DESIGNED METALLOPROTEINS: FROM STRUCTURALLY CHARACTERIZED SCAFFOLDS TO HELICAL BUNDLES

by

Saumen Chakraborty

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry)
In The University of Michigan
2011

Doctoral Committee:

Professor Vincent L. Pecoraro, Chair
Professor Carol A. Fierke
Professor Erik R. P. Zuiderweg
Assistant Professor Nicolai Lehnert

To my Parents

Acknowledgements

There are many people to whom I would like to express my sincere gratitude that has helped me go through the tough years of graduate school. Normally I would write an extensive acknowledgement. But due to my shoulder pain I am writing a minimal acknowledgement. First, I would thank Vince for his support, encouragement, and enthusiasm. I thank Jeanne, Lars and Erik for fruitful collaborations. I am thankful to Olga, Anna, and Marek. I thank Melissa and Joe for being good supportive friends. I thank Ted and Cecily for inviting me to their wedding. Kathy has been very helpful to me over these years. I thank Leela for her help in the last few months. I thank other members of the Pecoraro group whom I have overlapped with.

I am thankful to my parents for their constant support in my progress. During the days of thesis writing their support gave me mental solace as I was suffering for severe physical pain. I thank my wife Amrita for all of her understanding and support during these years. She has been a source of mental calmness to me in tough times. Unfortunately, we haven't been able to close to each other when we needed each other most.

I thank all of my other friends in Michigan as well as in other places for their company and encouragement.

Table of Contents

List of Figures

List of Tables

ABSTRACT

DESIGNED METALLOPROTEINS: FROM STRUCTURALLY CHARACTERIZED SCAFFOLDS TO HELICAL BUNDLES

by

Saumen Chakraborty

Chair: Vincent L. Pecoraro

Fundamental studies using two classes of *de novo* designed α-helical scaffolds (three-stranded coiled coils or a single chain three-helix bundle), spectroscopy, and X-ray crystallography are exploited to enhance our understanding of the biochemistry of heavy metals Cd(II), Hg(II), and Pb(II) in thiol-rich *de novo* designed peptides.

The X-ray structure of a parallel three-stranded coiled coil (**CoilSer** L9C) is determined to 1.36 Å resolution in its non-metallated form (R_w= 18.2%; R_f=

21.5%). Cys side chains show alternate conformations and are partially (15%) preorganized for metal binding at the polypeptide interior.

How the physical properties of two Cd(II) ions bound in “identical” first coordination spheres at different topological positions within the same three-stranded coiled coil peptide vary has been investigated. **Grand**L16PenL19IL23PenL26I sequesters two Cd(II) ions in trigonal planar CdS_3 geometries with similar physical properties; however, these two sites are not independent from each other. Metallation state (apo or $[Cd(pep)(Hpep)_2]^+$ or $[Cd(pep)_3]^-$) of one site perturbs the physical properties of the second site. **Grand**L12AL16CL26AL30C binds two Cd(II) ions as 4-coordinate CdS_3O (O = water molecule) at both the binding sites, but with dramatically different properties. The 4-coordinate sites display differences in thermodynamic (pK_{a2} for Cd(II) binding ranging from 9.9 to 13.9) and kinetic (L26AL30C is more dynamic than L12AL16C site) properties depending on their location.

Exchange studies have been performed to understand the mechanism of insertion of Cd(II) into and out of helical scaffolds. A multi-site exchange scheme is invoked. Free Cd(II) is believed to interact with surface Glutamates before being bound within the peptide interior. The process occurs on the slow exchange regime with rates ranging from 5.8 to 10.9 ms. CdS_3 sites do not exchange on the NMR time scale.

A tris(cysteine) metal binding site was engineered at the C-terminal end of an existing *de novo* designed three-helix bundle yielding α_3**DIV**, a well folded and

stable protein capable of binding heavy metals with high affinity ($>10^7$ M^{-1}). Hg(II) is complexed as a linear HgS_2 (low pH) / trigonal planar HgS_3 complex (high pH), Cd(II) as four coordinate CdS_3O/CdS_3N (N = His72), and Pb(II) as trigonal pyramidal PbS_3 species.

Chapter 1

De Novo Designed Metalloproteins

Introduction

How the primary amino acid sequence of a protein defines its three-dimensional structure and function is not yet well understood. Due to inherent complexity of proteins it is often difficult to discern the features in their sequences that confer structural stability and function to the polypeptide. One approach to simplify this problem in the *de novo* design of model proteins and peptides that are much simpler than the natural counterparts, yet which retain sufficient complexity and information in their sequences that help us understand protein structure–function relationships.[1-4] In the purest form the terminology *de novo* design applies to the design of proteins 'from scratch'. *De novo* protein design involves the construction of a protein intended to fold into a well defined three-dimensional structure, with a sequence that is not directly related to any naturally occurring proteins.[5] The term '*de novo'* design has also been frequently applied to design sequences that would fold into a given general fold, without defining any atomic level details of the target proteins.[6-8] In other instances, it has been used to a variety of strategies involving redesign of natural proteins. A sequence based strategy for *de novo* design involves applying a 'minimalist' approach

leading to construction of ‘minimalist’ sequences.[2, 3, 9] Alternatively, automated structure-based strategies that begin with experimentally determined three-dimensional structures of naturally occurring proteins, have been very useful to design proteins with enhanced stability and novel functions.[10, 11]

Metalloproteins account for almost one half of the naturally occurring proteins in nature,[12] where the metals play crucial roles in structural and/or functional integrity of the biopolymer. Metalloproteins catalyze some of the most important biological processes in nature, starting from photo synthesis, respiration, water oxidation to signal transduction and complex chemical transformations. Because of the vast importance of metalloproteins due to their crucial functions, scientists have devoted their efforts towards the design of *de novo* metalloproteins over the past two decades. The ultimate goal is to use the knowledge gained from the design process to design metalloproteins that reproduce the structures and functions of native metalloproteins.[13] *De novo* metalloprotein design has empowered chemists to understand biologically relevant metal binding sites at a molecular level and gain important fundamental knowledge on how the protein matrix influences and controls the metal ion properties. This approach allows one to elucidate, in a systematic manner, structural features and elements that play crucial roles in determining the coordination geometries and physical properties of metal ions, as well as their site specificity. This degree of knowledge is necessary to fully comprehend how metalloproteins work and fine tune the metal ion geometries and properties. However, designing metalloproteins is not just an intellectual exercise that mimics the biochemical functions of native metalloproteins. Equipped with the insight from the design process one hopes to devise novel metalloproteins with improved properties such as higher stability as well as exciting new catalytic transformations that are not found in nature.[12] Designing metalloproteins has been more challenging compared to non-metalloproteins. Both designs involve the construction of a sequence that would into a precisely determined three-dimensional structure. However, as metalloproteins involve introduction of metals

which are much more in number than the number of amino acids and have metal ion geometries that are much more variable, designing metalloproteins have proven to be more complex. On the other hand, most metal binding sites can be characterized by numerous metal-based spectroscopic techniques, rather than relying only either on X-ray crystallography or nuclear magnetic resonance spectroscopy. Thus, the field of metalloprotein design has seen significant amount of successes in recent years owing to advances in peptide synthesis, computational and structural biology. Current studies in metalloprotein design are focused on unveiling important structural features in natural protein systems with the goal of using this knowledge to design proteins that are tailor- made to carryout receptor, sensory and catalytic functions.[14-19] The field of metalloprotein design is also progressing from the design of primary coordination sphere of metals towards the design of secondary coordination sphere and beyond.

The most commonly used secondary structural motif in the field of *de novo* metalloprotein design is the α helix. Numerous metalloprotein systems have been designed over the past fifteen years using α helices, typically using unassociated peptides that assemble into three-stranded coiled coils or helix-loop-helix motifs that form anti-parallel four- stranded bundles. These helical aggregates result from the sequences containing a characteristic seven-residue [heptad] repeat, denoted [**abcdefg**]$_n$, that typically has hydrophobic residues at **a** and **d** positions, and polar or charged residues at **b, e, f** and **g** positions. Helix-inducing residues (Ala, in general) are placed at position **c**. The hydrophobic **a** and **d** residues point towards the interior of the coiled coil and form the hydrophobic core. The **a** and **a'** residues are oriented in such a way that they are 20° out of phase, leading to α helices that twist around each other to form the coiled coil.[20] The hydrophilic residues are involved in the formation of inter-helical electrostatic interactions at the helical interfaces that are critical for parallel or antiparallel configurations of the coiled coils. Despite the apparently simple pattern of seven amino acids, the heptad repeat approach has shown its high versatility and potential to address many fundamental problems is

chemistry. Figure 1-1 shows helical wheel diagrams of parallel two- and three-stranded coiled coils.

Due to the extensive amount of literature that exists on the *de novo* designed metalloprotein systems, I will discuss the advances in the field under two broad sub categories: 1) unassociated peptides that assemble into two, three or four- stranded coiled coils or helix-loop-helix motifs that form anti-parallel four-stranded bundles and 2) single polypeptide chains where the helices are tethered together in the form of two or three-helix bundles.

1) Unassociated peptides that assemble into two, three or four- stranded coiled coils or helix-loop-helix motifs that form anti-parallel four- stranded bundles

a) Metal Induced Self Assembly of Random Coils

Studies have been performed by many different research groups where metal-ion assisted folding and self assembly of random coils leading to the formation of well defined secondary and tertiary structures have been reported. Ghadiri *et al.*[21] introduced a His residue at i and a Cys / His residue at i+4 position of a 17 amino acid peptide Ac-AEAAAKEAAAKX_1AAAX_2A-CONH_2, where X_1 = His, X_2 = Cys / His. Binding of Cd(II) to the His/Cys derivative in a bidentate geometry promoted the formation of monomeric α-helix (90% α-helicity) from a disorder polypeptide (54% α-helicity) in the absence of Cd(II). Ni(II) binding to the same peptide induced higher α-helicity compared to Cd(II). The His / His derivative showed high affinity binding for Cu(II) and Zn(II) with concomitant increase in the α-helicity up to 90% upon metal binding to the monomeric unstructured polypeptide. They have shown the utility of an exchange inert metal complex in the formation of stable α-helical peptides.[22] Using the same His / His derivative of the 17 mer peptide they have shown that addition of $[Ru(NH_3)_5(OH_2)]^{2+}$ led to the formation of the macrocyclic complex *cis*-[Ru-

$(NH_3)_4(His)_2$peptide$]^{3+}$ afforded an 80% α-helical peptide which in the absence of metal was only 45% α-helical. They have also designed a 15-residue amphiphilic peptide with a 2,2'-bipyridine functionality at the N-terminus end which was shown to undergo a spontaneous self-assembly in the presence of transition metal ions Ni(II), Co(II), and Ru(II) to form a 45-residue triple-helical coiled coil (Figure 1-2).[23] In the absence of the metal ions this peptide was 30% α-helical, but in the presence of metals it exhibited >70% α-helicity. Furthermore, the $[Ni(bpy\text{-}peptide)_3]^{2+}$ led to the selective formation of one diastereomer where tris-bipyridyl moiety adopted a left-handed Λ-isomer. Subsequently, a heterodinuclear three-helix construct was designed which incorporated both Ru(II) and Cu(II) (Figure 1-2).[24] The $[Ru(bpy\text{-}peptide)_3]^{2+}$ construct was able to act as a template for incorporating Cu(II) leading to the formation of a $Cu(His)_3$ site. Spectroscopic features revealed that the bound Cu(II) displayed properties of a type-II copper protein. Another 15-residue amphiphilic polypeptide, equipped at the N-terminus with a pyridyl functionality, was designed and shown to undergo intermolecular self-assembly, upon Ru(II) complexation, to form a remarkably stable 60-residue parallel four-helix bundle metalloprotein (Figure 1-3).[25] Inter-subunit hydrophobic interactions were the major driving force to the formation, while the metal ion complexation to the pyridyl ligands protruding from each structural subunit provided the key element for controlling the overall topology and the number of subunits participating in the assembly process.

Ogawa *et al.* designed a 32 mer peptide C16C19-GGY (Ac-$K(IEALEGK)_2(CEACEGK)(IEALEGK)GGY\text{-}CONH_2$) where the presence of Cys-X-X-Cys metal binding domain of rubredoxin was used to place cysteine residues at the hydrophobic **a** and **d** positions upon formation of a homodimeric α-helical coiled coil. The apo peptide existed as a random coil and assembled into a dimeric metal-bridged coiled coil in the presence of Cd(II) (Figure 1-4).[26, 27] The two-stranded coiled coil contained a single Cd(II) ion. Intriguingly, in the presence of Cu(I), the random coil peptide C16C19-GGY self assembled to a 4-helix bundle metalloprotein and incorporated a tetranuclear Cu_4S_4 cofactor (Figure 1-

4).[28] They have also reported the use of metal coordination to alter the oligomerization state of a designed peptide structure. The 30-residue peptide AQ-Pal14Pal21 contained two metal-binding residues 4-pyridylalanine (Pal) on its solvent-exposed surface and existed as a two-stranded coiled coil on the absence of metal. Upon addition of $Pt(en)(NO_3)_2$, a significant conformational change was observed leading to the formation of a metal-bridged four-helix bundle.[29]

Tanaka *et al.* designed the native-like three-stranded coiled coil isoleucine zipper peptide IZ ($YGG(IEKKIEA)_4$).[30] They engineered a metal binding site in the hydrophobic core of the coiled coil by replacing two Ile groups, one from the **a** site and a second from the **d** site of the third heptad with two His residues, resulting in a peptide IZ-3adH.[31] The apo peptide had a random coil structure in the absence of metals. In the presence of different metal ions such as Co(II), Ni(II), Cu(II), and Zn(II), IZ-3adH formed parallel three-stranded α-helical coiled coil (Figure 1-5). K_d values obtained from metal titrations monitored by circular dichroism spectroscopy were 35 ± 1, 5.0 ± 0.3, 17 ± 1 and 23 ± 2 μM for Co(II), Ni(II), Cu(II), and Zn(II) respectively. Highest affinity was observed for Ni(II) and the metal ion binding site had an octahedral geometry. In a separate study they showed that the replacement of Ile18 and Ile22 from the IZ peptide by Ala and Cys residues (IZ-AC) resulted in disruption of the triple-stranded α-helix structure. The addition of Cd(II) and Hg(II) resulted in the reconstitution of the three-stranded α-helix bundle.[32] Additionally only one Ile (Ile22) was mutated to His in order to study metal ion specificity, resulting in the peptide IZ-3aH. This construct was also weakly α-helical in the absence of metal, but in the presence of Cu(II) and Zn(II) it formed a three-stranded coiled coil. The k_d values were determined to be 9 and 10 μM, for Cu(II) and Zn(II), respectively. Cu(II) was coordinated in a square planar geometry to three His ligands and an exogenus water molecule. Interestingly, this peptide did not bind Ni(II).[33]

Pecoraro *et al.* have reported studies involving encapsulation of Hg(II) to 23-residue Baby peptides that are weakly associated or unstructured in the absence of metal ion. In the absence of metal ions BabyL9C peptide (Leucine from 9 position is replaced to metal binding Cys residue) is only ~20% α-helical at pH 8.5.[34] Addition of Hg(II) induced α-helicity and BabyL9C formed well-structured three-stranded α-helical coiled coil encapsulating tristhiolato Hg(II) at the interior of the coiled coil (Figure 1-6).[34] Kinetic investigation of the encapsulation of Hg(II) to BabyL9C has been reported.[35] The mechanism involved a rapid collapse of two peptide strands forming linear dithiolato-Hg(II) complex within a two-stranded coiled coil. Addition of the third peptide strand, leading to the formation of three-stranded coiled coil was the rate limiting step and resulted in an intermediate state involving two peptide species. One of them, the properly folded intermediate underwent rapid deprotonation of the third Cys thiol yielding the trigonal thiolato Hg(II). The other species, termed as the misfolded intermediate rearranged to the properly folded intermediate (Figure 1-6) and then formed the trigonal HgS_3 complex.

Metal-induced folding of other proteins such as zinc finger and calmodulins[36] are also known. Zinc finger proteins have a simple ββα structure, that fold around a central Zn(II) ion tetrahedrally coordinated to a Cys_2, His_2 metal binding site. The two cysteine ligands are near a turn in the antiparallel β sheet, and the two histidines are in the C-terminal portion of the α helix (Figure 1-7).[37] Zinc finger proteins are characterized by the presence of one or more sequences that closely approximate the form (Tyr,Phe)-X-Cys-$X_{2,4}$-Cys-X_3-Phe-X_5-Leu-X_2-His-$X_{3,4}$-His, where X represents relatively variable amino acids.[38] Proteins that contain these domains have been shown to be specific nucleic acid binding proteins.[39] Peptides corresponding to single domains of this type have been used It has been demonstrated that these peptides are largely unfolded in the absence of metal ions such as Zn(II) or Co(II) but fold to unique three-dimensional structures in their presence.[40-48]

b) Cofactor Binding to Designed Proteins: Heme / Iron-Sulfur Cluster

Many research groups have been interested in incorporating cofactors such as hemes / iron sulfur clusters into designed proteins. A two-helix peptide α_2 [1] that dimerizes to form a four-helix bundle was engineered to bind a single heme parallel to the helices in the core of the bundle using two His side chains as ligands.[49] The resulting protein VAVH25(S-S) bound the heme with a k_d of 7 $\pm$ 1 $\times$ 10^{-7} M. To ensure an unambiguous dimeric aggregation state with both loops oriented at one end of the bundle, a cysteine was introduced at the N-terminus of the peptide oxidized to form a homodimeric, disulfide linked protein (Figure 1-8). Mihara and coworkers designed a 14-mer peptide to form a di-α-helical peptide His-2α that was shown to effectively bind one molecule of heme. A 31-residue α-helical peptide was designed by Dutton *et al.* to form a di-α-helical peptide via disulfide bond formation.[50] The resulting α-helical dimer was shown to accommodate two heme groups by axial ligation to two His residues at position 10 and 24 along the sequence. The peptide self assembled into a four-helix dimer with a 2-fold symmetry and four parallel hemes that closely resembled native hemes in their spectral and electrochemical properties, including heme-heme redox interaction (Figure 1-9). This protein was an essential intermediate in the synthesis of molecular 'maquettes', a novel class of simplified versions of the metalloproteins involved in redox catalysis and in energy conversion in respiratory and photosynthetic electros transfer. Subsequent studies involved design of the di-α-helical peptide as a continuous sequence instead of linking two individual α-helices by disulfide bonds.[51] The resulting peptide also self assembled in solution to form four-α-helix bundle that bound two molecules of Fe(III)-protoporphyrin IX and two molecules of Zn(II)-protoporphyrin IX with k_d values ranging from 35 to 135 nM. Other studies by Dutton, Gibney and coworkers have been reported involving the self assembly of Heme A and Heme B in designed four-helix bundles and addressing the thermodynamics of ferric and ferrous heme affinity of a separate four α-helix

protein.[52-54]. Benson *et. al.* have studied peptides that were designed to form α-helices when incorporated in novel hemoprotein model compounds, peptide-sandwiched mesohemes, which consist of two identical peptides covalently attached to an Fe(III) mesoporphyrin via the side chains of His and Lys.[55-57] Suslick *et al.* have made a series of studies with synthetic analogues of heme proteins to understand the factors that influence peptide binding to porphyrins. They had prepared a series of 15-mer peptides that form 1:2 metalloporphyrin-peptide complexes using a centrally located His residue and examined the effect of their sequence on binding constants, secondary structure, and electrochemical behavior.[58] In all of these cases metalloporphyrin binding induced α-helicity and outlined the importance of heme-peptide hydrophobic interactions to the stability of the resulting complexes. They discovered that noncoordinating hydrophobic residues flanking a central histidine stabilized the heme-peptide complexation by as much as 4.5 kcal/mol.[59] Subsequently, they incorporated disulfide bridges between the peptides to make hairpin and even cyclic peptides that bound heme extremely well, roughly 5×10^6 times more strongly than histidine itself (Figure 1-10).[60] They also characterized the cyclic peptide-porphyrin complexes and showed that the complex had native-like characteristics from thermal unfolding curves and well-dispersed NMR signals. DeGrado and coworkers have reported design of a *D2*-symmetric four-helix bundle that bound two heme units with high affinity at low micromolar concentrations (Figure 1-11).[61] Each heme was shown to be coordinated by two His residues from diagonally opposed helices. They have also designed a four-helix bundle that selectively bound two molecules of a nonbiological cofactor Diphenyl prophyrin FeDPP(III).[62] Recently they have reported the computational design of a single-chain four-helix bundle that noncovalently self-assembled with FeDPP(III).[63]

Dutton *et. al.* have investigated the rules for incorporation of cofactors such as $[4Fe\text{-}4S]^{2+/1+}$ in simplified systems to produce minimal, functional, synthetic proteins- the molecular maquettes. They designed a 16-residue peptide (NH_2-KLCEGGCIACGACGGW-$CONH_2$) derived from a consensus motif of

natural ferredoxins. This ferredoxin maquette (FdM) was shown to incorporate a tetranuclear iron sulfur cluster at pH 8.0 with the spectroscopic and electrochemical properties typical of natural bacterial ferredoxins. The four cysteines of FdM were systematically replaced with nonligating alanine residues resulting in lower yields (10-56% relative to a FdM control) demonstrating the necessity for each Cys residue in successful cluster incorporation.[64] In a subsequent study they investigated the role of the nonligating amino acids of FdM in the assembly of the iron-sulfur cluster.[65] In a stepwise fashion they truncated the FdM to a seven-amino acid peptide, FdM-7 (CIACGAC) which incorporated a cluster spectroscopically identical to FdM but in lower yield, 29% relative to FdM. This result showed that only three Cys residues are required for monomeric cluster ligation. When all the nonligating residues were mutated to either Gly or Ala, the resulting constructs incorporated the cluster to only 6-7%. A systematic study involving the placement of Ile in each of the four nonligating positions indicated that placement of Ile either in the second or sixth core motif positions of FdM-7, restored the cluster binding capacity to the level of FdM-7. This study showed that the identities of the nonligating amino acids are as important as the ligating Cys residues for incorporation and stabilization of the iron-sulfur cluster.

c) Dinuclear Metal Binding Sites in Designed Proteins: Artificial diiron Proteins

The diiron/dimanganese cluster is a versatile cofactor found in functionally diverse enzymes that are capable to catalyze both hydrolytic and oxygen-dependent reactions.[66] Structurally very similar diiron sites are capable of catalyzing the ferroxidase (ferritins),[67-69] hydroxylase (methane monooxygenase,[70-72] and toluene oxygenases),[73-75] Mn catalase,[76, 77] alkane desaturation reactions,[78, 79] or are capable of acting as a radical generator in ribonucleotide reductase.[80-82] These proteins have a common dimetal site consisting of two metal ions, bridged by a combination of oxo, hydroxo, or

carboxylate donors, and are housed within a very simple four-helix bundle motif. DeGrado *et. al.* have designed four helix bundle proteins to elucidate how the electrostatic environment, polarity, solvent accessibility of the metal-binding site, and electrochemical reduction potential influence the properties of these proteins, especially those of the diiron proteins.

Based on a retrostructural analysis of the active sites of the above proteins, they designed and prepared a series of models for the diiron proteins, called the DueFerri (DF) family of artificial diiron proteins. The first was DF1,[83] a minimal model for the Glu- Xxx-Xxx-His class of dinuclear metalloproteins. This protein was synthesized and structurally characterized as the di-Zn(II) complex by X-ray crystallography, to 2.5 Å resolution (Figure 1-12). This four-helix bundle protein was comprised of two noncovalently associated helix-loop-helix motifs. The dinuclear metal center is formed by two bridging Glu and two chelating Glu side chains, as well as two monodentate His ligands. The primary ligands were mostly buried in the protein interior, and their geometries are stabilized by a network of hydrogen bonds to second-shell ligands. In particular, a Tyr residue formed a hydrogen bond to a chelating Glu ligand, similar to a motif found in the diiron-containing R2 subunit of *E. coli* ribonucleotide reductase and the ferritins. DF1 also binds cobalt and iron ions and provided an attractive model for a variety of diiron proteins that use oxygen for processes including iron storage, radical formation, and hydrocarbon oxidation. Subsequently they solved the NMR structure of the apo form of DF1 which showed that the metal binding residues were almost entirely preorganized for metal binding.[84] As originally designed the DF1 protein had low water solubility (10 μM) as it contained too many hydrophobic residues on its surface, and the experiments had to be performed in DMSO. The hydrophobic surface residues were converted to polar residues in the next generation of DF proteins. Additionally a pair of Leu residues, Leu 13 and Leu 13' blocked the access of metal ions and small molecule substrates to the dimetal active site of DF1.[85] Therefore, in order to increase accessibility to the dimetal site of DF2, Leu 13 was mutated to Ala and Gly. The DF2 was

expressed in *E. coli* and had much better water solubility (0.5 mM). Although DF2 was less stable compared to DF1, it retained sufficient stability in solution to remain well folded even in the absence of metals.[86] The NMR and X-ray structure of the di-Cd(II) form DF2 was solved and were essentially the same as DF1.[87] They also solved the crystal structure of the dimanganous forms of L13A-DF1 and L13G-DF1, which is similar in sequence to DF2, which indeed showed a substrate access channel to the dimetal center.[85, 88] In the crystal structure of L13A-DF1, water molecules and a ligating dimethyl sulfoxide molecule, which forms a monatomic bridge between the metal ions, occupy the cavity. The structure of di-Mn(II)-L13G-DF1 displayed four crystallographically independent dimers in the asymmetric unit. The individual dimers were closely related to the structure of di-Mn(II)-L13A-DF1, with the exception that the water-filled access channel has been expanded in the expected manner. The most striking feature found in di-Mn(II)-L13G-DF1 was the presence of two different dimanganese coordination environments. In three of the four dimers a solvent molecule bridges the two metal ions, while in the fourth dimer two terminal solvent molecules are coordinated to the two manganese ions trans to the histidine ligands.

Although DF2 was significantly more soluble than DF1, DF2 had the tendency to aggregate above pH 7.0, and was expressed at relatively modest yields. Careful examination of the structure revealed that the interhelical turn had a strained conformation, which might account for these problems. Based on the structures available in the Protein Data Bank it appeared that the turn in DF2 was not optimal. Therefore, a longer loop was inserted between the two helices. This protein DF2t expressed in significantly better yield and did not have the tendency to aggregate. The di-Zn(II) structure of DF2t was solved by X-ray crystallography as well as NMR, and was shown to conform exactly to the design.[87] Subsequently, they engineered a four-chain heterotetramer. First, an A_2B_2 heterotetramer was generated (DFtet A_2B_2) and was designed to be an antiparallel coiled coil.[89] The helices were extended compared to those of DF1 (33 residues in DFtet vs. 24 residues in DF1) to increase the stability of the DFtet

system by increasing the stability of the DFtet system by increasing the size of the hydrophobic core. DF2tet A_2B_2 bound Co(II), Zn(II), and Fe(II) in the expected stoichiometry and showed ferroxidase activity under single turn over conditions. A redesigned derivative of DF2tet A_2B_2, where Leu 15 and Ala 19 of both the A chains were substituted to Gly, showed phenol oxidase activity.[19] The diferrous protein first undergoes oxidation to diferric species which then catalyzed the two-electron oxidation of 4-aminophenol (k_{cat}/K_M = 1500 $M^{-1}min^{-1}$) producing benzoquinone monoimine. The sequence of DF2tet A_2B_2 was further modified to create DF2tet $A_aA_bB_2$.[90] In this design one helix-helix interface was redesigned to provide high specificity for the desired three-component assembly. This four-helix bundle was constructed by the noncovalent self-assembly of three different chains (A_a, A_b, and B) that self assembled into the desired heterotetramer when mixed in a 1:1:2 molar ratio. On addition of ferrous ions and oxygen, the protein was shown to form a complex with a UV-visible spectrum closely resembling that of peroxo-bridged diferric species in natural proteins and model compounds. Then they characterized a single-chain version of the four-helix bundles, DFsc, consisting of 114 amino acid residues. The identities of 26 residues were predetermined, including the primary and secondary ligands in the active site, residues involved in active-site accessibility, and the turn sequences. The remaining 88 amino acids were determined using side-chain packing algorithm.[91] DFsc expressed well in *E. coli* and was highly soluble (2.5 mM).[92] The protein was stable in the apo form and stoichiometrically bound a variety of divalent metal ions including Zn(II), Co(II), Fe(II), and Mn(II) with micromolar affinities. The ^{15}N HSQC of both apo and metallated protein had well-dispersed chemical shifts with evidence of a significant structural change upon metal binding. An NMR structure of the Di-Zn(II)-DFsc confirmed the overall fold of the protein.

The metal sites in all of the structures of DF family of proteins are approximately 5-coordinate: two Glu residues bridge the two metal ions, two Glu interact with individual metal ions in an approximately bidentate interaction, and

two His residues interact with individual metal ions. There is a vacant site *trans* to the His ligands that potentially acts as a site for substrate binding. The inter-Zn(II) distance ranges from 3.9 to 4.0 Å, while the inter Cd(II) distance is ~ 3.7 Å. The greatest degree of variability was observed in the bridging Glu residues, whose carboxylate groups frequently lie in the same place as the two metal ions, but in some structures are rotated out of the plane. The chelating carboxylate groups also show a high degree of variability in the angle of the carboxylate relative to the bound metal ions. The orientation of His residues remains essentially invariable in the structures.

Recently DeGrado and coworkers have redesigned DF1 to DF3 which also showed phenol oxidase activity for the substrates 4-aminophenol (k_{cat}/K_M = 1380 $M^{-1}min^{-1}$), 3,5,-ditert-butyl-catechol (k_{cat}/K_M = 6315 $M^{-1}min^{-1}$) and *para*-phenylenediamine (k_{cat}/K_M = 83 $M^{-1}min^{-1}$).[93] L9 and L13 from the first helix of DF1 were mutated to Gly to allow access of the substrate to the active side. However, the introduction of glycines severely destabilized the protein. A single mutation caused destabilization by 10.8 kcal mol^{-1} $dimer^{-1}$. To increase conformational stability the loop residues were modified from Val-Lys-Leu in DF1 to Thr-His-Asn which formed stabilizing N- and C-capping interactions. Solution structure of Di-Zn(II)-DF3 has been solved and is nearly identical to the design.

De novo design of proteins and metalloproteins has been dominated by α-helical coiled coil and bundle domains. A significant amount of effort has been directed towards designing β sheet structures which has turned out to be more challenging due to aggregation, challenges in designing turn sequence, folding cooperativity, and overall stability of the designed construct. Bulky hydrophobic amino acids such as Tyr, Phe, Thr, Ile, Val, and Trp have turned out to have higher propensity to form β-sheet like structures. Overcoming these challenges a number of related molecular scaffolds have been designed such as, $\beta_2\alpha$,[48, 94, 95] small peptides that bind metal clusters at loop region,[96-98] γ turns,[99] β-hairpin peptides,[100, 101] and β- sheet structures.[102-105]

d) A Specific Example of *De Novo* Metallopeptide Design: Parallel Three-Stranded Coiled Coils

Using *de novo* designed parallel three-stranded coiled coil peptides as model systems, research in the Pecoraro group has focused on developing structural and spectroscopic models that provide a foundational understanding of interactions of heavy metals with different metalloregulatory proteins such as MerR,[106, 107] CadC/CmtR,[108-111] ArsR/SmtB,[112] and pbrR.[113, 114] Lower organisms such as bacteria, have evolved a detoxification mechanism to cope with heavy metals such as Hg(II), Cd(II), As(III) and Pb(II) by employing these different metalloregulatory proteins. In the absence of heavy metals these proteins bind to DNA, blocking the transcription of the genes that code for proteins that extrude heavy metals from cells. Upon metal binding the repressor protein either twists the DNA (MerR) or dissociates from DNA (CadC/CmtR, ArsR/SmtB) after undergoing a conformational change which then allow transcription to happen. In all of these cases, the proteins display exclusive Cys ligation around the metal centers (Figure 1-13). The peptide systems that are used by the Pecoraro group to understand the fundamentals of heavy metal biochemistry as well as metalloprotein design, are the TRI,[115] GRAND,[116] and Coil Ser[117] series of peptides, designed based on the well-known heptad repeat (Table 1 shows the sequences), and form parallel three-stranded coiled coils in solution and in the solid state. At low pH these peptides exist as two-stranded coiled coils, and at high pH they form parallel three-stranded coiled coils. The conversion from two-stranded to three-stranded coiled coils occurs with a pK_a of ~ 4.5 as the Glu side chains are deprotonated which form favorable salt bridge interactions with the Lys residues that are critical for the parallel orientation of the coiled coils.[115] Substitution of Leu residues from either **a** or **d** sites of the heptad to Cys or Penicillamine (Pen) generates homoleptic thiol-rich soft metal binding sites at the interior of coiled coils that are suitable for binding different heavy metals such as Hg(II), Cd(II), As(II), Pb(II) and Bi(III) with high affinity (10^7–10^8 M^{-1}) and predefined coordination geometry.[34, 35, 115, 116, 118-128] The group has

reported numerous successes using the *de novo* protein design strategy to understand heavy metal biochemistry using peptide systems based on the three-stranded coiled coil motif (TRI series), including the first water-soluble spectroscopic model[118] of the MerR protein, the first trigonal pyramidal As(III) bound to three cysteines as a model of ArsR[119] (Coil Ser series), and the first example of a trigonal thiolate Cd(II).[120-122] While effective protein based heavy metal sequestering agents, the TRI and GRAND series of peptides are not easily crystallizable. For structural studies, the group has used the structurally similar Coil Ser (CS) series of peptides which are also based on a heptad repeat, and can crystallize as parallel or antiparallel three-stranded coiled coils. Coil Ser was first crystallized by Lovejoy and DeGrado et al., as an antiparallel coiled coil at low pH.[117] It was suggested that the steric clash between the N-terminal Trp residues led to the antiparallel orientation. Subsequent NMR work demonstrated that Coil Ser formed parallel three-stranded coiled coils in slightly basic solution.[129] The Pecoraro group has subsequently characterized both metallated[119] As(**CS**L9C)$_3$ and apo derivatives of **CS**L16-L-Pen and **CS**L16-D-Pen peptides[130] which have been shown to be parallel three-stranded coiled coils. This orientation may reflect the pH, the sequence modification, the presence of metals bound in the hydrophobic interior, the presence of Zn(II) ions that bind to external His and Glu residues at the crystal packing interfaces or a combination of these factors. As(III) in As(**CS**L9C)$_3$ was shown to be coordinated to the cysteines in an *endo* conformation within the protein environment (Figure 1-14). Comparison of the structure of metallated As(**CS**L9C)$_3$ and apo **CS**L16-L-Pen showed that the sulfurs appeared to be preorganized for metal binding (Figure 1-15). In the Pen structure one set of methyl groups were oriented towards the interior of the coiled coil blocking the plane below Cys. The second set of methyl groups pointed away from the center in-between the helical interfaces. The sulfurs in the D-Pen structure were oriented towards the C-terminus and pointed away from the interior of the coiled coils towards the helical interface. From previous studies the group has demonstrated that the affinities

and selectivities of metal binding to the **a** vs. **d** site substituted peptides are different in solution.[126, 131] Cd(II) binds selectively to the **a** site and has a 10-fold higher affinity over **d** site.[124] Pb(II) on the other hand has a 4-fold higher affinity for **d** site.[132, 133] Clearly, structural information regarding **a** and **d** sites within the coiled coils are required to fully comprehend the differential behavior of **a** vs. **d** site peptides.

