# *De Novo* Design of Copper Metallopeptides Capable of Electron Transfer: From design to function

**by**

**Jefferson Plegaria**

**A dissertation submitted in partial fulfillment**

**of the requirements for the degree of**

**Doctor of Philosophy**

**(Chemistry)**

**in the University of Michigan**

**2015**

**Doctoral Committee:**

**Professor Vincent L. Pecoraro, Chair**
**Professor Carol A. Fierke**
**Associate Professor Nicolai Lehnert**
**Professor Stephen W. Ragsdale**
**Professor Timothy L. Stemmler, Wayne State University**

To my family

## Acknowledgements

This thesis would not have been possible without the guidance and support of many individuals both at and associated with our lab. I would like to give my deepest appreciation to my research advisor Professor Vincent L. Pecoraro, whose knowledge and enthusiasm for science encouraged me throughout graduate school. Thank you for allowing me to explore and learn a variety of scientific techniques, from protein expression and NMR to electrochemistry and photophysics. I am also thankful that you allowed me to work in Europe. I was fortunate enough to attend a conference in Spain, collect data at CERM and work in world-renowned labs in France. These experiences have enriched and strengthened my graduate school experience. Thank you for letting me GSI your undergraduate bioinorganic chemistry class. Both teaching, as well as our Friday scientific discussions after class elevated my understanding of this topic. I will miss these discussions. Finally, I would like to thank you for helping me maintain perspective. As an aside, I will also miss hearing about all your adventures with Peggy.

I would also like to thank my dissertation committee member, Professors Carol A. Fierke, Nicolai Lehnert, Stephen W. Ragsdale and Timothy L. Stemmler, for their comments and support.

I would like to thank all my research collaborators in Michigan and France. First, I would like to thank Professor Erik R. P. Zuiderweg for teaching me how to make, collect and process NMR data, as well as Professor Timothy L. Stemmler for helping me solve the structure of $\alpha_3$D**IV** and for his career advice. I am indebted to these two individuals as the NMR work makes up chapter 2 of my thesis. I would like to acknowledge and thank Dr. Anirrudha Deb and Professor James Penner-Hahn for analyzing and fitting my XAS data.

I am truly thankful for the labs at the University of Paris, Diderot and CEA Saclay. The work in both labs made the fourth chapter of my thesis possible. In Paris, I would like to thank Drs. Marc Robert and Cèdric Tard for allowing me to run electrochemistry experiments. Thank you Dr. Tard for your encouragement and friendship. I would like to offer special thanks Dr.

Matteo Duca who taught me everything I know about electrochemistry. At the CEA, I would like to thank Drs. Winfried Leibl and Annamaria Quaranta for providing me the opportunity to study the photophysical properties on my copper constructs. Without this collaboration, I would not be able to add the word "function" to my thesis title. I am extremely thankful to Dr. Quaranta for all your support, hard work in collecting photophysical data and teaching me the science behind laser flash photolysis.

I would like to acknowledge the Rackham Graduate School for funding. I thank Rackham for awarding me a Rackham Merit Fellowship. I was also fortunate to receive funding to return to France in 2013 and 2014 via the Rackham International Research Awards (Warner Linfield Ward) and Rackham Graduate Research Grant, respectively. I would also like to thank Rackham for funding the conferences I attended through the Rackham Conference Travel Grant.

I would like to thank former and current lab (permanent and visiting) members. I owe special thanks to Drs. Matteo Tegoni and Virginia Cangelosi for their mentorship and friendship. I am thankful to Dr. Alison Tebo, Cathy, Dr. David Chow, Dr. Evan Trivedi, Dr. Fangting Yu, Jake, Dr. Joseph Jankolovits, Kathy, Dr. Kosh Neupane, Leela, Dr. Saumen Chakraborty, Dr. Ted Boron and Dr. Tu Nguyen for all their help and encouragement. I would like to offer my appreciation to the undergraduates I was fortunate to mentor and work with, Jennifer, Sarah and Thomas.

Finally, I would like to thank my family and friends (in Michigan and back East) for all their love and support. I am especially thankful to Michael for my source of inspiration starting from our research job at CHOP and to our adventures here Michigan. Finally, I thank my parents (who are nicely retired in the Philippines) and Michael's parents for their continuous support.

## TABLE OF CONTENTS

## List of Tables

## List of Figures

**Figure I-11.** Metal binding site of plastocyanin (A), stellacyanin (B) and nitrosocyanin (C). Plastocyanin and CuT1 centers with a similar coordination environment have a trigonal pyramidal geometry where the copper ion and the core 2HisCys residues make the base of the pyramid and the Met is the apex. Stellacyanin and its phytocyanins family with a Gln at the axial position display a distorted tetrahedral geometry, where the bond between the Cu and the axial O(Gln) ligand shortens while the Cu-S(Cys) bond elongates. A more recent addition to the CuT1 family is nitrosocyanin, which was determined to complex a copper center in a square pyramidal geometry. The metal binding site of nitrosocyanin lacks S(Met) ligand but instead is comprised of a O(Glu) ligand in the axial position and a 2HisCys ligand and a $H_2O$ molecule that make up the base of the square. 23

**Figure I-12.** Overlay of the oxidized (light gray) and reduced (dark gray) CuT1 centers, which illustrates minimal change in the bond lengths and geometry. The bond lengths are adjacent to their corresponding residue and values inside parentheses belong to the reduced state. A) Plastocyanin from *Poplar nigra* [PDB 1PLC (oxidized) at pH 6.0 & 5PCY (reduced) at pH 7.0 and 1.80 Å resolution]. B) Horse radish umecyanin, which is part of the stellacyanin family [PDB 1X9U (oxidized) at pH 7.5 and 1.90 Å resolution & 1X9R (reduced) at pH 7.5 and 1.90 Å resolution]. C) Nitrosocyanin from *Nitrosomonas europea* [PDB 1IBY (oxidized) at pH 7.5 and 1.65 Å resolution & 1IBZ (reduced) at pH 7.5 and 2.3 Å resolution]. 24

**Figure I-13.** UV-VIS (A) and EPR (B) spectra of blue copper centers compared to normal copper sites. Blue copper proteins display a strong absorption band in the visible range, while normal copper centers show a strong feature in the UV range. Further, EPR spectra of blue copper sites have an $A_{||}$ value $<100 \times 10^{-4}$ $cm^{-1}$, whereas normal copper complexes often have

## Abstract

Structural and biophysical characterization on *de novo* designed three-helical bundles is presented. The structure of a designed construct was determined to establish the physical integrity of a peptide scaffold. Subsequently, spectroscopic, electrochemical and photophysical studies were employed to characterize designed redox-active copper sites.

$\alpha_3$D is *de novo* designed peptide that preassembles into an antiparallel three-helix bundle fold. This scaffold was previously functionalized with a triscysteine site to produce $\alpha_3$D**IV**. An integral step in *de novo* protein design is the establishment of a well-defined scaffold, which is a fundamental characteristic of native proteins. The NMR solution structure of apo $\alpha_3$D**IV** was solved to demonstrate this characteristic and to determine the effects of incorporating a metal binding site (Cys18, Cys28, and Cys67) into the framework of $\alpha_3$D. The structure of $\alpha_3$DIV (at pH 7.0) comprised 1067 NOE restraints derived from multinuclear multidimensional NOESY, as well as 138 dihedral angles ($\psi$, $\phi$ and $\chi_1$). The backbone and heavy atoms of the 20 lowest energy structures have an RMSD from the mean structure of 0.79 (0.16) Å and 1.31 (0.15) Å, respectively, demonstrating a high quality structure. Overall, the structure of apo $\alpha_3$D**IV** reveals a distorted triscysteine metal binding site, which offers a model for native proteins with thiol-rich ligands that function in regulating toxic heavy metals, as well as future designed constructs.

Recent work in the Pecoraro group produced derivatives of the $\alpha_3$D scaffold that served as excellent functional models of a native metalloregulatory and catalytic protein, with symmetric/same-ligand metal binding sites. Efforts to functionalize the $\alpha$-helical fold of $\alpha_3$D with a redox-active site, as well as an asymmetric/mixed-ligand site had not been attempted. The asymmetric metal binding site of electron transfer proteins, cupredoxins, which are naturally observed in a $\beta$-barrel fold, was incorporated in the helical fold of $\alpha_3$D. The preassembled fold of $\alpha_3$D provides a more direct incorporation of mixed-ligand/asymmetric metal center, which has

been a challenge in current *de novo* designed scaffolds. The overall objective is to examine whether an electron transfer site can be achieved in and whether the physical properties (absorption band at 600 nm, copper-sulfur bond of 2.1 – 2.2 Å and reduction potential of 180 – 800 mV vs. NHE) of a cupredoxin center can be recapitulated in a markedly unrelated fold.

The asymmetric 2HisCys(Met) residues of cupredoxins were incorporated in $\alpha_3$D to generate three distinct designs designated core ($\alpha_3$D-CR1), chelate ($\alpha_3$D-CH3 and $\alpha_3$D-CH4) and chelate-core ($\alpha_3$D-ChC2) constructs. The core construct incorporates the 2HisCys residues on three separate strands (His18, Cys28 and His67, as well as Met72). The chelate derivatives possess an L–$X_n$–L (L= His or Cys, n = #nonligating residues) chelate motif at the end of helix 1 (positions 18 and 21, as well His28 and Met72), while the chelate-core design contains a Cys14-$X_3$-His18 chelate motif (and His31) deeper within the three-helix bundle. Cu(II) binding to the core and chelate constructs displayed intense absorption bands between 380-400 nm (~2000 $M^{-1}$ $cm^{-1}$); whereas, the chelate-core construct showed two intense absorption bands at 401 (4429 $M^{-1}$ $cm^{-1}$) and 499 (2020 $M^{-1}$ $cm^{-1}$). Despite not retaining the spectroscopic profile for the Cu(II) state, XAS analysis on the Cu(I) adducts recapitulated the reduced state of cupredoxin proteins. These Cu(I) complexes revealed a short Cu-S(Cys) bond at 2.16 – 2.23 Å, as well as Cu-N(His) bond distances of 1.94 – 1.98 Å. Overall, these results showed that the lack of rigidity at the designed metal binding sites cannot enforce the proper constraints to achieve the appropriate Cu(II) chromophore, however the Cu(I) environment was retained within the α-helical bundle of $\alpha_3$D.

The redox activity of the designed constructs was tested using electrochemical and photophysical methods. The electrochemistry studies showed that the core, chelate and chelate-core derivatives possess reduction potentials of +362 – +462 mV (vs. NHE), which are in the range observed for native cupredoxins. This result demonstrates that a copper-based redox center with a high positive potential can be achieved within an α-helical bundle protein. Moreover, photophysical studies, including the transient spectra and kinetic trace of the core and chelate constructs, revealed intermolecular ET activity with a photo-oxidant, ruthenium(III)trisbipyridine. The first-order and bimolecular rate constants of these constructs were determined to be $10^5$ $s^{-1}$ and $10^8$ $s^{-1}$ $M^{-1}$, respectively.

Overall, this work illustrates that the redox function of a native copper center in a β-barrel fold can be achieved in the α-helical framework of $\alpha_3$D. Ultimately, this work provides a foundation for investigating long-range electron transfer reaction using *de novo* protein design.

# Chapter I. Protein Design: from design to function

Proteins acquire metal ions such as iron, copper or zinc to perform essential functions in biology, such as catalytic reactions, signal transduction, transport and storage of small molecules and redox chemistry.[1] Metalloproteins play a central role in various biological system including photosynthesis and respiration, two systems that sustain all life on earth. Recent initiatives by the scientific community to find alternatives to carbon-based fuels have turned to the bioenergetics processes in photosynthesis for inspiration to create a "greener planet." Before efficient artificial photosynthetic systems can be realized, fundamental research on the individual metalloproteins and metalloenzymes in this multifaceted system must be examined. Ultimately, the knowledge gained from this research will bring us one step closer to fully understanding and harvesting the rewards of many biological systems.

Protein design is a biologically relevant approach used to study the concept of structure-function relationship in native proteins.[2-7] This emerging approach has two central design strategies: the first is protein redesign[2, 5] and the second is *de novo* design.[3, 5-7] Protein redesign involves modifying or incorporating a novel function in an existing native protein scaffold. Existing native scaffolds offer a thermodynamically stable frame that is amenable to mutations, deletions and insertions. Further, this stable scaffold allows for facile spectroscopic and X-ray crystallographic characterization. The protein redesign process involves a close analysis of a solved protein structure. Followed by specific active site mutations yielding a desired construct- that hopefully can function. In a step-wise manner one can interchange functions between proteins with similar cofactors and ligand environment to discover the important amino acid residues (in the primary and secondary coordination sphere) and forces (electrostatic or hydrophobic) allowing for proper protein function. This approach also allows one to investigate

the evolutionary hierarchy of proteins that have similar functions but distinct structures (and vice versa).

The second approach, which is unquestionably the most challenging strategy, employs first principles to design peptide or protein scaffolds from scratch with an amino acid sequence not found in nature. It allows one to tailor-design an original sequence that forms the proper hydrophobic, electrostatic and hydrogen bonding interactions that will manifest into a well-defined peptide scaffold, an important characteristic of native proteins. *De novo* protein design offers a novel approach in studying the mechanisms behind protein folding and exploring the active sites of native proteins in a simplified or unrelated fold. The knowledge gained from this approach could ultimately provide insight into the fundamental processes of biological systems, thus allowing the possibility to produce new metalloproteins with higher stability and superior efficiency than native proteins for many biotechnological applications. Protein design strategies have been employed for a wide range of metalloproteins.

The Pecoraro group is actively involved in the *de novo* protein design of peptide scaffolds to understand the metal active sites of metalloenzymes (Figure I-1A), as well as metalloproteins involved in electron transfer (Figure I-1B),[8] metal regulated gene expression[1, 9] and metallochaperones.[1] These later systems are a subset of metalloproteins used by microorganisms to control the levels of essential transition metal ions (Fe, Cu, Zn and Mn), decrease levels of toxic metals (Hg, As, Pb and Cd) within their cells and to ensure proper trafficking and insertion of metals into enzymes or secretory vesicles. The first half of this chapter will focus on a *de novo* designed scaffold $\alpha_3$D,[10] a single polypeptide chain that folds into a three-helix bundle (THB) The latter half will transition into a description and protein design of copper electron transfer proteins, cupredoxins. Overall, this chapter is comprised of 7 sections. In section 1, I will briefly cover the work and progress with three stranded coiled-coil (3SCC) constructs. It will then be followed with a discussion of the achievement of DeGrado and coworkers in producing a THB scaffold. In section 2, I will discuss the Pecoraro and co-workers' first approach in functionalizing the $\alpha_3$D scaffold through the incorporation of symmetric metal binding sites to

yield peptides $\alpha_3$D**IV**[11] and $\alpha_3$D**H$_3$**.[12] Section 2 will also briefly introduce my work in incorporating the asymmetric metal binding site of native cupredoxins within the $\alpha_3$D scaffold. This work will be fully described in Chapter 3. Section 4 of this chapter will introduce and discuss the physical properties of cupredoxin proteins. The following two sections, 5 and 6, will cover my attempts in examining cupredoxin centers using protein design. For the sake of brevity, only those systems that have been developed in the Pecoraro group and other research groups to investigate cupredoxin proteins will be discussed in this introduction. Readers more generally interested in this topic are referred to recent reviews for more exhaustive coverage of the topic.[2, 5, 6]

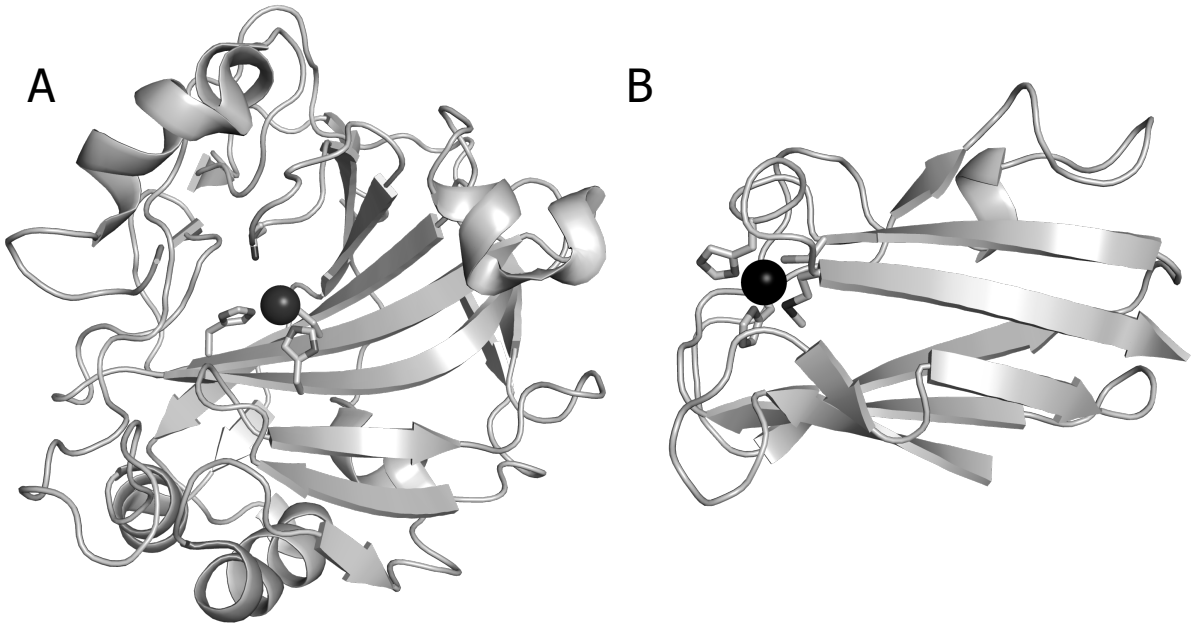



**Figure I-1.** A) Crystal structure of a zinc metalloenzyme human carbonic anhydrase II (PDB code 3KS3).[13] B) Crystal structure of a copper electron-transfer metalloprotein plastocyanin (PDB code 1PLC)[14], which is part of the cupredoxin family.

1. Development of *de novo* designed scaffolds

1.1. Three-stranded coiled-coil constructs

The most common *de novo* designed peptides use a heptad repeat sequence (abcdefg) that self-assemble into a parallel 3SCC tertiary structure (Figure I-2A).[15-17] Much of our effort and success in the *de novo* design field has been carried out using the 3SCC scaffold of TRI,[18-33] Grand (Gr)[26, 27, 34, 35] and BABY[20, 34, 36] peptides, as well as in CoilSer (CS),[23, 29, 33, 34, 37-39] which serves a

crystallographic analogue for TRI (Table I-1). A metal binding site is generated by incorporating a cysteine or penicillamine residue at the "a" (Figure I-2B) or "d" (Figure I-2C) position in a 3SCC scaffold.[38] It was discovered that the subtle difference between the "a" and "d" positions can produce distinctive outcomes in heavy metal binding affinity and geometry, which can be attributed to the preorganization of the sulfur ligands prior to metal binding.[16, 19] Overall, a deeper understanding in the metallobiochemistry of heavy metals was gained, including as As(III),[21, 33, 37] Cd(II),[16, 21-24] Hg(II)[18-21, 23, 25, 28, 29, 36] and Pb(II)[24, 33, 34, 40] in a tristhiolate site. This work demonstrated how to control the coordination number and geometry of Cd(II)[16, 26, 27, 35] and Hg(II),[20, 21, 25, 28] determined the affinity for Cd(II)[24] and Pb(II)[24, 36] based on site preferences for "a" or "d" sites of the Cys residues and elucidated the effects of the core aliphatic groups in the second coordination sphere on the molecular recognition of Cd(II)[16, 26, 27] and Pb(II).[34] This work on heavy metal chemistry in 3SCC is a culmination of over 10 years of research and has given us a solid foundation to model catalytic sites of natural metalloenzymes.

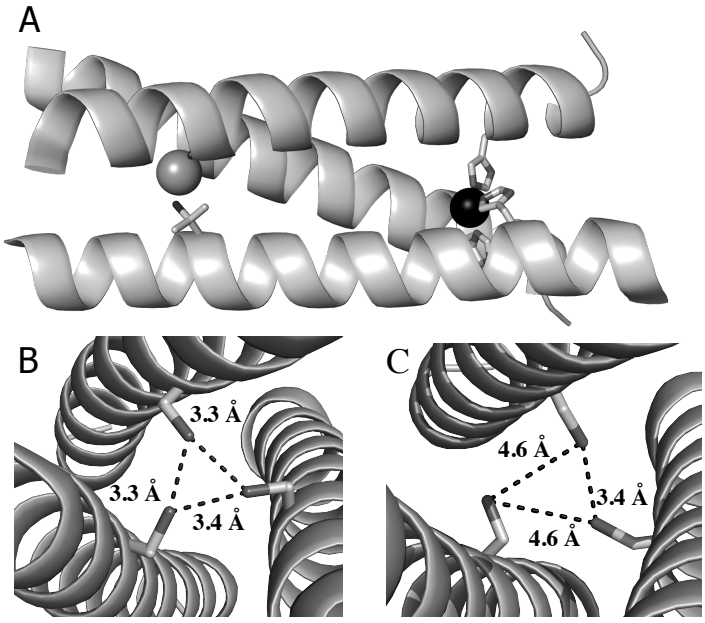



**Figure I-2**. X-ray crystal structure of *de novo* desgined peptides. A) Crystal structure of $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$ (PDB code 3PBJ),[29, 31] which is used as cystrallographic model for $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(TRIL9CL23H)_3^{n+}$ CA model. This bimetallic 3SCC construct contains a Zn(II) and Hg(II) atom bound to a His-$N_3$ and Pen-$S_3$ site, respectively. B) Symmetric "a" site Cys residues in CSL9C (PDB code 3LJM)[38] contain Sγ ligands that orient inside the core. C) Symmetric "d" site Cys residues in CSL19C (PDB code 2X6P)[38] include Sγ ligands that orient towards the interhelical interface forming a larger metal binding site than Cys "a" sites.

**Table I-1**. Amino acid sequence of *de novo* designed 3SCC analogues.

| Peptide | Sequence |
|---|---|
| | **abcdefg abcdefg abcdefg abcdefg abcdefg** |
| CoilSer | Ac-E WEALEKK LAALESK LQALEKK LEALEHG –$NH_2$ |
| Baby | Ac-G LKALEEK LKALEEK LKALEEK G-$NH_2$ |
| TRI | Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-$NH_2$ |
| Grand | Ac-G LKALEEK LKALEEK LKALEEK LKALEEK LKALEEK G-$NH_2$ |

Leucine residues at the "a" or "d" positions are mutated to metal binding residues such as cysteine, penicillamine or histidine. Mutation of Leu residues to Cys and His at the 9th and 23rd, respectively, is designated as TRIL9CL23H.

Using the 3SCC scaffold, our work has progressed into modeling the symmetric tris(histidine) metal binding site found in carbonic anhydrase (CA)[29, 31] and nitrite reductase (CuNiR).[30, 32] The $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(TRIL9CL23H)_3^{n+}$ CA model (Figure I-2A) of Zastrow *et. al.* contains a structural $Hg(II)S_3$ site towards the N-terminal end of the TRI fold (Figure I-3A) and a $Zn(II)N_3O$ catalytic site at the C-terminal end (Figure I-3B).[29] This construct is an artificial metalloenzyme that catalyzes the hydration of $CO_2$ with an efficiency faster than any other small molecule model and is within ~500-fold of CAII, the most active isoform of carbonic anhydrase. Further, the CuNiR models of Tegoni *et. al.*[30] and Yu *et. al.*[32] is capable of multiple turnover catalysis for the one electron reduction of nitrite using ascorbate, and is the first mononuclear redox enzyme that was isolated via *de novo* protein design. Nevertheless, the metal binding sites of native proteins, such as the ET site in curpredoxins, are often asymmetric and contain a mixed-ligand (O, N or S) environment (Figure I-3C).[14] Further, the coordination environment contains secondary residues that participate in hydrogen bonding networks (Figure I-3D) like in CA[13, 41] or electrostatic interactions, which are essential in the catalytic or redox activity of many metalloproteins. Even though a variety of metal centers were successfully modeled and performed the desired catalytic reactions in the 3SCC scaffolds, its self-assembling nature makes it challenging to obtain asymmetric constructs. Therefore, to achieve an asymmetric metal binding site, the work in the Pecoraro group was expanded to acquire, and then later developed, a single polypeptide sequence that folds into a preformed three-helix bundle (THB) fold.

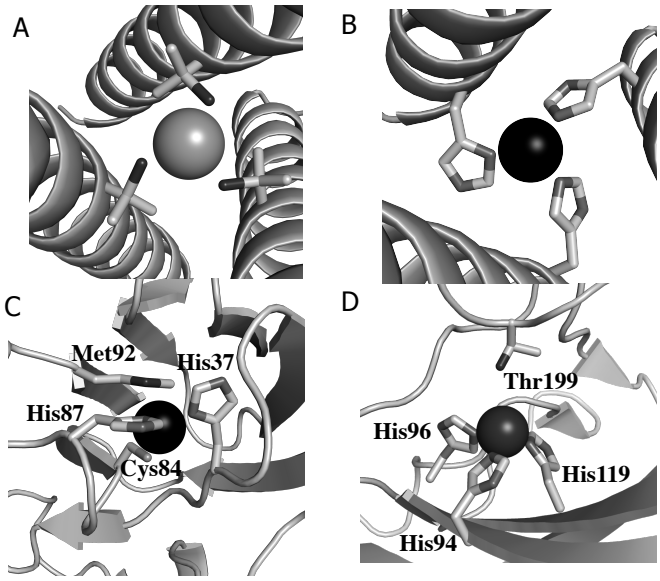



**Figure I-3.** Enlargement of metal binding sites in *de novo* designed and native proteins. A) Top-down view of the trigonal $Hg(II)S_3$ site in $[Hg(II)_SZn(II)_N(TRIL9CL23H)_3]$ that serves as a structural motif in the scaffold.[29] B) Top-down view of the $Zn(II)N_3O$ site in $[Hg(II)_SZn(II)_N(TRIL9CL23H)_3]$ that is capable of CA activity.[29] C) 2His, Cys and Met copper binding site in plastocyanin, illustrating an asymmetric metal center.[14] D) Zinc metal binding site in CAII, which contains a tris(histidine) site and hydrogen bonding residues such as Thr199 that are essential in catalysis.[13, 41] Reprinted with permission from ref 5. Copyright 2014 American Chemical Society.

1.2. From three stranded coiled-coil to a three-helix bundle fold

The THB fold is used as a molecular recognition domain found in many biological systems including immunoglobulin G, DNA binding proteins and various enzymes. Inspired by its universal presence and diverse function in nature, DeGrado and coworkers aimed to create a *de novo* designed antiparallel THB scaffold. Bryson *et. al*. used the sequence of CS as foundation because its X-ray crystal structure[42] was observed to pack in an antiparallel manner, where the helices orient in up-up-down manner, instead of a predicted parallel style.[43] Using a hierarchical approach, $\alpha_3D$[10] was the final product and isolated through a step-wise process (3 design rounds) of modifying helix-capping interactions that dictate the topology of the bundle, electrostatic interactions that orient the desired helix-helix pairing to avoid alternative states and hydrophobic interactions to achieve a well-packed core.

In the first round, Byrson *et. al*. shortened the sequence of CS by one heptad repeat to yield three 21-residue helices in $\alpha_3A$ (Table I-2).[43] To achieve an antiparallel strand, Glu and Lys

residues in helix 2 at the "e" and "g" positions were reversed from the original positions in CS. Next, two simple loops of Gly-Asn and Pro-Gly-Asn were incorporated between helix 1 and 2 and helix 2 and 3, respectively, to serve as hairpin loops. This motif is essential for directing the topology of the bundle, which can adopt clockwise or a counterclockwise orientation. $\alpha_3$A was observed to form monomer/dimer/trimer species in solution, indicating that the hairpin loops were not successful in stabilizing intermolecular interactions. Round two designs were a direct response to these issues.

Sequence $\alpha_3$B was designed to contain stronger helix stop signals to accurately direct the formation and conformation of the loops. This was achieved by lengthening residues in the loops and adding Asn residues as a helix stop signal in the form of helix capping boxes (Asn-Pro-Asp-Glu between helix 1 and 2 and Asn-Pro-Glu between helix 2 and 3). $\alpha_3$B is monomeric in solution but still retains some characteristics of a molten globule, an undefined folded state with several energetically equal conformations.

Lastly, the final round of design focused on repacking the hydrophobic core, reordering the residues involved in interhelical electrostatic interactions and further enhancing helix-capping interactions to prevent nonnative characteristics in $\alpha_3$B. First, an Asn was replaced with a Ser as the helix-capping residue between helix 1 and 2. Next, the positions of the Lys and Glu residues were redesigned to force a counterclockwise topology in the bundle (Figure I-4). The clockwise form was destabilized by careful placement of charged residues at the "e" and "g" positions. In helix 1, only positively (+) residues were placed both at the "e" and "g" positions; while in helix 3, negatively (-) charged residues were assigned in those corresponding sites. For helix 2, the "e" sites were given only - charged residues, whereas the "g" sites received + charged residues. Furthermore, some Lys residues were changed to Arg to reduce redundancy in the sequence and provide added stability gained from an Arg-Glu salt-bridge interaction. Lastly, using a genetic repacking algorithm, the hydrophobic residues at the "a" and "d" positions were altered to include various nonpolar residues such Ala, Val, Ile, Leu and Phe residues. The native-like property of $\alpha_3$C was characterized in tandem with $\alpha_3$B, and it was demonstrated to exhibit thermodynamic and spectroscopy properties of a well-defined and folded protein.

**Table I-2**. Amino acid sequence of THB analogues.

| Peptide | Sequence | Design Purpose/Function |
|---|---|---|
| | **abcdefg abcdefg abcdef loop** | |
| $\alpha_3$A | **E**WEALEKK L**N**ALESK LQALEK **G**<br>**N**WEAL**K**K**E** L**N**AL**K**S**E** LQAL**K**K **PG**<br>**E**WEALEKK L**N**ALESK LQALE**HG** | Iteration of CS |
| $\alpha_3$B | EWEALEKK L**A**ALESK LQALEK G**G**<br>**NPDE**W**A**ALKKE L**A**ALKSE LQALK **GKG**<br>**NP**EWEALEKK L**A**ALESK LQALEHG | Iteration of $\alpha_3$A |
| $\alpha_3$C | **S**W**AEFKER** LAA**IK**S**R** LQAL GG<br>**SEA**E**L**AA**FE**KE **I**AA**FE**SE LQA**Y**K GKG<br>NPE**V**EAL**R**K**E** **A**AA**IR**S**E** LQA**YR**H**N** | Iteration of $\alpha_3$B |
| $\alpha_3$D | **MG**SWAEFK**Q**R LAAIK**T**R LQAL GGS<br>EAELAAFEKE IAAFESE LQAY KGKG<br>NPEVEALRKE AAAIR**D**E LQAYRHN | Molecular recognition domain |
| $\alpha_3$D**IV** | MGSWAEFKQR LAAIKTR **C**QAL GGS<br>EAE**C**AAFEKE IAAFESE LQAY KGKG<br>NPEVEALRKE AAAIRDE **C**QAYRHN | Heavy metal peptide<br>Cd(II), Hg(II) and Pb(II) |
| $\alpha_3$D$\mathbf{H_3}$ | MGSWAEFKQR LAAIKTR **H**QAL GGS<br>EAE**H**AAFEKE IAAFESE LQAY KGKG<br>NPEVEALRKE AAAIRDE **H**QAYR**V**N**GSGA** | Carbonic anhydrase model |

The sequences are prepared in heptads. Residues that are underlined and bolded were changed from previous design. The $\alpha_3$A was altered from the CS sequence.

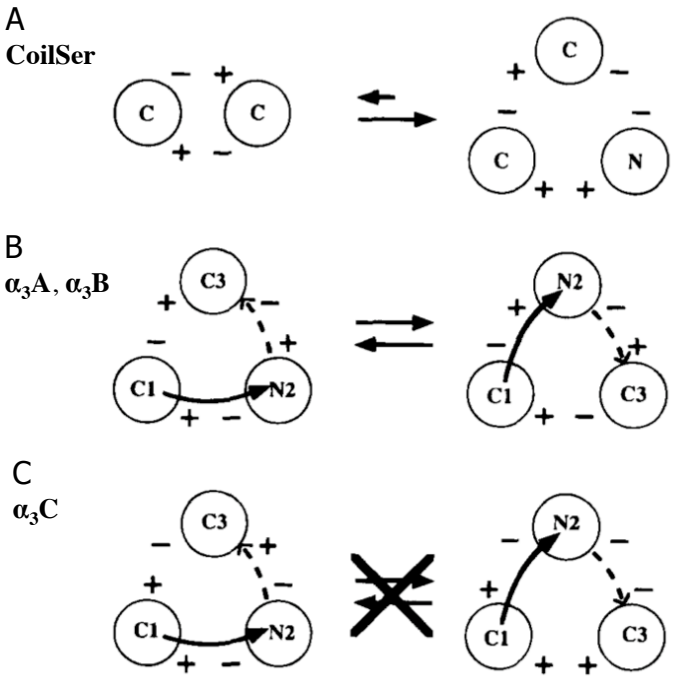



**Figure I-4**. A) Carton scheme representing interhelical electrostatic interactions between α-helices.[43] The circles symbolize α-helices and illustrated from either the N- or C- terminal end. The numbers correspond to their sequential positions in the THB. The "–" signs denote negatively charged Glu residues in either the "e" or "g" positions, while "+" signs indicate positively charged Lys and Arg residues at the same positions. Solid and dashed lines represent loops that connect helices. a) CS was designed to form a parallel dimer, instead it was observed to pack into an antiparallel trimer with unfavorable interactions with like charges. B) In $\alpha_3$A and $\alpha_3$B, the arrangement of the interhelical electrostatic interactions allow for both topologies to be possible. C) The clockwise form was destabilized by careful placement of charged residues at the "e" and "g" positions. In helix 1, only positively (+) residues were placed both at the "e" and "g" positions; while in helix 3, negatively (-) charged residues were assigned in those corresponding sites. For helix 2, the "e" sites were given only - charged residues, whereas the "g" sites received + charged residues. Reprinted with permission from ref 43. Copyright 1998 The Protein Society.

## 1.3. Solution structure of $\alpha_3$D, the final iteration in the THB design

DeGrado and coworkers work had a significant impact on the field of *de novo* protein design through the design, preparation and characterization of $\alpha_3$D, a 73-residue peptide with a single conformation in solution and a unique native-like fold.[10] Its well-packed core and single topology led to a solution structure, at the time this was a very challenging feat to achieve (Figure I-5). In contrast to its predecessors ($\alpha_3$A, $\alpha_3$B and $\alpha_3$C), which were chemically synthesized, Walsh *et. al*. expressed $\alpha_3$D in *E. coli*. Met1, Gly2, Gln9, Thr16, and Asp65 were changed in $\alpha_3$C to generate the sequence of $\alpha_3$D (Table I-2). Thermodynamic studies (chemical and thermal denaturation) showed a scaffold that is fully folded at room temperature (pH 3 -7) with a melting

temperature ($T_m$) in the range of 80 – 95 °C, heat capacity ($\Delta C_p$) of 10 - 12 cal $mol^{-1}$ $K^{-1}$ per residue, Gibbs free energy of unfolding ($\Delta G_U$) of 5.1 kcal $mol^{-1}$, and an enthalpy ($\Delta H_{DSC}$) value of -44 kcal $mol^{-1}$.

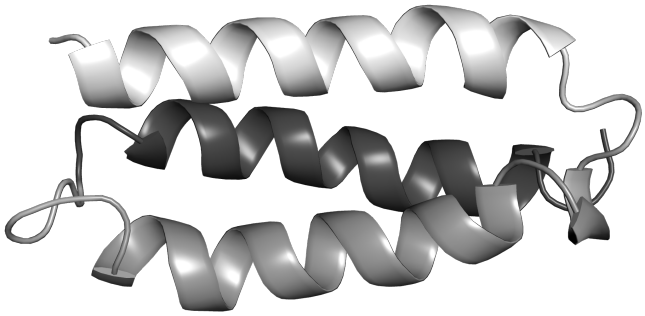


**Figure I-5**. Solution structure of $\alpha_3$D (PDB code 2A3D) demonstrating a THB fold.[10]

The solution structure of $\alpha_3$D was obtained from several three-dimensional (triple resonance, TOCSY and NOESY) NMR experiments at pH 5.5. About 1260 experimental restraints were used in solving the structure, which comprises of 1191 distances and 69 dihedral angles ($\phi$ and $\chi_1$). The ensemble of 13 structures demonstrate a high quality structure with RMSD values of 1.06, 0.75 and 1.61 Å for the backbone atoms (residues 1-73, N, $C^\alpha$, C), backbone atoms in the structure regions (residues 4-21, 24-45, 51-70, N, $C^\alpha$, C) and heavy atoms (residues 1-73), respectively. The helical bundle adopts a counterclockwise topology, which is confirmed by the interhelical tilt angles of the lowest energy structure ($\Omega_{1,2}$ = -165°, $\Omega_{1,3}$ = 17° and $\Omega_{2,3}$ = -171°). Further, the $\chi_1$ torsional angles of 14 core residues assumed a single conformation in the ensemble demonstrating a well-packed apolar core. Ultimately, the success of $\alpha_3$D demonstrates that the *de novo* design strategy could now serve as a practical method in constructing complex and multistranded scaffolds from a single sequence.

Furthermore, additional solution studies revealed that the backbone $^{15}N$ and $^{13}C$ atoms are well-ordered with restrictive motion on the pico to nanosecond scale,[44] the relative hydration of the backbone amides[45] and the folding time scale of $\alpha_3$D to be in the 1 – 5 μs range.[46] In addition,

mutation studies showed that replacing Ala60 to a Leu or an Ile resulted in 1.5 kcal $mol^{-1}$ net gain in stability.[47]

## 2. Functionalizing the $\alpha_3$D framework

### 2.1. Construction and structure of a symmetric heavy metal binding peptide

$\alpha_3$D offers a novel opportunity to add function to a well-defined *de novo* designed scaffold (Figure I-6A). The sequence of $\alpha_3$D was redesigned by introducing a tris(cysteine) motif to emulate the type of $MS_3$ environments that have been proposed for the metalloregulatory proteins MerR,[48-50] ArsR/SmtB[51] and CadC/CmtR[51-53] (an $MS_4$ or $MS_3O$ environment). At the C-terminal end of the bundle, three Leu residues at positions 18, 28 and 67 are inside a "hydrobic box" which is formed by Ile14, Leu21, Phe31, Ile63 and Tyr70. Chakraborty *et. al.* functionalized $\alpha_3$D by mutating the "a" site Leu residues to Cys (Leu18Cys, Leu28Cys and Leu67Cys) to produce $\alpha_3$D**IV** (Figure I-6B).[11] This forms a metal binding site with two "a" $S_\gamma$ ligands and one pseudo "a" site at the 28th position in the antiparallel strand.

