z0 interacted with other proteins and could not bind to Ni²⁺ ions under these conditions.

Cation exchange chromatography was even less effective when trying to purify Z0. The chimera was able to bind to the column, but co-eluted with *ftr* and another contaminating protein (**Figure 2.10**). Z0 can faintly be seen in fractions eluted in the second half of the salt gradient. Again, Z0 was extremely dilute and contaminated. Size exclusion was also attempted, but again, Z0 co-eluted with other contaminating proteins (**Figure 2.11**). Furthermore, it eluted in early fractions (5 - 7) where large proteins typically appear. It was expected that Z0 would elute slowly from such a column because it should have been able to interact with the resin's pores. This early elution indicated that under these conditions, Z0 was either sticking to other larger proteins in solution, aggregating into larger protein species, or that its hydrodynamic radius was greatly increased (*i.e.* due to denaturation). Regardless of the reason for the early elution, it indicated that Z0 was mis-folded and probably interacting with itself and other

contaminating proteins. While not shown here, anion exchange and other combinations of chromatography methods also produced low-yield, contaminated Z0.

2.3.2.3 Purification of Z0 Using an Un-tagged System

Purification of Z0 using untagged protein is shown in **Figure 2.12**. Expression yields were very low, which is indicative of production of mis-folded proteins. However, some Z0 could be extracted from precipitates. After application to a heparin column, pure Z0 could be detected. While Z0 was successfully purified and concentrated, it was in conditions that were either unsuitable for metal-binding studies or that would have left Z0 denatured. In this particular trial, Z0 was in 20 mM Tris, 1 M NaCl, pH 8 buffer.

Any attempts to lower the salt concentration led to precipitation of protein. If denaturants were used in purification, their removal resulted in protein precipitation. Changes in pH, whether performed slowly or quickly also caused precipitation of Z0. The only way to stabilize this protein, was to place it in buffers meant to stabilize misfolded proteins.

2.4 Conclusions

Previous, successful chimeric designs in the Franklin laboratory substituted hydrophilic loops containing a 12 amino acid, lanthanide-binding sequence for the turn connecting helices 2 and 3 of *engrailed*. These metal-binding residues were mostly acidic aspartates. In contrast, part of this *engrailed*-astacin chimera's substitution needed to maintain its α -helical structure. Furthermore, residues from a buried metal-binding site were used instead of those from a surface-accessible loop. Lastly, the ligand set was changed to neutral histidine residues in this design. At biologically relevant pH, these residues would be in an assortment of protonation states rather than homogeneously negatively charged as in aspartate.

Z0 could never purified in conditions useful for metal-binding studies (see section 2.3.2). Protein was only soluble in the presence of high concentrations of NaCl or in the

presence of mild denaturants (1 M urea) or surfactants (1% TritonX-100). Any attempts to change the pH of buffers or to remove solubolizing additives caused Z0 to precipitate out of solution. Although this behavior is also common for many naturally occurring proteins, empirical evidence supports the hypothesis that introduction of a hydrophobic patch greatly reduced protein solubility. This loss of solubility could be due to protein destabilization or to aggregate formation.

The Glu37 – Arg15 hydrogen bond was the only such bond connecting helix 2 to helix 1. With the mutation to isoleucine, no apparent hydrogen bonds could be made between the two helices to stabilize overall tertiary structure. Even if Z0 could have properly folded itself into the correct tertiary structure, it probably would have aggregated with itself and other proteins *via* the introduced hydrophobic residues. Aggregation could be seen when Z0 could not be purified from other contaminating proteins. This patch would have tried to bury itself in other proteins in order to shield itself from water. If aggregates became too large, they could form protein precipitates as seen throughout protein purification.

Due to the changes in approach (where a Zn^{2+} -binding site from a buried, α helical, histidine-rich protein region was chosen to build a chimera with *engrailed* homeodomain), this first foray into building a soluble Zn^{2+} -binding chimera was most probably doomed to fail. Specifically, it was concluded that a chimeric approach to protein design based solely on structural considerations is not as robust as originally hoped. Although previous designs showed that successful chimeras could be made after comparing protein geometries, the homeodomain scaffold cannot accommodate such drastic changes to its hydrophilic/hydrophobic balance without significant compensation from charges (salts) or surfactants. In naturally-occurring systems, such "patchwork-proteins" (*i.e.* membrane proteins) are often not isolated without co-purified binding partners, salts, or surfactants. More than just geometry must be taken into account to design a Zn^{2+} -binding chimera from *engrailed* homeodomain. Despite lack of success in making pure Z0, this project did provide a lot of knowledge that could be applied to future projects. Instead of "borrowing" complete metal-binding sites from other proteins, they could be built onto protein scaffolds. In other words, instead of substituting one entire site of a protein for another, point mutations could be made or parts of sites could be borrowed. By being judicious when choosing residues to mutate, one could avoid creating hydrophobic patches or destroying hydrogen-bonding networks. If a new Z-series protein were to be designed from *engrailed*, extending the loop joining helices 2 and 3 with histidine and spacer residues such as glycine could be done. Point mutations around the turn joining helices 2 and 3 could be made, but this region of the protein is very compact, and sterics would need to be taken into account. In an approach that was not purely chimeric, point mutations could be made elsewhere on the homeomomain in regions close in 3° structure even if they were not near each other in sequence. Such new Zn^{2+} -binding sites could stabilize the entire protein 3° structure instead of stabilizing a turn.

Making Zn^{2+} -binding mutants of a protein requires the use of many *de novo* design principles. It is necessary not only to conserve secondary structure elements, but also maintain hydrogen-bonding networks and hydrophobic cores in order to stabilize a protein's tertiary structure. As an additional hurdle when designing a Zn^{2+} -binding protein, steric allowances and charge repulsions must also be considered. While *engrailed* homeodomain may not be the most straight-forward scaffold to use when building metal-binding sites, the *de novo* design principles learned by working with Z0 could be applied to work with other scaffolds.



Figure 2.1 Proposed ribbon structure of Z0 (center) as constructed from *engrailed* homeodomain (left) and astacin (right). The DNA-binding motif is shown in green while the Zn^{2+} -binding site is shown in yellow. The zinc ion is shown as a grey sphere bound to three histidine residues also shown in yellow.



Figure 2.2 Ribbon diagrams showing deviation of astacin's β -turn from meeting helix-3 of the homeodomain if a residue is added or subtracted from the N-terminus of the substituted sequence. The middle figure shows the correct alignment of 2° structures while the top and bottom figures show resulting models due to subtraction (top) or addition (bottom) to the N-terminus of the exchanged sequence. The homeodomain motif is shown in green while astacin's Zn²⁺-binding site is shown in yellow. The zinc ion is shown as a grey sphere bound to three histidine residues also shown in yellow. Helix-1 of *engrailed* has been omitted for clarity.



Figure 2.3 Two ribbon diagram views showing the overlay of *engrailed* homeodomain (1ENH) and the Zn^{2+} -binding site of astacin (1AST) when sequence Z0 is used as shown at the right. Conserved residues from the homeodomain motif are shown in green while those that are substituted by asticin are shown in grey. Astacin's Zn^{2+} -binding site is shown in yellow along with its zinc-binding histidines. Helices are numbered 1 - 3, and the zinc ion is shown as a grey sphere. Protein sequences are shown in one-letter amino acid abbreviations.



Figure 2.4 PLOP results from calculations with Z0. Residues from the homeodomain motif are shown in green while astacin's Zn^{2+} -binding site is shown in yellow or pink. Yellow residues were constrained while those in pink (shown in input structure only) were not. Zn^{2+} -binding histidines are shown in blue (shown in input and first two outputs only).



Figure 2.5 Representations of Z0 and *engrailed* homeodomain comparing a structural view with that of one showing polarity assignments. In the top-left, structural view, homeodomain residues are shown in green while astacin's Zn^{2+} -binding site is shown in yellow. In the middle two views, Z0 residues are labeled by type where basic residues are blue, acidic are red, polar are yellow, and non-polar are grey. Helices are numbered 1 – 3, and the zinc ion is shown as a grey sphere. In the bottom-right, polarity-labeled view, *engrailed* homeodomain is shown for comparison.



Figure 2.6 SDS-PAGE analysis of Z0-CBD expression using the IMPACT system. The arrow indicates where Z0-CBD should appear in gel lanes according to the molecular weight standards shown in the middle lane.



Figure 2.7 SDS-PAGE analysis of Z0-ftr expression using the Thermovector system. The arrow indicates where Z0-ftr appears in gel lanes (in agreement with molecular weight standards shown in the right-most lane). The first five lanes show soluble lysate heated to 50°C in the presence of 5% PEG 4000, 200 mM NaCl over time. The amount of time which the sample was heated increases from left to right.







Figure 2.9 SDS-PAGE analysis of Z0 purification using the Thermovector system and a second round of IMAC. The arrows indicate where Z0-ftr, *ftr*, His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the left-most lane. Column washes are shown in the next three lanes. Proteins eluted with imidazole are shown in the next lanes where imidazole concentration increases from left to right.



Figure 2.10 SDS-PAGE analysis of Z0 purification using the Thermovector system and cation exchange. The arrows indicate where Z0-ftr, *ftr*, His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the right-most lane. Proteins eluted with NaCl are shown in the next lanes where [NaCl] increases from left to right.



Figure 2.11 SDS-PAGE analysis of Z0 purification using the Thermovector system and size exclusion. The arrows indicate where Z0-ftr, *ftr*, His-TEV, and Z0 appear in gel lanes. The reaction mixture applied to the column is shown in the right-most lane. Proteins were eluted from the column over time as shown in lanes from left to right.



Figure 2.12 SDS-PAGE analysis of Z0 purification using an untagged system. The arrow indicates where Z0 appears in gel lanes. Purified Z0 can be seen in the right-most lane. All samples are labeled according to how they were purified, and lanes 1 and 8 show identical samples.
Notes

- (1) Franklin, S. J.; Welch, J. T. Comm. Inorg. Chem. 2005, 26, 127-164.
- (2) Harris, K. L.; Lim, S.; Franklin, S. J. *Inorg. Chem.* **2006**, *45*, 10002-10012.
- (3) Kim, Y.; Franklin, S. J. *Inorg. Chim. Acta* **2002**, *341*, 107-112.
- (4) Kim, Y.; Welch, J. T.; Lindstrom, K. M.; Franklin, S. J. J. Biol. Chem. 2001, 6, 173-181.
- (5) Kovacic, R. T.; Welch, J. T.; Franklin, S. J. J. Am. Chem. Soc. **2003**, *125*, 6656-6662.
- (6) Lim, S., The University of Iowa, 2006.
- (7) Lim, S.; Franklin, S. J. Prot. Sci. 2006, 15, 2159-2165.
- (8) Sirish, M.; Franklin, S. J. J. Inorg. Biochem. 2002, 91, 253-258.
- (9) Welch, J. T.; Kearney, W. R.; Franklin, S. J. PNAS **2003**, *100*, 3725-3730.
- (10) Welch, J. T.; Sirish, M.; Lindstrom, K. M.; Franklin, S. J. *Inorg. Chem.* **2001**, *40*, 1982-1984.
- (11) Wong-Deyrup, S. W., The University of Iowa, 2007.
- (12) Wong-Deyrup, S. W.; Kim, Y.; Franklin, S. J. J. Biol. Inorg. Chem. 2006, 11, 17-25.
- (13) Hegg, E. R.; Burstyn, J. N. Coord. Chem. Rev. 1998, 173, 133-165.
- (14) Vallee, B. L.; Auld, D. S. *Biochem.* **1990**, *29*, 5647-5659.
- (15) Ledneva, R. K.; Alexeevskii, A. V.; Vasil, S. A.; Sprin, S. A.; Karyagina, A. S. *Mol. Biol.* **2001**, *35*, 647-659.
- (16) Laughon, A. Biochem. 1991, 30, 11357-11367.
- (17) Kornberg, T. B. J. Biol. Chem. 1993, 268, 26813-26816.
- (18) Draganescu, A.; Tullium, T. C. J. Mol. Biol. 1998, 276, 529-536.
- (19) Fraenkel, E.; Rould, M. A.; Chambers, K. A.; Pabo, C. O. *J. Mol. Biol.* **1998**, 284, 351-361.
- (20) Hanes, S. D.; Brent, R. Science 1991, 251, 426-430.
- (21) Hanes, S. D.; Brent, R. Cell 1989, 57, 1275-1283.
- (22) Dragan, A. I.; Li, Z.; Makeyeva, E. N.; Millgotina, E. I.; Liu, Y.; Crane-Robinson, C.; Privalov, P. L. *Biochem.* **2006**, *45*, 141-151.

- (23) Bode, W.; Gomis-Ruth, F.-X.; Stockler, W. FEBS 1993, 331, 134-140.
- (24) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714-2723.
- (25) Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. *Proteins* **2004**, *55*, 351-367.
- (26) Sambrook, J.; Russell, D. W. In *Molecular Cloning, A Laboratory Manual*; Third ed.; Argentine, J., Ed.; Cold Spring Harbor Laboratoratory Press: New York, 2001; Vol. 2, p 8.18-8.24.
- (27) DeMarco, A.; Casatta, E.; Savaresi, S.; Geerlof, A. J. Biotech. 2004, 107, 125-133.

CHAPTER 3: METAL-BINDING TRP-CAGE PEPTIDES, 3HT-C AND 3HT-CB

3.1 Introduction

As stated previously, the Trp-cage is a compact, globular mini-protein composed of a 20 amino acid motif optimized from exendin-4. Its topology is defined by the following residues: Tyr22, Trp25, Gly30, Pro31, Pro37, and Pro38 (exendin-4 numbering).¹ The interwoven hydrophobic core orients Trp25 between the two proline residues at positions 31 and 38. The importance of the conformation of the proline residue at position 31 was highlighted by proline-editing experiments in which the residue was stabilized in either its exo or endo conformation.² When this residue was locked into its exo-conformation, α -helicity and stability increased. Conversely, when the residue was locked in its endo-conformation, α -helicity and stability decreased. It was determined that hydrophobic collapse alone does not define the stability of the Trpcage motif; constraints determined by proline conformation can contribute much to peptide stability. In another study examining the interactions between the tryptophan residue packed between proline residues, mutants exchanging Phe, His and naphthalene residues for Trp were made.³ It was found that the tryptophan residue is unique in its ability to form the necessary hydrophobic cluster and hydrogen bond to a remote backbone carbonyl (from Arg35). There is also a strong interaction between Gly30 and the caged tryptophan residue; these glycine protons have the most upfield shift observed for non-heme proteins.⁴

Typically, small peptides need other stabilizing factors such as disulfide bonds or metal chelation to stabilize their structures, but the Trp-cage folds autonomously without such post-translational modifications. In this motif, side-chains separated by up to thirteen amino acids in sequence are brought into close proximity upon folding. Furthermore, research indicates that 90 - 95% of the peptide is properly folded in water

at 2 °C and that folding proceeds *via* a two-state process.¹ In other words, this peptide is well suited to serve as a model, miniprotein scaffold.

The Trp-cage was originally designed for applications in computational studies because calculations on such a small, fast-folding (4 μ s) peptide could be more easily modeled than for larger systems.^{1,4} In fact, this 4 μ s folding time makes the Trp-cage the fastest folding protein described in the literature.¹ Since the tryptophan provides a spectroscopic handle and the well-defined structure was determined by 2-D NMR studies, calculated structures can be easily compared with empirical findings.

Several groups have tried to improve the stability and folding speed of the Trpcage. In an attempt to reach the estimated "speed limit" for protein folding (1 μ s), computationally designed mutants were designed.⁵ The resulting design incorporated a second tryptophan residue able to stack with the originally-caged tryptophan in an edgeto-face orientation. This mutant was slightly more stable with a T_m of ~57 °C and had a folding rate of ~ 1 μ s at 23 °C. In an attempt to form another stable Trp-cage mutant, this thesis work incorporated a metal-binding site onto this scaffold.

3.2 Materials and Methods

3.2.1 Mutant Design

Since the hydrophobic core is absolutely crucial to Trp-cage folding and stability, changes to the peptide were made so as to limit their influence upon that core. Therefore, two His point mutations were made on the surface of the α -helix at Asn20 and Gln24 (exendin-4 numbering). A third Histidine was added to the N-terminus *via* a 4-Gly flexible linker. This initial design was named "3HT-C" for "<u>3-His Trp-C</u>age". In another attempt to improve the peptide's stability, an Ala residue was added between the Gly linker and the α -helix. This Ala residue was added to serve as a helical cap to aid in helix collapse/folding.⁶ This second design was named "3HT-Cb" for "<u>3-His Trp-C</u>age, design <u>b</u>". 3HT-C has 25 amino acids and a calculated molar mass of 2.6 kDa while

3HT-Cb has 26 amino acids and a calculated mass of 2.7 kDa. Their calculated pI is 8.9 (DNAstar, Inc). An extinction coefficient of 6760 cm⁻¹ M⁻¹ previously reported in the literature (for the originally designed Trp-cage) was used for these designs.⁴ The calculated extinction coefficient was $6970 \pm 5\%$ (DNAstar, Inc.) and is in good agreement with the experimentally determined value. A proposed structure for 3HT-C, 3HT-Cb, and their peptide sequences are shown in **Figure 3.1**.