In the second chapter of my thesis I describe the X-ray structures of two apo peptides. **CS**L9C, and **a** substituted peptide and **CS**L19C, a **d** site peptide to 1.36 and 2.15 Å resolutions, respectively. Data collection and refinement for the letter peptide was done by Dr. Debra Touw, a former graduate student from the group. With **CS**L9C peptide, I discuss local/global structural differences with the metallated peptide As(**CS**L9C)$_3$, effect of As(III) binding on the overall structure of the peptide, and address whether As(III) binding changes the conformations and the extent of the degree of preorganization of the metal binding cysteine residues. With the apo **CS**L19C structure I have addressed whether there are differences in the orientation of the Cys residues in the **a** vs. **d** site structures, size of the thiol pocket generated by the two types of structures, packing of the adjacent hydrophobic Leu residues. On the basis of the observed structural difference between the two sites, I also hypothesize the possible binding modes in which Cd(II), Hg(II), and Pb(II) would coordinate in the two types of structures which would help us understand the observed differences in physical properties for metal binding to **a** vs. **d** substituted three-stranded coiled coil peptides.

The Pecoraro group has been very interested in controlling coordination number of metals such as Cd(II) within the protein interior. By combining ^{113}Cd NMR and ^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopy, the group has shown how the peptide **TRI**L16C binds Cd(II) as a mixture of pseudotetrahedral (CdS_3O) and trigonal planar (CdS_3) structures.[124] Complete control of the Cd(II) coordination number was achieved by modifying the sterics in the metal ion binding site through changes in the amino acids of either the first

or second coordination sphere as well as harnessing the steric effects associated with the spatial orientations of the side chains by incorporating D-amino acids. Removal of steric bulk directly above the Cys plane by replacement of a Leu with an Ala generated the peptide **TRI**L12AL16C that binds Cd(II) exclusively in the CdS_3O geometry.[120, 134] On the other hand, the pure CdS_3 geometry was achieved by increasing steric constrains around the metal ion binding site. Two different approaches were used in the latter case, both employing non-natural amino acids. In the first approach, Cys was replaced by Penicillamine (Pen) which has bulky methyl groups replacing the β-methylene hydrogen atoms (peptide **TRI**L16Pen).[120] In the second strategy, the chirality of the Leu directly above the Cys was modified by replacing L-Leu by D-Leu. As a consequence, the side chain of the Leu was reoriented towards the C termini and towards the metal ion binding site (peptide **TRI**L12L_DL16C**)** leading to a lower coordination number of Cd(II) by blocking the access of exogenous water molecules to the metal ion.[121] This knowledge has driven the design strategy for the heterochromic peptides, **GRAND**L16PenL26AL30C and **GRAND**L12L_DL16CL26AL30C.[120, 121] These peptides, containing two binding sites in close proximity (~ 20 Å) within the same three-stranded coiled coil, are capable of binding two Cd(II) ions with different coordination geometries (three coordinate, trigonal planar as CdS_3 and four coordinate, pseudotetrahedral as CdS_3O) and thus diaplay different physical properties for the metal centers. Furthermore, these peptides show site-selective Cd(II) recognition, where in both peptides, Cd(II) binds selectively to the 4-coordinate site (CdS_3O) regardless of the pH conditions.

While heterochromic peptides allowed us to discriminate between 3 and 4-coordination sites that were at first glance nearly identical, this success raised an important new question: Are all 3-coordinate or 4-coordinate sites in the same peptide identical? Said differently, can we refer to, as an example, an environment simply as a 4-coordinate site or must we specify whether this sulfur rich environment is in the center of the helix versus an end? And, if there is a

difference, will the physical properties of the center be modified, the selectivity for binding sites or the dynamics of the structures be different? Ultimately, one can address whether the position of these sites in the coiled coil play any role in fine tuning the physical properties of the bound Cd(II).

In the third chapter of my thesis, I address these issues by sets of designed single peptides capable of binding two equivalents of Cd(II) with the same coordination geometry. A construct such as this has two 3-coordinate sites (CdS_3) or two 4-coordinate sites (CdS_3O) at different positions in the coiled coil. These peptides are used to assess the selectivity of metal binding between two equivalent positions within the peptide, whether binding of one metal would influence the physical properties of the second bound metal and if different physical properties can be observed for the same metal in two similar first coordination environments but in different topological positions within the coiled coil. Furthermore, I address the important issues of how the specific position of these sites could modify the intrinsic properties of the bound Cd(II).

Metal ion homeostasis and transport are important cellular processes for maintenance of different biochemical activities. Essential metal ions are safely delivered to their partners by assistance with metallochaperones.[135-139] Toxic metal ions such as Hg(II), Cd(II), As(III), and Pb(II) that a cell encounters must be detoxified before they are released back to cellular environment. One important factor that defines the exchange processes between metal ions and biological macromolecules is the dynamics involving metal exchange. The rates of metal incorporation and removal are essential parts for proper functioning of these proteins. However, little is known about the fundamental rates of metal ion exchange and a complete understanding of the mechanistic pathways in which metals insert in and out of the proteins has been elusive and is of fundamental interest to the bioinorganic community. In the fourth chapter of my thesis I attempt to understand the exchange of Cd(II) using designed GRAND peptides as models for more complicated native proteins such as metallochaperones, and

metalloregulatory proteins such as CmtR which belongs to the ArsR/SmtB metal sensing family and senses Cd(II) and Pb(II). The Pecoraro group has previously investigated the insertion of Hg(II), and Cd(II) into well-structures TRI peptides that exist as parallel three-stranded coiled coils even in the absence of metals. Studies with insertion of Hg(II) **TRI**L9C was shown to initially follow the "Dissociation Mechanism" followed by the "StepAD" mechanism.[116] In case of Cd(II) the insertion mechanism followed the "Breathing" mechanism.[116] Although these studies were significant, the following questions need to be answered to obtain a better molecular level understanding of the metal insertion process into helical scaffolds: Do the metals first interact with amino acids located at the surface of the peptide before being inserted into the coiled coil interior? Do the cysteines at the designed metal binding site play any role in assisting the insertion of the metal from the surface to the hydrophobic interior once the metal is bound to the surface residues? What are the rates of metal exchange? Do the rates depend on the location of the metal binding site at the interior of the coiled coils (middle of the helix vs. towards the helical terminus)? Is there any impact of the metal's coordination number on the exchange dynamics (4-coordinate CdS_3O, vs. 3-coordinate CdS_3)? To address these questions, in the fourth chapter I have investigated the exchange of Cd(II) using the GRAND series of peptides containing both dual and single metal binding sites.

2) Designing metal binding sites in single-chain helical bundles.

As discussed above, numerous metalloprotein systems have been designed over the past 15 years typically using unassociated peptides that self-assemble into three- or four- stranded coiled coils and helix-loop-helix motifs that form antiparallel four-stranded bundles. It is often difficult to design non-symmetrical metal sites using these systems due to the symmetry systems relying on homo-oligomerization. Preparing metalloproteins with asymmetric metal binding sites would allow us to mimic active sites of many natural proteins

such as the blue copper proteins plastocyanin,[140, 141] azurin,[142-144] introducing a single specific hydrogen bonding interaction to the metal site such as in carbonic anhydrase,[145, 146] as well as making catalytic metalloproteins. There has been an attempt to design a heterotrimeric coiled coil intended to form a two-His, one-Cys trigonal site by DeGrado *et. al*.[147] Even though the intended heterotrimer was formed, the construct was unable to incorporate metal ions. The only successful incorporation of metal ions in a heterooligomeric system is the heterotetrameric assembly of DF2tet $A_aA_bB_2$ as described earlier. Due to very limited success in preparing heterooligomeric systems involving single self-associating peptides, preparing a single polypeptide chain capable of binding metals in predefined coordination geometries has been a highly sought-after objective by which to impart proteins with novel properties and activities.[11, 148-151] However, preparing novel metal binding sites into *de novo* designed single-chain helical bundles have proven to be challenging.

A three-His Zn(II) site has been introduced into the designed protein α_4, an idealized four-helix bundle protein in which four identical helices were connected by three identical loops.[1] Three surface His residues were introduced into the protein, two on a single helix, with the His-X_3-His arrangement, and the third on a neighboring helix.[152] In the presence of Zn(II) the protein was stabilized against chemical denaturation as well as tightening the somewhat molten protein structure, as evidenced by ^{1}H NMR spectroscopy.[153] Metal chelation by all three His residues were confirmed by observation of chemical shift changes of the ring protons of His residues, however, the coordination geometry of the site was not reported. A 61-residue all-β protein has been designed where 6 β-strands were connected by loops. A three-His metal binding site was designed in two neighboring loop regions (Figure 1-16), resulting in a folded, compact construct designated as ‘minibody’. This construct was able to bind metal ions such as Cu(II), Zn(II), Cd(II), and Co(II).[105] Detailed characterization of the system was inhibited due to low solubility of the protein (10 μM). Subsequent redesign of the protein led to increased solubility of up to

milimolar level.[154] A tetrahedral metal binding site formed by two Cys and two His ligands on adjacent helices of the *de novo* designed four-helix bundle α_4 was introduced by Regan *et. al* (Figure 1-17).[155] The protein Zα_4 was shown to bind Zn(II), and Co(II) with high affinity [$k_d = 2.5 \times$ M for Zn(II), and $1.6 \times$ M for Co(II)] in tetrahedral coordination geometry. Circular dichroism studies showed no significant changes in secondary structure between the metal-bound and apo protein. However, metal binding significantly stabilized the protein to chemical denaturation. Subsequent studies with different mutants showed that all four ligands contributed to the high affinity for metal binding.[156] They also designed a tetrahedral His_3Cys Zn(II) binding site in a small protein of known structure: the B1 domain of protein G, a 56-resdiue domain which is composed of four strands of β-sheet crossed by a single α-helix.[157] The metal site was formed by two His residues from a β-strand (His-X-His), and one His and one Cys residue from the α-helix (His-X_2-Cys).[158] Molecular modeling identified a potential steric clash between a metal ligand and the Leu 5. Three variants were therefore prepared with Leu 5 either retained (Zβ1L) or substituted to Ala (Zβ1A) or Met (Zβ1M). Figure 1-18 shows a model of Zβ1M. All three variants bound Co(II) tightly, with k_d values ranging from 4 ± 1 to 18 ± 3 μM. Zn(II) bound 1000 times more tightly compared to Co(II). The presence of Zn(II), and Cd(II) induced folding of the Ala, and Met variants but not of the Leu variant, which was well-folded even in the absence of metals.

Hellinga and colleagues used computer search of the protein thioredoxin to identify a set of backbone positions arranged such that, if appropriate side chains were substituted, a Type-1 blue copper site could be formed.[159] In Type I blue copper sites, a Cu(II) ion is chelated by one Cys (Sγ), two His (Nδ) and one Met (Sδ) residue.[160] The affinity of the re-designed thioredoxin for Cu(II) was determined by monitoring the change in protein's intrinsic Trp fluorescence induced by metal binding. Cu(II) was shown to bind weakly to the re-designed protein ($k_d = 4 \times 10^{-5}$ M).[161] Typical Type I blue copper sites have characteristic

intense band at 600 nm for the resulting from charge transfer transitions from Cys thiolate/ Met thioether to Cu(II).[160] The re-designed protein showed an unusually broad absorbance over the entire range of the visible region, suggesting that either the Cu(II)-thiolate and Cu(II)-thioether bonds were missing, or that the protein-metal complex had several different forms.[161] Electron paramagnetic resonance spectroscopic studies suggested that the Cu(II) might be coordinated to two nitrogen (His) and two oxygen ligands (acidic side chains from protein / water molecules).

The existing examples as discussed above, where metal binding sites have been incorporated into *de novo* designed single-chain polypeptides involves either the α/β structural framework or four-helix bundles including the DFsc protein[92] mentioned earlier in this chapter. However, there are no examples where novel metal binding sites have been incorporated in three-helix bundles. An attractive scaffold is the *de novo* designed three-helix bundle α_3D, the solution structure of which was solved by DeGrado and coworkers.[162] It was shown that the helices of this 73 amino acid protein were oriented in a counter-clockwise topology (Figure 1-19). Although the α_3D protein originated from a coiled-coil, its helices were shortened to the point where it might be better considered as a globular protein whose repetitive structure makes each of the heptads very similar to one another. The stability of α_3D (5.3 kcal mole^{-1}) is similar to that of natural proteins and thus α_3D should be tolerant to mutations, allowing this protein to serve as an excellent framework to engineer specific metal binding sites. Additionally, with this protein scaffold one can study the effect of the ligating residue located on the second helix which is antiparallel to the first and third helices of the bundle.

Studies have been performed to test the malleability and adaptability of α_3D by mutating Ala 60 in its core to larger, hydrophobic side chains Leu and Ile.[163] Such changes introduce strain into the structure of natural proteins, and therefore generally destabilize the native state. By contrast, these mutations

were slightly stabilizing (~1.5 kcal mole^{-1}) to the tertiary structure of α_3D. The value of ΔC_p for unfolding of these mutants was not greatly affected relative to the wild type protein, indicating that the change in solvent accessibility for unfolding was similar. However, two-dimensional heteronuclear single quantum coherence spectra indicated that the protein adjusts to the introduction of steric bulk in different ways. A60L mutant showed serious erosion in the dispersion of both the amide backbone as well as the side-chain methyl chemical shifts. By contrast, A60I showed excellent dispersion of the backbone resonances, and selective changes in dispersion of the aliphatic side-chains proximal to the site of mutation. Together, these data suggested that the α_3D, although folded into a unique three-dimensional structure, is nevertheless more malleable and flexible than most natural, native proteins. The backbone and side-chain dynamics of α_3D were investigated using ^{15}N, ^{13}C, and ^{2}H nuclear magnetic resonance relaxation methods with the aim of assessing the character of the internal motions of this native-like protein of *de novo* design.[164] At the backbone level, both ^{15}N and $^{13}C_\alpha$ relaxation studies indicate highly restrictive motion on the picosecond to nanosecond time scale in the α-helical regions of α_3D, with increasing mobility at the ends of the α-helices and in the two loop regions. This is largely consistent with what is seen in proteins of natural origin. Overall, the view provided by both ^{2}H and ^{13}C methyl relaxation methods suggest that the side chains of α_3D are more dynamic compared to natural proteins. Regions of relative flexibility bound clusters of rigid methyl-bearing side-chain groups that are interspersed with aromatic and β-branched amino acids. The time scale of motions associated with methyl-bearing side chains of α_3D are significantly longer than that seen in natural proteins. These results indicate that the strategies underlying the design of α_3D have largely, but not completely, captured both the structural and dynamic character of natural proteins. Hydration of amides of α_3D was examined by molecular dynamics simulation, IR, and NMR.[165] Molecular dynamics calculations show that the amide carbonyls on the surface of the protein tilt away from the helical axis to interact with solvent water,

resulting in a lengthening of the hydrogen bonds on this face of the helix. Water molecules are bonded to these carbonyl groups with partial occupancy (~ 50%–70%), and their interaction geometries show a large variation in their hydrogen bond lengths and angles on the nsec time scale. This heterogeneity is reflected in the carbonyl stretching vibration (amide I' band) of a group of surface Ala residues. The surface-exposed amides are broad, and shift to lower frequency (reflecting strengthening of the hydrogen bonds) as the temperature is decreased. By contrast, the amide I' bands of the buried ^{13}C-labeled Leu residues are significantly sharper and their frequencies are consistent with the formation of strong hydrogen bonds, independent of temperature. The rates of hydrogen-deuterium exchange and the proton NMR chemical shifts of the helical amide groups also depend on environment. The partial occupancy of the hydration sites on the surface of helices suggests that the interaction is relatively weak, on the order of thermal energy at room temperature. One unexpected feature that emerged from the dynamics calculations was that a Thr side chain subtly disrupted the helical geometry 4–7 residues N-terminal in sequence, which was reflected in the proton chemical shifts and the rates of amide proton exchange for several amides that engage in a mixed $3_{10}/\alpha/\pi$-helical conformation. The folding/unfolding kinetics of α_3D, has been described.[166] Both IR temperature jump and ultrafast fluorescence mixing methods reveal a single-exponential process consistent with a minimal folding time of 3.2 ± 1.2 μs (at ~ 50℃), indicating that a protein can fold on the 1- to 5-μs time scale. Furthermore, the single-exponential nature of the relaxation indicates that the prefactor for transition state (TS)-folding models is probably ≥1 $(\mu s)^{-1}$ for a protein of this size and topology. Molecular dynamics simulations and IR spectroscopy provide a molecular rationale for the rapid, single-exponential folding of this protein. α_3D shows a significant bias toward local helical structure in the thermally denatured state. The molecular dynamics-simulated transition state ensemble is highly heterogeneous and dynamic, allowing access to the transition state via multiple pathways.

Before attempting to prepare asymmetric metal coordination environments or site specific H-bonds, it is important to first redesign α_3D so as to introduce symmetric metal binding sites involving cysteine residues. This approach would allow us to exploit the extensive body of work defining heavy metal complexation using the TRI and Coil Ser peptides, to assess whether a specific metal structure could be achieved in the modified α_3D construct. It would also allow us to probe the physical properties of metals such as Cd(II), Hg(II) and Pb(II) in a more natural antiparallel helical system. In the fifth chapter of my thesis I have described successful design and preparation of a triple-Cys derivative of α_3D, designated as α_3D**IV**. Based on visual inspection of the α_3D structure, four potential sites along the bundle were identified where three cysteines, one from each helix, could be introduced. Out of these four mutants, the sites in one such mutant, α_3D**IV**, located at the C-terminal end of the bundle, seemed to be optimal based on the properties of the starting protein as described earlier.[163, 164, 166, 167] The three-Cys site has been engineered at the C-terminal end of the bundle by mutating a Leu residue from each helix to a Cys. Previous NMR structural and dynamic investigations showed a gradient in the dynamic behavior and malleability of the protein, with the C-terminal end of the bundle being most amenable to amino acid substitutions. The selected location has a well-ordered backbone conformation, but the side chains of the mutated residues are less well ordered than residues in other locations of the bundle. The 3-Cys site, which is largely sequestered from solvent, occupies a “box” (Figure 1-20); the sides are formed by the backbone of the helices and the bottom by the apolar side chains of Phe31, Ile14, and Ile63. The aromatic residue of Phe31 lies directly over the predicted metal-binding site and lines most of the bottom of the box. The top is formed by the main-chain and side chains of residues in the non-helical loops including Leu21, as well as the apolar portion of Tyr 70 at the terminus of helix 3. His72, which was entirely disordered in the NMR solution structure, also lies proximal to the site. Moreover, after introduction of the Cys side chains in one of the two preferred rotamers for Cys, the thiol S^G atoms formed a nearly equilateral

triangle with the side chains well oriented to form the desired site (inter S^G distances = 3.5 -4.5 Å). Overall, the location is ideal to explore the effects of hydrophobic sequestration in the present work. Characterization of apo α_3D**IV** as well as different metallated complexes of this protein is described in the presence of Hg(II), Cd(II) and Pb(II).

In the sixth chapter of this thesis I conclude the key achievements from my work presented in each chapter and proposing future directions of the project.

References:

[1] L. Regan, W. F. DeGrado, *Science* **1988**, *241*, 976.
[2] W. F. DeGrado, Z. R. Wasserman, J. D. Lear, *Science* **1989**, *243*, 622.
[3] J. W. Bryson, S. F. Betz, Z. X. Lu, D. J. Suich, H. X. Zhou, *Science* **1995**, *270*, 935.
[4] L. Baltzer, H. Nilsson, J. Nilsson, *Chem. Rev.* **2001**, *101*, 3153.
[5] W. F. DeGrado, C. M. Summa, V. Pavone, F. Nastri, A. Lombardi, *Annu. Rev. Biochem.* **1999**, *68*, 779.
[6] S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, in *Science, Vol. v262*, **1993**, p. p1680(6).
[7] S. Kamtekar, M. H. Hecht, *Faseb Journal* **1995**, *9*, 1013.
[8] Y. Fezoui, D. L. Weaver, J. J. Osterhout, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 3675.
[9] D. Riddle, J. V. Santiago, S. T. Bray-Hall, N. Doshi, V. P. Grantcharova, Q. Yi, D. Baker, *Nature Structural Biology* **1997**, *4*, 805.
[10] J. R. Desjarlais, N. D. Clarke, *Current Opinion in Structural Biology* **1998**, *8*, 471.
[11] H. W. Hellinga, *Folding & Design* **1998**, *3*, R1.
[12] Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855.
[13] Y. Lu, S. M. Berry, T. D. Pfister, *Chem. Rev.* **2001**, *101*, 3047.
[14] G. Tuchscherer, L. Scheibler, P. Dumy, M. Mutter, *Biopolymers* **1998**, *47*, 63.
[15] L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355.
[16] W. F. DeGrado, *Nature* **2003**, *423*, 132.
[17] Y. Lu, *Inorg. Chem.* **2006**, *45*, 9930.
[18] R. L. Koder, P. L. Dutton, *Dalton Trans.* **2006**, 3045.
[19] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 11566.
[20] D. Ghosh, Ph.D. thesis, University of Michigan (Ann Arbor), **2006**.
[21] M. R. Ghadiri, C. Choi, *J. Am. Chem. Soc.* **1990**, *112*, 1630.
[22] M. R. Ghadiri, K. Fernholz, *J. Am. Chem. Soc.* **1990**, *112*, 9633.
[23] M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, *114*, 825.
[24] M. R. Ghadiri, M. A. Case, *Angewandte Chemie International Edition in English* **1993**, *32*, 1594.
[25] M. R. Ghadiri, C. Soares, C. Choi, *J. Am. Chem. Soc.* **1992**, *114*, 4000.
[26] O. A. Kharenko, M. Y. Ogawa, *J. Inorg. Biochem.* **2004**, *98*, 1971.
[27] J. Hong, O. A. Kharenko, M. Y. Ogawa, *Inorg. Chem.* **2006**, *45*, 9974.
[28] O. A. Kharenko, D. C. Kennedy, B. Demeler, M. J. Maroney, M. Y. Ogawa, *Journal of the American Chemical Society* **2005**, *127*, 7678.
[29] M. V. Tsurkan, M. Y. Ogawa, *Inorganic Chemistry* **2007**, *46*, 6849.
[30] K. Suzuki, H. Hiroaki, D. Kohda, T. Tanaka, *Protein Eng.* **1998**, *11*, 1051.
[31] K. Suzuki, H. Hiroaki, D. Kohda, H. Nakamura, T. Tanaka, *J. Am. Chem. Soc.* **1998**, *120*, 13008.

[32] X. Q. Li, K. Suzuki, K. Kanaori, K. Tajima, A. Kashiwada, H. Hiroaki, D. Kohda, T. Tanaka, *Protein Sci.* **2000**, *9*, 1327.
[33] T. Kiyokawa, K. Kanaori, K. Tajima, M. Koike, T. Mizuno, J. Oku, T. Tanaka, *J. Pept. Res.* **2004**, *63*, 347.
[34] B. T. Farrer, N. P. Harris, K. E. Balchus, V. L. Pecoraro, *Biochemistry* **2001**, *40*, 14696.
[35] B. Farrer, V. L. Pecoraro, *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3760.
[36] M. Siedlecka, G. Goch, A. Ejchart, S. H., A. Bierzynski, *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 903.
[37] C. O. Pabo, E. Peisach, R. A. Grant, *Annual Review of Biochemistry* **2001**, *70*, 313.
[38] J. M. Berg, *Annual Review of Biophysics and Biophysical Chemistry* **1990**, *19*, 405.
[39] B. A. Krizek, B. T. Amann, V. J. Kilfoil, D. L. Merkle, J. M. Berg, *J. Am. Chem. Soc.* **1991**, *113*, 4518.
[40] A. D. Frankel, J. M. Berg, C. O. Pabo, *Proceedings of the National Academy of Sciences of the United States of America* **1987** *84* 4841.
[41] M. Lee, G. Gippert, K. Soman, D. Case, P. Wright, *Science* **1989** *245* 635.
[42] G. PÃ¡rraga, S. Horvath, L. Hood, E. T. Young, R. E. Klevit, *Proceedings of the National Academy of Sciences of the United States of America* **1990** *87* 137.
[43] R. E. Klevit, J. R. Herriott, S. J. Horvarth, *Proteins: Structure, Function, and Bioinformatics* **1990**, *7*, 215.
[44] M. A. Weiss, K. A. Mason, C. E. Dahl, H. T. Keutmann, *Biochemistry* **1990**, *29*, 5660.
[45] J. M. Berg, H. A. Godwin, *Annu. Rev. Biophys. Biomol. Struct.* **1997**, *26*, 357.
[46] J. M. Berg, *Acc. Chem. Res.* **1995**, *28*, 14.
[47] J. Berg, *J. Biol. Chem.* **1990**, *265*, 6513.
[48] B. A. Krizek, D. L. Merkle, J. M. Berg, *Inorg. Chem* **1993**, *32*, 937.
[49] C. T. Choma, J. D. Lear, M. J. Nelson, P. L. Dutton, D. E. Robertson, W. F. DeGrado, *Journal of the American Chemical Society* **1994**, *116*, 856.
[50] D. E. Robertson, R. S. Farid, C. C. Moser, J. L. Urbauer, S. E. Mulholland, R. Pidikiti, J. D. Lear, A. J. Wand, W. F. Degrado, P. L. Dutton, *Nature* **1994**, *368*, 425.
[51] R. E. Sharp, J. R. diers, D. F. Bocian, P. L. Dutton, *J. Am. Chem. Soc.* **1998**, *120*, 7103.
[52] B. R. Gibney, Y. Isogai, F. Rabanal, K. S. Reddy, A. M. Grosset, C. C. Moser, P. L. Dutton, *Biochemistry* **2000**, *39*, 11041.
[53] C. J. Reedy, M. L. Kennedy, B. R. Gibney, *Chem. Commun.* **2003**, 570.
[54] C. J. Reedy, B. R. Gibney, *Chem. Rev.* **2004**, *104*, 617.
[55] P. A. Arnold, D. R. Benson, D. J. Brink, M. P. Hendrich, G. S. Jas, M. L. Kennedy, D. T. Petasis, M. X. Wang, *Inorg. Chem* **1997**, *36*, 5306.
[56] D. R. Benson, B. R. Hart, X. Zhu, M. B. Doughty, *Journal of the American Chemical Society* **1995**, *117*, 8502.

[57] D. H. Liu, D. A. Williamson, M. L. Kennedy, T. D. Williams, M. M. Morton, D. R. Benson, *J. Am. Chem. Soc.* **1999**, *121*, 11798.
[58] D. L. Huffman, M. M. Rosenblatt, K. S. Suslick, *J. Am. Chem. Soc.* **1998**, *120*, 6183.
[59] D. L. Huffman, K. S. Suslick, *Inorg. Chem* **2000**, *39*, 5418.
[60] M. M. Rosenblatt, D. L. Huffman, X. Wang, H. A. Remmer, K. S. Suslick, *J. Am. Chem. Soc.* **2002**, *124*, 12394.
[61] G. Ghirlanda, A. Osyczka, W. Liu, M. Antolovich, K. M. Smith, P. L. Dutton, A. J. Wand, W. F. DeGrado, *Journal of the American Chemical Society* **2004**, *126*, 8141.
[62] F. V. Cochran, S. P. Wu, W. Wang, V. Nanda, J. G. Saven, M. J. Therien, W. F. DeGrado, *Journal of the American Chemical Society* **2005**, *127*, 1346.
[63] G. M. Bender, A. Lehmann, H. Zou, H. Cheng, H. C. Fry, D. Engel, M. J. Therien, J. K. Blasie, H. Roder, J. G. Saven, W. F. DeGrado, *Journal of the American Chemical Society* **2007**, *129*, 10732.
[64] S. E. Mulholland, B. R. Gibney, F. Rabanal, P. L. Dutton, *J. Am. Chem. Soc.* **1998**, *120*, 10296.
[65] S. E. Mulholland, B. R. Gibney, F. Rabanal, P. L. Dutton, *Biochemistry* **1999**, *38*, 10442.
[66] J. R. Calhoun, F. Nastri, O. Maglio, V. Pavone, L. A.;, W. F. DeGrado, *Peptide Sci.* **2005**, *80*, 264.
[67] F. Frolow, A. J. Kalb (Gilboa), J. Yariv, **1994**, *1*, 453.
[68] E. C. Thiel, *Ferritin and Iron Biomineralization*, Elsevier Science, New York, **1995**.
[69] M. A. Carrondo, **2003**, *22*, 1959.
[70] A. C. Rosenzweig, C. A. Frederick, S. J. Lippard, P. Nordlund, *Nature* **1993**, *366*, 537.
[71] J. D. Lipscomb, *Annual Review of Microbiology* **1994**, *48*, 371.
[72] M.-H. Baik, M. Newcomb, R. A. Friesner, S. J. Lippard, *Chemical Reviews* **2003**, *103*, 2385.
[73] G. M. Whited, D. T. Gibson, *J. Bacteriol.* **1991**, *173*, 3010.
[74] J. D. Pikus, J. M. Studts, C. Achim, K. E. Kauffmann, E. Munck, R. J. Steffan, K. McClay, B. G. Fox, *Biochemistry* **1996**, *35*, 9106.
[75] M. H. Sazinsky, J. Bard, A. Di Donato, S. J. Lippard, *Journal of Biological Chemistry* **2004** *279* 30600.
[76] V. V. Barynin, M. M. Whittaker, S. V. Antonyuk, V. S. Lamzin, P. M. Harrison, P. J. Artymiuk, J. W. Whittaker, *Structure (London, England : 1993)* **2001**, *9*, 725.
[77] A. J. Wu, J. E. Penner-Hahn, V. L. Pecoraro, *Chemical Reviews* **2004**, *104*, 903.
[78] Y. Lindqvist, W. J. Huang, G. Schneider, J. Shanklin, *Embo Journal* **1996**, *15*, 4081.
[79] B. G. Fox, K. S. Lyle, C. E. Rogge, *Accounts of Chemical Research* **2004**, *37*, 421.