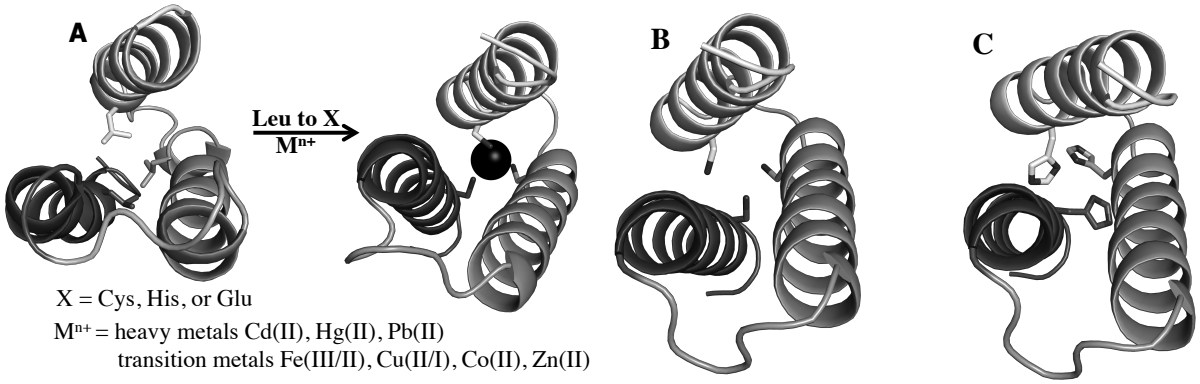



**Figure I-6**. A) Schematic representation of designing a metal center into the $\alpha_3$D scaffold. B) Top-down view of the tris(cysteine) site in $\alpha_3$D**IV** modeled from the $\alpha_3$D structure. C) Top-down view of the tris(histidine) site in $\alpha_3$D$\mathbf{H_3}$ modeled from the $\alpha_3$D structure.

### 2.2. Characterizing heavy metal binding properties of $\alpha_3$DIV

Chakraborty *et. al.* determined that apo $\alpha_3$**DIV** is well-folded in solution between pH 6-9 via circular dichroism studies and has a chemically-induced $\Delta G_U$ of 2.5 kcal $mol^{-1}$.[11] This $\Delta G_U$ is

half of the reported value for $\alpha_3$D, exhibiting a lost in stability after removing packing Leu residues. $\alpha_3$**DIV** stoichiometrically binds Hg(II), Pb(II) and Cd(II) in a pH dependent manner. The tris(cysteine) site forms a linear $[Hg(II)S_2(SH)]$ below pH 6.0 and a trigonal $[Hg(II)S_3]^-$ complex above pH 8.5; and a mixture of both species were observed under intermediate pH conditions (~pH 7.5). Above pH 5.0, Pb(II) and Cd(II) bind $\alpha_3$D**IV** to generate a trigonal pyramidal $[Pb(II)S_3]^-$ and pseudotetrahedral $[Cd(II)S_3(N/O)]^-$ geometry, respectively. These coordination modes were determined using various spectroscopic methods and compared to the physical properties of 3SCC analogues (Table I-3). The absorption features of all three metallated species were characterized via UV/Vis spectroscopy. $^{113}$Cd and $^{199}$Hg NMR and $^{111m}$Cd and $^{199m}$Hg Perturbed Angular Correlation Spectroscopy (PAC) were obtained for Cd(II)– and Hg(II)–$\alpha_3$D**IV** to confirm their binding modes in solution. These NMR[54, 55] techniques allow us to study the coordination environment at the millisecond timescale, while the PAC[56] techniques can further confirm and elucidate speciation behavior at the micro to nanosecond timescale. The chemical shift environment of $^{113}$Cd– and $^{199}$Hg–NMR are especially sensitive to the coordination environment. The combination of metal-NMR and PAC provides a powerful tool in identifying the primary ligand environment, as well additional coordinating ligands, such as solvent molecules or residues that are several layers removed. Furthermore, the development of $^{207}$Pb-NMR[34, 40] was significantly advanced using the TRI and CS peptides by proving how extremely sensitive this nucleus is to subtle changes in the apolar layer above the $Pb(II)S_3$ complex, with a chemical shift range of 5800 – 5500 ppm.

**Table I-3**. Physical parameters of metallated $\alpha_3$D**IV** compared to 3SCC constructs.

| Complex | λ [nm] ($\Delta\varepsilon$ [$M^{-1}$ $cm^{-1}$]) | δ (ppm) $^{113}$Cd | $^{199}$Hg | $^{207}$Pb | $\omega_0$ (rad/ns), η $^{111m}$Cd PAC | $\nu_Q$(GHz), η $^{199m}$Hg PAC |
|---|---|---|---|---|---|---|
| | | 583 | | | 0.350(6), 0.00(1) | |
| Cd($\alpha_3$D**IV**) | 232 (18 200) | 595 | | | 0.268(4), 0.18(7) | |
| | | | | | 0.170(2), 0.50(2) | |
| Hg($\alpha_3$D**IV**) | | | | | | |
| 3-coordinate | 247 (12 500), 265 (8400), 295 (3900), 240 (850) | | –244 | | | 1.11(2), 0.40(3) |
| 2-coordinate | 240 (850) | | –938 | | | 1.48(2), 0.15(5) |
| Pb($\alpha_3$D**IV**) | 236 (18 000), 260 (14 400), 278 (9100), 346 (3150) | | | | | |
| Cd(TRIL12C)$_3$ | 231 (20 600) | 619 | | | 0.233(8), 0.25(12) | |
| | | | | | 0.468(9), 0.12(10) | |
| Cd(TRIL16C)$_3$ | 232 (22 600) | 625 | | | 0.337(2), 0.23(2) | |
| | | | | | 0.438(4), 0.20(3) | |
| Hg(TRIL9C)$_3$ | | | –185 | | 1.164(5), 0.25(2) | |
| Hg(TRIL9C)$_2$ | | | –908 | | 1.558(7), 0.23(1) | |
| Hg(TRIL12C)$_3$ | 230 (21 300) | | –316 | | | |
| Hg(TRIL16C)$_3$ | 247 (16 800), 265 (10 600), 295 (5000) | | –179 | | | |
| Hg(TRIL16C)$_2$ | 240 (2700) | | –834 | | | |
| Pb(TRIL12C)$_3$ | 238 (17 000), 278 (12 300), 343 (3700) | | | 5814[a] | | |
| Pb(TRIL16C)$_3$ | 236 (18 500), 260 (16 500), 278 (14 500), 346 (3400) | | | 5612[b] | | |

UV/Vis, metal ($^{113}$Cd, $^{199}$Hg and $^{207}$Pb) NMR and ($^{111m}$Cd and $^{199m}$Hg). L9C and L16C are "a" site constructs, while L12C are "d" sites. [a] $^{207}$Pb chemical shift for Pb(CSL12C)$_3$. [b] $^{207}$Pb chemical shift for Pb(CSL16C)$_3$.

The ligand-to-metal charge transfer (LMCT) bands of metallated $\alpha_3$D**IV**[11] in the UV/Vis studies exhibited metal-thiolate transitions that are comparable to the reported values for its TRI counterparts (Table I-3). The Cd(II)–$\alpha_3$D**IV** species has a $\lambda_{max}$ at 232 nm (18,200 $M^{-1}$ $cm^{-1}$) and Pb(II)–$\alpha_3$D**IV** exhibits four absorption bands with a $\lambda_{max}$ at 236 nm (18,000 $M^{-1}$ $cm^{-1}$) and a characteristic band at 346 nm (3150 $M^{-1}$ $cm^{-1}$). The trigonal complex that forms in Hg(II)-$\alpha_3$D**IV** demonstrates three bands with a $\lambda_{max}$ at 247 nm (12,500 $M^{-1}$ $cm^{-1}$), while the linear species contains one $\lambda_{max}$ at 240 nm (850 $M^{-1}$ $cm^{-1}$). From further UV/Vis work, the $[Pb(II)S_3]^-$ complex was determined to have a lower limit binding constant of 2.0 x $10^7$ $M^{-1}$, while the $[Cd(II)S_3(N/O)]^-$ complex has a lower limit value of 3.1 x $10^7$ $M^{-1}$.

The linear complex of $^{199}$Hg(II)-$\alpha_3$D**IV** has a chemical shift of –938 ppm, whereas the trigonal form experiences a downfield shift to –244 ppm. The $^{199}$Hg-NMR spectrum at an intermediate

pH (7.5) contains both the linear and trigonal planar species. The $^{199m}$Hg–$\alpha_3$D**IV** spectra from the PAC analysis confirm this pH dependent speciation behavior (Figure I-7). The PAC parameter $\nu_Q$ was determined to be 1.48(2) and 1.11(2) GHz for the linear and trigonal complex, respectively, which matches well with the reported $\nu_Q$ values for 2 and 3 coordinate Hg(II) in 3SCC constructs.[25] These $\nu_Q$ values were also observed at pH 7.5, demonstrating the presence of both linear and trigonal planar species. In hindsight, these results show a peptide scaffold that can finely control the coordination environment of Hg(II) ions. Additionally, Chakraborty *et. al.* observed that the addition of 2 equivalents of Hg(II) induced the formation of a dimer species, where Hg(II) atoms bridge two peptides through a linear complex.[57] This dimer species was confirmed with both $^{199}$Hg(II)–NMR and $^{199m}$Hg(II)–PAC.

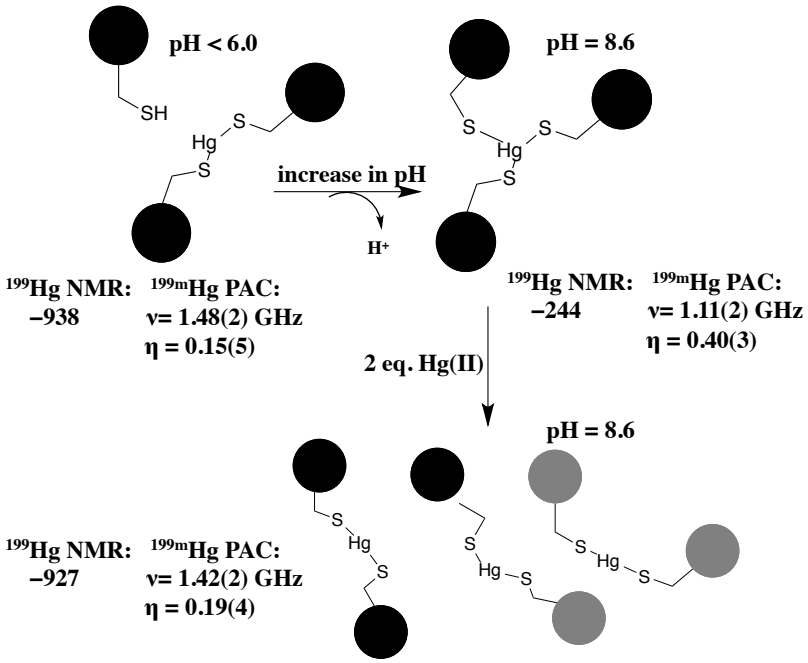



**Figure I-7**. Schematic representation of the pH and stoichiometric dependent behavior of Hg(II) species in $\alpha_3$D**IV**.[11, 57]

The $^{113}$Cd-NMR spectrum of $^{113}$Cd-$\alpha_3$D**IV** shows overlapping resonance peaks at 583 and 595 ppm. Based on these chemical shift positions, this result indicates the presence of two 4 coordinate Cd(II) species. $^{111m}$Cd-PAC was again collected to supplement the $^{113}$Cd-NMR result, but in this case, it was used to define the two species observed in the NMR time scale. The

$^{111m}$Cd-PAC showed three nuclear quadrupole interactions at 0.35, 0.27 and 0.17 rad ns$^{-1}$. The peaks at 0.35 and 0.27 rad ns$^{-1}$ agree well with a $CdS_3O$ complex with two conformations, *endo* and *exo*. The frequency value at 0.17 rad ns$^{-1}$ fitted well with $CdS_3N$ species, where the N ligand was determined to originate from the imidazole ring of His72. Overall, not only did the combination of metal–NMR and PAC provide a way to accurately characterize our *de novo* designed peptides, it also gives powerful insights into how toxic heavy metals may interact with native proteins.

2.3. Constructing a symmetric metalloenzyme site in $\alpha_3$D

Our work with $\alpha_3$D was further expanded to incorporate a tris(histidine) metal binding site reminiscent of carbonic anhydrase (Figure I-6C). This metalloenzyme plays a vital role in respiration, vision, cancer metathesis, regulation of acid-base equilibria and other processes in animals, plants and bacteria. Human carbonic anhydrase II is one the most efficient enzymes (approaching the diffusion limit) catalyzing the reversible interconversion between $CO_2$ and $HCO_3^-$.[58] Even though the mechanism, structure and inhibition have been previously studied, *de novo* protein design still offers a novel approach to study and replicate an important function of a native metalloenzyme in a simplified peptide system. A carbonic anhydrase (CA) model was previously demonstrated in a bimetallic 3SCC construct $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(TRIL9CL23H)_3^{n+}$ and Zastrow *et. al.* reported this model to be within 500-fold of the fastest isozyme (CAII), which is the fastest CA-model to date (Figure I-2A).[29, 31] Nevertheless, CAII contains residues that participate in hydrogen bonding networks and the self-associating nature of our 3SCC constructs limits its use in preparing asymmetric sites. Thus, the zinc catalytic site of CA was modeled into $\alpha_3$D to recapitulate CA activity. Now that this was established, future work will focus on adding hydrogen bonding residues in the second coordination sphere of the Zn(II) complex in CA designs of $\alpha_3$D.

Cangelosi *et. al.* incorporated a tris(histidine) site in $\alpha_3$D to yield $\alpha_3$D**H$_3$** (Figure I-6C), a *de novo* designed metalloenzyme model that exhibited CA activity.[12] The sequence of $\alpha_3$D**H$_3$** (Table I-2) was expanded by four residues (77 in $\alpha_3$D**H$_3$**), which led to a peptide with increased yields during expression from ~100 to 230 mg/L. His residues were substituted at positions 18, 28 and

67, while a His72Val mutation to eliminate a competing ligand was completed. At pH 9.0, the apo form folds well in solution (82% folded) according to the 208 and 222 nm bands that were observed in the CD spectrum and has a chemically induced $\Delta G_U$ of 3.1 kcal $mol^{-1}$. The Zn(II)-$\alpha_3D\mathbf{H_3}$ complex was characterized using UV/Vis and X-ray absorption spectroscopies, while its CA activity was determined with Khalifah's stopped-flow indicator technique.[59]

Using a UV/Vis Zincon colorimetric assay,[60] the apparent Zn(II) binding constant to $\alpha_3D\mathbf{H_3}$ was determined to be 150 ± 40 nM at pH 7.5 and strengthens to 59 ± 9 nM at pH 9.0. When compared to CA, these affinity values are only two orders of magnitude weaker than the recent value determined for CAII (0.45 nM)[58] and stronger than the $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(TRIL9CL23H)_3^{n+}$ (0.8 ± 0.1 μM at pH 7.5 and 0.22 ± 0.06 μM at pH 9.0)[31] CA model. From extended X-ray absorption fine structure spectroscopy, the Zn(II)-$\alpha_3D\mathbf{H_3}$ coordination environment (at pH 9.0) fitted well to a site that contains 1 oxygen (from an exogenous $H_2O$ or OH molecule) and 3 nitrogens from each His residue bound to a Zn(II) atom at 1.90 and 1.99 Å, which matches well with the Zn complex in CAII with Zn-N/O of 1.98 Å (pH 7.0).[58]

To demonstrate the success of $\alpha_3D\mathbf{H_3}$ as a metalloenzyme, Cangelosi *et. al.* performed a stopped-flow $CO_2$ hydration assay using Khalifah's indicator technique. The maximal catalytic efficiency ($k_{cat}/K_M$), which was derived from the $k_{cat}/K_M$ values for pH 8-9.5, and kinetic p$K_a$ for the deprotonation of Zn(II)-bound water to yield the active $[Zn(II)N_3O]^{1+}$ hydroxide complex were worked out to be 6.9 x $10^4$ L $mol^{-1}$ $s^{-1}$ and 9.4, respectively. When compared to two small molecule models Zn(II)(tris(4,5-di-*n*-propyl-2-imidazolyl)-phosphine)[61] and Zn(II)nitrilotris(2-benzimidazolylmethyl-6-sulfonate)[62] that both have a $Zn(II)N_3O$ complex and show CA activity, Zn(II)-$\alpha_3D\mathbf{H_3}$ significantly outperformed these models exhibiting a second-order rate constant ($k_2$) that is 14-fold higher. The $CO_2$ hydration efficiency of Zn(II)-$\alpha_3D\mathbf{H_3}$ is 2.6 fold slower than its 3SCC counterpart,[29] 1400-fold less efficient than CAII[63] but only 11-fold slower than CAIII.[64] The decrease in the catalytic activity as compared to our 3SCC CA model could be as a result of a weaker dipole, less-symmetric environment for the imidazole rings and difference in the electrostatics at the metal binding site in the antiparallel bundle of $\alpha_3D\mathbf{H_3}$. Furthermore, a

product inhibition assay was performed on Zn(II)-$\alpha_3$D**H$_3$** using acetate since it serves a more probable mimic of bicarbonate. At pH 8.5, the first-order rate constant ($k_{cat}$) experienced a modest decrease from 82 ± 6 to 66 ± 4 $s^{-1}$. This inhibition result indicates no significant loss in catalytic activity and illustrates a CA model capable of preventing product inhibition, which is a major problem in small molecule models of enzymes. Overall, Zn(II)-$\alpha_3$D**H$_3$** exhibits a *de novo* designed model that is successful in recapitulating the primary active site and the function of carbonic anhydrase in a simplified antiparallel THB scaffold, a metalloenzyme that is found in a twisted β-sheet fold in nature.

## 2.4. Introduction to the design of a copper electron transfer site in $\alpha_3$D

Electron transfer (ET) is the simplest chemical transformation, involving the transfer of electrons from one molecule to another.[8, 65-78] This reaction is an essential step in numerous biological processes: from collagen synthesis, the immune response and the nitrogen cycle to two fundamental bioenergetics processes that sustain life on earth, aerobic respiration and photosynthesis.[79-82] The latter two processes are complementary, that is photosynthetic organisms release oxygen ($O_2$) gas that is then consumed by aerobic bacteria and animals. In turn, the final products of aerobic respiration ($CO_2$ and $H_2O$) are used by photosynthetic organisms. This continues the global cycle of vital elements C, H and O. Furthermore, electron transfer centers are central components in multifaceted electron transport chains responsible for energy production and serve as an intramolecular electron source in enzymatic reactions. The reactions involve the chemical transformation of a substrate bound to the catalytic site, often a metalloprotein (e.g., biosynthesis of DNA catalyzed by ribonucleotide reductase and monoxygenase reactions in drug activation by cytochrome P-450s).

Cupredoxins or Type 1 copper proteins (CuT1) are biological electron transfer agents isolated from *Bacteria, Archaea and Eukarya*. These copper centers were discovered to function in bioenergetic electron transport chains (e.g. photosynthesis) as a freely diffusible electron carrier (Figure I-8A) or membrane-bound via a N-terminal lipid tether or C-terminal glycosylphosphatidylinositol (GPI) anchor.[83] Further, they take part in the chemical reactions of larger proteins (Figure I-8B), where the copper center provides an electron sink to the catalytic

site (e.g. copper nitrite reductase and laccase).[84-86] CuT1 proteins have been extensively studied to understand their fundamental function in nature.[87] The copper ion is coordinated to a well-defined and pre-organized asymmetric metal binding site that is encapsulated by a β-barrel fold. It contains two histidines, one cysteine and one or two more weakly bound residues to form a 3-5 coordinate copper complex. The preassembled $\alpha_3$D offers a viable framework in achieving an asymmetric cupredoxin site. Now knowing that cysteine ligands in $\alpha_3$D**IV** and bulky His residues in $\alpha_3$D$\mathbf{H_3}$ can be accommodated by this framework, I designed several CuT1 constructs to investigate if the properties and function of an electron transfer site that is naturally found in β-barrel fold can be achieved in an unrelated α-helical fold. Ultimately, these designed constructs will provide a novel approach for investigating biological long-range electron transfer. This work will be fully described in Chapter 3 and 4 of this thesis.

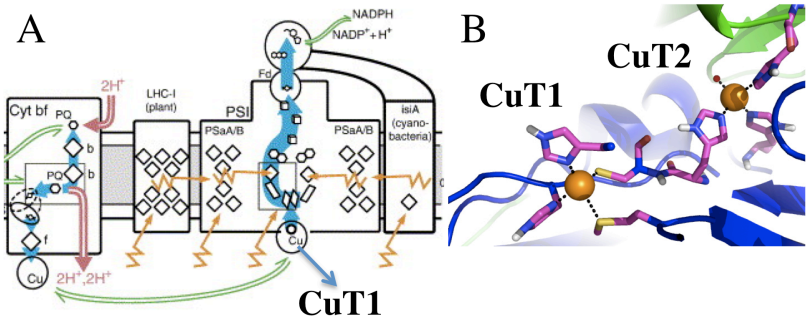



**Figure I-8** Electron transfer role of copper centers in biological reactions. A) CuT1 protein relays electron in in photosynthesis.[78] Reprinted with permission from Ref 78. Copyright 2005 Elsevier B.V.B) CuT1 provides an electron to a CuT2 catalytic site that reduces nitrite in copper nitrite reductase.

3. Cupredoxin proteins

The CuT1 superfamily can be subdivided into three classes. Class 1 is comprised of a single CuT1 copper-binding domain in the mature state and is known as cupredoxins (which is analogous to iron ferredoxins). This class includes plastocyanin,[14, 88-92] azurin,[93-95] pseudoazurin[96, 97] and rusticyanin.[98-100] Cupredoxins are soluble proteins or bound to the cell surface. Plastocyanin[89] and azurin[93] were the first two to be structurally characterized through X-ray crystallography in 1978, setting the stage for further research that characterized the structure-function relationships of these proteins. Further, the azurin gene was the first to be cloned[101] and became the platform for many mutagenesis studies that aim to understand the coordination

environment of CuT1 proteins and electron transfer further.[102-104] The second class is composed of multidomain blue copper proteins such as nitrite reductase,[85] laccase,[86] ascorbate oxidase,[105] and ceruloplasmin.[106] These copper proteins also referred to as "blue copper oxidases" because CuT1 centers, which provide electrons to copper catalytic sites that react with dioxygen, dominate the spectroscopic feature of these proteins. Finally, the third class, the phytocyanins, shows a chimeric domain organization, where the blue copper domain is attached to structurally unrelated sequence motifs. This plant derived CuT1 class is bound to the cytoplasmic membrane through a GPI-anchor and includes stellacyanin,[107-111] umecyanin[110, 112] and plantacyanin.[110, 111, 113] Stellacyanin was one of the first of the phytocyanins to be characterized using biochemical and biophysical studies and was determined to possess spectroscopic and structural features that uniquely differ from plastocyanin and azurin. By and large, CuT1 proteins function as ET proteins but their redox partners, with the exception of photosynthetic plastocyanin, are unclear (Table I-4).

**Table I-4**. Copper electron transfer proteins.

| Protein | Family | Source | Redox Partners |
|---|---|---|---|
| Plastocyanin[114, 115] | Class 1: Cupredoxins | Plants | Cytochrome *f* & Cytochrome *b6/f* complex |
| Azurin[116, 117] | | Bacteria | C-type Cytochrome *c* and Nitrite Reductase[a] |
| Pseudoazurin[97] | | Bacteria | Blue copper site of nitrite reductase |
| Rusticyanin[98, 100, 118] | | Bacteria | Iron proteins[a] |
| Nitrosocyanin[119, 120] | | Bacteria | Ammonia-oxidizing proteins[a] |
| Stellacyanin, Uclacyanin, Plantacyanin[110, 111] | Class 3: Phytocyanins | Plants | Unknown |

[a]Proposed redox partners.

3.1. Structure of native cupredoxin

Overall, CuT1 proteins are single polypeptide chains that are composed of 91-155 residues. This single chain folds into a Greek $\beta$-barrel with eight parallel or antiparallel $\beta$-strands that divides into two $\beta$-sheets (Figure I-9A). CuT1 proteins that contain more residues posses one or two more $\beta$-strand such as amicyanin and rusticyanin, respectively. The $\beta$-barrel fold is ubiquitous in nature. It is adopted and modified by native metalloproteins, including copper-zinc superoxide dismutase, lacacase, beta-amylase and immunoglobin. The copper center (Figure I-

9B) is encompassed within the hydrophobic environment of the $\beta$-barrel fold, and the metal binding ligands are in a pocket between loops, protected from solvent access. The copper center is located at the "northern," C-terminal end of the $\beta$-barrel. The amino acid sequence, the number of strands, as well as the make up of the loops and the hydrogen-bonding network surrounding the copper center vary among the CuT1 proteins. However, structural interactions provided by the loop elements and the hydrogen-bonding network are conserved across the family, including the metal binding residues. Furthermore, in many CuT1 proteins, particularly in plastocyanin, the "northern" end of the fold includes a hydrophobic patch (Figure I-9C) that was proposed to serve as a recognition site for redox partners.[114, 115, 121-127] This patch consists of 5-10 apolar residues that surround one of the His ligands bound to the copper ion. Figure I-10 displays the X-ray crystal structures of several CuT1 proteins, including plastocyanin,[14] azurin,[128] stellacyanin,[109] cucumber basic protein[113] and nitrosocyanin,[119] demonstrating the shared Greek $\beta$-barrel fold and a variety in the length of the sequence.

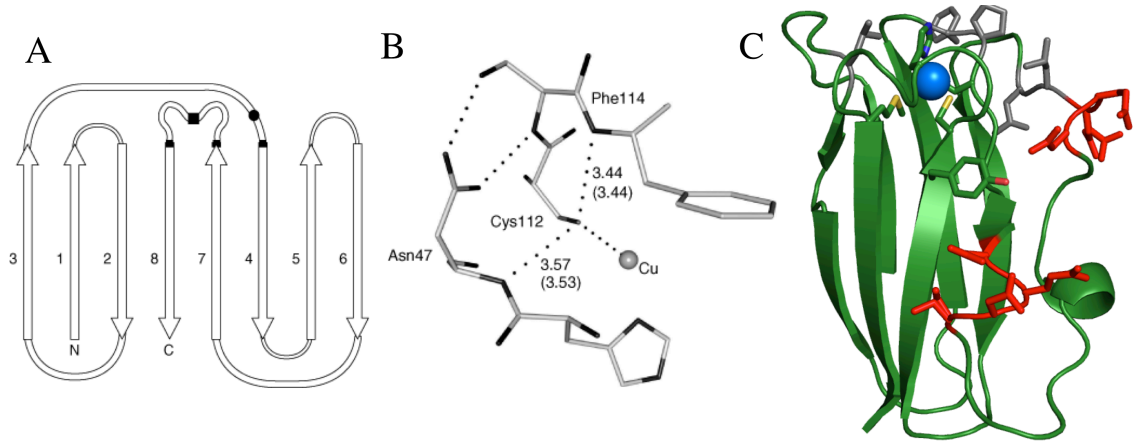



**Figure I-9.** Structure of cupredoxin proteins. A) Topology illustration that show the connections of the eight extended polypeptide strands in the β-barrel fold of CuT1 proteins. The filled squares demonstrate the locations of the common ligands and the filled circle signifies the carbonyl oxygen found in azurin.[83] B) The constraint applied on the Cys residue in azurin by the protein fold via hydrogen bonding residues from neighboring residues or residues located on a nearby loop. The distances are for the oxidized form, while the distances within the parentheses are for the reduced state. [83] The minimal change in these values exhibits the rigidity of the Cys ligand. Figure A and B were reprinted with permission from Ref 83. Copyright 2006 John Wiley & Sons, Ltd. C) The structure of *Poplar nigra* plastocyanin (PDB 1PLC at 1.33 Å resolution),[14] where the metal center, hydrophobic patch (residues in gray) and acidic patch (residues in red) are highlighted.

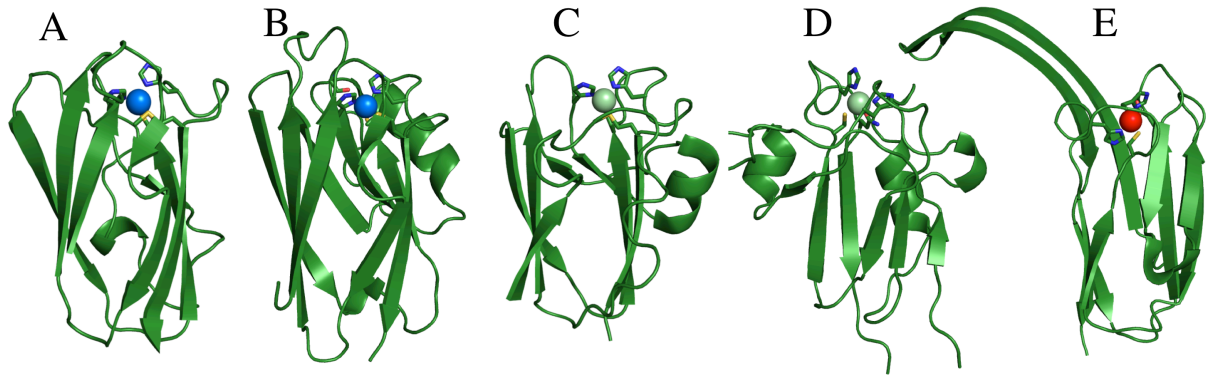



**Figure I-10.** Varying β-barrel fold and size of CuT1 proteins, where the color of the copper ion signify their spectroscopic feature in the Cu(II) state. A) *Poplar nigra* plastocyanin (PDB 1PL at 1.33 Å resolution).[14] B) *Alcaligenes denitrificans* azurin (PDB 2AZA at 1.80 Å resolution).[128] C) Cucumber stellacyanin (PDB 1JER at 1.60 Å resolution).[113] D) Cucumber basic protein (PDB 2CBP at 1.80 Å resolution). E) Nitrosocyanin (1 of 3) monomer from *Nitrosomonas europea* (PDB 1IBY at 1.65 Å resolution).[119]

The copper metal ion centers in CuT1 proteins are coordinated to three equatorial ligands: to two imidazole nitrogens (N) of histidine (His) and one thiolate sulfur (S) of cysteine (Cys) to form a pseudo-trigonal plane with a $CysHis_2$ geometry that is conserved across the CuT1 family.[66, 83, 87, 129, 130] The positions of these ligands vary between proteins; for example, the copper metal binding site in poplar plastocyanin is composed of His 37, Cys 84 and His 87 (Figure I-11A), whereas azurin contains a His 46, Cys 112 and His 117 in the metal center. From the crystal structure of the oxidized state ($Cu^{2+}$), the copper center has average Cu-N bond lengths at ~2.0 Å, a typical distance observed for imidazole ligands. However, the Cu-S(Cys) bond is remarkably short with an average distance range at ~2.07 – ~2.26 Å, which is usually between 2.3 – 2.9 Å. In addition to the core $CysHis_2$ ligands, one or two longer axial ligands can be bound to generate a four or five coordinate geometry. With the exception of some plantacyanins (which only has $CysHis_2$ ligands), the most common axial ligand is a thioether sulfur from a Met residue to yield an $N_2SS^*$ ligand set that forms a distorted tetrahedral geometry. The $N_2SS^*$ ligand set is found in plastocyanin, cucumber basic protein, amicyanin, auracyanin, as well as rusticyanin. This Cu-S (thioether) bond has a length at ~2.8 Å and displaces the copper atom from the trigonal plane by about 0.2 – 0.4 Å. Azurins also contain the same $N_2SS^*$ ligands but the backbone carbonyl oxygen of Gly 45 provides a second axial ligand, forming a five-coordinate, distorted trigonal bipyramidal copper center. An equivalent Gly

residue is also present but it is not considered to be a ligand because it is ~3.8 Å from the copper complex. The Cu-S (Met) and Cu-O (Gly) have bond lengths at ~3.3 and ~2.6 Å, respectively. Unlike its counterparts, the stellacyanin family (Figure I-11B) and nitrosocyanin (Figure I-11C) do not posses an S(Met) axial ligand and are examples of "perturbed" blue copper centers.[131, 132] Stellacyanin has an O(Gln 99) axial ligand with an ~ 2.4 Å bond length, which is much shorter (~0.5 Å) than the Cu-S(Met) bond length of plastocyanin and azurin and is ~0.2 Å shorter than the Cu-O (Gly) reported above. Nitrosocyanin is the most perturbed blue copper site, possessing an even stronger axial bond at ~2.1 Å for coordination to O(Glu 60).[119] As a result of a shorter axial bond, the Cu-S(Cys) bond experiences a modest increase to ~2.3 Å from 2.1 – 2.2 Å.

The metal binding site of CuT1 proteins are provided by two loops that connect $\beta$-strands 3 & 4 and 7 & 8. The first His ligand (and the carbonyl oxygen of a Gly) is the terminal residue of the loop that joins $\beta$-strands 3 & 4. Further, the 3 & 4 loop is longer in azurin and rusticyanin than in plastocyanin and pseudoazurin and even shorter in plantacyanin and stellacyanin. As a result, azurin and rusticyanin have the most buried copper center; whereas, the proteins with shorter loops are the most exposed. The final three ligands are found on the loop between $\beta$-strands 7 & 8. The Cys residue is the final residue of strand 7; the second His is at the center of the loop; and the Met (or Gln) is the leading residue in strand 8. In addition to containing 3 of the 4 ligands, the loop between $\beta$-strands 7 & 8 includes apolar residues that make up the hydrophobic patch around the second His ligand. Moreover, the number of residues between these last three residues, Cys and His and His and Met, differs from two to four. In nitrosocyanin, the copper ligands are located in the loops that connect $\beta$-strands 4 & 5 and 8 & 9, which contains a Glu (60) residue and ligands Cys (95), His (98) and His (103), respectively. Further, nitrosocyanin show a characteristic cupredoxin, but it has an unusual $\beta$-hairpin structure that is prolonged, but does not take part in the $\beta$-barrel fold (Figure I-10E).

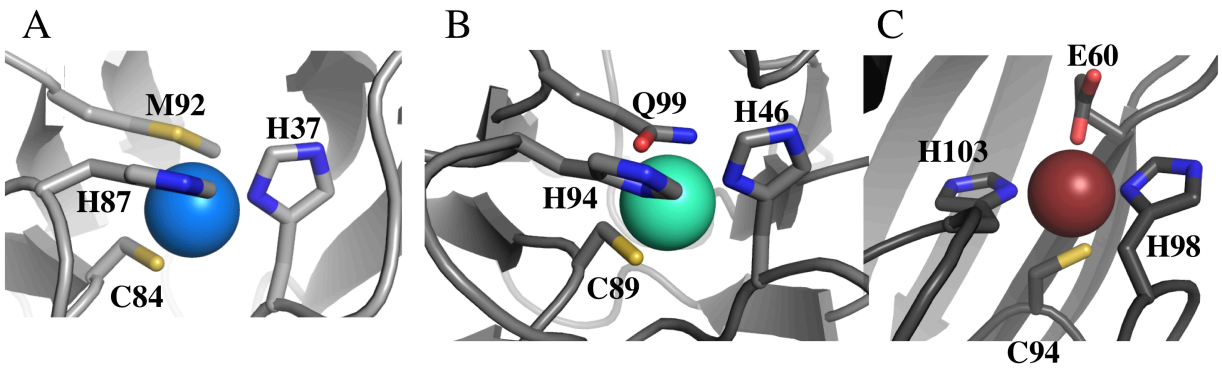



**Figure I-11.** Metal binding site of plastocyanin[14] (A), stellacyanin[109] (B) and nitrosocyanin[119] (C). Plastocyanin and CuT1 centers with a similar coordination environment have a trigonal pyramidal geometry where the copper ion and the core 2HisCys residues make the base of the pyramid and the Met is the apex. Stellacyanin and its phytocyanins family with a Gln at the axial position display a distorted tetrahedral geometry, where the bond between the Cu and the axial O(Gln) ligand shortens while the Cu-S(Cys) bond elongates. A more recent addition to the CuT1 family is nitrosocyanin, which was determined to complex a copper center in a square pyramidal geometry. The metal binding site of nitrosocyanin lacks S(Met) ligand but instead is comprised of a O(Glu) ligand in the axial position and a 2HisCys ligand and a $H_2O$ molecule that make up the base of the square.

The copper coordination environment in CuT1 proteins is extremely rigid, and this is attributed to the restraint that the protein fold applies to the metal bind site. In azurin, cucumber stellacyanin and rusticyanin, the Cys residue is (N-H-S) hydrogen bonded to the backbone NH group of a residue neighboring the first His in loop 3 & 4 and to a NH group of a residue in loop 7 & 8 two residues removed towards the C-terminal end (Figure I-9B). The latter interaction is absent in plastocyanin, pseudoazurin and amicyanin because a Pro residue resides at the appropriate position and only has one N-H-S interaction. Further, the Cu-S(Cys) bond is highly covalent, which produces a short bond. This feature is highly conserved among CuT1 proteins. Similarly, the His and Met residues, as well as the second coordination sphere are also finely controlled by the protein fold. The N$\delta$1 of the first His ligand is directly bound to the copper ion, while the N$\varepsilon$2 in the imidazole ring forms a hydrogen bond to a backbone carbonyl oxygen or to the oxygen sidechain group of a Glu (amicyanin) or Asn (rusticyanin) residue. In phytocyanins, the N$\varepsilon$2 commonly forms a hydrogen bond with an exogenous water molecule. The second His ligand does not have hydrogen bond interactions with other residues, instead it is packed in place by the hydrophobic patch and establishes a hydrogen bond with a water molecule that is, in turn, hydrogen bonded to surrounding protein atoms. Likewise, the Met (or Gln) ligand, as well as the Gly carbonyl oxygen in azurins is packed in place by aromatic residues (Phe, Tyr or Trp).

Finally, conserved residues outside the metal binding site have been shown to play a crucial role in the stability of the copper complex. A conserved Asn residue (Asn 38 in plastocyanin and Asn 47 in azurin but Ser in rusticyanin) that follows the first His ligand forms multiple hydrogen bonds to the residue, typically a conserved Ser or Thr, after the Cys ligand. These hydrogen bonds create a bridge between the copper-binding loops.