Peptides were ordered from GenScript, Corp. at $\geq 95\%$ purity. They were received as lyophilized powders which were resuspended in ddi H₂O, aliquotted, and lyophilized again to afford stable powders that were stored over drierite at -20 °C. For experiments, powders were resuspended in appropriate, chelex-treated buffers and the concentration determined by A₂₈₀.

3.2.2 Folding and Structure Preservation of Apo-3HT-Cb

Folding of 3HT-C folding was qualitatively inspected using 2-D ¹H-NMR (TOCSY and NOESY). Spectra were collected on a Bruker Advance 600 MHz spectrometer (NMR facility, The University of Iowa, Dept. of Chemistry) at 4 °C. A 0.95 mM sample of 3HTC in 90% H₂O:10% D₂O and 20 mM MOPS, 100 mM NaCl, pH 7 buffer was prepared with a 5 μ M DSS internal standard. WATERGATE or presaturation pulses were used to suppress the H₂O resonance. 1.11 equivalents of ZnCl₂ were added to the peptide, but the addition of this salt caused so much peak broadening that correlations could not be easily detected.

3.2.3 Zn²⁺-binding by Modified Trp-Cages

 Zn^{2+} -binding was investigated by fluorescence and CD. In fluorescence studies, Trp quenching upon addition of Zn^{2+} was monitored. In CD studies, helix structure induction upon Zn^{2+} -binding was monitored.

3.2.3.1 Fluorescence Studies

Fluorescence spectra were collected on an Aminco-Bowman Series 2 fluorimeter. 25-35 μ M peptide samples in 20 mM MOPS, 100 mM NaCl, pH 7 buffer were excited at 285 nm and their emission scanned from 295-395 nm at a scan rate of 3 nm/sec. The voltage sensitivity was set to 60 or 70% to allow changes in peak intensity to remain within the detection limits of the fluorimeter. Samples were scanned at room temperature. After Zn²⁺ addition (as the ZnCl₂ salt diluted in buffer), samples were allowed to incubate for 5 minutes before scanning. Zn²⁺ was added at ratios of 0 – 10:1, Zn²⁺:peptide. For each data point, 3 scans were averaged and the emission at 350 nm was recorded. Error bars for the Zn²⁺ titration of 3HT-C were generated from calculation of the standard deviation of data from two separate trials. No error bars were generated for the Zn²⁺ titration of 3HT-Cb since only one trial was performed. The K_d was determined by fitting data to a non-linear, least squares, single binding site algorithm.⁷

3.2.3.2 CD Studies

CD spectra were collected on an Olis Cary-17 DS Conversion Spectrophotometer. Spectra were collected (in 1 mm path-length quartz cuvettes) from 200 - 290 nm at 1 nm *per* point resolution and a bandwidth of 3 nm with a variable integration time determined from the PMT voltage. Peptide was diluted to ~80 µM in 5 mM MOPS, 25 mM NaCl, pH 7 buffer. Spectra were baseline-corrected using the same buffer as the reference. Peptide folding was increased by the addition of TFE or by lowering the temperature to 2 °C during titrations.⁸ ZnCl₂ diluted in buffer was added in aliquots at Zn:peptide ratios varying from 0 – 10. Samples were allowed to equilibrate for at least 5 minutes before scanning. Three spectra were averaged and then smoothed (11 point, moving, weighted average using least squares calculations) using the software provided by Olis. Spectra were normalized for concentration and number of amino acid residues before zeroing at 260 nm. Error bars were generated by calculating the instrument noise at 222 nm. K_d values were determined by fitting $[\theta]_{222}$ data to a non-linear, least squares, single binding site algorithm.⁷

3.2.3.3 HPLC Studies

Reversed-phase HPLC analyses were performed on an Agilent 1100 series highperformance liquid chromatograph using a PRP-3 analytical column (Hamilton, Co.) and the following gradient (where solvent A was 0.1% TFA in ddi water and solvent B was 0.1% TFA in acetonitrile): 0 - 2 minutes = isocratic at 10% solvent B; 2 - 22 minutes = linear gradient, 10 - 50% solvent B; 22 - 33 minutes = linear gradient, 50 - 90% solvent B; 33 - 35 minutes = isocratic at 90% solvent B. 150 µL samples were injected and their chromatographs were produced from UV-Vis analyses at 214 and 280 nm.

3.3 Results and Discussion

3.3.1 Folding and Structure of Apo-Trp-Cage Mutants

Dispersion of peaks in the 2-D ¹H-NMR spectra (TOCSY and NOESY) of 3HT-C indicated that the peptide was well-folded under conditions used in this experiment (data not shown). Peaks near 2.7, 2.9, and 4.6 ppm were buried beneath the resonances of water and MOPS buffer. However, patterns of ¹H interactions can be seen in other areas of the spectra. This data supports findings from low-temperature CD studies.

CD spectra of apo-peptides incubated at low temperatures also showed that they were folded (see later CD spectra). From literature, it is known that the Trp-cage is stable, but that its structure is more easily determined if the temperature is lowered or if TFE is added.⁸ The original developers of this peptide noted that its CD spectrum was dominated by α -helical characteristics, and had minima at ~ 205 and 222 nm (typical α -helix bands are at 209 and 222 nm).⁴ Like the original peptide, cooled 3HT-C and 3HT-C balso had minima at 204 and 222 nm, but the 222 nm band was not as strong in intensity. Together, these spectral characteristics describe a protein that has more random

coil contribution (minimum at 199 nm) than the parent Trp-cage. This finding is not surprising since the last metal-binding histidine residue was added to this peptide at the end of a very flexible, N-terminal tail. Even upon metal-binding, this loop was not expected to assume a rigid structure. Also, the Trp-cage motif was highly optimized, and introduction of mutations could negatively affect stability.

Not shown are investigations of structure at room temperature or in the presence of 30% TFE, a fluoroalcohol commonly used to stabilize α -helices.⁸ Addition of TFE caused the spectrum of 3HT-C to adopt a curve shape more like that of its parent where the intensities at ~205 and 222 nm were more similar. Data collected at room temperature was very noisy because the signal was so weak. In other words, the peptide was not folded well enough (or homogeneously enough) to give a measurable CD signal.

3.3.2 Zn²⁺-binding by Modified Trp-Cages

3.3.2.1 Fluorescence Studies

There is an obvious, systematic decrease in tryptophan emission upon addition of Zn^{2+} . Emission decreased sharply until addition of one equivalent of Zn^{2+} , implying a 1:1 binding stoichiometry. K_d values in the micromolar range were determined in these studies with ZnCl₂. Such values were calculated from titration data as shown in **Figure 3.2**. When 3HT-C was titrated with Zn^{2+} in the presence of 20 mM MOPS, 100 mM NaCl, pH 7 buffer at room temperature, a K_d of $16 \pm 5 \mu$ M (2 separate trials) was determined. When 3HT-C's sister peptide, 3HT-Cb, was titrated with Zn^{2+} in the presence of the same buffer at the same temperature, the K_d was 18 μ M. This determination was made form only one trial.

Classically, tryptophan fluorescence quenching has been seen as indicative of the residue entering a more polar environment (*i.e.* exposure to solvent upon unfolding). However, quenching mechanisms can be very complex and other entities besides solvent can cause quenching (disulfide bonds, other aromatic residues, or even the peptide backbone).⁹⁻¹¹ Despite the Trp-cage containing an intrinsic tryptophan residue, only one study was found where the intrinsic fluorescence of the peptide was described, but this description was not thorough.¹² (Most researches have studied this peptide's folding mechanism by CD or NMR.) The authors noted that as the peptide was heated (to induce denaturation) from ~20 - ~40 °C, fluorescence intensity subtly increased. Further heating caused quenching as expected in unfolded proteins. Another study using UV-resonance Raman spectroscopy showed that there is a folding maximum at ~20 °C. Temperature deviations above and below this temperature caused Trp-cage to "relax". The combination of unusual fluorescence spectral characteristics with complexity of Trp-cage compactness during folding precludes using fluorescence intensity to formulate hypotheses as to the tryptophan exposure to solvent. In another proposed quenching mechanism, decreases in fluorescence emission could be due to changes in the apparent, local polarity due to proximity of the Trp ring's edge to the ionic, Zn²⁺ environment.

Typically, a titration curve levels off upon saturation of binding sites with metal. With these peptides, the emission at 350 nm continues to decrease even after the addition of one equivalent of metal, albeit more slowly. Addition of more metal ions probably causes additional quenching due to collisions of the protein with metal ions. Without further study as to quenching mechanisms, it is difficult to say whether all quenching is due to addition of Zn^{2+} ions. As the peptide is initially loaded with Zn^{2+} , the tryptophan residue is quenched markedly. This quenching proceeds semi-linearly because fluorescence studies only allow observation of the average of protein states. As more proteins bind Zn^{2+} , these holo-proteins' emission dominate the spectra. After saturation of the binding sites, further quenching could be strictly collisional and therefore, strictly concentration dependent (with a linear correlation).

Since fluorescence spectra gave reproducible and well-behaved data, but the mechanism of quenching could not be explained, this method for investigating binding by folded peptide was not reliable standing alone. However, these data, in conjunction with

CD studies (below), serve as a first indication that a metal-binding Trp-cage having micromolar affinity for Zn^{2+} was made.

3.3.2.2 CD Studies

By monitoring changes in ellipticity at 222 nm, Zn^{2+} -binding induction of secondary structure could be investigated (**Figure 3.3**). Data indicate a K_d of 7.3 ± 1.5 μ M when 3HT-Cb is titrated with Zn^{2+} (ZnCl₂ diluted in buffer) in the presence of 5 mM MOPS, 25 mM NaCl, pH 7 buffer at 2 °C (2 trials). The temperature was lowered so it could be assumed that metal-ions were bound by a more folded peptide rather than just three histidines spaced out over a large, unfolded, peptidic chelate. Again, spectra indicate a 1:1 stoichiometry when 3HT-Cb binds to Zn^{2+} . Furthermore, the presence of an isosbestic point (at ~215 nm) suggests that binding is a two-state process, but the identity of these states has not been fully elucidated.¹³

Traditionally, induction of α -helix formation is measured by an increase in $[\theta]_{222}$ magnitude.¹⁴ While addition of Zn²⁺ causes an increase in magnitude of this band, a quantitative measure of α -helicity is impossible due to a reinforcing chromophore contribution from the tryptophan.⁴ Even with this caveat, it is clear that Zn²⁺ addition causes 3HT-Cb to assume a more α -helical structure. Two of the designed zinc-binding histidines are located in the peptide's helix and it is not difficult to imagine metal-binding assisting in nucleation of this α -helix. Interestingly, the binding affinity determined by CD studies performed at low temperature is approximately twice that of the affinity calculated by fluorescence studies at room temperature. This finding is consistent with structural differences observed as a function of temperature, wherein the 2-His binding site on one side of the helix (H20 and H24) is more pre-organized for Zn²⁺-binding at 2 °C than at 25 °C.

These CD results provide better support for the hypothesis that a metal-binding Trp-cage was successfully designed. Although the reporter is indirect (measuring metalbinding by observing secondary structure), the data are much more easily rationalized than for the fluorescence data. The proposed folding mechanism easily explains changes in CD spectra: histidine binding induces α -helix formation.

3.3.2.3 HPLC Studies

Addition of Zn^{2+} seems to encourage sample homogeneity. When apo-3HT-C is analyzed by RP-HPLC, two dominant species appear to be in solution while upon the addition of Zn^{2+} , only one species is observed (Figure 3.4). Much research has been performed investigating the folding mechanism of Trp-cage, and most evidence indicates that the peptide folds in a simple 2-state mechanism. Under these HPLC conditions (room temperature and elution with organic solvent), perhaps both native and "relaxed" peptide are observed in the chromatogram of the apo-peptide. Metal-binding could shift the equilibrium toward natively-folded peptide in solution. Another postulate is that the N-terminal, glycine-rich arm of the peptide added to the Trp-cage adopts open and closed fluxional structures. Upon binding of Zn^{2+} to the three introduced histidines, this arm could be forced into the closed position. Once a metal ion is bound to all three histidines, this flexible arm can no longer open to explore other structural states. A final explanation is that addition of a zinc ion changes the charge state of the peptide and causes a different retention time on the hydrophobic column (that is coincident with an apo-protein retention time). This change could be due to ion binding to either denatured or folded peptide. If the explanation is this simple, it still supports the hypothesis that the Zn^{2+} is actually bound to the protein and not just weakly associated.

As with the fluorescence data, the mechanism explaining this induction of homogeneity is not easily explained, and this experiment is in no way definitive. However Zn²⁺ addition causes a marked affect upon the homogeneity of the sample. This finding supports CD data indicating the presence of only two principle species in solution during binding reactions.

3.4 Conclusions

Ultimately, data indicate that these Trp-cage mutants bind Zn^{2+} and that binding shifts equilibrium to homogeneity. Fluorescence data showed that addition of Zn^{2+} to the peptide caused a change in tryptophan environment. CD data showed that addition of Zn^{2+} induced α -helix formation. As observed in HPLC data, holo-mutants are homogeneously structured peptides while apo-protein is a mixture of two species. These data support the hypothesis that a metal-binding Trp-cage was designed that binds Zn^{2+} with an affinity in the low micromolar range.

The Trp-cage was originally designed for use as a tool in computational chemistry. It was thought that such a small, fast-folding peptide could be more easily modeled than other, larger proteins. Moreover, modeling of metal-binding sites in proteins has proven especially difficult. If algorithms could be optimized to properly model the empirical Trp-cage data, then modeling of a metal-binding mutant could further test the efficacy of such algorithms.

Another application of this peptide could be as a peptidic fusion tag for other molecules. Such fusion molecules could be used to seek a variety of targets *in vivo*. The Trp-cage is a small, autonomously folding peptide that could cause limited interference with its fusion partner or other biomolecules. Zn^{2+} -bindings sites have been shown to bind other metals with high affinity. While not shown in this thesis work, 3HT-C or 3HT-Cb could be used to deliver therapeutic metal ions such as platinum, ruthenium, and rhenium to tissues *via* its fusion partner. Technetium-99m could also be bound by such peptides and be used to image specific tissues depending on the fusion partner.

3HT-C and 3HT-Cb readily serve as examples of *de novo*, redesigned peptides. Further characterization could lead to more optimized structures, and such an optimization process could test knowledge of protein folding mechanisms and proteinstabilizing features. Already, the Trp-cage has been shown to have relatively unique tryptophan-proline interactions that are highly stabilizing.³ Difficulty in *de novo* design of proteins has already proven to researchers that we do not have this area of science "all figured out". A small system such as 3HT-Cb containing several secondary structure elements and a metal-binding site could serve as a model protein for future work in *de novo* design research.



Figure 3.1 Proposed peptide structure for Trp-cage mutants. The sequences of mutants can be compared to that of "wt" Trp-cage, TC5b. Peptide sequences are shown is single-letter amino acid abbreviations.



Figure 3.2 Titrations of Trp-cage mutants monitored by dependence of fluorescence emission at 350 nm upon addition of Zn^{2+} . Top: titration of 3HT-C. Bottom: titration of 3HT-Cb. Raw data is shown as closed circles and non-linear least squares fittings are shown as dashed lines. Error bars for the Zn^{2+} titration of 3HT-C were generated from calculation of the standard deviation of data from two separate trials. No error bars were generated for the Zn^{2+} titration of 3HT-Cb since only one trial was performed. Insets show representative fluorescence spectra.



Figure 3.3 Representative titration of 3HT-Cb monitored by dependence of molar residue ellipticity at 222 nm upon addition of Zn^{2+} . Raw data is shown as closed circles, and the non-linear least squares fitting is shown as a dashed line. Error bars were generated by calculating the instrument noise at 222 nm. Inset shows CD spectra.



Figure 3.4 HPLC analysis of absorbance at 280 nm of Zn^{2+} -binding by 3HT-C. Apopeptide is shown in red while Zn^{2+} -tritrated peptide (10 equivalents of metal ion) is shown in blue.