[80] P. Nordlund, H. Eklund, *Journal of Molecular Biology* **1993**, *232*, 123.
[81] J. Stubble, *Current Opinion in Chemical Biology* **2003**, *7*, 183.
[82] M. Kolberg, K. R. Strand, P. Graff, K. K. Andersson, *Biochimica Et Biophysica Acta-Proteins and Proteomics* **2004**, *1699*, 1.
[83] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci., USA* **2000**, *97*, 6298.
[84] O. Maglio, F. Nasti, V. Pavone, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3772.
[85] L. Di Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2001**, *123*, 12749.
[86] A. Pasternak, J. Kaplan, J. D. Lear, W. F. Degrado, *Protein Science* **2001**, *10*, 958.
[87] S. J. Lahr, D. E. Engel, S. E. Stayrook, O. Maglio, B. North, S. Geremia, A. Lombardi, W. F. DeGrado, *J . Mol. Biol.* **2005**, *346*, 1441.
[88] W. F. DeGrado, L. Di Costanzo, S. Geremia, A. Lombardi, V. Pavone, L. Randaccio, *Angewandte Chemie International Edition* **2003**, *42*, 417.
[89] C. M. Summa, M. M. Rosenblatt, J.-K. Hong, J. D. Lear, W. F. DeGrado, *Journal of Molecular Biology* **2002**, *321*, 923.
[90] E. N. G. Marsh, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **2002**, *99*, 5150.
[91] J. G. Saven, *Current Opinion in Structural Biology* **2002**, *12*, 453.
[92] J. R. Calhoun, H. Kono, S. Lahr, W. Wang, W. F. DeGrado, J. G. Saven, *Journal of Molecular Biology* **2003**, *334*, 1101.
[93] M. Faiella, C. Andreozzi, R. T. M. de Rosales, V. Pavone, O. Maglio, F. Nastri, W. F. DeGrado, A. Lombardi, *Nat. Chem. Biol* **2009**, *5*, 882.
[94] M. D. Struthers, R. P. Cheng, B. Imperiali, *J. Am. Chem. Soc.* **1996**, *118*, 3073.
[95] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342.
[96] C. L. Hill, D. J. Steenkamp, R. H. Holm, T. P. Singer, *Proc. Natl. Acad. Sci., U.S.A.* **1977**, *74*, 547.
[97] B. R. Gibney, F. Rabanal, J. J. Skalicky, A. J. Wand, P. L. Dutton, *Journal of the American Chemical Society* **1999**, *121*, 4952.
[98] R. G. Daugherty, T. Wasowicz, B. R. Gibney, V. J. DeRose, *Inorg. Chem.* **2002**, *41*, 2623.
[99] R. P. Bonomo, L. Casella, L. D. Gioia, H. Molinari, G. Impellizzeri, T. Jordan, G. Pappalardo, E. Rizzarelli, *J. Chem. Soc. Dalton Trans.* **1997**, 2387.
[100] R. P. Cheng, S. L. Fisher, B. Imperiali, *J. Chem. Soc. Dalton Trans.* **1996**, *118*, 11349.
[101] M. M. Rosenblatt, J. Wang, K. S. Suslick, *Proceedings of the National Academy of Sciences of the United States of America* **2003** *100* 13140.
[102] J. P. Schneider, J. W. Kelly, *Journal of the American Chemical Society* **1995**, *117*, 2533.
[103] G. Platt, C. Chung, M. Searle, *Chem. Commun.* **2001**, 1162.

[104] J. Venkatraman, G. A. Naganagowda, R. Sudha, P. Balaram, *Chem. Commun.* **2001**, 2660.
[105] A. Pessi, E. Bianchi, A. Crameri, S. Venturini, A. Tramontano, M. Sollazzo, **1993**, *362*, 367.
[106] J. G. Wright, M. J. Natan, F. M. MacDonnell, D. M. Ralston, T. V. O'Halloran, *Progress in Inorganic Chemistry:Bioinorganic Chemistry* **1990**, *38*, 323.
[107] J. G. Wright, H.-T. Tsang, J. E. Penner-Hahn, T. V. O'Halloran, *J. Am. Chem. Soc.* **1990**, *112*, 2434.
[108] J. Ye, A. Kandegedara, P. Martin, B. P. Rosen, *J. Bacteriol.* **2005**, *187*, 4214.
[109] L. Banci, I. Bertini, F. Cantini, S. Ciofi-Baffoni, J. S. Cavet, C. Dennison, A. I. Graham, D. R. Harvie, N. J. Robinson, *Journal of Biological Chemistry* **2007** *282* 30181.
[110] L. S. Busenlehner, T. C. Weng, J. E. Penner-Hahn, D. P. Giedroc, *J. Mol. Biol.* **2002**, *319*, 685.
[111] L. S. Busenlehner, N. J. Cosper, R. A. Scott, B. P. Rosen, M. D. Wong, D. P. Giedroc, *Biochemistry* **2001**, *40*, 4426.
[112] W. Shi, J. Dong, R. A. Scott, M. Y. Ksenzenko, B. P. Rosen, *J. Biol. Chem.* **1996**, *271*, 9291.
[113] Y. Sun, M. Wong, C. Stalhandske, R. A. Scott, B. P. Rosen, *FASEB J.* **1999**, *13*, A1464.
[114] B. Borremans, J. L. Hobman, A. Provoost, N. L. Brown, D. van der Lelie, *J. Bacteriol.* **2001**, *183*, 5651.
[115] G. R. Dieckmann, D. K. McRorie, J. D. Lear, K. A. Sharp, W. F. DeGrado, V. L. Pecoraro, *J. Mol. Biol.* **1998**, *280*, 897.
[116] D. Ghosh, V. L. Pecoraro, *Inorg. Chem.* **2004**, *43*, 7902.
[117] B. Lovejoy, S. Choe, D. Cascio, D. McRorie, W. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288.
[118] G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, V. L. Pecoraro, *J. Am. Chem. Soc.* **1997**, *119*, 6195.
[119] D. S. Touw, C. E. Nordman, J. A. Stuckey, V. L. Pecoraro *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969.
[120] K.-H. Lee, C. Cabello, L. Hemmingsen, E. N. G. Marsh, V. L. Pecoraro, *Angew. Chem., Int. Ed.* **2006**, *45*, 2864.
[121] A. F. A. Peacock, L. Hemmingsen, V. L. Pecoraro, *Proceedings of the National Academy of Sciences* **2008**, *105*, 16566.
[122] O. Iranzo, C. Cabello, V. L. Pecoraro, *Angew. Chem., Int. Ed.* **2007**, *46*, 6688.
[123] B. Farrer, C. McClure, J. E. Penner-Hahn, V. L. Pecoraro, *Inorg. Chem* **2000**, *39*, 5422.
[124] M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2002**, *124*, 8042.

[125] M. Matzapetakis, D. Ghosh, T.-C. Weng, J. E. Penner-Hahn, V. L. Pecoraro, *J. Biol. Inorg. Chem.* **2006**, *11*, 876.
[126] M. Matzapetakis, V. L. Pecoraro, *J. Am. Chem. Soc.* **2005**, *127*, 18229.
[127] O. Iranzo, Lee, K.H., Jakusch, T., Hemmingsen, L., Pecoraro, V. L., *Chem. Eur. J.,* **2009**, *15*, 3761.
[128] O. Iranzo, P. V. Thulstrup, S.-B. Ryu, L. Hemmingsen, V. L. Pecoraro, *Chem. Eur. J.* **2007**, *13*, 9178.
[129] H. Wendt, C. Berger, A. Baici, R. M. Thomas, H. R. Bosshard, *Biochem.* **1995**, *34*, 4097.
[130] A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, *Angewandte Chemie International Edition* **2009**, *48*, 7371.
[131] M. Matzapetakis, University of Michigan (Ann Arbor), **2005**.
[132] D. S. Touw, *Ph.D Thesis* **2007**, *University of Michigan*.
[133] K. P. Neupane, V. L. Pecoraro, *Angewandte Chemie International Edition* **2010**, *49*, 8177.
[134] K. H. Lee, M. Matzapetakis, S. Mitra, E. N. G. Marsh, V. L. Pecoraro, *J. Am. Chem. Soc.* **2004**, *126*, 9178.
[135] A. C. Rosenzweig, *Accounts of Chemical Research* **2000**, *34*, 119.
[136] A. C. Rosenzweig, T. V. O'Halloran, *Curr. Opin. Chem. Biol.* **2000**, *4*, 140.
[137] T. V. O'Halloran, V. C. Culotta, *J. Biol. Chem.* **2000**, *275*, 25057.
[138] A. K. Wernimont, D. L. Huffman, A. L. Lamb, T. V. O'Halloran, A. C. Rosenzweig, *Nat. Struct. Biol.* **2000**, *7*, 766.
[139] W. Xia, H. Li, K.-H. Sze, H. Sun, *Journal of the American Chemical Society* **2009**, *131*, 10031.
[140] J. M. Guss, P. R. Harrowell, M. Murata, V. A. Norris, H. C. Freeman, *Journal of Molecular Biology* **1986**, *192*, 361.
[141] J. Mitchell Guss, H. C. Freeman, *Journal of Molecular Biology* **1983**, *169*, 521.
[142] E. N. Baker, *Journal of Molecular Biology* **1988**, *203*, 1071.
[143] E. T. Adman, R. E. Stenkamp, L. C. Sieker, L. H. Jensen, *Journal of Molecular Biology* **1978**, *123*, 35.
[144] Z. R. Korszun, *Journal of Molecular Biology* **1987**, *196*, 413.
[145] W. S. Sly, P. Y. Hu, *Annual Review of Biochemistry* **1995**, *64*, 375.
[146] H. F. Deutsch, *International Journal of Biochemistry* **1987**, *19*, 101.
[147] A. Lombardi, J. W. Bryson, W. F. DeGrado, *Biopolymers* **1997**, *40*, 495.
[148] L. Regan, *Trends in Biochemical Sciences* **1995**, *20*, 280.
[149] L. Regan, *Annual Review of Biophysics and Biomolecular Structure* **1993**, *22*, 257.
[150] H. W. Hellinga, *Current Opinion in Biotechnology* **1996**, *7*, 437.
[151] J. A. Tainer, V. A. Roberts, E. D. Getzoff, *Current Opinion in Biotechnology* **1992**, *3*, 378.
[152] T. Handel, W. F. DeGrado, *Journal of the American Chemical Society* **1990**, *112*, 6710.
[153] T. M. Handel, S. A. Williams, W. F. DeGrado, *Science* **1993**, *261*, 879.

[154] E. Bianchi, S. Venturini, A. Pessi, A. Tramontano, M. Sollazzo, *Journal of Molecular Biology* **1994**, *236*, 649.
[155] L. Regan, N. D. Clarke, *Biochemistry* **1990**, *29*, 10878.
[156] M. Klemba, L. Regan, *Biochemistry* **1995**, *34*, 10094.
[157] A. M. Gronenborn, D. R. Filpula, N. Z. Essig, A. Achari, M. Whitlow, P. T. Wingfield, G. M. Clore, *Science*, *253*, 657.
[158] M. G. Klemba, Kevin H.|Marino,Stephen|Clarke,Neil D.|Regan,Lynne, in *Nature structural biology, Vol. 2*, **1995**, p. 368.
[159] H. W. Hellinga, F. M. Richards, *J . Mol. Biol.* **1991**, *222*, 763.
[160] R. H. Holm, P. Kennepohl, E. I. Solomon, *Chemical Reviews* **1996**, *96*, 2239.
[161] H. W. Hellinga, J. P. Caradonna, F. M. Richards, *Journal of Molecular Biology* **1991**, *222*, 787.
[162] S. T. R. Walsh, H. Cheng, J. W. Bryson, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **1999**, *96*, 5486.
[163] S. T. R. Walsh, V. I. Sukharev, S. F. Betz, N. L. Vekshin, W. F. DeGrado, *Journal of Molecular Biology* **2001**, *305*, 361.
[164] S. T. R. Walsh, A. L. Lee, W. F. DeGrado, A. J. Wand, *Biochemistry* **2001**, *40*, 9560.
[165] S. T. R. Walsh, R. P. Cheng, W. W. Wright, D. O. V. Alonso, V. Daggett, J. M. Vanderkooi, W. F. DeGrado, *Protein Science* **2003**, *12*, 520.
[166] Y. Zhu, D. O. V. Alonso, K. Maki, C.-Y. Huang, S. J. Lahr, V. Daggett, H. Roder, W. F. DeGrado, F. Gai, *Proceedings of the National Academy of Sciences of the United States of America* **2003** *100* 15486.
[167] S. T. Walsh, R. P. Cheng, W. W. Wright, D. O. Alonso, V. Daggett, J. M. Vanderkooi, W. F. DeGrado, *Protein Science* **2003**, *12*, 520.

Table 1-1. Peptide sequences used in this study.

Peptide	Sequence
BABY	Ac-G LKALEEK LKALEEK LKALEEK G-NH_2
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
Coil Ser	Ac-EWEALEKK-LAALESK-LQALEKK-LEALEHG-NH_2
GRAND	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
BABYL9C	Ac-G LKALEEK **C**KALEEK LKALEEK G-NH_2
TRIL16C	Ac-G LKALEEK LKALEEK **C**KALEEK LKALEEK G-NH_2
TRIL12AL16C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LKALEEK G-NH_2
TRIL16Pen	Ac-G LKALEEK LKALEEK **X**KALEEK LKALEEK G-NH_2
GRANDL26AL30C	Ac-G LKALEEK LKALEEK LKALEEK LKA**A**EEK **C**KALEEK G-NH_2
GRANDL16PenL26AL30C	Ac-G LKALEEK LKALEEK **X**KALEEK LKA**A**EEK **C**KALEEK G-NH_2
GRANDL16PenL19IL23PenL26I	Ac-G LKALEEK LKALEEK **X**KA**I**EEK **X**KA**I**EEK CKALEEK G-NH_2
GRANDL12AL16CL26AL30C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LK**A**LEEK **C**KALEEK G-NH_2
GRANDL12L_D16CL26AL30C	Ac-G LKALEEK LKA**L_D**EEK **C**KALEEK LK**A**LEEK **C**KALEEK G-NH_2
α_3D**IV**	MGSWAEFKQRLAAIKTR**C**QALGGSEAE**C**AAFEKEIAAFESELQAYK GKGNPEVEALRKEAAAIRDE**C**QAYRHN

Residues in red indicate modifications to the parent peptide. X = Penicillamine. L_D = D-Leu.

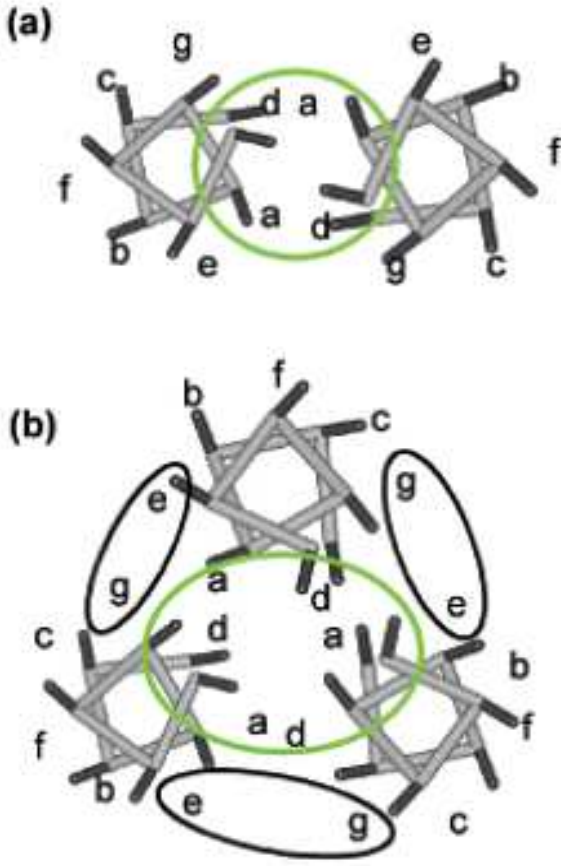


Figure 1-1. Helical wheel diagrams of parallel two (a) and three-stranded (b) coiled coils. Hydrophobic cores of the coiled coils are formed by the hydrophobic **a** and **d** residues. Hydrophilic **e** and **g** residues form electrostatic interactions at the helical interfaces that lead to the formation of parallel three-stranded coiled coils. This figure is taken from ref 116.

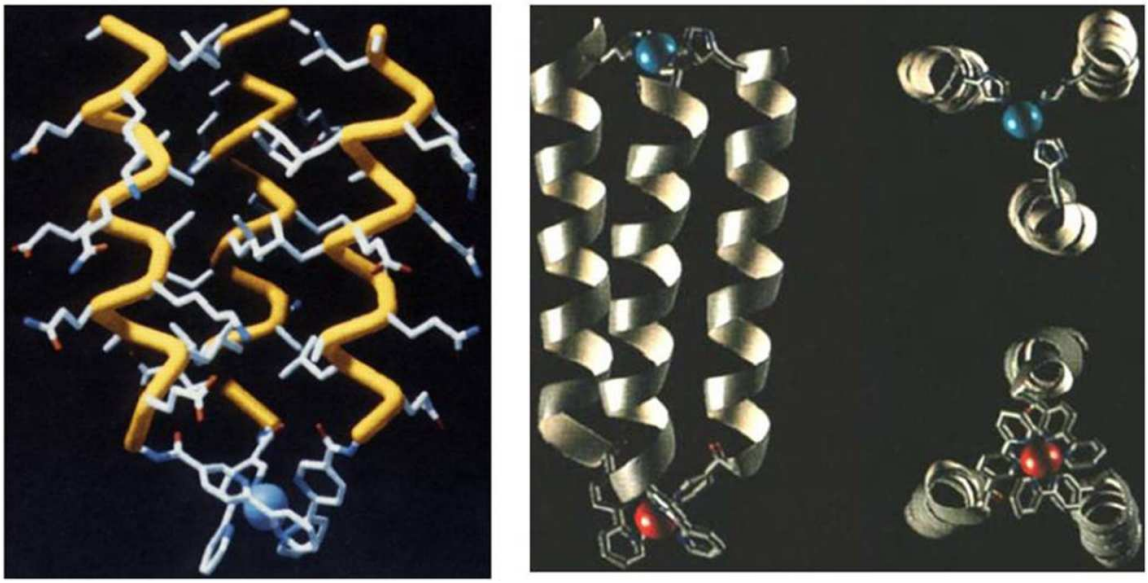


Figure 1-2. Computer generated models of (left): metal ion-assisted triple-helical coiled-coil protein with a 2,2'-bipyridine functionality at the N-terminus. Shown is the side view of the structure with the helical backbone and the bipyridyl metal binding site; (right): parallel three-helix bundle $Cu^{II}Ru^{II}$ bound heterodimetallic protein. Ru(II) is shown as red sphere, bound to 2,2'-bpy. Cu(II) (blue sphere) is shown to be bound to His residues. The figures are taken from ref 23 and 24, respectively.

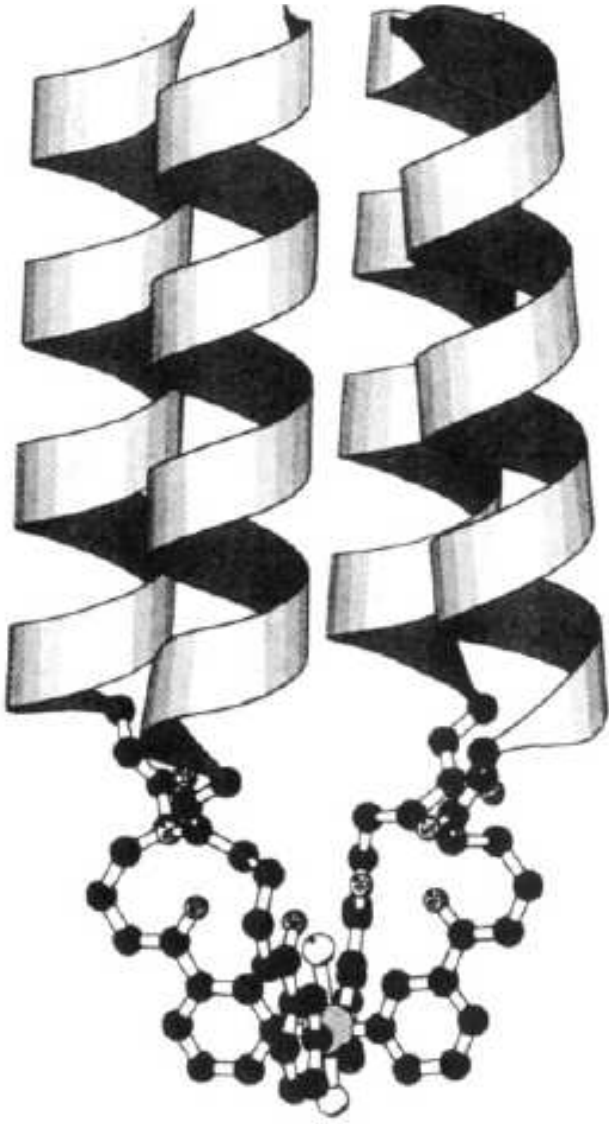


Figure 1-3. Computer-generated model of the parallel four-helix bundle metalloprotein. A 15-residue amphipathic peptide with a pyridyl functionality at the N-terminus was shown to undergo intermolecular self assembly upon Ru(II) complexation. This figure is taken from ref 25.

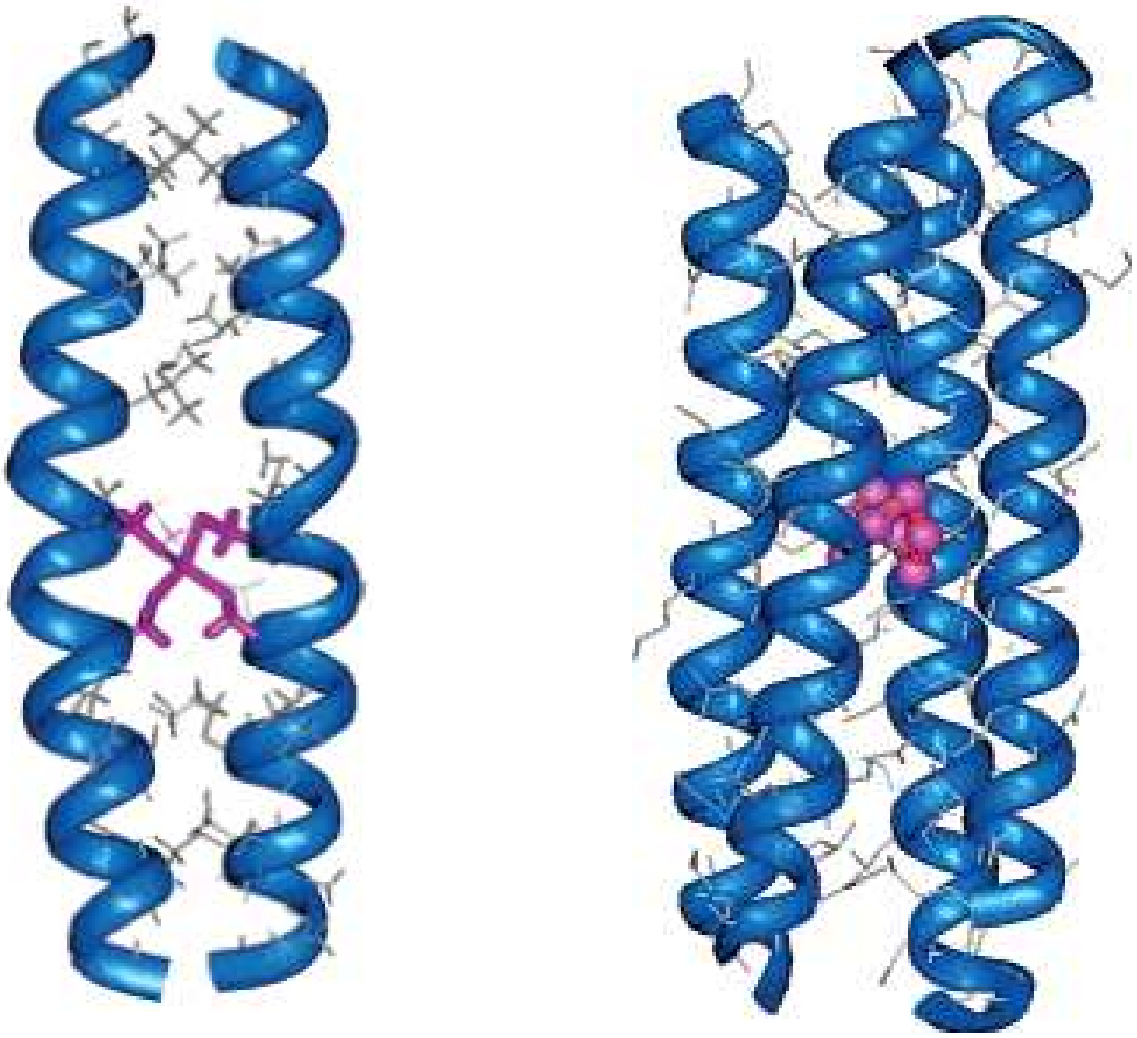


Figure 1-4. Energy-minimized computer models (left): The Cd-bridged C16C19 peptide dimer in which Cys residues located at postions 16 and 19 of each peptide chain bind the Cd(II) center in a tetrahedral geometry; (right): the tetrameric Cu(I)-C16C19-GGY metalloprotein. Figures are taken from ref 27, and 28, respectively.

Figure 1-5. Tertiary model of the Ni(II) complex of the IZ-3adH. A side view (right) and a bottom view from the C-termini (left) of the peptide-metal complex are shown. The six His side chains are shown by the sticks with the three helix backbones. The Ni(II) is indicated by a sphere. This figure is taken from ref 31.

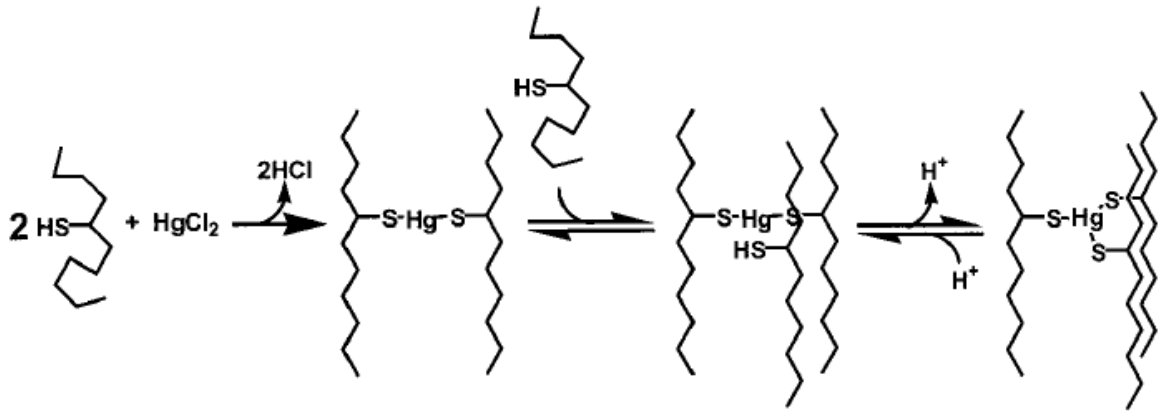


Figure 1-6. Scheme showing the encapsulation of Hg(II) by Baby L9C which is unstructured in the absence of Hg(II). This figure is taken from ref 34.

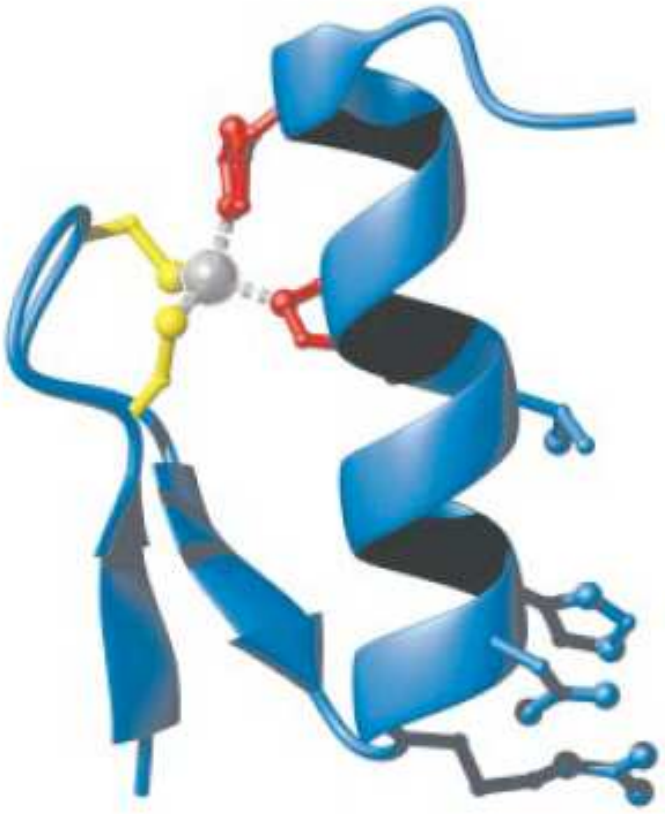


Figure 1-7. A ribbon diagram of the Cys_2His_2 zinc finger motif, including the two cysteine side chains (yellow) and two histidine side chains (red) that coordinate the zinc ion (silver sphere). This figure is taken from ref 37.

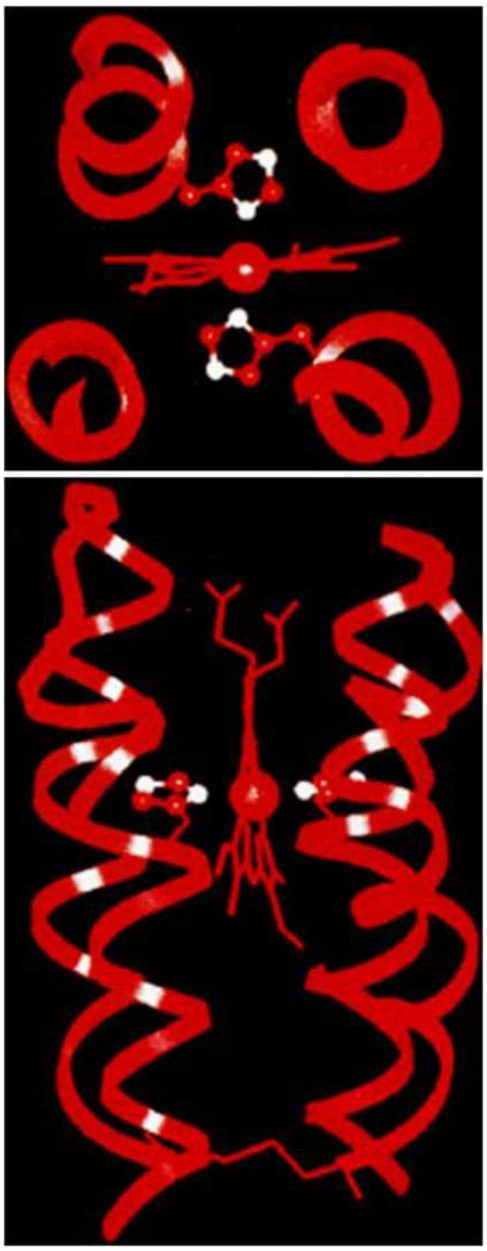


Figure 1-8. Energy minimized model of VAVH25(S–S) with bound heme. Ribbon diagram of the backbone is shown along with the His residues shown as sticks. Nitrogens on the imidazoles are highlighted in white. Top: view looking down the core of each bundle at the heme-binding site. Bottom: lateral view of the protein, with the loops positioned at the top and the disulfide positioned at the bottom of the bundles. This figure is taken from ref 49.

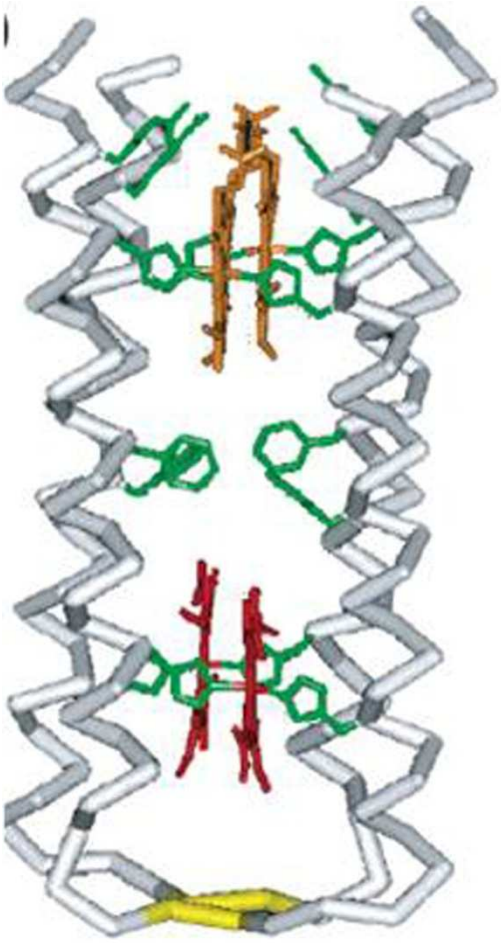


Figure 1-9. Molecular model of the four-α-helix bundle heme maquette. A di-α-helical peptide self assembled to the four-helix bundle incorporating four units of heme. Protein backbone is shown as white, His and Phe as green, hemes as orange and red sticks. The disulfide linkage is shown as yellow. The figure is taken from ref 54.

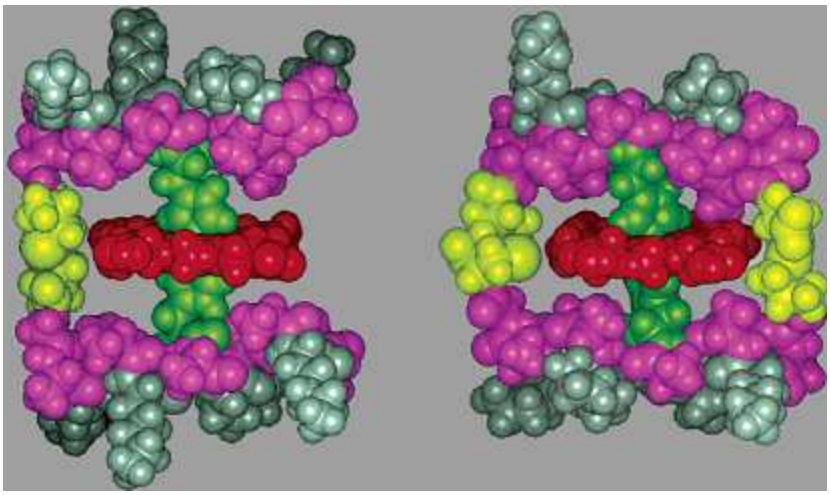


Figure 1-10. Computer generated models of hairpin peptide (15-17 mer) (left) and cyclic peptide bound to heme. Hydrophobic residues are shown as magenta, polar residues as gray, histidine as green, cysteine as yellow, and porphyrin as red. This figure is taken from ref 60.