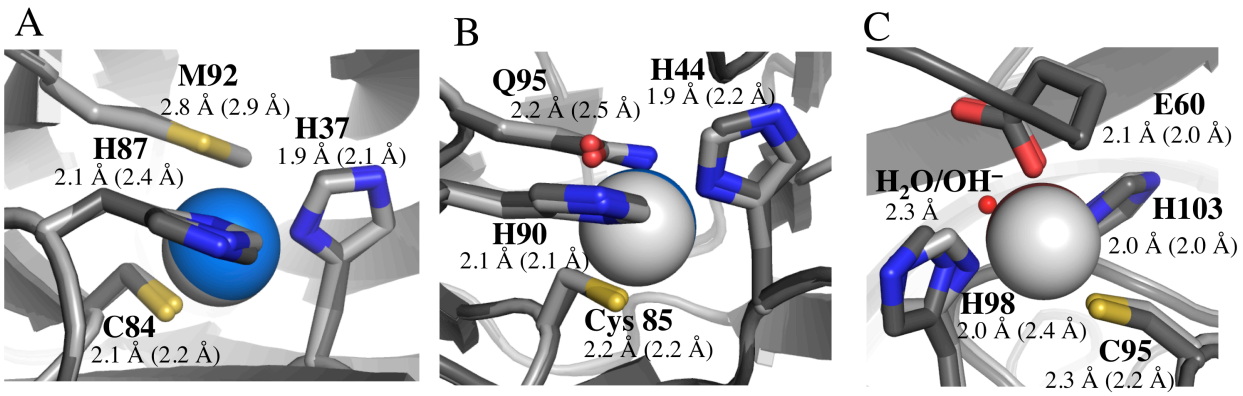



**Figure I-12.** Overlay of the oxidized (light gray) and reduced (dark gray) CuT1 centers, which illustrates minimal change in the bond lengths and geometry. The bond lengths are adjacent to their corresponding residue and values inside parentheses belong to the reduced state. A) Plastocyanin from *Poplar nigra* [PDB 1PLC (oxidized) at pH 6.0 & 5PCY (reduced) at pH 7.0 and 1.80 Å resolution].[14, 133] B) Horse radish umecyanin, which is part of the stellacyanin family [PDB 1X9U (oxidized) at pH 7.5 and 1.90 Å resolution & 1X9R (reduced) at pH 7.5 and 1.90 Å resolution].[112] C) Nitrosocyanin from *Nitrosomonas europea* [PDB 1IBY (oxidized) at pH 7.5 and 1.65 Å resolution & 1IBZ (reduced) at pH 7.5 and 2.3 Å resolution].[119]

The structural features of CuT1 proteins provide them with unusual spectroscopic and redox properties. This structural property is often described using the rack-induced bonding model,[103, 131, 134, 135] which suggests that the ligand environment in a β-barrel fold forces the copper ion, regardless of the oxidation state, into a strict geometry. The constraint applied on the copper complex is evident in the overlay of metal binding site of the oxidized and reduced states of plastocyanin,[14, 133] umecyanin[112] and nitrosocyanin[119] in Figure I-12. Upon reduction of the copper ion, a minor increase in ligand-metal bond lengths are observed, causing minimal change in the copper coordination environment during electron transfer that minimizes inner sphere reorganization energy. The latter example, which is a more recent addition to the CuT1 family, was observed to lack a water molecule during reduction. It was argued that this loss of a water

molecule implicates its role as a catalytic and not an ET center.[119, 120] Nevertheless, the copper ion in nitrosocyanin is also strongly constrained by the protein fold as shown in Figure I-12C.

### 3.2. Spectroscopic properties of cupredoxin centers

The oxidized state of the first purified CuT1 proteins displayed an extraordinarily intense blue color, a trait that intrigued and inspired inorganic chemists and biochemists to elucidate the mechanism behind this physical feature. The work on CuT1 proteins, as well as iron proteins (iron-sulfur and cytochromes) have helped develop and shape the field of bioinorganic chemistry. The CuT1 family now contains a spectrum of color from the classic blue copper proteins (plastocyanin and azurin) to green (stellacyanin), red copper (nitrosocyanin) and yellow (model compounds and designed proteins). The work of various groups, especially Solomon and coworkers, using various biophysical methods (UV-VIS, EPR, MCD, RR and XAS), as well as computational studies have provided great insight into the spectroscopic properties and structure-function relationship of the copper complex in CuT1 proteins.[87, 88, 107, 129, 131, 132, 136-157] Solomon and coworkers have led the characterization of the physical features as it relate to function of classic blue copper proteins[87, 140, 147, 148, 150, 153, 156-161] with respect to a copper chloride salt, $CuCl_4$,[87, 161, 162] that represents the features of CuT2 or "normal" copper complexes with a square planar or tetragonal geometry in solution. CuT1 sites with a coordination environment that deviates from the classic CuT1 geometry of a distorted trigonal planar with a weak axial ligand are termed perturbed blue copper sites.[120, 131, 132, 163, 164] These sites can be subdivided into proteins with a $CysHis_2Met$ ligand set but with a more tetragonal or square planar geometry (e.g., cucumber basic proteins, nitrite reductase, nitrosocyanin); a Gln ligand in the axial position instead of a Met (e.g., stellacyanin) and has been given a CuT1.5 term by Lu and coworkers[66]; and an apolar residue (Leu, Phe, Ile or Val) in place of a fourth axial ligand (e.g., plant and fungal laccases). These studies have been previously reviewed in detail.[87, 146, 157, 161, 165]

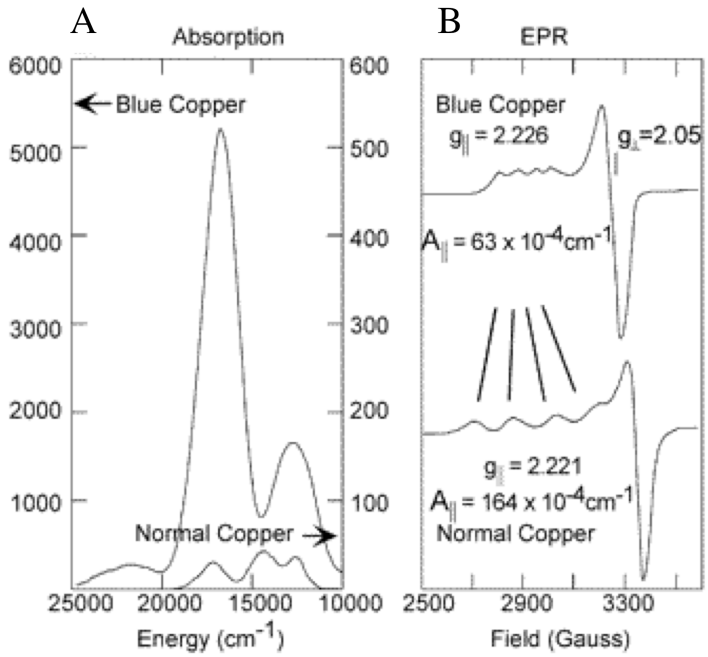



**Figure I-13.** UV-VIS (A) and EPR (B) spectra of blue copper centers compared to normal copper sites.[87] Blue copper proteins display a strong absorption band in the visible range, while normal copper centers show a strong feature in the UV range. Further, EPR spectra of blue copper sites have an $A_{||}$ value <100 x $10^{-4}$ $cm^{-1}$, whereas normal copper complexes often have values >100 x $10^{-4}$ $cm^{-1}$. Reprinted with permission from Ref 87. Copyright 2004 American Chemical Society.

The classic CuT1 proteins plastocyanin and azurin exhibit a deep blue color as a result of a strong electronic absorption band at ~600 nm (~3000 - 6000 $M^{-1}$ $cm^{-1}$) in a UV-VIS spectra (Figure I-13A). Blue copper proteins also contain a less intense band at ~450 nm (~300 $M^{-1}$ $cm^{-1}$) and yield an $R_\varepsilon$ value (the absorption ratio around ~400 and ~600 nm) < 0.1, a designation used by Lu and coworkers to demonstrate the perturbation from the $CysHis_2$ trigonal plane.[66] This strong absorption band was assigned to a ligand-to-metal charge transfer (LMCT) from the S(Cys)$\pi$ → Cu $3d_{x^2-y^2}$ charge transfer transition, which is made possible by a 45° rotation of the $d_{x2-y2}$ that increases the overlap with the $p_y$ orbital of S(Cys).[140, 158, 161] A weak band at ~530 nm was assigned to the S(Cys)$\sigma$ → Cu $3d_{x^2-y^2}$ transition.[87] "Normal" or CuT2 centers have an inverted assignment, exhibiting weak $\pi$ charge transfer (CT) but intense $\sigma$ CT transitions. Low-temperature magnetic circular dichroism (LT-MCD) studies were used to confirm these assignments as the absorption features in a LT-MCD spectrum have the opposite intensities in a UV-VIS spectrum. Further, bands at ~430 and 450 correspond to S(Met) → Cu and S(His) → Cu bands, respectively, and bands between 650 – 1050 nm are either ligand field (LF) or d → d

transitions. CuT1 centers tomato plantacyanin and laccases lack an axial ligand and display a similar electronic absorption as blue copper proteins.

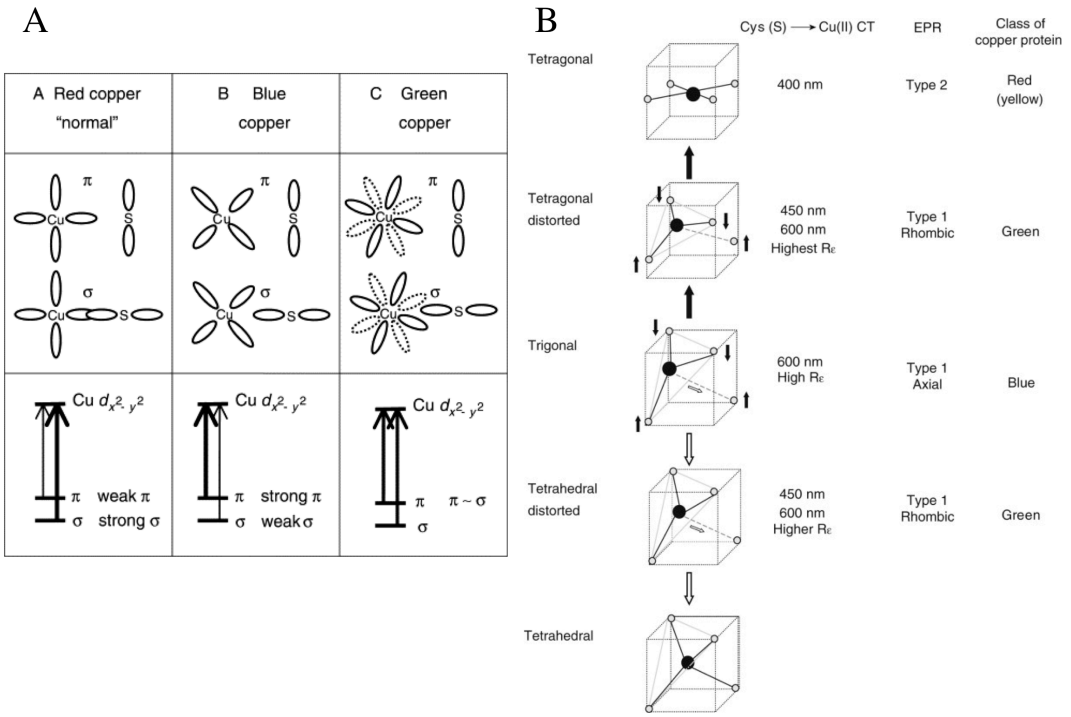



**Figure I-14.** A) Illustration of the relationship between the perturbation of the copper center and spectroscopic properties.[66] As the interaction between the copper ion and the axial ligand(s) increases, the coordination environment shifts towards a tetragonal (red/yellow) or tetrahedral geometry (green). B) Exhibits the rotation and interaction of the $Cud_{x^2-y^2}$ orbital with the $Sp_{x,y}$ orbitals. Blue copper centers have a strong π and weak σ overlap, while in green copper centers have slightly stronger σ and weaker π. Red or "normal" copper centers posses strong σ bond. Reprinted with permission from Ref 66. Copyright 2003 Elsevier B.V.

Conversely, in the perturbed CuT1 sites[87, 131, 132, 163], an increase in the intensity of near UV bands is observed, which is due to a rotation of the Cu $3d_{x2-y2}$ orbital that leads to a stronger $\sigma$ bond overlap as a consequence of a greater interaction with an axial ligand (Figure I-15). Stellacyanin (a CuT1.5 center or green copper center) contains a stronger axial ligand (Gln) than plastocyanin and has a distorted tetrahedral geometry. This slight perturbation from a trigonal planar geometry causes a rotation of the Cu $3d_{x2-y2}$ orbital that increases its interaction with the S(Cys) pseudo-$\sigma$ orbital and decreases its interaction with the S(Cys)-$\pi$ orbital. Consequently, stellacyanins show an additional intense absorption band at ~450 nm (~1100 $M^{-1}$ $cm^{-1}$) to

produce a green color in solution and a $R_\varepsilon$ value 0.3 – 0.5. The copper complex in nitrosocyanin, and to a lesser effect in nitrite reductase and cucumber basic protein, has the most perturbed site, signifying a rotation of the Cu $3d_{x2-y2}$ orbital away from a $\pi$ to a more $\sigma$ bonding interaction with the highest-occupied molecular orbital (HOMO).[164] Nitrosocyanin, a red copper protein, was previously included in the CuT1 family because it contains the conserved core residues. Its absorption spectrum mirrors CuT2 centers in a tetragonal geometry ("normal" copper) with a strong absorption band at 390 nm (~7000 $M^{-1}$ $cm^{-1}$) and a $R_\varepsilon$ value ~3.0. Furthermore, Resonance Raman spectroscopy studies on CuT1 centers corroborate well with the perturbation trend. Plastocyanin has an effective Cu-S stretching frequency of 403 $cm^{-1}$, while stellacyanin, cucumber basic protein and nitrite reductase have values of ~386, ~394 and 383 $cm^{-1}$, respectively. A decrease in the value demonstrates a weakening in the Cu-S bond (Figure I-15C).

The EPR spectra of CuT1 proteins also significantly deviate from "normal" or CuT2 centers (Figure I-13B & I-15D). Both copper centers show $g_{||} > g_{\perp} > 2.0023$, which is indicative of a $3d_{x2-y2}$ ground state. The EPR spectra of CuT1 proteins, however, display extremely small $A_{||}$ values < 100 x $10^{-4}$ $cm^{-1}$, with respect to CuT2 centers. Plastocyanin and azurin have nearly identical EPR spectra with axial EPR signals, and this demonstrates that O carbonyl Gly ligand has a very small influence on the electronic structure of azurin. The EPR spectra of the perturbed CuT1 centers can vary from rhombic ($\Delta g_{\perp} = g_x - g_y > 0.01$) (stellacyanin and basic cucumber protein) to axial (fungal laccase). Nitrosocyanin has an axial EPR signal, but again exhibits a CuT2 EPR characteristic with an $A_{||}$ value of ~144 x $10^{-4}$ $cm^{-1}$. The narrow $A_{||}$ was proposed to stem from a highly covalent Cu-$S_{Cys}$ bond, which delocalizes the unpaired copper electron spin towards the sulfur and reduces the interaction of the unpaired electron with the nuclear spin of $Cu^{2+}$.[87, 87, 140, 147, 148, 158, 161, 162] This highly covalent hypothesis was confirmed by polarized Cu K-edge X-ray absorption (XAS) studies that revealed $4p_{x,y}$ mixing into the $3d_{x2-y2}$ orbital.[147, 162] In addition, combined XAS studies at the Cu L-edge[148] and S K-edge supported a highly covalent Cu-S bond and further defined the electronic structure of the blue copper centers. From the Cu L-edge work, plastocyanin was determined to have ~41% Cu 3d character in the HOMO, while the

"normal" copper $CuCl_4$ reference has 61% Cu 3d character. Moreover, these experimental values agree well with computational studies on the electronic structure of plastocyanin.[87]

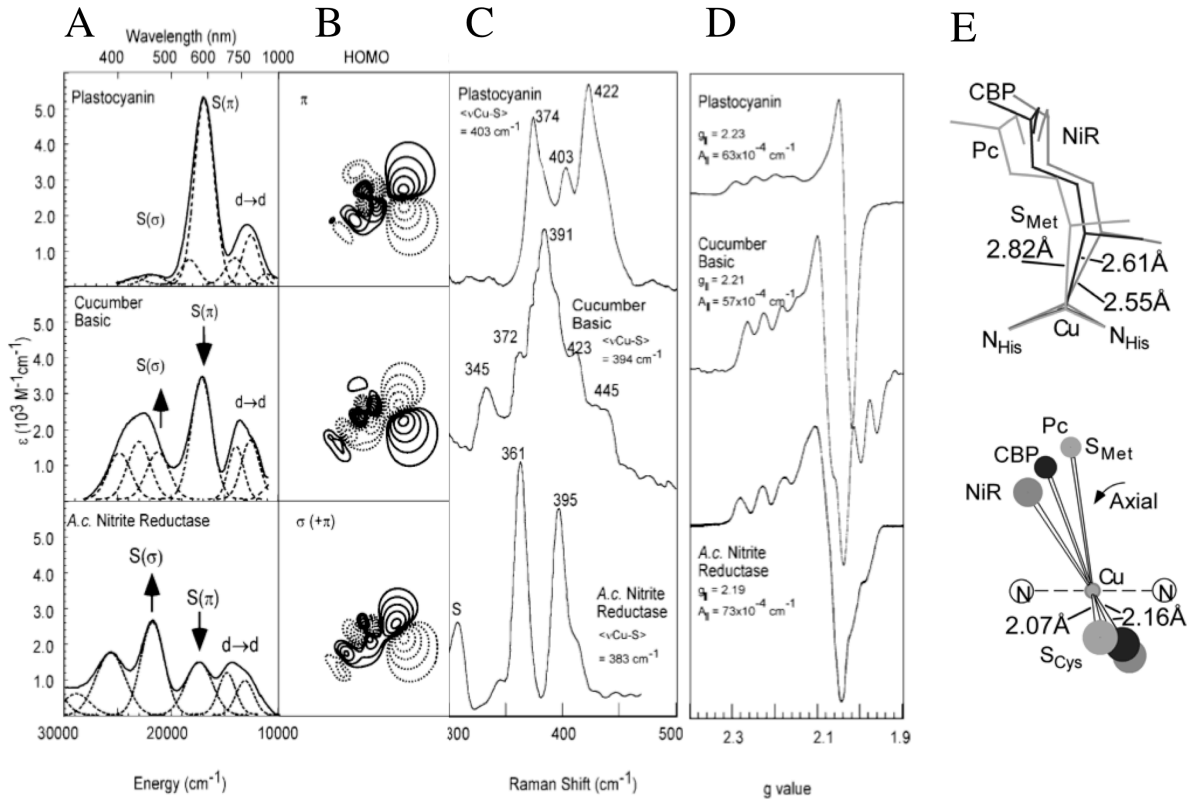



**Figure I-15.** Spectroscopic and structural comparison of between blue and perturbed CuT1 centers.[87] A) Low-temperature absorption spectra of plastocyanin, cucumber basic protein (CBP) and *Achromobacter cycloclastes* nitrite reductase, which highlights Cys π and σ bands. B) The π bonding interaction in plastocyanin and the rotation that occurs in nitrite reductase to generate σ and π mixture of redox active orbitals calculated with Xα-scattered waves. C) Resonance Raman spectra comparison showing weakening of the Cu-S(Cys) bond by the shift towards lower energy. D). EPR spectra of each center, where nitrite reductase has a modestly higher $A_{||}$ value. E) Top figure demonstrates a decrease of the Cu-axial ligand bond in CBP and nitrite reductase and an increase in the Cu-S(Cys) bond as indicated in the bottom figure. Reprinted with permission from Ref 87. Copyright 2004 American Chemical Society.

4. Protein redesign of cupredoxin centers

CuT1 centers have a unique coordination environment with characteristic spectroscopic and structural properties. Elucidating how the primary and the secondary coordination environments relate to the function of CuT1 as an electron transfer center is a major goal in protein design. The following section covers the transformation of a CuT2 into a CuT1 site in

copper-zinc superoxide dismutase and the incorporation of a CuT1 site into an unrelated native scaffold, thioredoxin.

4.1. Redesign of a cupredoxin center within a related fold

Copper–zinc superoxide dismutases (CuZnSOD) perform an essential role in alleviating the levels of reactive oxygen species in the cell by catalyzing the conversion of a superoxide radical into oxygen and hydrogen peroxide (Figure I-16A).[166-168] The zinc and copper ions are confined within a Greek β-barrel fold, similar to overall fold of CuT1 proteins, where the Zn(II) ion plays a structural role and the copper ion performs the dismutase reaction. Although there is no sequence or active site homology between the two proteins, researchers aimed to introduce a CuT1 site in CuZnSOD to determine if incorporating a Cys in the metal sites is sufficient in recapitulating the physical properties of CuT1 proteins.[145, 169-171] The crystal structure of CuZnSOD revealed that there are three types of His residues coordinating to the metal ions: His residues coordinated to copper (His46, 48, and 120); His that are coordinated to Zn(II) (His71 and His80); and a bridging His residue that binds to both copper and zinc ions (His63). To obtain a CuT1 center, Lu et al. mutated the copper-His in *Saccharomyces cerevisiae* CuZnSOD and produced mutants SOD-H46C and SOD-H120C.[169, 171] Upon copper binding, both mutants displayed spectroscopic characteristics of a CuT2 center instead of a CuT1 center, even though sulfur to copper LMCT was observed. The EPR hyperfine coupling constants of the mutants His46Cys and His120Cys were significantly larger by ~ 70 x $10^{-4}$ $cm^{-1}$ than that of a typical CuT1 center. A second mutation, substituting a Zn(II)–His with Cys at the structural site, was also reported by Lu et al by substituting His80 with Cys at the $Zn^{2+}$ site.[170] Upon the addition of $Cu^{2+}$ (before $Zn^{2+}$) or only $Cu^{2+}$, the copper complex formed in SOD-H80C generated a site that exhibits intense absorption bands at 459 and 595 nm analogous to green or CuT1.5 centers (stellacyanin), with an $R_\varepsilon$ of 1.03. These transitions were assigned to the sulfur-to-copper LMCT excitations. Further, the $A_{||}$ constant from the newly formed CuT1 site was estimated to ~15 x $10^{-4}$ $cm^{-1}$, which is comparable to stellacyanin. These studies indicate that the simple incorporation of a Cys ligand at the copper center in CuZnSOD is not sufficient to construct a CuT1 site; however, when Cys was placed in the more rigid structural site, a CuT1.5 site was achieved, which demonstrates the strict geometric constraints that are required for a CuT1 center.

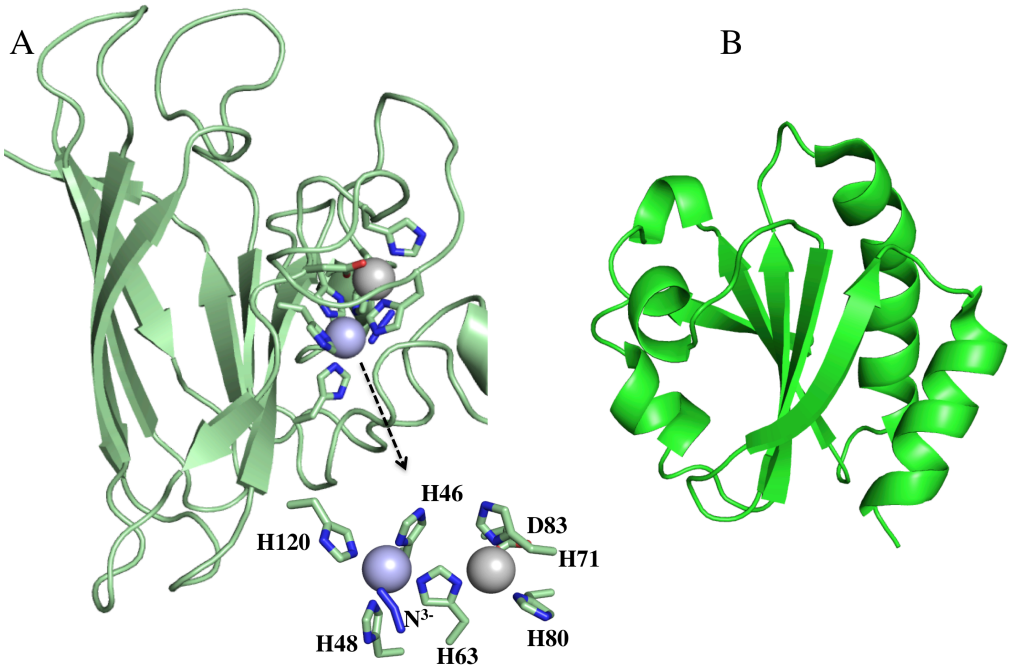



**Figure I-16**. A) X-ray crystal structure of yeast copper-zinc superoxide dismutase (PDB 1YAZ at 1.70 Å resolution).[166] B) X-ray crystal structure of thioredoxin expressed in *Escheriachia coli* (PDB 2TRX at 1.68 Å resolution).[172]

4.2. Redesign of a cupredoxin center within an unrelated fold.

Using thioredoxin (Trx)[172] (Figure I-16B), a small protein with a β/α fold that is devoid of a metal center as a scaffold, Hellinga constructed a $CysHis_2Met$ center within Trx to explore if an unrelated fold is able to adopt a CuT1 center. Hellinga's design approach utilized an automated rational protein design program, Dezymer.[173] Trx consists of five β-strands surrounded by four α-helices,[172] providing multiple possible sites for the construction of a CuT1 center. At the same time, the β/α fold is distinctly different from the native cupredoxin scaffold, which provides an excellent opportunity to examine the impact of the secondary coordination sphere on the properties of the T1Cu center. After several rounds of design cycles to examine the location of the site, possible competing coordinating ligands, the coordination geometry, solvent access, and equatorial versus axial ligands, a few mutants were prepared that exhibited strong sulfur-to-copper LMCT excitations upon binding to Cu(II). Nevertheless, these sites displayed a strong absorption band at ~400 nm and more closely resemble CuT2 centers than CuT1, which was attributed to the strong equatorial ligand and a weak axial ligand. Although these efforts

failed to introduce a classic CuT1 center with a tetrahedral coordination, this work still provided very important and interesting insights in the design of a CuT1 center. That is, the authors emphasized the importance of the negative design approach, which involves the incorporation of design elements that destabilizes any competing structures and deters access by exogenous molecules, such as $H_2O$.

5. *De Novo* designed cupredoxin centers

CuT1 centers are undeniably one of the most difficult metal centers to model. The oxidized copper form has been observed to form three different four-coordinate geometries: (1) tetragonal/square planar, (2) trigonal planar or (3) a distorted tetrahedral, in native proteins with a Cu(II)-S(Cys) binding site.[169, 174, 175] Further, a highly favored redox reaction between the Cu(II) ion and thiolate ligand (Equation I-2) has been observed in previous small molecule and protein models.[176-178] This redox reaction involves a reduction of the Cu(II) ion by the thiolate ligand to produce a Cu(I) ion and a disulfide bond, which is an indication of an unstable Cu(II)-S(Cys) bond. CuT1 proteins have evolved to contain a copper complex with a unique geometry that is finely controlled by secondary elements including hydrogen bonds and sterics provided by apolar residues, as well as the β-barrel overall fold that limits the access of exogenous molecules. The highly covalent Cu(II)-S(Cys) bond avoids unwanted redox reaction mentioned above. Therefore, the biological role of CuT1 proteins and their extraordinary copper complex makes them a desirable target to model via the *de novo* design approach. To date, the most successful CuT1 models (by Schnepf et al. and Tanaka and coworkers) utilize an α-helical fold.[179-182] *De novo* design offers further validation and critical understanding of how the geometric constraints affect the spectroscopic features of CuT1 centers. Ultimately, using an unrelated fold allows one to examine whether β-sheets are obligatory in an electron transfer reaction.

$$2RS^- + 2Cu(II) \rightleftharpoons RSSR + 2Cu(I) \qquad \textbf{Equation I-2}$$

Schnepf et al. reported the first *de novo* designed CuT1 center in an antiparallel four-helix bundle using a template-assembled synthetic (TASP) approach (Figure I-17A).[179, 180] This technique uses topological templates to covalently attach peptide strands to a predetermined

packing arrangement on a surface. The helices were derived from the backbone structure of the natural supercoiled four-helix bundle Repressor of Primer (Rop) protein.[183] Several modular protein (Mop) variants were synthesized to determine if a CuT1 center would be achieved. One of the first variants, which buries the copper binding site in the hydrophobic core, underwent the unwanted redox reaction between the Cu(II) ion and the Cys ligand. The UV-VIS absorption spectra of Cu(II) bound Mop5, Mop6, and Mop7 displayed strong absorption bands in the range at 410, 401, and 379 nm, respectively, which is an indication of a tetragonal coordination geometry. The authors assigned the absorption bands to a S(Cys) to Cu(II) LMCT and was further confirmed by Resonance Raman spectroscopy.

Based on these initial results that only focused on the primary $His_2Cys$ ligands, Schnepf et al. modified the secondary coordination sphere to examine its influence on the overall stability of the center, the copper coordination geometry and consequent spectroscopic properties.[180] They synthesized a library of proteins mutated from Mop5, the most stable copper-binding protein in the first design cycle,[179] and examined the copper-binding properties. The design of the second generation proteins[180] focused on modifying the secondary coordination sphere by introducing residues of differing polarity, bulkiness and flexibility to achieve two goals: to stabilize the copper center and to enforce the variation of copper coordination by changing the sterics around the copper binding site. Three types of copper centers were found in the initial screening step. Specifically, nonpolar residues such as Leu or a combination of Leu/Ala were placed above the $His_2Cys$ site to yield Mop23 and Mop21, respectively. In Mop22, a Met residue was introduced above both of the His ligands, providing a putative weak axial ligand as found in CuT1 centers. The Cu(II) absorption, EPR, Resonance Raman and Magnetic Circular Dichroism spectra were collected for these Mop variants. The Cu(II)-Mop21 species was determined to contain a tetragonal Cu(II) center; while the Cu(II)-Mop23 exhibited a CuT1 center with a distorted tetrahedral geometry, which was attributed to the difference in the steric hindrance above the copper site. Intriguingly, the spectroscopic features of Mop22, which contains an additional Met substitution, resulted in a binuclear copper, a purple $Cu_A$, center. In particular, the complex displayed an absorption band at 774 nm, indicative of a Cu–Cu bond. The Resonance Raman marker band at 345.5 $cm^{-1}$ matches well with the native $Cu_A$ center[184] and a relatively small $A_{||}$

supported a delocalized unpaired electron at a binuclear site. The MCD spectrum of Cu(II)-Mop22 was also remarkably similar to that of the native $Cu_A$ site.[185]

Taking a *de novo* design approach, Tanaka and coworkers built a $His_2Cys$ core in a four-stranded α-helical coiled coil scaffold to produce AM2C (Figure I-17B).[181, 182] Like many of the α-helical *de novo* design scaffolds, the amino acid sequence of AM2C was based on a heptad approach, containing Lys and Glu residues to form electrostatic interactions and nonpolar residues such as Ala and Leu to establish a hydrophobic interior. The $His_2Cys$ residues replaced Leu residues in three of the four helices, forming a pre-organized trigonal planar binding site.[181] The AM2C-copper complex exhibited a strong absorption feature at 616 nm, which was assigned to a S(Cys)-to-copper LMCT, and a weak LMTC band at 474 nm. The $A_{||}$ constant of AM2C-copper complex was unusually small and was not resolved in the X-band EPR. Using cyclic voltammetry, the reduction potential was determined to be +328 mV (vs NHE), a value that falls in the 180–800 mV range of CuT1 proteins. Moreover, XAS techniques were employed to gain further information on the copper coordination. The edge structures of Cu(II)-AM2C were similar to those of the CuT1 center in Az, and the EXAFS fittings resulted in a 2.3 Å copper sulfur bond, which is 0.1-0.2 Å longer than in native CuT1 centers, and 2.66 Å copper chloride bond when chloride was added as an exogenous ligand. This coordination environment mimics the unusually short Cu–S(Cys) and long axial Cu–S(Met) bond observed in native CuT1 centers.

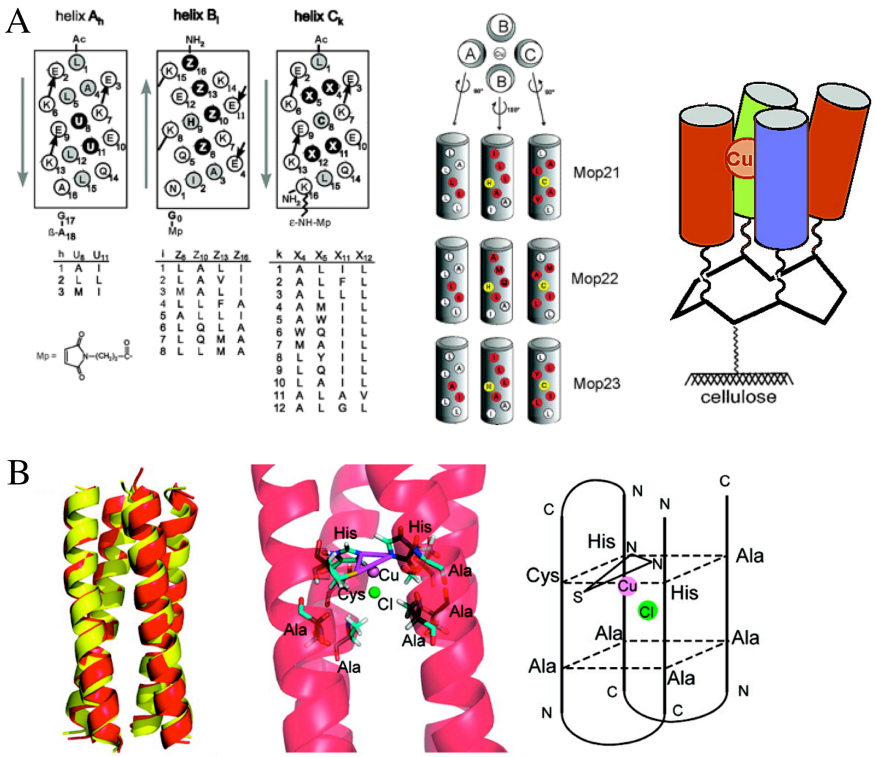



**Figure I-17**. A) Left figure is a helical net representation of the helices $A_h$, $B_i$ and $C_k$. Filled circles designated U, Z and X residues that were varied according to the tables shown below, while the filled-gray circles and open circles correlate to hydrophobic and polar residues, respectively.[179, 180] Middle figure represents the helical content of Mop21 ($A_2(B_5)_2C_{11}$), Mop22 ($A_2(B_7)_2)C_4$) and Mop23 ($A_1(B_5)_2C_8$). Right figure displays the preformed fold of a Mop scaffold, where the helices are attached to a cyclic decapeptide that is in turn tethered to a cellulose membrane. Reprinted with permission from Ref 179. Copyright 2004 American Chemical Society. B) Model of minimized (red) and initial (yellow) structure of AM2C-$Cu^{2+}$.[181] Middle figure is a view of the metal binding site of the minimized model and left figure represents the copper center. Reprinted with permission from Ref 181. Copyright 2010 American Chemical Society.

## 6. Thesis outline

From design to function, the central focus of this thesis entails the incorporation of a cupredoxin-inspired electron transfer center within a *de novo* designed antiparallel three-helix bundle of $\alpha_3$D (Figure I-18). Even though the physical properties of native CuT1 proteins have been established, as highlighted in this Introduction, *de novo protein* design nevertheless offers novel avenues for examining the structure-function relationship of ET centers, as well as other functional metal sites as follows: (1) It provides a framework for studying native functions in a completely unrelated protein fold. (2) *De novo* design allows for the examination of the primary coordination sphere (which includes only the ligands and the metal ion) to test whether that is

sufficient in recapitulating the desired function. (3) Subsequently, this approach affords a step-wise strategy in incorporating important secondary elements, such as steric-clash or H-bonding interactions. (4) Ultimately, the knowledge gained from these processes will inspire the design of multi-metal and functional designs that aim to study multi-faceted reactions in nature.

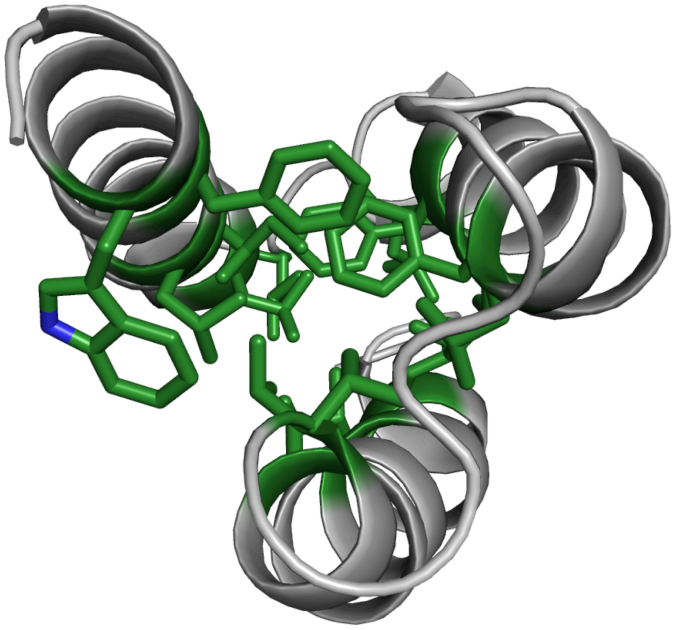


**Figure I-18**. The NMR solution structure of $\alpha_3$D (PDB 2A3D), an antiparallel three-helix bundle peptide (core residues designated in green).[10]

This thesis contains five chapters: an introduction and conclusion chapter and three research chapters. Chapter II will describe the apo structure of $\alpha_3$D**IV**. This iteration $\alpha_3$D possesses a thiol-rich metal binding site and has provided further insight into the metallobiochemistry of toxic heavy metals Cd, Hg and Pb. The determination of the apo structure demonstrated a well-defined and stable scaffold that can be utilized in achieving next generation metallopeptides Subsequently, using the structure of $\alpha_3$D**IV** in the design process, Chapter III will discuss the characterization of CuT1 constructs with several biophysical techniques, including CD, UV-vis, EPR, XAS and metal-NMR. Chapter IV will reveal the redox properties and ET reactivity of the designed constructs using electrochemical and photophysical techniques. The last chapter will summarize the major achievements of this work and future directions.

## References


[1] Waldron, K. J., Rutherford, J. C., Ford, D., and Robinson, N. J. (2009) Metalloproteins and metal sensing, *Nature 460*, 823-830.

[2] Lu, Y., Berry, S. M., and Pfister, T. D. (2001) Engineering Novel Metalloproteins: Design of Metal-Binding Sites into Native Protein Scaffolds, *Chem. Rev. 101*, 3047-3080.