Notes

- (1) Streicher, W. W.; Makhatadze, G. I. *Biochem.* **2007**, *46*, 2876-2880.
- (2) Naduthambi, D.; Zondlo, N. J. J. Am. Chem. Soc. 2006, 128, 12430-12431.
- (3) Barua, B.; Andersen, N. H. Letters Peptide Sci. 2002, 8, 221-226.
- (4) Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. *Nat. Struct. Biol.* **2002**, *9*, 425-430.
- (5) Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. J. Phys. Chem. B 2006, 110, 3759-3763.
- (6) Fersht, A. R. In *Structure and Mechanism in Protein Science*; Julet, M. R., Ed.; W.H. Freeman and Company: New York, 1999, p 523-532.
- (7) Ye, Y.; Lee, H.-W.; Yang, W.; Shealy, S. J.; Yang, J. J. J. Am. Chem. Soc. 2005, 127, 3743-3750.
- (8) Kelly, S. M.; Jess, R. J.; Price, N. C. *Biochim. Biophys. Acta* **2005**, *1751*, 119-139.
- (9) Engelborghs, Y. J. Fluorescence **2003**, *13*, 9-16.
- (10) Royer Chem. Rev. 2006, 106, 1769-1784.
- (11) Stollar, E. J.; Gelpi, J. L.; Valankar, S.; Golovin, A.; Orozco, M.; Luisi, B. *Proteins* **2004**, *57*, 1-8.
- (12) Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. J. Am. Chem. Soc. 2002, 124, 12952-12953.
- (13) Harris, D. C. In *Quantitative Chemical Analysis*; Julet, M. R., Ed.; W.H. Freeman and Company: New York, 1999, p 549.
- (14) Woody, R. W. Meth. Enzymol. 1995, 246, 34-71.

CHAPTER 4: METAL-BINDING UBIQUITIN MUTANTS, 3HPU AND 3HIU

4.1 Introduction

Ubiquitin is a small protein that can be reversibly attached to other proteins and serves as a signaling moiety. Its mode of action has been compared to that of phosphorylation.^{1,2} However, due to the added complexity of ubiquitin, this signaling mechanism is more versatile than phosphorylation. The mode of ubiquitination of target proteins serves as a method to assign molecules to a variety of destinations.³ These assignments are read by a variety of dedicated Ubq-binding proteins. Not only are they capable of distinguishing the mode of ubiquitination, but they also bind to different surfaces of the molecule (Ile44, D58-centered hydrophobic face, C-terminal tail, Ile36centered surface).³ The known modes of ubiquitination are mono-, multi-, and oligoubiquitination.⁴ In mono-ubiquitination, a single ubiquitin protein is attached to its target. Multi-ubiquitination attaches ubiquitin monomers to more than one location on its target. Both mono- and multi-ubiquitination serve as signals for endocytosis pathways.¹ Polyubiquitination, attachment of chains of ubiquitin oligomers to targets, sends proteins to the proteasomal degradation or DNA repair pathways. Aberrations to ubiquitin signaling networks lead to disease and other malignancies (i.e. oncogenesis of several cancers).^{1,4,5} Such pathways have been proposed as targets for pharmaceuticals, and one proteasomal inhibitor (bortenzomib) has been successful in causing apoptosis of plasma cells.^{2,5}

Nature has optimized the stability of this protein so that it can be used as a small, robust, soluble, versatile signaling molecule. If its structure was not highly stable, then its topology could not be properly recognized by Ubq-binding proteins and signals could be misread. In the studies described in this thesis, ubiquitin's stability was exploited when building metal-binding sites onto it. In a sense, nature has provided a perfect scaffold for building metal-binding sites to be studied in a protein context. Furthermore,

many structural studies have been performed on this protein, and comparisons to literature reports for wt Ubq can be easily made.

4.2 Materials and Methods

4.2.1 Ubiquitin Mutant Design

Since the 9-10 loop of ubiquitin had been identified as a site amenable to insertion mutations^{6,7}, this site was chosen for insertion of a Gly-His-Gly-His-Gly-His, 6 amino acid loop. Glycine was chosen as a flexible spacer due to its ability to access dihedral angles unattainable in other protein residues. This mutant was named "3HIU" for "<u>3-H</u>is Insertion Mutant of <u>U</u>biquitin". It has 82 residues, a calculated mass of 9.1 kDa, a calculated pI of 7.5 and a calculated extinction coefficient of 1280 cm⁻¹ M⁻¹ (DNAstar, Inc).

The 35-36 loop of ubiquitin cannot accommodate insertion mutations as easily as the 9-10 loop.^{6,7} Therefore, point mutations were made at residues flanking this loop. To choose mutation sites, the crystal structure of ubiquitin (PDB accession code 1UBQ) was loaded into Swiss PDB viewer⁸, and acids were systematically mutated to His using the mutation function provided in the viewer software. Residues Ala28, Gln31, and Asp39 were chosen for mutation to Histidine because they created a site for *fac*-metal-binding. This mutant was named "3HPU" for "<u>3-His Point Mutant of Ubiquitin</u>". It has 76 residues, a calculated mass of 8.7 kDa, a calculated pI of 8.5 and a calculated extinction coefficient of 1280 cm⁻¹ M⁻¹ (DNAstar, Inc).

Proposed structures and their protein sequences are shown in Figure 4.1.

4.2.2 Plasmid DNA Mutations

Mutations were made to wt ubiquitin DNA received in a pRS plasmid (received from Debra Ferraro, Ph.D., Robertson Laboratory, The University of Iowa, Dept. of Biochemistry). This plasmid also carries a T7 promoter and a gene for Amp resistance. All DNA primers were ordered from IDT (Coralville, IA) with desalting but no further purification. Mutations were made by PCR with *Pfu turbo* polymerase and the manufacturer's instructions (Strategene, Agilent Technologies, Co.). During PCR, the pRS plasmid containing DNA coding for the mutant of interest was amplified. *E. coli* DH5α cells (Invitrogen, Corp.) were transformed with mutated DNA and their plasmids purified following a mini-prep kit manual (Qiagen, Inc.) in order to make DNA plasmid stocks. All DNA sequences were verified by sequencing performed by the DNA Facility (The University of Iowa).

4.2.2.1 Cloning of 3HIU

Forward and reverse DNA primers were designed to make the loop insertion mutation while annealing to 13 or 14 bases on either side of the insertion. Forward: 5'-CAA GAC ATT AAC CGG CCA CGG CCA CGG CCA CGG TAA AAC CAT AAC. Reverse: 5'-GTT ATG GTT TTA CCG TGG CCG TGG CCG TGG CCG GTT AAT GTC TTG. (45 b.p., $T_m = 79$ °C, Hairpin $T_m = 50$ °C.) A total of 18 PCR cycles were used with the annealing temperature for the first 5 cycles lowered by 5 °C. (The DNA mutation entailed a large insertion of 18 b.p., and the T_m of the primer when pairing to wt DNA before it had been mutated was low and had to be accommodated).

4.2.2.2 Cloning of 3HPU

Because the three amino acid residues of interest were coded in DNA spread out over 36 b.p., point mutations had to be performed in two rounds. In the first round, Ala28 and Gln31, were mutated, and in the second round, Asp39, was mutated. Forward and reverse DNA primers were designed to make four point mutations for the first round of PCR. Forward: 5'-CCA TCG AAA ACG TGA AAC ACA AAA TCC ACG ATA AAG AAG GCA TCC CGC. Reverse: 5'-GCG GGA TGC CTT CTT TAT CGT GGA TTT TGT GTT TCA CGT TTT CGA TGG. (48 b.p., $T_m = 78$ °C, Hairpin $T_m = 46$ °C). A total of 18 PCR cycles were used with the annealing temperature for the first 5 cycles lowered by 5 °C (only because this reaction was performed at the same time as that to make 3HIU). The product of this PCR reaction was called "2HPU" for <u>2-His Point</u> Mutant of <u>U</u>biquitin." Forward and reverse DNA primers were designed to make two point mutations for the second round of PCR. Forward: 5'-GGC ATC CCG CCG CAC CAG CAG CGT CTG. Reverse: 5'-CAG ACG CTG CTG GTG CGG CGG GAT GCC. (27 b.p., $T_m = 80$ °C, Hairpin $T_m = 49$ °C). A total of 16 PCR cycles were used with the annealing temperature for the first 5 cycles lowered by 5 °C since making this mutation proved difficult.

4.2.3 3HIU and 3HPU Expression and Purification

Both ubiquitin mutants were expressed and purified in the same way. BL21(DE3)star competent cells (Invitrogen, Corp.) were transformed with the appropriate, purified plasmid DNA and plated on LB agar + 200 µg/mL Amp plates. Plates were incubated at 37 °C for at least 18 hours. A 5 mL LB + 200 µg/mL Amp subculture was grown from a single colony and incubated at 37 °C for 8-10 hours. An aliquot was then diluted 1:1000 in 150 mL of LB + 200 µg/mL Amp media to make a larger subculture that was incubated at 37 °C overnight. Finally, an aliquot of the overnight culture was diluted 1:60 in 1.5 L of LB + 100 µg/mL Amp media. This culture was incubated at 37 °C until the OD₆₀₀ reached ~0.4. The culture was then induced by the addition of IPTG to a final concentration of 0.5 mM, and protein expression continued for 3 hours under the same incubation conditions. Cells from each 1.5 L culture were harvested separately by centrifugation at 6000 x g for 10 minutes at 4 °C. Cell pellets were stored at -80 °C.

Purification was performed using slight modifications to a previously reported protocol.⁶ Briefly, cell pellets were resuspended in 30 mL of cold (4 °C) 50 mM MOPS, 10 mM EDTA, 1 mM DTT, 1 mM PMSF, pH 7.65 buffer before lysis with a French Press. Cells were passed through the pressure cell at least 3 times at a pressure of ~1600

psi. Insoluble cell particles were removed by centrifugation (20 000 x g for 45 minutes at 4 °C) before passing the solution through a 0.45 μ m filter (Millipore, Corp.). Streptomycin sulfate was added to a final concentration of 5% (w/v), and the solution was stirred for 30 minutes at 4 °C. Precipitated nucleic acids were removed by centrifugation (16 000 x g for 20 minutes at 4 °C). (NH₄)₂SO₄ was added to 40% saturation, and the solution was stirred for 15 minutes at 4 °C. Precipitated proteins were removed by centrifugation (12 000 x g for 25 minutes at 4 °C. More (NH₄)₂SO₄ was added to bring its concentration to 85% saturation, and the solution was stirred for another 15 minutes at 4 °C. This ubiquitin mutant-containing protein fraction was harvested by centrifugation (12 000 x g for 25 minutes at 4 °C), and the pellet was resuspended in a minimal volume (typically ~ 5-10 mL) of cold (4 °C) 50 mM NH₄C₂H₃O₂, pH 4.5 buffer. The protein was dialyzed extensively against the same buffer at 4 °C.

Desalted protein was applied to a 10 mL (height = 3 cm) CM52 (Whatman, Plc.) cation exchange column equilibrated with 50 mM NH₄C₂H₃O₂, pH 4.5 buffer. Elution was performed by altering the pH in a step gradient: 5 CV of 50 mM NH₄C₂H₃O₂, pH 4.5; 8 CV of 100 mM NH₄C₂H₃O₂, pH 5.5; 5 CV of 100 mM MES, pH 6.5; 5 CV of 100 mM MOPS, pH 7.5; 8 CV of 100 mM MOPS, 1 M NaCl, pH 7.5. 3HIU and 3HPU eluted upon application of the pH 6.5 buffer. Further purification was performed by passing the 3HIU or 3HPU-containing fractions through a YM30 membrane (30 kDa NMWL, Millipore, Corp.) using an ultrafiltration stirred cell. 3HIU or 3HPU protein was then concentrated using a stirred cell fitted with a YM3 membrane (3 kDa NMWL, Millipore, Corp.). Concentrated protein was dialyzed extensively against 20 mM MOPS, 100 mM NaCl, pH 7 buffer. In a final purification step to remove any metal ions from solution, protein was dialyzed against the same buffer containing 5 g/L chelex resin (Bio-Rad Laboratories, Inc.). Purity of protein samples was determined by Coomassie-stained SDS-PAGE, and the identity was confirmed by ESI-MS.

4.2.4 SDS-PAGE Analysis

Laemmli buffer systems were used in the preparation and running of SDS-PAGE experiments.⁹ Samples were thermally denatured in loading buffer by heating at 95 °C for at least 10 minutes. Samples were loaded onto 4%/25% (discontinuous stacking/separating) gels of 0.75 mm thickness. The gels were run at 300 V for the first 5 minutes and at 200 V until the dye-front left the gel. Gels were stained with Coomassie R-250 using standard techniques.¹⁰

4.2.5 ESI-MS Analysis

ESI-MS data were collected on a Thermo Finnigan LCQ Deca Spectrometer (University of Iowa Mass Spectrometry Core Facility) in positive ion mode. Formic acid was added to samples to a final concentration of 0.1% before injection into the instrument and infusion at for ~ 10 minutes at 10 μ L/minute. Data was collected while the sample flowed at a rate of 3 or 8 μ L/minute. The following conditions were optimized: spray voltage, 4.52 kV; sheath gas, 33.93 units; capillary temperature, 250 °C; capillary voltage, 26.64 V. Deconvolution was performed using Bioworks 3.2 software (Thermo, Fisher Scientific, Inc.).

4.2.6 Folding and Structure Preservation of Apo-3HPU

Qualitative determination of this mutant's folding was inspected using 2-D ¹H-NMR (DQF-COSY and NOESY). Spectra were collected on a Bruker Advance 600 MHz spectrometer (NMR Facility, The University of Iowa, Dept. of Chemistry) at 25 °C. A 0.95 mM sample of 3HPU in 90% H₂O:10% D₂O and 20 mM MOPS, 100 mM NaCl, pH 7 buffer was prepared with a 5 μ M DSS internal standard. WATERGATE or presaturation pulses were used to suppress the H₂O resonance. 1.11 equivalents of ZnCl₂ were added to the protein, but the addition of this salt caused so much peak broadening that correlations could not be detected.

4.2.7 Ubiquitin Mutant Stability

The structure of 3HIU and 3HPU at various temperatures was investigated by CD. Spectra were collected using an Olis Cary-17 DS Conversion Spectrophotometer. Spectra were collected (of samples in 1 mm path-length quartz cuvettes) from 200 - 290 nm at 1 nm *per* point resolution and a bandwidth of 3 nm with a variable integration time determined from the PMT voltage. Protein was diluted to ~40 μ M in 5 mM MOPS, 25 mM NaCl, pH 7 buffer. Spectra were baseline-corrected using the same buffer as the reference. Temperature (25 - 95 °C) was controlled by the temperature controller provided with the instrument. Samples were allowed to equilibrate to set temperatures for 10 minutes before scanning. Three spectra were averaged and then smoothed (11-point, moving, weighted-average using least squares calculations) using the software provided by Olis. Spectra were normalized for concentration and number of amino acid residues before zeroing at 260 nm.

4.2.7 Zn²⁺-binding by Ubiquitin Mutants

The effect of Zn^{2+} addition upon 3HIU and 3HPU structure was investigated by CD. Spectra were collected on the same spectrophotometer described above under the same conditions except that the temperature was held constant at 25 °C for these experiments. ZnCl₂ diluted in buffer was added in aliquots at Zn:protein ratios varying from 0 – 5. Samples were allowed to equilibrate for at least 5 minutes before scanning. Error bars were generated by calculating the instrument noise at 207 nm. K_d values were determined by fitting [θ]₂₀₇ data to a non-linear, least squares, single binding site algorithm.¹¹

4.3 Results and Discussion

4.3.1 Expression and Purification of Ubiquitin Mutants

Both proteins can be purified to $\geq 90\%$ purity (judged by Coomassie-stained SDS-PAGE gels) relatively easily (**Figure 4.2**). These mutants are so over-expressed that bulk purification by (NH₄)₂SO₄ precipitation removes a large proportion of contaminants. While it looks like cation exchange completely cleans the sample (as seen in lane 8), it is evident upon concentration of these fractions with a YM-30 membrane that high-mass contaminants persist after this purification step. Therefore, this crude size-exclusion step is needed to attain high purity. In these experiments, it was found that concentrated ubiquitin mutants (800 µM) were stable at pH 4.5 – 7.5 and at NaCl concentrations ranging from 25 mM to 1 M at 4 °C for at least 6 months. Buffer changes are trivial with this protein due to its high solubility.

Once mutants were purified, their identity was confirmed by ESI-MS (**Figure 4.3**). Deconvolution of spectra indicated an apparent mass of 9146.3 Da for 3HIU and 8662.0 Da for 3HPU. The calculated, isotopic masses for these proteins are 9141.9 Da and 8656.7 Da for 3HIU and 3HPU, respectively.¹² The uncertainty for such deconvolution-determined masses is ≥ 2 Da (Lynn Teesch, Mass Spectrometry Core Facility Director, The University of Iowa, personal communication). Furthermore, the monoisotopic mass calculators are most reliable for small proteins (<3000 Da). ESI-MS confirms the identity of these proteins since the agreement between mass determinations is within 0.04% for 3HIU and within 0.06% for 3HPU. Other peaks in the spectrographs' peak envelopes can be attributed to multiple adducts of sodium. They appear as artifacts in these deconvoluted spectra.