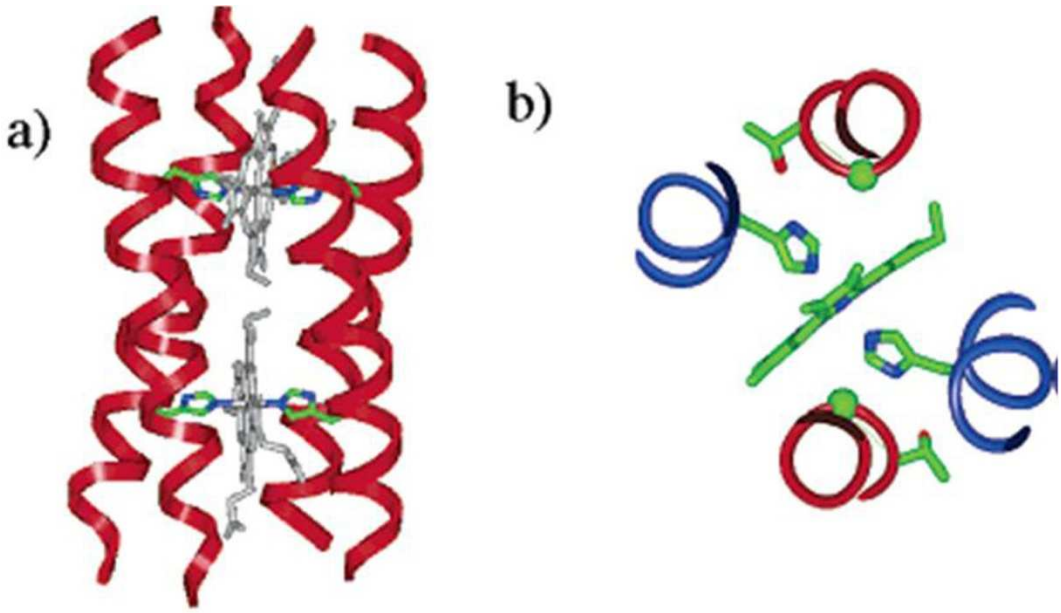


Figure 1-11. Model of (a) heme binding site of *D*2-symmetric four-helix bundle. The heme is axially coordinated by the Nε of two histidine side chains contained in helices 1 and 3 (blue ribbon); the orientation of each His residue is fixed by hydrogen bonding between its Nδ and a Thr residue on the neighboring helices 2 and 4 (red ribbon), as shown in b. The figure is taken from ref 61.

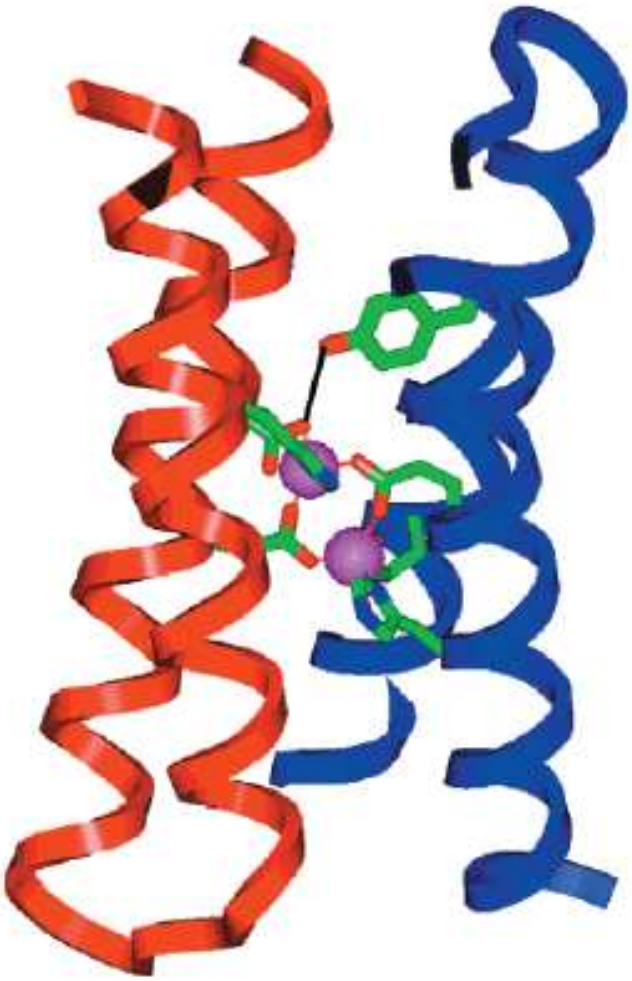
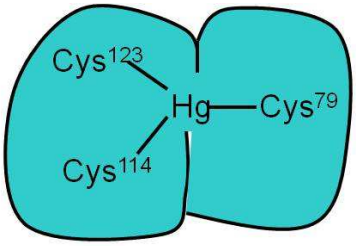
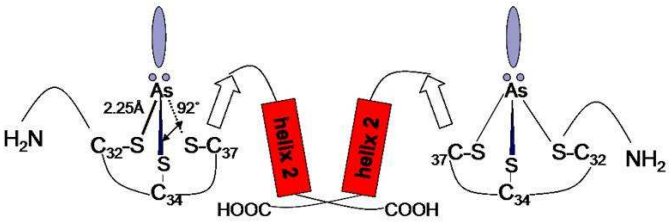


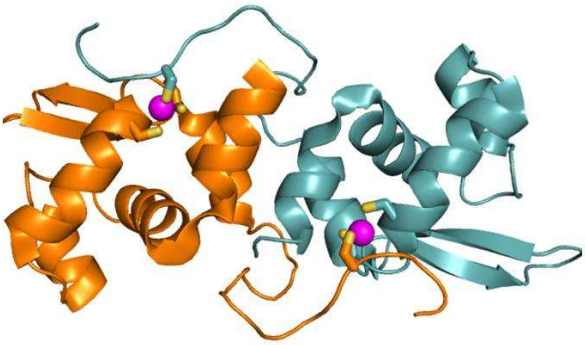
Figure 1-12. X-ray structure of the di-Zn(II) form of DF1 (2.5 Å resolution). The backbone of the structure along with the ligands are shown. A Tyr phenolic group hydrogen bonds to a Glu carboxylate. The figure is taken from ref 83.



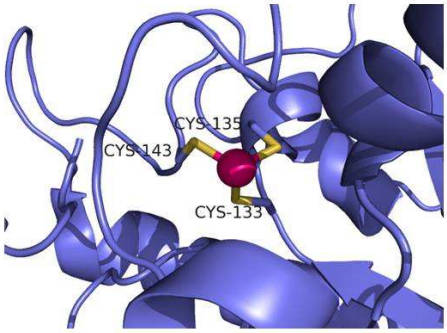
MerR Metalloregulatory Protein



ArsR Metalloregulatory Protein



Cd(II) Regulatory Protein CmtR



Pb(II) inhibited form of δ-ALAD

Figure 1-13. Heavy metal-thiolate interactions of different regulatory proteins such as MerR, ArsR, CmtR. Pb(II) inhibited form of ALAD is also shown. All of these proteins display exclusive Cys ligation around the metal sites.

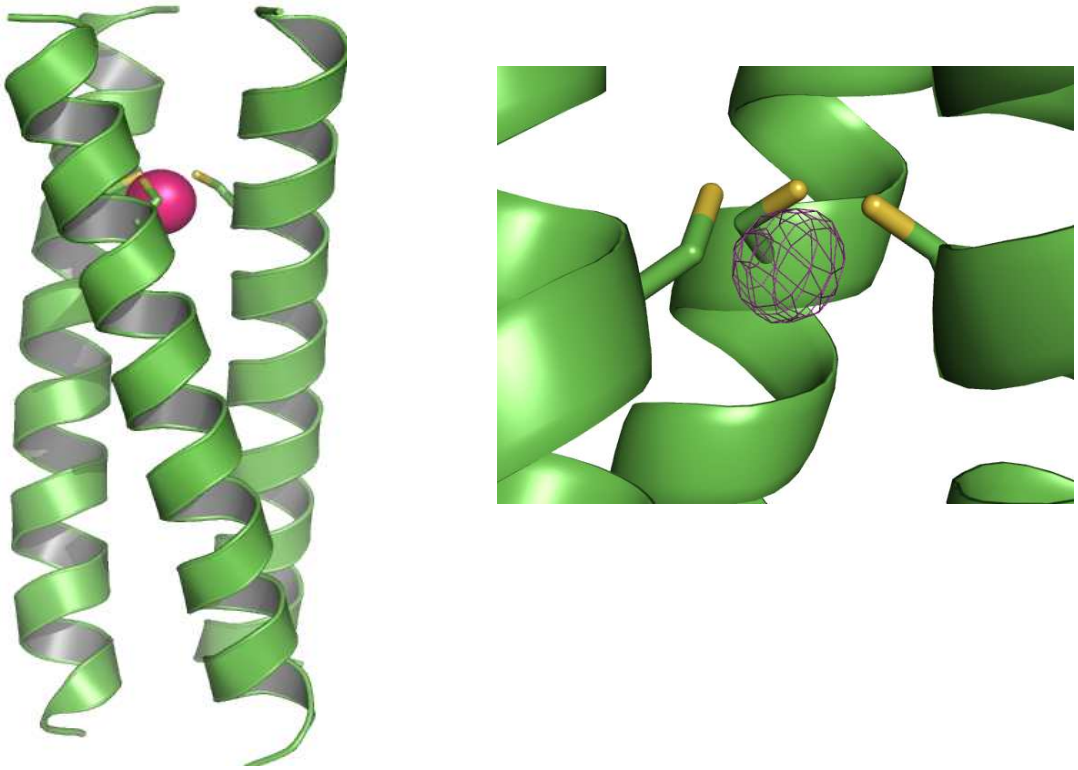


Figure 1-14. X-ray structure of As(III) bound to parallel three-stranded coiled coil **CS**L9C (PDB: 2JGO). Left: peptide backbone is shown as green ribbon, Cys side chains as sticks, As(III) as magenta sphere. Right: Electron density of As(III) shows that the metal bound in *endo* conformation below the thiol plane (Figure taken from ref 119).

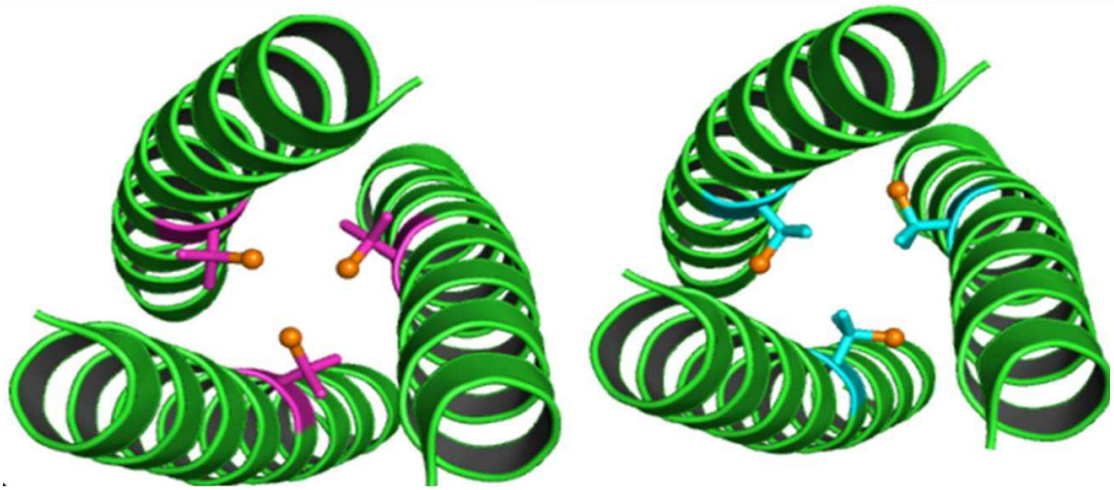


Figure 1-15. X-ray structures of apo **CS**L16-L-Pen (left) (PDB: 3H5F) and **CS**L16-D-Pen (right) (3H5G). Sulfurs (shown as orange balls) in **CS**L16-L-Pen are oriented towards the interior of coiled coils and are preorganized for metal binding. On the other hand the sulfurs in **CS**L16-D-Pen are oriented away from the interior of the coiled coil towards the helical interface.

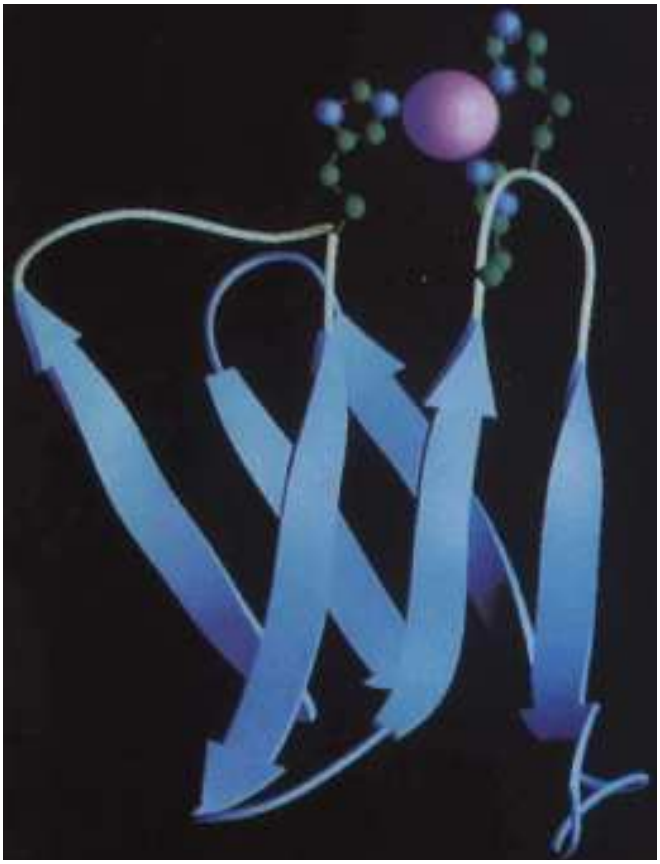


Figure 1-16. A computer model of the designed ‘minibody’ protein comprising of 6 beta strands. Metal binding site involves two His residues from one loop and one His from a nearby loop. His residues are shown as sticks and Zn(II) is shown as sphere. This figure is taken from ref 105.

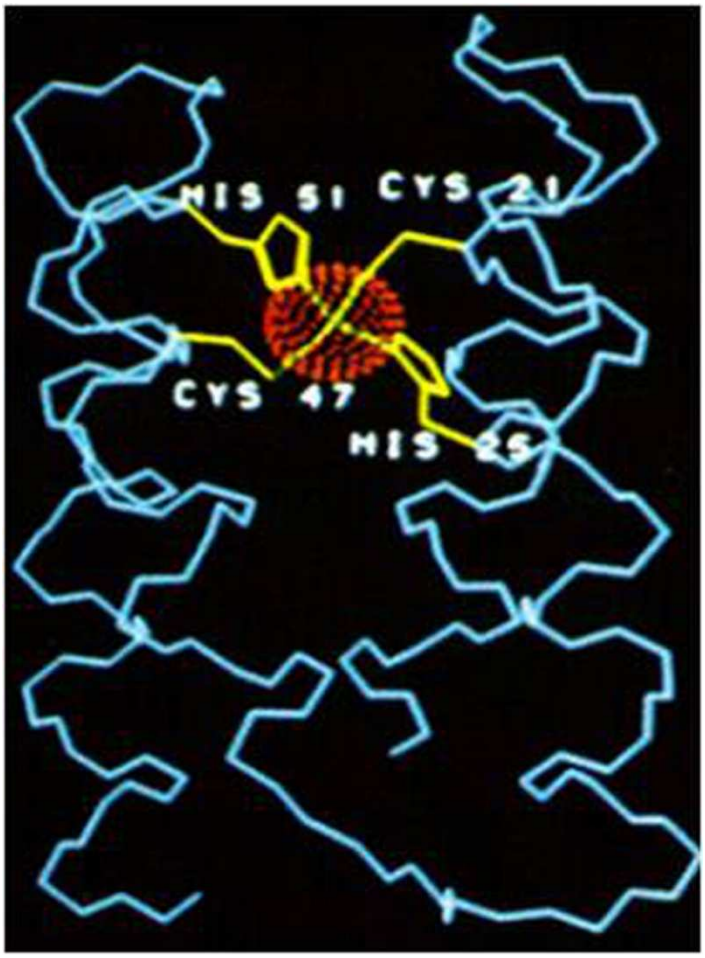


Figure 1-17. Model of the His_2Cys_2 metal binding variant of α_4. The backbone of the four helix bundle is shown in blue, the side chains of the four amino acid ligands are labeled and shown in yellow, Zn(II) is indicated by a sphere of red dots, and bonds from ligands to Zn(II) are shown by lines within the sphere. The image is taken from ref 155.

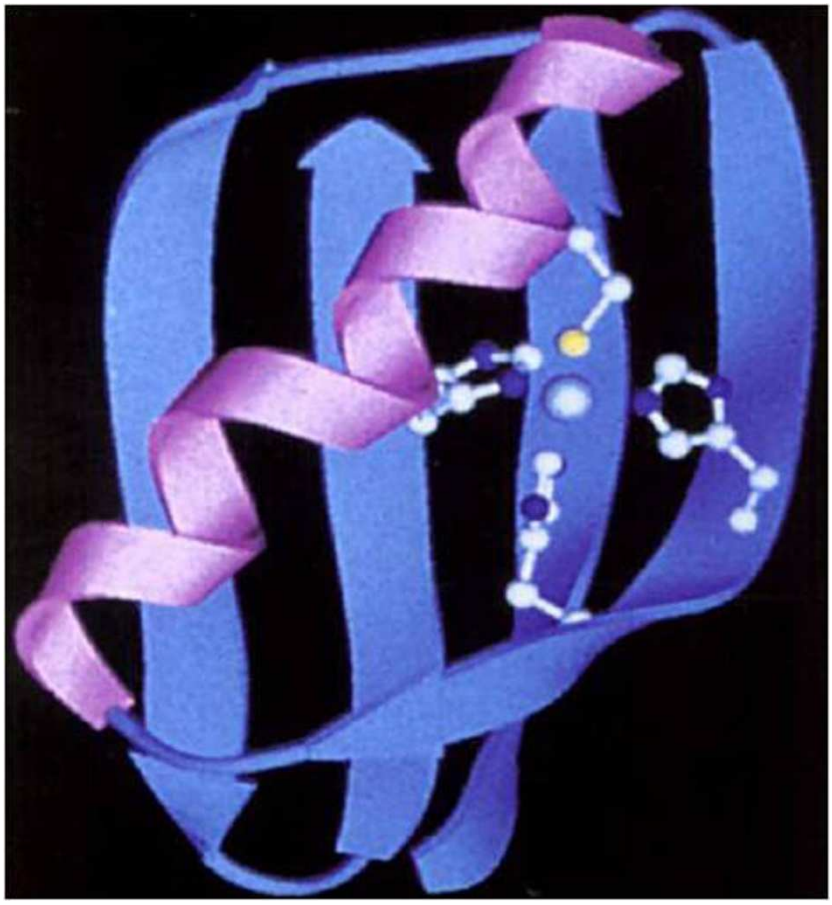


Figure 1-18. Ribbon representation of Zβ1M. The metal binding residues are shown as sticks while the Zn(II) ion is shown as grey sphere. This image is taken from ref 158.

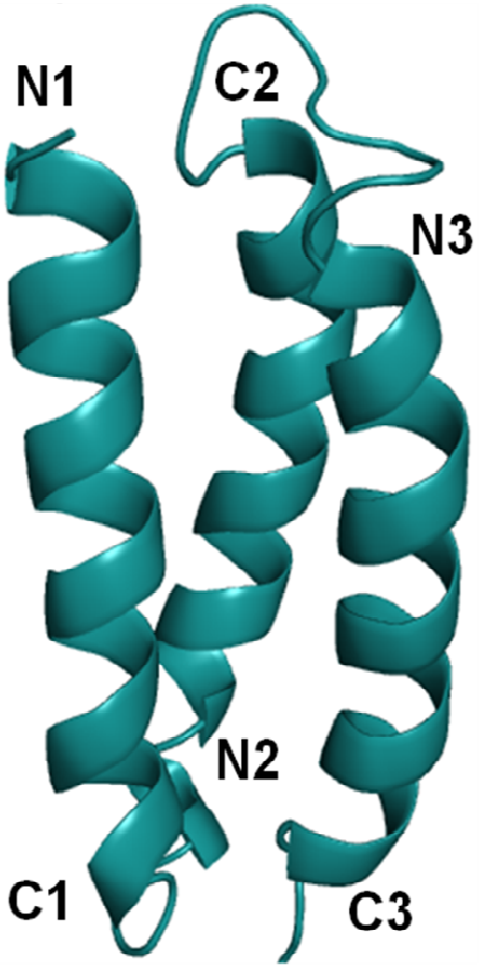


Figure 1-19. Solution structure of the *de novo* designed three-helix bundle protein α_3D. N and C-termini are designated as N1-C3.

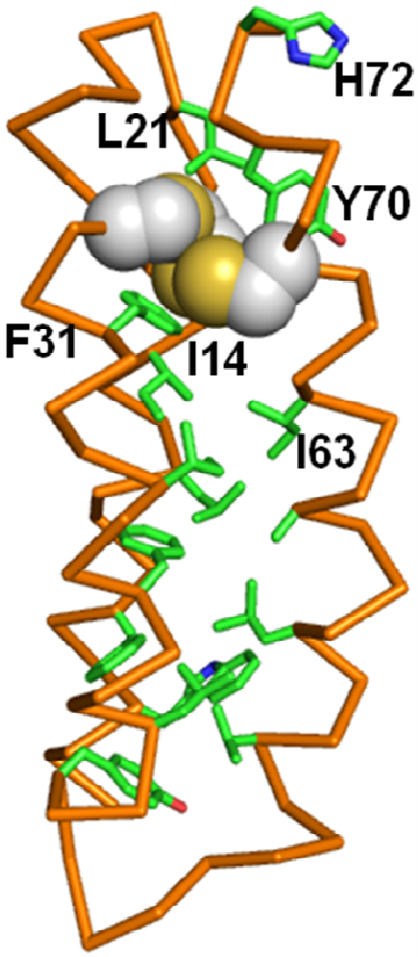


Figure 1-20. PyMol model of α_3D**IV** generated from the NMR structure of α_3D. Cys residues, located at the C-terminal end of the bundle are shown as spheres. Protein backbone is shown in orange. The Cys site can be considered to be located in a hydrophobic “box” formed by the hydrophobic residues F31, I14, I63, L21 and Y70, shown as sticks.

Chapter 2

Structural Comparisons of Apo- and Metallated Three Stranded Coiled Coils Clarify Metal Binding Determinants in Thiolate Containing Designed Peptides

Introduction

Due to the inherent complexity of proteins, it is often difficult to discern the sequence determinants that confer structural stability and function to the biopolymer. Hence, a complete understanding of how a protein's primary amino acid sequence defines its three dimensional structure and function has eluded many researchers for decades and still it is not very well understood. One approach to simplify this problem is the *de novo* design of model proteins and peptides which are less complex than their natural counterparts, yet contain sufficient information in their sequences to clarify protein structure-function relationships.[1-3] Current studies in protein design are focused on unveiling important structural features in natural protein systems with the ultimate goal of using this knowledge to design novel proteins that are tailor-made to carry out receptor, sensory, and catalytic functions.[4-9] *De novo* design of proteins and metalloproteins has been dominated by α-helical coiled coil and bundle domains; however, secondary structural motifs other than α helices have also been used

such as $\beta_2\alpha$,[10-12] small peptides that bind metal clusters at loop region,[13-15] γ turns,[16] β-hairpin peptides[17, 18] and β sheet structures.[19-21]

Research in the Pecoraro group focuses on the broad areas of heavy metal biochemistry and metalloprotein design. The systems that are used to understand these fundamental issues are the **TRI**,[22] **GRAND**[23] and **Coil Ser**[24] series of peptides, designed based on the well known heptad repeat (Table 1 shows the sequences) and form well folded three-stranded parallel coiled coils in solution and solid state. One of the major goals has been to develop models that provide a foundational understanding of the interactions of heavy metals with different metalloregulatory proteins such as MerR, ArsR/Smtb and to develop effective heavy meal sequestering agents. The Pecoraro group has reported numerous successes using the *de novo* protein design strategy to further understand heavy metal biochemistry using peptide systems based on the three-stranded coiled coil motif (**TRI** series), including the first water soluble spectroscopic model of the MerR protein[25], the first trigonal pyramidal As(III) bound to three cysteines as a model of ArsR[26] (**Coil Ser** series) and the first example of a trigonal thiolate Cd(II).[27],[28]

From a metalloprotein design perspective the group has provided a quantitative thermodynamic evaluation of Hg(II) binding demonstrating cooperative metalloprotein folding[23]. Using the **GRAND** series of peptides it has been possible to design for the first time, heterochromic peptides that sequester Cd(II) as 4-coordinate CdS_3O at one site and as trigonal planar CdS_3 structure at a second site within the same peptide construct.[29],[28]

The **TRI** and **GRAND** series of peptides are related to Coil V_aL_d, in which valine occupies all **a** positions and leucine all **d** positions, which was shown,[30] using X-ray crystallography to be a parallel three stranded coiled coil. Substitution of one or more Leu residues from **a** and/or **d** position of **TRI** or **GRAND** peptides with Cys or Penicillamine (Pen) generate thiol-rich sites at the

interior of the coiled coil which bind heavy metals such as Hg(II), Cd(II), As(III), Pb(II) and Bi(III).[22, 23, 25-27, 29, 31-36] While effective protein based heavy metal sequestering agents, these series of peptides are not easily crystallizable. For structural studies the structurally similar substituted **Coil Ser** (**CS**) series of peptides, (Table 1) which are also based on a heptad repeat, and which can crystallize as parallel or anti-parallel three-stranded coiled coils, have been used.

Coil Ser was first crystallized by Lovejoy and DeGrado *et al*, as an anti parallel coiled coil at low pH.[24] It was suggested that the steric clash between the N-terminal Trp residues led to the antiparallel orientation. Subsequent NMR work demonstrated that **Coil Ser** formed parallel three stranded coiled coils in slightly basic solution.[37] Subsequently both metallated and apo derivatives of **Coil Ser** peptides which have been shown to be parallel three stranded coiled coils have been characterized. This orientation may reflect the pH, the sequence modification, the presence of metals bound in the hydrophobic interior, the presence of Zn(II) ions that bind to external His and Glu residues at the crystal packing interfaces or a combination of these factors.[26, 38]

Despite the fact that there are a significant number of studies using *de novo* designed peptide systems by many different groups, there are only few examples of structurally characterized *de novo* designed constructs available in literature. The most well characterized systems are four helix bundle motifs consisting of dimers of helix-turn-helix peptides known as Dueferri (DF), designed by DeGrado. These peptides serve as paradigms for di-iron, di-zinc and di-manganese containing natural proteins and contain carboxylate (Glu) bridged dinuclear metal sites at the center of the four helix bundle.[39] Various Zn(II), Cd(II) and Mn(II) derivatives of DF have been characterized crystallographically.[40-42] To determine the extent to which the active site is preorganized in the absence of metals, NMR structure of apo DF1 was solved which suggested that the native fold and metal binding site are largely preorganized in the apo protein.[43] Active sites of metalloproteins are frequently

preorganized in the absence of metals, which requires burial of polar amino acids at the interior at the expense of folding free energy.[44] Preorganization requires that proteins impart their own structural preference on metals to bind in a specific geometry.

Our group has been directing the efforts towards the structural characterization of various three-stranded coiled coils. Recently, Touw *et. al.* reported the X-ray structure[26] of As(**CS**L9C)$_3$ where As(III) is coordinated to three Cys residues in an **endo** conformation within a protein environment. This construct served as the first structural model for As(III) binding to ArsR or ArsD.[45, 46] The structures of both **L**- and **D**- versions of unmetallated (apo) Pen derivatives (**CS**L16**L**Pen and **CS**L16**D**Pen) which are both substituted in the **a** heptad position, were successfully obtained.[38] The sulfur atoms of **L**- thiol containing residues appear to be orientated in a similar way when the mutation occurs at an **a** site and appear to be independent of both the place of substitution in the coiled coil, 9 vs. 16, the nature of the coordinating ligand, L-Cys vs. L-Pen, as well as the state of metallation. A comparison of the structures of apo (**CS**L16**L**Pen)$_3$ and metallated As(**CS**L9C)$_3$ show that the sulfurs appear to be pre-organized for metal binding. The major differences in the Pen vs. Cys structures are the packing of the two methyl groups on the Pen. The plane below the Cys is blocked by one set of methyl groups which point towards the interior of the coiled coil below the sulfurs, the second set of methyl groups point away from the center in-between the helical interfaces. The **D**-Pen points down towards the C-terminus and, most interestingly, appears to point away from the interior of the coiled coil towards the helical interfaces. The dramatically different Cd(II) binding to **CS**L16**L**Pen and **CS**L16**D**Pen has been correlated with the structural differences. Furthermore my group has demonstrated that the affinities and selectivities of metal binding to **a** vs. **d** site substituted peptides are different in solution.[36, 47] Cd(II) binds selectively to **a** site and has 10 fold higher binding affinity over **d** site.[34] Pb(II) on the other hand has 4 times higher affinity for

binding to **d** site over **a** site.[48] Evidently, structural information about **a** and **d** sites within these coiled coils is sorely-required.

In this chapter I present the structures of two apo peptides, (**CS**L9C)$_3$ which is an **a** site peptide, and (**CS**L19C)$_3$ which is a **d** substituted peptide. With the apo structure of the **a** site peptide (**CS**L9C)$_3$ I address local/global structural differences with the metallated analogue As(**CS**L9C)$_3$.[26] Additionally I address the effect of arsenic binding on the overall structure of the peptide and answer questions as to whether As(III) binding changes the conformation of the coordinating Cys residues significantly. Furthermore, the extent of the degree of preorganization of the coordinating residues at the metal binding site and whether large amounts of water are extruded on metal coordination are also investigated. The apo (**CS**L19C)$_3$ structure allows me to address whether there are major conformational differences between the orientation of Cys sulfurs in **a** vs. **d** sites, which thiol site generates more void space in the hydrophobic interior, how different are the packing of hydrophobes above and below the Cys site and what are the consequences of the observed differences. Based on the structural analysis of these two apo peptides I also hypothesize possible modes in which Cd(II), Hg(II) and Pb(II) would coordinate to these peptides. The knowledge gained from this study would allow us to have access to the rules for the *de novo* design of metal binding sites within the interior of coiled coils, as well as a deeper understanding of heavy metal binding in designed proteins, potentially providing significant insight into metal binding in metalloregulatory proteins and metallochaperones.

Materials and Methods.

Peptide Synthesis and Purification. Peptides **CS**L9C and **CS**L19C were synthesized on an Applied Biosystems 433A automated peptide synthesizer using standard Fmoc protocols,[49] and purified by reverse-phase HPLC on a C18 column at a flow rate of 10 mL/min using a linear gradient that was varied from

0.1% TFA in water to 0.1% TFA in 9:1 CH_3CN:H_20 over 50 minutes. Purified peptides were characterized by electro spray mass spectroscopy.

Crystallization. Crystals of **CS**L9C were grown by vapor diffusion on a hanging drop at 20°C containing equal volumes (2 µL) of peptide (19.15 mg/mL **CS**L9C, 17 mM $Zn(OAc)_2$, 100 mM sodium cacodylate buffer at pH 6.0) and precipitant (100 mM $Ca(OAc)_2$ and 24% PEG 3350). Figure 2-1 shows a picture of a single crystal of **CS**L9C. Crystals were soaked in the mother liquor and cryo protectant (20% ethylene glycol) prior to freezing in liquid N_2. **CS**L19C was crystallized by vapor diffusion in a sitting drop with equal volumes of peptide (10 mg/mL in 20 mM $Zn(OAc)_2$) and a precipitant solution (100 mM MES pH 6.5 and 40 % PEG 200). The hexagonal plate-like crystals obtained were mounted in a 0.15 mm robotic loop and frozen in their mother liquor in liquid N_2 for data collection.

Data Collection and Refinement. Data were collected at the Advanced Photon Source of the Argonne National Laboratory on the LS-CAT Beamline 21-ID, equipped with a Mar 225 CCD detector. 360° frames of data were collected with a rotation of 1° and exposure time of 1s. The data were processed and scaled with HKL 2000.[50] The space group of **CS**L9C was determined to be C2. The structure of **CS**L9C was solved by molecular replacement using Phaser in CCP4 suite of programs[51, 52] by using 29 amino acids of **CS**L16**L**-Pen (PDB code 3H5F)[38] as the search model. Leu 9 of **CS**L16**L**-Pen were mutated to Cys and the side chains of Pen 16 were modified in such a way that they resembled Leu residues containing only Cβ and Cγ atoms before using it as a search model for Molecular Replacement. The generated model was refined with restrained refinement by Refmac 5 in CCP4 suite of programs[53] and built in and completed in Coot[54] by using the $2F_o$-F_c and F_o-F_c electron density maps generated in Refmac 5. After one round of refinement the methyl groups of Leu were built in using the density present in F_o-F_c electron density map. The final structure was refined to 1.36 Å resolution ($R_{working}$ = 18.2%, R_{free} = 21.5 %). The coordinates

and structure factors of **CS**L9C have been deposited in the Protein Data Bank with the ID code 3LJM.