[3] Lu, Y., Yeung, N., Sieracki, N., and Marshall, N. M. (2009) Design of functional metalloproteins, *Nature 460*, 855-862.

[4] Zastrow, M. L., and Pecoraro, V. L. (2014) Designing hydrolytic zinc metalloenzymes, *Biochemistry 53*, 957-978.

[5] Yu, F., Cangelosi, V. M., Zastrow, M. L., Tegoni, M., Plegaria, J. S., Tebo, A. G., Mocny, C. S., Ruckthong, L., Qayyum, H., and Pecoraro, V. L. (2014) Protein design: toward functional metalloenzymes, *Chem. Rev. 114*, 3495-3578.

[6] Plegaria, J. S., and Pecoraro, V. L. (2015) Sculpting Metal-binding Environments in *De Novo* Designed Three-helix Bundles, *Isr. J. Chem. 55*, 85-95.

[7] DeGrado, W. F., Summa, C. M., Pavone, V., Nastri, F., and Lombardi, A. (1999) *De novo* design and structural characterization of proteins and metalloproteins, *Annu. Rev. Biochem. 68*, 779-819.

[8] Williams, R. J. P. (1989) Overview of Biological Electron Transfer, In *Electron Transfer in Biology and the Solid State* (Johnson, M., King, R. B., Kurtz Jr., D. M., Kutal, C., Norton, M. L., and Scott, R. A., Eds.), pp 3-23, American Chemical Society, Washington, DC.

[9] Giedroc, D. P., and Arunkumar, A. I. (2007) Metal sensor proteins: nature's metalloregulated allosteric switches, *Dalton Trans. 0*, 3107-3120.

[10] Walsh, S. T., Cheng, H., Bryson, J. W., Roder, H., and DeGrado, W. F. (1999) Solution structure and dynamics of a *de novo* designed three-helix bundle protein, *Proc. Natl. Acad. Sci. U. S. A. 96*, 5486-5491.

[11] Chakraborty, S., Kravitz, J. Y., Thulstrup, P. W., Hemmingsen, L., DeGrado, W. F., and Pecoraro, V. L. (2011) Design of a three-helix bundle capable of binding heavy metals in a triscysteine environment, *Angew. Chem. Int. Ed. Engl. 50*, 2049-2053.

[12] Cangelosi, V. M., Deb, A., Penner-Hahn, J. E., and Pecoraro, V. L. (2014) A de novo designed metalloenzyme for the hydration of $CO_2$, *Angew. Chem., Int. Ed. Engl. 53*, 7900-7903.

[13] Krebs, J. F., Ippolito, J. A., Christianson, D. W., and Fierke, C. A. (1993) Structural and functional importance of a conserved hydrogen bond network in human carbonic anhydrase II, *J. Biol. Chem. 268*, 27458-27466.

[14] Guss, J. M., Bartunik, H. D., and Freeman, H. C. (1992) Accuracy and precision in protein structure analysis: restrained least-squares refinement of the structure of poplar plastocyanin at 1.33 Å resolution, *Acta Crystallogr., Sect. B: Struct. Sci. 48*, 790-811.

[15] Ghosh, D., and Pecoraro, V. L. (2005) Probing metal-protein interactions using a *de novo* design approach, *Curr. Opin. Chem. Biol. 9*, 97-103.

[16] Peacock, A. F., Iranzo, O., and Pecoraro, V. L. (2009) Harnessing natures ability to control metal ion coordination geometry using de novo designed peptides, *Dalton Trans.* 2271-2280.

[17] Ghosh, D., and Pecoraro, V. L. (2004) Understanding metalloprotein folding using a *de novo* design strategy, *Inorg. Chem. 43*, 7902-7915.

[18] Dieckmann, G. R., McRorie, D. K., Tierney, D. L., Utschig, L. M., Singer, C. P., O'Halloran, T. V., Penner-Hahn, J. E., DeGrado, W. F., and Pecoraro, V. L. (1997) *De Novo* Design of Mercury-Binding Two- and Three-Helical Bundles, *J. Am. Chem. Soc. 119*, 6195-6196.

[19] Dieckmann, G. R., McRorie, D. K., Lear, J. D., Sharp, K. A., DeGrado, W. F., and Pecoraro, V. L. (1998) The role of protonation and metal chelation preferences in defining the properties of mercury-binding coiled coils, *J. Mol. Biol. 280*, 897-912.

[20] Farrer, B. T., Harris, N. P., Balchus, K. E., and Pecoraro, V. L. (2001) Thermodynamic Model for the Stabilization of Trigonal Thiolato Mercury(II) in Designed Three-Stranded Coiled Coils, *Biochemistry 40*, 14696-14705.
[21] Matzapetakis, M., Farrer, B. T., Weng, T. C., Hemmingsen, L., Penner-Hahn, J. E., and Pecoraro, V. L. (2002) Comparison of the binding of cadmium(II), mercury(II), and arsenic(III) to the *de novo* designed peptides TRI L12C and TRI L16C, *J. Am. Chem. Soc. 124*, 8042-8054.
[22] Lee, K. H., Matzapetakis, M., Mitra, S., Marsh, E. N., and Pecoraro, V. L. (2004) Control of metal coordination number in de novo designed peptides through subtle sequence modifications, *J. Am. Chem. Soc. 126*, 9178-9179.
[23] Iranzo, O., Ghosh, D., and Pecoraro, V. L. (2006) Assessing the integrity of designed homomeric parallel three-stranded coiled coils in the presence of metal ions, *Inorg. Chem. 45*, 9959-9973.
[24] Matzapetakis, M., Ghosh, D., Weng, T. C., Penner-Hahn, J. E., and Pecoraro, V. L. (2006) Peptidic models for the binding of Pb(II), Bi(III) and Cd(II) to mononuclear thiolate binding sites, *J. Biol. Inorg. Chem. 11*, 876-890.
[25] Iranzo, O., Thulstrup, P. W., Ryu, S. B., Hemmingsen, L., and Pecoraro, V. L. (2007) The application of $^{(199)}$Hg NMR and $^{(199m)}$Hg perturbed angular correlation (PAC) spectroscopy to define the biological chemistry of Hg(II): a case study with designed two- and three-stranded coiled coils, *Chemistry 13*, 9178-9190.
[26] Iranzo, O., Cabello, C., and Pecoraro, V. L. (2007) Heterochromia in designed metallopeptides: geometry-selective binding of Cd(II) in a *de novo* peptide, *Angew. Chem. Int. Ed. Engl. 46*, 6688-6691.
[27] Peacock, A. F., Hemmingsen, L., and Pecoraro, V. L. (2008) Using diastereopeptides to control metal ion coordination in proteins, *Proc. Natl. Acad. Sci. U. S. A. 105*, 16566-16571.
[28] Luczkowski, M., Stachura, M., Schirf, V., Demeler, B., Hemmingsen, L., and Pecoraro, V. L. (2008) Design of thiolate rich metal binding sites within a peptidic framework, *Inorg. Chem. 47*, 10875-10888.
[29] Zastrow, M. L., Peacock, A. F., Stuckey, J. A., and Pecoraro, V. L. (2012) Hydrolytic catalysis and structural stabilization in a designed metalloprotein, *Nat. Chem 4*, 118-123.
[30] Tegoni, M., Yu, F., Bersellini, M., Penner-Hahn, J. E., and Pecoraro, V. L. (2012) Designing a functional type 2 copper center that has nitrite reductase activity within α-helical coiled coils, *Proc. Natl. Acad. Sci. U. S. A. 109*, 21234-21239.
[31] Zastrow, M. L., and Pecoraro, V. L. (2013) Influence of active site location on catalytic activity in *de novo* designed zinc metalloenzymes, *J. Am. Chem. Soc. 135*, 5895-5903.
[32] Yu, F., Penner-Hahn, J. E., and Pecoraro, V. L. (2013) *De novo* designed metallopeptides with type 2 copper centers: modulation of reduction potentials and nitrite reductase activities, *J. Am. Chem. Soc. 135*, 18096-18107.
[33] Zampella, G., Neupane, K. P., De Gioia, L., and Pecoraro, V. L. (2012) The importance of stereochemically active lone pairs for influencing Pb(II) and As(III) protein binding, *Chemistry 18*, 2040-2050.
[34] Neupane, K. P., and Pecoraro, V. L. (2010) Probing a homoleptic $PbS_3$ coordination environment in a designed peptide using $^{207}$Pb NMR spectroscopy: implications for understanding the molecular basis of lead toxicity, *Angew. Chem. Int. Ed. Engl. 49*, 8177-8180.
[35] Iranzo, O., Chakraborty, S., Hemmingsen, L., and Pecoraro, V. L. (2011) Controlling and fine tuning the physical properties of two identical metal coordination sites in *de novo* designed three stranded coiled coil peptides, *J. Am. Chem. Soc. 133*, 239-251.
[36] Farrer, B. T., and Pecoraro, V. L. (2003) Hg(II) binding to a weakly associated coiled coil nucleates an encoded metalloprotein fold: a kinetic analysis, *Proc. Natl. Acad. Sci. U. S. A. 100*, 3760-3765.
[37] Touw, D. S., Nordman, C. E., Stuckey, J. A., and Pecoraro, V. L. (2007) Identifying important structural characteristics of arsenic resistance proteins by using designed three-stranded coiled coils, *Proc. Natl. Acad. Sci. U. S. A. 104*, 11969-11974.

[38] Chakraborty, S., Touw, D. S., Peacock, A. F., Stuckey, J., and Pecoraro, V. L. (2010) Structural comparisons of apo- and metalated three-stranded coiled coils clarify metal binding determinants in thiolate containing designed peptides, *J. Am. Chem. Soc. 132*, 13240-13250.
[39] Peacock, A. F., Stuckey, J. A., and Pecoraro, V. L. (2009) Switching the chirality of the metal environment alters the coordination mode in designed peptides, *Angew. Chem. Int. Ed. Engl. 48*, 7371-7374.
[40] Neupane, K. P., and Pecoraro, V. L. (2011) Pb-207 NMR spectroscopy reveals that Pb(II) coordinates with glutathione (GSH) and tris cysteine zinc finger proteins in a $PbS_3$ coordination environment, *J. Inorg. Biochem. 105*, 1030-1034.
[41] Avvaru, B. S., Kim, C. U., Sippel, K. H., Gruner, S. M., Agbandje-McKenna, M., Silverman, D. N., and McKenna, R. (2010) A short, strong hydrogen bond in the active site of human carbonic anhydrase II, *Biochemistry 49*, 249-251.
[42] Lovejoy, B., Choe, S., Cascio, D., McRorie, D., DeGrado, W. F., and Eisenberg, D. (1993) Crystal structure of a synthetic triple-stranded α-helical bundle, *Science 259*, 1288-1293.
[43] Bryson, J. W., Desjarlais, J. R., Handel, T. M., and DeGrado, W. F. (1998) From coiled coils to small globular proteins: design of a native-like three-helix bundle, *Protein Sci. 7*, 1404-1414.
[44] Walsh, S. T. R., Lee, A. L., DeGrado, W. F., and Wand, A. J. (2001) Dynamics of a De Novo Designed Three-Helix Bundle Protein Studied by $^{15}N$,$^{13}C$, and $^{2}H$ NMR Relaxation Methods, *Biochemistry 40*, 9560-9569.
[45] Walsh, S. T., Cheng, R. P., Wright, W. W., Alonso, D. O., Daggett, V., Vanderkooi, J. M., and DeGrado, W. F. (2003) The hydration of amides in helices; a comprehensive picture from molecular dynamics, IR, and NMR, *Protein Sci. 12*, 520-531.
[46] Zhu, Y., Alonso, D. O., Maki, K., Huang, C.-Y. Y., Lahr, S. J., Daggett, V., Roder, H., DeGrado, W. F., and Gai, F. (2003) Ultrafast folding of $\alpha_3$D: a de novo designed three-helix bundle protein, *Proc. Natl. Acad. Sci. U. S. A. 100*, 15486-15491.
[47] Walsh, S. T., Sukharev, V. I., Betz, S. F., Vekshin, N. L., and DeGrado, W. F. (2001) Hydrophobic core malleability of a de novo designed three-helix bundle protein, *J. Mol. Biol. 305*, 361-373.
[48] O'Halloran, T., and Walsh, C. (1987) Metalloregulatory DNA-binding protein encoded by the merR gene: isolation and characterization, *Science 235*, 211-214.
[49] Wright, J. G., Tsang, H. T., Penner-Hahn, J. E., and O'Halloran, T. V. (1990) Coordination chemistry of the Hg-MerR metalloregulatory protein: evidence for a novel tridentate mercury-cysteine receptor site, *J. Am. Chem. Soc. 112*, 2434-2435.
[50] Wright, J. G., Natan, M. J., MacDonnel, F. M., Ralston, D. M., and O'Halloran, T. V. (1990) Mercury(II)—Thiolate Chemistry and the Mechanism of the Heavy Metal Biosensor MerR, *Inorg. Chem. 38*, 323-412.
[51] Busenlehner, L. S., Weng, T.-C., Penner-Hahn, J. E., and Giedroc, D. P. (2002) Elucidation of Primary (α3N) and Vestigial (α5) Heavy Metal-binding Sites in *Staphylococcus aureus* pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins, *J. Mol. Biol. 319*, 685-701.
[52] Ye, J., Kandegedara, A., Martin, P., and Rosen, B. P. (2005) Crystal structure of the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor, *J. Bacteriol. 187*, 4214-4221.
[53] Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Cavet, J. S., Dennison, C., Graham, A. I., Harvie, D. R., and Robinson, N. J. (2007) NMR structural analysis of cadmium sensing by winged helix repressor CmtR, *J. Biol. Chem. 282*, 30181-30188.
[54] Öz, G., Pountney, D. L., and Armitage, I. M. (1998) NMR spectroscopic studies of I= 1/2 metal ions in biological systems, *Biochem. Cell Biol. 76*, 223-234.
[55] Ian, M. A., Torbjörn, D., and Brian, R. (2013) Use of $^{(113)}$Cd NMR to probe the native metal binding sites in metalloproteins: an overview, *Met. Ions Life Sci. 11*, 117-144.
[56] Hemmingsen, L., Stachura, M., Thulstrup, P. W., Christensen, N. J., and Johnston, K. (2010) Selected applications of perturbed angular correlation of γ-rays (PAC) spectroscopy in biochemistry, *Hyperfine Inter. 197*, 255-267.

[57] Chakraborty, S. (2011) Designed Metalloproteins: from Structurally Characterized Scaffolds to Helical Bundles, In *Chemistry*, pp 1-270, University of Michigan, Ann Arbor, MI.
[58] Song, H., Wilson, D. L., Farquhar, E. R., Lewis, E. A., and Emerson, J. P. (2012) Revisiting zinc coordination in human carbonic anhydrase II, *Inorg. Chem. 51*, 11098-11105.
[59] Khalifah, R. G. (1971) The carbon dioxide hydration activity of carbonic anhydrase I. Stop-flow kinetic studies on the native human isoenzymes B and C, *J. Biol. Chem. 246,* 2561-2573.
[60] Rush, R. M., and Yoe, J. H. (1954) Colorimetric Determination of Zinc and Copper with 2-Carboxy-2'-hydroxy-5'-sulfoformazylbenzene, *Anal. Chem.26*, 1345-1347.
[61] Slebocka-Tilk, H., Cocho, J. L., and Frackman, Z. (1984) Carbonic anhydrase models, *J. Am. Chem. Soc. 106,* 2421-2431.
[62] Nakata, K., Shimomura, N., Shiina, N., Izumi, M., Ichikawa, K., and Shiro, M. (2002) Kinetic study of catalytic $CO_2$ hydration by water-soluble model compound of carbonic anhydrase and anion inhibition effect on CO2 hydration, *J. Inorg. Biochem. 89*, 255-266.
[63] Jackman, J. E., Merz, K. M., Jr., and Fierke, C. A. (1996) Disruption of the active site solvent network in carbonic anhydrase II decreases the efficiency of proton transfer, *Biochemistry 35*, 16421-16428.
[64] David, A. J., Chingkuang, T., Shanthi, R. P., Susan, M. T., Philip, V. L., Philip, J. L., and David, N. S. (1991) Enhancement of the catalytic properties of human carbonic anhydrase III by site-directed mutagenesis, *Biochemistry 30*.
[65] Winkler, J. R. (2011) Long-Range Electron Transfer in Biology, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[66] Lu, Y. (2003) Electron Transfer Cupredoxin, In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology* (Que, J., L., and Tolman, W. B., Eds.), pp 91-122, Elsevier Ltd., San Diego, CA, USA.
[67] Rodgers, K. R., and Lukat-Rodgers, G. S. (2003) Electron Transfer Cytochromes, In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology* (Que, J., L., and Tolman, W. B., Eds.), pp 17-60, Elsevier Ltd., San Diego, CA, USA.
[68] Holm, R. H. (2003) Electron Transfer Iron–Sulfur Clusters, In *Comprehensive Coordination Chemistry II: From Biology to Nanotechnology* (Que, J., L., and Tolman, W. B., Eds.), pp 61-90, Elsevier Ltd., San Diego, CA, USA.
[69] Rorabacher, D. B. (2004) Electron transfer by copper centers, *Chem. Rev. 104*, 651-697.
[70] Dmitry, M. (2011) Electron Transfer Reactions: Theory, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[71] Holwerda, R. A., Wherland, S., and Gray, H. B. (1976) Electron transfer reactions of copper proteins, *Annu. Rev. Biophys. Bioeng. 5*, 363-396.
[72] Chapman, S. K., Davies, D. M., Watson, A. D., and Sykes, A. G. (1983) Metalloproteins and Electron Transfer, In *Inorganic Chemistry, ACS Symposium Series* (Chisholm, M.) *211*, pp 177-197, American Chemical Society, Washington DC, USA.

[73] Marcus, R. A., and Sutin, N. (1985) Electron transfers in chemistry and biology, *Biochim. Biophys. Acta, Rev. Bioenerg. 811*, 265-322.

[74] Onuchic, J. N., Beratan, D. N., Winkler, J. R., and Gray, H. B. (1992) Pathway analysis of protein electron-transfer reactions, *Annu. Rev. Biophys. Biomol. Struct. 21*, 349-377.
[75] Marcus, R. A. (1993) Electron transfer reactions in chemistry. Theory and experiment, *Rev. Mod. Phys. 65*, 599-610.
[76] Jay, R. W. (2011) Long-Range Electron Transfer in Biology, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[77] Warren, J. J., Ener, M. E., Vlcek, A., Jr., Winkler, J. R., and Gray, H. B. (2012) Electron hopping through proteins, *Coord. Chem. Rev. 256*, 2478-2487.

[78] Noy, D., Moser, C. C., and Dutton, P. L. (2006) Design and engineering of photosynthetic light-harvesting and electron transfer using length, time, and energy scales, *Biochim. Biophys. Acta 1757*, 90-105.
[79] Christopher, R. S., and Robert, E. B. (2011) Photosynthesis, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[80] Babcock, G. T., and Wikstrom, M. (1992) Oxygen activation and the conservation of energy in cell respiration, *Nature 356*, 301-309.
[81] Chen, J., and Strous, M. (2013) Denitrification and aerobic respiration, hybrid electron transport chains and co-evolution, *Biochim. Biophys. Acta 1827*, 136-144.
[82] McIntosh, C. L., and Blankenship, R. E. (2014) Photosynthesis, 1-27.
[83] Hart, P. J., Nersissian, A. M., and George, S. D. (2011) Copper Proteins with Type 1 Sites, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[84] Kakutani, T., Watanabe, H., Arima, K., and Beppu, T. (1981) A blue protein as an inactivating factor for nitrite reductase from *Alcaligenes faecalis* strain S-6, *J. Biochem. 89*, 463-472.
[85] Adman, E. T., and Murphy, M. E. P. (2011) Copper Nitrite Reductase, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[86] Davies, G. J., and Ducros, V. (2011) Laccase , *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[87] Solomon, E. I., Szilagyi, R. K., DeBeer George, S., and Basumallick, L. (2004) Electronic structures of metal sites in proteins and models: contributions to function in blue copper proteins, *Chem. Rev. 104*, 419-458.
[88] Katoh, S., and Takamiya, A. (1964) Nature of Copper-Protein Binding in Spinach Plastocyanin, *J. Biochem. 55*, 378-387.
[89] Colman, P. M., Freeman, H. C., Guss, J. M., Murata, M., Norris, V. A., Ramshaw, J. A. M., and Venkatappa, M. P. (1978) X-ray crystal structure analysis of plastocyanin at 2.7 Å resolution, *Nature 272*, 319-324.
[90] Guss, J. M., and Freeman, H. C. (1983) Structure of oxidised poplar plastocyanin at 1.6 Å resolution, *J. Mol. Biol. 169*, 521-563.
[91] Redinbo, M. R., Yeates, T. O., and Merchant, S. (1994) Plastocyanin: structural and functional analysis, *J. Bioenerg. Biomembr. 26*, 49-66.
[92] Freeman, H. C., and Guss, J. M. (2011) Plastocyanin, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[93] Adman, E. T., Stenkamp, R. E., Sieker, L. C., and Jensen, L. H. (1978) A crystallographic model for azurin at 3 Å resolution, *J. Mol. Bio. 123*, 35-47.
[94] Adman, E. T., and Jensen, L. H. (1981) Structural Features of Azurin at 2.7 Å Resolution, *Isr. J. Chem. 21*, 8-12.
[95] Nar, H., Messerschmidt, A., Huber, R., van de Kamp, M., and Canters, G. W. (1991) Crystal structure analysis of oxidized *Pseudomonas aeruginosa* azurin at pH 5.5 and pH 9.0, *J. Mol. Biol. 221*, 765-772.
[96] Petratos, K., Dauter, Z., and Wilson, K. S. (1988) Refinement of the structure of pseudoazurin from *Alcaligenes faecali*s S-6 at 1.55 Å resolution, *Acta Crystallogr., Sect. B: Struct. Sci. 44*, 628-636.
[97] Adman, E. T. (2011) Pseudoazurin.
[98] Cox, J. C., and Boxer, D. H. (1978) The purification and some properties of rusticyanin, a blue copper protein involved in iron(II) oxidation from *Thiobacillus ferro-oxidans*, *Biochem J. 174*, 497-502.
[99] Walter, R. L., Ealick, S. E., Friedman, A. M., Blake Ii, R. C., Proctor, P., and Shoham, M. (1996) Multiple Wavelength Anomalous Diffraction (MAD) Crystal Structure of Rusticyanin: a Highly Oxidizing Cupredoxin with Extreme Acid Stability, *J. Mol. Biol. 263*, 730-751.
[100] Shoham, M. (2011) Rusticyanin, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[101] Arvidsson, R. H. A., Nordling, M., and Lundberg, L. G. (1989) The azurin gene from *Pseudomonas aeruginosa*. Cloning and characterization, *Eur. J. Biochem. 179*, 195-200.

[102] Farver, O., and Pecht, I. (1989) Long-range intramolecular electron transfer in azurins, *Proc. Natl. Acad. Sci. U. S. A. 86*, 6968-6972.
[103] Karlsson, B. G., Aasa, R., Malmström, B. G., and Lundberg, L. G. (1989) Rack-induced bonding in blue copper proteins: Spectroscopic properties and reduction potential of the azurin mutant Met-121 → Leu, *FEBS Lett. 253*, 99-102.
[104] Karlsson, B. G., Nordling, M., Pascher, T., Tsai, L. C., Sjolin, L., and Lundberg, L. G. (1991) Cassette mutagenesis of Met121 in azurin from *Pseudomonas aeruginosa*, *Protein Eng. 4*, 343-349.
[105] Messerschmidt, A. (2011) Ascorbate Oxidase, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[106] Lindley, P. F. (2011) Ceruloplasmin, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[107] Peisach, J., Levine, W. G., and Blumberg, W. E. (1967) Structural properties of stellacyanin, a copper mucoprotein from *Rhus vernicifera*, the Japanese lac tree, *J. Biol. Chem. 242*, 2847-2858.
[108] Peisach, J., Powers, L., Blumberg, W. E., and Chance, B. (1982) Stellacyanin - Studies of the Metal-Binding Site Using X-Ray Absorption-Spectroscopy, *Biophys. J. 38*, 277-285.
[109] Hart, P. J., Nersissian, A. M., Herrmann, R. G., Nalbandyan, R. M., Valentine, J. S., and Eisenberg, D. (1996) A missing link in cupredoxins: crystal structure of cucumber stellacyanin at 1.6 A resolution, *Protein Sci. 5*, 2175-2183.
[110] Nersissian, A. M., Hart, P. J., and Valentine, J. S. (2011) Stellacyanin, A Member of the Phytocyanin Family of Plant Proteins, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[111] Nersissian, A. M., Immoos, C., Hill, M. G., Hart, P. J., Williams, G., Herrmann, R. G., and Valentine, J. S. (1998) Uclacyanins, stellacyanins, and plantacyanins are distinct subfamilies of phytocyanins: plant-specific mononuclear blue copper proteins, *Protein Sci. 7*, 1915-1929.
[112] Koch, M., Velarde, M., Harrison, M. D., Echt, S., Fischer, M., Messerschmidt, A., and Dennison, C. (2005) Crystal structures of oxidized and reduced stellacyanin from horseradish roots, *J. Am. Chem. Soc. 127*, 158-166.
[113] Guss, J. M., Merritt, E. A., Phizackerley, R. P., and Freeman, H. C. (1996) The structure of a phytocyanin, the basic blue protein from cucumber, refined at 1.8 Å resolution, *J. Mol. Biol. 262*, 686-705.
[114] Ubbink, M. (2011) Cytochromef/Plastocyanin Complex, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[115] Sykes, A. G. (1985) Tilden Lecture. Structure and electron-transfer reactivity of the blue copper protein plastocyanin, *Chem. Soc. Rev. 14*, 283.
[116] Augustin, M. A., Chapman, S. K., Davies, D. M., Sykes, A. G., Speck, S. H., and Margoliash, E. (1983) Interaction of cytochrome c with the blue copper proteins, plastocyanin and azurin, *J. Biol. Chem. 258*, 6405-6409.
[117] Kolczak, U., Dennison, C., Messerschmidt, A., and Canters, G. W. (2011) Azurin and Azurin Mutants, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[118] Lappin, A. G., Lewis, C. A., and Ingledew, W. J. (1985) Kinetics and mechanisms of reduction of rusticyanin, a blue copper protein from *Thiobacillus ferrooxidans*, by inorganic cations, *Inorg. Chem. 24*, 1446-1450.
[119] Lieberman, R. L., Arciero, D. M., Hooper, A. B., and Rosenzweig, A. C. (2001) Crystal structure of a novel red copper protein from *Nitrosomonas europaea*, *Biochemistry 40*, 5674-5681.
[120] Arciero, D. M., Pierce, B. S., Hendrich, M. P., and Hooper, A. B. (2002) Nitrosocyanin, a red cupredoxin-like protein from *Nitrosomonas europaea*, *Biochemistry 41*, 1703-1709.
[121] Kyritsis, P., Lundberg, L. G., Nordling, M., Vänngård, T., Young, S., Tomkinson, N. P., and Sykes, A. G. (1991) The reactivity of spinach plastocyanin mutants with inorganic oxidants $[Fe(CN)_6]_3$ and $[Co(phen)_3]^{3+}$, *J. Chem. Soc., Chem. Commun.*, 1441.

[122] Kyritsis, P., Dennison, C., McFarlane, W., Nordling, M., Vanngard, T., Young, S., and Sykes, A. G. (1993) Reactions of five spinach plastocyanin PCu(I) mutants with $[Fe(CN)_6]_3$ and $[Co(phen)_3]^{3+}$ (phen = 1,10-phenanthroline) and related studies, *J. Chem. Soc., Dalton Trans.*, 2289.
[123] He, S., Modi, S., Bendall, D. S., and Gray, J. C. (1991) The surface-exposed tyrosine residue Tyr83 of pea plastocyanin is involved in both binding and electron transfer reactions with cytochrome f, *EMBO J. 10*, 4011-4016.
[124] Gong, X.-S., Wen, J. Q., Fisher, N. E., Young, S., Howe, C. J., Bendall, D. S., and Gray, J. C. (2000) The role of individual lysine residues in the basic patch on turnip cytochrome f for electrostatic interactions with plastocyanin in vitro, *Eur. J. .Biochem. 267*, 3461-3468.
[125] Gong, X.-S., Wen, J. Q., and Gray, J. C. (2000) The role of amino-acid residues in the hydrophobic patch surrounding the haem group of cytochrome f in the interaction with plastocyanin, *Eur. J. Biochem. 267*, 1732-1742.
[126] Nordling, M., Sigfridsson, K., Young, S., Lundberg, L. G., and Hansson, Ö. (1991) Flash-photolysis studies of the electron transfer from genetically modified spinach plastocyanin to photosystem I, *FEBS Lett. 291*, 327-330.
[127] Ejdebäck, M., Bergkvist, A., Karlsson, B. G., and Ubbink, M. (2000) Side-Chain Interactions in the Plastocyanin–Cytochromef Complex, *Biochemistry 39*, 5022-5027.
[128] Baker, E. N. (1988) Structure of azurin from *Alcaligenes denitrificans* refinement at 1·8 Å resolution and comparison of the two crystallographically independent molecules, *J. Mol. Biol. 203*, 1071-1095.
[129] Adman, E. T. (1991) Copper protein structures, *Adv. Protein Chem. 42*, 145-197.
[130] Nersissian, A. M., and Shipp, E. L. (2002) Blue copper-binding domains, *Adv. Protein Chem. 60*, 271-340.
[131] LaCroix, L. B., Shadle, S. E., Wang, Y., Averill, B. A., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1996) Electronic Structure of the Perturbed Blue Copper Site in Nitrite Reductase: Spectroscopic Properties, Bonding, and Implications for the Entatic/Rack State, *J. Am. Chem. Soc. 118*, 7755-7768.
[132] LaCroix, L. B., Randall, D. W., Nersissian, A. M., Hoitink, C. W. G., Canters, G. W., Valentine, J. S., and Solomon, E. I. (1998) Spectroscopic and Geometric Variations in Perturbed Blue Copper Centers: Electronic Structures of Stellacyanin and Cucumber Basic Protein, *J. Am. Chem. Soc. 120*, 9621-9631.
[133] Guss, J. M., Harrowell, P. R., Murata, M., Norris, V. A., and Freeman, H. C. (1986) Crystal structure analyses of reduced (CuI) poplar plastocyanin at 6 pH values, *J. Mol. Biol. 192*, 361-387.
[134] Malmstrom, B. G. (1994) Rack-induced bonding in blue-copper proteins, *Eur. J. Biochem. 223*, 711-718.
[135] Messerschmidt, A., Prade, L., Kroes, S. J., Sanders-Loehr, J., Huber, R., and Canters, G. W. (1998) Rack-induced metal binding vs. flexibility: Met121His azurin crystal structures at different pH, *Proc. Natl. Acad. Sci. U.S.A. 95*, 3443-3448.
[136] Malkin, R., and Malmström, B. G. (1970) *The State and Function of Copper in Biological Systems*, Interscience Publishers, New York.
[137] Visser, J. W., Amesz, J., and Van Gelder, B. F. (1974) EPR signals of oxidized plastocyanin in intact algae, *Biochim. Biophys. Acta 333*, 279-287.
[138] Fee, J. A. (1975) copper proteins systems containing the "Blue" copper center, In *Biochemistry*, pp 1-60, Springer Berlin Heidelberg.
[139] Solomon, E. I., Hare, J. W., and Gray, H. B. (1976) Spectroscopic studies and a structural model for blue copper centers inproteins, *Proc. Natl. Acad. Sci. U. S. A. 73*, 1389-1393.
[140] Penfield, K. W., Gewirth, A. A., and Solomon, E. I. (1985) Electronic structure and bonding of the blue copper site in plastocyanin, *J. Am. Chem. Soc. 107*, 4519-4529.

[141] Ainscough, E. W., Bingham, A. G., Brodie, A. M., Ellis, W. R., Gray, H. B., Loehr, T. M., Plowman, J. E., Norris, G. E., and Baker, E. N. (1987) Spectrochemical studies on the blue copper protein azurin from *Alcaligenes denitrificans*, *Biochemistry 26*, 71-82.
[142] Feiters, M. C., Dahlin, S., and Reinhammar, B. (1988) The metal site of stellacyanin: EXAFS studies of the Cu(II), Cu(I), Ni(II) and Co(II) derivatives, *Biochim. Biophys. Acta 955*, 250-260.
[143] Sykes, A. G. (1991) Active-Site Properties Of The Blue Copper Proteins, In *Adv. Inorg. Chem.* (Sykes, A. G., Ed.), pp 377-408, Elsevier Ltd.
[144] Han, J., Adman, E. T., Beppu, T., Codd, R., Freeman, H. C., Huq, L. L., Loehr, T. M., and Sanders-Loehr, J. (1991) Resonance Raman spectra of plastocyanin and pseudoazurin: evidence for conserved cysteine ligand conformations in cupredoxins (blue copper proteins), *Biochemistry 30*, 10904-10913.
[145] Han, J., Loehr, T. M., Lu, Y., Valentine, J. S., Averill, B. A., and Sandersloehr, J. (1993) Resonance Raman Excitation Profiles Indicate Multiple Cys-] Cu Charge-Transfer Transitions in Type-1 Copper Proteins, *J. Am. Chem. Soc. 115*, 4256-4263.
[146] Solomon, E. I., Lowery, M. D., LaCroix, L. B., and Root, D. E. (1993) Electronic absorption spectroscopy of copper proteins, *Methods Enzymol. 226*, 1-33.
[147] Shadle, S. E., Penner-Hahn, J. E., Schugar, H. J., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1993) X-ray absorption spectroscopic studies of the blue copper site: metal and ligand K-edge studies to probe the origin of the EPR hyperfine splitting in plastocyanin, *J. Am. Chem. Soc. 115*, 767-776.
[148] George, S. J., Lowery, M. D., Solomon, E. I., and Cramer, S. P. (1993) Copper L-edge spectral studies: a direct experimental probe of the ground-state covalency in the blue copper site in plastocyanin, *J. Am. Chem. Soc. 115*, 2968-2969.
[149] Grossmann, J. G., Ingledew, W. J., Harvey, I., Strange, R. W., and Hasnain, S. S. (1995) X-ray absorption studies and homology modeling define the structural features that specify the nature of the copper site in rusticyanin, *Biochemistry 34*, 8406-8414.
[150] Guckert, J. A., Lowery, M. D., and Solomon, E. I. (1995) Electronic Structure of the Reduced Blue Copper Active Site: Contributions to Reduction Potentials and Geometry, *J. Am. Chem. Soc. 117*, 2817-2844.
[151] Solomon, E. I., Lowery, M. D., Root, D. E., and Hemming, B. L. (1996) Electronic Structures of Active Sites in Copper Proteins and Their Contributions to Reactivity, In *Mechanistic Bioinorganic Chemistry* (Thorp, H. H., and Pecoraro, V. L., Eds.), pp 121-164, American Chemical Society.
[152] Zaitseva, I., Zaitsev, V., Card, G., Moshkov, K., Bax, B., Ralph, A., and Lindley, P. (1996) The X-ray structure of human serum ceruloplasmin at 3.1 Å: nature of the copper centres, *J. Biol. Inorg. Chem. 1*, 15-23.
[153] Solomon, E. I., Penfield, K. W., Gewirth, A. A., Lowery, M. D., Shadle, S. E., Guckert, J. A., and LaCroix, L. B. (1996) Electronic structure of the oxidized and reduced blue copper sites: contributions to the electron transfer pathway, reduction potential, and geometry, *Inorg. Chim. Acta 243*, 67-78.
[154] Palmer, A. E., Randall, D. W., Xu, F., and Solomon, E. I. (1999) Spectroscopic Studies and Electronic Structure Description of the High Potential Type 1 Copper Site in Fungal Laccase: Insight into the Effect of the Axial Ligand, *J. Am. Chem. Soc. 121*, 7138-7149.
[155] DeBeer, S., Randall, D. W., Nersissian, A. M., Valentine, J. S., Hedman, B., Hodgson, K. O., and Solomon, E. I. (2000) X-ray Absorption Edge and EXAFS Studies of the Blue Copper Site in Stellacyanin: Effects of Axial Amide Coordination, *J. Phys. Chem. B 104*, 10814-10819.
[156] Solomon, E. I. (2006) Spectroscopic methods in bioinorganic chemistry: blue to green to red copper sites, *Inorg. Chem. 45*, 8012-8025.
[157] Solomon, E. I., and Hadt, R. G. (2011) Recent advances in understanding blue copper proteins, *Coord. Chem. Rev. 255*, 774-789.