Qualitative determination of folding was inspected by 2D ¹H-NMR (DQF-COSY and NOESY) of 3HPU. This protein was chosen for NMR analysis because previous studies with more 3HIU-like proteins showed that insertion mutants at the 9-10 loop conserved a wt ubiquitin structure. Not as much is known about point mutants at the 35-36 turn. Of the two mutants, 3HPU was considered the more likely candidate to be misfolded. Spectra show very disperse peaks indicative of a well folded protein (data not shown). Peaks near 2.7, 2.9, and 4.6 ppm were buried beneath the resonances of water and MOPS buffer. However, patterns of ¹H interactions can be seen in other areas of the spectra. Further investigations into protein folding by CD spectroscopy are described below.

4.3.2 Stability of Ubiquitin Mutants

To more definitively determine if mutations caused destabilization of the ubiquitin mutants' structures, changes in CD spectra were measured at several temperatures (**Figure 4.4**). CD spectra of ubiquitin mutants are very similar to that of wt Ubq and are dominated by α -helical characteristics (minimum at 207 with a shoulder at 222 nm) despite only being ~20% α -helical.¹³ Also like wt Ubq, the T_m of these mutants in solution could not be determined using thermal denaturation since their melting points are > 95 °C.¹⁴ When the mutants were studied, their spectra never adopted a random coil signature (minimum at 199 nm). For 3HIU, [θ]₂₂₂ only changed by 3% and [θ]₂₀₇ changed by 4%. For 3HPU, [θ]₂₂₂ changed by 13% and [θ]₂₀₇ changed by 10%. While it seems that 3HPU's structure is more affected by higher temperatures, it is shown that it still retains its structure. As described above, it was expected that 3HPU could be slightly less stable than 3HIU. Overall, these studies showed that mutations made in order to build metal-binding sites onto this scaffold did not greatly destabilize either protein.

4.3.3 Zn²⁺-Binding by Ubiquitin Mutants

The most efficient way to study metal binding was by CD although this method only allows indirect observation of metal-binding (**Figure 4.5**). Unlike for the Trp-cage

mutants, these spectra could be collected at room temperature since these mutants maintain their structure very well.

Upon addition of metal, it appeared that these mutants lost some α -helical structure since the magnitude of the α -helix bands at 207 and 222 nm decreased. While metal-binding may indeed cause a loss of helical structure, it does not mean that the overall structure is destabilized. Zn²⁺-binding could be stabilizing other 2° structure elements such as β sheets, turns, and loops that would cause a decrease in magnitude of ellipticity values at these wavelengths. CD spectra are an average of all structures present in a protein. Therefore, it is not a perfect tool for observing effects of Zn²⁺-binding upon protein structure.

On the other hand, these titrations did serve as a first line of evidence supporting the hypothesis that metal-binding ubiquitin mutants were made. Addition of Zn^{2+} caused changes in $[\theta]_{207}$ by ~20%. Furthermore, binding curves support a 1:1 stoichiometry binding model. When calculating K_d values for the Trp-cage mutants, (which are much more helical in structure), $\Delta[\theta]_{222}$ values were used. While both ubiquitin mutants have shoulders at 222nm, it was decided that $\Delta[\theta]_{207}$ values would be used for calculations of these K_d values. Rationalization of this choice is described below. 3HIU binds Zn²⁺ with a K_d of 16 ± 15 µM (two trials), and 3HPU binds Zn²⁺ with a K_d of 36 ± 2 µM (two trials). The K_d for Zn²⁺-binding by 3HIU when inputting $[\theta]_{222}$ data was 15 ± 2 µM, which is in good agreement with the K_d calculated with $[\theta]_{207}$ data. The K_d for Zn²⁺binding by 3HPU when inputting $[\theta]_{222}$ was 256 ± 81 µM and is different by an order of magnitude. Intrinsic noise of CD data combined with the subtlety of change in $[\theta]_{207}$ values made K_d determination difficult, but it can be said that 3HIU binds Zn²⁺ stronger than 3HPU. (This same finding was observed in Tc¹-binding studies as shown in chapter 6.)

Ellipticity at 207 nm would be more greatly affected by formation of other secondary structures such as β -turns and random coils/loops. β -turns typically show

positive absorbances between 195 and 215 nm but no absorbances near 222 nm. Random coils show a negative absorbance at ~207 nm. Whether metal-binding induces β -turn-like binding loops or causes unwinding of nearby structures into more random coils, it does cause measurable changes in secondary structure without completely unfolding the protein.

4.4 Conclusions

These ubiquitin mutants support findings in the Robertson Laboratory that ubiquitin is a useful scaffold for investigating the effect of mutations. This scaffold is extremely robust; neither mutant could be thermally denatured. Not only did these mutations not destabilize the protein, but they succeeded in introducing metal-binding sites as inferred from CD data. Unfortunately, neither Zn^{2+} nor ubiquitin offer many spectroscopic handles for observing metal-binding. Perturbation of d-orbitals upon ligand binding cannot be investigated with Zn^{2+} because it is a d¹⁰ ion. The abundant nucleus is not NMR-active. Ubq has no tryptophan residues to monitor by fluorescence. Therefore, CD was the most efficient method for investigating metal binding. Since ubiquitin is well characterized (*i.e.* by NMR and X-ray crystallography), metal-binding by these mutants could be examined more thoroughly in further studies.

Mutant 3HIU shows that loop mutagenesis can be used to make a metal-binding site. Previous work in the Franklin Laboratory successfully applied this approach to the design of *engrailed*/calmodulin chimeras. However, the previously borrowed Ca²⁺-binding site was chosen because it kept its orthogonal turn geometry when placed into a new scaffold. When a robust scaffold such as ubiquitin is used, a much more flexible loop can be introduced. Mutant 3HPU was designed using a grafting approach. Instead of inserting an entire new loop, an orthogonal turn was found and histidine residues were introduced to make an astacin-like binding site. Both mutants utilize a variety of *de novo* design principles.

Besides using these mutants for models to examine re-design of metal-binding loops into proteins, there are other applications much like those discussed for the Trpcage mutants. Like that peptide, these mutants could be fused to other proteins as a tag. There is already a precedent for using Ubq as a tag, although for a different purpose. Ubq has been fused to other proteins to increase their bacterial expression with great results.¹⁵ Such an expression system could allow for high-throughput manufacture of therapeutic metal-binding Ubq mutants in which the tag targets specific tissues *in vivo*. As shown in chapters 5 and 6, 3HIU has been shown to have an affinity for ruthenium and technetium complexes.

Ubiquitin is a useful scaffold for *de novo* metalloprotein design and could be used to test site-designs for a variety of metal ions from hard ions like Ca²⁺ to softer ions like Pb²⁺. Furthermore, effects of mutations could be used to determine stabilizing elements in proteins. Previous work in the Robertson laboratory showed that there is a stability dependence upon site of mutation. Making changes to other parts of this protein could serve *de novo* design research by mapping responses of proteins to a variety of mutations including incorporation of metal-binding sites.



Figure 4.1 Proposed structures of 3HIU (top) and 3HPU (bottom). Residues conserved from wt ubiquitin are shown in green while mutated residues are shown in yellow. The sequences at the right are shown using single-letter amino acid abbreviations.



Figure 4.2 Representative SDS-PAGE analysis of 3HPU purification. The arrow indicates where 3HPU appears in gel lanes. Purified 3HPU can be seen in the right-most lane. All samples are labeled according to how they were purified. 3HIU purifies similarly.



Figure 4.3 Deconvoluted ESI-MS data collected for 3HIU (top) and 3HPU (bottom). Arrows indicate positions of apo-protein peaks and their apparent masses.



Figure 4.4 CD spectra of ubiquitin mutant melting studies. Top: melting of 3HIU. Bottom: melting of 3HPU.



Figure 4.5 Representative titrations of Ubq mutants monitored by dependence of molar residue ellipticity at 207 nm upon addition of Zn^{2+} . Top: titration of 3HIU. Bottom: titration of 3HPU. Raw data is shown as closed circles, and non-linear least squares fittings are shown as dashed lines. Error bars were generated by calculating the instrument noise at 207 nm. Insets show CD spectra

Notes

- (1) Schnell, J. D.; Hicke, L. J. Biol. Chem. 2003, 278, 35857-35860.
- (2) Dikic, I.; Crosetto, N.; Calatroni, S.; Bernasconi, P. *Eur. J. Cancer* **2006**, *42*, 3095-3102.
- (3) Harper, J. W.; Schulman, B. A. *Cell* **2006**, *124*, 1133-1136.
- (4) Mukhopadhyay, D.; Riezman, H. *Science* **2007**, *12*, 201-205.
- (5) Crosetto, N.; Bienko, M.; Dikic, I. Mol. Cancer Res. 2006, 4, 899-904.
- (6) Ferraro, D. M.; Hope, E. K. R., A.D. J. Mol. Biol. 2005, 352, 575-584.
- (7) Ferraro, D. M., Ferraro, D.J., Ramaswamy, S., Robertson, A.D. *J. Mol. Biol.* **2006**, *359*, 390-402.
- (8) Guex, N.; Peitsch, M. C. *Electrophoresis* **1997**, *18*, 2714-2723.
- (9) Laemmli, U. K. *Nature* **1970**, *227*, 680-685.
- (10) Sambrook, J.; Russell, D. W. In *Molecular Cloning, A Laboratory Manual*; Third ed.; Argentine, J., Ed.; Cold Spring Harbor Laboratory Press: New York, 2001; Vol. 3, p A8.46-A8.47.
- (11) Ye, Y.; Lee, H.-W.; Yang, W.; Shealy, S. J.; Yang, J. J. J. Am. Chem. Soc. 2005, *127*, 3743-3750.
- (12) Hoogland, G. E.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press: 2005.
- (13) Ermolenko, D. N.; Dangi, B.; BGvritishvili, A.; Gronenborn, A.; Makhatadze, G. I. *Biophys. Chem.* **2007**, *126*, 25-35.
- (14) Ibarra-Molero, B.; Loladze, V. V.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. *Biochem.* **1999**, *38*, 8138-8149.
- (15) Catanzariti, A.-M.; Soboleva, T. A.; Jans, D. A.; Board, P. G.; Baker, R. T. *Prot. Sci.* **2004**, *13*, 1331-1339.
CHAPTER 5: REACTIONS OF RUTHENIUM COMPLEXES WITH UBIQUITIN MUTANTS

5.1 Introduction

Ruthenium complexes have been shown to have antitumor and anti-metastatic properties by virtue of their ability to create DNA lesions.^{1,2} Design of ruthenium complexes has been proposed as an objective for pharmaceuticals because these complexes are stable, well characterized and have predictable structures in solid and solution.³ Currently, two Ru-complexes have completed phase I clinical trials.⁴ The first was NAMI-A (imidazolium *trans*-[tetrachloro(dimethylsulfoxo)(imidazole) ruthenate(III)]) followed by KP1019 (indazolium *trans*-[tetrachlorobis(1*H*-indazole) ruthenate(III)]). However, *cis/trans*- [Ru(dmso)₄X₂] (where X = Br or Cl) had been screened for antitumor activity as early as 1988.⁴

It is not known whether Ru^{III} or Ru^{II} is the active species in these systems. Some researchers have proposed that Ru^{III} complexes are reduced to relatively more reactive Ru^{II} species upon entering hypoxic, more acidic tumor tissues *in vivo*.^{1,2} One proposed mode of cancer cell delivery for these ions is through transferrin.¹ In cancer cells, there is a greater need for iron and transferrin is directed to these cells. What is known about Ru-complexes is that Ru^{III} and Ru^{III} ions are capable of binding nitrogen-containing heterocycles such as DNA. Like cisplatin, complexes preferentially bind to N7of guanine bases in DNA, but do not seem to form the characteristic intra-strand cross-links like cisplatin.¹ Ru-complexes are typically octahedral, and steric hindrance is thought to provent such cross-linking. However, Ru-pharmaceuticals can form inter-strand DNA or protein-DNA crosslinks effective at blocking DNA and RNA synthesis.² Therefore, DNA-binding is still assumed to be responsible for the effectiveness of ruthenium pharmaceuticals against cancer.

On a completely different note, Ruthenium is also used in catalysts such as those developed by Noyori to catalyze asymmetric transfer hydrogentation reactions.⁵ While

small-molecule catalysts are typically easy to characterize and resistant to harsh conditions, biocatalysts are more regio-, stereo- and enantioselective due to their preorganized active sites optimized for substrate recognition. They also perform at ambient temperatures and pressures, offer higher turnovers and are active under aqueous, environmentally-friendly conditions.⁶ Introduction of non-native metal ions (such as Ru^{II}) into proteins can lead to development of biocatalysts able to perform chemistries inaccessible to physiologically-relevant metals.⁷

Despite the rich chemistry and spectroscopy of Ru-complexes, they are not commonly studied in protein contexts because of the difficulty of incorporating these ions into biomolecules. These kinetically-inert ions can cause protein misfolding and aggregation by forming inappropriate cross-links between protein residues. Therefore, Ru-binding proteins must be designed with pre-organized, surface-accessible binding sites. Study of such sites built onto robust scaffolds can aid elucidation of principles governing Ru-binding to physiological proteins.

5.2 Materials and Methods

5.2.1 Synthesis of Ruthenium Complexes

5.2.1.1 Synthesis of Ru(dmso)₄Cl₂ - "RuA"

 $Ru(dmso)_4Cl_2$ was synthesized from $RuCl_3 \cdot xH_2O$ and DMSO using a previously reported procedure.⁸ Briefly, 1.011 g of $RuCl_3 \cdot xH_2O$ (4.14 mmol) and 20 mL of DMSO were combined in a flask, degassed, and refluxed under nitrogen for 5 minutes. The volume was reduced by half by vacuum distillation of DMSO. After cooling the solution, acetone was added to initiate precipitation. The yellow solid was removed by filtration, washed with acetone and diethyl ether, and allowed to dry in air. When the filtrate was cooled to 4 °C, more solid was formed overnight. This solid was removed by filtration and washed as described above. Solid $Ru(dmso)_4Cl_2$ was recrystallized from hot DMSO.

The chemical identity of the yellow solid was verified by elemental analysis. (Calculated: C, 19.83; H, 4.99; Cl, 14.63; S, 26.47. Found: C, 19.97; H, 5.13; Cl, 14.61; S, 26.55.) The yield was 1.396 g (69%), which is comparable to the reported yield of 72%.⁸

5.2.1.2 Synthesis of [Ru(tacn)(dmso)₂Cl]Cl – "RuB"

[Ru(tacn)(dmso)₂Cl]Cl was synthesized using a previously reported procedure.⁹ Briefly, 576 mg of Ru(dmso)₄Cl₂ (1.2 mmol) was reacted with 170 mg of tacn (1.3 mmol, ~1.1 equivalents) in 20 mL of toluene for 1 hour under reflux in a $N_{2(g)}$ atmosphere. The reaction vessel was cooled on ice for 30 minutes before the pale grey/yellow solid was removed by filtration. This solid was washed with toluene and diethyl ether.

The chemical identity of the solid was verified by elemental analysis. (Calculated: C, 26.26; H, 5.95; N, 9.19; S, 14.02; Cl, 15.50. Found: C, 26.48; H, 6.10; N, 9.18; S, 13.95; Cl, 15.66.) The yield was 520 mg (96%). A yield for this reaction was not reported in the literature since this product was used directly in a further synthesis. However, the reported overall yield was 94%. Therefore, the yield of this intermediate must have also been very high.⁹

5.2.1.3 Preparation of $[Ru(tacn)(dmso)_2(OTf)]^{n+}$ - "RuC"

In an attempt to exchange the chloride ions of $[Ru(tacn)(dmso)_2Cl]Cl$ with triflate (to increase the reactivity of the $Ru^{II}(tacn)$ complex toward protein binding), this complex was treated with Me₃SiOTf.¹⁰ In this preparation, 51.5 mg of $[Ru(tacn)(dmso)_2Cl]Cl$ (0.11 mmol) and ~120 mg Me₃SiOTf (0.55 mmol, ~5 equivalents) were allowed to react in ~5 mL of acetonitrile under an inert atmosphere at room temperature for 5 minutes. The solvent was removed under vacuum until the solid was dry. A minimal amount of DMSO was added to afford an orange/brown solution. Upon addition of acetone and

diethyl ether, a red/orange oil developed. After letting the oil stand at room temperature for ~12 hours, a yellow solid developed. This solid was washed with acetone then diethyl ether.