X-ray data collection, solution of the data, and initial refinement of the **CS**L19C structure was done by Dr. Debra Touw, a former graduate student of the group. Final refinements and structural analysis / comparisons with **CS**L9C structure were done by me. The space group of **CS**L19C was determined to be $P6_1$ or $P6_5$. Molecular replacement with both potential space groups was performed in the program Phaser in the CCP4 suite of programs[51, 52] with the first 26 amino acids of a single helix of Coil V_aL_d (PDB code 1COI)[30] as a search model. The top 65% of rotation function solutions were included in the translation search. In two separate trials three and six helices were searched for with zero allowed backbone clashes. The molecular replacement was successful for three helices per asymmetric unit in the space group $P6_1$ with a Z-score of 14.9 and a Log Likelihood Gain of 195. The Matthews coefficient was determined to be 3.65 which correspond to a solvent content of 65 %. The molecular replacement solution was used as input for the Auto Build feature of Phenix,[55] followed by refinement in Refmac.[53] F_o-F_c and $2F_o$-F_c electron density maps generated with the CCP4 map utility FFT and with the mtz phase file in Coot[56] were used for iterative building in of side chains, Zn(II) ions, and water molecules. The structure was refined to 2.15 Å resolution ($R_{working}$ = 21.9 %, R_{free} = 26.4 %). The validity of the model was verified with a composite omit map generated in CNS.[57] The coordinates and structure factors of **CS**L19C have been deposited in the Protein Data Bank with the ID code 2X6P. The data processing and refinement statistics for **CS**L9C and **CS**L19C are given in Tables 1-2 and 1-3, respectively. Figures were generated in Pymol.[58] Angles between thiol plane of **CS**L19C and planes formed by the C_γ atoms of Leu 16 and Leu 23 were measured using the program Mercury from Cambridge Crystallographic Data Center.[59]

UV-Vis and ^{199}Hg NMR Spectroscopy. To remove ubiquitous Zn(II) from solution, peptide stock solutions and buffers were prepared in deionized water

(Milli Q) treated with Chelex-100 resin (BioRad). All plasticware and glassware were soaked in 5 mM EDTA overnight followed by several thorough rinses with deionized water. Only metal free pipette tips (Fisher) were used. All solutions were purged with argon before use to minimize chances of Cys oxidation. Peptide concentration was determined by quantization of Cys thiols using Ellman's test.[60] UV-Vis experiments were performed at room temperature on a Cary 100 Bio spectrophotometer. Quartz cuvettes of path length 1 mm and 1 cm were used for the experiments at pH 6.5 and 8.6, respectively. Zn(II) titrations were performed by adding aliquots of 1 M stock solution of $Zn(OAc)_2$ to solutions containing 2.5 mM **CS**L9C and 50 mM MES buffer at pH 6.5 and 1 equivalents of $HgCl_2$ (per trimer). Appropriate background spectrum was subtracted prior to addition of $HgCl_2$. Hg(II) titrations were performed by adding aliquots of 7.37 mM stock solution of $HgCl_2$ to solutions containing 2.5 mM **CS**L9C, 50 mM MES buffer at pH 6.5 in the presence of 21 equivalents of $Zn(OAc)_2$ (per monomer) and with no Zn(II) present. For experiments at pH 8.6 same stock solution of $HgCl_2$ was added to solutions containing 10 μM **CS**L9C, 50 mM CHES buffer in the presence and absence of 30 μM $Zn(OAc)_2$. Direct Hg(II) titration data were analyzed by non-linear least squares fits to the equation similar to what has been described previously[35, 61]:

$$A = \frac{1}{2K_b}\left\{K_b\mathrm{M}_0 + K_b\mathrm{P}_t + 1 - (K_b{}^2\mathrm{M}_0{}^2 - 2K_b{}^2\mathrm{M}_0\mathrm{P}_t + 2K_b\mathrm{M}_0 + K_b{}^2\mathrm{P}_t{}^2 + 2K_b\mathrm{P}_t + 1)^{1/2}\mathrm{b}\Delta\varepsilon\right\}$$

where A is the observed absorbance on addition of Hg(II) to peptide solution, K_b is the binding constant of Hg(II) to be determined, P_t is the total concentration of peptide trimer, M_0 is the concentration of Hg(II) used during titration, $\Delta\varepsilon$ is the difference molar extinction coefficient of the complex formed during titration and b is the path length of the cuvette. As the absorbance at 247 nm decreases in the presence of more than stoichiometric amount of Hg(II) due to formation of HgS_2 in two stranded coiled coil, it was necessary to assume that the absorbance stays constant in order to fit the binding data. ^{199}Hg NMR experiments were performed at room temperature on a Varian Inova 500 spectrometer (89.5 MHz for ^{199}Hg) equipped with a 5 mm broadband probe as previously described.[62]

The samples for ^{199}Hg NMR experiments contained a) 12 mM **CS**L9C, 4 mM $^{199}Hg(NO_3)_2$ (prepared from 91% isotopically enriched ^{199}HgO, Oak Ridge), 252 mM $Zn(OAC)_2$ and b) 14 mM **CS**L9C and 4.6 mM $^{199}Hg(NO_3)_2$. pH of both the samples were adjusted to 5 using concentrated solutions of KOH and HCl. pH of the solutions were also measured after the experiment. The data were analyzed using the software MestRe-C.[63] The free induction decay was processed by applying a linear prediction from 1-32 points out of 120 basis points with a coefficient of 32 followed by zero filling and exponential line broadening of 200 Hz prior to Fourier transformation.

Results

The nomenclature used to define the “apo” and “metallated” peptides is given in ref [[64]].

Overall Structure. The X-ray structures of well folded three-stranded parallel coiled coil peptides (**CS**L9C)$_3$ and (**CS**L19C)$_3$ are determined to 1.36 Å and 2.15 Å resolution, respectively. The final model of (**CS**L9C)$_3$ contains 816 protein atoms, 59 water molecules, 5 Zn(II) ions, 1 Ca(II) ion and 3 ethylene glycol molecules. (**CS**L19C)$_3$ contains 699 protein atoms, 19 water molecules and six Zn(II) ions. The RMSD between the 29 amino acids of **CS**L16L-Pen (PDB code 3H5F)[38] as the search model and the final refined structure of (**CS**L9C)$_3$ is 0.024 Å. Leu 9 of CSL16L-Pen were mutated to Cys and the side chains of Pen 16 were modified in such a way that they resembled Leu residues containing only C_β and C_γ atoms before using it as a search model for Molecular Replacement. After one round of refinement the methyl groups of Leu were built in using the density present in F_o-F_c electron density map. Using the first 26 amino acids of the single helix of Coil V_aL_d (PDB code 1COI)[30] as the search model the RMSD for the final refined structure of (**CS**L19C)$_3$ is 0.025 Å. All non-glycine residues for both the structures fall in the right handed α-helical region of the Ramachandran plot.[65] The carbonyls of N-terminal capping acetyl groups in both structures are

hydrogen bonded to Ala 4 of respective main chains. In (**CS**L9C)$_3$ the amide NH_2 groups of chains A and C are hydrogen bonded to the main chain CO of Ala 25 from respective chains. All side chains in both the structures except those involved in metal binding or hydrogen bonding to solvent molecules are present in their preferred rotameric conformation as analyzed by rotamer check utility in Coot.[56]

Cysteine Sites. Side chains of all three cysteine residues of **CS**L9C display alternate conformations. The cysteine side chains in all three helices are oriented in essentially equal proportion between the coiled coil interior and the helical interface (in chain C the cysteine side chain is oriented towards the interior and helical interface at what is best modeled as a 60:40 ratio). Thus, there is an ensemble of different combinations of cysteine side chain orientations between three peptide helices. Out of these different orientations of three Cys side chains, 15% (0.5x0.5x0.6x100%) of the time all the cysteine side chains are exclusively oriented towards the interior of the coiled coil and 10% of the time all of them are oriented towards the helical interface. The orientations of the two conformers of Cys residues are shown in Figure 2-2. The separation between the S_γ atoms (S-S) of the three cysteines for the interior conformation is 3.32 Å, whereas the S_γ atoms for the exterior conformation are separated at a distance of 6.85 Å with respect to each other. It can be seen from Figure 2-2 that the thiol groups of cysteine side chains in the interior conformation are directing towards the N-termini of the helices, whereas those in the exterior conformation are pointing towards the C-termini. The distance between the planes formed by the S_γ atoms of the two conformers of cysteines is 1.11 Å. The β-methylene carbons of all Cys residues are pointing towards the N-terminus of the helical scaffold. The distance between the C_α and C_β planes of cysteines is 1.07 Å. The S_γ atoms in the interior orientation of cysteines are in a plane which is at a distance of 0.59 Å above from the C_β plane of cysteines. On the other hand, the S_γ atoms in the exterior orientation of cysteines are in a plane which is 0.52 Å below from the C_β plane of cysteines. The χ_1 dihedral angles (N-C_α-C_β-S_γ) for the interior orientation

of cysteines are −70.3 ± 3.9 which are close to the value of −65.2° for most common cysteine rotamers.[66] The χ_1 values for the outer conformer are 171.4° ± 6.2. The leucine layers above (Leu 5) and below (Leu 12) Cys 9, both of which are **d** sites of the heptad, are oriented differently (Figure 2-3). Leu 5 is oriented more towards the center of the coiled coil with the Leu 12 layer being oriented away from the center. One methyl group of both Leu 5 and Leu 12 is pointing towards the N termini and the other one is directing towards the C termini. The distances between the planes formed by the C_γ atom of Leu 5 and Leu 12 and the S_γ of interior cysteines conformer are 3.85 Å and 5.29 Å, respectively. The distances for the exterior cysteine conformer are 4.99 Å and 4.24 Å from Leu 5 and Leu 12 planes formed by the C_γ atoms, respectively.

The orientation of the cysteine side chains in (**CS**L19C)$_3$ is depicted in Figure 2-4. β-methylene groups of three cysteines are pointing towards the N-termini. The χ angles for the three cysteines are −90.53°, −165.92° and 164.33° for chains A, B, and C, respectively. All three cysteines are present as a single rotamer. Thiol groups of both Cys B and C are oriented toward the interior of the helices and towards the C-termini, while Cys A is oriented toward the helical interface and almost perpendicular to the helical axis. The latter orientation towards the helical interface is in accordance with previous predictions[22] which were based on the Coil V_aL_d structure.[30] Electron density in the center of the coiled coil corresponding to a heavy atom was not found in F_o-F_c maps down to a contour level of 2 σ. The cysteine sulfhydryl groups were found to be greater than 3 Å from one another, excluding the possibility of disulfide bridge formation, which would not be likely at pH 6.5 where the cysteines are more likely to be protonated than deprotonated. The S-S separations for Cys A - Cys B, Cys B - Cys C and Cys C - Cys A are 4.64 Å, 3.40 Å and 4.59 Å, respectively.

Non-covalent Interactions. The structure of (**CS**L9C)$_3$ has seven interhelical salt bridge interactions. Glu 3 B forms an intrahelical salt bridge with Lys 7. In addition to interhelical salt bridge interactions between side chains of **e**

and **g** residues Glu and Lys commonly found in coiled coils, there is also an unusual interhelical salt bridge present, between **b** and **g** residues involving Glu 24 A and Lys 22 C. There is an interhelical water mediated hydrogen bonding interaction between Lys 22 A / Glu 20 B. The indole NH of Trp 2 B is involved in hydrogen bonding with Glu 6 B side chain. There are many hydrogen bonding interactions involving side chain and main chains such as, Glu 3 A and Glu 3 C to main chain NH of same residues, Ser 14 C to the main chain CO of Ala 10 and Ser 14. Finally, there are some intrahelical hydrogen bonds involving side chains such as Glu 20 C to Gln 17 C. There are five inter-helical electrostatic interactions in (**CS**L19C)$_3$ as reported previously.[48] Glu 24 B and Lys 22 B are an example of the **e**-**g** salt bridge typically found in three-stranded coiled coils. An interaction between Glu 1 C and Glu 6 B is mediated by a water molecule. Glu 1 C is hydrogen bonded to Trp 2 B and Glu 1 B is hydrogen bonded to Trp 2 A. Helical wheel diagrams of (**CS**L9C)$_3$ and (**CS**L19C)$_3$ summarize these interactions in Figure 2-5 and 2-6.

External Metal Binding Sites. Four out of the five Zn(II) ions in (**CS**L9C)$_3$ are bound at the crystal packing interfaces (Figure 2-7). One of them (ZN1) has its ligands as Glu 24 B (Oε2), His 28 B (Nε2), a water molecule and Glu 24 A (Oε2) from a symmetry related peptide trimer. T-he other Zn(II) (ZN2) is bound to His 28 A (Nε2) and Glu 3 B (Oε1,A / Oε2,B), Glu 20 B (Oε1) and Glu 24 B (Oε1) from a symmetry related peptide molecule. ZN3 is bound to Glu 1 B (Oε2), Glu 6 C (Oε2) and two water molecules. ZN4 has two water molecules, Glu 6 A (Oε1) and His 28 C (Nε2) from a symmetry related molecule as its ligands. ZN5 is bound to one water molecule, His 28 B (Nδ1) and Glu 20 A (Oε2) from a symmetry related trimer. Figure 2-8 shows the ligands to the four Zn(II) ions present at the crystal packing interfaces of each trimer molecule of **CS**L9C. The Ca(II) ion is bound to Glu 1 A (Oε2) and Glu 6 B (Oε1, Oε2).

Zn(II) ion sites at the crystal packing interfaces of (**CS**L19C)$_3$ are very different than that of (**CS**L9C)$_3$. The six fold screw axis and crystal packing of the

structure is shown in Figure 2-9. There are six Zn(II) ions per asymmetric unit, which form two dinuclear Zn(II) sites and two other Zn(II) ions not directly connected by side chain ligands that are 7.6 Å apart. Detailed views of the Zn(II) binding sites are shown in Figure 2-10. The first site, shown in Panel B which is not dinuclear, has one Zn(II) ion coordinated by Glu 27 C, Glu 24 C, and Lys 22 B from one molecule and Glu 3A from another trimer. The neighboring Zn(II) ion is coordinated by His 28 C from one trimer and Glu 1 C and Glu 6 A from another trimer plus one water molecule. The dinuclear Zn(II) site in Panel C contains a bridging glutamate, Glu 6 B, and the coordination sphere of one Zn(II) ion is filled by His 28 A and Glu 24 A from one helix as well as Glu 1 A from the same trimer as Glu 6 B. The other Zn(II) ion of this cluster is coordinated by Glu 3 B, and weakly coordinated by Glu 24 A and the other carbonyl oxygen of Glu 6 B. All three tryptophan residues are closely associated with the Zn(II) sites on the exterior, and are hydrogen bonded to the Zn(II) ligands. The indole nitrogen of Trp 2 A, Trp 2 B Trp 2 C are all 3.8-4.2 Å from a Zn(II) ion.

Discussion

Over the past two decades, designed metallopeptides have held the promise for understanding a variety of fundamental questions in metallobiochemistry; however, these dreams have not yet been realized because of a lack of structural data to elaborate the protein scaffolds before metal complexation and the resultant metallated structures which ultimately exist. This is because there are few reports of structural characterization of such systems either in their metallated or non-metallated forms and no examples where an apo structure and the corresponding metallated peptide assembly have both been defined by X-ray crystallography. In this chapter I present X-ray structures of two *de novo* designed parallel three-stranded coiled coils (designed using the heptad repeat (**a→g**)) **CS**L9C (**CS = Coil Ser**) and **CS**L19C in their non-metallated forms, determined to 1.36 Å and 2.15 Å resolutions, respectively. Leucines from either position 9 (**a** site) or 19 (**d** site) are replaced by cysteine to generate the

constructs **CS**L9C and **CS**L19C, respectively, yielding thiol-rich pockets at the hydrophobic interior of these peptides, suitable to bind heavy metals such as As(III), Hg(II), Cd(II) and Pb(II). These structures are used to understand the inherent structural differences between **a** and **d** sites to clarify the basis of the observed differential spectroscopic behavior of metal binding in these types of peptides. Cys side chains of (**CS**L9C)$_3$ show alternate conformations and are partially preorganized for metal binding, whereas cysteines in (**CS**L19C)$_3$ are present as a single conformer. Zn(II) ions, which do not coordinate or influence Cys residues at the designed metal sites but are essential for forming X-ray quality crystals, are bound to His and Glu residues at the crystal packing interfaces of both structures. These "apo" structures are used to clarify the changes in metal site organization between metallated As(**CS**L9C)$_3$ and to speculate on the differential basis of Hg(II) binding in **a** versus **d** peptides. Thus, attempts are made to establish general rules for heavy metal binding to Cys-rich sites in designed proteins which may provide insight for understanding how heavy metals bind to metallochaperones or metalloregulatory proteins.

Comparison Between (**CS**L9C)$_3$ **and** As(**CS**L9C)$_3$**.** The α helical backbones of apo and As(III) bound **CS**L9C are extremely similar and, as shown in Figure 2-11, overlay very well. This observation demonstrates that As(III) binding does not perturb the overall secondary structure of the coiled coil. Both the apo and As(**CS**L9C)$_3$ crystallize with 4 Zn(II) ions at the crystal packing interfaces coordinated to His and Glu ligands. In the As(**CS**L9C)$_3$ structure all four Zn(II) ions in the asymmetric unit are present at the crystal packing interfaces; whereas in apo (**CS**L9C)$_3$, four (ZN1, ZN2, ZN4 and ZN5) out of five Zn(II) ions in the asymmetric unit are bound at the crystal packing interfaces. The same number of interhelical electrostatic interactions are present in (**CS**L9C)$_3$ as there are in As(**CS**L9C)$_3$ except the interaction between Glu 6 B and Glu 1 C which is present only in the latter structure.[26] The β methylene carbons of both the apo and metallated (**CS**L9C)$_3$ structures are pointing towards the N termini of the helices. The thiol groups of As(**CS**L9C)$_3$ and the major interior conformer of

(**CS**L9C)$_3$ are pointing towards the N termini, with the S-S separation for the major conformer of apo (**CS**L9C)$_3$ being 3.32 Å, which is 0.07 Å longer than the metallated peptide (3.25 Å). The χ_1 dihedral angles for the metallated construct are $-60° \pm 1$ as opposed to $-70.3° \pm 3.9$ for the major conformer of the apo structure. Smaller S-S separation and χ_1 dihedral angles in As(**CS**L9C)$_3$ compared to the apo structure where all cysteines are oriented towards the interior, suggests that the metal binding induces only minor changes in the geometry of the cysteine residues, making them slightly more constrained upon metal binding. The major change, however, is observed in the conformations of cysteine side chains in the metallated structure compared to the apo peptide where the cysteines are orientated towards the helical interface. The apo structure shows sufficient conformational flexibility of the cysteine side chains as out of an ensemble of many different orientations of Cys conformations, only 15% of the time all the cysteines are exclusively oriented in an orientation where they are preorganized for metal binding. The rapid motion between different cysteine conformers results in favorable entropic contribution in the apo structure. Upon metal binding this local conformational motion is restricted to a single cysteine conformer which makes the process entropically unfavorable. Moreover, there is reorganization energy required to orient the thiols to the interior for metal binding. The entropic penalty and the cost of required reorganization energy is overcome by the enthalpic gain upon formation of bonds between the soft thiolate ligands and the As(III) ion. [As(III)-thiolate bond enthalpy associated with binding of monomethyl arsenite (MMA), to dimercaptosuccinic acid (DMSA), dihydrolipoic acid (DHLA), and dithiothreitol (DTT) have been reported to be 13.2, 17, and 15.9 kcal mole^{-1}, respectively.][67] The smaller size of As(III) with ionic radius of 0.72 Å fits well within the smaller cavity generated by the interior cysteine conformer resulting in As-S bond distance of 2.28 Å and mean S-As-S angle of 89°.[26] Because of these geometrical restraints, As(III) binding is not suitable to the exterior cysteine conformer. It may be noted that when Cys in an **a** site is replaced by its bulkier analogue Pen (β, β dimethyl cysteine) in **CS**L16L-Pen the thiol groups of the

major rotamer (95%) are found to be positioned towards the interior of the coiled coil preorganized for metal binding. Improved packing of the methyl groups and restricted rotation due to the additional sterics of the bulky methyl groups, are believed to lead to this enhanced preorganization and rigidity of the thiols.[38]

The degree of preorganization of metal binding ligands in metalloproteins varies depending on the functional requirement of the protein active site. Zinc finger proteins feature a metal binding site where Zn(II) binding induces overall folding of the protein.[68] On the contrary, the R2 subunit of ribonucleotide reductase displays significantly smaller geometrical changes around metal binding region in the apo versus different dimetallated forms.[69-71] Structural comparison between As(**CS**L9C)$_3$ and (**CS**L9C)$_3$ shows that As(III) binding does not induce folding of the overall secondary structures of the peptide but the metal binding thiolate ligands in the apo structure show significant local conformational motion which is restricted upon metal binding and locked into a single conformation. The observed dynamics of Cys side chains in the apo structure suggests a mechanism for insertion of metals into the hydrophobic interior of coiled coils and helical bundles. Initial binding of a metal to a ligand such as cysteine located at the helical interface followed by a breathing motion of the helices coupled with the movement of the amino acid side chain to the interior conformer would internalize the metal within the protein's hydrophobic core.

Intriguingly, the Leu layers above and below the Cys 9 site are oriented in a similar way in both the apo and metallated (**CS**L9C)$_3$ structures. The Leu 5 layer is tucked in towards the center of the helical axis and the Leu 12 layer is oriented away from the center, thus opening up space below the metal binding site. It is believed that As(III) binds in a hemidirected **endo** conformation[26] to (**CS**L9C)$_3$ so that the stereochemically active lone pair can be housed in the open space generated by the outward orientation of the Leu 12 layer below Cys 9.

Comparison Between (**CS**L9C)$_3$ **and** (**CS**L19C)$_3$**.** Once again the peptide backbones of the two apo structures **CS**L9C and **CS**L19C overlay very well as shown in Figure 2-12. The C_β atoms of cysteines in both the structures are pointing towards the N termini of the helical scaffold. Cysteine side chains of **CS**L19C appear to be present in single conformers and show sufficient rigidity as opposed to **CS**L9C. The S-S separations in **CS**L19C vary from each other; however, they do generate a larger cavity with which to bind metals with larger ionic radii such as Pb(II). In this structure of **CS**L19C the cysteines are oriented in such a way that the thiol plane formed by S_γ atoms is not parallel to the planes formed by the C_γ atoms of Leu 16 and Leu 23. The angle between the thiol plane and Leu 16, Leu 23 planes are 8.28° and 3.02°, respectively (Figure 2-13). On the contrary the thiol plane for **CS**L9C is parallel to Leu 5 and 12. Leucines above and below Cys 19 site are at positions 16 and 23 of the peptide chain and both correspond to **a** sites of the heptad. The Leu residues at the 16 layer are positioned away from the center of the helical axis whereas those at layer 23 are directing towards the center of the helical axis, see Figure 2-14. Thus, the Leu 16 layer generates more open space compared to that of the Leu 23 layer. The orientation of these Leu layers above and below the thiol site is exactly opposite compared to those of Leu 5 and Leu 12 of **CS**L9C as mentioned earlier. These differential orientations of Leu layers above and below the cysteine **a** and **d** sites represent one of the important factors that define site selectivity and binding mode preferences of some metals to such sites. This issue has been addressed in more detail later in this chapter in the section “Implications of Binding Other Metals”.

Cys vs. Pen Structures. Recently, the X-ray structures of **CS**L16L-Pen and its diastereomeric counterpart **CS**L16D-Pen where the chirality of one amino acid is altered from L- to D- within the coiled coil, were reported.[38] Figure 2-15 shows the overlay of the α-helical backbones of the two Cys structures (**CS**L9C, **CS**L19C) and two Pen structures (**CS**L16L-Pen, **CS**L16D-Pen). The β-methylene carbons and thiol groups of **CS**L16L-Pen are oriented towards the N-termini

similar to **CS**L9C. Thiol groups of the major rotamer (95%) of each Pen residue in **CS**L16L-Pen are positioned towards the interior of the coiled coil with S-S separation of 3.7 Å as if preorganized for metal binding like the major rotamer of **CS**L9C (15%, for all Cys side chains). Enhanced degree of preorganization in the Pen structure compared to the Cys structure is believed to be incurred by improved packing of the methyl groups above and below the sulfur plane (one methyl group is pointing towards the N-termini and helical interface whereas the other one is directed towards the C-termini and the interior of the coiled coil) and restricted rotation of the bulky methyl groups. The thiols in the lower abundance rotamers in Pen (5%, each side chain) and Cys (10%, for all side chains) structures are oriented away from the helical axis, towards the helical interface. In **CS**L16D-Pen the thiols are directing towards the C-termini similar to the two thiol side chains in **CS**L19C. All the thiols in **CS**L16D-Pen are positioned away from the helical axis, towards the helical interface and thus this site is not preorganized for metal binding within the coiled coil interior. With one thiol group being oriented towards the helical interface in **CS**L19 the resulting thiol pockets in both D-Pen and **CS**L19 structures are large, in general. This analysis suggests that there is some similarity in side chain orientation and the resulting thiol site in **CS**L19C with an L-amino acid in a **d** site and in **CS**L16D-Pen with a D-amino acid in an **a** site. Because of opposite chirality of D-Pen and L-Cys, the C_β atoms in these two structures are oriented differently. Additionally, the effect of methyl groups in the D-Pen structure might have a significant impact on the observed differential orientation of the thiol groups in the two structures.

Implications for Binding Other Metals: One of the most important questions we have investigated in our lab using these designed peptides is how site selectivity can be achieved utilizing the same set of first coordination sphere ligands to the metal. One can envision multiple ways that this objective might be realized. It could be that some sites bind metals with a different goodness of fit for the metal within the group of endogenously supplied donor atoms from the protein. Alternatively, selectivity might be achieved by recognition/selection

based on exogenous ligands or stereochemically active lone pairs. It appears that such lone pair recognition may be a contributing feature for As(III) recognition as discussed above.

Implications for Cd(II) Binding. It has been observed that Cd(II) binding to **TRI** peptides yields an equilibrium mixture of 3- and 4- coordinate Cd(II) species regardless whether the Cd(II) is located in the **a** or **d** sites in the peptide. The three coordinate complex is assigned as a CdS_3 species with a planar geometry while the 4-coordinate compound is thought to have a CdS_3O structure with the fourth ligand being an exogenous water. The proportions of 3- and 4-coordinate metal are roughly similar in both **a** and **d** sites. This indicates that the leucine layers above and below the cysteines are unable to fully exclude a coordinated water in either site.

First I will discuss how the leucine layer packing may influence the accommodation of a CdS_3O species within the hydrophobic interior of these coiled coils. The first thing that one may notice is that the orientation of packing layers around **a** and **d** sites is not equivalent. The heptad repeat requires that an **a** site have the N-terminal side leucine layer four residues above the substituted cysteine position and a C-terminal side leucine layer three residues below the substituted cysteine position. In contrast, the **d** sites reverses this orientation with an N-terminal side leucine layer three residues above the substituted cysteine position and a C-terminal side leucine layer four residues below the substituted cysteine position. Hence, we can see that cysteine substituted **a** and **d** sites are inherently unequal. Examination of Figures 2-3 and 2-14 illustrates that one side of the cysteine sulfur plane is well packed by the leucines, fully blocking a potential coordinated water. However, the second layer for both **a** and **d** sites has significant rotation of the leucine side chains so that a small ‘hole’ is formed. For **CS**L9C, this opening is oriented towards the C terminus whereas for **CS**L19C the small cavity is located towards the N-terminal side of the sulfur plane. Unless metal binding should in some way lead to a rearrangement of these packing

layers (which was not observed for As(III) coordination) one would anticipate that the exogenous water would be accommodated within these leucine cavities and would suggest the cadmium atom should be located on the water containing side of the plane formed by the three cysteine sulfur donor atoms. Thus, the conclusion would be that the Cd(II) is displaced towards the C terminus for an **a** site bound metal and towards the N terminus for a **d** site coordination. These predictions would require that both the **a** site and the **d** site structures would be the **endo** isomers (an **exo** isomer has the Cd(II) ion on the opposite face of the plane formed by the three sulfur atoms from the C_β atoms while **endo** isomers have the C_β and metal on the same side of the plane as is seen for As(**CS**L9C)$_3$). Substitution of a Leu above the metal binding site to a less sterically demanding Ala generates a 'hole' which potentially allows access of water to the metal binding site on either side of the sulfur plane. The peptide **TRI**L12AL16C, therefore, binds Cd(II) solely as a four coordinate species CdS_3O.[27] If the Cd(II) binds towards the largest internal cavity (the 12 Ala layer rather than the 19 Leu layer) this would require that Cd(II) in **TRI**L12AL16C form an **exo** conformer, consistent with a longer Cd(II)-thiolate bond distance of 2.54 Å.[35]

The 3-coordinate Cd(II) structures differ markedly from the conformations seen for As(III) and predicted for CdS_3O as the CdS_3 should be planar, without a stereochemically active lone pair and an exogenous water molecule. Thus, one must consider the affinity of the metal to be within a sulfur plane. With an ionic radius of 0.92 Å, Cd(II) may be better suited to bind at the **a** site compared to the **d** site due to the greater separation of the sulfur atoms in the layer formed for the **d** site.

As mentioned earlier, Cd(II) has been shown to have a 10 fold preference for binding to the **a** site over the **d** site,[34] independent of the location of the **a** or **d** site within the hydrophobic interior (i.e., along the primary sequence of the peptide). Given the structural considerations just described, this higher affinity could be because the **a** site is better capable of tolerating a planar 3-coordinate

Cd(II). Smaller Cd-S bond distance of 2.43 Å for trigonal CdS_3 compared to that of pseudotetrahedral CdS_3O (2.54 Å) would support this. Furthermore, this would be consistent with the better fit of the Cd(II) within the **a** site plane and also due to the fact that one of the cysteine side chains in **d** site is exclusively pointed towards the helical interface, away from the interior, which would disfavor the complete reorientation of the cysteine side chain within the hydrophobic interior. Additionally, there could be a preference for the pure four coordinate structures in the **a** site. Moreover, as the relative affinities for 3-coordinate CdS_3 structures in the **a** vs. **d** site substituted peptides are not known, the present analysis does not provide a basis for predicting a significant difference in affinity for CdS_3O in these two types of sites. Clearly, metallated structures containing Cd(II) are required to further understand these structural intricacies.

Implications of Hg(II) Binding. Previous research using ^{199}Hg NMR and ^{199m}Hg PAC spectroscopic techniques has shown that Hg(II) coordinates to **a** site **TRI** peptides as a linear dithiolate HgS_2 complex at the interior of a three stranded coiled coil at pH 6.5 in the presence of 1:3 ratio of Hg(II):peptide.[72] This dithiolate-Hg(II) complex is different than $Hg(peptide)_2$ as the presence of a third helix with a protonated cysteine breaks the axial symmetry of the $Hg(peptide)_2$ complex making the Hg(II) adopt a T-shaped geometry. Studies with **d** site peptides also indicate the formation of T-shaped Hg(II) complex at the interior of three-stranded coiled coils under similar conditions (V.L.Pecoraro *et al.,* unpublished results). However, with a ^{199}Hg NMR chemical shift δ = –895 ppm the **d** site is believed to be more T-shaped than the corresponding **a** site which has a ^{199}Hg NMR chemical shift δ = –908 ppm. When the pH is raised, a trigonal $Hg(peptide)_3^-$ complex is formed. The conversion from the Hg-dithiolate to Hg-trithiolate complex occurs with a pK_a of 8.6 for the **d** site which is one pK_a unit higher than the corresponding **a** site value of 7.6. The analysis of apo structures of (**CS**L9C)$_3$ and (**CS**L19C)$_3$ reveal that two thiols in each structure are at optimal distances of 4.58 Å and 4.64 Å, respectively, to accommodate Hg(II) in linear coordination, the ideal Hg-S distance for linear HgS_2 being 2.32 Å.[34] One

of these thiols in each structure is oriented towards the interior of the helical scaffold and the other thiol to the helical interface. It is possible that linear coordination of Hg(II) in either **a** or **d** substituted peptides consists of these two types of thiols as ligating residues forming a preformed 2-coordinate site for Hg(II). Figure 2-16 shows the models of T-shaped Hg(II) complexes with **CS**L9C and **CS**L19C. It can be seen that the third thiol is at a distance of 4.31 Å and 3.33 Å from linear Hg(II) in (**CS**L9C)$_3$ and (**CS**L19C)$_3$, respectively. This difference in distances indicates that the geometry of the two T-shaped Hg-complexes are not same and may serve as structural models to explain the observed difference in ^{199}Hg NMR δ values between **a** and **d** sites. The models from this study suggest that Hg(II) fits better in an **a** site as a trigonal HgS_3 at high pH than in a **d** site which is confirmed by ^{199}Hg NMR (δ = –185 ppm for **a** site and δ = –313 ppm for **d** site).[61, 72] This observation may also serve as a plausible explanation for the higher pK_a for HgS_2-SH → HgS_3 complex formation for the **d** site compared to the **a** site. These structures provide important insight into explaining observed differences in metal binding to these sites in solution.