[158] Gewirth, A. A., and Solomon, E. I. (1988) Electronic structure of plastocyanin: excited state spectral features, *J. Am. Chem. Soc. 110*, 3811-3819.
[159] Randall, D. W., Gamelin, D. R., LaCroix, L. B., and Solomon, E. I. (1999) Electronic structure contributions to electron transfer in blue Cu and $Cu_A$, *J. Biol. Inorg. Chem. 5*, 16-19.
[160] Randall, D. W., George, D. S., Holland, P. L., Hedman, B., Hodgson, K. O., Tolman, W. B., and Solomon, E. I. (2000) Spectroscopic and Electronic Structural Studies of Blue Copper Model Complexes. 2. Comparison of Three- and Four-Coordinate Cu(II)−Thiolate Complexes and Fungal Laccase, *J. Am. Chem. Soc. 122*, 11632-11648.
[161] Solomon, E. I., Baldwin, M. J., and Lowery, M. D. (1992) Electronic structures of active sites in copper proteins: contributions to reactivity, *Chem. Rev. 92*, 521-542.
[162] Shadle, S. E., Hedman, B., Hodgson, K. O., and Solomon, E. I. (1994) Ligand K-Edge X-ray Absorption Spectroscopy as a Probe of Ligand-Metal Bonding: Charge Donation and Covalency in Copper-Chloride Systems, *Inorg. Chem.33*, 4235-4244.
[163] Augustine, A. J., Kragh, M. E., Sarangi, R., Fujii, S., Liboiron, B. D., Stoj, C. S., Kosman, D. J., Hodgson, K. O., Hedman, B., and Solomon, E. I. (2008) Spectroscopic studies of perturbed T1 Cu sites in the multicopper oxidases *Saccharomyces cerevisiae* Fet3p and *Rhus vernicifera* laccase: allosteric coupling between the T1 and trinuclear Cu sites, *Biochemistry 47*, 2036-2045.
[164] Basumallick, L., Sarangi, R., DeBeer George, S., Elmore, B., Hooper, A. B., Hedman, B., Hodgson, K. O., and Solomon, E. I. (2005) Spectroscopic and density functional studies of the red copper site in nitrosocyanin: role of the protein in determining active site geometric and electronic structure, *J. Am. Chem. Soc. 127*, 3531-3544.
[165] Solomon, E., and Lowery, M. (1993) Electronic structure contributions to function in bioinorganic chemistry, *Science 259*, 1575-1581.
[166] Hart, P. J., Balbirnie, M. M., Ogihara, N. L., Nersissian, A. M., Weiss, M. S., Valentine, J. S., and Eisenberg, D. (1999) A structure-based mechanism for copper-zinc superoxide dismutase, *Biochemistry 38*, 2167-2178.
[167] Kitagawa, Y., Tanaka, N., Hata, Y., Kusunoki, M., Lee, G.-p., Katsube, Y., Asada, K., Aibara, S., and Morita, Y. (1991) Three-Dimensional Structure of Cu,Zn-Superoxide Dismutase from Spinach at 2.0 A Resolution, *J. Biochem. 109*, 477-485.
[168] Djinovic Carugo, K., Battistoni, A., Carri, M. T., Polticelli, F., Desideri, A., Rotilio, G., Coda, A., Wilson, K. S., and Bolognesi, M. (1996) Three-dimensional structure of *Xenopus laevis* Cu,Zn superoxide dismutase b determined by X-ray crystallography at 1.5 Å resolution, *Acta Crystallogr. D Biol. Crystallogr. 52*, 176-188.
[169] Lu, Y., Gralla, E. B., Roe, J. A., and Valentine, J. S. (1992) Redesign of a Type-2 into a Type-1 Copper Protein - Construction and Characterization of Yeast Copper-Zinc Superoxide-Dismutase Mutants, *J. Am. Chem. Soc. 114*, 3560-3562.
[170] Lu, Y., Lacroix, L. B., Lowery, M. D., Solomon, E. I., Bender, C. J., Peisach, J., Roe, J. A., Gralla, E. B., and Valentine, J. S. (1993) Construction of a Blue Copper Site at the Native Zinc Site of Yeast Copper-Zinc Superoxide-Dismutase, *J. Am. Chem. Soc. 115*, 5907-5918.
[171] Lu, Y., Roe, J. A., Bender, C. J., Peisach, J., Banci, L., Bertini, I., Gralla, E. B., and Valentine, J. S. (1996) New Type 2 Copper-Cysteinate Proteins. Copper Site Histidine-to-Cysteine Mutants of Yeast Copper-Zinc Superoxide Dismutase, *Inorg. Chem. 35*, 1692-1700.
[172] Katti, S. K., LeMaster, D. M., and Eklund, H. (1990) Crystal structure of thioredoxin from Escherichia coli at 1.68 Å resolution, *Journal of Molecular Biology 212*, 167-184.
[173] Hellinga, H. W. (1998) Construction of a blue copper analogue through iterative rational protein design cycles demonstrates principles of molecular recognition in metal center formation, *J. Am. Chem. Soc. 120*, 10055-10066.
[174] Canters, G. W., and Gilardi, G. (1993) Engineering type 1 copper sites in proteins, *FEBS Lett. 325*, 39-48.

[175] Faham, S., Mizoguchi, T. J., Adman, E. T., Gray, H. B., Richards, J. H., and Rees, D. C. (1997) Role of the active-site cysteine of *Pseudomonas aeruginosa* azurin. Crystal structure analysis of the Cu(II) (Cys112Asp) protein, *J. Biol. Inorg. Chem. 2*, 464-469.
[176] Kitajima, N. (1992) Synthetic Approach to the Structure and Function of Copper Proteins, In *Adv. Inorg. Chem.*, pp 1-77, Elsevier, Ltd.
[177] Mizoguchi, T. J., Di Bilio, A. J., Gray, H. B., and Richards, J. H. (1992) Blue to type 2 binding. Copper(II) and cobalt(II) derivatives of a Cys112Asp mutant of *Pseudomonas aeruginosa* azurin, *J. Am. Chem. Soc. 114*, 10076-10078.
[178] Siluvai, G. S., Mayfield, M., Nilges, M. J., Debeer George, S., and Blackburn, N. J. (2010) Anatomy of a red copper center: spectroscopic identification and reactivity of the copper centers of *Bacillus subtilis* Sco and its Cys-to-Ala variants, *J. Am. Chem. Soc. 132*, 5215-5226.
[179] Schnepf, R., Horth, P., Bill, E., Wieghardt, K., Hildebrandt, P., and Haehnel, W. (2001) De novo design and characterization of copper centers in synthetic four-helix-bundle proteins, *J. Am. Chem. Soc. 123*, 2186-2195.
[180] Schnepf, R., Haehnel, W., Wieghardt, K., and Hildebrandt, P. (2004) Spectroscopic identification of different types of copper centers generated in synthetic four-helix bundle proteins, *J. Am. Chem. Soc. 126*, 14389-14399.
[181] Shiga, D., Nakane, D., Inomata, T., Funahashi, Y., Masuda, H., Kikuchi, A., Oda, M., Noda, M., Uchiyama, S., Fukui, K., Kanaori, K., Tajima, K., Takano, Y., Nakamura, H., and Tanake, T. (2010) Creation of a Type 1 Blue Copper Site within a *de Novo* Coiled-Coil Protein Scaffold, *J. Am. Chem. Soc. 132*, 18191-18198.
[182] Shiga, D., Hamano, Y., Kamei, M., Funahashi, Y., Masuda, H., Sakaguchi, M., Ogura, T., and Tanaka, T. (2012) Tuning the geometries of a *de novo* blue copper protein by axial interactions, *J. Biol. Inorg. Chem. 17*, 1025-1031.
[183] Banner, D. W., Kokkinidis, M., and Tsernoglou, D. (1987) Structure of the ColE1 Rop protein at 1.7 Å resolution, *J. Mol. Biol.196*, 657-675.
[184] Andrew, C. R., and SandersLoehr, J. (1996) Copper-sulfur proteins: Using Raman spectroscopy to predict coordination geometry, *Acc. Chem. Res. 29*, 365-372.
[185] Farrar, J. A., Neese, F., Lappalainen, P., Kroneck, P. M. H., Saraste, M., Zumft, W. G., and Thomson, A. J. (1996) The Electronic Structure of CuA: A Novel Mixed-Valence Dinuclear Copper Electron-Transfer Center, *J. Am. Chem. Soc. 118*, 11501-11514.

# Chapter II. Apoprotein Structure and Metal Binding Characterization of a *De Novo* Designed Peptide, $\alpha_3$DIV, that Sequesters Toxic Heavy Metals

## Introduction

Chakraborty and co-workers redesigned the sequence of $\alpha_3$D[1] by introducing a triscysteine motif that is found in the metal binding site of metalloregulatory proteins MerR,[2-4] ArsR/SmtB[5] and CadC/CmtR.[5-7] Three Leu to Cys mutations (Leu18Cys, Leu28Cys and Leu67Cys) at the C-terminal end of $\alpha_3$D were performed to produce $\alpha_3$DIV.[8] Utilizing several spectroscopic methods, we demonstrated that apo $\alpha_3$D**IV** is stable and well folded in solution at a pH range between 5-9. The coordination mode of Hg(II), Pb(II) and Cd(II) bound $\alpha_3$D**IV** is pH dependent (Figure I-1). From a linear mercury complex, $[Hg(II)S_2(SH)]$, a thiol group is deprotonated to form a trigonal $[Hg(II)S_3]^-$ and thiol group has a $pK_a$ of 7.1 (0.1). Therefore, a pH condition below and above this $pK_a$ value leads to the linear or a trigonal complex, respectively. The formation of a trigonal pyramidal $[Pb(II)S_3]^-$ and pseudotetrahedral $[Cd(II)S_3(N/O)]^-$ complex was determined to require a simultaneous deprotonation of two Cys thiol groups, yielding $pK_{a2}$ values of 10.6 (0.1) and 10.2 (0.1), respectively. A pH condition greater than 6 results in the formation of a $[Pb(II)S_3]^-$ and $[Cd(II)S_3(N/O)]^-$ complex. The $[Pb(II)S_3]^-$ complex was determined to have a lower limit binding constant of $2.0 \times 10^7\ M^{-1}$ and the $[Cd(II)S_3(N/O)]^-$ complex has the corresponding value of $3.1 \times 10^7\ M^{-1}$. Subsequently, Cangelosi and co-workers introduced a tris(histidine) zinc site within the $\alpha_3$D framework ($\alpha_3$D**H$_3$**) that is able to recapitulate the function of native carbonic anhydrase.[9] $\alpha_3$D**IV** and $\alpha_3$D**H$_3$** serve as excellent functional models of native metalloregulatory proteins and a metalloenzyme. However, fundamental understanding of $\alpha_3$D**IV** and metal site structures was lacking. To resolve this issue, I solved the structure of apo $\alpha_3$D**IV** to understand how the original $\alpha_3$D framework can tolerate modification in order to serve as a structural/functional scaffold for future designed

The work in this chapter was recently published, see reference 10.
NMR experiments were collected with Dr. Erik R. P. Zuiderweg from Department of Biological Chemistry, University of Michigan. NMR structural analysis were performed with the assistance of Dr. Timothy L. Stemmler from Department of Pharmaceutical Sciences, Wayne State University.

metalloproteins and metalloenzymes.[10] In this chapter, I present the solution structure of apo $\alpha_3$DIV, solved at pH 7.0. This is the first reported structure that incorporates a triscysteine metal binding site in an antiparallel three-helix bundle fold of $\alpha_3$D through the modifications of stabilizing core hydrophobic residues to introduce a new function. The solution structure of $\alpha_3$D**IV** possesses the same overall topology and counterclockwise bundle as $\alpha_3$D, and the incorporation of Cys residues increased the helical content of the scaffold. Overall, the $\alpha_3$D**IV** structure provides evidence that the framework of $\alpha_3$D is amenable to mutations that involve removing Leu residues that were thought to be essential in inducing hydrophobic interactions. This structure offers insight into how the protein is preorganized before metal binding, which is essential for utilizing this framework in designing future functional metallopeptides.

**Table II-1**: Amino acid sequence of *de novo* designed peptides

| Peptide | Sequence | | | |
|---|---|---|---|---|
| | **abcdefg** | **abcdefg** | **abcdef** | **loop** |
| $\alpha_3$D | MGSWAEFKQR | LAAIKTR | LQAL | GGS |
| | EAELAAFEKE | IAAFESE | LQAY | KGKG |
| | NPEVEALRKE | AAAIRDE | LQAYRHN | |
| $\alpha_3$D**IV** | MGSWAEFKQR | LAAIKTR | **C**QAL | GGS |
| | EAE**C**AAFEKE | IAAFESE | LQAY | KGKG |
| | NPEVEALRKE | AAAIRDE | **C**QAYRHN | |
| $\alpha_3$D**H$_3$** | MGSWAEFKQR | LAAIKTR | **H**QAL | GGS |
| | EAE**H**AAFEKE | IAAFESE | LQAY | KGKG |
| | NPEVEALRKE | AAAIRDE | **H**QAYR**V**N | **GSGA** |

The sequences are prepared in heptads. Residues that are underlined and bolded were changed from the previous design.

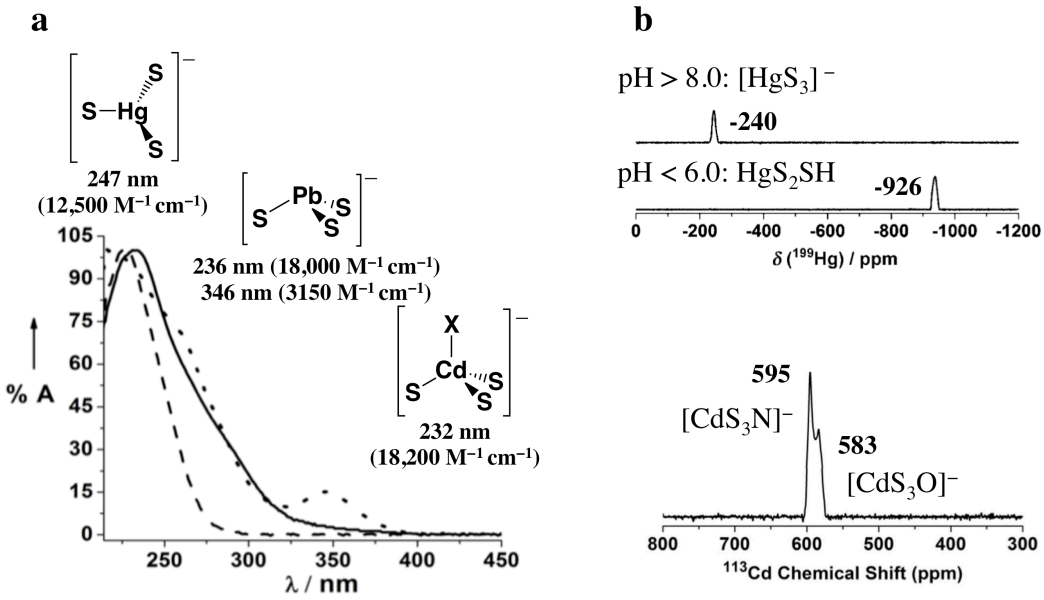



**Figure II-1** A) Uv-vis absorption spectra of metallated $\alpha_3$D**IV** species: Hg(II)- (solid line), Pb(II)- (dotted line) and Cd(II)-$\alpha_3$D**IV** (dashed line). B) Solution speciation behavior of Hg(II)-$\alpha_3$D**IV** as function of pH conditions and Cd-$\alpha_3$D**IV** observed by $^{199}$Hg-NMR and $^{113}$Cd-NMR, respectively.[8]

## Materials and Methods

*Protein Expression and Purification*. *Protein Expression and Purification*. A pET15b recombinant DNA plasmid (Celtek Genes) that contains the gene for $\alpha_3$D**IV** was transformed and expressed in *E. coli* BL21(DE3) (Life Technologies). To overexpress $^{15}$N-labeled and $^{15}$N,$^{13}$C- labeled $\alpha_3$D**IV**, *E. coli* cells were grown in M9 minimum media that contained 1 g/L $^{15}NHCl_4$ (Cambridge Isotope Laboratories) or 1 g/L $^{15}NHCl_4$ and 2 g/L [U-$^{13}$C]glucose (Cambridge Isotope Laboratories), respectively. The cells were grown at 37 °C to an $OD_{600}$ of 0.6, induced with 1 mM Isopropyl β-D-1-thiogalactopyranoside (Life Technologies) and incubated at 18 °C overnight. The cells were lysed by sonication, and the soluble protein was obtained after heat denaturation (50 °C) and lyophilization. The protein powder was redissolved in 10% acetic acid and purified on a reversed-phase C18 HPLC, using a flow rate of 20 mL/min and a linear gradient of 0.1% TFA in water to 0.1% TFA in 9:1 $CH_3CN$:$H_2O$ over 45 min. Using an electronspray mode on a Micromass LCT Time-of-Flight mass ionization spectrometer, the peptide mass was determined to be 7946.9 Da, which accounts for 72 of the 73 amino acids as Met1 is cleaved post-translation.

*NMR Sample Preparation* Stock peptide concentrations were determined from the absorption band of the aromatic residues (Trp4, Tyr45, and Ty70) at 280 nm. Experiments that involved coherence transfer from the backbone amide protons required a 9:1 $H_2O$-to-$^2H_2O$ solution, which contained 1.0 mM $^{15}N$, $^{13}C$ labeled peptide, 100 mM sodium chloride, 0.8 mM tris(2-carboxyethyl) phosphine (Fisher), 0.05 mM phenylmethylsulfonyl fluoride (Fisher) and 0.5% sodium azide (Sigma-Aldrich). Samples for carbon-filtered aromatic to aliphatic NMR experiments were prepared in a 100% $^2H_2O$ solution and incubated overnight, containing the reagents listed above. The pH of apoprotein samples used in the structure was adjusted at 7. The Pb(II)-$^{15}N$-$\alpha_3$D**IV** sample was prepared by adding 1.0 equivalent of $PbCl_2$ to a 1 mM apo $^{15}N$-$\alpha_3$D**IV** solution, containing 20 mM BIS-TRIS buffer set at pH 7.0 in a 9:1 $H_2O$-to-$^2H_2O$ solution. The Hg(II)-$^{15}N$-$\alpha_3$D**IV** sample was prepared under similar conditions but contained no buffer and the pH was adjusted to 8.6 after the addition of 1.0 equivalent of $HgCl_2$. Control apo samples were also prepared to match the sample conditions of their metallated counterparts.

*NMR Data Collection and Processing*. NMR experiments for the structure determination were performed at 25 °C on a Varian (Agilent) INOVA 800 MHz NMR spectrometer, equipped with a triple resonance cold probe. The NMR acquisition parameters for the 800 MHz spectrometer are reported in Table II-2. $^{15}N$-HSQC spectra were collected for the metallated $\alpha_3$D**IV** species and their corresponding apo controls. These experiments were performed on a Bruker Avance 500 MHz with room temperature triple resonance probes and a Varian VNMRS 500 MHz equipped with a switchable probe. The $^{15}N$-HSQC experiments for the Pb(II)-$^{15}N$-$\alpha_3$D**IV** and its apo control were collected at 25 °C, while the Hg(II)-$^{15}N$-$\alpha_3$D**IV** and its apo control were obtained at 9 °C. The $^{15}N$-HSQC experiments were also performed at 25 °C for apo $^{15}N$-$\alpha_3$D**IV** with pH conditions of 5.8 and 8.6. The spectra were processed using NMRPIPE[11] and then visualized and analyzed with SPARKY.[12]

**Table II-2**: Acquisition Parameters for NMR Experiments Performed on $\alpha_3$D**IV**.

| | no. of acquired data points (nucleus) | | | spectral width (Hz) | | | |
|---|---|---|---|---|---|---|---|
| **experiment** | **$t_1$** | **$t_2$** | **$t_3$** | **$F_1$** | **$F_2$** | **$F_3$** | **nt** |
| 2D $^1$H–$^{15}$N TROSY | 400 ($^{15}$N) | 4102 ($^1$H) | | 3000 | 12019 | | 2 |
| 3D HNCO TROSY | 150 ($^{13}$C) | 40 ($^{15}$N) | 2396 ($^1$H) | 2750 | 2500 | 12019 | 2 |
| 3D HN(CA)CO | 150 ($^{13}$C) | 80 ($^{15}$N) | 2368 ($^1$H) | 2750 | 2500 | 12019 | 4 |
| 3D HNCA | 150 ($^{13}$C) | 80 ($^{15}$N) | 2396 ($^1$H) | 5500 | 2500 | 12019 | 4 |
| 3D HN(CO)CA | 150 ($^{13}$C) | 80 ($^{15}$N) | 2396 ($^1$H) | 5500 | 2500 | 12019 | 2 |
| 3D HNCACB | 300 ($^{13}$C) | 80 ($^{15}$N) | 2396 ($^1$H) | 15001 | 2500 | 12019 | 8 |
| 3D HN(CO)CACB | 300 ($^{13}$C) | 80 ($^{15}$N) | 2396 ($^1$H) | 15001 | 2500 | 12019 | 8 |
| 3D HN(CA)HA[a] | 200 ($^1$H) | 60 ($^{15}$N) | 4102 ($^1$H) | 2500 | 2500 | 12019 | 4 |
| 3D HC(C)H–COSY[a] | 128 ($^1$H) | 200 ($^{13}$C) | 2048 ($^1$H) | 5500 | 16000 | 15060 | 4 |
| 3D HC(C)H–TOCSY | 200 ($^{13}$C) | 100 ($^{13}$C) | 2048 ($^1$H) | 16000 | 16000 | 15060 | 4 |
| 3D $^{15}$N NOESY TROSY | 200 ($^1$H) | 80 ($^{15}$N) | 4102 ($^1$H) | 8000 | 2500 | 12019 | 8 |
| 3D HMQC NOESY TROSY | 200 ($^{13}$C) | 80 ($^{15}$N) | 4102 ($^1$H) | 14000 | 2500 | 12019 | 12 |
| 3D $^{15}$N NOESY TROSY[a] | 200 ($^1$H) | 100 ($^{13}$C) | 2048 ($^1$H) | 9000 | 16000 | 15060 | 4 |
| 3D NOESY $^{13}$CHSQC[a] | 150 ($^1$H) | 200 ($^{13}$C) | 2048 ($^1$H) | 7500 | 16000 | 15060 | 4 |
| 3D NOESY $^{13}$C TROSY | 72 ($^{13}$C) | 80 ($^{15}$N) | 4102 ($^1$H) | 7000 | 2500 | 12019 | 4 |
| 3D NOESY ($^{13}$C) TROSY | 72 ($^1$H) | 80 ($^{15}$N) | 4102 ($^1$H) | 2000 | 2500 | 12019 | 4 |

[a]In-phase and anti-phase experiment

*Assignments*. Backbone assignments were determined from a series of complimentary three-dimensional (3D) triple resonance[13] Transverse Relaxation Optimized Spectroscopy (TROSY)[14-16] experiments: HNCO/HN(CA)CO,[17] HN(CO)CA/HNCA,[18] HN(CO)CACB/HNCACB[19, 20] and HN(CA)HA.[21] Since the sequence of $\alpha_3$D**IV** is known, the backbone assignments were manually completed and then successively confirmed with an auto-assignment program SAGA.[22] Ser24, Thr16 and Ser40 were used as landmarks in the manual assignment because the β-carbons of Ser and Thr residues are downfield of their α-carbons. Next, aliphatic assignments were obtained from the combined inphase/antiphase spectra of 3D HC(C)H-Correlation Spectroscopy (COSY)[23] and (H)CCH-Total Correlation Spectroscopy (TOCSY)[24, 25] experiments. These assignments were further confirmed with 3D ($^1$H, $^{15}$N, $^1$H) Nuclear Overhauser Effect Spectroscopy (NOESY)-TROSY[26] and ($^{13}$C, $^{15}$N, $^1$H) Heteronuclear Multiple-Quantum Coherence (HMQC)-TROSY[27] experiments, providing a through-space validation of the COSY and TOCSY assignments. Chemical shift assignments for residues 3-73 were then compiled and used to assign sequential-intra and inter-residue NOE upper distance limits (upl).

*NOESY Experiments*. In order to attain structural distance restraints, several 3D-$^{13}$C-edited and -resolved NOESY experiments were completed. Aliphatic-aliphatic NOEs were acquired from ($^{13}$C-edited) HC(C)H-NOESY (inphase/antiphase), which were collected at 60 and 200 ms mixing times. Aromatic to aliphatic NOEs were obtained from a 3D NOESY $^{13}$CHSQC experiment, while aromatic to amide NOEs were collected from 3D NOESY-$^{13}$C(resolved)-TROSY and the 3D NOESY ($^{13}$C resolved)-TROSY. Further, non-intraresidue amide proton NOEs were acquired from 3D ($^{1}$H, $^{15}$N, $^{1}$H) NOESY-TROSY and ($^{13}$C, $^{15}$N, $^{1}$H) HMQC-NOESY-TROSY experiments. The upper-limit NOE restraints, which contain intraresidue, intrahelical and interhelical NOEs, were then determined from the manual chemical shift assignments made on the COSY and TOCSY spectra.

*Dihedral Angles and Hydrogen Bonds Restraints*. TALOS-N[28] was used to generate $\phi$, $\psi$ and $\chi1$ dihedral angle restraints from $^{1}$H, $^{1}$H$_{\alpha}$, $^{14}$N, $^{13}$C, $^{13}$C$_{\alpha}$ and $^{13}$C$_{\beta}$ chemical shift assignments. Dihedral $\phi$ and $\psi$ angles that were classified as "strong" (residues 2-20, 25-45 and 45, 48, 51-70) were included in all the structure calculations and $\chi1$ angles for selected aliphatic residues were then included in the final rounds of calculations. TALOS-N also predicted the secondary structure and order parameters for each amino acid residue, which provided the structured regions: residues 5-20, 25-44 and 51-70. Furthermore, backbone hydrogen bonds in most of the structured regions were later incorporated in the upl restraints, excluding residues in the loop-turn regions. Upper limit distance restraints of 2.0 and 3.0 Å were given to $O_i$ to $H_{i+4}$ and $O_i$ to $N_{i+4}$, respectively.

*Structure Calculation*. CYANA 2.1[29] was used in the structure calculations and the input files comprised of the upper limit restraints obtained from 3D $^{15}$N- and $^{13}$C-edited/resolved-NOESY spectra, dihedral angle restraints ($\phi$, $\psi$ and $\chi1$) and backbone hydrogen bond restraints in the structured regions of the sequence. In the initial rounds of structure calculations, $\phi$ and $\psi$ dihedral angle restraints were used along with the NOE upper limit distances, which were set at 3 or 5 Å. These values were then adjusted in the later rounds of calculations. That is, intraresidue upper limit distances were modified to have a 3-7 Å range, increasing the restraints between pseudoatom-pseudoatom contacts. For sequential-intrahelical distances, a 3-4.5 Å range was

applied following Wuethrich's $^{1}H$-$^{1}H$ short-to-medium-range distances for an α-helical secondary structure.[30] Long sequential-intrahelical distances, $^{1}H_{i}$ to $^{1}H_{i+4}$, were given a 5-7 Å range. After, interhelical upl distances were adjusted to a 4-7 Å range, where the lower ends distances were based off of the intensity in the NOESY spectra and higher ends were set to again compensate for pseudoatom-pseudoatom correlations. In the final rounds of calculations, χ1 dihedral angle restraints were added and were followed by the incorporation of backbone hydrogen bonds. No lower limit distances were used. Of the 100 calculated structures, only the 20 lowest structures are reported here.

*Circular dichroism experiments*. Circular dichorism (CD) spectra were collected on an AVIV CD spectrometer at 25 °C and each sample were scanned between 260-195 nm. CD samples contained 10 μM peptide and 10 mM potassium phosphate buffer. The metallated samples were prepared by adding 1.0 equivalent of $HgCl_2$, $CdCl_2$ or $PbCl_2$ to an apo solution. The pH was adjusted at 8.2 for the apo, Cd(II)-, Pb(II)- and Hg(II)-$\alpha_3$D**IV**. The pH conditions were chosen so that the metallated samples formed the appropriate complex: $[Hg(II)S_3]^-$, $[Pb(II)S_3]^-$ and $[Cd(II)S_3(N/O)]^-$. All the samples were purged with $Ar_{(g)}$ and prepared in triplicates. The mean residue ellipticities (MRE) were determined using equation:

$$[\theta] = \theta_{obs}/10lcn$$

where $\theta_{obs}$ is the measured ellipticity in millidegrees, $l$ is the cell pathlength in centimeters, $c$ is the concentration in M, and $n$ is the number of residues in the structured regions. Fifty-five and 59 residues were used in the MRE calculation for $\alpha_3$D**IV** (apo and metallated species) and $\alpha_3$D, respectively.

*Thermal denaturation experiments*. Thermal denaturation studies were performed on a Nano Differential Scanning Calorimetry (N-DSC TA Instruments model 602001). The samples contained 0.13 mM peptide (1.0 mg/mL) and 30 mM MOPS or HEPES buffer solution that contained NaCl setting the ionic strength at 100 mM. The pH for the samples containing MOPS was adjusted at 7.0, while the HEPES solution had a pH of 8.2. The metallated samples were prepared by adding 1.0 equivalent of $HgCl_2$, $CdCl_2$ or $PbCl_2$ to an apo solution. All the samples were prepared in an anaerobic environment and were degassed prior to injection. In each

experiment, 300 µL of a peptide-buffer solution and its corresponding buffer solution were injected in the sample cell and reference cell, accordingly. Experiments were performed in duplicates or triplicates. The experimental methods involved a heating cycle that originated at 25 °C and ended at 110 °C, using a 1 or 2 °C/min scan rate, and spectra were collected in an anaerobic Coy Box environment. Thermograms were blanked with the appropriate control (buffer or metal buffer) scans, baseline corrected and normalized with protein concentration. NanoAnalyze Data Analysis (version 2.4.1 by TA Instruments) was used to fit the thermograms to determine the melting temperatures and thermodynamic parameters (Figure II-2).

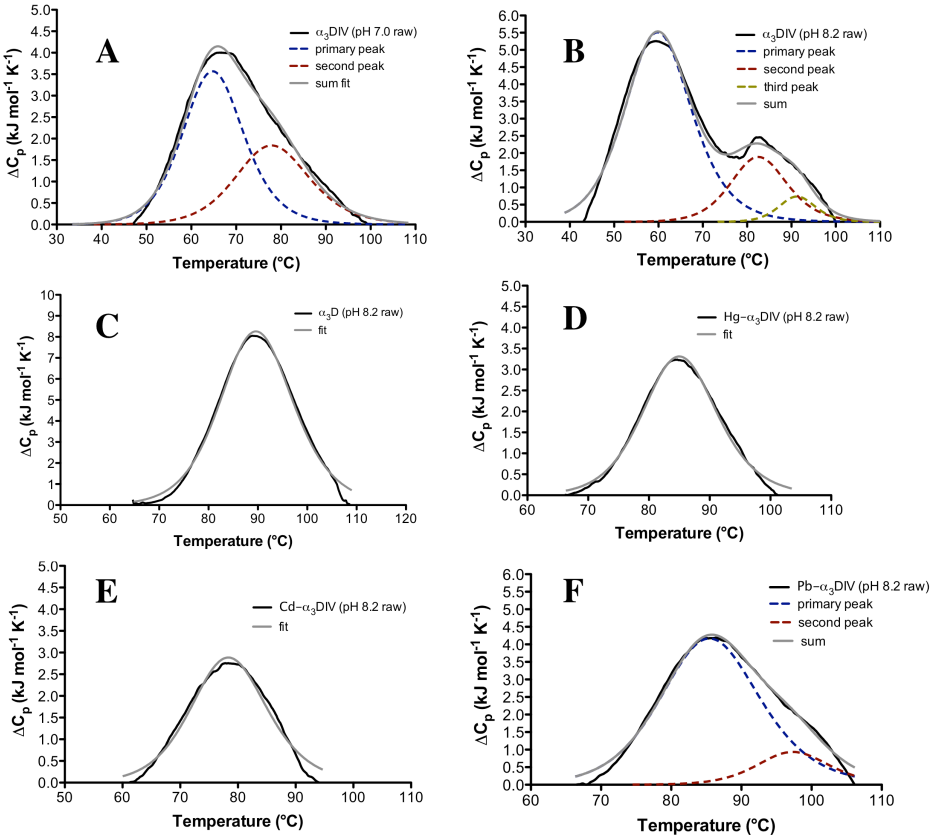



**Figure II-2**: A) Thermal denaturation curve of apo $\alpha_3$D**IV** at pH 7.0 fitted to a two-peak model. B) Thermal denaturation curve of apo $\alpha_3$D**IV** at pH 8.2 fitted to a three-peak model. C) Thermal denaturation curve of $\alpha_3$D at pH 8.2 fitted to a one-peak model. D) Thermal denaturation curve of Hg-$\alpha_3$D**IV** at pH 8.2 fitted to a one-peak model. E) Thermal denaturation curve of Cd-$\alpha_3$D**IV** at pH 8.2 fitted to a one-peak model. F) Thermal denaturation curve of Pb-$\alpha_3$D**IV** at pH 8.2 fitted to a two-peak model.

*X-Ray Absorption Spectroscopy (XAS).* Samples for XAS were prepared with final concentrations of 2.5 mM $HgCl_2$, 2.7 mM $\alpha_3$D**IV**, 30 mM TRIS buffer and 40% glycerol at pH 8.7. Samples were then loaded into lucite XAS sample cells wrapped in Kapton tape, flash frozen in liquid $N_2$ and stored at 77 K until data collection. XAS data were collected at the National Synchrotron Light Source (NSLS) on beamline X3-A. This beamline utilized a Si[220] single crystal monochromator equipped with a harmonic rejection mirror. During data collection, samples were maintained at 24° K using a He Displex Cryostat. Protein fluorescence excitation spectra were collected using a 13-element solid-state Ge detector with a Ga fluorescence filter (0.3 μM in width) placed between the cryostat and detector to remove lower energy photons. XAS spectra were measured in 5 eV increments in the pre-edge regions (12,200-12270 eV), 0.25 eV increments in the edge regions (12,270-12350 eV) and 0.05 $Å^{-1}$ increments in the extended X-ray absorption fine structure (EXAFS) region (to k = 14 $Å^{-1}$), integrating from 1s to 25 s in a $k^3$ weighted manner for a total scan length of approximately 50 minutes. X-ray energies were individually calibrated by collecting Au foil absorption spectra simultaneously with protein data. The first inflection point of the Au foil spectrum was assigned to 11,919 eV. Each fluorescence channel of each scan was examined for spectral anomalies and data represent the average of 20 scans.

XAS data were processed using the Macintosh OS X version of the EXAFSPAK program suite[31] integrated with the Feff v8 software[32] for theoretical model generation. Data reduction utilized a Gaussian function in the pre-edge region and a three-region cubic spline throughout the EXAFS region. Data were converted to k-space using a mercury $E_0$ value of 12,284 eV. The k cubed weighted EXAFS was truncated at 1.0 and 12 $Å^{-1}$ for filtering purposes. This k range corresponds to a spectral resolution of ca. 0.14 Å for all mercury-ligand interactions; therefore only independent scattering environments > 0.14 Å were considered resolvable in the EXAFS fitting analysis.[33] EXAFS simulation analysis was performed on filtered and then on raw/unfiltered data; results listed in Table II-3 were from simulating unfiltered data EXAFS data were fit using single scattering amplitude and phase functions calculated with the program Feff v8. Single scattering theoretical models were calculated for Hg-oxygen and Hg-sulfur coordination to simulate mercury nearest-neighbor ligand environments. Scale factors (Sc) and

$E_0$ values used during the simulations were calibrated by fitting crystallographically characterized Hg models; specific values include a Scale Factor of 0.95, and $E_0$ values for O and S of 0 and 1.5 eV, respectively. During the simulation, only the bond lengths and Debye-Waller factors were allowed to freely vary, adjusting coordination number values during the simulations in a non-varied incremental fashion. Criteria for judging the best-fit simulation utilized both the lowest mean square deviation between data and fit (F'), corrected for the number of degrees of freedom and a reasonable Debye-Waller factor.[33, 34]

**Table II-3**: Summary of Hg EXAFS fitting analysis for Hg-$\alpha_3$D**IV**.

| | Nearest Neighbor Ligand Environment[a] | | | | |
|---|---|---|---|---|---|
| **Complex** | ***Atom*[b]** | ***R* (Å)[b]** | ***C. N.*[d]** | **$\sigma^2$ [e]** | ***F*[f]** |
| Hg-$\alpha_3$D**IV** | S | 2.36 | 2.0 | 3.54 | 2.24 |
| | **S** | **2.36** | **2.5** | **4.97** | 2.21 |
| | S | 2.36 | 3.0 | 6.20 | 2.21 |

Data fit over a k range of 1 to 12 Å-1. Best fit simulation parameters are in bold. [a]Independent metal-ligand scattering environment at R < 3.0 Å. [b]Scattering atoms: S (Sulfur). [c]Average metal-ligand bond length for 2 independent samples. [d]Average metal-ligand coordination number for 2 independent samples. [e]Average Debye-Waller factor in $Å^2$ x $10^3$ for 2 independent samples. [f]Number of degrees of freedom weighted mean square deviation between data and fit.

## Results

### Structure of Apo $\alpha_3$DIV

*Overall structure*. Apo $\alpha_3$D**IV** has an α-helical fold as indicated by the pattern in the chemical shift dispersion in the $^{15}$N-TROSY spectrum (Figure II-3A) and sequential NOE correlations (Figure II-3B). The $^{15}$N-TROSY spectrum of apo $\alpha_3$D**IV** at pH 7.0 exhibits core residues with well-dispersed chemical shifts and a pattern that is typical for a well-folded α-helical protein.[35, 36] Chemicals shift assignments for 69 of the 72 residues, determined from triple resonance NMR experiments, are indicated in Figure II-3A and the total number of peaks validates the amino acid sequence of $\alpha_3$D**IV**. Resonances for residues Met1, Gly2 and Ser3 are not observed in the spectrum. Met1 is cleaved post-translation, Gly2 and Ser3 are in the dynamic region of the structure. Even after three Leu residues in $\alpha_3$D were substituted for Cys, the $^{15}$N-TROSY spectrum contains single peaks for every residue, indicating that $\alpha_3$D**IV** exhibits a single

conformation in solution or an ensemble of conformations interconverting at a sub-microsecond timescale.

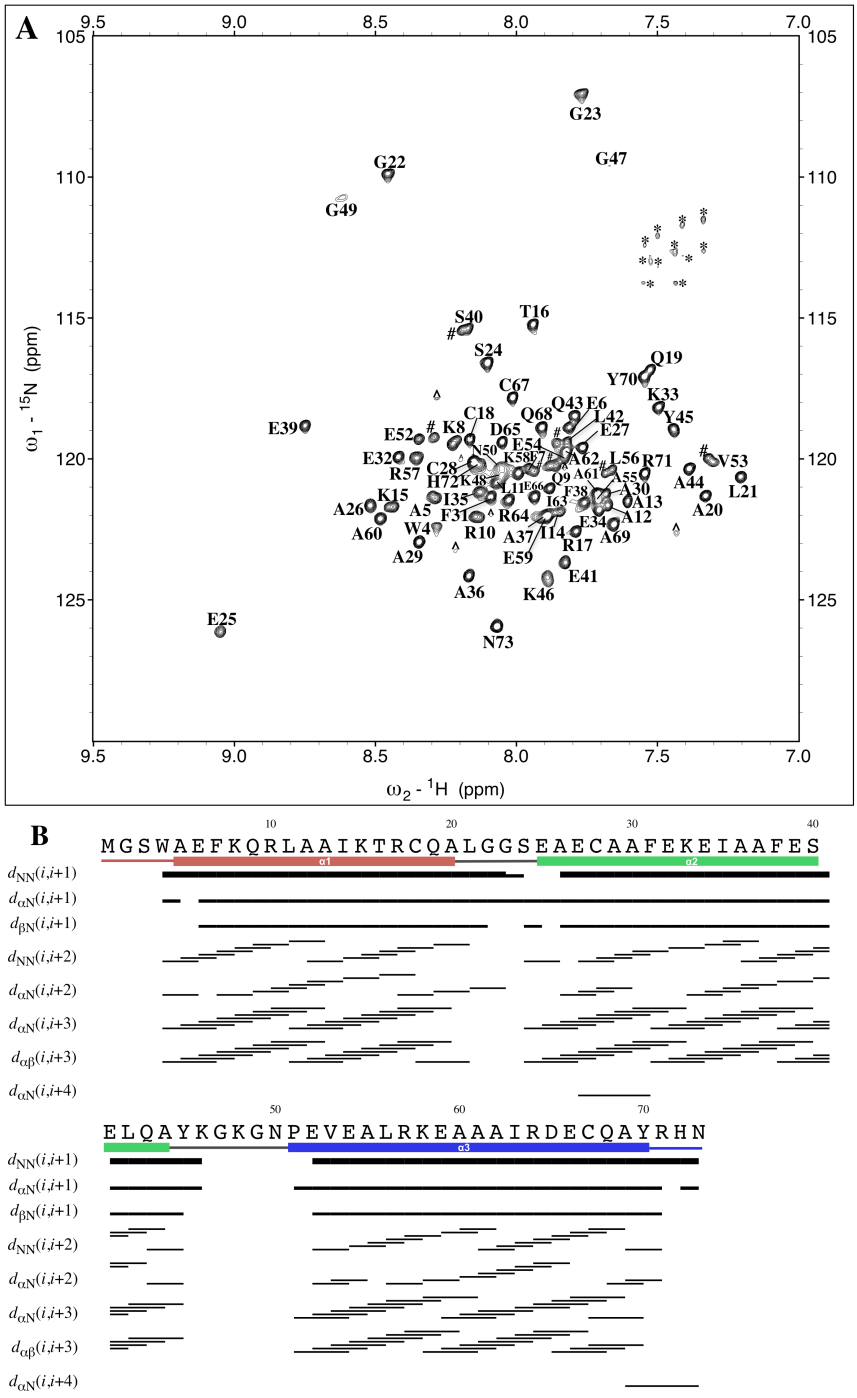



**Figure II-3**: A) $^{15}$N-TROSY spectrum of $^{15}$N-labeled $\alpha_3$D**IV**. The assignments are adjacent to their corresponding peaks. Of the 73 residues, 69 were assigned, with residues 1-3 and Pro51 not observed in the spectrum. The residual Gln and Asn side-chain peaks were not assigned and are marked with an asterisk. Further, aliased and noise peaks are respectively given a pound and circumflex symbol. B)

Summary of sequential NOEs for apo $\alpha_3$D**IV**, which were determined from 3D $^{15}$N- and $^{13}$C-NOESY-TROSY spectra.