Elemental analysis indicated that the product contained the following elemental contributions: C, 23.50, H, 4.76, N, 7.31, F, 10.44, S, 17.08. Since the molar ratio of F:S is approximately 1:1, the ratio of triflate:dmso in this compound must be 1:2. Consequently, the compound likely contains only one triflate anion and two DMSO molecules (consistent with the C:F:S ratio of 11:3:3). The other counterions/ligands must be other C, N, F, S-free anions such as chloride or hydroxide. The percent values are consistent with RuC being one of the following compounds containing the Ru(tacn) fragment: a Ru^{II} monomer with a chloride or hydroxide counterion, a Ru^{III} monomer with hydroxide counterions, or an oxo-bridged Ru^{III} dimer with either two hydroxide counterions and purification of the products is still under investigation, the initially obtained material was used for preliminary protein binding studies.

5.2.2 Reaction of Ubiquitin Mutants with Ruthenium Compounds

5.2.2.1 Method 1, Reactions in air at 37 °C

Powders of RuA, RuB, and RuC, were solubolized in 20 mM MOPS, 100 mM NaCl, pH 7 buffer. 3HIU and 3HPU (in the same buffer) were mixed with ruthenium compounds in 0.6:1 and 1.2:1 Ru:protein ratios at a final protein concentration of 100 μ M in a volume of 300 μ L. Imidazole was also mixed with ruthenium compounds at Ru:imidazole ratios of 0.2:1 and 0.4:1 at a final imidazole concentration of 300 μ M in a volume of 300 μ L. Trp-cage mutants were not used in this study because the stability of these peptides at 37 °C for long periods of time was questionable. Samples were incubated in plastic tubes in a heat block at 37 °C for at least 34 days while monitoring UV-Vis spectra (250 – 750 nm). After determining that the reaction was nearing

completion by monitoring A₄₅₀₋₄₆₀ values, samples were analyzed by SDS-PAGE, HPLC, and CD without further purification. Unsuccessful attempts were made to further characterize these reaction products by MALDI-TOF-MS and ESI-MS. Sample heterogeneity due to protein oligomerization prevented analysis of species by mass spectrometry techniques.

5.2.2.2 Method 2, Reactions under N_{2(g)} at 75 °C

Reaction conditions were modified from Ghadiri, *et.al.*¹¹ Briefly, RuB and RuC powders were solubolized in 20 mM MOPS, 100 mM NaCl, pH 7 buffer. 3HIU and 3HPU (in the same buffer) were mixed with the ruthenium compounds in a Ru:protein ratio of approximately 1:1 at a final concentration of 300 μ M in a volume of 800 μ L. (Since the development of a suitable procedure for substitution of the chloride ions is still under investigation, the identity of RuC used here was not known. However, the elemental analysis indicated the formation of a material whose molecular mass is similar to that of the targeted formulation of [Ru^{II}(tacn)(dmso)₃][OTf]₂.) N_{2(g)} was gently bubbled through the samples for 30 seconds before incubating them in 1 mL quartz cuvettes at 75 °C for 6 hours while monitoring UV-Vis spectra (250 – 750 nm). Trp-cage mutants were not used in this study because the reaction temperature was much greater than the melting point of the parent Trp-cage design.

After terminating the reaction, the Ru-protein adduct was purified using methods in Ghadiri *et.al.* and Muheim, *et.al*^{11,12}. After cooling to room temperature, the reaction was stopped by desalting on a 5 mL (height ~ 5 cm) sephadex G-25 (Sigma-Aldrich, Co.) column equilibrated with 5 mM MOPS, 25 mM NaCl, pH 7 buffer. Protein was eluted with the same buffer, and the 1 mL fractions were analyzed for Ru-protein content by UV-Vis. Samples were frozen (-20 $^{\circ}$ C) until further purification could be performed.

Thawed Ru-protein fractions were then purified by cation-exchange using a 2 mL (height ~ 1.5 cm) CM-52 (Whatman, Plc.) column equilibrated with 20 mM ammonium

acetate, 25 mM NaCl, pH 4.5 buffer. Before pooled fractions could be applied to the column, their pH was adjusted to 4.5 by addition of dilute acetic acid. The following gradient (exploiting changes in pH and then ionic strength) was used to elute proteins from the column: 4 CV of 20 mM ammonium acetate, 25 mM NaCl, pH 4.5; 2.5 CV of 20 mM ammonium acetate, 25 mM NaCl, pH 4.5; 2.5 CV of 20 mM ammonium acetate, 25 mM NaCl, pH 5.0; 2.5 CV of 20 mM ammonium acetate, 25 mM NaCl, pH 5.5; 2.5 CV of 20 mM MES, 25 mM NaCl, pH 6.0; 2.5 CV of 20 mM MES, 25 mM NaCl, pH 6.5; 2.5 CV of 20 mM MOPS, 25 mM NaCl, pH 7.0; 2.5 CV of 20 mM MOPS, 25 mM NaCl, pH 7.5; 2.5 CV of 20 mM MOPS, 25 mM NaCl, pH 7.5; 2.5 CV of 20 mM MOPS, 253 mM NaCl, pH 7.5; 2.5 CV of 20 mM MOPS, 1.013 M NaCl, pH 7.5; and 4 CV of 20 mM MOPS, 2.0 M NaCl, pH 7.5. The 1 mL fractions were analyzed for Ru-protein content by UV-Vis. Ru-protein eluted after application of 20 mM MOPS, 253 mM NaCl, pH 7.5 buffer and was used in further purification.

A suspension of Co-NTA (133 μ L at 75% v/v in water) immobilized on agarose beads (modified Ni-NTA Superflow resin from Qiagen, Inc.) was added to Ru-proteincontaining CM-52 fraction(s). The sample was allowed to incubate with gentle rotation for 10 minutes at room temperature before centrifuging at 16 000 x g for 10 minutes. The supernatant was collected, and Co-NTA was added again as described above. This supernatant was used in the next step.

In order to exchange the buffer for one suitable for characterization techniques, samples were centrifuged in Microcon YM-3 filter units (3 kDa NMWL) according to the company's protocol (Millipore, Corp.) while diluting with OptimaTM-grade water (Thermo Fisher Scientific, Inc.). The buffer was diluted by a factor of ~19 by the addition of 1.16 mL of water to 500 μ L of sample over the course of 5 rounds of buffer dilution. Presence of Ru-protein was verified by UV-Vis after purification. Samples were then analyzed by SDS-PAGE, ESI-MS, and CD.

5.2.3 SDS-PAGE Analysis

Laemmli buffer systems¹³ were used in the preparation and running of SDS-PAGE experiments. Reductant-free loading buffer (250 mM Tris, 40% (v/v) glycerol, 8% (w/v) SDS, 0.4% (w/v) bromophenol blue, pH ~6.8) was diluted 1:4 in Ru-protein samples. If samples were further thermally denatured, they were heated at 95 °C for at least 10 minutes. Samples were loaded onto 4%/25% (discontinuous stacking/separating) gels of a thickness of 0.75 mm. The gels were run at 300 V for the first 5 minutes and at 200 V until the dye-front left the gel. Gels were stained with either Coomassie R-250 using standard techniques¹⁴ or with the Silver Stain Plus kit (Bio-Rad Laboratories, Inc.) using the manufacturer's instructions.

5.2.4 HPLC Analysis

Reversed-phase HPLC analyses were performed on an Agilent 1100 series highperformance liquid chromatograph using a PRP-3 analytical column (Hamilton, Co.) and the following gradient (where solvent A was 0.1% TFA in ddi water and solvent B was 0.1% TFA in acetonitrile): 0 - 2 minutes = isocratic at 25% solvent B; 2 - 25 minutes = linear gradient, 25 - 65% solvent B; 25 - 27 minutes = linear gradient, 65 - 95% solvent B; 27 - 45 minutes = isocratic at 95% solvent B. 50 µL samples were injected and their chromatographs produced from UV-Vis analyses at 214, 276, and 457 nm.

5.2.5 MALDI-TOF-MS Analysis

MALDI-TOF-MS data were collected on a Bioflex III, Bruker Daltonics spectrometer (Molecular Analysis Facility, The University of Iowa). Samples were desalted by dilution or by using zip-tips following the manufacturer's instructions (Millipore, Corp.) prior to analysis. The instrument was calibrated at two points every 4 – 8 samples with external standards. CHCA was used as the matrix. Spectra were collected from 3000 – 21 000 m/z in linear mode.

5.2.6 ESI-MS Analysis

ESI-MS data were collected on a Thermo Finnigan LCQ Deca Spectrometer (Mass Spectrometry, The University of Iowa) in positive ion mode. Formic acid was added to samples to a final concentration of 0.1% before injection into the instrument and infusion for ~ 10 minutes at 10 μ L/minute. Data were collected while the sample flowed at a rate of 3 or 8 μ L/minute. The following conditions were optimized: spray voltage, 4.52 kV; sheath gas, 33.93 units; capillary temperature, 250 °C; capillary voltage, 26.64 V. Deconvolution was performed using Bioworks 3.2 software (Thermo, Fisher Scientific, Inc.)

5.2.7 CD Analysis

CD spectra were collected using an Olis Cary-17 DS Conversion Spectrophotometer. Spectra were collected (of samples in 1 mm path-length quartz cuvettes at 25 $^{\circ}$ C) from 200 – 290 nm at 1 nm *per* point and a bandwidth of 3 nm with a variable integration time determined by the PMT voltage. Spectra were baselinecorrected using water as the reference. Three spectra were averaged and then smoothed (11-point, moving, weighted-average using least squares calculations) using the software provided by Olis. Spectra are reported in millidegrees because they could not be corrected for concentration since it was unknown.

5.3 Results and Discussion

5.3.1 Reactions of Ubq Mutants with Ru-Compounds, Method 1, Reactions in air at 37 °C

Panel C in **Figure 5.1** shows representative spectra of reactions of 3HIU with ruthenium compounds over a 34-day time period. Reactions with 3HPU protein produce similar spectra. RuA-protein reactions showed no distinct peaks in the visible range of the spectra, but all reactions of mutants with RuB or RuC showed peaks in this range at ~460 nm. In contrast, apo-proteins and Ru-complexes alone did not show peaks in this range. The change in absorbance at 460 nm over time implied that reactions with RuA were slowing as the reaction time approached 34 days and that reactions with RuB and RuC were nearing completion (Panel A, **Figure 5.1**). Final spectra of RuB and RuC with 3HIU and 3HPU are shown in panel B of the same figure. Ru-3HIU reactions have an absorbance at 460 nm while Ru-3HPU reactions have an absorbance at 450 nm.

For comparison, reactions of Ru-complexes with imidazole were performed. They show an absorbance similar to that of Ru-3HIU with a λ_{max} at ~457 nm (Panel B, **Figure 5.1**). An additional feature at 320 nm in the imidazole reactions is not easily seen in the protein reactions. However, it cannot be said that this peak does not exist. It could be buried under peaks of other absorbance bands.

A dependence of absorbance intensity upon the amount of ruthenium complex added was observed (data not shown). When twice as much ruthenium complex was added, the absorbance was twice as intense. For example, the absorbance of 3HIU + 1.2equivalents of RuB after 34 days was 0.20 which was twice as great as the absorbance (0.10) when only 0.6 equivalents of the same ruthenium complex was added to 3HIU.

At first, it was thought that this result implied 1:1 binding since absorbance was proportional to the amount of ruthenium compound added. However, SDS-PAGE analyses showed that there were many protein products made when allowing Ubq mutants to react with RuA, RuB, and RuC (**Figure 5.2**). Induction of protein oligomerization upon Ru-binding can be seen by comparing lanes containing apo-protein, where only one band is observed, to other lanes, where at least four distinguishable bands are present. It was thought that making ruthenium complexes with Ru(tacn) cores would prevent oligomerization by steric hindrance, but this organic ligand was not successful at stopping oligomer formation. While addition of ruthenium causes oligomerization of protein (probably through oxo-bridged species), the size of these oligomers is finite. In other words, random aggregation does not seem to occur. Instead, distinct 2, 3, 4, and 5mers are made (**Figure 5.3**). For this analysis, the dependence of gel mobility upon masses of protein standards (assumed to be proportional to size) was determined and plotted. Then, the mobilities of the oligomers formed in reactions and their expected masses were plotted on the same graph. Oligomer sizes/mobilities were in good agreement with those determined for the standards. Furthermore, when predicting the mobility of 6, 7, and 8-mers from the best fit line, corresponding bands were not observed in gels. Therefore, the mobilities of protein species seen in the gels indicate the formation of oligomers with up to five units.

As expected, these Ru-bonded oligomers are extremely stable. SDS-PAGEobserved oligomer species were identical whether samples were completely denatured by extensive heating or not. Oligomerization is similar regardless of the type or amount of ruthenium compound added. This result implies that stoichiometry does not greatly impact these reactions if they are allowed to continue for long periods of time.

HPLC analyses reaffirmed SDS-PAGE findings. Apo-protein eluted as a distinct peak. Upon addition of Ru-complexes, another broad, multi-shouldered peak appeared (**Figure 5.4**). This new peak is probably due to elution of oligomers as observed by SDS-PAGE. While SDS-PAGE separates proteins by size, RP-HPLC separates them based on polarity. If ruthenium complexes are bridging proteins in oligomers, these species would all share similar charge-to-size ratios and elute together as seen in the broad HPLC peaks.

Another method of analysis that proved inconclusive was mass spectrometry, whether by MALDI or ESI-MS (data not shown). MALDI was initially used to examine RuA-protein reactions, and species could be detected but their signals were very low, of poor resolution, and not reproducible. With ESI, the presence of oligomers made deconvolution and assignment in spectra impossible. Both data sources support the hypothesis that ruthenium compounds reacted with ubiquitin mutants, but describing the products was not feasible. Purified Ru-protein samples are necessary for further characterization since there are so many oligomerization products upon addition of ruthenium compounds. (See below for similar reactions performed at 75 °C where monomeric Ru-proteins were successfully purified and characterized.)

The last characterization method employed to investigate the effect of Ru-binding on these proteins was CD. This technique was used to determine if complexation of RuB and RuC (and consequent oligomerization) caused protein unfolding. Signals were weak, but the characteristic bands at 207 and 222 nm were still present (**Figure 5.5**). Therefore, it can be said with confidence that these mutants retain structures similar to that of the apo-protein despite binding to ruthenium and undergoing oligomerization.

5.3.2 Reactions of Ubq Mutants with Ru-Compounds, Method 2, Reactions under $N_{2(g)}$ at 75 °C

Reactions were monitored by UV-Vis spectroscopy (**Figure 5.6**). As seen in Panels B and C, a peak at 320 nm was observed in addition to the one at ~460 nm observed for reactions performed at 37 °C in Method 1. It is an interesting observation that depending upon reaction conditions, peaks at 460 nm or 320 nm or both wavelengths can be observed. From these three cases, it is known that further elucidation of species giving rise to the different peaks (A_{320} and A_{460}) requires performance of a more systematic investigation of reaction conditions and stringent control of temperature and oxygen levels

Though weaker in intensity, the peak at 460 nm grows at a rate similar to that of the peak at 320 nm. From the slopes of these peaks, (Panel A, **Figure 5.6**) it was determined that the reaction was still progressing at 6 hours. However, the reaction was stopped to prevent formation of possible oligomers or decomposition products.

5.3.2.1 Ru-protein Purification

As noted above, multiple Ru-protein species can be formed during reactions, but these heterogeneous samples are difficult to study. Therefore, Ru-protein monomers were isolated in order to study these reactions more thoroughly. Protein purification was followed by SDS-PAGE (**Figure 5.7**). Reactions at high temperatures in the presence of limited amounts of oxygen reduce the amount of oligomers made (as seen in the sephadex fractions). It also seems that addition of Ru-complexes may have caused some protein degredation as seen in the 4th lane. The appearance of smaller-mass, broad bands was observed in SDS-PAGE gels used to analyze purification of RuC-3HIU, too. Ru^{II} starting materials could cause oxidative cleavage of proteins. Coincidently, the SDS-PAGE analysis of the RuC-3HPU reaction (that only showed an absorbance at 320 nm as described above) did not seem to have these low-mass bands (data not shown).

Ru-protein monomers were selectively purified using the three chromatography methods described above. Gel filtration removed unreacted Ru-compounds. Cation exchange separated species of different charges. Co-NTA removed any un-reacted protein by exploiting its metal-binding site. Although samples were diluted during purification, Coomassie staining suggests that the samples were of \geq 90% purity. Additionally, characterization by ESI-MS showed the presence of only one species *per* sample.