While I am writing my thesis, Melissa Zastrow, a graduate student in the group has solved the structure of **CS**L9PenL23H at pH 7.5 where one Hg(II) is bound to the Pen site (**a** site) in a T-shaped geometry at the interior on the coiled coil. The final refinement of the structure is in progress. Two Pen thiols are bound as linear HgS_2 with the Hg(II)-S distances of 1.79, and 2.02 Å while the third thiol is at 2.91 Å away from the Hg(II) ion. These distances are shorter than the predicted distances for the apo **a** site structure of **CS**L9C. The differences may be possibly due to the fact that the metal binding residues are different (Cys vs. Pen) in the two structures. Additionally the difference in pH (6.0 vs. 7.5) for the two structures as well as the fact that complete refinement of the Pen structure is in progress, might account for the differences in the observed distances form predictions.

Implications of Pb(II) Binding. Unlike the observations with As(III), Cd(II) and Hg(II) which have a distinctly higher affinity for **a** sites, Pb(II) exhibits a slight

(~2 to 4-fold) preference for complexation to the **d** site. Pb(II) with an ionic radius of 1.12 Å is much larger than As(III) and has a typical Pb(II)-thiolate bond distance of 2.64 Å.[35] Furthermore, as a consequence of its greater size and lower charge, Pb(II) has a larger, more diffuse lone pair which may interact with nearby leucine side chains. The difference in size and subsequent S-Pb-S angles may not allow Pb(II) to adopt as easily coordination as an **endo** conformer in an **a** site (unlike As(III)). Because Pb(II) has a bulky stereochemically active lone pair it is also unlikely that it can accept an **exo** conformation in an **a** site as the packing of Leu layers above the metal binding plane does not allow much space to accommodate the lone pair (and subsequently would require reorganization). In addition, the larger size of Pb(II) may perturb the helix, causing it to pucker out so as to fit Pb(II) in an **a** site. From the structural analysis of two apo structures presented here one can infer that Pb(II) could be accommodated in the larger site provided by the **d** substituted peptides. The thiols in a **d** site are placed further apart from each other than those at the **a** site which would simultaneously allow for a longer Pb(II)-thiolate bond length and less constraint of the S-Pb-S bond angles. It is more likely that the stereochemically active lone pair of Pb(II) would occupy the open space towards the N-termini generated by the outward packing of Leu layers above the thiol layer (see Figure 2-17), making this the most plausible mode (**endo**) of Pb(II) coordination to a **d** site in the interior of a coiled coil. It is possible that nature uses these differences discussed above to distinguish between divalent and trivalent heavy metals in metalloregulatory systems. ArsR/SmtB family of transcriptional regulator proteins detect As(III) and Sb(III) whereas CadC is used for detecting Cd(II) and Pb(II).[73]

Presence of Zn(II) at the Crystal Packing Interfaces. Crystallization of **CS**L9C and **CS**L19C was not successful in the absence of Zn(II). Binding of Zn(II) to His and Glu residues at the crystal packing interfaces links the coiled coils together and provides a third dimension to the crystals and thus generates better diffraction quality crystals for X-ray studies. In the absence of Zn(II), however, 2D plate-like crystals are obtained which are not suitable for X-ray structural studies.

As Zn(II) has intermediate polarizability (hard-soft character), it uses both hard ligands such as His, Glu, Asp and water molecules and soft ligands such as Cys in naturally occurring metalloproteins. However, in the majority (65%) of the Zn(II) containing metalloproteins, Zn(II) uses amino acids such as His, Glu, Asp and water molecules as all of its first coordination sphere ligands.[74] Zn(II) binding to His and Glu residues and waters at the crystal packing interfaces of the peptides **CS**L9C and **CS**L19C is not an exception of this later situation. Additionally, under the low pH conditions where the crystals were grown (6 for **CS**L9C, and 6.5 for **CS**L19C), Zn(II) may not prefer to bind to the cysteines at the interior of the coiled coil.

To examine the effect of external Zn(II) ions on the physical and spectroscopic properties of the bound metal at the designed metal binding site in solution, I have performed studies with **CS**L9C peptide and monitored the physical properties of Hg(II) in the presence of a varying concentration of Zn(II). Measures taken to remove ubiquitous Zn(II) ions from solution is described in the Materials and Methods. These studies were performed both at low pH, where the HgS_2 complex is formed within the three-stranded coiled and at high pH where the HgS_3 complex is formed. Given that the solubility product[75] of $Zn(OH)_2$ is 2.1×10^{-16} and the high peptide concentrations (milimolar) used in the solution studies at low pH , experiments were not possible at pH values higher than 6.5 for the UV-Vis and 5 for ^{199}Hg NMR experiments due to hydrolysis of Zn(II) leading to the formation of $Zn(OH)_2$ as precipitate. For the same reason, experiments at pH 8.6 where the HgS_3 complex is formed, were not possible to perform above 10 μM CSL9C and 30 μM Zn(II) concentrations. Due to the low concentrations of peptide used at high pH experiments, it was not possible to obtain ^{199}Hg NMR measurements.

The UV-Vis titration in Figure 2-18 shows that addition of up to 21 equivalents of Zn(II) (per monomer) to 2.5 mM **CS**L9C in the presence of 1 eq of Hg(II) (per trimer) at pH 6.5 does not seem to cause any changes in the HgS_2

signal at 247 nm as at this low pH condition Hg(II) coordinates to the peptide as linear HgS_2 within a three-stranded coiled coil.[34, 61, 72, 76] The overall UV-Vis spectra overlay very well in the presence of varying concentrations of Zn(II) except in the Trp absorbance region ~280 nm. The crystallographic conditions for **CS**L9C : Zn(II) is 5.1 mM : 17 mM, i.e. 1 : 3. For the UV-Vis experiments the **CS**L9C : Zn(II) ratio is 2.5 mM : 52.5 mM, i.e. 1 : 21. Although the absolute concentration of the peptide is half that which is present in crystals, this data shows that a 7 fold relative excess of Zn(II) compared to the crystallographic conditions has no effect on binding of Hg(II) to **CS**L9C. Titration of Hg(II) to **CS**L9C in the presence of 21 equivalents Zn(II) (per monomer) (Figure 2-19, A) and in the absence of Zn(II) (Figure 2-19, B) show that the stoichiometry of **CS**L9C : Hg(II) is unchanged (3 : 0.95) in the presence of Zn(II). Additionally, the molar extinction coefficient (2050 $M^{-1}cm^{-1}$ at 247 nm) in the presence of Zn(II) is similar to the value of 2000 $M^{-1}cm^{-1}$ observed in the absence of Zn(II). Similar values of molar extinction coefficient have been observed for linear HgS_2 within a three-stranded coiled coil for TRI peptides.[61, 76] Hg(II) binding titration data both in the presence and absence of Zn(II) were analyzed by non-linear least squares fits. Because of strong binding, the minimum affinities obtained for Hg(II) binding to **CS**L9C in the presence and absence of Zn(II) are 1.4×10^7 M^{-1} and 2.3×10^7 M^{-1}, respectively. ^{199}Hg NMR chemical shift of 12 mM **CS**L9C and 252 mM Zn(II) (1 : 21 **CS**L9C : Zn(II), a 6 fold absolute excess of Zn(II) compared to crystallographic conditions) in the presence of 1 equivalent Hg(II) (per trimer) at pH 5 is –926 ppm (Figure 2-20, A), expected for a T-shaped Hg(II) complex within a three-stranded coiled coil. The same chemical shift is observed under similar conditions in the absence of Zn(II) (Figure 2-20, B). These studies show that the presence of Zn(II) up to a concentration that is several fold in excess compared to what is present under crystallographic conditions does not influence the stoichiometry, affinity and coordination number of Hg(II) bound at the designed metal site.

Titration of $HgCl_2$ to 10 μM **CS**L9C in 50 mM CHES buffer in the presence (Figure 2-21, A) and absence (Figure 2-21, B) of 30 μM $Zn(OAC)_2$ at pH 8.6

shows that the stoichiometry of **CS**L9C : Hg(II) is not influenced by the presence of 3 fold excess of Zn(II). The molar extinction coefficient of HgS_3 complex is similar in the presence and absence of Zn(II) at pH 8.6, the values being 18000 and 19500 M^{-1} cm^{-1} , respectively, at 247 nm. Hg(II) titration data were fit using non-line least square analysis. Due to strong binding the minimum values of binding constants obtained are 6.6×10^7 M^{-1} and 4.3×10^7 M^{-1} in the presence and absence of Zn(II), respectively. These results show that the presence of a 3 fold excess of Zn(II) does not influence the stoichiometry and affinity of Hg(II) bound at the Cys site as HgS_3.

Conclusions

Having solved the crystal structures of apo **CS**L9C, an **a** site, apo **CS**L19C, a **d** site peptide, as well as the previously reported apo L-Pen (an analogue of Cys) and D-Pen **a** site structures by our group, we are now in a unique position to evaluate the structural implications on modifying backbone chirality vs. the subtle differences we have just reported between **a** vs. **d** sites. At first glance, there is a striking similarity between the D-Pen in an **a** site and a L-amino acid (Cys) in a **d** site. Though these sites are not identical, (most significantly the beta-C points towards the C-terminus for the D-amino acid) qualitatively, the orientation of thiols leads to larger thiol pockets in both the structures. From this similarity one can envision that a D-amino acid in a **d** site may resemble an **a** site with L- amino acid. Furthermore, this approach may allow us to alter the various types of **a** vs. **d** site specificity we have previously reported. The structure of apo (**CS**L9C)$_3$ now provides a comprehensive basis for a direct comparison of the metallated and apo three-stranded coiled coils in these types of designed polypeptide systems. The (**CS**L9C)$_3$ structure shows that the Cys side chains have conformational flexibility and the orientation is partitioned between the coiled coil interior and helical interface. The apo (**CS**L19C)$_3$ structure shows that the Cys sulfurs are present as a single conformer and the thiol site generates a larger pocket in the hydrophobic interior. The

overall secondary structure of both the (**CS**L9C)$_3$ and (**CS**L19C)$_3$ are unperturbed from one another. However, the packing of hydrophobes above and below the Cys site are opposite in the two types of structures. Using **CS**L9C peptide I have also shown that several fold excess of Zn(II) at the crystal packing interfaces does not influence the physical properties of linear (dithiolate) or trigonal (trithiolate) Hg(II) bound at the interior of the three-stranded coiled coil. Using the lessons learned from the structures of apo **CS**L9C and **CS**L19C, in this chapter I have also discussed the possible coordination modes in which heavy metals such as As(III) Cd(II) Hg(II) and Pb(II) can be sequestered by these peptides. With the knowledge gained from this study we now have access to the rules for the *de novo* design of metal binding sites within the interior of coiled coils, as well as a deeper understanding of heavy metal binding in designed proteins, potentially providing significant insight into metal binding in metalloregulatory proteins and metallochaperones. Clearly, additional structures of metallated constructs would give direct evidence of the possible ways of metal coordination to these peptides. Studies addressed in this chapter, which demonstrate the salient features of two *de novo* designed apo structures, can now serve as a basis for understanding solution behavior of metal encapsulation and physical properties of metallo derivatives of **Coil Ser** and related *de novo* designed family of peptides.

References:

[1] L. Regan, W. F. DeGrado, *Science* **1988**, *241*, 976.
[2] W. F. DeGrado, Z. R. Wasserman, J. D. Lear, *Science* **1989**, *243*, 622.
[3] J. W. Bryson, S. F. Betz, Z. X. Lu, D. J. Suich, H. X. Zhou, *Science* **1995**, *270*, 935.
[4] G. Tuchscherer, L. Scheibler, P. Dumy, M. Mutter, *Biopolymers* **1998**, *47*, 63.
[5] L. Baltzer, J. Nilsson, *Curr. Opin. Biotechnol.* **2001**, *12*, 355.
[6] W. F. DeGrado, *Nature* **2003**, *423*, 132.
[7] Y. Lu, *Inorg. Chem.* **2006**, *45*, 9930.
[8] R. L. Koder, P. L. Dutton, *Dalton Trans.* **2006**, 3045.
[9] J. Kaplan, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 11566.
[10] B. A. Krizek, D. L. Merkle, J. M. Berg, *Inorg. Chem* **1993**, *32*, 937.
[11] M. D. Struthers, R. P. Cheng, B. Imperiali, *J. Am. Chem. Soc.* **1996**, *118*, 3073.
[12] M. D. Struthers, R. P. Cheng, B. Imperiali, *Science* **1996**, *271*, 342.
[13] C. L. Hill, D. J. Steenkamp, R. H. Holm, T. P. Singer, *Proc. Natl. Acad. Sci., U.S.A.* **1977**, *74*, 547.
[14] B. R. Gibney, F. Rabanal, J. J. Skalicky, A. J. Wand, P. L. Dutton, *J. Am. Chem. Soc.* **1999**, *121*, 4952.
[15] R. G. Daugherty, T. Wasowicz, B. R. Gibney, V. J. DeRose, *Inorg. Chem.* **2002**, *41*, 2623.
[16] R. P. Bonomo, L. Casella, L. D. Gioia, H. Molinari, G. Impellizzeri, T. Jordan, G. Pappalardo, E. Rizzarelli, *J. Chem. Soc. Dalton Trans.* **1997**, 2387.
[17] R. P. Cheng, S. L. Fisher, B. Imperiali, *J. Chem. Soc. Dalton Trans.* **1996**, *118*, 11349.
[18] M. M. Rosenblatt, J. Y. Wang, K. S. Suslick, *Proc. Natl. Acad. Sci. USA* **2002**, *100*, 13140.
[19] J. P. Schneider, J. W. Kelly, *J. Am. Chem. Soc.* **1995**, *117*, 2533.
[20] G. Platt, C. Chung, M. Searle, *Chem. Commun.* **2001**, 1162.
[21] J. Venkatraman, G. A. Naganagowda, R. Sudha, P. Balaram, *Chem. Commun.* **2001**, 2660.
[22] G. R. Dieckmann, D. K. McRorie, J. D. Lear, K. A. Sharp, W. F. DeGrado, V. L. Pecoraro, *J. Mol. Biol.* **1998**, *280*, 897.
[23] D. Ghosh, V. L. Pecoraro, *Inorg. Chem.* **2004**, *43*, 7902.
[24] B. Lovejoy, S. Choe, D. Cascio, D. McRorie, W. DeGrado, D. Eisenberg, *Science* **1993**, *259*, 1288.
[25] G. R. Dieckmann, D. K. McRorie, D. L. Tierney, L. M. Utschig, C. P. Singer, T. V. O'Halloran, J. E. Penner-Hahn, W. F. DeGrado, V. L. Pecoraro, *J. Am. Chem. Soc.* **1997**, *119*, 6195.
[26] D. S. Touw, C. E. Nordman, J. A. Stuckey, V. L. Pecoraro *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969.

[27] K.-H. Lee, C. Cabello, L. Hemmingsen, E. N. G. Marsh, V. L. Pecoraro, *Angew. Chem., Int. Ed.* **2006**, *45*, 2864.
[28] A. F. A. Peacock, L. Hemmingsen, V. L. Pecoraro, *Proceedings of the National Academy of Sciences* **2008**, *105*, 16566.
[29] O. Iranzo, C. Cabello, V. L. Pecoraro, *Angew. Chem., Int. Ed.* **2007**, *46*, 6688.
[30] N. L. Ogihara, M. S. Weiss, W. F. DeGrado, D. Eisenberg, *Protein Sci.* **1997**, *6*, 80.
[31] B. Farrer, C. McClure, J. E. Penner-Hahn, V. L. Pecoraro, *Inorg. Chem* **2000**, *39*, 5422.
[32] B. T. Farrer, N. P. Harris, K. E. Balchus, V. L. Pecoraro, *Biochemistry* **2001**, *40*, 14696.
[33] B. Farrer, V. L. Pecoraro, *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3760.
[34] M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2002**, *124*, 8042.
[35] M. Matzapetakis, D. Ghosh, T.-C. Weng, J. E. Penner-Hahn, V. L. Pecoraro, *J. Biol. Inorg. Chem.* **2006**, *11*, 876.
[36] M. Matzapetakis, V. L. Pecoraro, *J. Am. Chem. Soc.* **2005**, *127*, 18229.
[37] H. Wendt, C. Berger, A. Baici, R. M. Thomas, H. R. Bosshard, *Biochem.* **1995**, *34*, 4097.
[38] A. F. A. Peacock, J. A. Stuckey, V. L. Pecoraro, *Angewandte Chemie International Edition* **2009**, *48*, 7371.
[39] J. R. Calhoun, F. Nastri, O. Maglio, V. Pavone, L. A.;, W. F. DeGrado, *Peptide Sci.* **2005**, *80*, 264.
[40] A. Lombardi, C. M. Summa, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, *Proc. Natl. Acad. Sci., USA* **2000**, *97*, 6298.
[41] S. J. Lahr, D. E. Engel, S. E. Stayrook, O. Maglio, B. North, S. Geremia, A. Lombardi, W. F. DeGrado, *Journal of Molecular Biology* **2005**, *346*, 1441.
[42] L. D. Costanzo, H. Wade, S. Geremia, L. Randaccio, V. Pavone, W. F. DeGrado, A. Lombardi, *J. Am. Chem. Soc.* **2001**, *123*, 12749.
[43] O. Maglio, F. Nasti, V. Pavone, A. Lombardi, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3772.
[44] A. Pasternak, J. Kaplan, J. D. Lear, W. F. Degrado, *Protein Science* **2001**, *10*, 958.
[45] W. Shi, J. Dong, R. A. Scott, M. Y. Ksenzenko, B. P. Rosen, *J. Biol. Chem.* **1996**, *271*, 9291.
[46] Y.-F. Lin, J. Yang, B. P. Rosen, *J. Biol. Chem.* **2007**, *282*, 16783.
[47] M. Matzapetakis, University of Michigan (Ann Arbor), **2005**.
[48] D. S. Touw, *Ph.D Thesis* **2007**, *University of Michigan*.
[49] W. C. Chan, and White, P. D., Oxford University Press, New York, **2000**.
[50] Z. Otwinowski, W. Minor, in *Methods in Enzymology: Macromolecular Crystallography, part A, Vol. 276* (Ed.: R. M. Sweet), Academic Press, New York, **1997**, pp. 307.
[51] E. Potterton, Briggs, P., Turkenburg, M., Dodson, E, *Acta Crystallographica Section D* **2003**, *59*, 1131.

[52] A. J. McCoy, Grosse-Kunstleve, R.W., Adams, P. V., Winn, M.D., Storoni, L.C. Read, R.J, *Journal of Applied Crystallography* **2007**, *40*, 658.
[53] G. N. Murshudov, Vagin, A.A., Dodson, E.J., *Acta Crystallographica Section D* **1997**, *53*, 240.
[54] P. Emsley, Cowtan, K., *Acta Crystallographica Section D* **2004**, *60*, 2126.
[55] P. D. Adams, R. W. Grosse-Kunstleve, L.-W. Hung, T. R. Ioerger, A. J. McCoy, N. W. Moriarty, R. J. Read, J. C. Sacchettini, N. K. Sauter, T. C. Terwilliger, *Acta Crystallographica Section D* **2002**, *58*, 1948.
[56] P. Emsley, K. Cowtan, *Acta Cryst. D* **2004**, *60*, 2126.
[57] A. T. Brunger, P. D. Adams, G. M. Clore, W. L. DeLano, P. Gros, R. W. Grosse-Kunstleve, J.-S. Jiang, J. Kuszewski, M. Nilges, N. S. Pannu, e. al., *Acta Cryst. D* **1998**, *54*, 905.
[58] W. L. DeLano, *The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, California, USA* **2005**, http://www.pymol.org.
[59] C. F. Macrae, I. J. Bruno, J. A. Chisholm, P. R. Edgington, P. McCabe, E. Pidcock, L. Rodriguez-Monge, R. Taylor, J. van de Streek, P. A. Wood, *Journal of Applied Crystallography* **2008**, *41*, 466.
[60] G. L. Ellman, *Archives of Biochemistry and Biophysics* **1959**, *82*, 70.
[61] O. Iranzo, D. Ghosh, V. L. Pecoraro, *Inorg. Chem.* **2006**, *45*, 9959.
[62] M. Luczkowski, M. Stachura, V. Schirf, B. Demeler, L. Hemmingsen, V. L. Pecoraro, *Inorganic Chemistry* **2008**, *47*, 10875.
[63] C. Cobas, J. Cruces, F. J. Sardina, 2.3 ed., Universidad de Santiago de Compostela, Spain, **2000**.
[64] **Nomenclature**, *While previously describing in detail the nomenclature for peptide modifications, we have not fully provided descriptors for metal complexation to these peptides. We will provide these definitions here with respect to Cd(II) and the exogenous ligand water; however, the nomenclature is general for any metal or non-protein bound ligand. We refer to an "apo peptide" as one which does not have metal in the designed metal coordination site within the hydrophobic interior of the coiled coils. We do not refer to an "apo peptide" to represent situations without the presence of any cofactors. Metals that may be bound non-specifically to the exterior hydrophilic residues are not considered to be metallated proteins/peptides in this context. Cd(II)(TRIL16C)$_3^-$ indicates that Cd(II) is bound to the three stranded coiled coil; however, the metal coordination environment is either unknown or is a mixture of species (e.g., Cd(II)(TRIL16C)$_3^-$ contains a 55:45 mixture of CdS$_3$O and CdS$_3$, respectively). When the metal is placed within brackets, as in [Cd(II)(H$_2$O)](TRIL12AL16C)$_3^-$, we specify the coordination environment (in this case CdS$_3$O) and the possible exogenous ligand (here water). If the site has only protein ligands, then it would be represented as [Cd(II)](TRIL16Pen)$_3^-$. At low pH, one can encounter the situation where the cysteine sulfurs are protonated and may, or may not, be directly coordinated to the metal. In this case, we indicate this situation as [Cd(II)](TRIL16Pen[S,(SH)$_2$])$_3^-$ in which one cysteine is bound as a thiolate and the remaining two cysteines are protonated. When a peptide contains*

more than one binding site, we use the above nomenclature, but specify which site is being discussed by adding a superscript outside of the bracket to indicate the amino acid residues coordinating the specific metal (e.g., $[Cd(II)]^{16}[Cd(II)(H_2O)]^{30}(GRANDL16PenL26AL30C)_3^{-2}$). Finally for cases where dual site peptides exhibit specificity for metal complexation, we will utilize [apo] to designate an empty site (e.g., $[apo]^{16}[Cd(II)(H_2O)]^{30}(GRANDL16PenL26AL30C)_3^{-}$).

[65] G. N. Ramachandran, Ramakrishnan, C., Sasisekharan, V., *J. Mol. Biol.* **1963**, *7*, 95.

[66] J. W. Ponder, F. M. Richards, *Journal of Molecular Biology* **1987**, *193*, 775.

[67] A. M. Spuches, H. G. Kruszyna, A. M. Rich, D. E. Wilcox, *Inorg. Chem.* **2005**, *44*, 2964.

[68] J. M. Berg, *Acc. Chem. Res.* **1995**, *28*, 14.

[69] D. T. Logan, F. deMare, B. O. Persson, A. Slaby, B.-M. Sjoberg, P. Nordlund, *Biochemistry* **1998**, *37*, 10798.

[70] M. E. Andersson, M. Hogbom, A. Rinaldo-Matthis, K. K. Andersson, B.-M. Sjoberg, P. Nordlund, *Journal of the American Chemical Society* **1999**, *121*, 2346.

[71] M. Atta, P. Nordlund, A. Aberg, H. Eklund, M. Fontecave, *Journal of Biological Chemistry* **1992**, *267*, 20682.

[72] O. Iranzo, P. V. Thulstrup, S.-B. Ryu, L. Hemmingsen, V. L. Pecoraro, *Chem. Eur. J.* **2007**, *13*, 9178.

[73] M. A. Pennella, D. P. Giedroc, *BioMetals* **2005**, *18*, 413.

[74] I. S. Bertini, A.; Sigel, H., *Handbook of Metalloproteins* **2001**, *Marcel Dekker Inc.: New York*.

[75] D. D. Ebbing, S. D. Gammon, *General Chemistry*, Eighth ed., Houghton Mifflin, New York, Boston, **2008**.

[76] V. L. P. Pecoraro, Anna F. A.; Iranzo, Olga; Luczkowski Marek., in *Bioinorganic Chemistry*, American Chemical Society, Washington, DC, **2009**, pp. 183.

Table 2-1. Peptide sequences used in this study.

Peptide	*Sequence*
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
Grand	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
Coil Ser	Ac-E WEALEKK LAALESK LQALEKK LEALEHG-NH_2
CSL9C	Ac-E WEALEKK **C**AALESK LQALEKK LEALEHG-NH_2
CSL19C	Ac-E WEALEKK LAALESK LQA**C**EKK LEALEHG-NH_2
CSL16Pen	Ac-E WEALEKK LAALESK **X**QALEKK LEALEHG-NH_2
CSL16DPen	Ac-E WEALEKK LAALESK **X**QALEKK LEALEHG-NH_2

X = Penicillamine. Residues in red indicate substitutions from parent peptides
D-Pen = D- Penicillamine

Table 2-2. Data Collection, Phasing and Refinement statistics for CSL9C

Data Collection Statistics	
Data Set	Native
SpaceGroup	C2
Unit Cell	a = 77.940, b = 29.856, c = 44.361 $\alpha = 90, \beta = 119.15, \gamma = 90$
Wavelength (Å)	0.97872
Resolution (Å)[1]	50-1.36 (1.38-1.36)
R_{sym} (%)[2]	6.9 (35.0)
$<I/\sigma I>$[3]	20 (5)
Completeness (%)[4]	99.4 (100)
Redundancy	6.7 (6.0)
Beamline	LS-CAT 21-ID-F
Refinement Statistics	
Resolution (Å)	1.36 (34.03-1.36)
R-Factor (%)[10]	18.2
R_{free} (%)[11]	21.5
Protein atoms[12]	816
Water Molecules	59
Unique Reflections	18278
R.m.s.d.[13]	
Bonds (Å)	0.024
Angles (˚)	2.364

[1]Statistics for highest resolution bin of reflections in parentheses.
[2]$R_{sym} = \Sigma_h \Sigma_j \, l \, I_{hj} - <I_h> \, l \, / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $<I_h>$ is the mean intensity for multiply recorded reflections.
[3]Intensity signal-to-noise ratio.
[4]Completeness of the unique diffraction data.
[5]Resolution cut-off used during heavy-atom refinement and phase calculations.
[6]$R_{iso} = \Sigma \, l \, lF_{ph}I - IF_pI \, I \, / \, \Sigma \, IF_pI$, where F_p and F_{ph} are the native and heavy-atom structure factor amplitudes, respectively.
[9]Number of heavy-atom sites per asymmetric unit.
[10]R-factor = $\Sigma_h \, I \, IF_oI - IF_cI \, I \, / \, \Sigma_h IF_oI$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.
[11]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement. [12]Total number of protein atoms refined in the asymmetric unit.
[13]Root mean square deviation of bond lengths and bond angles.

Table 2-3. Data Collection, Phasing and Refinement statistics for CSL19C:

Data Collection Statistics	
Data Set	Native
SpaceGroup	$P6_1$
Unit Cell	a =b=38.08, c = 178.75 $\alpha = \beta = 90, \gamma = 120$
Wavelength (Å)	1.00
Resolution (Å) [1]	2.00 (2.07-2.0)
R_{sym} (%)[2]	6.3 (35.0)
$<I/\sigma I>$[3]	5
Completeness (%)[4]	97.3 (89.1)
Redundancy	8 (7)
Beamline	LS-CAT 21-ID-D
Refinement Statistics	
Resolution (Å)	2.15 (32.99-2.15)
R-Factor (%)[10]	21.9
R_{free} (%)[11]	26.4
Protein atoms[12]	699
Water Molecules	19
Unique Reflections	7281
R.m.s.d. [13]	
Bonds (Å)	0.025
Angles (˚)	2.318

[1]Statistics for highest resolution bin of reflections in parentheses.
[2]$R_{sym} = \Sigma_h \Sigma_j \, l \, I_{hj} - <I_h> \, l \, / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $<I_h>$ is the mean intensity for multiply recorded reflections.
[3]Intensity signal-to-noise ratio.
[4]Completeness of the unique diffraction data.
[5]Resolution cut-off used during heavy-atom refinement and phase calculations.
[6]$R_{iso} = \Sigma \, l \, lF_{ph}I - IF_pI \, I \, / \, \Sigma \, IF_pI$, where F_p and F_{ph} are the native and heavy-atom structure factor amplitudes, respectively.
[9]Number of heavy-atom sites per asymmetric unit.
[10]R-factor = $\Sigma_h \, I \, IF_oI - IF_cI \, I \, / \, \Sigma_h IF_oI$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.
[11]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement. [12]Total number of protein atoms refined in the asymmetric unit.
[13]Root mean square deviation of bond lengths and bond angles.

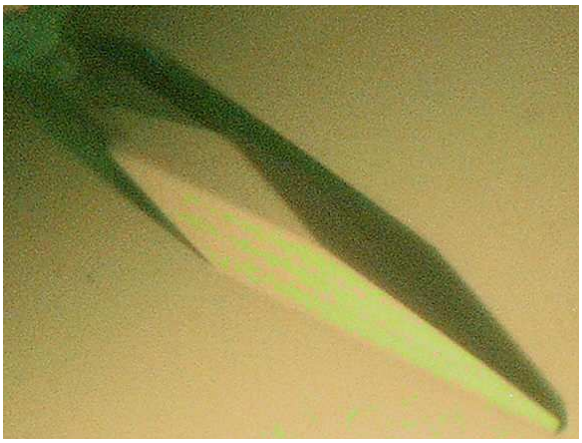


Figure 2-1. A single crystal of **CS**L9C. Cystal growing conditions are 19.15 mg/mL **CS**L9C, 17 mM $Zn(OAc)_2$, 100 mM sodium cacodylate buffer at pH 6.0, and 100 mM $Ca(OAc)_2$, and 24% PEG 3350 as precipitant.

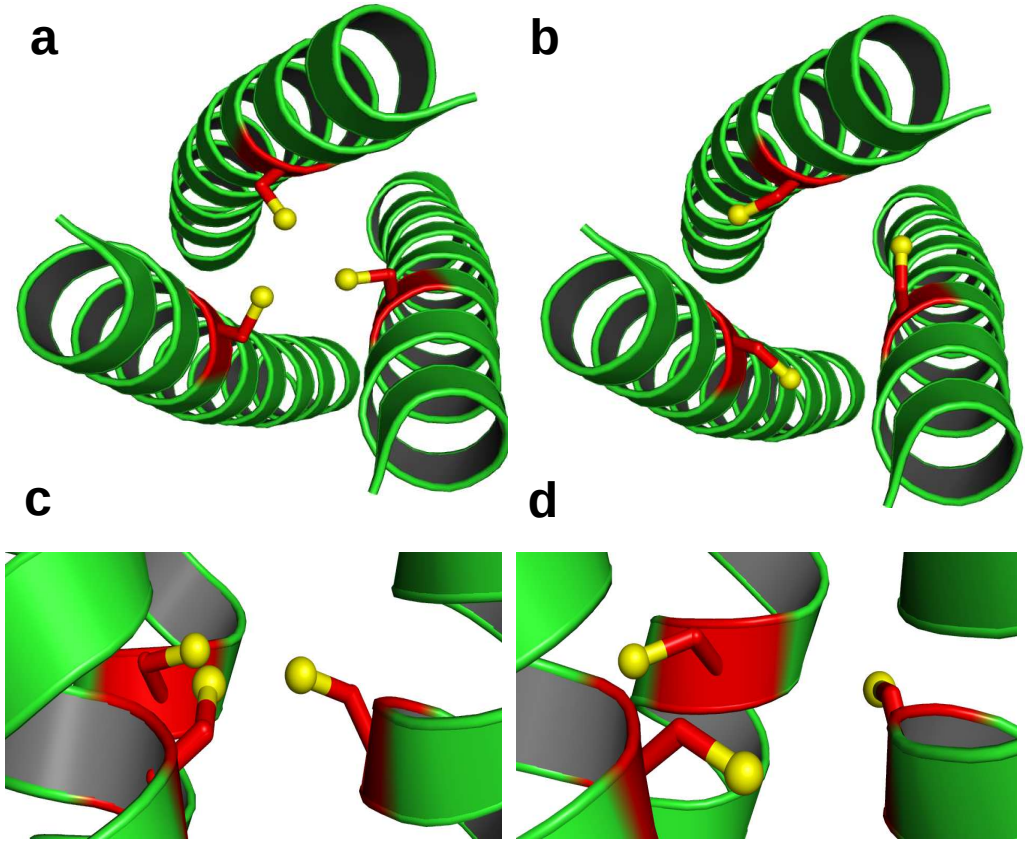


Figure 2-2. Ribbon diagrams of (**CS**L9C)$_3$ showing the orientation of Cys side chains. Main chain atoms are shown as green ribbon and the Cys side chain as red stick with the thiol groups being shown as yellow spheres. Top-down view from N termini showing the exclusive orientation of the a) major conformer (15%) where all the Cys side chains are pointing towards the interior of the coiled coil and b) minor conformer (10%) where all Cys side chains are oriented towards the helical interface. Side views show that the thiol groups are pointing towards the N- termini and C- termini of the helical scaffold in the interior c) and d) exterior Cys conformers, respectively.