An ensemble of the 20 lowest energy structures is visualized in Figure II-4A using PYMOL.[37] The structure of $\alpha_3$D**IV** encompasses 1067 NOE restraints, which include 395 intraresidue, 367 short-to-medium and 305 long range NOEs (Table II-4). There are ~15 NOE restraints per residue for 70 residues. In addition, 138 dihedral angle restraints were utilized in the structure, which were derived from the chemical shifts of $^{1}$H, $^{1}H_{\alpha}$, $^{14}$N, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ atoms for residues 4-73 using TALOS-N. These restraints included 60 $\phi$ and 61 $\psi$ angle restraints, as well as 17 $\chi_1$ angles. The majority of the $\chi_1$ restraints (10 restraints) were designated to core aliphatic and aromatic residues. The rotamers largely had a trans or gauche(-) conformation, characteristic rotamers of nonpolar-aliphatic residues. Finally, 39 hydrogen bond restraints (total of 78) were added in the structured regions. The ensemble has an averaged CYANA energy function of 1.9 (0.4) kcal mol$^{-1}$. Two distance restraints were greater than the 0.35 Å cut off, however no dihedral angles >5° or van der Waals contacts >0.35 Å were violated. The Ramachandran statistics for the 20 structures show that 90.1% of the backbone stereochemistry is in the favored regions (Figure II-5) and agrees well with the Ramachandran analysis by Protein Structure Validation Suite[38] (PSVS) (Table II-5). There are 55 residues in the structured regions (residues 5-20, 25-44, and 51-70) of $\alpha_3$D**IV**. The backbone (N, $C^{\alpha}$, C) RMSD of the 20 structures for residues 3-73 and the structured regions is 0.79 (0.16) and 0.49 (0.12) Å, respectively. Likewise, these RMSD values corroborate with the structure quality analysis by PSVS (see Table II-6 for the global quality scores).

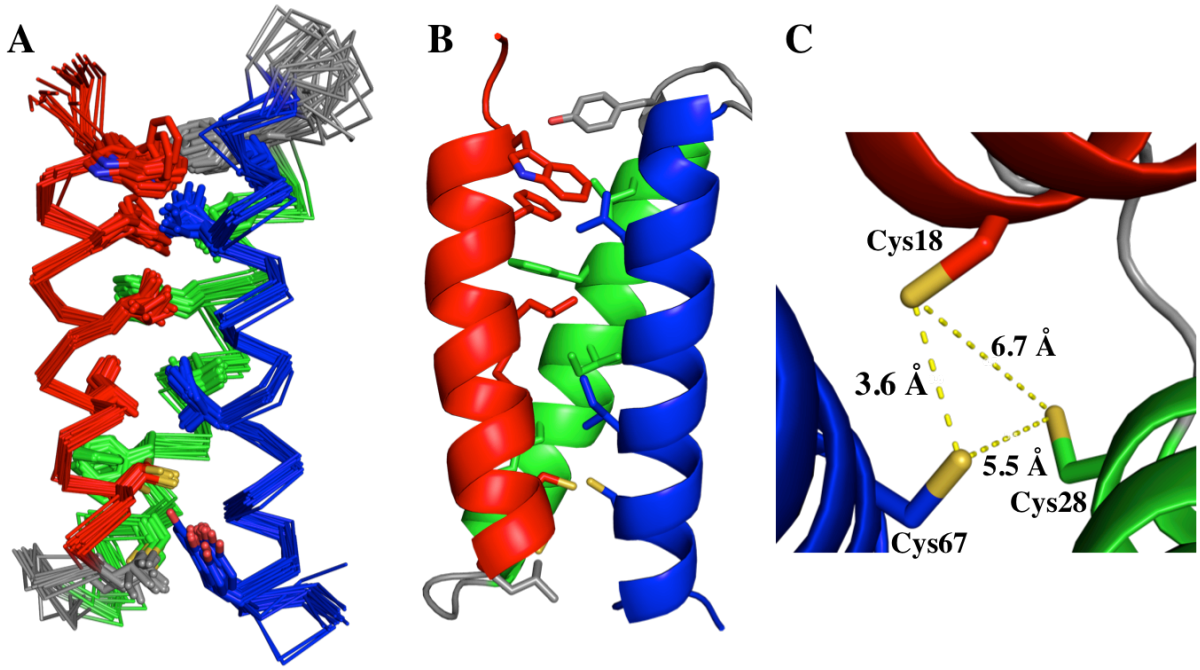



**Figure II-4**: A) An overlay of the 20 lowest-energy structures of apo $\alpha_3$D**IV**, calculated with CYANA[29] 2.1 and visualized with PYMOL.[37] Showing residues 1-73 (red: helix1, green: helix 2, blue: helix 3 and gray: loops 1 & 2). The structures were calculated from 1067 experimental NOE experimental restraints, 138 dihedral angles generated from the chemical shift index, and 78 backbone hydrogen bonds, added after the initial structure was obtained. The backbone and heavy atom RMSD values for residues 3-73 are 0.79 (0.16) Å and 1.31 (0.15) Å, respectively. For the structured regions residues 5-20, 25-44, and 51-70, the backbone and heavy atom RMSD values are 0.49 (0.12) Å and 0.97 (0.11) Å, respectively. B) Lowest energy structure, where the side chains of core residues are visualized. C) Top down view of the metal binding site: Cys18, Cys28 and Cys67.

**Table II-4**: Structural statistics for apo $\alpha_3$D**IV**.

| **Restraints** | | |
|---|---|---|
| Total NOE | 1067 | |
| Intraresidue | 395 | |
| Short-to-medium range ($1 < \|i - j\| < 5$) | 367 | |
| Long range ($\|i - j\| \geq 5$) | 305 | |
| Total dihedral angles[a] | 138 | |
| Φ | 60 | |
| Ψ | 61 | |
| $\chi^1$ | 17 | |
| Hydrogen bonds[b] | 39 x 2 | |
| Average CYANA target function (kcal mol$^{-1}$) | 1.9 (0.4) | |
| **Residual distance restraint violations**[c] | | |
| average no. of violations >0.35 Å | 0.00 (0.00) | |
| average of maximal violations (Å) | 0.24 (0.12) | |
| **Residual dihedral angle restraint violations**[c] | | |
| average no. of violations >5.0° | 0.00 (0.00) | |
| average of maximal violations (°) | 1.94 (0.67) | |
| **Van der Waals violations**[c] | | |
| average no. of violations >0.35 Å | 0.00 (0.00) | |
| average of sum violations (Å) | 0.28 (0.07) | |
| **Ramachandran statistics** | CYANA[d] | PSVS[e] |
| favored | 90.1% | 90.1% |
| allowed | 7.8% | 7.8% |
| generously allowed | 2.0% | 2.0% |
| disallowed | 0.1% | 0.1% |
| **RMSD from the mean structure (Å)** | CYANA | PSVS |
| Backbone (N, $C^\alpha$, C) | 0.79 (0.16)[f] | 0.8[g] |
| Heavy atoms (all non-H atoms) | 1.31 (0.15)[f] | 1.3[g] |
| Ordered backbone | 0.49 (0.12)[h] | 0.6[i] |
| Ordered heavy atoms | 0.97 (0.11)[h] | 1.0[i] |
| **Close Contacts and Deviations from Ideal Geometry (PDB validation software)** | | PSVS |
| Number of close contacts within 1.6 Å for H atoms & 2.2 Å for heavy atoms | | 0 |
| RMS deviation for bond angles (°) | | 0.2 |
| RMS deviation for bond lengths (Å) | | 0.001 |

Summary from CYANA[29] structure calculation. The ensemble of structures did not exhibit distance violations of >0.60 Å or dihedral angle violations >5°. [a]Dihedral angle restraints were obtained from a TALOS-N[28] analysis. [b]Upper-limit hydrogen bond distance restraints were used in the ordered regions of the sequence. [c]Violations were obtained from CYANA. [d]Ramachandran plot summary of Procheck-style analysis on CYANA, 1-73 residues. [e]Ramachandran plot summary for residues 1-73 from PSVS[38] analysis. [f]Residues 3-73. [g]Residues 1-73. [h]Structured regions: residues 5-20, 25-44, and 51-70. [i]Structured regions from PSVS analysis: residues 4-20, 25-42, and 51-70.

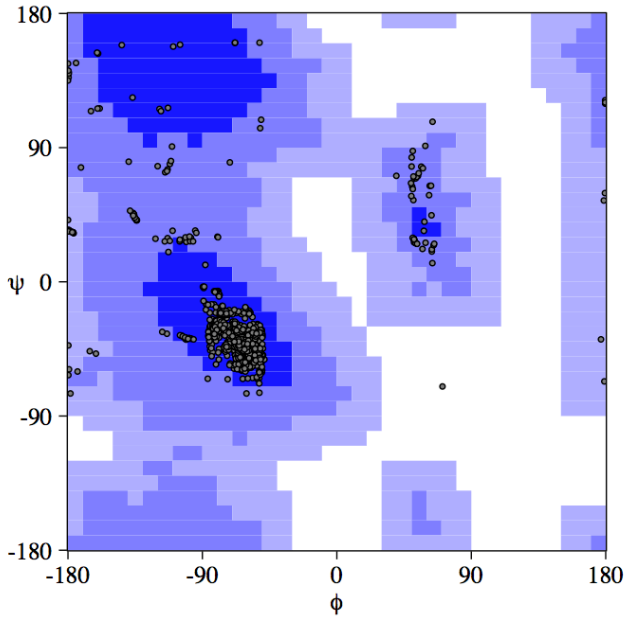



**Figure II-5**: Procheck-NMR Ramachandran plot of the 20 lowest structures exhibits that 90.1% of the backbone stereochemistry is located in the most favored, 7.8% in additionally allowed, 2.0% in generously allowed and 0.1% in disallowed regions.

**Table II-5**. Ramachandran plot summary for residues 1-73 from PSVS analysis.

| | Procheck | | Richardson Lab's Molprobity | |
|---|---|---|---|---|
| | $\alpha_3$D**IV** | $\alpha_3$D | $\alpha_3$D**IV** | $\alpha_3$D |
| Most favored regions | 89.2% | 86.2% | 93% | 81.7% |
| Allowed regions | 9.2% | 10.8% | 4.2% | 15.5% |
| Generously allowed regions | 1.5% | 1.5% | | |
| Disallowed regions | 0.0% | 1.5% | 2.8% | 2.8% |

**Table II-6**. PSVS global quality scores for 20 structures of $\alpha_3$D**IV**, structure 1 of $\alpha_3$D**IV** and structure 1 of $\alpha_3$D.

| Program | Verify3D | ProsaII (-ve) | Procheck (phi-psi)[a] | Procheck (all)[a] | Molprobity Clashscore | |
|---|---|---|---|---|---|---|
| 20 structures: $\alpha_3$D**IV** | | | | | | |
| Raw Score | 0.47 | 1.35 | 0.09 | -0.40 | 15.39 | |
| Z-Score[b] | 0.16 | 2.89 | 0.67 | -2.37 | -1.12 | |
| Structure 1: $\alpha_3$D**IV** | | | | | | |
| Raw Score | 0.47 | 1.45 | 0.02 | -0.37 | 17.95 | |
| Z-Score[b] | 0.16 | 3.31 | 0.39 | -2.19 | -1.55 | |
| **Close Contacts and Deviations from Ideal Geometry (PDB validation software)** | | | | | | |
| Number of close contacts (within 1.6 Å for H atoms & 2.2 Å for heavy atoms) | | | | | | 0 |
| RMS deviation for bond angles | | | | | | 0.2° |
| RMS deviation for bond lengths | | | | | | 0.001 Å |
| Structure 1: $\alpha_3$D | | | | | | |
| Raw Score | 0.55 | 1.29 | -0.64 | -0.98 | 68.48 | |
| Z-Score[b] | 1.44 | 2.65 | -2.20 | -5.80 | -10.23 | |
| **Close Contacts and Deviations from Ideal Geometry (PDB validation software)** | | | | | | |
| Number of close contacts (within 1.6 Å for H atoms & 2.2 Å for heavy atoms | | | | | | 0 |
| RMS deviation for bond angles | | | | | | 2.6° |
| RMS deviation for bond lengths | | | | | | 0.020 Å |

[a]For all residues. [b]With respect to mean and standard deviation for a set of 252 X-ray structures <500 residues of resolution of 1.80 Å, R-factor of 0.25 and R-free of 0.28; a positive value indicates a "better" score.

The lowest energy structure of $\alpha_3$DIV is illustrated in Figure 2B and has a CYANA energy function of 1.2 kcal $mol^{-1}$. The three-helix bundle adopts a counterclockwise orientation, which is in complete agreement with $\alpha_3$D. The interhelical-tilt angles, which were calculated using QHELIX,[39] further confirmed this counterclockwise topology: $\Omega_{1,2}$ = -149°, $\Omega_{1,3}$ = 21°, and $\Omega_{2,3}$ = -156. These angles decreased by 16°, 7° and 15° from the $\Omega_{1,2}$, $\Omega_{1,3}$ and $\Omega_{2,3}$ values reported for $\alpha_3$D, respectively. However, when the same QHELIX analysis was performed on $\alpha_3$D and compared again to $\alpha_3$DIV, a much more modest change was observed. The tilt angles of $\alpha_3$D reduced by 11° ($\Omega_{1,2}$), 8° ($\Omega_{1,3}$) and 8° ($\Omega_{2,3}$).

The triscysteine metal binding site has $S_\gamma$ – $S_\gamma$ distances of 6.7 Å, 5.5 Å and 3.6 Å between Cys18 – Cys28, Cys28 – Cys68 and Cys18 – Cys67, respectively (Figure II-4C). ). It should be noted that these distances were not restrained in the calculations. This hydrophobic plane has an area of 9.82 $Å^2$ and can accommodate large heavy metals like Cd(II), Hg(II) and Pb(II). In addition, the $\chi_1$ dihedral angles (N-$C_\alpha$-$C_\beta$-S$\gamma$) for the Cys residues are 168°, -52.7° and -68.1° for Cys18, Cys28 and Cys67, respectively. Both Cys28 and Cys67 have $\chi_1$ angles that are close to the most common rotamer for Cys (-65°).[40]

Analysis of metallated $\alpha_3$DIV

*$^{15}$N-HSQC spectra of metallated $\alpha_3$**DIV***. Our subsequent objective was to determine the chemical shift perturbation in the $^{15}$N-HSQC spectrum of $\alpha_3$D**IV** after the addition of Hg(II) and Pb(II). These results provide a qualitative assessment of the change in the peptide fold in the presence of a metal-ligand complex, as well as whether 3D NMR experiments can be collected in order to solve a metallated structure of $\alpha_3$D**IV**. However, the chemical shift peak dispersions in the $^{15}$N-HSQC spectrum of Pb(II)- and Hg(II)-$\alpha_3$D**IV** were observed to be significantly perturbed. Therefore, 3D NMR experiments were not obtained for the metallated species.

In Figure II-6, a superposition of the $^{15}$N-HSQC spectrum of an apo control and Pb(II)-$\alpha_3$D**IV** shows a significant decrease in the number of cross peaks for the Pb(II) spectrum. The pH of the apo and metallated species was set at 7.0 in order to achieve a trigonal pyramidal lead complex.[8] The apo spectrum displays 68 of the 69 expected peaks (Trp4 not observed), again demonstrating a well-defined conformation of the apo-peptide. Upon the addition of 1.0 equivalent Pb(II), only about 57 peaks could be observed in the Pb(II)-$\alpha_3$D**IV** spectrum and they were compared to corresponding peaks in the assigned apo spectrum. Of these identified peaks, several core residues, including Cys18, Phe38, Val53 and Ile63 overlay well with or slightly deviate by about ±0.2 ppm in $^1$H from their corresponding apo peaks. The peaks for other residues, such as Leu21, Phe31, Leu56 and Cys67, are either broadened beyond detection or have significantly shifted, resulting to an incomplete assignment of the chemical shifts for Pb(II)-$\alpha_3$D**IV**. Dramatic perturbation was also observed for the $^{15}$N-HSQC spectrum of Hg-$\alpha_3$D**IV** at pH 8.6 (Figure II-7), which forms a trigonal complex at this condition. The $^{15}$N-HSQC spectrum of Hg-$\alpha_3$D**IV**

contained only 47 observable peaks, while the spectrum of its apo counterpart only has 55 of the 68 expected resonance peaks at the same pH. The resonances in the Hg-$\alpha_3$D**IV** spectrum were not assigned.

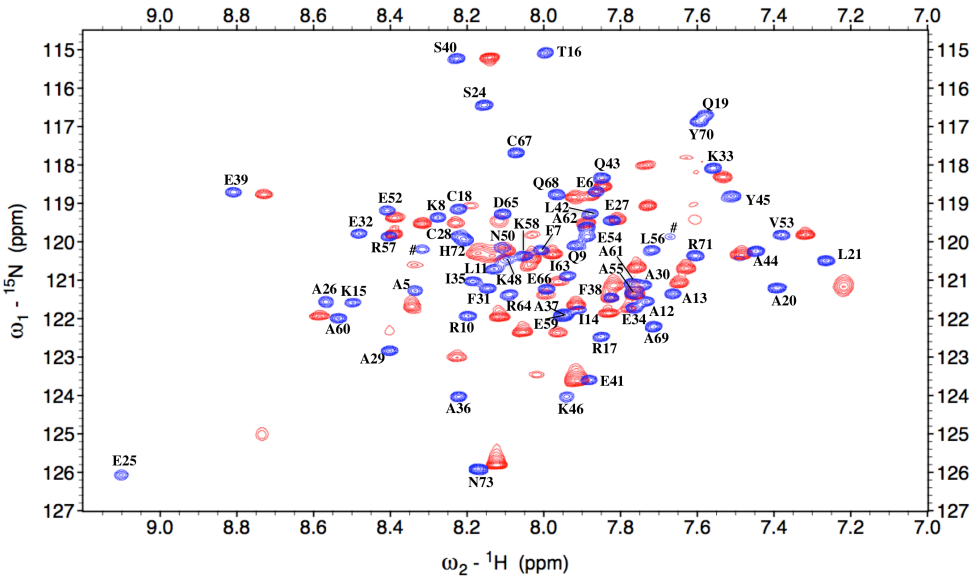



**Figure II-6**: Enlargement of the $^{15}$N-HSQC spectra of apo $\alpha_3$D**IV** (blue) and Pb(II)-$\alpha_3$D**IV** (red), both at pH 7.0. Spectra were collected on a 500 MHz Bruker Avance NMR spectrometer at 25 °C. The assignments in the apo spectrum are adjacent to their corresponding peak. In this view, the apo spectrum displays 63 of 68 assigned peaks, while the Pb spectrum $\alpha_3$D**IV** shows 54 of its 68 expected peaks.

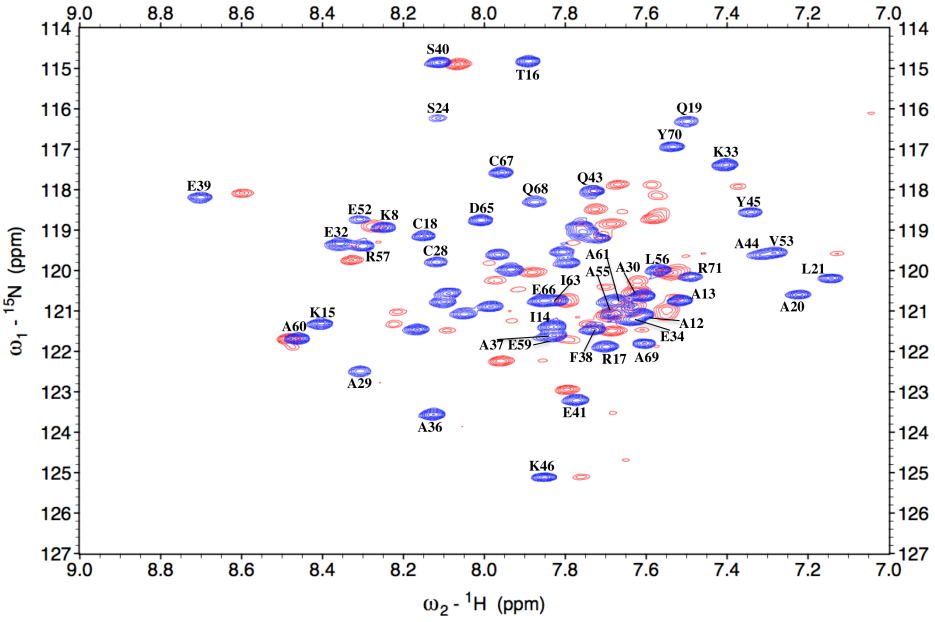



**Figure II-7**: Enlargement of the $^{15}$N-HSQC spectra of apo $\alpha_3$D**IV** (blue) and Hg(II)-$\alpha_3$D**IV** (red), both at pH 8.6. Spectra were collected on a 500 MHz Varian VNMRS NMR system at 9 °C. The peaks that were

assigned in the apo spectrum are labeled. At this high pH condition, a reduction in the chemical shift peaks for both spectra were expected as the backbone proton kinetic exchange rate increases with pH. Thereby, these experiments were collected at 9 °C in order to decrease the exchange rate and regain missing resonance peaks. In this view, the apo spectrum displayed 54 peaks of 55 that were identified, whereas the Hg-$\alpha_3$D**IV** spectrum contained 47 peaks.

*CD analysis of metallated $\alpha_3$DIV*. The CD spectra of Hg(II)-, Pb(II)- and Cd(II)-$\alpha_3$D**IV** at pH 8.2 are compared to apo $\alpha_3$D**IV**, as well as to $\alpha_3$D in Figure II-8A. Like its apo counterpart and $\alpha_3$D (both at pH 8.2), the metallated spectra display double minima at 208 and 222 nm, displaying a CD profile of a well-folded α-helical system. This is further supported by the large negative molar ellipticity values (Table II-7). Apo $\alpha_3$D**IV** has the largest $-[\theta]_{222\ nm}$ value of 29231 (672) deg $cm^2\ dmol^{-1}\ res^{-1}$; while Hg(II)-$\alpha_3$D**IV**, Pb(II)-$\alpha_3$D**IV** and Cd(II)-$\alpha_3$D**IV** had values of 27589 (421), 26940 (2420) and 27600 (487) deg $cm^2\ dmol^{-1}\ res^{-1}$, respectively. $\alpha_3$D had the lowest $-[\theta]_{222\ nm}$ value of 25213 (306) deg $cm^2\ dmol^{-1}\ res^{-1}$, indicating that $\alpha_3$D**IV** has a higher helical content than its parent structure. Furthermore, the $[\theta_{222}]/[\theta_{208}]$ ratio for all the samples were ~1.0, which is representative of a bundled or coiled-coil tertiary structure.[41] These results show that binding heavy metal ions to $\alpha_3$D**IV** does not lead to the unfolding or destabilization of its overall structure; instead $\alpha_3$D**IV** still retains an α-helical fold.

*DSC studies of metallated $\alpha_3$DIV*. The thermograms were collected on apo, Hg(II)-$\alpha_3$D**IV**, Pb(II)-$\alpha_3$D**IV**, Cd(II)-$\alpha_3$D**IV** and $\alpha_3$D with a Differential Scanning Calorimetry (DSC) apparatus. The thermal denaturation profile of $\alpha_3$D, apo samples (at pH 7.0 and 8.2) and metallated $\alpha_3$DIV are overlaid in Figure II-8B, while their respective thermodynamic parameters are listed in Table II-7. Apo $\alpha_3$D**IV** at pH 7.0 displays a broad melting curve, which is resolved in the thermogram of apo $\alpha_3$D**IV** at pH 8.2. The pH 7.0 thermograms fitted well to a two-peak model and the higher pH to a three-peak model. The first peak was identified to be the primary melting temperature of $\alpha_3$D**IV**. The primary peak at pH 7.0 has a melting temperature of 64.4 (0.6) °C, while the melting maxima at pH 8.2 decreased to 60.2 (0.1) °C, exhibiting a pH-dependence on the thermal-induced denaturation. Both species yielded similar $\Delta H_{cal}$ (excess heat capacity) values of 50.0 (0.2) and 49.9 (4.6) kcal $mol^{-1}$ for pH 7.0 and 8.2, respectively. The $\Delta H_{van't\ Hoff}$ was determined (SI Table II-8), and the $\Delta H_{cal}/\Delta H_{van't\ Hoff}$ ratio is listed in Table II-7. A value of 1 is indicative of a two-state unfolding model, while a value above or below 1 signifies self-association (e.g., as

dimer, trimer etc.) or unfolding through one or more intermediate states, respectively. This ratio for the parent structure, $\alpha_3$D, is 1.2. Apo $\alpha_3$D**IV** has value of 2.4 and 1.9 at pH 7.0 and 8.2, respectively. The metallated structures have $\Delta H_{cal}/\Delta H_{van't\ Hoff}$ ratio values > 2.0.

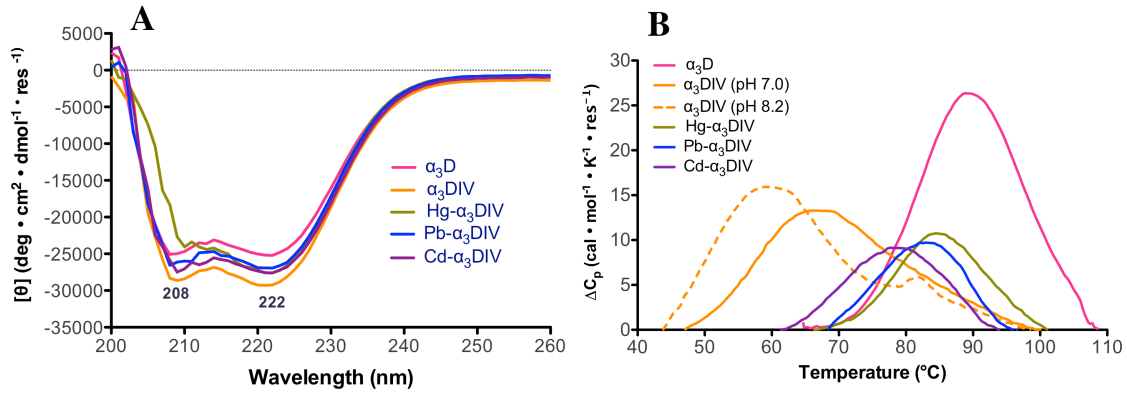



**Figure II-8**: Circular dichroism spectra and thermograms of $\alpha_3$D, $\alpha_3$D**IV** and metallated $\alpha_3$D**IV**. A) CD spectrum of $\alpha_3$D (pink), $\alpha_3$D**IV** (orange), Hg(II)-$\alpha_3$D**IV** (green), Pb(II)-$\alpha_3$D**IV** (blue) and Cd(II)-$\alpha_3$D**IV** (purple). Each spectrum contains a double minima at 208 and 222 nm and [θ] molar ellipticity values typical for a well-folded α-helical structure. B) Thermograms of $\alpha_3$D (pink), $\alpha_3$D**IV** at pH 7.0 (orange), $\alpha_3$D**IV** at pH 8.2 (dashed orange), Hg(II)-$\alpha_3$D**IV** (green), Pb(II)-$\alpha_3$D**IV** (blue), and Cd(II)-$\alpha_3$D**IV** (purple). The metallated species had melting temperatures ~20 °C higher than the apo.

**Table II-7**: Circular dichroism and thermal denaturation parameters.

| **Sample** | **$-[\theta]_{222\ nm}$ (deg cm$^2$ dmol$^{-1}$ res$^{-1}$)** | **$[\theta]_{222\ nm}/[\theta]_{208\ nm}$** | **Tm (°C)** | **$\Delta H_{cal}$[a] (kcal mol$^{-1}$)** | **$\Delta H_{cal}/\Delta H_{van't\ Hoff}$** |
|---|---|---|---|---|---|
| $\alpha_3$D[b] | 25213 (306) | 1.01 | 89.6 (0.3) | 49.9 (4.6) | 1.2 (0.3) |
| $\alpha_3$D**IV** | 29231 (672)[b] | 1.03 | 64.4 (0.6)[c]<br>60.2 (0.1)[b] | 50.0 (0.2)<br>44.9 (1.2) | 2.4 (0.4)<br>1.9 (0.2) |
| Hg-$\alpha_3$D**IV**[b] | 27589 (421) | 1.51 | 84.0 (1.7) | 65.1 (3.5) | 6.7 (2.4) |
| Pb-$\alpha_3$D**IV**[b] | 26940 (2420) | 1.01 | 83.4 (3.3) | 59.2 (6.9) | 4.5 (2.3) |
| Cd-$\alpha_3$D**IV**[b] | 27600 (487) | 1.07 | 78.1 (0.7) | 52.7 (4.3) | 4.2 (0.5) |

All the averaged values were determined from triplicate or duplicate experiments. [a]Calculated excess heat capacity. [b]pH 8.2. [c]pH 7.0.

**Table II-8**: Thermal denaturation parameters.

| Protein | $\alpha_3$D[a] | $\alpha_3$D**IV**[b] | $\alpha_3$D**IV**[a] | Hg-$\alpha_3$D**IV**[a] | Pb-$\alpha_3$D**IV**[a] | Cd-$\alpha_3$D**IV**[a] |
|---|---|---|---|---|---|---|
| **$T_m$ (°C)** | 89.6 (0.3) | **64.4 (0.6)**<br>76.1 (1.9) | **60.2 (0.1)**<br>82.2 (0.8)<br>89.9 (2.1) | 84.0 (1.7) | **83.4 (3.3)**<br>92.0 (7.7) | 78.1 (0.7) |
| **$\Delta H_{calc}$ (kcal mol$^{-1}$)** | 49.9 (4.6) | **50.0 (0.2)**<br>46.1 (1.4) | **44.9 (1.2)**<br>60.0 (0.5) | 65.1 (3.5) | **59.2 (6.9)**<br>104.2 (35.1) | 52.7 (4.3) |
| **$\Delta H_{van't\ Hoff}$ (kcal mol$^{-1}$)** | 41.2 (4.8) | **20.8 (3.1)**<br>15.8 (2.7) | **23.5 (2.1)**<br>8.5 (2.1) | 10.3 (3.1) | **14.6 (6.0)**<br>5.9 (1.5) | 12.6 (0.6) |

All the averaged values were determined from triplicate or duplicate experiments. Bolded values indicate primary denaturation species. [a]pH 8.2. [b]pH 7.0.

Significant shifts in the melting temperatures were observed for the metallated species of $\alpha_3$D**IV**, demonstrating that the metal-ligand complex provides extra stability against thermal denaturation. The melting curve of Hg(II)-$\alpha_3$D**IV** and Cd(II)-$\alpha_3$D**IV** fitted well to a one-peak model, whereas Pb(II)-$\alpha_3$D**IV** was fitted with two peaks. The $T_m$ values for the metallated species increased by a range of 18 – 24 °C, with Hg(II)-$\alpha_3$D**IV** exhibiting the largest $T_m$. This trend was also present in the enthalpy values, where an 8 – 20 kcal mol$^{-1}$ growth was observed. In addition, the thermogram of $\alpha_3$D at pH 8.2 was collected and compared with $\alpha_3$D**IV**. The $T_m$ value for $\alpha_3$D was determined to be 89.6 (0.3), which is consistent with previously reported values. This $T_m$ is ~30 °C higher than apo $\alpha_3$D**IV** but only about 6 – 12 °C greater than the metallated species. Nevertheless, the $\Delta H_{cal}$ values for apo $\alpha_3$D**IV** and $\alpha_3$D are within a similar range and, most importantly, the metallated $\alpha_3$D**IV** generated values that were 3 – 15 kcal mol$^{-1}$ greater than $\alpha_3$D. Overall, these stability studies using CD and DSC signify that binding a metal ion in the triscysteine binding site of $\alpha_3$D**IV** does not destabilize or unfold its structure. In fact, from the DSC analysis, we illustrate that metal binding further stabilizes the structure of $\alpha_3$D**IV**.

*XAS analysis of Hg-$\alpha_3$DIV*. XAS was utilized to determine the metrical parameters for mercury bound to $\alpha_3$D**IV**. Several spectra were collected on independent reproducible samples and spectra were averaged independently to verify reproducibility. The spectrum presented in Figure II-9 represents an average of the data for optimal signal/noise ratio. The X-ray absorption near edge structure (XANES) spectrum in Figure II-9 is consistent with those typically observed for Hg(II)-S complexes.[42] The Hg EXAFS could only be fit with a single Hg-S scattering environment constructed by a coordination number between 2.0 and 3.0 (±0.5) and centered at a

bond length of 2.36 Å (Figure II-10). There is no evidence for Hg-O/N scattering in this area. Long-range scattering could not be deconvoluted from noise signals in that region of the data. The averaged mercury-sulfur bond length in Hg-$\alpha_3$D**IV** is compared to relevant Hg(II)-S complexes, including protein systems and model compounds[43-58] in (Table II-9, Table II-10 and Figure II-11).

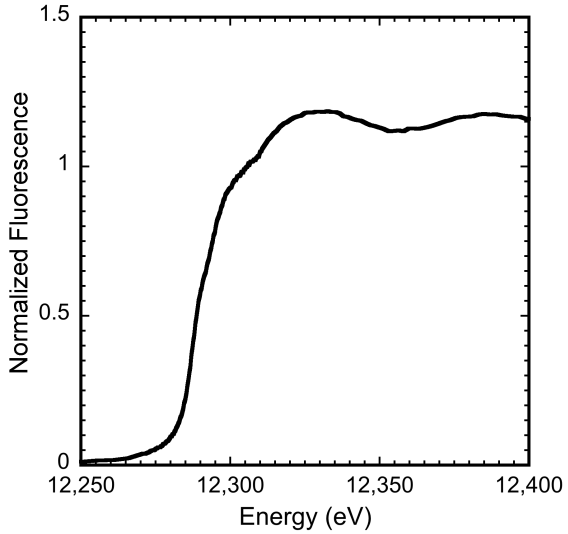



**Figure II-9**: Normalized XANES spectrum for average of all Hg-$\alpha_3$D**IV** spectra.

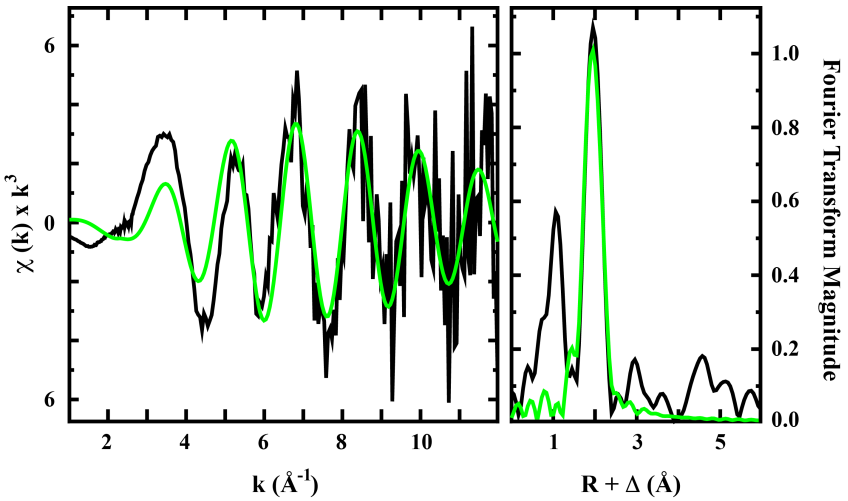



**Figure II-10**: EXAFS and Fourier Transform for Hg-$\alpha_3$D**IV**. (Left) Raw unfiltered EXAFS data (black) and simulations (green) for Hg-$\alpha_3$D**IV**. (Right) Fourier transforms of the raw EXAFS (black) and best fit simulation (green) for Hg-$\alpha_3$D**IV**. A noise peak at 1 Å from the Ge detector was not fitted in simulation.

**Table II-9**: Hg EXAFS fitting analysis of $\alpha_3$D**IV** compared to Hg-S bonds of relevant model compounds and proteins.

| Complex | CN | Geometry | Hg-S R (Å) | σ x $10^3$ (Å$^2$) | F | pH |
|---|---|---|---|---|---|---|
| Hg-$\alpha_3$D**IV**[a] | 2 | | 2.36 | 3.54 | 2.24 | 8.7 |
| | 2.5 | | 2.36 | 4.97 | 2.21 | |
| | 3 | | 2.36 | 6.20 | 2.21 | |
| Hg-TRIL16C[b] | 2 | | 2.324 | | | 5.5 |
| | 3 | | 2.443 | | | 9.5 |
| Hg-MerR[b] | 3 | | 2.43 | | | 7.5 |
| Average of 5 models[c] | 2 | Linear | 2.348 | | | |
| Average of 11 models[c] | 3 | Trigonal | 2.462 | | | |
| Average of 3 models[c] | 3 | Trigonal T-shaped | 2.497 Avg. 2 shortest bonds: 2.372 | | | |

[a]Data fit over a k range of 1 to 12 Å-1. Independent metal-ligand scattering environment at R < 3.0 Å. Scattering atoms: S (Sulfur). F = Number of degrees of freedom weighted mean square deviation between data and fit. [b]Hg(II)-S bond lengths determined from EXAFS. [c]Hg(II)-S bond lengths determined from an X-ray crystal structure.