5.3.2.2 Ru-Protein Characterization

Purified samples were first analyzed by ESI-MS (**Figure 5.8**). Peaks could be deconvoluted and products of these reactions appear at the following masses: 3HIU + RuB = 9374.8, 3HIU + RuC = 9375.0, 3HPU + RuC = 8890.0. When compared to the mass of the parent apo-proteins as determined by ESI-MS, the Δ mass values are 228.5, 228.7, and 228.0, respectively. The isotopic mass of the proposed Ru(tacn) fragment is 231.0 Therefore, the apparent masses of these products are in good agreement for metal binding reactions in which the protein loses three protons upon binding of the Ru(tacn) fragment. As with ESI-MS data collected for the apo-proteins, peaks appear as envelopes in which artifacts can be seen in the spectra (sodium adducts). Therefore, the prominent

peak of lowest mass was assigned as the parent peak. Even in the presence of these broad envelopes, no apo-protein was detected.

From this data, it can be shown with confidence that monomeric Ru-protein was successfully purified. Furthermore, all Ru-products show binding of a species with a mass of ~ 228.5 whether the reactant was RuB or RuC. These data support the hypothesis that RuC is a ruthenium complex with a Ru(tacn) core even though the other ligands and counter-ions have not been identified. Such a core complex can selectively bind to the 3-His sites built into Ubq.

As before with unpurified Ru-protein products, CD analysis was used to assess the protein structure upon Ru-binding. Spectra of purified Ru-protein products are shown in **Figure 5.9**. The characteristic bands at 207 and 222 nm were clearly present. As expected from previous data, the Ubq mutants remain in a native structure when bound to Ru(tacn) fragments.

5.4 Conclusions

These experiments demonstrate the efficacy of using ubiquitin as a *de novo* design scaffold for building Ru-binding sites. Binding of a kinetically-inert metal ion such as Ru^{II} is difficult in protein systems. First, binding is slow, and second, binding can form kinetically-stable, yet misfolded products. Therefore, 3HIU and 3HPU were designed with surface-accessible, pre-organized 3-His binding sites purposely constructed to complement a Ru-complex with exchangeable *fac*-ligands. Reactions of 3HPU and 3HIU with Ru^{II}(tacn) complexes yield different products (monomers *vs*. oligomers) depending upon reaction conditions, and their UV-Vis spectra can not be fully described at this time. However, monomeric Ru-proteins have been isolated and characterized by ESI-MS, and this technique shows evidence of binding of a Ru(tacn) fragment to 3HIU and 3HPU monomers. Furthermore, the Ubq mutants are able to maintain their structure upon Ru-binding as seen in their CD spectra.

Binding of Ru^{II} or Ru^{III} by these proteins incorporates new spectroscopicallyactive elements into their structures. UV-Vis bands for ruthenium complexes are well described in the literature,¹⁵⁻²⁰ and there are two NMR-active nuclei (⁹⁹Ru and ¹⁰¹Ru).²¹ Ru^{II} is diamagnetic, and its protein binding environment could be investigated directly by NMR. Ru^{III} is paramagnetic, and shifts induced in ¹H-NMR spectra could be used to examine affected protein residues when binding ruthenium complexes.²²

Complexes of Ru^{II} and Ru^{III} ions are kinetically inert and typically adopt octahedral ligand geometries. By making pre-organized metal-binding sites for these ions, the balance in energy demands of the protein *versus* the ion can be investigated. In other words, the contribution to metal-binding by site pre-organization can be investigated. Such studies can be useful when designing binding sites for other ions with specific ligand-geometry preferences.

By introducing metal ions such as Ru^{II} or Ru^{III} into proteins, they can perform new functions. Unlike Zn²⁺, these ions can perform redox chemistry and have been used in many catalysts such as those developed by Noyori for asymmetric transfer hydrogentation reactions.⁵ Ru-protein complexes could be the next generation of chemoand enantioselective catalysts since protein specificity is unmatched by small organic compounds and ligands.

Ubiquitin plays many roles *in vivo* as previously described. In fact, new roles and signaling mechanisms for ubiquitin are still being discovered. While the processes have not been fully elucidated, it is known that ubiquitin is used in oncogenic signaling networks. Ruthenium compounds such as NAMI-A and KP1019 have been shown to be effective in cancer treatment. By combining these two ideas, a chemotherapeutic metal ion could be delivered to these signaling pathways. Cross-linking of Ubq with other proteins *via* ruthenium ions would irreversibly affect signaling. Even if not useful as a therapy, Ru-Ubq species could be used to more thoroughly investigate this protein's

interactions with other biomolecules by examining ruthenium cross-linking or delivery to other proteins.

This work demonstrates that Ru-bound mutants of ubiquitin can be easily made and that their isolation is required for complete characterization. 3HIU and 3HPU are the first proteins designed to bind to Ru^{II} *via* three, natural amino acids (histidine residues) in a pre-organized site. While these experiments investigated metal-binding and *de novo* design principles, applications of this research could reach beyond design of Ru-binding proteins and into the fields of medicine and catalysis.



Figure 5.1 UV-Vis Spectra of Ru-Ubq mutant reactions. Panel A shows ΔA at 450 or 460 nm during reaction progression, and Panel B shows the final spectra. Red curves are reactions with 3HIU and blue curves are reactions with 3HPU. Circles are reactions with RuB and squares are reactions with RuC. Panel C shows representative spectra for reactions of 3HIU with RuA, RuB, and RuC from left to right. Reactions with 3HPU produce similar spectra.



Figure 5.2 SDS-PAGE Analysis of Ru-Ubq mutant reactions. Samples in the gel on the left were not completely denatured while samples in the gel on the right were heated extensively. Differences in gel color are due to different staining techniques used: silver on the left and Coomassie on the right. The arrow indicates where Ubq mutant monomers appear in gel lanes.



Figure 5.3 Mobility analysis of species present after reactions of 3HIU and 3HPU with $Ru^{II}(tacn)$ complexes *vs.* protein standards



Figure 5.4 HPLC A_{214} chromatograms for reactions of RuB and RuC with Ubq mutants. Reactions with 3HIU are shown at top while those with 3HPU are shown at the bottom. Apo-protein in shown in black while reactions with RuB are shown in red, and reactions with RuC are shown in blue.



Figure 5.5 CD Spectra of unpurified Ru + Ubq mutant reactions performed in air at 37 °C for 34 days.



Figure 5.6 UV-Vis Spectra of Ru-Ubq mutant reactions. Panel A shows ΔA during reaction progression with ΔA_{320} on the left and ΔA_{460} on the right. Panel B shows the final spectra of all reactions. In the upper two panels, 3HIU + RuB is shown in red, 3HIU + RuC in gold, and 3HPU + RuC in blue. Panel C shows raw data for reactions of Ru compounds with Ubq mutants.



Figure 5.7 Representative SDS-PAGE analysis of RuB-3HIU purification. The arrow indicates where 3HIU monomers appear in gel lanes. RuC-3HIU and RuC-3HPU purified similarly.



Figure 5.8 ESI-MS analysis of purified Ru-protein products. 3HIU + RuB is shown at top, 3HIU + RuC in the middle and 3HPU + RuC at the bottom. Arrows indicate positions of Ru(tacn)-Ubq mutant species and their apparent masses.



Figure 5.9 CD Spectra of purified Ru + Ubq mutant reactions performed at 75 °C for 6 hours in the presence of limited oxygen.

Notes

- (1) Guo, Z.; Sadler, P. J. Angew. Chem. Int. Ed. **1999**, 38, 1512-1531.
- (2) Kostova, I. Curr. Med. Chem. 2006, 13, 1085-1107.
- (3) Hartinger, C. G.; Zorbas-Seifried, S.; Jakupec, M. A.; Kynast, B.; Zorbas, H.; Keppler, B. J. In. Biochem. 2006, 100, 981-904.
- (4) Ang, W. H.; Dyson, P. J. Eur. J. Inorg. Chem. 2006, 4003-4018.
- (5) Noyori, R.; Hashiguchi, S. Acc. Chem. Res. **1997**, *30*, 97-102.
- (6) Lu, Y. Angew. Chem. Int. Ed. 2006, 45, 5588-5601.
- (7) Lu, Y. Inorg. Chem. 2006, 45, 9930-9940.
- (8) Evans, P.; Spencer, A.; Wilkinson, G. J. Chem. Soc. Dalton Trans. 1973, 204-209.
- (9) Geilenkirchen, A.; Neubold, P.; Schneider, R.; Wieghardt, K.; Florke, U.; Haupt, H. J.; Nuber, B. J. Chem. Soc. Dalton Trans. **1994**, 457-464.
- (10) Arnold, J.; Hoffman, C. G.; Dawson, D. Y.; Hollander, F. J. *Organometallics* **1993**, *12*, 3645-3654.
- (11) Ghadiri, M. R.; Soares, C.; Choi, C. J. Am. Chem. Soc. 1992, 114, 4000-4002.
- (12) Muheim, A.; Todd, R. J.; Casimiro, D. R.; Gray, H. B.; Arnold, F. H. J. Am. *Chem. Soc.* **1993**, *115*, 5312-5313.
- (13) Laemmli, U. K. *Nature* **1970**, *227*, 680-685.
- (14) Sambrook, J.; Russell, D. W. In *Molecular Cloning, A Laboratory Manual*; Third ed.; Argentine, J., Ed.; Cold Spring Harbor Laboratory Press: New York, 2001; Vol. 3, p A8.46-A8.47.
- (15) Henn, M.; Alessio, E.; Mestroni, G.; Calligaris, M.; Attia, W. M. *Inorg. Chim. Acta* **1991**, *187*, 39-50.
- (16) Serrano, S.; Sala, X.; Plantalech, E.; Rodriguez, M.; Romero, I.; Jansat, S.; Gomez, M.; Parella, T.; Stoeckli-Evans, H.; Solans, X.; Font-Bardia, M.; Vidjayacoumar, B.; Llobet, A. *Inorg. Chem.* **2007**, *46*, 5381-5389.
- (17) Yamaguchi, R.; Harada, K.; Sunatusuki, Y.; Kojima, M.; Nakajima, K.; Matsumoto, N. *Eur. J. Chem.* **2006**, 3326-3243.
- (18) Batista, A. A.; Polato, E. A.; Queiroz, S. L.; Nascimento, O. R.; James, B. R.; Rettig, S. J. *Inorg. Chim. Acta* **1995**, *230*, 111-114.
- (19) Baird, I. R.; Rettig, S. J.; James, B. R.; Skov, K. A. *Can. J. Chem.* **1998**, *76*, 1379-1388.

- (20) Sundberg, R. J.; Bryan, R. F.; Taylor, I. F.; Taube, H. J. Am. Chem. Soc. **1974**, 96, 381-392.
- (21) Ebsworth, E. A. V.; Rankin, D. W. H.; Cradock, S. In *Structural Methods in Inorganic Chemistry* Second ed.; Blackwell Scientific Publications: London, 1991, p 33.
- (22) Bertini, I.; Luchinat, C.; Parigi, G. *Solution NMR of Paramagnetic Molecules*; Elsevier Science: Amsterdam, 2001.

CHAPTER 6: REACTIONS OF TECHNETIUM-99M COMPLEXES WITH UBIQUITIN MUTANTS

6.1 Introduction

^{99m}Tc is a high-intensity γ -emitter (140 keV) with a 6 hour half-life used in radioimaging.¹ The energy emitted by these gamma rays is high enough that low doses of Tccomplexes can be administered to patients and still be observable.² The radionuclide is eluted from ⁹⁵Mo generators as TcO₄⁻. This metal can access eight different oxidation states and bind with a variety of ligands ranging from pure electron donors (oxo species) to back-donating ligands (CO).² One such complex, Tc¹(CO)₃(H₂O)₃ is an organometallic, kinetically inert complex with three exchangeable water ligands in a facial arrangement. This Tc¹(CO)₃ core has a high affinity for aromatic amine ligands.²

^{99m}Tc imaging agents have been designed in three generations.³ As early as 1961, TcO₄⁻ was used to image the thyroid because it mimicked iodide.² Subsequent firstgeneration molecules were simple chelates that were distributed in tissues based on their charge density. In the next generation, receptor specific agents were designed. These agents mimicked steroids or dopamine inhibitors in order to localize in tissues. The most recent generation aims at higher target specificity by using biomolecule-chelate conjugates. This strategy is employed to meet the challenge of preparing biomolecules with sufficient binding affinity for Tc. If the ion is not tightly bound, then the detected radioactivity signal could be from other species since the ion can be transferred to other naturally occurring biomolecules (*i.e.* transferrin or albumin).

While ^{99m}Tc is a useful nucleus for medical imaging, its cognate ¹⁸⁶Re is also useful in medicine. This nucleus is a β -emitter with an E_{max} of 1.1 MeV and a half life of 89.2 hours.² Due to its similar coordination chemistry preferences to Tc, it can be placed in complexes structurally-analogous to designed Tc-complexes and deliver therapeutic doses of radiation to sites previously imaged with the Tc-complex.

6.2 Materials and Methods

6.2.1 Synthesis of $Tc^{I}(CO)_{3}(H_{2}O)_{3}$

 $Tc^{1}(CO)_{3}(H_{2}O)_{3}$ was synthesized from TcO_{4}^{-} generator eluate (The University of Iowa, Dept. of Nuclear Medicine) using a previously reported procedure with some modifications.⁴ Briefly, 6 mg of NaBH₄, 4 mg of Na₂CO₃ and 15 mg of NaK-tartrate were combined in a 3-necked vial. The vial was flushed with $CO_{(g)}$ for ~1 minute before adding 2.5 mL of 3.3 mCi/mL TcO_{4}^{-} generator eluate. Bubbling of $CO_{(g)}$ through the reaction solution was continued during heating to 75-80 °C for at least 1 hour. The solution was cooled on ice for ~15 minutes before addition of 0.3 mL of 1 M phosphate buffer, pH 7.4. The reaction was allowed to neutralize for at least 15 minutes before use in any other reaction/application. To test for ligand exchangeability, L-histidine was added to the unpurified product to a final concentration of 100-200 µM and allowed to incubate at room temperature.

6.2.2 Reactions of $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ with Proteins

Unpurified, neutralized $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ was added directly to ~ 50 µM protein diluted in isotonic saline. As a control, $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ was also added to untagged, wt ubiquitin in the same way. While the concentration of protein was in the micromolar range, the final concentration of the $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ in solution was in the pico- and femtomolar range. Therefore, proteins and other ligands were added to Tc^{+} in large stoichiometric excess. To investigate the stability of the Tc-protein complexes, free Lhistidine was added to the reaction mixtures. All reactions were allowed to incubate at room temperature unless otherwise noted.

6.2.3 HPLC Methods

Reversed-phase HPLC analyses were performed on a Waters high-performance liquid chromatograph using an Altima C18 analytical column (Alltech Associates, Inc.).

Chromatographs were produced from both A_{280} and radioactivity data after injection of 20 µL samples. Two previously reported solvent systems were used in this project.⁵ In Method I, solvent A was 0.1% TFA in ddi H₂O, and solvent B was 0.1% TFA in acetonitrile. The following gradient was used: 0 - 3 minutes = isocratic at 0% solvent B; 3 - 13 minutes = linear gradient, 0 - 50% solvent B; 13 - 17 minutes = isocratic at 50% solvent B; 17 - 20 minutes = linear gradient, 50 - 100% solvent B. In Method II, solvent A was 50 mM TEAP buffer, pH 2.25, and solvent B was methanol. The following gradient was used: 0 - 3 minutes = isocratic at 100% solvent B. HPLC Method I was used for analyzing Tc^I(CO)₃(H₂O)₃ production and reactions with proteins. Method II was only used for analyzing Tc^I(CO)₃(H₂O)₃ production. This solvent system was not effective for eluting proteins species.

6.3 Results and Discussion

6.3.1 Synthesis of $Tc^{1}(CO)_{3}(H_{2}O)_{3}$

 $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ starting material was successfully prepared (**Figure 6.1**). This chromatograph shows elution profiles for starting materials when HPLC Method II (where starting materials produce sharper peaks) is used. All subsequent studies discussed herein use Method I. Upon changing solvent systems, the elution peak for $Tc^{I}(CO)_{3}$ broadens as observed in the remaining chromatograms shown in this work. Furthermore, the order of elution for the TcO_{4}^{-} , $Tc^{I}(CO)_{3}$, and $Tc^{I}(CO)_{3}(L-His)_{3}$ species changes with the new solvent system. In Method I, the order of elution is $Tc^{I}(CO)_{3}$, TcO_{4}^{-} , and then $Tc^{I}(CO)_{3}(L-His)_{3}$. Use of acetonitrile when performing HPLC studies of $Tc^{I}(CO)_{3}$ complexes causes sample heterogeneity because the solvent can act as a ligand for the metal complex by exchanging with the bound water molecules.⁵ Further broadening can also be caused by presence of various charge states due to possible deprotonation of bound water molecules. The pK_a of these bound water molecules is not known. TFA was added to the solvent system, but addition of acid does not guarantee that all the ligands are protonated homogeneously if the pH is close to the pKa of one or more ligands.