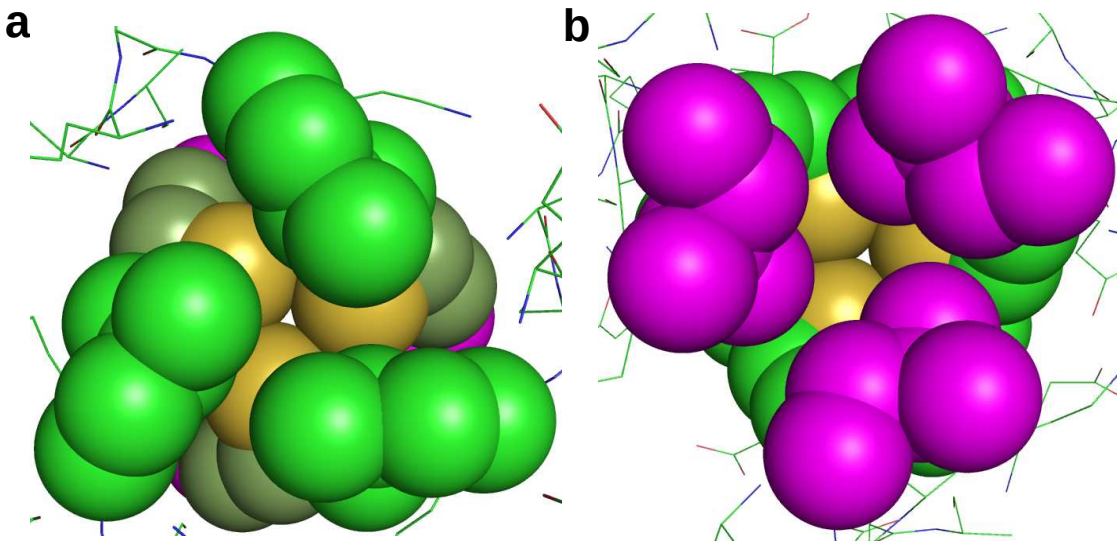


Figure 2-3. Packing of Leu layers above and below the Cys site of **CS**L9C trimer shown as spheres. a) Top-down view from N–termini, showing the Leu 5 layer (green) and b) Leu 12 layer (magenta), viewing from C-termini up the helical axis. It shows that metals with stereochemically active lone pair such as As(III), would bind to a coiled coil in an **a** site such that the lone pair would be housed in the open space towards the C-termini.

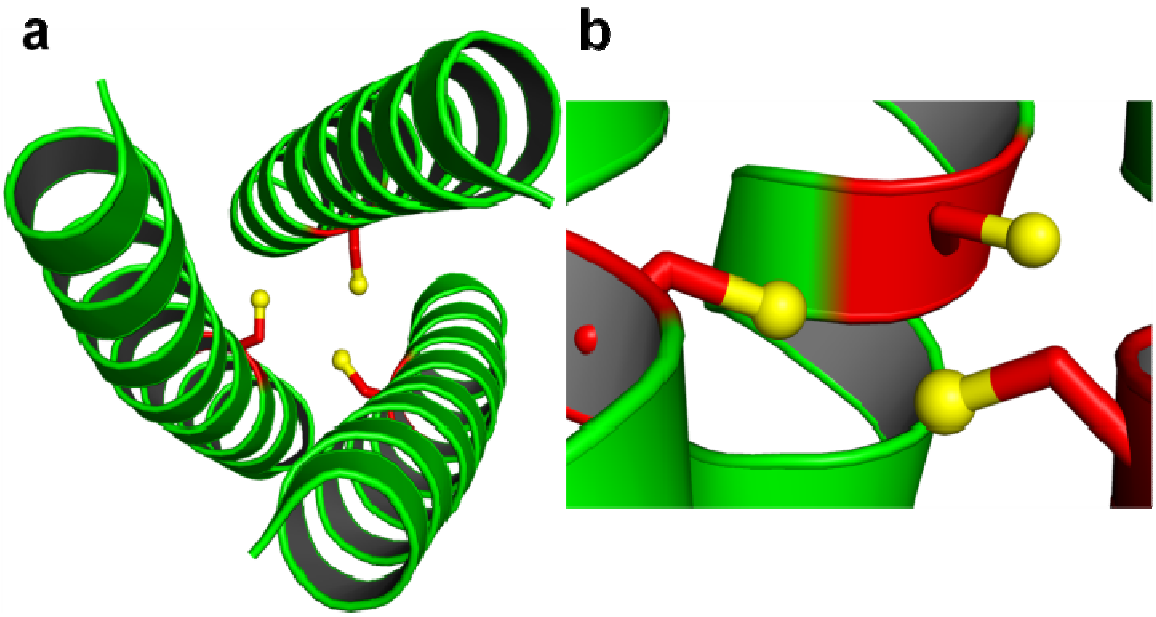


Figure 2-4. Ribbon diagrams of (**CS**L19C)$_3$ showing the orientation of Cys side chains. Main chain atoms are shown as green ribbon and the Cys side chain as red sticks with the thiol groups being shown as yellow spheres. a) Top-down view from N termini showing that two Cys side chains are pointing towards the interior of the coiled coil with the third thiol directed towards the helical interface. b) Side view shows that two thiol groups are pointing towards the C- termini of the helical scaffold and the third thiol is almost perpendicular to the helical axis.

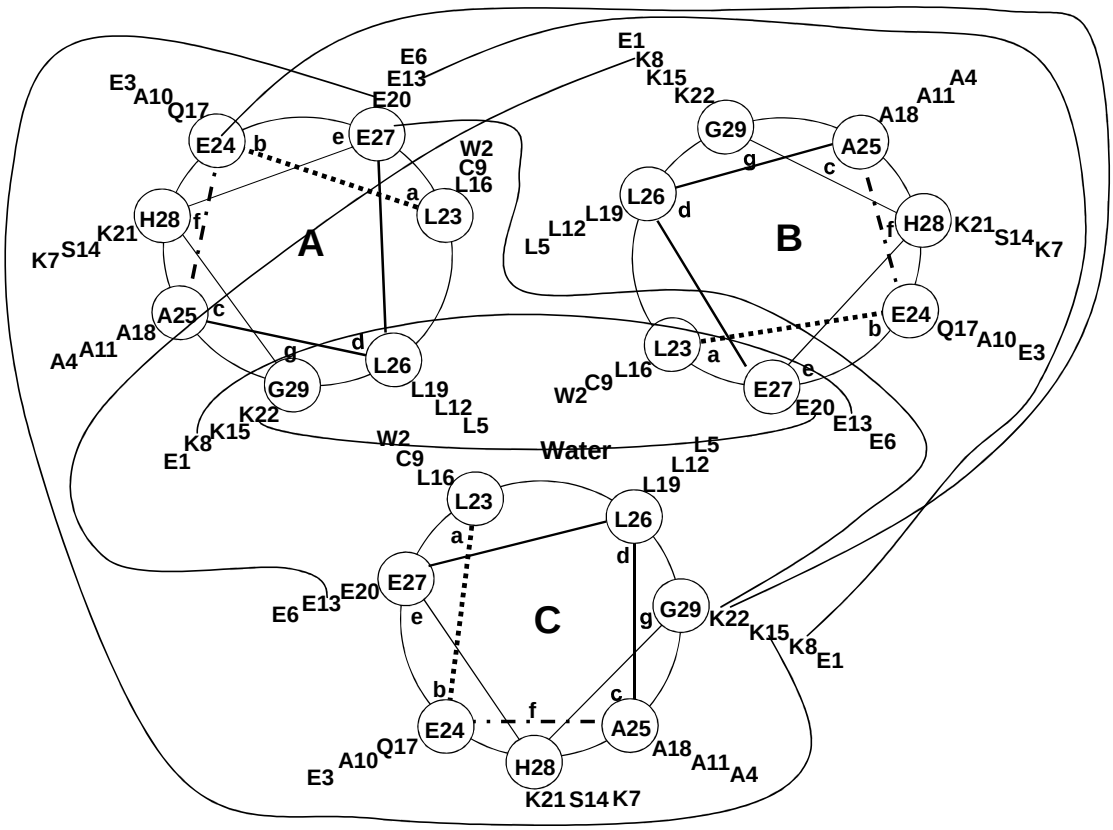


Figure 2-5. Helical wheel diagram of (**CS**L9C)$_3$. Solid lines between different residues from helices A, B and C represent interhelical electrostatic interactions.

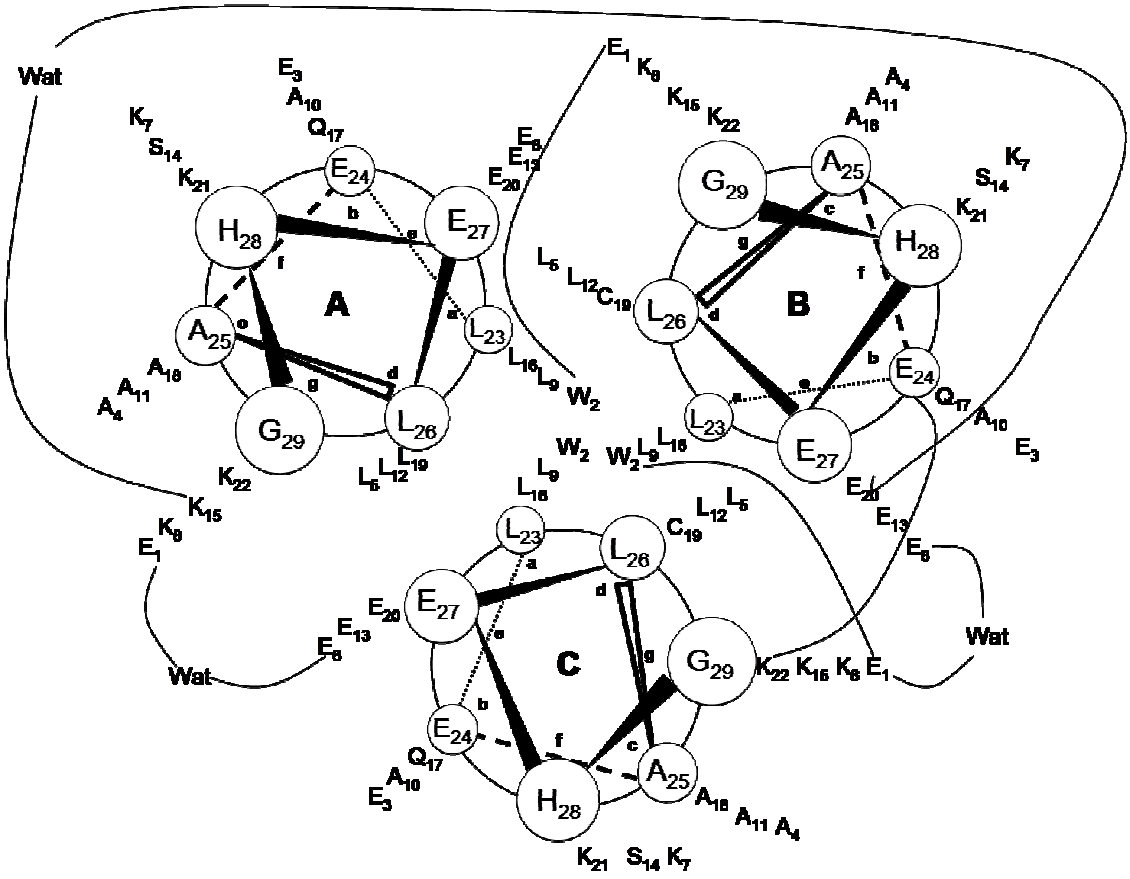


Figure 2-6. Helical wheel diagram of (**CS**L19C)$_3$. Solid lines between different residues from helices A, B and C represent interhelical electrostatic interactions. This figure is taken from Touw, D.S., Ph.D. Thesis (2006), University of Michigan, Ann Arbor.

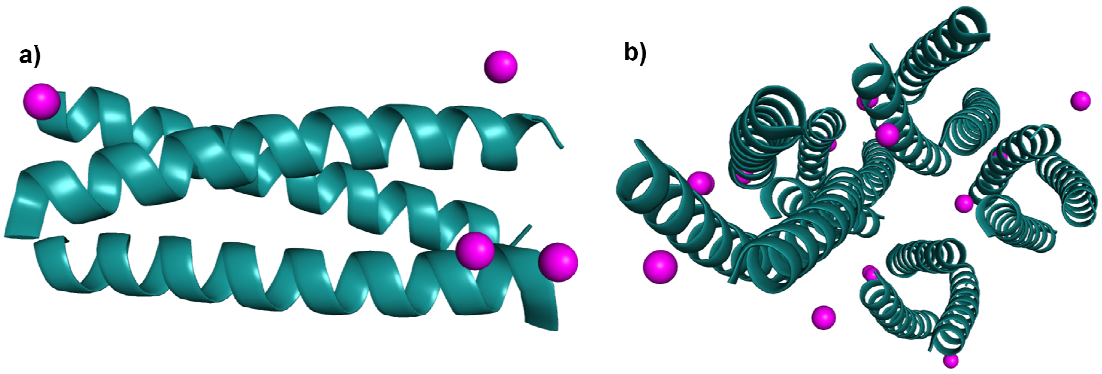


Figure 2-7. Zn(II) ions (magenta) at the crystal packing interfaces of **CS**L9C. Shown are a) the four Zn(II) ions present at the crystal packing interfaces of each trimer and b) a view of the trimers that coordinate to all the four Zn(II) ions.

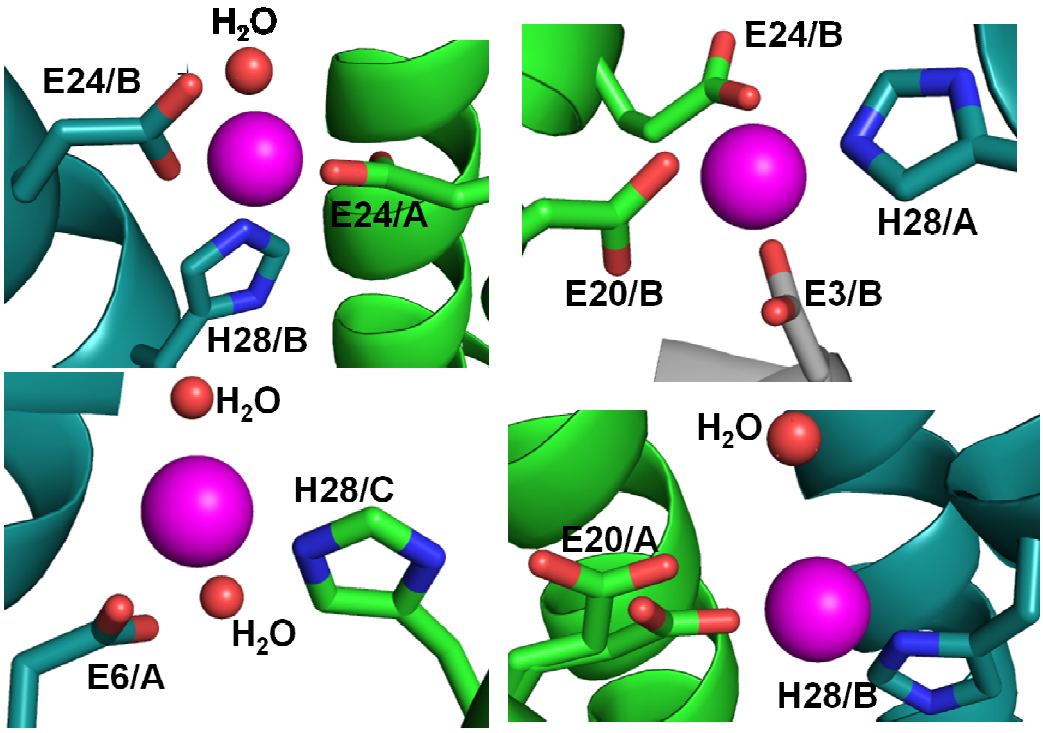


Figure 2-8. Detailed view showing the ligating residues to the four Zn(II) ions (magenta) present at the crystal packing interfaces of each trimer. Respective amino acids (Glu/His) from different symmetry related peptide molecules are shown as separate colors. Waters are shown as red spheres.

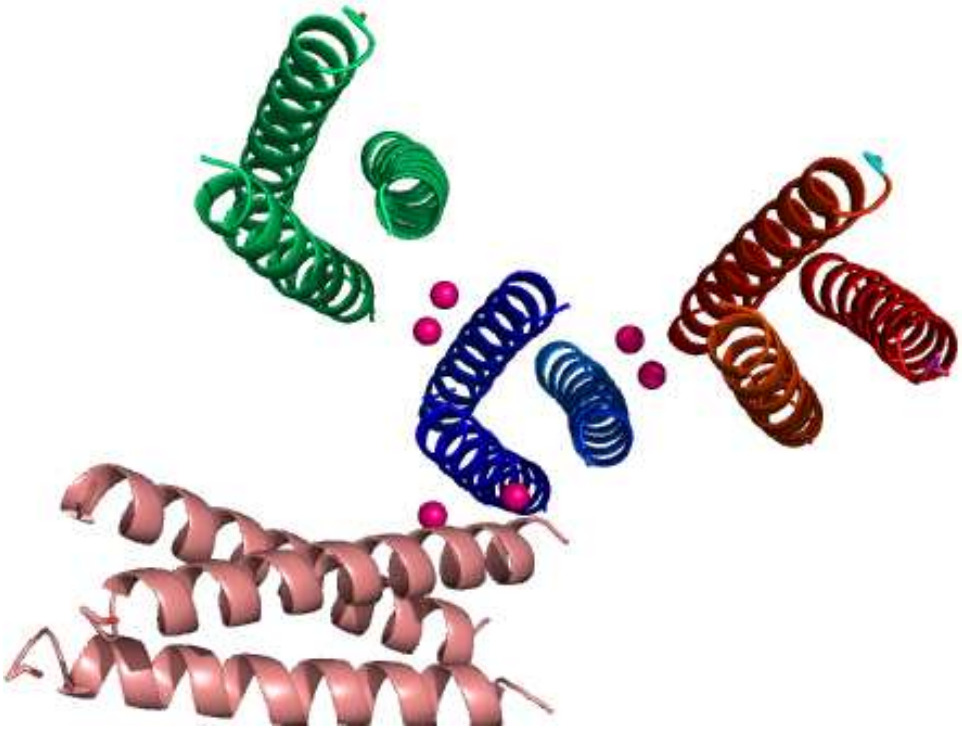


Figure 2-9. Illustration of the location of the three pairs of Zn(II) ions shown in magenta spheres at crystal packing interfaces of **CS**L19C trimers. This figure is taken from Touw, D.S., Ph.D. Thesis (2006), University of Michigan, Ann Arbor.

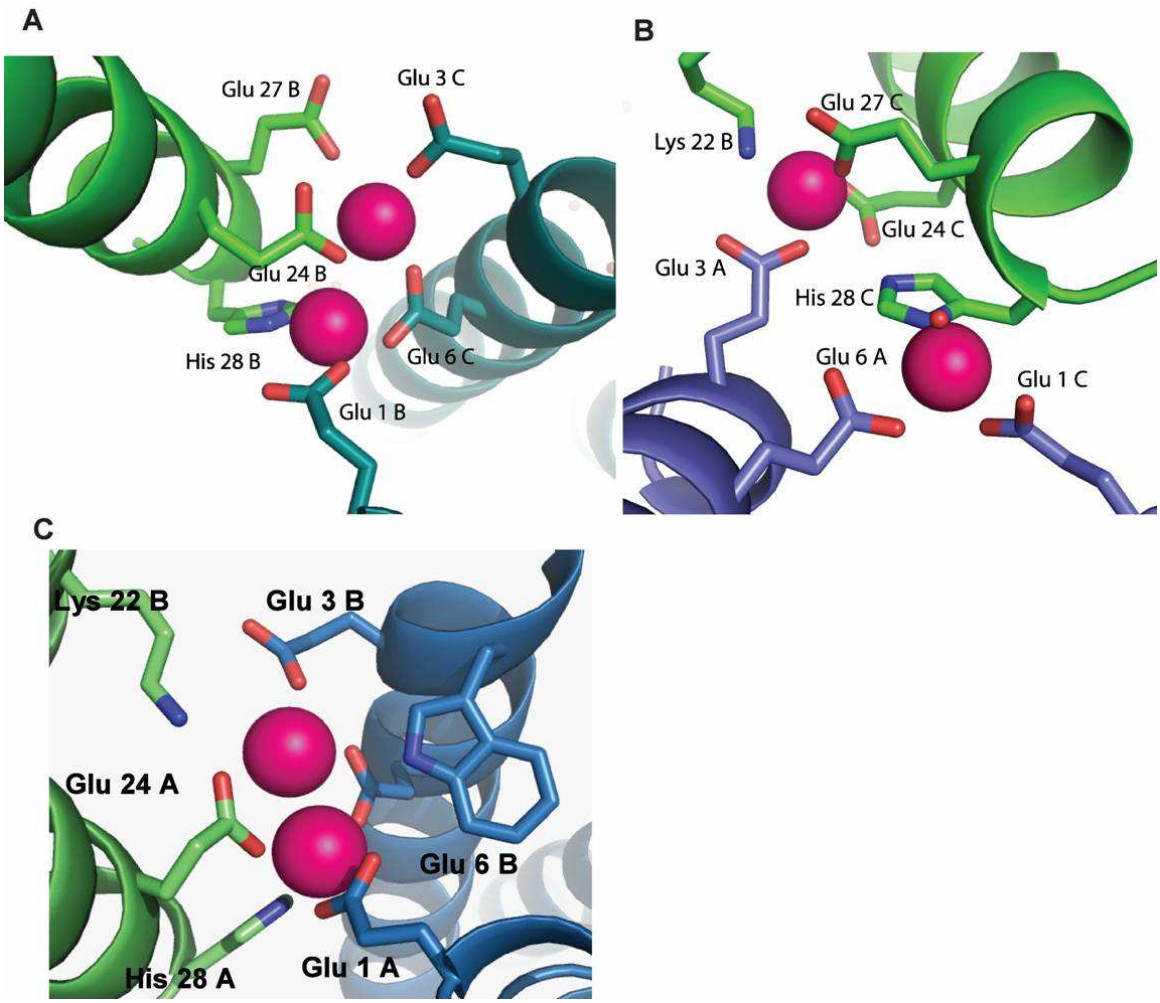


Figure 2-10. A detailed view of the six Zn(II) ions at crystal packing interfaces of (**CS**L19C)$_3$. The Zn(II) ions shown in magenta are coordinated tetrahedrally by glutamate, histidine and lysine residues and water molecules. This figure is taken from Touw, D.S., Ph.D. Thesis (2006), University of Michigan, Ann Arbor.

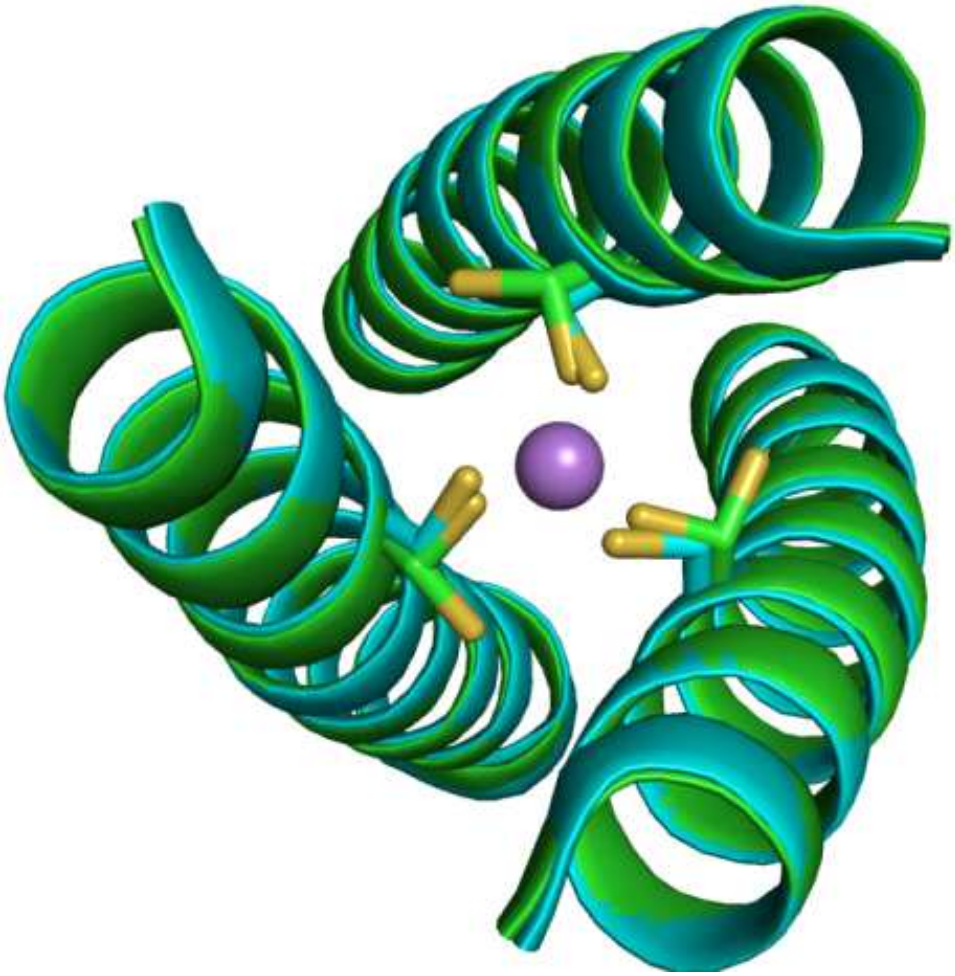


Figure 2-11. Ribbon diagram showing top-down view of an overlay of apo (**CS**L9C)$_3$ (green) and As(**CS**L9C)$_3$ (cyan) structures. Cysteine side chains are shown in stick form with As(III) as purple sphere. It can be noted that there is no change in the overall secondary structures of apo and metallated peptide. The major change can be noted from the apo structure where Cys side chains are present as alternate conformers which is locked into a single conformer in As(**CS**L9C)$_3$ upon metal binding.

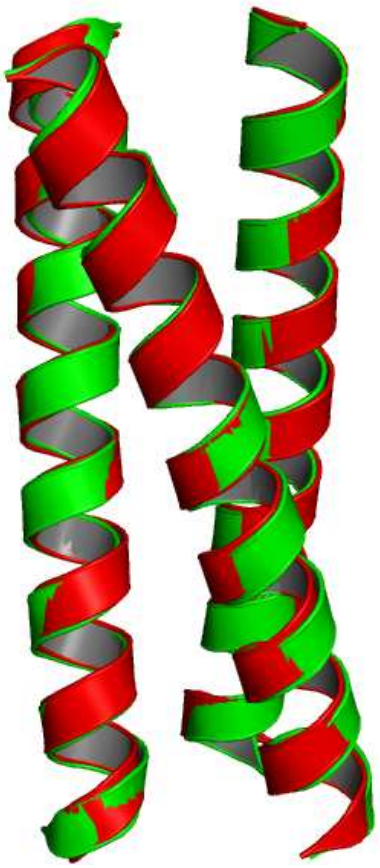


Figure 2-12. Ribbon diagram showing the overlay of apo (**CS**L9C)$_3$ (green) and (**CS**L19C)$_3$ (red). The secondary structures in both the constructs are unperturbed.

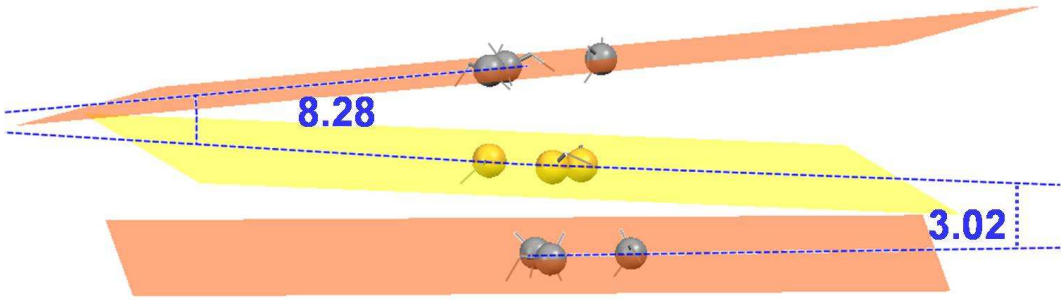


Figure 2-13. Shown are the planes formed by S_γ atoms (yellow spheres) of Cys 19 and C_γ atoms (grey spheres) of Leu 16 and Leu 23 of (**CS**L19C)$_3$. The angles formed by the thiol plane with Leu 16 and Leu 23 planes are shown to be 8.28° and 3.02°, respectively.

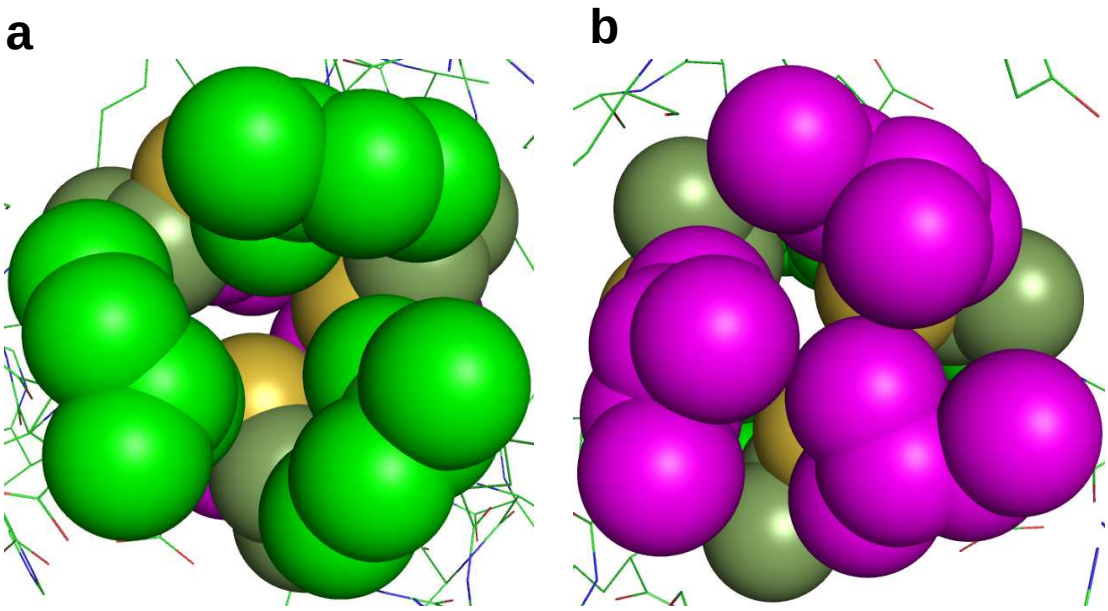


Figure 2-14. Packing of Leu layers above and below the Cys site of **CS**L19C trimer shown as spheres. a) Top-down view from N–termini of Leu 16 layer (green) and b) Leu 23 layer (magenta), viewing from C-termini up the helical axis.

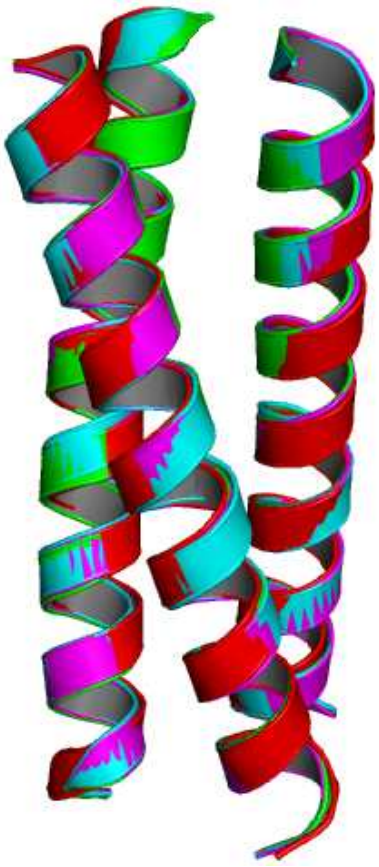


Figure 2-15. Ribbon diagram showing the overlay of (**CS**L9C)$_3$ (green), (**CS**L19C)$_3$ (red), (**CS**L16L-Pen)$_3$ (magenta) and (**CS**L16D-Pen)$_3$ (cyan) structures.

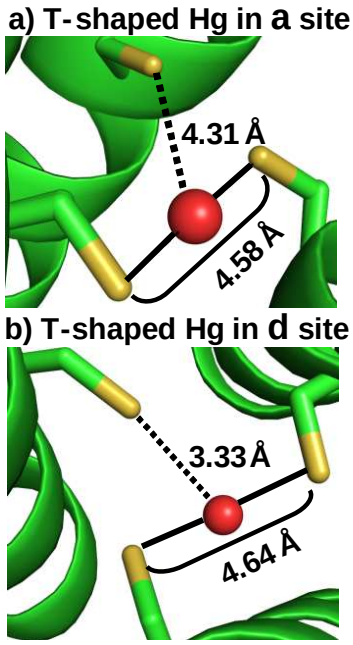


Figure 2-16. Models of T-shaped Hg(II) complexes inside three-stranded coiled coils. Shown are possible coordination modes of Hg(II) with the two thiols in a) **CS**L9C and b) **CS**L19C where one thiol is pointing towards the interior and the other towards the helical interface. Peptide backbone is shown as green ribbon, Cys residues as ball and stick and Hg(II) as red sphere. Solid lines between sulfur atoms and Hg(II) in the two models represent S-Hg-S bonds; whereas the dotted lines from the third sulfur to the Hg(II) represent situations where this third sulfur may or may not be coordinated to the Hg(II) ion. With S-S separation of 4.58 Å and 4.64 Å these two thiols are well suited to bind Hg(II) in linear coordination. The third thiol in **CS**L19C being at a shorter distance than in **CS**L9C may explain why Hg(II) in **d** sites is believed to be more T-shaped than in **a** site. [Structure of **CS**L9PenL23H, solved by Melissa Zastrow, a graduate student in the group shows that the Hg(II) is bound to the **a** site Pen residues in a T-shaped geometry. Hg(II)-S distances for linear HgS_2 are 1.79 and 2.02 Å, respectively; while the third thiol is at 2.91 Å from the Hg(II) ion. The differences in the observed Hg(II)-S distances from predictions may be due to the difference in the coordinating amino acids residues (Cys in **CS**L9C vs. Pen in **CS**L9PenL23H), pH (6 for Cys structure vs. 7.5 for Pen structure). Refinement of the Pen structure is in progress].