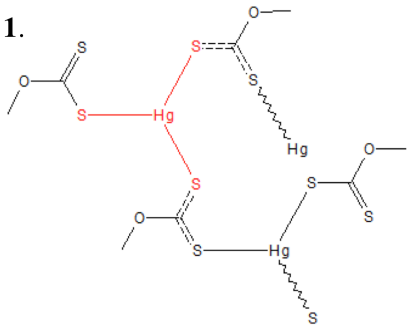
**Table II-10**: Metrical parameters mercury-sulfur model compounds.

| Compound | Hg-S R (Å) | Hg-S R (Å) | Hg-S R (Å) |
|---|---|---|---|
| **$HgS_3$ Trigonal T-shaped coordination** | | | |
| **1**. catena-(bis(O-methyldithiocarbonato-S)-mercury(ii)) | 2.365 | 2.383 | 2.924 |
| **2**. tetra-n-butylammonium (3-ethoxycarbonylthiolato-4,5-diphenylthiophene-2-thiolato)-(4,5-diphenylthiophene-2,3-dithiolato)-mercury(ii) | 2.373 | 2.388 | 2.495 |
| **3**. (thiocyanato)-bis(4-trimethylammoniobenzenethiolato)-mercury(ii) hexafluorophosphate | 2.353 | 2.369 | 2.823 |
| **Average 3 M-L** | | **2.497 (0.071)** | |

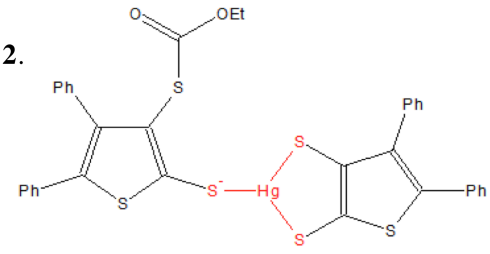
| | | | |
|---|---|---|---|
| **$HgS_3$ trigonal coordination** | | | |
| **4**. tetra-n-butylammonium tris(phenylthiolato-S)-mercury(ii) | 2.407 | 2.432 | 2.507 |
| **5**. catena-(ethylenediammonium tris(m2-sulfido)-sulfido-mercury-tin) | 2.396 | 2.435 | 2.653 |
| **6**. tetraethylammonium tris(t-butylthiolato-S)-mercury(ii) | 2.436 | 2.438 | 2.451 |
| **7**. tetrakis(n-propylammonium) tris((2,4,6-tri-isopropyl)benzenethiolato)-mercury methanol solvate | 2.398 | 2.46 | 2.47 |
| **8**. tetraphenylphosphonium tris(2,3,5,6-tetramethylbenzenethiolato-S)-mercury acetonitrile solvate | 2.397 | 2.404 | 2.493 |
| **9**. tetramethylammonium bis(tris(m2-thiobenzoato-O,S)-mercury(ii))-sodium | 2.443 | 2.468 | 2.485 |
| **10**. bis(tetraethylammonium) (m2-benzene-1,2-dithiolato-S,S')-bis(benzene-1,2-dithiolato)-di-mercury(ii) | 2.382 | 2.429 | 2.437 |
| **11**. tetraethylammonium tris(cyclohexylthiolato-S)-mercury | 2.403 | 2.455 | 2.487 |
| **Average 3 M-L** | | **2.449 (0.023)** | |
| **$HgS_2$ linear coordination** | | | |
| **12**. bis(carboxymethylthiolato)-mercury(ii) | 2.339 | 2.339 | |
| **13**. bis(n-pentanethiolato)-mercury(ii) | 2.304 | 2.304 | |
| **14**. catena-(bis(m2-bromo)-bis(m2-N,N-diethyldithiocarbamato-S,S')-di-mercury) | 2.364 | 2.385 | |
| **15**. bis(2-mercaptobenzoato-S)-mercury(ii) dioxane solvate | 2.363 | 2.363 | |
| **16**. bis(4-t-butylbenzenethiolato-S)-mercury(ii) | 2.358 | 2.363 | |
| **Average 2 M-L** | | **2.348 (0.028)** | |

The Hg(II)-S bond lengths represent one complex within a network.

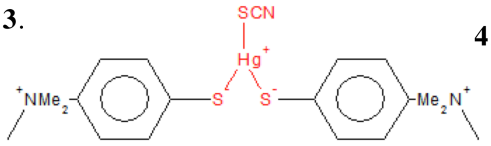
**1.**



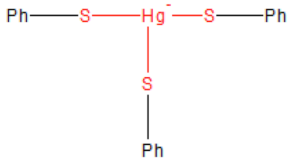
**2.**



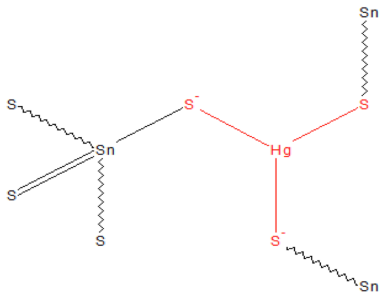
**3.**



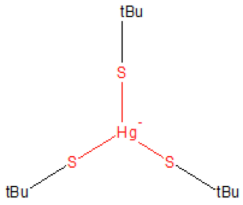
**4.**



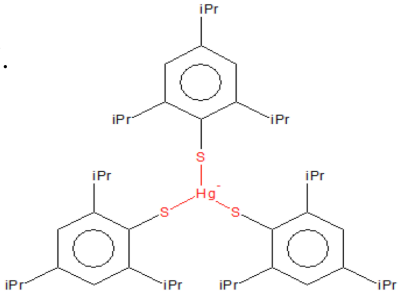
**5.**



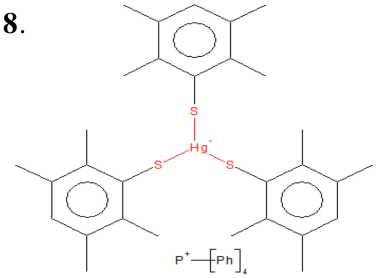
**6.**



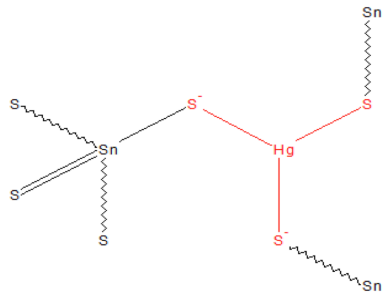
**7.**



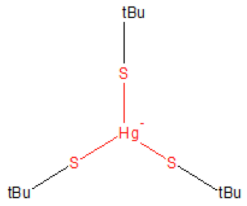
**8.**



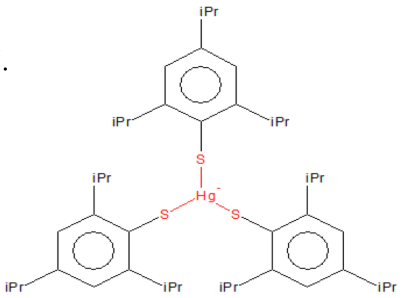
**5.**



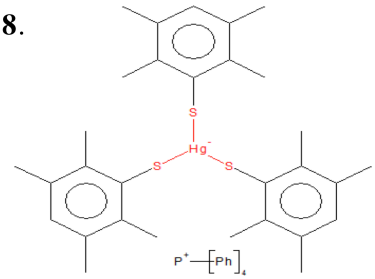

**6.**



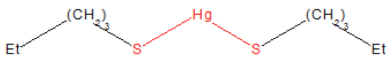

**7.**



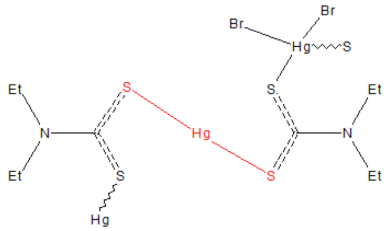

**8.**



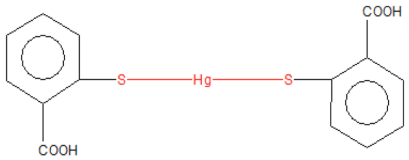

**13.**




**14.**




**15.**




**16.**

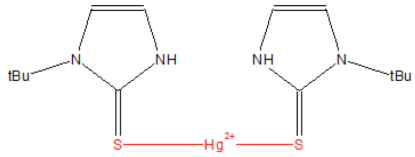



**Figure II-11**: ChemDraw representation of model compounds, which are labeled according to Table II-10.

**Discussion**

Peptide $\alpha_3$D**IV** was prepared to functionalize the framework of $\alpha_3$D by incorporating a metal binding site.[8] NMR and stability studies that examined the dynamic behavior and malleability of $\alpha_3$D revealed that the C-terminal end was the most amenable to mutations. In addition, this region of the bundle contained apolar residues that formed a hydrophobic "box" region, which offered a suitable location for a metal binding site. For these reasons, Leu residues were converted to Cys to produce a triscysteine motif and demonstrated its binding properties to the heavy metal ions Hg(II), Pb(II), and Cd(II). Subsequently, an $\alpha_3$D**IV** related construct that substituted the same leucine residues with histidine was prepared in order to form a $ZnN_3(H_2O)$ that served as an excellent analogue of carbonic anhydrase (Table II-1).[9] Despite these successes, it is critical to assess how amino acid substitution at this important position impacted the structure of the apo protein (e.g., was a stable, preformed metal binding site realized by these changes).

*Comparison to the structural statistics of $\alpha_3$D*. The structural statistics comparison of $\alpha_3$D**IV** and its parent structure, $\alpha_3$D, are found in Table II-11. There are notable differences in the NOE and dihedral angle restraints used in the structure calculation of both structures and the empirical origins of these restraints. The solution structure of $\alpha_3$D was based on 1143 NOE restraints,[8] incorporating 70 more NOE restraints than the structure of $\alpha_3$D**IV**. The majority of its NOEs for $\alpha_3$D were classified in the short-to-medium range (45%), whereas both structures have roughly the same percentage (~30%) of short-to-medium and long-range restraints. Unlike the structure of $\alpha_3$D, $\alpha_3$D**IV** included $\phi$, $\psi$ and $\chi1$ dihedral angle restraints that were generated from a TALOS-N analysis, which employed the chemical shift assignments of $^{1}H$, $^{1}H_{\alpha}$, $^{14}N$, $^{13}C$, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ atoms to predict dihedral angle restraints. $\alpha_3$D only contained $\phi$ angle restraints that were derived from a triple resonance HNHA experiment, as well as $\chi1$ angles, which were determined from NOE patterns. Moreover, the comparison of the backbone and heavy atoms root mean square deviation (RMSD) from the mean structure, for the same residue range and number of calculated structures, illustrated that the ensemble of the 13 lowest energy structures for $\alpha_3$D**IV** is lower than $\alpha_3$D. The RMSD for $\alpha_3$D**IV** is ~0.3 Å lower than the 13 structures reported for $\alpha_3$D,

for both the backbone and heavy atoms in the structured regions and the backbone atoms for residues 1-73.

The α-helical regions for the two structures slightly differ, with $\alpha_3$D (residues 4-21, 24-45 and 51-70) covering 2 more residues than $\alpha_3$D**IV** (5-20, 25-44 and 51-70) in helix 1 and 2. The structure of $\alpha_3$D incorporated Trp4, Leu21, Ser24 and Tyr45 in its structured regions, but these residues, from a TALOS-N analysis, were designated to have a coiled (Trp4) or loop (Leu21, Ser24 and Tyr45) secondary structure.

**Table II-11**: Structural statistics of $\alpha_3$D**IV** compared to $\alpha_3$D.

| **RMSD from the mean structure (Å)** | $\alpha_3$D | $\alpha_3$D**IV** | |
|---|---|---|---|
| Backbone (residues 1-73, N, $C^{\alpha}$, C) | 1.06[a] | 0.78 (0.15)[a] | 1.08 (0.31)[b] |
| Backbone (residues 4-21, 24-45, and 51-70, N, $C^{\alpha}$,C) | 0.75[a] | 0.48 (0.11)[a] | 0.53 (0.15)[b] |
| Heavy atoms (residues 1-73) | 1.61[a] | 1.25 (0.13)[a] | 1.58 (0.32)[b] |
| Heavy atoms (residues 4-21, 24-45, and 51-70) | | 0.92 (0.10)[a] | 0.99 (0.10)[b] |

[a]13 structures. [b]ψ angles excluded from the structure calculation.

*Comparison of the 13 lowest energy structures*. The 13 lowest energy structures of $\alpha_3$D**IV** were calculated in order to properly compare the minimized model to the reported structural statistics of $\alpha_3$D, and it was observed that the overall ensemble is better ordered. The backbone (N, $C_{\alpha}$, $C_O$) and heavy atom RMSD values for the 13 structures of $\alpha_3$D**IV** were observed to be lower by ~0.3 Å, even though the $\alpha_3$D contained more intraresidue and short-to-medium range NOE restraints (Table II-11). This difference was attributed to the dihedral angle and NOE sequential restraints used in the structure calculations. Specifically, the $\alpha_3$D**IV** structure contains ϕ and ψ backbone dihedral angles, while $\alpha_3$D only incorporated ϕ angle restraints. The presence of both dihedral angles limit the range of interaction between $N_i$-$N_{i+1}$ and $C_i$-$C_{i+1}$ atoms in the structured regions, thereby eliminating unfavorable sterics and could result to better ordered α-helical chains. This was tested in an analysis that involved removing ψ angle restraints in the structure calculation of $\alpha_3$D**IV**. It was determined that the RMSD from the mean structure values of the backbone and heavy atoms for residues 1-73 in $\alpha_3$D**IV** increased to 1.08 (0.31) from 0.78 (0.15) Å and to 1.58 (0.32) from 1.25 (0.13), respectively, which are almost equal to the same

parameters determined for $\alpha_3$D. However, for the structured regions, a very modest increase of 0.02 Å was observed, which is well within the statistical error (~± 0.10 Å).

In addition, the increased unity in the 13 structures of $\alpha_3$D**IV** can be further justified by the amount of sequential NOE restraints between $^1$H backbone atoms used in the calculations. The $\alpha_3$D structure has a proper extent of $H^N_i$-$H^N_{i+1}$ and $H^\alpha_i$-$H^\beta_{i+3}$ sequential NOE correlations in all three helices and strong $H^\beta_i$-$H^N_{i+1}$ NOE correlations in helix 3 for residues 50-70. Nevertheless, it significantly lacked $H^\alpha$-$H^N_{i+1}$, $H^\alpha$-$H^N_{i+3}$ and $H^\beta_i$-$H^N_{i+1}$ in helix 1 and 2 correlations, NOEs that are typically observed in α-helical structures since the distances between these $^1$H atoms are within the experimental limit of 5 Å. On the other hand, in the structured regions, $\alpha_3$D**IV** contained most of the NOE correlations mentioned above in a sequential manner. Additionally, the sequential NOE pattern in the structure of $\alpha_3$D**IV** can be viewed as compensating restraints since only a modest increase in the RMSD values were observed in the structured regions after ψ dihedral angle restraints were removed.

*Comparison to the lowest energy structure of $\alpha_3$D*. The lowest energy structure of $\alpha_3$D**IV** and $\alpha_3$D are superimposed in Figure II-12A. The two structures were aligned via backbone (N, $C^\alpha$, C, O) atoms of residues 1-73, and this alignment yielded an overall heavy atom RMS value of 1.75 Å. From the overlay, $\alpha_3$D**IV** possesses the same overall topology as $\alpha_3$D, with a counterclockwise bundle that was expected in the design process of $\alpha_3$D. Figure II-12B is an overlay of the Leu-to-Cys mutation site, positions 18, 28 and 67, of both structures. The $C_\beta$ distances between the Leu-Leu and Cys-Cys were measured and the triangular plane that forms between these residues has an area of 15.5 $Å^2$ in $\alpha_3$D and 14.1 $Å^2$ in $\alpha_3$D**IV**. This slight deviation demonstrates that the fold at the C-terminal end was not significantly affected even after removing Leu residues that provide stabilizing and packing interactions in $\alpha_3$D. Regardless, incorporating Cys residues with polar, uncharged thiol groups to form a metal-binding site in $\alpha_3$D**IV** led to two significant structural differences as a result of a more packed core.

First, the $\alpha_3$D**IV** structure has lower $\Omega_{1,2}$, $\Omega_{2,3}$ and $\Omega_{1,3}$ angles by 16°, 15° and 7°, respectively. This deviation from $\alpha_3$D could stem from better packing of the apolar layers above the triscysteine site (Figure II-13). The first layer, which is composed of a plane between the $C_\beta$ atoms of Ile14, Phe31 and Ile63, has an area of 12.9 $Å^2$ in $\alpha_3$D**IV**, whereas in $\alpha_3$D this plane is 9.0 $Å^2$ greater. This trend is also observed in the two subsequent layers 2 (Leu11, Ile35 and Ala60) and 3 (Phe7, Phe38 and Leu56), which has an area of 7.8 $Å^2$ (16.1 $Å^2$) and 7.3 $Å^2$ (17.7 $Å^2$) lower than $\alpha_3$D.

Next, it was observed that the incorporation of Cys residues did not disrupt the overall α-helical framework of $\alpha_3$D, as demonstrated in the $^{15}$N-TROSY and CD studies. Unexpectedly, the alignment of $\alpha_3$D and $\alpha_3$D**IV** revealed that the addition of a Cys residue in the 28$^{th}$ position in helix 2 improved its helical content (Figure II-12C). In $\alpha_3$D, the helicity in the second strand breaks after Ala26 and becomes continuous again at Ala29. Residues Glu27 and Leu28 appear to have a more coil-like secondary structure. In contrast, this disruption in the helicity between residues 26-28 is not observed in the structure of $\alpha_3$D**IV**, which demonstrates an increase in the α-helical content. Additionally, the analysis from two web-based structural determination programs, TALOS-N and VADAR (Volume Area Dihedral Angle Reporter),[59] as well as the CD comparison of apo $\alpha_3$D**IV** with $\alpha_3$D supports this observation. TALOS-N uses a database of 580 proteins with almost complete backbone NMR chemical shifts ($^1$H, $^1H_\alpha$, $^{14}$N, $^{13}$C, $^{13}C_\alpha$ and $^{13}C_\beta$) and an additional database of 9,523 high-resolution X-ray crystal structures (without experimental chemical shifts) to predict dihedral angles ϕ, ψ, and χ1, as well as the secondary structures from experimental NMR chemical shifts. Similar to the dihedral angle prediction, the TALOS-N secondary structure classification system involves a two-system artificial neural network (ANN) and categorized residues 26-28 into an α-helical secondary structure, with a probability score of ~0.96. VADAR analysis on the PDB coordinates of the lowest energy structure of $\alpha_3$D**IV** structure also characterized residues 26-28 to be in an α-helical structure, while the same analysis for $\alpha_3$D resulted in a coiled secondary structure. Furthermore, the molar ellipticity value for apo $\alpha_3$DIV is about -4200 deg $cm^2$ $dmol^{-1}$ $res^{-1}$ greater than $\alpha_3$D. This again indicates that the replacement of Leu residues with a Cys at the 28$^{th}$ position may have increased the overall helical content of the $\alpha_3$D framework.

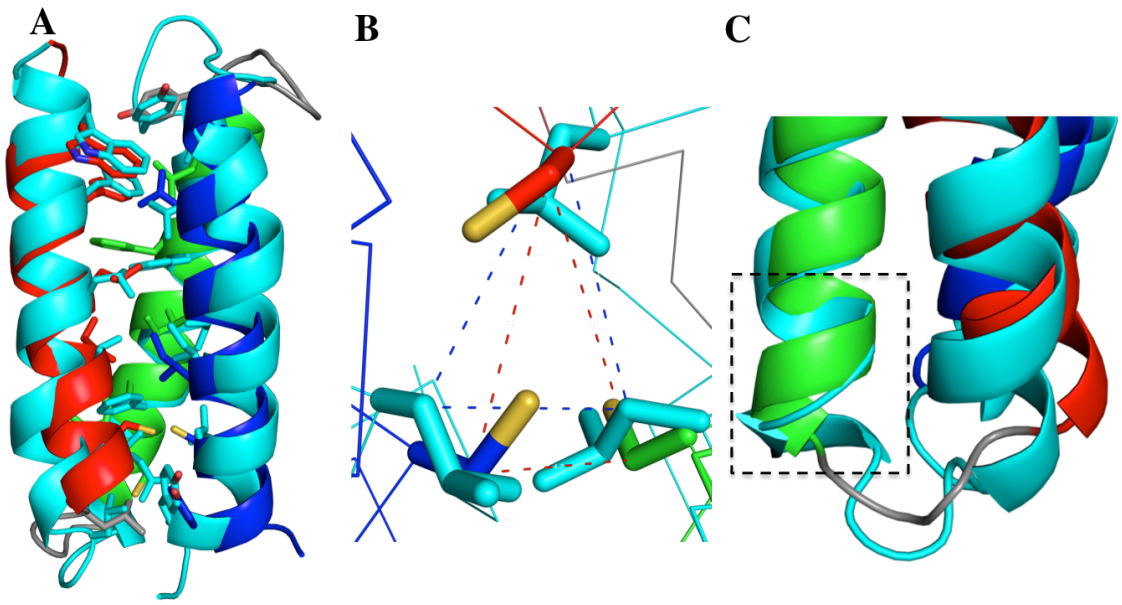



**Figure II-12**: A) An overlay of $\alpha_3$D**IV** (helix 1: red, helix 2: green & helix 3: blue) and $\alpha_3$D (cyan). The backbone (N, $C^{\alpha}$, C,O) rms was determined from PYMOL to be 1.75 Å. B). Top-down view of the mutation site (18, 28, and 67), displaying superimposed Cys/Leu residues. C) Gain of helical content in helix 2 for residues 26-28.

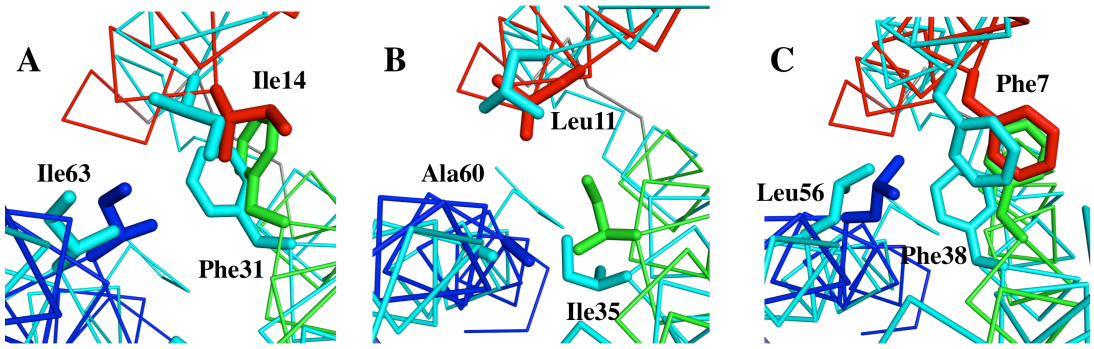



**Figure II-13**: A) First layer of apolar groups ($\alpha_3$D**IV** red: helix1, green: helix2, and blue: helix3; $\alpha_3$D: cyan) above positions 18, 28, and 67, which involves Ile14, Phe31, and Ile63. B) Second layer of apolar groups, Leu11, Ile35, and Ala60. C) Third layer of apolar groups, Phe7, Phe38, and Leu56.

*Preorganization of the triscysteine site*. Native metalloproteins have evolved to possess metal coordination environments that are unique to their function, and two limiting extremes of metal binding sites have been described. The first group is typified by zinc finger proteins, which adopt a stable fold only after a metal has bound to the endogenous ligands of the sequence.[60] In this case, metal binding defines protein structure. Less extreme, but still within this category, are cases where metal binding leads to important protein conformational changes that define function. Examples of such behavior can be seen with metalloregulatory proteins such as MerR

or ArsR.[1,2] The second group is illustrated by the cupredoxins that have a well-defined, pre-organized binding site. The concept of a racked-induced bonding or entatic state model has been proposed for such systems to elucidate the unique properties of these electron transfer (ET) proteins.[61] The rack-induced bonding model explicates that the ligand environment ($CysHis_2$ site and one or two axial weakly bound axial ligands) in a cupredoxin fold forces the copper ion, regardless of the oxidation state, into a strict distorted-tetrahedral geometry. This feature in cupredoxins and redox-active metalloproteins was discovered to be one of the fundamental driving forces in attaining efficient electron transfer activity. Ultimately, it is of great importance in the *de novo* design of metallopeptides to be able to predict accurately the extent of metal binding site preorganization prior to the introduction of a metal ion in order to achieve the desired affinity, selectivity and function.

The sequence of $\alpha_3$D was designed to contain diverse sets of heptad repeats (defined as "abcdefg"), where nonpolar residues are incorporated at the "a" and "d" positions to provide core stabilizing interactions and charged residues at the "e" and "g" to form salt-bridges between helices.[9] Previous work on *parallel* three-stranded coiled-coil (TRI and Coil-Ser) systems showed that the subtle difference between the "a" and "d" positions can produce distinctive outcomes in heavy metal binding affinity and geometry, which can be attributed to the preorganization of the sulfur ligands prior to metal binding.[62, 63] It was determined that the substitution of Leu to Cys residues in "a" sites tend to favor the formation of a trigonal metal binding site, where three $S_\gamma$ atoms point into the core (Figure II-14A). For instance, apo CSL9C (PDB 3LJM)[64] comprise of "a" site $S_\gamma$ atoms with distances of 3.3 and 3.4 Å. Conversely, "d" site substitutions orient $S_\gamma$ atoms away from the core and towards the helical interface forming a much larger cavity (Figure II-14B). Apo CSL19C (PDB 2X6P)[64] contains "d" site Cys residues with $S_\gamma$ - $S_\gamma$ distances of 3.4 and 4.6 Å. Initial examination of the $\alpha_3$D NMR structure indicated that a preorganized metal binding site could be carved into a hydrophobic box region of the protein. Therefore, it was for these structural reasons that leucines in the "a" positions of $\alpha_3$D were targeted for modification; however, one should recognize that the side chain orientations are expected to differ in this antiparallel $\alpha_3$D**IV** framework.

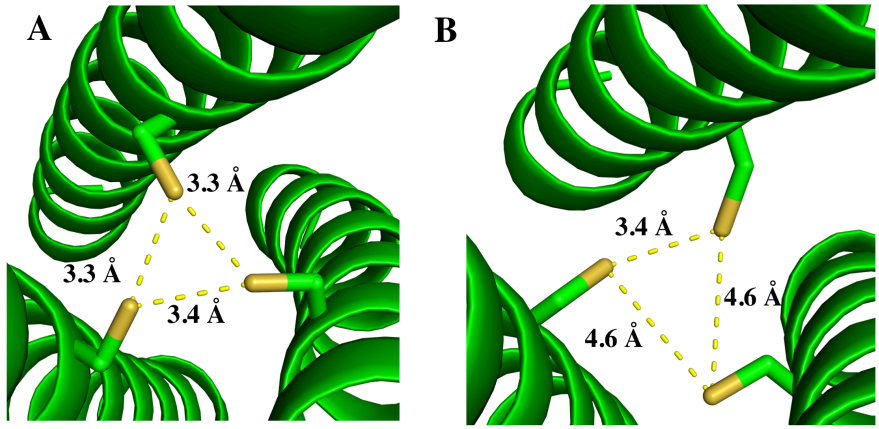



**Figure II-14**: A) "a" site Cys in apo CSL9C (PDB 3LJM),[64] in which the Sγ atoms position inside the core. The alternate conformation of the Sγ atoms was removed in order to illustrate the orientation of "a" site Cys residues. B) "d" site Cys in apo CSL19C (PDB 2X6P),[64] which contain Sγ atoms that point at the interhelical interface.

In $\alpha_3$D**IV**, "a" site Cys residues were incorporated and predicted that the $S_\gamma$ atom of Cys18 and Cys67 could orient into the core to initiate the generation of a preformed trigonal plane of sulfurs starting with these two ligands. Cys28 was intended to provide a more adaptable ligand, which was predicted to orient towards the C-terminal end since this residue is on the antiparallel strand. The main objective with the design of $\alpha_3$D**IV** was to create a single-polypeptide scaffold that produces a more asymmetric or a distorted triscysteine metal binding site than the symmetric Cys sites found in TRI or Coil-ser analogues. As a result, the more native-like properties of $\alpha_3$D**IV** could provide a better model for native metalloregulatory proteins or metallochaperones that sequester toxic heavy metals in nature.

The triscysteine metal binding site of the 20 lowest energy structures of apo $\alpha_3$D**IV** is superimposed in Figure II-15A. Overall, the 20 conformations of the three Cys residues illustrate a well-defined metal binding site. For instance, the $C_\beta$ and Sγ atoms of Cys28 are uniformly oriented towards the C-terminal end of the bundle (Figure II-15A, side-on view), while the $S_\gamma$ atoms of Cys18 are pointing to the interhelical interface between helix 1 and 3 in 18 of the 20 structures. It should be pointed out that the Cys18 χ1 dihedral angle restraint was not constrained but this corresponding angle were restrained in Cys28 and Cys67. Additionally, the $H_\beta$ atoms of Cys18 have five NOE interactions, four and one intra- and inter-helical (Phe31)

NOEs, respectively. The corresponding atoms of Cys28 contain a single intra-helical NOE and three inter-helical (Cys18, Phe31, and Arg71) NOEs; whiles, the $H_\beta$ atoms of Cys67 only possess two intra-helical NOEs. Despite this description of the averaged structure of the Cys residues, it is proposed that the site is not rigid, because the amide proton $^{15}$N-TROSY signals for such residues and their sequential neighbors are ~ 30% more intense than the average, indicating motion on the nano-second timescale for these residues. With this caveat, the (time) averaged structure of the metal binding site is described below.

The sulfur atoms of Cys28 and Cys67 are observed to have set orientations, while Cys18 contains two conformations that diverge from its ensemble. The $C_\beta$ and Sγ atoms of Cys28 on the antiparallel strand are directed towards the C-terminal end of the bundle (Figure II-15A, side-on view). This orientation was also predicted to occur for the $C_\beta$ atoms of non-coded D-amino acids, such as D-Pen or D-Leu, in 3SCC analogues.[65, 66] The change in chirality in these two D derivatives were expected to direct their respective thiol and isopropyl moeity towards the C-terminal end. Consequently, it is shown here in the $\alpha_3$D**IV** structure that an L-Cys in an antiparrallel strand provides the same effect as a D-amino acid. For Cys67, its $S_\gamma$ atoms are oriented inside the core of the bundle, which was predicted for this "a" site substitution. Lastly, the collective $S_\gamma$ atoms of Cys18 result in a more "d" site-like orientation, where the $S_\gamma$ atoms are directed at the interhelical interface between helix 1 and 3. Cys18 possesses two structures (16 and 18) that deviate from the ensemble. In structure 14, the "d" site conformation of the Sγ atom of Cys18 is oriented towards the opposite direction, at helix 1 and 2. Structure 16 contains a Cys18 $S_\gamma$ atom that points inside the core, providing a structure that presents a metal binding site with two "a" site ligands at Cys18 and Cys67 (Figure II-15B). When compared to apo CSL9C (Figure II-15C), which contain "a" site residues (with two conformations), the $C_\beta$ and $S_\gamma$ atoms of Cys18 and Cys67 in structure 16 overlay well with two Cys9 residues. As a result, structure 16 demonstrates a partially preformed trigonal cavity for metal ions that coordinate in a 3 or 4 coordination environment, such as Hg(II), Pb(II) and Cd(II).

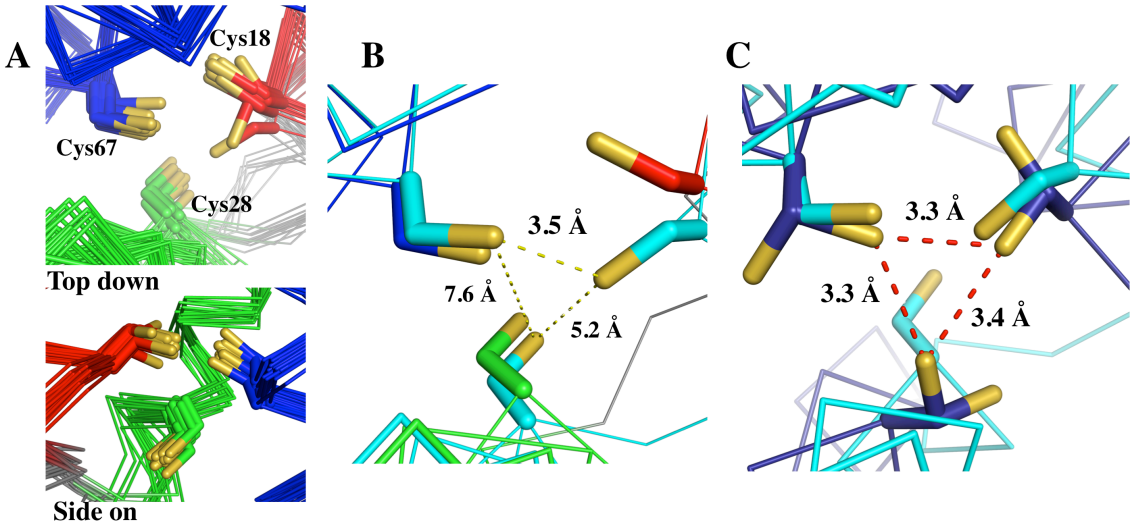



**Figure II-15**: A) Preorganization of the triscysteine metal binding site of the 20 lowest energy structures (Cys18: red, Cys28: green and Cys67: blue). B) A partially preformed trigonal site for metal binding in structure 16 (cyan) of 20 aligned with the lowest energy structure, structure 1. The Sγ – Sγ distances are measured for structure 16 of $\alpha_3$D**IV**. C) Structure 16 (cyan) aligned with the Cys "a" sites of CSL9C (PDB: 3LJM) (blue).[61] The Sγ – Sγ distances are measured for CSL9C.

*Implications on metal binding*. With a solution structure in hand of apo $\alpha_3$D**IV**, one may next consider the consequence of metal binding to the system. The apo structure allowed us to investigate the global, as well as, local structural changes that occur when a protein scaffold is functionalized with Cys residues to form a metal binding site. A metallated structure would afford a concise comparison with the apo structure to assess the extent of preorganization of the site, which could provide a mechanistic model for metal ion acquisition in native systems. In terms of *de novo* protein design, both structures offer a base for future design of a metal binding site with mixed residues/ligands or with identical residues that require secondary interactions (such as hydrogen bonding) in able to achieve redox or catalytic activity with transition metals Fe(II/III), Cu(I/II) or Zn(II).

The spectroscopic properties of Hg(II), Pb(II) and Cd(II) bound $\alpha_3$D**IV** were previously reported[8] to form the predicted complexes but a much closer inspection of the solution stability of metallated $\alpha_3$D**IV** was necessary to illustrate that the addition of a metal ion does not have a unfavorable effect on its fold. The $^{15}$N-HSQC spectra of metallated $\alpha_3$D**IV** were collected and

performed circular dichroism and thermal denaturation studies to illustrate that heavy metal binding reinforces and improves the solution stability of $\alpha_3$D**IV**. In addition, the Hg-$\alpha_3$D**IV** species was analyzed by XAS to obtain structural metrics on a metallated species.

Circular dichroism studies have been a central tool in investigating and confirming metal-induced (transition and/or heavy metal ions) folding in native proteins, such as zinc finger proteins[67] and calmodulin,[68] as well as in *de novo* designed peptides including isoleucine zipper[69] and BABY[70] 3SCC peptides. The CD fingerprint of α-helical proteins and peptides exhibit two minima at 208 and 222 nm, which arise from $\pi$-$\pi^*$ and n-$\pi^*$ transitions in the amide atoms.[71] When compared to apo $\alpha_3$D**IV**, the CD spectra of the metallated species also contain a double band profile at 208 and 222 nm (Figure II-8A). These two absorption bands are a characteristic feature in α-helical folds. The calculated molar ellipticity values ($[\theta_{222}]$) for the metallated species fall within the range of the apo and these large negative [θ] values are typical for a well-folded α-helical fold. Furthermore, the CD results show that the apo form has a high α-helical content, which was the intended target in the original design of $\alpha_3$D. In this peptide system, metal binding to the C-terminal end of the bundle was not necessarily predicted to increase the helicity of the peptide because the overall fold is highly α-helical. Metal binding was expected to perturb that region of the peptide (positions 18, 28 and 67 at the C-terminal end), resulting in a slight decrease in the helicity of the metallated species. This corroborates well with DSC results (which is discussed subsequently) because, although a lost in helicity was observed, metal binding provides M-thiolate bonds that enhance the tertiary stability of the structure, thus much higher melting temperatures when compared to the apo. In addition, the replacement of Leu to Cys residues, to build a metal binding site, removes stabilizing hydrophobic interaction within the $\alpha_3$D fold and binding a metal ion into this site counters this loss of stability.

A thermal denaturation analysis was performed to further assess the stability of metallated $\alpha_3$D**IV** and to explore the concept that metal binding to proteins could lead to greater stability. This rise in thermal stability has been previously observed in ligand-binding studies on native proteins. For example, the melting temperature ($T_m$) of a copper metallochaperone BsSco was determined to increase by 23 °C when bound to Cu(II).[72] Ca(II) and Zn(II) ion binding to

coagulation factors IX binding protein increased its melting temperature by 4 – 5 °C and calculated enthalpic parameter ($\Delta H_{cal}$) by more than 100 kcal mol$^{-1}$.[73] The DSC results illustrate that metal bound $\alpha_3$D**IV** is more stable than its apo counterpart in thermally induced denaturations. The metallated species have much higher melting temperatures, by a range of 18 – 24 °C, and the calculated enthalpic parameters ($\Delta H_{cal}$) are greater by 8 – 20 kcal mol$^{-1}$ (Table II-7). The enhancement of both thermodynamic values shows that the metal-thiolate bonds (Cd-S, Hg-S or Pb-S) supplements the weak-nonbonding forces (hydrophobic interactions, hydrogen bonds, and salt-bridges) that are essential in protein folding and disruptive during the denaturation process. In addition, the $T_m$ value of parent structure $\alpha_3$D is still greater than apo but slightly higher than the metallated $\alpha_3$D**IV** species, which indicates that the incorporation of a triscysteine site in its apo or metal-bound form cannot entirely reproduce the hydrophobic interactions that are provided by Leu residues in a protein fold. This finding is further supported by a chemical denaturation performed on apo $\alpha_3$D**IV**, where a 2.5 kcal mol$^{-1}$ decrease was observed in the $\Delta G_{UF}$ from the reported value of $\alpha_3$D (5.0 kcal mol$^{-1}$).[8] Moreover, the $\Delta H_{cal}/\Delta H_{van't\ Hoff}$ ratios reveal the thermal-induced denaturation process and a value of 1 indicates a two-state model. $\alpha_3$D has a ratio closest to this model of 1.2, which matches well with the reported value. Apo and metallated $\alpha_3$D**IV** have values > 2.0 and the denaturation curve of these species are much broader than the corresponding curve of $\alpha_3$D, demonstrating a self-associating unfolding process. This result was not unexpected as the metal binding site contains three thiolate ligands that could induce the formation of external disulfide bonds at higher temperatures.