6.3.2 Reactions of $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ with Proteins

Reactions of designed proteins with Tc-99m can be seen in **Figure 6.2**. $TcO_4^$ elutes at 2.5 minutes while $Tc^{I}(CO)_3$ elutes between 5 and 8.5 minutes The $Tc^{I}(CO)_3(L-His)_3^+$ control elutes at 11.5 minutes. Upon addition of designed proteins to $Tc^{I}(CO)_3(H_2O)_3$, elution shifts to ~ 12.6 minutes. Upon addition of Ubq to $Tc^{I}(CO)_3$, elution shifts to ~13 minutes, most probably due to binding through the one naturally-occurring histidine residue near its N-terminus.

Neither stoichiometry nor affinity of metal binding could be determined due to the fact that protein was added at concentrations at least 1×10^{10} that of the Tc-complex. Adding protein at such excess concentrations would push the equilibrium to Tc-protein complexes instead of unbound Tc^I(CO)₃ (which at least gives a lower limit to the affinity). It is assumed that the Tc-protein complex is kinetically stable due to the nature of the metal center, though it is clear the ligand exchange can occur on the minute time-scale as seen with L-His competition experiments (as discussed below).

Addition of proteins and L-His to TcO_4^- eluate did not cause any shifting of radioactive peaks (data not shown). Therefore, the proteins discussed here (3HIU, 3HPU, 3HT-Cb, and wt Ubq) and free L-His do not bind to TcO_4^- under used conditions.

6.3.2.1 Competition Assays

By adding free L-histidine to Tc-protein products, one can investigate the inertness toward ligand exchange, presumably correlated to thermodynamic stability (**Figure 6.3**). However, it should be noted that these experiments do not differentiate thermodynamic stability from changes in kinetic inertness due to the chelate effect. Upon addition of 1 equivalent of L-His, 3HIU retains its bound Tc ion. The stability of the Tc-

3HPU complex is much less than that of the complex with 3HIU since some $Tc^{I}(CO)_{3}(L-His)_{3}$ is formed under similar conditions. In previous studies, addition of $Tc^{I}(CO)_{3}$ to 3HT-Cb caused protein aggregation. Therefore, initial competition studies were performed at 0 °C. At this temperature, the Tc-3HT-Cb complex is stable. However, lability of the ion at room temperature could not be studied and compared with room temperature stabilities of the other Tc-Ubq mutant complexes. Therefore, binding by 3HT-Cb will not be discussed further here. Note, all designs contain a 3-His binding site. Therefore, addition of 1 equivalent of free L-His is actually the addition of 1 L-His *per* 3 protein ligands.

As shown above Tc-3HIU complexes are more stable than complexes with 3HPU. Both proteins have 3-His binding sites on the same protein scaffold, but yet their binding capabilities are very different in these studies. 3HIU was designed to have a highly flexible 3-His *fac*-binding site in a protein location that could easily accommodate loop mutations. On the other hand, 3HPU was designed to have a more rigid *fac*-binding site by introduction of point mutations. The effect of point mutations at this site upon protein stability has not been quantitatively determined, but thermal denaturation studies with this mutant indicated that 3HPU was slightly less stable than 3HIU. Tc-complexes have geometrical requirements for ligand binding. 3HIU is probably better at accommodating this ion while retaining its overall structural stability. While the 3-His site in 3HPU was designed to bind to metal ions in a facial orientation, these bond angles may not be appropriate for Tc^I, causing the observed decrease in binding stability.

Since Tc-3HIU complexes were most stable, they were used in further competition studies. For comparison, stability of Tc-binding by wt Ubq was also studied. In these experiments, 4 equivalents of free-L-histidine were added to protein reactions (**Figure 6.4**). It was shown that $Tc^{I}(CO)_{3}$ -binding by 3HIU was stable in the presence of L-His for almost 5 hours. In other experiments where competition was investigated for even longer time periods, binding was stable even at 12 hours. (Subtle shifts in retention time will be discussed below). In comparison, wt Ubq begins losing its Tc ion to a 3-L-His complex within the first 10 minutes of reaction. By 30 minutes later, wt Ubq had lost almost all of its ^{99m}Tc to free L-His complex formation. This lack of binding stability supports the hypothesis that wt Ubq cannot affectively bind Tc ions through its single natural His residue.

6.3.2.2 Kinetics Assays

While not defined quantitatively, qualitative comparisons were made amongst protein reactions to see which species reacted most quickly with $Tc^{I}(CO)_{3}$. 3HIU binding compared to that by wt Ubq and free L-His is shown in **Figure 6.5**. Binding of $Tc^{I}(CO)_{3}$ to 3HIU is observed within the first 5 minutes of reaction. Binding is complete within half an hour (measured by disappearance of the $Tc^{I}(CO)_{3}$ peak at 6.5 minutes). Even in 12.3 hours, binding of $Tc^{I}(CO)_{3}$ to wt Ubq is not complete. While not shown here, binding of Tc ions by L-His does not occur within 5 minutes like with 3HIU, but it can be observed at ~15 minutes. Some free L-His-bound Tc can still be observed after 2 hours, but the reaction is complete by 6.5 hours.

Therefore the order for binding kinetics is 3HIU > L-His > wt Ubq. This order of reaction kinetics shows that the presence of more histidine in solution (3 times as many histidine residues are present in the 3HIU and free His samples than in wt Ubq) makes the reaction proceed faster. Further, having those 3 histidines in a pre-formed binding site rather than free in solution speeds up the reaction as well. These experiments serve as more evidence that a protein chelate has been produced.

On a side note, it is observed that the Tc-3HIU complex subtly changes over time. Within the 8 hours of data shown, peak maxima shift between 12.5 and 12.9 minutes (by ~ 3%). At some times, even two species can be resolved. Furthermore, peaks do not shift systematically. Obviously, the thermodynamics of protein binding are not as simple as initially envisioned. Upon mixing 3HIU with this kinetically inert metal ion, a variety of associations can be made: with negatively-charged residues on the protein surface, with the natural histidine near the N-terminus, or at the designed binding site. Therefore, initial binding would be heterogeneous. Soon though, equilibrium would be pushed to the favored 3-His binding site. Even after binding at this single site, the protein structure probably changes to accommodate the ions' geometrical preferences. It is also not known how Tc-binding could change the charge state at such binding sites. While the first histidine proton would be lost upon binding, it not known if the pKa of the second proton is significantly lowered. Such sources of heterogeneity could account for the slight shifts in elution profiles.

6.4 Conclusions

Like in the Ru-binding studies, these experiments show that ubiquitin is a very useful protein scaffold for protein re-design. Ru-binding studies further suggest that binding of this scaffold to a different inert metal such as Tc^{I} is probably similar in nature. Full characterization of the ^{99m}Tc complex is not possible, but effects of binding upon structure are likely very limited just like in Ru-binding. It has been shown that 3HIU effectively binds $Tc^{I}(CO)_{3}$, and that the complex is stable to L-His substitution. From comparison with wt Ubq, it can be assumed that binding occurs at the designed binding site and not at any intrinsic features of the protein.

As described above, ^{99m}Tc is a much-used nucleus in medical imaging. Design of ^{99m}Tc-complexes has evolved in order to target specific tissues in a sophisticated manner. Currently, research has attempted to utilize Tc-labeled biomolecules with varying degrees of success. In recent work, non-natural chelating moieties are coupled to these biomolecules *via* molecular spacers. The approach described in this thesis is proof in principle that non-natural chelators and spacers are not necessary to design a Tc-binding biomolecule. A protein chelate can be produced from bacteria inexpensively and used without further modification. Imaging of ubiquitin regulation in tissues could greatly benefit research elucidating this molecule's various signaling mechanisms. It is already known that ubiquitin plays roles in cancer networks and discovering regulation by ubiquitin in cancers could lead to new drug targets. If a molecule like 3HIU localizes in cancerous tissues, a therapeutic metal such as ¹⁸⁶Re could be substituted for ^{99m}T to deliver radiation to such tissues.

Full characterization of the Tc-protein complexes discussed in this thesis is not possible with ^{99m}Tc, but analogous Re complexes could be used in its place. Nonradioactive isotopes of Re are easily available in the quantities needed for characterization studies. 3HIU is the first protein designed to bind to ^{99m}Tc^I without using the bioconjugate approach. Principles learned in these studies could be applied to design of other, more biorelevant ^{99m}Tc-delivering biomolecules.



Figure 6.1 Radioactivity HPLC chromatograms using Method I showing production of $Tc^{I}(CO)_{3}(H_{2}O)_{3}$ from TcO_{4}^{-} generator eluate. The complexation product with L-histidine is also shown.



Figure 6.2 HPLC Analysis of ^{99m}Tc complexation products with 3HIU, 3HPU, and 3HT-Cb. All reactions with proteins are shown as blue curves with data points. Data points for the reaction with 3HIU are shown at squares, with 3HPU as diamonds and with 3HT-Cb as circles.


Figure 6.3 Initial L-His competition assays with designed proteins. Tc-3HIU is shown at the top, Tc-3HPU is shown in the middle, and Tc-3HT-Cb is shown at the bottom.



Figure 6.4 L-His competition assays with 3HIU (top) and wt Ubq (bottom).



Figure 6.5 Timed binding assays for $Tc^{I}(CO)_{3} + 3HIU$ reactions (top), + wt Ubq (middle) and + L-His (bottom)

Notes

- (1) Guo, Z.; Sadler, P. J. Angew. Chem. Int. Ed. 1999, 38, 1512-1531.
- (2) Abram, U.; Alberto, R. J. Braz. Chem. Soc. 2006, 17, 1486-1500.
- (3) Ferro-Flores, G.; de Murphy, C. A.; Melendez-Alafort, L. *Curr. Pharm. Anal.* **2006**, *2*, 339-352.
- (4) Egli, A.; Alberto, R.; Tannahill, L.; Schibli, R.; Abram, U.; Schaffland, A.; Waibel, R.; Tourwe, D.; Jeannin, L.; Iterbeke, K.; Schubier, P. A. *J. Nuc. Med.* **1999**, *40*, 1913-1917.
- (5) Chen, X.; Guo, Y.; Liu, B. J. Pharm. Biomed. Anal. 2007, 43, 1576-1579.

CHAPTER 7: SUMMARY AND FUTURE DIRECTIONS

7.1 Zn-Binding by Protein Designs

7.1.1 Engrailed Homeodomain/Astacin Chimera: Z0

The chimeric approach used to design an astacin-like Zn^{2+} -binding site into *engrailed* homeodomain was not as robust as first thought. Previous research led to design of stable, soluble, Ln^{3+} -binding proteins. However, in the case of Z0, several bacterial expression and protein purification strategies failed to isolate this protein. The helix in the Zn^{2+} -binding site of astacin also included a hydrophobic patch aiding in proper folding of the parent protein. However, when this patch was introduced into a surface-accessible region of the homeodomain motif, it reduced the solubility of this chimera either due to protein misfolding or to aggregation. The behavior of Z0 is similar to that of other proteins with hydrophobic regions on their surfaces (*i.e.* membrane proteins) in that it was very difficult to keep this protein in solution and stabilize its monomers.

To make a soluble, stable, Zn^{2+} -binding homeodomain, histidine residues would need to be incorporated into *engrailed* more judiciously. Such residues could be incorporated in flexible loops or as separate point mutations. If such a soluble protein could be made, it would be capable of bringing a catalytic metal-ion into close proximity of DNA.

7.1.2 Trp-cage Mutants: 3HT-C and 3HT-Cb

Metal-binding sites were successfully incorporated onto this scaffold to make small (25 amino acid), Zn^{2+} -binding peptides. Mutants had similar structures to their parent peptide as seen by CD but were more flexible. Intrinsic-tryptophan fluorescence quenching and circular dichroism studies indicate 1:1 Zn^{2+} :peptide binding with affinity in the low micromolar range. RP-HPLC indicated that Zn^{2+} -binding increased

homogeneity of the sample. This evidence shows that an extremely small, metal-binding Trp-cage was successfully made.

Other methods could be applied to fully characterize this Zn^{2+} -binding peptide. Many Trp-cage researchers use 2-D ¹H-NMR to study these small peptides. Sample preparation and NMR methods used could be optimized so that the apo- and holopeptides' structures could be determined. Upon full characterization, the structural effect of metal binding could be elucidated. Metal-substitution studies with Cu²⁺ and Co²⁺binding monitored by UV-Vis spectroscopy could determine the ion coordination environment. These peptides' structural integrity needs future study and optimization. Initial results indicated that these peptides were poorly folded at room temperature, but the T_m was not quantified. Removal of the N-terminal tail would probably increase stability and still provide a 2-His mutant. Such a design would not be useful for binding labile metal ions, but could effectively bind ions such as Pt^{II}. Using a different approach, this peptide could be further stabilized by introducing histidines at each terminus, thereby locking the peptide into a more rigid structure upon metal-binding.

7.1.3 Ubiquitin Mutants: 3HIU and 3HPU

Metal-binding sites were also successfully incorporated onto this scaffold using two different approaches. In 3HIU, a loop insertion was used to make a 3-His binding site, and in 3HPU, three His point mutations were made. Mutants' stability was investigated by circular dichroism. Both designs retained structural characteristics of wt ubiquitin, and they could not be completely denatured by heat. However, 3HIU was slightly more stable than its sister protein, 3HPU. Their Zn^{2+} -binding affinities were determined by CD by monitoring structure induction upon addition of metal. Results indicate 1:1 Zn^{2+} :protein binding with micromolar affinity.

As with 3HT-C and 3HT-Cb, 2-D ¹H-NMR parameters could be optimized and metal-substitution studies performed so that 3HIU and 3HPU could be more fully

125

characterized. The efficacy of ubiquitin as a scaffold for metalloprotein design could be tested by building an entire library of 3-His mutants. 3HIU and 3HPU introduced histidine residues at sites near in primary sequence. In other mutants, histidines could be placed at various locations on ubiquitin's surface regardless of sequence proximity. This protein could even be used as a scaffold upon which other types of binding sites could be designed to target other metals or even small organic molecules. Ubiquitin is so stable that it easily serves as a protein scaffold, and tryptophan mutants have been successfully made in other laboratories. This residue could be placed near binding sites to introduce a spectroscopic reporter of binding.

7.2 Ruthenium-Binding by Ubiquitin Mutants

Performing Ru + Ubq mutant reactions under different conditions (37 °C in air for 34 days *vs.* 75 °C in presence of limited oxygen for 6 hours) leads to formation of different products. Lower temperature reactions performed in air for long periods of time produce Ru-protein oligomers of finite size. UV-Vis spectra of the products show a band at ~460 nm. Reactions performed at higher temperature for shorter times and in the presence of limited oxygen produce fewer oligomers. UV-Vis spectra of these products show the appearance of the 460 nm band and another at one ~320 nm. Purified, monomeric Ru-protein products were studied by ESI-MS to confirm metal-binding. CD spectra indicated that neither Ru-binding nor protein oligomerization caused denaturation of the ubiquitin scaffold. These designs are the first examples of Ru-binding by a 3-His binding site. Such sites are used in many bioinorganic systems, but substitution of their natural metal ions for kinetically-inert ones is often difficult because these metals cannot reach their proper binding sites before incorrectly cross-linking residues at the proteins' surfaces. 3HIU and 3HPU were designed to overcome this difficulty by having their 3-His sites pre-organized in a surface-accessible location on ubiquitin.

Ru-protein complexes need to be further characterized after a systematic investigation of reaction conditions for Ru^{II}(tacn)-binding by 3HIU and 3HPU. Variables include temperature (37 °C *vs.* 75 °C), environment (aerobic *vs.* anaerobic), and time (hours *vs.* days). Performing Ru + protein reactions under conditions chosen to vary parameters in a more methodical fashion may aid in elucidation of the identity of compounds giving rise to the different UV-Vis absorbances observed at 450 nm and/or 320 nm. Together with studies of well-defined reference compounds such as $[Ru^{II}(tacn)(im)_3]^{2+}$, the species associated with the observed UV-Vis features can be identified.

⁹⁹Ru or ¹⁰¹Ru-NMR of Ru^{II}-protein complexes could lead to determination of the metal-binding environment. A library of ruthenium compounds can be made to test differences in their binding to proteins. Larger facially-coordinated caps could be used to sterically hinder protein oligomer formation and perhaps prevent the need for purification after reactions. Other protein scaffolds can be used to make Ru-binding sites capable of delivering the ion to a variety of biological targets. For instance, a DNA and Ru-binding protein could be developed to crosslink genes of interest.

7.3 Technetium-Binding by Ubiquitin Mutants

3HIU binds to $Tc^{I}(CO)_{3}$ and is stable to free L-histidine substitution. Tc-3HIU complexes are more inert than Tc-3HPU or Tc-wt Ubq ones. These studies show that proteins can bind ^{99m}Tc^I ions without introduction of non-natural chelates and spacers, an exciting step for the development of novel, targeted PET imaging agents.

Stability of Tc-3HIU *vs*. human plasma needs to be established in order to prove that protein chelates are strong enough to prevent Tc^I transfer to other naturally-ocurring proteins and small-molecules. Rhenium analogs of Tc^I(CO)₃(H₂O)₃ can be made to enable further characterization of such metal-protein complexes. As with Ru-binding proteins, new scaffolds capable of targeting specific tissues could be designed and used in medical imaging. If ¹⁸⁶Re were substituted for ^{99m}Tc, protein complexes could deliver a potentially therapeutic metal to very specific targets.

BIBLIOGRAPHY

- Abram, U.; Alberto, R. J. Braz. Chem. Soc. 2006, 17, 1486-1500.
- Ahmed, Z.; Beta, I. A.; Mikhonin, A. V.; Asher, S. J. Am. Chem. Soc. 2005, 127, 10943-10950.
- Ang, W. H.; Dyson, P. J. Eur. J. Inorg. Chem. 2006, 4003-4018.
- Arnold, J.; Hoffman, C. G.; Dawson, D. Y.; Hollander, F. J. Organometallics 1993, 12, 3645-3654.
- Axup, A.; Albin, M.; Mayo, S. L.; Crutchley, R. J.; Gray, H. B. J. Am. Chem. Soc. 1988, 110, 435-439.
- Baird, I. R.; Rettig, S. J.; James, B. R.; Skov, K. A. Can. J. Chem. 1998, 76, 1379-1388.
- Barua, B.; Andersen, N. H. Letters Peptide Sci. 2002, 8, 221-226.
- Batista, A. A.; Polato, E. A.; Queiroz, S. L.; Nascimento, O. R.; James, B. R.; Rettig, S. J. *Inorg. Chim. Acta* 1995, 230, 111-114.
- Berg, J. M.; Godwin, H. A. Annu. Rev. Biophys. Biomol. Struct. 1997, 26, 357-371.
- Bertini, I.; Luchinat, C.; Parigi, G. Solution NMR of Paramagnetic Molecules; Elsevier Science: Amsterdam, 2001.
- Bode, W.; Gomis-Ruth, F.-X.; Stockler, W. FEBS 1993, 331, 134-140.
- Brennan, R. G.; Matthew, B. W. Trends Biochem. Sci. 1989, 14, 286-290.
- Bunagan, M. R.; Yang, X.; Saven, J. G.; Gai, F. J. Phys. Chem. B 2006, 110, 3759-3763.
- Catanzariti, A.-M.; Soboleva, T. A.; Jans, D. A.; Board, P. G.; Baker, R. T. *Prot. Sci.* **2004**, *13*, 1331-1339.
- Chen, X.; Guo, Y.; Liu, B. J. Pharm. Biomed. Anal. 2007, 43, 1576-1579.
- Chevalier, B. S.; Kortemme, T.; Chadsey, M. S.; Baker, D.; Monnat, R. J.; Stoddard, B. L. *Mol. Cell* **2002**, *10*, 895-905.
- Costantini, D. L.; Hu, M.; Reilly, R. M. Cancer Biother. Radiopharm. 2008, 23, 3-24.
- Cowan, J. A. Chem. Rev. 1998, 98, 1067-1088.
- Crosetto, N.; Bienko, M.; Dikic, I. Mol. Cancer Res. 2006, 4, 899-904.
- DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. Annu. Rev. Biochem. **1999**, 68, 779-819.
- DeMarco, A.; Casatta, E.; Savaresi, S.; Geerlof, A. J. Biotech. 2004, 107, 125-133.
- Dikic, I.; Crosetto, N.; Calatroni, S.; Bernasconi, P. Eur. J. Cancer 2006, 42, 3095-3102.

Dragan, A. I.; Li, Z.; Makeyeva, E. N.; Millgotina, E. I.; Liu, Y.; Crane-Robinson, C.; Privalov, P. L. *Biochem.* **2006**, *45*, 141-151.

Draganescu, A.; Tullium, T. C. J. Mol. Biol. 1998, 276, 529-536.

- Ebsworth, E. A. V.; Rankin, D. W. H.; Cradock, S. In *Structural Methods in Inorganic Chemistry* Second ed.; Blackwell Scientific Publications: London, 1991, p 33.
- Egli, A.; Alberto, R.; Tannahill, L.; Schibli, R.; Abram, U.; Schaffland, A.; Waibel, R.; Tourwe, D.; Jeannin, L.; Iterbeke, K.; Schubier, P. A. J. Nuc. Med. **1999**, 40, 1913-1917.
- Engelborghs, Y. J. Fluorescence 2003, 13, 9-16.
- Ermolenko, D. N.; Dangi, B.; BGvritishvili, A.; Gronenborn, A.; Makhatadze, G. I. *Biophys. Chem.* **2007**, *126*, 25-35.
- Evans, P.; Spencer, A.; Wilkinson, G. J. Chem. Soc. Dalton Trans. 1973, 204-209.
- Ferraro, D. M., Ferraro, D.J., Ramaswamy, S., Robertson, A.D. J. Mol. Biol. 2006, 359, 390-402.
- Ferraro, D. M.; Hope, E. K. R., A.D. J. Mol. Biol. 2005, 352, 575-584.
- Ferro-Flores, G.; de Murphy, C. A.; Melendez-Alafort, L. Curr. Pharm. Anal. 2006, 2, 339-352.
- Fersht, A. R. In *Structure and Mechanism in Protein Science*; Julet, M. R., Ed.; W.H. Freeman and Company: New York, 1999, p 523-532.
- Fraenkel, E.; Rould, M. A.; Chambers, K. A.; Pabo, C. O. J. Mol. Biol. 1998, 284, 351-361.
- Franklin, S. J.; Welch, J. T. Comm. Inorg. Chem. 2005, 26, 127-164.
- Freemont, P. S.; Lane, A. N.; Sanderson, M. R. Biochem. J. 1991, 278, 1-23.
- Garayoa, W. G.; Ruegg, D.; Blauenstein, P.; Zwimpfer, M.; Khan, I. U.; Maes, V.; Blanc, A.; Beck-Sickinger, A. G.; Tourwe, D. A.; Schubiger, A. *Nuc. Med. Biol.* 2007, *34*, 17-28.
- Geilenkirchen, A.; Neubold, P.; Schneider, R.; Wieghardt, K.; Florke, U.; Haupt, H. J.; Nuber, B. J. Chem. Soc. Dalton Trans. **1994**, 457-464.
- Ghadiri, M. R.; Case, M. A. Angew. Chem. Int. Ed. 1993, 32, 1594-1597.
- Ghadiri, M. R.; Fernholz J. Am. Chem. Soc. 1990, 112, 9633-9635.
- Ghadiri, M. R.; Soares, C.; Choi, C. J. Am. Chem. Soc. 1992, 114, 4000-4002.
- Ghadiri, M. R.; Soares, C.; Choi, C. J. Am. Chem. Soc. 1992, 114, 825-831.
- Ghosh, D.; Pecararo, V. L. Inorg. Chem. 2004, 43, 7902-7915.

Guex, N.; Peitsch, M. C. Electrophoresis 1997, 18, 2714-2723.

Guo, Z.; Sadler, P. J. Angew. Chem. Int. Ed. 1999, 38, 1512-1531.

- Hanes, S. D.; Brent, R. Cell 1989, 57, 1275-1283.
- Hanes, S. D.; Brent, R. Science 1991, 251, 426-430.
- Harper, J. W.; Schulman, B. A. Cell 2006, 124, 1133-1136.
- Harris, D. C. In *Quantitative Chemical Analysis*; Julet, M. R., Ed.; W.H. Freeman and Company: New York, 1999, p 549.
- Harris, K. L.; Lim, S.; Franklin, S. J. Inorg. Chem. 2006, 45, 10002-10012.
- Hartinger, C. G.; Zorbas-Seifried, S.; Jakupec, M. A.; Kynast, B.; Zorbas, H.; Keppler, B. *J. In. Biochem.* **2006**, *100*, 981-904.
- Hegg, E. R.; Burstyn, J. N. Coord. Chem. Rev. 1998, 173, 133-165.
- Henn, M.; Alessio, E.; Mestroni, G.; Calligaris, M.; Attia, W. M. *Inorg. Chim. Acta* **1991**, *187*, 39-50.
- Hong, J.; Kharenko, O. A.; Ogawa, M. Y. Inorg. Chem. 2006, 45.
- Hong, J.; Miao, Y.; Miao, R.; Yang, G.; Tang, H.; Guo, Z.; Zhu, L. J. Mass Spec. 2005, 40, 91-99.
- Hoogland, G. E.; Gattiker, A.; Wilkins, M. R.; Appel, R. D.; Bairoch, A. In *The Proteomics Protocols Handbook*; Walker, J. M., Ed.; Humana Press: 2005.
- Ibarra-Molero, B.; Loladze, V. V.; Makhatadze, G. I.; Sanchez-Ruiz, J. M. *Biochem.* **1999**, *38*, 8138-8149.
- Ippolito, J. A.; Baird, T. T.; McGee, S. A.; Christianson, D. W.; Fierke, C. A. *PNAS* **1995**, *92*, 5017-5021.
- Jacobson, M. P.; Pincus, D. L.; Rapp, C. S.; Day, T. J. F.; Honig, B.; Shaw, D. E.; Friesner, R. A. Proteins 2004, 55, 351-367.
- Jain, S.; Welch, J. T.; Horrocks, W. D.; Franklin, S. J. Inorg. Chem. 2003, 42, 8098-8104.
- Kelly, S. M.; Jess, R. J.; Price, N. C. Biochim. Biophys. Acta 2005, 1751, 119-139.
- Kim, Y.; Franklin, S. J. Inorg. Chim. Acta 2002, 341, 107-112.
- Kim, Y.; Welch, J. T.; Lindstrom, K. M.; Franklin, S. J. J. Biol. Chem. 2001, 6, 173-181.
- Kornberg, T. B. J. Biol. Chem. 1993, 268, 26813-26816.
- Kostova, I. Curr. Med. Chem. 2006, 13, 1085-1107.
- Kovacic, R. T.; Welch, J. T.; Franklin, S. J. J. Am. Chem. Soc. 2003, 125, 6656-6662.

- Laemmli, U. K. Nature 1970, 227, 680-685.
- Laughon, A. Biochem. 1991, 30, 11357-11367.
- Ledneva, R. K.; Alexeevskii, A. V.; Vasil, S. A.; Sprin, S. A.; Karyagina, A. S. *Mol. Biol.* **2001**, *35*, 647-659.
- Lee, H.-W.; Matzapetakis, M.; Mitra, S.; Neil, E.; Marsh, G.; Pecararo, V. L. J. Am. Chem. Soc. 2004, 126, 9178-9179.
- Li, C.; Banfield, M. J.; Dennison, C. J. Am. Chem. Soc. 2007, 129, 709-718.
- Lim, S., The University of Iowa, 2006.
- Lim, S.; Franklin, S. J. Prot. Sci. 2006, 15, 2159-2165.
- Lipscomb, W. N.; Strater, N. Chem. Rev. 1996, 96, 2375-2433.
- Lu, Y. Angew. Chem. Int. Ed. 2006, 45, 5588-5601.
- Lu, Y. Inorg. Chem. 2006, 45, 9930-9940.
- Lu, Y.; Berry, S. M.; Pfister, T. D. Chem. Rev. 2001, 101, 3047-3080.
- Lu, Y.; Gralla, E. B.; Roe, J. A.; Valentine, J. S. J. Am. Chem. Soc. **1992**, 114, 3560-3562.
- Lu, Y.; Roe, J. A.; Bender, C. J.; Peisach, J.; Banci, L.; Bertini, I.; Gralla, E. B.; Valentine, J. S. *Inorg. Chem.* **1996**, *35*, 1692-1700.
- Maret, W. J. Trace Elem. Med. Biol. 2005, 19, 7-12.
- Montclare, J. K.; Schepartz, A. J. Am. Chem. Soc. 2003, 125, 3416-3417.
- Muheim, A.; Todd, R. J.; Casimiro, D. R.; Gray, H. B.; Arnold, F. H. J. Am. Chem. Soc. **1993**, *115*, 5312-5313.
- Mukhopadhyay, D.; Riezman, H. Science 2007, 12, 201-205.
- Naduthambi, D.; Zondlo, N. J. J. Am. Chem. Soc. 2006, 128, 12430-12431.
- Neidigh, J. W.; Fesinmeyer, R. M.; Andersen, N. H. Nat. Struct. Biol. 2002, 9, 425-430.
- Nomura, A.; Sugiura, Y. J. Am. Chem. Soc. 2004, 126, 15374-15375.
- Nomura, A.; Sugiura, Y. Inorg. Chem. 2004, 43, 1708-1713.
- Noyori, R.; Hashiguchi, S. Acc. Chem. Res. 1997, 30, 97-102.
- Patikoglou, G.; Burley, S. K. Annu. Rev. Biophys. Biomol. Struct. 1997, 26, 289-325.
- Pokala, N.; Handel, T. M. J. Struct. Biol. 2001, 134, 269-281.

- Qiu, L.; Pabit, S. A.; Roitberg, A. E.; Hagen, S. J. J. Am. Chem. Soc. 2002, 124, 12952-12953.
- Reedijk, I. PNAS 2003, 100, 3611-3616.
- Regan, L. Trends Biochem. Sci. 1995, 20, 280-285.
- Royer Chem. Rev. 2006, 106, 1769-1784.
- Sambrook, J.; Russell, D. W. In *Molecular Cloning, A Laboratory Manual*; Third ed.; Argentine, J., Ed.; Cold Spring Harbor Laboratory Press: New York, 2001; Vol. 3, p A8.46-A8.47.
- Sambrook, J.; Russell, D. W. In *Molecular Cloning, A Laboratory Manual*; Third ed.; Argentine, J., Ed.; Cold Spring Harbor Laboratoratory Press: New York, 2001; Vol. 2, p 8.18-8.24.
- Schnell, J. D.; Hicke, L. J. Biol. Chem. 2003, 278, 35857-35860.
- Sera, T.; Uranga, C. Biochem. 2002, 41, 7074-7081.
- Serrano, S.; Sala, X.; Plantalech, E.; Rodriguez, M.; Romero, I.; Jansat, S.; Gomez, M.; Parella, T.; Stoeckli-Evans, H.; Solans, X.; Font-Bardia, M.; Vidjayacoumar, B.; Llobet, A. *Inorg. Chem.* 2007, 46, 5381-5389.
- Shields, S. B., The University of Iowa, 2005.
- Shields, S. B.; Franklin, S. J. Biochem. 2004, 43, 16086-16091.
- Shields, S. B.; Franklin, S. J. J. Inorg. Biochem. 2007, 101, 783-788.
- Sirish, M.; Franklin, S. J. J. Inorg. Biochem. 2002, 91, 253-258.
- Stollar, E. J.; Gelpi, J. L.; Valankar, S.; Golovin, A.; Orozco, M.; Luisi, B. Proteins 2004, 57, 1-8.
- Stollar, E. J.; Mayori, U.; Lovell, S. C.; Federici, L.; S.M.V., F.; Fersht, A. R.; Luisi, B. F. J. Biol. Chem. 2003, 278, 43699-43708.
- Streicher, W. W.; Makhatadze, G. I. Biochem. 2007, 46, 2876-2880.
- Sundberg, R. J.; Bryan, R. F.; Taylor, I. F.; Taube, H. J. Am. Chem. Soc. 1974, 96, 381-392.
- Szymkowski, D. E. Curr. Opin. Drug Discov. Devel. 2005, 8, 590-600.
- Vallee, B. L.; Auld, D. S. Biochem. 1990, 29, 5647-5659.
- Welch, J. T.; Kearney, W. R.; Franklin, S. J. PNAS 2003, 100, 3725-3730.
- Welch, J. T.; Sirish, M.; Lindstrom, K. M.; Franklin, S. J. Inorg. Chem. 2001, 40, 1982-1984.
- Wilcox, D. E. Chem. Rev. 1996, 96, 2435-2458.

Wong-Deyrup, S. W., The University of Iowa, 2007.

Wong-Deyrup, S. W.; Kim, Y.; Franklin, S. J. J. Biol. Inorg. Chem. 2006, 11, 17-25.

- Woody, R. W. Meth. Enzymol. 1995, 246, 34-71.
- Woolfson, D. N. Adv. Protein Chem. 2005, 70, 79-112.
- Yamaguchi, R.; Harada, K.; Sunatusuki, Y.; Kojima, M.; Nakajima, K.; Matsumoto, N. *Eur. J. Chem.* **2006**, 3326-3243.
- Ye, Y.; Lee, H.-W.; Yang, W.; Shealy, S. J.; Yang, J. J. J. Am. Chem. Soc. 2005, 127, 3743-3750.