The transition from a two-coordinate (2 C. N.) [Hg(II)$S_2$(SH)] to a three-coordinate (3 C. N.) [Hg(II)$S_3$]$^-$ complex in $\alpha_3$D**IV** has an apparent p$K_a$ of 7.1,[8] and this transition was examined by $^{199}$Hg-NMR and UV-vis. It was determined that $\alpha_3$D**IV** forms a linear [Hg(II)$S_2$(SH)] 2 C.N. below pH 6.0 and a 3-coordinate trigonal [Hg(II)$S_3$]$^-$ complex above pH 8.5 with a single chemical shift at -938 and -244 ppm, respectively. Further, under intermediate pH conditions (~pH 7.5), two chemical shifts were observed at -240 and -926 ppm, which demonstrate a mixture of both species. The 3 coordinate Hg-$\alpha_3$D**IV** species (at pH 8.7) was further analyzed with XAS to obtain structural parameters on metallated $\alpha_3$D**IV**, as well as to confirm previous

findings. The EXAFS analysis on the Hg(II)-$\alpha_3$D**IV** is dominated by nearest-neighbor scattering typical of sulfur ligation (Figure II-10) and yielded an average Hg(II)-S bond length at 2.36 Å and a 2.5 C. N. This bond length falls in between reported values for 2 C. N. and 3 C. N. mercury-sulfur complexes. Therefore, in able to correlate this EXAFS analysis to the coordination environment of the $[Hg(II)S_3]^-$ complex of $\alpha_3$D**IV**, these results were compared to the Hg(II)-S distances of protein and *de novo* designed peptide systems derived from EXAFS analysis and model compounds measured from X-ray crystal structures in Table II-9. The native protein, MerR, which was determined to form a trigonal mercury complex, has an EXAFS derived Hg(II)-S bond length at 2.43 Å.[4] Similarly, the EXAFS analysis on Hg(II)-TRIL16C provides Hg(II)-S distances of 2.32 and 2.44 Å for the corresponding 2 and 3 coordinate species. Table II-9 also compares the EXAFS result to 2 C. N and 3 C. N. compounds from the Cambridge Structural Database (Table II-10).[74] The five model compounds with a 2 C. N. have an average Hg(II)-S bond length of 2.348 (0.023) Å.[54-58] The three-coordinate compounds (11 total) produced an average Hg(II)-S distance of 2.462 (0.044) Å.[43-53] Three of these compounds exhibit a T-shaped coordination environment, which consists of two short bonds with an average bond length of 2.372 (0.01) Å and a single longer bond at 2.497 (0.071) Å.[43-45] These structure based Hg(II)-S bond lengths corroborate well with the EXAFS analysis of Hg(II)-TRIL16C and Hg(II)-MerR.

The two most straightforward interpretations of the observed 2.36 Å value for Hg-$\alpha_3$D**IV** is either an approximately equimolar mixture of 2 and 3 coordinate mercury sites or a single, distorted T-shaped structure similar to those of small molecule models. The mixed speciation behavior was previously observed in an EXAFS analysis of Hg(II)-TRIL16C. A mixture of 2 and 3 C. N. species has a 2.37 Å Hg(II)-S bond length. However, this mixture of 2 and 3-coordinate Hg(II)-TRIL16C[75] simultaneously show $^{199}$Hg-NMR resonance peaks at ~-180 ppm and ~-830 indicative of both species, whereas Hg-$\alpha_3$D**IV** has a single resonance at -244 ppm. Furthermore, the apparent p$K_a$ for the formation of $Hg(II)(TRIL16C)_3$ was calculated to be 7.8, and a 9.5 pH condition was used to obtain a pure 3 C. N species in the EXAFS analysis. This p$K_a$ value is nearly 1 log unit higher than what was determined for the $[Hg(II)S_3]^-$ complex of Hg(II)-$\alpha_3$D**IV**. Thereby, the pH condition of 8.7, which is 1.6 log unit higher than its apparent p$K_a$, favors a 3 C.

N. species instead of a mixture with a 2 C. N. complex as observed in Hg(II)-TRIL16C. Finally, the absorbance values in the ultraviolet between 2 – 4 coordinate mercury-sulfur chromophores are distinctive, where $Hg(SR)_2$ complexes have a $\lambda_{max}$ band <220 nm and $Hg(SR)_3$ and $Hg(SR)_4$ compounds exhibit absorption $\lambda_{max}$ bands between 230 – 340 nm.[76] The absorption characteristics of Hg(II)-$\alpha_3$D**IV** at pH 8.6 is more consistent with a 3 C. N. complex. For these reasons, it is concluded that the best interpretation of the XAS data collectively with the $^{199}$Hg NMR and UV-vis is to assign a T-shaped geometry with two short Hg(II)-S bonds and one longer (~2.8 – 3.0 Å) bond.

Together, the results from the CD and DSC studies demonstrate that metal binding did not disrupt the overall secondary structure of $\alpha_3$D**IV**; instead it provides further stability to the framework. In addition, the EXAFS analysis validates Hg(II) binding to the triscysteine site and provides structural metrics on the $[HgS_3]^-$ complex. Therefore, the loss of cross peaks in the $^{15}$N-HSQC spectrum of Pb(II)- and Hg(II)-$\alpha_3$D**IV** is not due to an unfolding of $\alpha_3$D**IV**. Instead, it indicates that metal binding quenches the more rapid dynamics of the metal-site (see above) and selects conformations that interconvert on a milli-microsecond time scale, leading to line-broadening and loss of the corresponding $^{15}$N-HSQC NMR cross peaks. Nonetheless, in the case of Pb(II)-$\alpha_3$D**IV**, 57 of the 68 resonance peaks were still identified and the perturbation of many residues, such as K8, Cys18, Phe38, Ala44 and Val53, can be identified. The missing peaks are due to residues localized near the metal binding site or at the C-terminal end, which was already determined to be more dynamic than the rest of the fold in $\alpha_3$D. For example, the chemical shift change for residues that are located at this end of the bundle, involving Gln19, Ser24, Ala29, Gln68 and Tyr70, cannot be easily assigned suggesting severe line broadening or overlap with a neighboring signal. For Hg(II)-$\alpha_3$D**IV**, the $^{15}$N-HSQC spectrum has 7 less peaks than Pb(II)-$\alpha_3$D**IV** and has 8 less than its apo counterpart. In contrast to the Pb spectrum, the loss in peaks seems to be global and cannot be assigned to any specific region of the sequence. Like the Pb spectrum, peak loss can be attributed to Hg(II) binding but the more dramatic effect could be largely associated with the pH environment. As pH increases from 6 to 9, the average lifetime of the exchangeable hydrogen groups, including backbone amide protons, increases from ms to μs range,[77] thereby decreasing the observable resonance peaks in a $^{15}$N-HSQC spectrum. This pH

effect is exhibited in the comparison of the $^{15}$N-HSQC spectra of apo $\alpha_3$D**IV** at pH 5.8 and 8.6 (Figure II-16). The higher pH spectrum contains 13 fewer peaks than the lower pH, demonstrating that the significant loss in cross peaks in the Hg(II)-$\alpha_3$D**IV** spectrum is mainly due to the pH condition.

An alternative explanation for the significant change in $^{15}$N-HSQC results for the metallated species could stem from the asymmetric nature of the triscysteine site leading to structural dynamics on micro-milli second time scale resulting in line broadening and peak loss. The side-on view of the triscysteine site of structure 1 in Figure II-15A shows that the $C_\beta$ and $S_\gamma$ atoms of Cys28 points toward the C-terminal end of the bundle. Upon metal binding, Cys28 requires a rotation at the N-$C_\alpha$-$C_\beta$ bonds of about 56° towards the N-terminal end in order to achieve the proper trigonal plane for Hg(II) or Pb(II). In the same manner, the "d" site-like conformation of Cys18 will have to adopt an "a" site conformation as seen in structure 16 of $\alpha_3$D**IV** or translate towards the core (Figure 6B). To further illustrate this process, apo $\alpha_3$D**IV** was overlaid to the X-ray crystal structures of metallated *de novo* designed peptides, As(III)(CSL9C)$_3$ (at 1.81 Å resolution)[78] and [Hg(II)]$_S$[Zn(II)($H_2O$/$OH^-$)]$_N$(CSL9PenL23H)$_3^{n+}$ (at 2.20 Å resolution)[79] in Figure II-17.

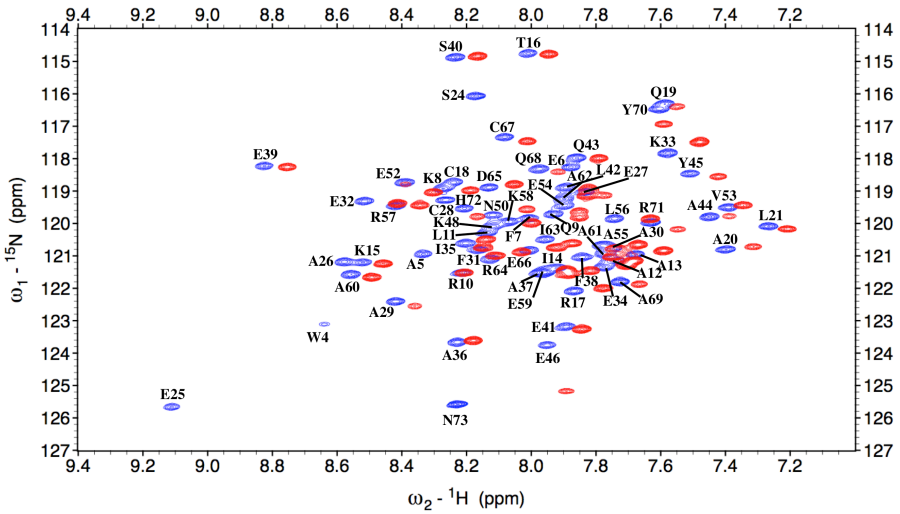



**Figure II-16**: Enlargements of $^{15}$N-HSQC spectra of apo $\alpha_3$D**IV** pH 5.8 (blue) and pH 8.6 (red). Spectra were collected on a 500 MHz Varian VNMRS NMR instrument at 25 °C. The pH 5.8 spectrum contains 64 of the 68 total peaks and the assignments are adjacent to its peak, whereas the pH 8.6 spectrum has 54 of the 55 identified peaks (not assigned). The peaks at pH 8.6 shift upfield of pH 5.8 and also contain 15 less identifiable peaks, demonstrating a pH effect on the hydrogen exchange rates.

Similar to Pb(II), As(III) has a sterochemically active lone pair of electrons and binds to a triscysteine environment in a trigonal pyramidal geometry. There are several possible orientations for the As(III) ion, with the most symmetric being *endo* and *exo*.[80] In the energetically preferred *endo* conformation, the As(III) ion and all three $C_\beta$ atoms are on the same side of the plane of S atoms. In the *exo* conformation, As(III) and the $C_\beta$ atoms are on opposite sides of this plane. It is also possible to have a mixed conformation, where one or two $C_\beta$ atoms are on the same side and one or two are on the opposite side. The crystal structure of As(III)$(CSL9C)_3$ reveals an *endo* conformation (Figure II-17A). For Pb(II)-$\alpha_3$D**IV** in solution, it is possible that the Pb atom forms a distorted, asymmetric triscysteine site that undergoes a exchange between these conformations; thereby, influencing the chemical shift dispersion in the $^{15}$N-HSQC spectrum. In addition, the $^{207}$Pb-NMR spectrum of $^{207}$Pb(II)-$\alpha_3$D**IV** contained no resonances in the range observed for the more symmetric Pb(II)-$S_3$ complexes in Coil-Ser, BABY and GRAND parallel three-stranded coiled coils.[81] UV-vis studies confirm that Pb(II) does bind in a trigonal pyramidal geometry,[8] so the absence of the $^{207}$Pb NMR signal supports the idea that Pb(II) is in a very dynamic environment. $^{207}$Pb-NMR is much more sensitive to the motion in the coordination environment than UV-vis spectroscopy and can provide insight into the dynamic behavior of a metal-ligand complex.

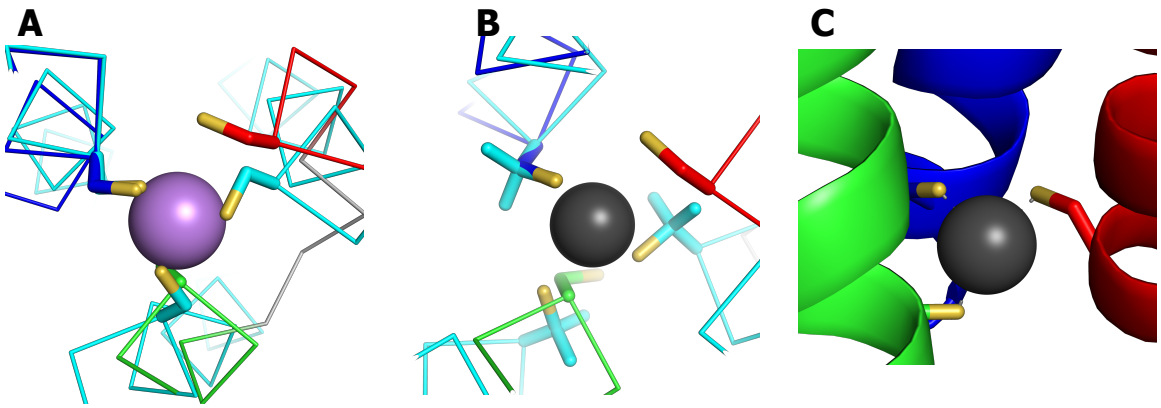



**FIGURE II-17**: A) Structure 1 of $\alpha_3$D**IV** superimposed over the As-$Cys_3$ site of $As(CSL9C)_3$ (PDB 2JGO at 1.81 Å resolution).[78] The $C^{\alpha}$ and $C^{\beta}$ of Cys67 was visually aligned with a Cys residue. B) Structure 1 of $\alpha_3$D**IV** superimposed over the Hg-$Penicillamine_3$ site of $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$ (PDB 3PBJ at 2.20 Å resolution).[79] The $C^{\alpha}$ and $C^{\beta}$ of Cys67 was also visually aligned with a penicillamine residue. C) Model of trigonal T-shaped complexed derived from the EXAFS analysis, where Cys18 and Cys67 form short Hg-S bonds and Cys28 provides a Hg-S long bond.

Hg(II) can form two- and three-coordinate sulfur complexes depending on the pH. In the 3 C. N. complex, two geometries are possible: a trigonal planar complex with equal Hg-$S_\gamma$ bond lengths (~2 Å), and a T-shaped complex with two short Hg-$S_\gamma$ bonds (2.1 – 2.3 Å) and one long (~3 Å) bond. The EXAFS analysis on the 3 C. N. complex of Hg(II)-$\alpha_3$D**IV** illustrates a T-shaped coordination environment, therefore the triscysteine site of $\alpha_3$D**IV** was aligned with the X-ray crystal structure of the T-shaped Hg(II)-trispenicillamine of $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$ (Figure II-17B). This alignment shows that the $S_\gamma$ atom of and Cys67 has the proper orientation to form a short Hg-$S_\gamma$ bond, however Cys18 and Cys28 would require ample conformational changes to align with this T-shaped mercury-complex in $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$. Thus, like the Pb(II) complex, the distorted T-shaped geometry of Hg(II)-$\alpha_3$D**IV** could form different conformations, where the one long bond can alternate between Cys18 and Cys28 (Figure II-17C). The resonance peak in the $^{199}$Hg-NMR spectrum of $^{199}$Hg-$\alpha_3$D**IV** can elucidate this dynamic process. This nucleus is exceptionally sensitive to a slight perturbation in the Hg-ligand bond length and angle. A 0.01 Å deviation in a Hg-Cl bond length and 10° change in Cl-Hg-Cl bond angle was calculated to change their corresponding chemical shift values by ~50 ppm and 100 ppm.[82] For Hg(II)-$\alpha_3$D**IV** in a 3 C. N. complex, its $^{199}$Hg-NMR chemical shift value was experimentally determined to be -240 ppm. This value falls in between the $^{199}$Hg-NMR signals reported for symmetric Cys "a" and "d" sites

in 3SCC peptides with chemical shifts of ~-185 and ~-316 ppm respectively. This deviation by +55 ppm from an "a" and -76 ppm from a "d" site could be representative of a distorted trigonal thiolato $[HgS_3]^-$ T-shaped complex of Hg(II)-$\alpha_3$D**IV** with metal-ligand bond lengths and angles that deviate from its 3SCC counterparts.

**Conclusions**

*De novo* protein design is an emerging and biologically relevant approach in studying the metal binding sites of metalloregulatory proteins, as well as, metalloproteins. Toxic heavy metals, such as Hg(II), Pb(II), As(III) and Cd(II), threaten all forms of life. By understanding the mechanism by which bacteria control expression of genes to detoxify heavy metals through metalloregulatory proteins, one can develop strategies to combat this environmental issue. It offers a novel approach of modeling metal centers using simple peptide scaffolds. In this chapter, I described the solution structure of apo $\alpha_3$D**IV**, a *de novo* designed three-helix bundle peptide with a triscysteine metal binding that sequesters heavy metals Hg(II), Pb(II) and Cd(II). This thiol-rich environment is a ubiquitous motif in the metal binding site of metalloregulatory proteins, including ArsR, MerR and CadC. I showed that $\alpha_3$D**IV** has the same overall topology as, and aligns well with, its parent structure $\alpha_3$D. I found that incorporating Cys in place of Leu residues, which provide stabilizing packing interactions in the core, increased the helical content of the $\alpha_3$D framework and resulted in a more packed core. In addition, heavy metal binding to $\alpha_3$D**IV** induces further stability. Ultimately, this structure provides a stable framework for designing future metallopeptides that could perform specific catalytic or redox reactions.

## References


[1] Walsh, S. T., Cheng, H., Bryson, J. W., Roder, H., and DeGrado, W. F. (1999) Solution structure and dynamics of a *de novo* designed three-helix bundle protein, *Proc. Natl. Acad. Sci. U. S. A. 96*, 5486-5491.

[2] O'Halloran, T., and Walsh, C. (1987) Metalloregulatory DNA-binding protein encoded by the merR gene: isolation and characterization, *Science 235*, 211-214.

[3] Wright, J. G., Natan, M. J., MacDonnel, F. M., Ralston, D. M., and O'Halloran, T. V. (1990) Mercury(II)—Thiolate Chemistry and the Mechanism of the Heavy Metal Biosensor MerR, *Inorg. Chem. 38*, 323-412.

[4] Wright, J. G., Tsang, H. T., and Penner-Hahn, J. E. O. H., T. V. (1990) Coordination chemistry of the Hg-MerR metalloregulatory protein: evidence for a novel tridentate mercury-cysteine receptor site, *J. Am. Chem. Soc. 112*, 2434–2435.

[5] Busenlehner, L. S., Weng, T.-C., Penner-Hahn, J. E., and Giedroc, D. P. (2002) Elucidation of Primary (α3N) and Vestigial (α5) Heavy Metal-binding Sites in *Staphylococcus aureus* pI258 CadC: Evolutionary Implications for Metal Ion Selectivity of ArsR/SmtB Metal Sensor Proteins, *J. Mol. Biol. 319*, 685-701.

[6] Ye, J., Kandegedara, A., Martin, P., and Rosen, B. P. (2005) Crystal structure of the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor, *J. Bacteriol. 187*, 4214-4221.

[7] Banci, L., Bertini, I., Cantini, F., Ciofi-Baffoni, S., Cavet, J. S., Dennison, C., Graham, A. I., Harvie, D. R., and Robinson, N. J. (2007) NMR structural analysis of cadmium sensing by winged helix repressor CmtR, *J. Biol. Chem. 282*, 30181-30188.

[8] Chakraborty, S., Kravitz, J. Y., Thulstrup, P. W., Hemmingsen, L., DeGrado, W. F., and Pecoraro, V. L. (2011) Design of a three-helix bundle capable of binding heavy metals in a triscysteine environment, *Angew. Chem. Int. Ed. Engl. 50*, 2049-2053.

[9] Cangelosi, V. M., Deb, A., Penner-Hahn, J. E., and Pecoraro, V. L. (2014) A *de novo* designed metalloenzyme for the hydration of $CO_2$, *Angew. Chem., Int. Ed. Engl. 53*, 7900-7903.

[10] Plegaria, J. S., Dzul, S., Zuiderweg, E. R., Stemmler, T. L., and Pecoraro, V. L. (2015) Apoprotein Structure and Metal Binding Characterization of a *de Novo* Designed Peptide, $\alpha_3$D**IV**, that Sequesters Toxic Heavy Metals, *Biochemistry*.

[11] Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: a multidimensional spectral processing system based on UNIX pipes, *J. Biomol. NMR 6*, 277-293.

[12] Goddard, T. D., and Kneller, D. G. SPARKY 3.

[13] Bax, A., and Grzesiek, S. (1993) Methodological advances in protein NMR, *Acc. Chem. Res. 26*, 131-138.

[14] Loria, J. P., Rance, M., and Palmer, A. G., 3rd. (1999) Transverse-relaxation-optimized (TROSY) gradient-enhanced triple-resonance NMR spectroscopy, *J. Magn. Reson. 141*, 180-184.

[15] Pervushin, K., Riek, R., Wider, G., and Wuthrich, K. (1997) Attenuated $T_2$ relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution, *Proc. Natl. Acad. Sci. U. S. A. 94*, 12366-12371.

[16] Salzmann, M., Pervushin, K., Wider, G., Senn, H., and Wuthrich, K. (1998) TROSY in triple-resonance experiments: new perspectives for sequential NMR assignment of large proteins, *Proc. Natl. Acad. Sci. U. S. A. 95*, 13585-13590.

[17] Muhandiram, D. R., and Kay, L. E. (1994) Gradient-enhanced triple-resonance three-dimensional NMR experiments with improved sensitivity, *J. Mag. Reson., Ser. B 103*, 203-216.

[18] Grzesiek, S., and Bax, A. (1992) Improved 3D triple-resonance NMR techniques applied to a 31 kDa protein, *J. Mag. Reson. 96*, 432-440.

[19] Grzesiek, S., and Bax, A. (1992) An efficient experiment for sequential backbone assignment of medium-sized isotopically enriched proteins, *J. Mag. Reson. 99*, 201-207.

[20] Grzesiek, S., and Bax, A. (1992) Correlating backbone amide and side chain resonances in larger proteins by multiple relayed triple resonance NMR, *J. Am. Chem. Soc. 114*, 6291-6293.
[21] Kuboniwa, H., Grzesiek, S., Delaglio, F., and Bax, A. (1994) Measurement of HN-H J couplings in calcium-free calmodulin using new 2D and 3D water-flip-back methods, *J. Biomol. NMR 4*, 871-878.
[22] Crippen, G. M., Rousaki, A., Revington, M., Zhang, Y., and Zuiderweg, E. R. (2010) SAGA: rapid automatic mainchain NMR assignment for large proteins, *J. Biomol. NMR 46*, 281-298.
[23] Bax, A., Clore, G. M., and Gronenborn, A. M. (1990) 1H-1H correlation via isotropic mixing of $^{13}C$ magnetization, a new three-dimensional approach for assigning $^{1}H$ and $^{13}C$ spectra of $^{13}C$-enriched proteins, *J. Mag. Reson 88*, 425-431.
[24] Fesik, S. W., Eaton, H. L., Olejniczak, E. T., Zuiderweg, E. R. P., McIntosh, L. P., and Dahlquist, F. W. (1990) 2D and 3D NMR spectroscopy employing carbon-13/carbon-13 magnetization transfer by isotropic mixing. Spin system identification in large proteins, *J. Am. Chem. Soc. 112*, 886-888.
[25] Wang, H., and Zuiderweg, E. P. (1995) HCCH-TOCSY spectroscopy of $^{13}C$-labeled proteins in $H_2O$ using heteronuclear cross-polarization and pulsed-field gradients, *J. Biomol. NMR 5*, 207-211.
[26] Fesik, S. W., and Zuiderweg, E. R. P. (1988) Heteronuclear three-dimensional nmr spectroscopy. A strategy for the simplification of homonuclear two-dimensional NMR spectra, *J. Mag. Reson. 78*, 588-593.
[27] Fischer, M. W. F., Zeng, L., and Zuiderweg, E. R. P. (1996) Use of $^{13}C$-$^{13}C$ NOE for the Assignment of NMR Lines of Larger Labeled Proteins at Larger Magnetic Fields, *J. Am. Chem. Soc. 118*, 12457-12458.
[28] Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks, *J. Biomol. NMR 56*, 227-241.
[29] Guntert, P. (2004) Automated NMR structure calculation with CYANA, *Methods in molecular biology 278*, 353-378.
[30] Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York.
[31] George, G. N., George, S. J., and Pickering, I. J. (2001) EXAFSPAK.
[32] Ankudinov, A., and Rehr, J. (1997) Relativistic calculations of spin-dependent x-ray-absorption spectra, *Phys. Rev. B 56*, 1712-1715.
[33] Bencze, K. Z., Kondapalli, K. C., and Stemmler, T. L. (2011) X-Ray Absorption Spectroscopy, *Encyclopedia of Inorganic and Bioinorganic Chemistry*.
[34] Cotelesage, J. J., Pushie, M. J., Grochulski, P., Pickering, I. J., and George, G. N. (2012) Metalloprotein active site structure determination: synergy between X-ray absorption spectroscopy and X-ray crystallography, *J. Inorg. Biochem. 115*, 127-137.
[35] Wishart, D. S., Sykes, B. D., and Richards, F. M. (1991) Relationship between nuclear magnetic resonance chemical shift and protein secondary structure, *J. Mol. Biol. 222*, 311-333.
[36] Yao, J., Dyson, H. J., and Wright, P. E. (1997) Chemical shift dispersion and secondary structure prediction in unfolded and partly folded proteins, *FEBS Lett., 419*, 285-289.
[37] The PyMOL Molecular Graphics System, Version 1.7.4 Schrödinger, LLC.
[38] Bhattacharya, A., Tejero, R., and Montelione, G. T. (2007) Evaluating protein structures determined by structural genomics consortia, *Proteins 66*, 778-795.
[39] Lee, H. S., Choi, J., and Yoon, S. (2007) QHELIX: a computational tool for the improved measurement of inter-helical angles in proteins, *Protein J. 26*, 556-561.
[40] Ponder, J. W., and Richards, F. M. (1987) Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes, *J. Mol. Biol. 193*, 775-791.
[41] Hodges, R. S. (1996) De novo design of $\alpha$-helical proteins: basic research to medical applications, *Biochem. Cell Biol. 74,* 133-154.
[42] Tesmer, J. J., Stemmler, T. L., Penner-Hahn, J. E., Davisson, V. J., and Smith, J. L. (1994) Preliminary X-ray analysis of Escherichia coli GMP synthetase: determination of anomalous scattering factors for a cysteinyl mercury derivative, *Proteins 18*, 394-403.

[43] Tiekink, E. R. T. (1987) Bis(O-methyldithiocarbonato)mercury(II), *Acta Crystallogr., Sect. C: Cryst. Struct. Commun. 43*, 448 – 450.

[44] Kim, Y.-Y., Lee, H.-H., Nam, H.-J., and Noh, D.-Y. (2001) Mononuclear three-coordinate mercury(II)-thiolate complex: synthesis and X-ray crystal structure, *Bull. Korean Chem. Soc. 22*, 17 – 18.

[45] Chen, J.-X., Zhang, W.-H., Tang, X.-Y., Ren, Z.-G., Zhang, Y., and Lang, J. P. (2006) Assembly of a New Family of Mercury(II) Zwitterionic Thiolate Complexes from a Preformed Compound [Hg(Tab)2](PF6)2 [Tab ) 4-(Trimethylammonio)benzenethiolate], *Inorg. Chem. 45*, 2568 – 2580.

[46] Christou, G., Folting, K., and Huffman, J. C. (1984) Mononuclear, three-coordinate metal thiolates: Preparation and crystal structures of $[NBun_4][Hg(SPh)_3]$ and $[NPrn_4]$ $[Pb(SPh)_3]$, *Polyhedron 3*, 1247 – 1253.

[47] Wang, Y.-Y., Baiyin, M.-h., Liu, X.-y., An, Y.-l., and Nin, G.-l. (2006) Solvothermal Synthesis and Characterization of $(NH_3CH_2CH_2NH_3)HgSnS_4$ with a Column Structure, *Chem. Res. Chin. Univ. 22*, 411 – 414.

[48] Watton, S. P., Wright, J. G., MacDonnell, F. M., Bryson, J. W., Sabat, M., and O'Halloran, T. V. (1990) Trigonal mercuric complex of an aliphatic thiolate: a spectroscopic and structural model for the receptor site in the mercury(II) biosensor MerR, *J. Am. Chem. Soc. 112*, 2824 – 2826.

[49] Gruff, E. S., and Koch, S. A. (1990) Trigonal-planar $[M(SR)_3]1-$ complexes of cadmium and mercury. Structural similarities between mercury-cysteine and cadmium-cysteine coordination centers. , *J. Am. Chem. Soc. 112*, 1245 – 1247.

[50] Santos, R. A., Gruff, E. S., Koch, S. A., and Harbison, G. S. (1991) Solid-state mercury-199 and cadmium-113 NMR studies of mercury- and cadmium-thiolate complexes. Spectroscopic models for [Hg(SCys)n] centers in the bacterial mercury resistance proteins, *J. Am. Chem. Soc. 113*, 469 – 475.

[51] Deivaraj, T. C., Dean, P. A. W., and Vittal, J. J. (2000) Synthesis, Structure, and Multi-NMR Studies of $(Me_4N)[A\{M(SC\{O\}Ph)_3\}_2]$ (A = Na, M = Hg; A = K, M = Cd or Hg), *Inorg. Chem. 39*, 3071 – 3074.

[52] Baba, K., Okamura, T., Yamamoto, H., Yamamoto, T., and Ueyama, N. (2008) Zinc, Cadmium, and Mercury 1,2-Benzenedithiolates with Intramolecular NH···S Hydrogen Bonds, *Inorg. Chem. 47*, 2837 – 2848.

[53] Alsina, T., Clegg, W., Fraser, K. A., and Sola, J. (1992) Homoleptic cyclohexanethiolato complexes of mercury(II), *J. Chem. Soc. Dalton Trans. 8*, 1393 – 1399.

[54] Bramlett, J. M., Im, H.-J., Yu, X.-H., Chen, T., Cai, H., Roecker, L. E., Barnes, C. E., Dai, S., and Xue, Z.-L. (2004) Reactions of thioether carboxylic acids with mercury(II). Formation and X-ray crystal structure of mercury(II) mercaptoacetate, *Inorg. Chim. Acta 357*, 243-249.

[55] Hoffmann, G. G., Steinfatt, I., Brockner, W., Kaiser, V., and Naturforsch, Z. (1999) Preparation, characterization, crystal and molecular structure; reactivity towards organic thiols, *B: Chem.Sci. 54*, 887 – 894.

[56] Chieh, C. (1978) Crystal structure of mercury(II) bromide dithiocarbamate, $[HgBr_2(S_2CNEt_2)Hg(S_2CNEt_2)]n$, an inorganic polymer, *Can. J. Chem. 56*, 564 – 566.

[57] Alsaadi, B. M., and Sandstrom, M. (1982) Crystal and molecular structure of bis(2-mercaptobenzoato-S)mercury(II) monodioxane, *Acta Chem. Scand. A 36*, 509 – 512.

[58] Melnick, J. G., Yurkerwich, K., and Parkin, G. (2009) Synthesis, structure, and reactivity of two-coordinate mercury alkyl compounds with sulfur ligands: relevance to mercury detoxification, *Inorg. Chem. 48*, 6763-6772.

[59] Willard, L., Ranjan, A., Zhang, H., Monzavi, H., Boyko, R. F., Sykes, B. D., and Wishart, D. S. (2003) VADAR: a web server for quantitative evaluation of protein structure quality, *Nucleic Acids Res. 31*, 3316-3319.

[60] Lee, M. S., Gippert, G. P., Soman, K. V., Case, D. A., and Wright, P. E. (1989) Three-dimensional solution structure of a single zinc finger DNA-binding domain, *Science 245,* 635-637.

[61] Malmstrom, B. G. (1994) Rack-induced bonding in blue-copper proteins, *Eur. J. Biochem. 223*, 711-718.
[62] Peacock, A. F., Iranzo, O., and Pecoraro, V. L. (2009) Harnessing natures ability to control metal ion coordination geometry using de novo designed peptides, *Dalton Trans.* 2271-2280.
[63] Dieckmann, G. R., McRorie, D. K., Lear, J. D., Sharp, K. A., DeGrado, W. F., and Pecoraro, V. L. (1998) The role of protonation and metal chelation preferences in defining the properties of mercury-binding coiled coils, *J. Mol. Biol. 280*, 897-912.
[64] Chakraborty, S., Touw, D. S., Peacock, A. F., Stuckey, J., and Pecoraro, V. L. (2010) Structural comparisons of apo- and metalated three-stranded coiled coils clarify metal binding determinants in thiolate containing designed peptides, *J. Am. Chem. Soc. 132*, 13240-13250.
[65] Peacock, A. F., Hemmingsen, L., and Pecoraro, V. L. (2008) Using diastereopeptides to control metal ion coordination in proteins, *Proc. Natl. Acad. Sci. U. S. A. 105*, 16566-16571.
[66] Peacock, A. F., Stuckey, J. A., and Pecoraro, V. L. (2009) Switching the chirality of the metal environment alters the coordination mode in designed peptides, *Angew. Chem. Int. Ed. Engl. 48*, 7371-7374.
[67] Frankel, A. D., Berg, J. M., and Pabo, C. O. (1987) Metal-dependent folding of a single zinc finger from transcription factor IIIA, *Proc. Natl. Acad. Sci. U. S. A. 84*, 4841-4845.
[68] Siedlecka, M., Goch, G., Ejchart, A., Sticht, H., and Bierzynski, A. (1999) α-helix nucleation by a calcium-binding peptide loop, *Proc. Natl. Acad. Sci. U. S. A. 96*, 903-908.
[69] Suzuki, K., Hiroaki, H., Kohda, D., Nakamura, H., and Tanaka, T. (1998) Metal Ion Induced Self-Assembly of a Designed Peptide into a Triple-Stranded α-Helical Bundle: A Novel Metal Binding Site in the Hydrophobic Core, *J. Am. Chem. Soc. 120*, 13008-13015.
[70] Farrer, B. T., Harris, N. P., Balchus, K. E., and Pecoraro, V. L. (2001) Thermodynamic model for the stabilization of trigonal thiolato mercury(II) in designed three-stranded coiled coils, *Biochemistry 40*, 14696-14705.
[71] Chen, Y. H., Yang, J. T., and Chau, K. H. (1974) Determination of the helix and beta form of proteins in aqueous solution by circular dichroism, *Biochemistry 13*, 3350-3359.
[72] Davidson, D. E., and Hill, B. C. (2009) Stability of oxidized, reduced and copper bound forms of Bacillus subtilis Sco, *Biochim. Biophys. Acta 1794*, 275-281.
[73] Wu, H., Xu, X., Shen, D., Peng, L., Song, J., and Zhang, Y. (2011) Binding of $Ca^{2+}$ and $Zn^{2+}$ to factor IX/X-binding protein from venom of *Agkistrodon halys Pallas*: stabilization of the structure during GdnHCl-induced and thermally induced denaturation, *J. Biol. Inorg. Chem. 16*, 69-79.
[74] Matzapetakis, M., Farrer, B. T., Weng, T. C., Hemmingsen, L., Penner-Hahn, J. E., and Pecoraro, V. L. (2002) Comparison of the binding of cadmium(II), mercury(II), and arsenic(III) to the *de novo* designed peptides TRI L12C and TRI L16C, *J. Am. Chem. Soc. 124*, 8042-8054.
[75] Dieckmann, G. R., McRorie, D. K., Tierney, D. L., Utschig, L. M., Singer, C. P., O'Halloran, T. V., Penner-Hahn, J. E., DeGrado, W. F., and Pecoraro, V. L. (1997) *De Novo* Design of Mercury-Binding Two- and Three-Helical Bundles, *J. Am. Chem. Soc. 119*, 6195-6196.
[76] Pecoraro, V. L., Peacock, A. F. A., and Iranzo, O. (2009) Understanding the biological chemistry of mercury using a *de novo* protein design strategy, In Bioinorganic Chemistry, *ACS symposium Series* (Long, E. and Baldwin, M. J.) pp 183-197, American Chemical Society, Washington DC.
[77] Rule, G. S., and Hitchens, K. T. (2006) Fundamentals of Protein NMR spectroscopy, In *Fundamentals of Protein NMR spectroscopy*, pp 313–351, Springer Netherlands, Netherlands.
[78] Touw, D. S., Nordman, C. E., Stuckey, J. A., and Pecoraro, V. L. (2007) Identifying important structural characteristics of arsenic resistance proteins by using designed three-stranded coiled coils, *Proc. Natl. Acad. Sci. U. S. A. 104*, 11969-11974.
[79] Zastrow, M. L., Peacock, A. F., Stuckey, J. A., and Pecoraro, V. L. (2012) Hydrolytic catalysis and structural stabilization in a designed metalloprotein, *Nat. Chem. 4*, 118-123.
[80] Zampella, G., Neupane, K. P., De Gioia, L., and Pecoraro, V. L. (2012) The importance of stereochemically active lone pairs for influencing Pb(II) and As(III) protein binding, *Chem.-Eur. J. 18*, 2040-2050.

[81] Neupane, K. P., and Pecoraro, V. L. (2010) Probing a homoleptic $PbS_3$ coordination environment in a designed peptide using $^{207}Pb$ NMR spectroscopy: implications for understanding the molecular basis of lead toxicity, *Angew. Chem. Int. Ed. Engl. 49*, 8177-8180.

[82] Wolff, S. K., Ziegler, T., van Lenthe, E., and Baerends, E. J. (1999) Density functional calculations of nuclear magnetic shieldings using the zeroth-order regular approximation (ZORA) for relativistic effects: ZORA nuclear magnetic resonance, *J. Chem. Phys. 110*, 7689.