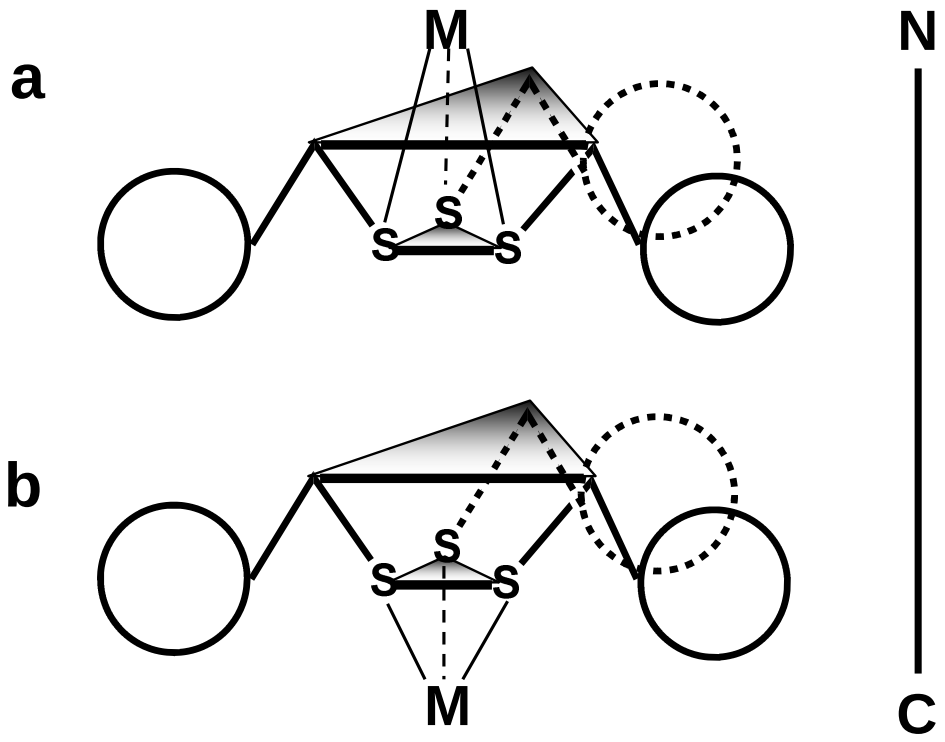


Figure 2-17. Schematic showing a) endo and b) exo configuration of a metal (M) bound to cysteines in a **d** site of a three stranded coiled coil peptide. The β methylenes of cysteines are shown to be directing towards the N-termini and S_γ atoms towards the C-termini. In endo configuration the metal and β-carbon plane (top plane) of cysteines are on the same site of the sulfur atom plane (bottom plane). In exo configuration the metal and β-carbon plane are on the opposite side of the sulfur atom plane.

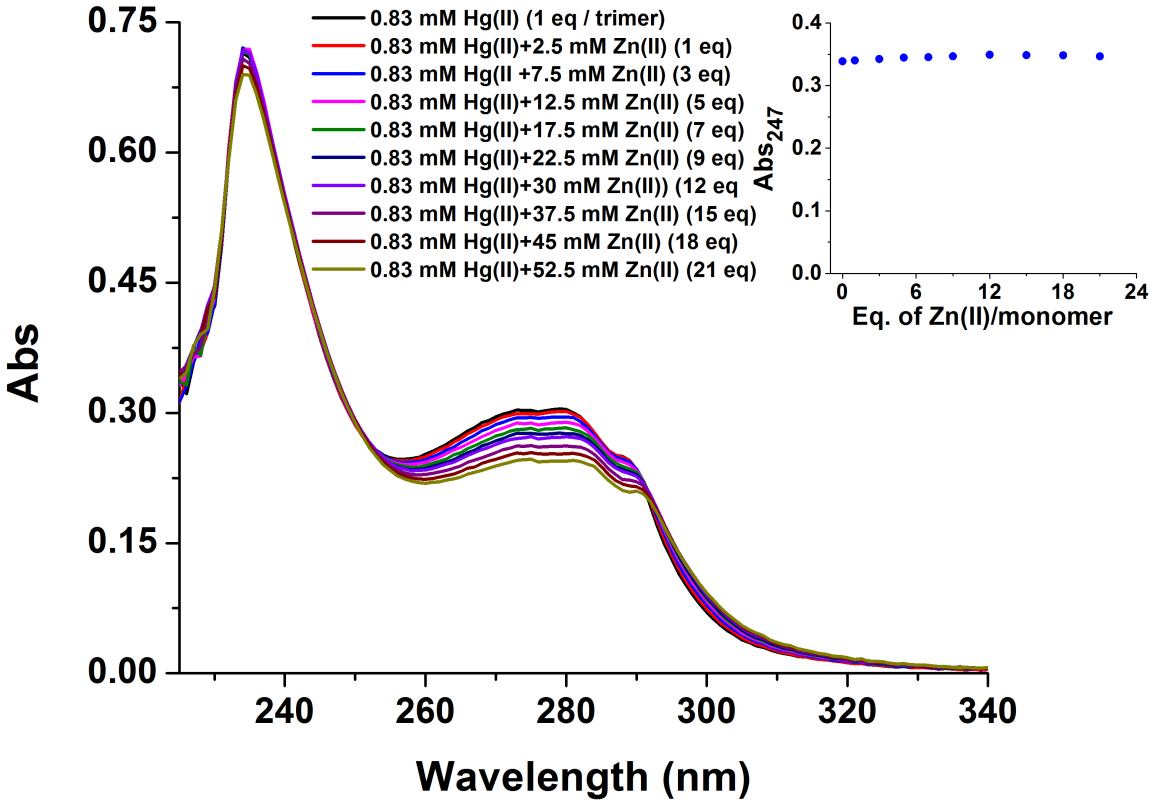


Figure 2-18. UV-Vis spectra of solutions containing 2.5 mM **CS**L9C and 0.8 mM $HgCl_2$ (1 eq/trimer) in 50 mM MES buffer at pH 6.5 with gradual addition of up to 52.5 mM $Zn(OAc)_2$ (21 eq/monomer). The inset shows the plot of absorbance at 247 nm (corresponding to HgS_2 within a three-stranded coiled coil) vs equivalents of added Zn(II)/monomer.

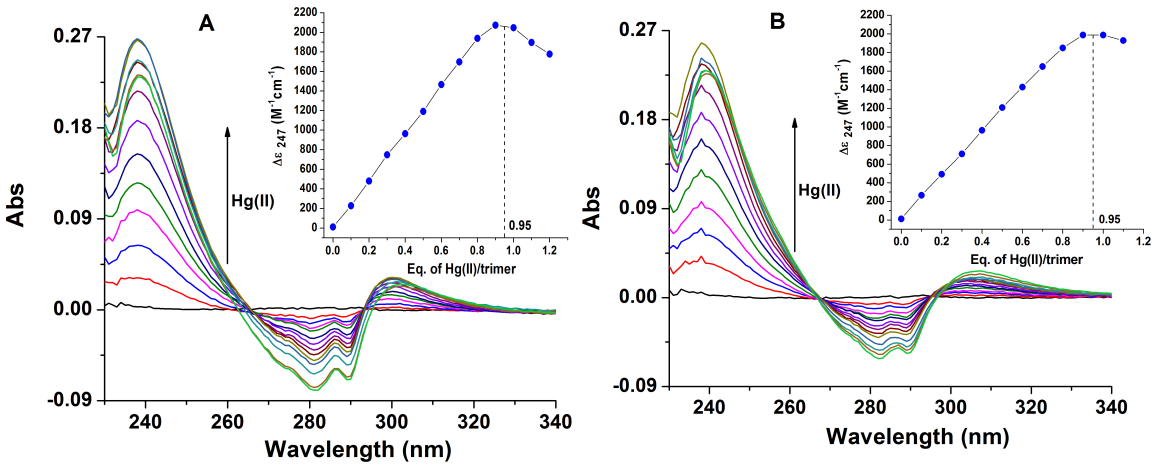


Figure 2-19. UV-Vis spectra showing the titration of $HgCl_2$ to a solution containing 2.5 mM **CS**L9C in 50 mM MES buffer at pH 6.5 in the presence of 52.5 mM $Zn(OAc)_2$ (A) and in the absence of Zn(II) (B). The insets of the figures show the titration curves obtained by plotting the change in extinction coefficient at 247 nm (corresponding to HgS_2 within a three-stranded coiled coil) as a function of added equivalents of Hg(II)/trimer of peptide. The negative absorbance at 280 nm region in both the presence and absence of Zn(II) corresponds to the Trp chromophore which is most likely being affected by the addition of Hg(II).

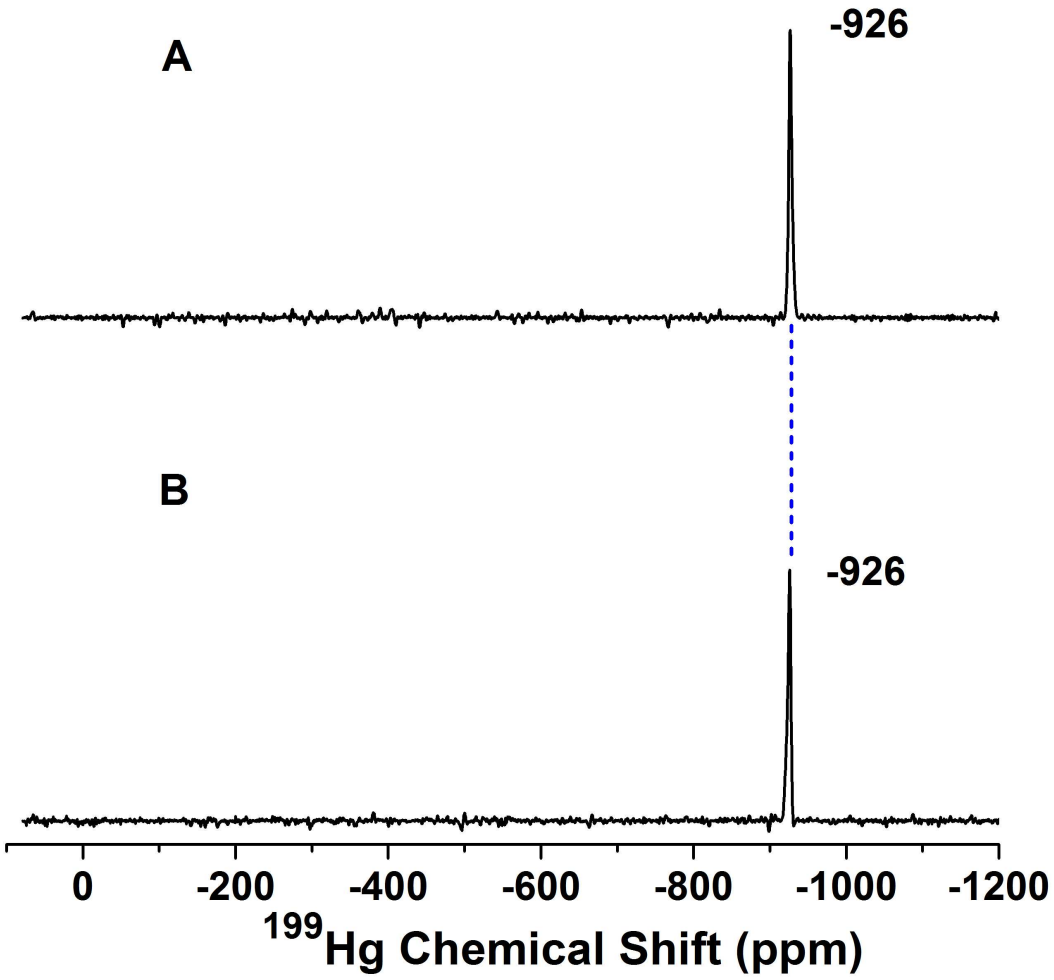


Figure 2-20. ^{199}Hg NMR spectra of 12 mM **CS**L9C and 4 mM $^{199}Hg(NO_3)_2$ at pH 5 in the presence of 252 mM $Zn(OAc)_2$ (A) and in the absence of $Zn(OAc)_2$ (B).

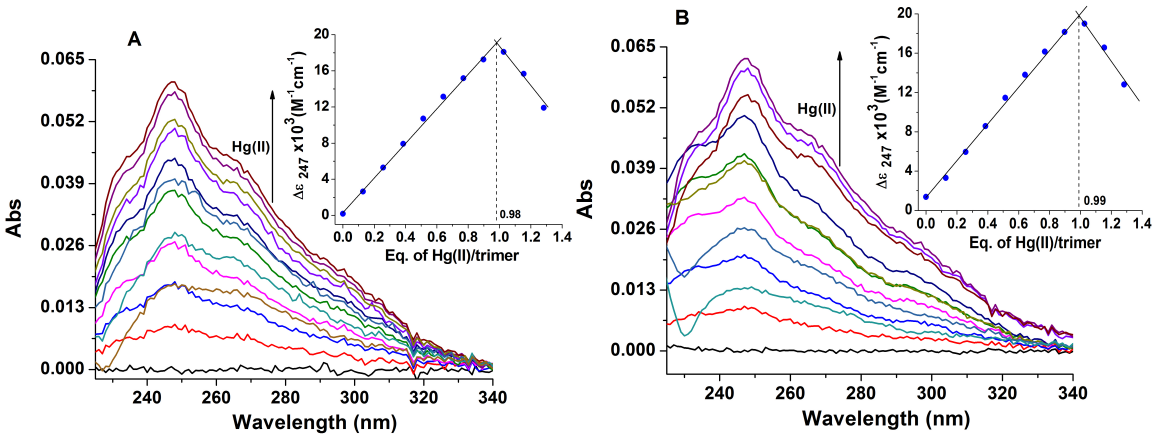


Figure 2-21. UV-Vis spectra showing the titration of $HgCl_2$ to a solution containing 10 μM **CS**L9C in 50 mM CHES buffer at pH 8.6 in the presence of 30 μM $Zn(OAc)_2$ (A) and in the absence of Zn(II) (B). The insets of the figures show the titration curves obtained by plotting the change in extinction coefficient at 247 nm (corresponding to HgS_3) as a function of added equivalents of Hg(II)/trimer of peptide.

Chapter 3

Controlling and Fine Tuning the Physical Properties of Two Identical Metal Coordination Sites in *De Novo* Designed Three Stranded Coiled Coil Peptides

Introduction

Over the last two decades the emerging field of *de novo* metalloprotein design has empowered chemists to understand biologically relevant metal binding sites at a molecular level and gain important fundamental knowledge on how the protein matrix influences and controls the metal ion properties.[1-3] Owing to significant advances in peptide synthesis, computational and structural biology it has become possible to design linear amino acid sequences that generate scaffolds with a well defined secondary and tertiary structure. This approach allows one to elucidate, in a systematic manner, structural features and elements that play crucial roles in determining the coordination geometries and physical properties of metal ions, as well as their site specificity. This degree of knowledge is necessary not only to comprehend fully how metalloproteins work and fine tune the metal ion geometries and properties, but also for the design of new bio-molecules with improved properties and 'state of the art' functionalities. Given this promise, it is not surprising that the field of *de* novo metalloprotein design is attracting a significant amount of attention as is shown by the reviews that have appeared recently.[4]

One of the most targeted secondary structural motifs in the *de novo* peptide design field is the α-helix and in general it has been used to generate α-

helical coiled coils and bundle structures. These helical aggregates result from sequences containing a seven-residue [heptad] repeat, denoted [**abcdefg**]$_n$, that typically has hydrophobic residues at **a** and **d** positions, and polar or charged residues at **e, f** and **g** positions. Despite this apparently simple pattern, the heptad repeat approach has shown its high versatility and potential. A wide variety of *de novo* designed structures has already been reported which differ in the nature (homomers/heteromers), number (from dimers to heptamers) and alignment (parallel/antiparallel) of the helices.[5] From an inorganic chemical point of view, these structures represent perfect scaffolds where diverse metal ion binding sites can be engineered. Indeed, this strategy has been employed by several research groups to design metallopeptides with specified characteristics.

Two major strategies have been employed to introduce metal ion binding sites into these helical scaffolds: a) the use of cofactors, mainly heme units, or b) the replacement of the amino acids in the hydrophobic cores by amino acids with metal ion coordinating properties. The Dutton and Gibney groups have used four-helix bundles as maquettes to study heme containing proteins.[6-9] Gibney and coworkers also used the maquette motifs to develop a prototype ferredoxin maquette, a peptide-based synthetic analogue of natural [4Fe-4S]$^{2+/+}$ proteins.[10] DeGrado,[11,12] Ghirlanda,[13] Lombardi and Pavone,[14] Mihara,[15,16] and their co-workers also used this approach to prepare different coiled coils and helical bundles which can bind between one and four heme units depending on the design. The second strategy was used by DeGrado and co-workers to design the dueferri (DF) peptides, a family of four-helix bundles, that contain a Glu_4-His_2 metal binding site capable of coordinating 2 equivalents of Fe(II), Mn(II) and Zn(II). These *de novo* designed metallopeptides are paradigms for di-iron, di-manganese and di-zinc proteins.[17,18]. The Ogawa group has designed Cys containing peptides which assemble into coiled coils upon binding Cd(II) and Cu(I), forming a CdS_4 and a Cu_4S_4 center, respectively.[19,20] On the same line, using Cys or His as donor ligands, Tanaka and co-workers designed peptides

that bound Hg(II) and Cd(II),[21] and Ni(II), Zn(II), Co(II) and Cu(II),[22-24] respectively. These last systems[19-23] and others[25,26] displayed metal-ion induced self-assembly.

Before one can use *de novo* design for the more lofty goal of chemical catalysis it is important to understand how to control the coordination environment of what would be the resting state of the active site. To this end, the Pecoraro group has used the *de novo* design approach to define the chemistry of Cd(II), Hg(II), As(III) and Pb(II) in thiolate rich environments, with particular emphasis on controlling the coordination number of the bound metal ions. Using the heptad repeat approach, we have described how (above pH 5.5) the parent peptide **TRI** (Table 3-1)[27] associates in aqueous solution to form parallel, three-stranded coiled coils. These aggregates bind metals in a sulfur rich environment ranging from HgS_2 (linear), HgS_3 (trigonal planar) and HgS_4 (tetrahedral) to $As(III)S_3$ (trigonal pyramidal).[28-31] We have also prepared peptides that contain either Cys or the non-coding amino acid analogue Pen (penicillamine) in their interiors, binding Cd(II) with high affinity and different coordination geometries. Substitution of the Leu by Cys or Pen generates partially or fully preorganized homoleptic thiol-rich binding sites inside the hydrophobic core of these coiled coils.[32,33] By combining ^{113}Cd NMR and ^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopy, we have shown how the peptide **TRI**L16C binds Cd(II) as a mixture of pseudotetrahedral (CdS_3O) and trigonal planar (CdS_3) structures.[34] Complete control of the Cd(II) coordination number was achieved by modifying the sterics in the metal ion binding site through changes in the amino acids of either the first or second coordination sphere. Namely, removal of steric bulk directly above the Cys plane by replacement of a Leu with an Ala generated the peptide **TRI**L12AL16C that binds Cd(II) exclusively in the CdS_3O geometry.[35,36] On the other hand, the pure CdS_3 geometry was achieved by increasing steric constrains around the metal ion binding site. Two different approaches were used in the latter case, both employing non-natural amino acids, a key benefit of the chemical peptide synthesis. In the first approach, Cys was replaced by Pen which has bulky methyl groups replacing the β-methylene hydrogen atoms

(peptide **TRI**L16Pen).[36] In the second strategy, the chirality of the Leu just above the Cys was modified (L-Leu was replaced by D-Leu) and as a consequence, the side chain of the Leu was reoriented towards the C termini and towards the metal ion binding site (peptide **TRI**L12L_DL16C**)**.[37] This knowledge has driven our design strategy for the heterochromic peptides, **GRAND**L16PenL26AL30C and **GRAND**L12L_DL16CL26AL30C.[37,38] These peptides, containing two binding sites in close proximity (~ 20 Å) within the same three-stranded coiled coil, are capable of binding two Cd(II) ions with different coordination geometries (three coordinate, trigonal planar as CdS_3 and four coordinate, pseudotetrahedral as CdS_3O) and thus different physical properties. Furthermore, these peptides show site-selective Cd(II) recognition, where in both peptides, Cd(II) binds selectively to the 4-coordinate site (CdS_3O) regardless of the pH conditions. These heterochromic peptides are important achievements towards our goal of preparing designed metalloproteins for achieving site-selective molecular recognition of metal ions.

While the heterochromic peptides allowed us to discriminate between 3- and 4-coordination sites that were at a first glance nearly identical, this success raised an important new question: Are all 3-coordinate or 4-coordinate sites in the same peptide identical? Said differently, can we refer to, as an example, an environment simply as a 4-coordinate site or we must specify whether this sulfur rich environment is in the center of the helix versus towards the helical terminus or whether the Cys resides are in **a** vs. **d** heptad positions? If there is a difference, will the physical properties of the center be modified, will there be any selectivity for binding of the same metal in two sites or the dynamics of the structures be different? Ultimately, we can address whether the position of these sites in the coiled coil play any role in fine tuning the physical properties of the bound Cd(II).

To address these issues we have designed sets of single peptides capable of binding two equivalents of Cd(II) with the same coordination geometry. A construct such as this will have two 3-coordinate sites (CdS_3) or two 4-coordinate sites (CdS_3O) at different positions within the coiled coil. Figure 3-1

shows models for these two constructs created in Pymol.[39] These peptides would allow us to directly assess the selectivity of metal binding between two equivalent positions within the peptide, whether binding of one metal would influence the physical properties of the second bound metal and if different physical properties can be observed for the same metal in two similar first coordination environments but in different topological positions within the coiled coil. With the constructs described herein, the important issues of how the specific position of these sites could modify the intrinsic properties of the bound Cd(II) are addressed.

Experimental Section

The nomenclature used to describe the metal complexes prepared in this study is provided in ref [[40]].

Peptide Synthesis and Purification. The **TRI** and **GRAND** peptides (see Table 3-1 for sequence nomenclature) were synthesized on an Applied Biosystems 433A peptide synthesizer using standard Fmoc protocols,[41] and purified and characterized as described previously.[29]

Ultraviolet-Visible (UV-Vis) Spectroscopy. All solutions were purged with argon before use to minimize the chances of oxidation of the peptides and formation of disulfide bonds. Fresh stock solutions of the purified peptides were prepared for each experiment in doubly distilled water and their concentrations determined by quantization of the Cys thiol groups using a known assay with 4,4'-dipyridyl disulphide.[42]

a) Metal Binding Titrations. Cd(II) into peptide titrations were performed at room temperature on a Cary 100 Bio UV-Vis spectrometer using a 1-cm quartz cuvette. Aliquots of 8.93 mM $CdCl_2$ stock solution were added into a 3-mL solution containing 60 μM peptide and 50 mM appropriate buffer (TRIS for pH 8.5 and CHES for pH 9.0 and 9.5). In each case, the difference spectra were obtained by subtracting the background spectrum of the peptide in the absence

of metal (60 μM peptide and 50 mM appropriate buffer). Difference molar absorption coefficients were calculated based on the total metal concentrations.

b) pH Titrations. UV-Vis pH titrations were carried out at room temperature on an Ocean Optics SD 2000 fiber optic spectrometer and the pH measured using a mini-glass combination pH electrode (Hamilton Biotrode) coupled to a Fisher Accumet digital pH meter model 805 MP. pH Titrations were performed by adding small aliquots of concentrated solution of KOH to unbuffered solutions containing $CdCl_2$ (20 – 40 μM) and peptide (60 – 120 μM), and monitoring the change in absorbance at 235 nm as a function of pH. Equilibration time was always allowed before reading the final pH. In all cases, reverse titrations were carried out by adding small aliquots of a concentrated solution of HCl to verify the reversibility of the process. The UV-Vis pH titration curves of the peptides containing a single binding site were fit using the model and procedure published previously for the release of two protons upon Cd(II) binding to the three Cys thiolates.[43,44] The experimental data for the peptides **GRAND**L12AL16CL26AL30C and **GRAND**L16PenL19IL23PenL26I, that contain two metal binding sites, were fit using a model that was derived from the single binding site model. The model from the single binding site was modified so as to incorporate two metal binding sites where Cd(II) binds non-specifically to both the binding sites. Non-specific binding of Cd(II) to both the binding sites was observed in the ^{113}Cd NMR experiments (see the Results section for the ^{113}Cd NMR experiments). The model used to derive the fitting equations for analyzing the pH titration curves of the peptides containing two metal binding sites is presented in Scheme 3-1. According to the model, at low pH Cd(II) is bound to both the binding sites as $(CdS(SH)_2)^+$ which do not have UV absorbance. Thus, considering both the sites, the two Cd(II) centers at low pH can be represented as $(Cd_2S_2(SH)_4)^{2+}$ (Scheme 3-1). With an increase in pH both the sites release two protons with similar pK_a values leading to the formation of the tris(thiolato) CdS_3 complex at both the binding sites. The first equilibrium step leading to the formation of Cd-tris(thiolato) species from $(CdS(SH)_2)^+$ in one of the sites is

represented as K_2. At this point the Cd(II) species that has UV absorbance is thus ($Cd_2S_4(SH)_2$) (Scheme 3-1) where Cd(II) in one site is present as tris(thiolato) and in the second site as mono(thiolato) species. The second step, leading to the formation of Cd-tris(thiolato) species at the second metal binding center, is represented with an equilibrium constant of K_1. At this point Cd(II) is present in both sites as the this(thiolato) species $(Cd_2S_6)^{2-}$ (Scheme 3-1). Using this model the derivation of the fitting equations for analyzing the pH titration curves for the peptides containing two metal binding sites is shown in the derivation section. The same model was applied to fit the experimental data for the peptide **GRAND**L12AL16CL26AL30CL33I.

CD Spectroscopy. All Circular Dichroism data were collected using an Aviv model 202 thermostatically controlled Circular Dichroism Spectrometer equipped with a 450 watt Xenon arc lamp. GuHCl titration experiments were carried out using a Microlab 500 series syringe pump automatic titrator controlled by Aviv software. Titrations were carried out by mixing two separate solutions of peptide containing 0.0 and 7.63 M GuHCl. Refractive index measurements were used to determine the concentration of the stock GuHCl solution.[45] Each of these solutions contained 10 μM monomer peptide and 10 mM phosphate buffer. The pH values were adjusted to 6.5 by adding KOH or HCl. Titrations were performed at 25°C in a rectangular open top quartz cell of 1 cm path length. The observed ellipticity in millidegrees were plotted vs. increasing GuHCl concentration. Whenever necessary, the data were fit by nonlinear least-square methods to a two-state trimer-monomer equilibrium[46] using the software Igor Pro (Wavemetrics, Inc). Midpoint values of the experimental curves were calculated from the concentration of GuHCl where the ellipticity is 50% from the starting ellipticity in the absence of denaturant.

^{113}Cd NMR Spectroscopy. All the spectra were collected at room temperature on a Varian Inova 500 spectrometer (110.92 MHz for ^{113}Cd) equipped with a 5 mm broadband probe. ^{113}Cd NMR spectra were externally

referenced to a 0.1 M $Cd(ClO_4)_2$ solution in D_2O. A spectral width of 847 ppm (93,897 Hz) was sampled using a 5.0 µs 90^o pulse and 0.05 second acquisition time with no delay between scans. Samples were prepared under a flow of argon by dissolving 30 – 35 mg of the lyophilized and degassed peptides in 450 - 500 µL 15% D_2O/H_2O solution. The peptide concentrations were determined by using the assay with 4,4'-dipyridyl disulphide,[42] and the concentrations range from 9 to 18 mM peptide, which corresponds to 3 – 6 mM three-stranded coiled coil. The final samples were prepared by the addition of the appropriate amount of 250 mM $^{113}Cd(NO_3)_2$ solution (prepared from 95 % isotopically enriched ^{113}CdO obtained from Oak Ridge National Laboratory) and the adjustment of the pH with KOH or HCl solutions. The pH value was measured both before and after the experiment. An argon atmosphere was maintained when possible but the samples came in contact with O_2 during addition of $^{113}Cd(NO_3)_2$, pH adjustment and acquisition. The data were analyzed using the software MestRe-C.[47] All free induction decays (FID's) were zero filled to double the original points and were processed by application of 100 Hz line broadening prior to Fourier transformation.

^{1}H NMR Spectroscopy. 2D ^{1}H NMR experiments were performed at room temperature on a 500 MHz Varian Inova spectrometer equipped with an inverse detection probe. Samples were prepared by dissolving 30-35 mg of lyophilized and degassed peptide in 500-600 uL 10% D_2O/H_2O under a flow of argon. Final solutions contained 3.0 – 4.0 mM of peptide trimer, as determined by the assay with 4,4'-dipyridyl disulphide.[42] Whenever necessary ^{113}Cd(II) was added in the form of 250 mM $^{113}Cd(NO_3)_2$ stock solution and pH was adjusted with concentrated KOH or HCl solutions as described in the previous section. Attempts were made to keep an argon atmosphere throughout, but the samples did come into contact of air during addition of $^{113}Cd(NO_3)_2$, pH adjustment and data acquisition. Nuclear Overhauser Effect Spectroscopy (NOESY) experiments were performed using standard Varian pulse sequences with mixing time of 100 ms.[48] For water suppression, pre-saturation was used for 1.5 s with saturation

power of 5 dB.[49] A spectral window of 11.2 ppm was sampled with 2048 points and 256 increments in the indirect dimension with 8 scans. Z-gradient Total Correlation Spectroscopy (ZTOCSY) experiment was performed only for the apo peptide under similar conditions with spin lock field strength of 7 kHz for the 90° pulse.[50] An 80 ms spin lock was used during the experiment. All the data were processed by MestRe Nova.[51]

^{111m}Cd Perturbed Angular Correlation (PAC) Spectroscopy. All ^{111m}Cd PAC experiments and sample preparation were performed as described previously.[44,52,53] Data collection and analysis was done by Dr. Lars Hemmingsen and coworkers at the University of Copenhagen, Denmark. The final volume of the samples ranged between 0.05 and 0.5 mL with concentrations of 250 - 300 μM peptide, 20 mM appropriate buffer (MES for pH values 6.5 and 7.0, and CHES for pH values between 8 and 9.3) and different Cd(II)/peptide ratios (Table 3-3). All fits were carried out with 300 data points, disregarding the first 5 points due to systematic errors in these. All PAC spectra were initially analyzed with four NQIs (vide infra), and NQIs were included in the fit if the amplitude was larger than two times the standard deviation. Each nuclear quadrupole interaction (NQI) was modeled using a separate set of parameters that includes ω_0, η, $\Delta\omega_0/\omega_0$, $1/\tau_c$ and A. The parameter ω_0 ($\omega_0 = 12\pi|eQVzz|/(40h)$, where Q is the nuclear electric quadrupole moment and V_{zz} is the numerically largest component of the diagonalized electric field gradient tensor) is associated with the strength of the interaction between the surrounding charge distribution and the Cd nucleus, η is the so-called asymmetry parameter which is 0 in an axially symmetric complex and has a maximal value of 1; $\Delta\omega_0/\omega_0$ describes static structural variations from one Cd(II) site to the next, and is as such a measure of the structural variability; τ_c is the rotational correlation time; and A is the amplitude of the signal.[34] The rotational correlation time was constrained to be the same for all NQIs in a given spectrum. The parameters fitted to the PAC data are presented in Table 3-3. In two cases $\Delta\omega_0/\omega_0$ was fixed to 15 %, denoted with

"d" in Table 3-3. Setting the parameter free in the fit led to very large values, and setting it to a fixed value did not alter the reduced chi-square.

Results

Cd(II) Binding to the GRAND peptides. The stoichiometry and the pH dependence of the binding of Cd(II) to the **GRAND** peptides were monitored by UV-Vis spectroscopy using the ligand-to-metal charge-transfer (LMCT) band at 235 nm corresponding to the formation of Cd-S bonds in CdS_3 or CdS_3O complexes.

a) Stoichiometry. The addition of Cd(II) aliquots into the different solutions containing 60 µM (20 µM trimer) **GRAND** peptides at pH values where, based on the pH titrations the peptides are capable of binding Cd(II) solely as CdS_3/CdS_3O, generated the appearance of the characteristic LMCT for this chromophore. Shown in Figure 3-2 are the titration curves obtained by plotting the increase in absorbance vs the equivalents of Cd(II) added per trimer of **GRAND** peptides. In all cases the curves plateau at 2 (±0.14) equivalents of Cd(II). The different observed final absorbance values reflect the higher extinction coefficient of the 3-coordinated CdS_3 species (54,772 $M^{-1}cm^{-1}$) compared to the 4-coordinate CdS_3O species (39,766 $M^{-1}cm^{-1}$) at 235 nm.[36,38]

b) pH dependence. The increase in pH of a solution containing 40 µM Cd(II) and 60 µM (20 µM trimer) **GRAND**L16PenL19IL23PenL26I enhances the 235 nm LMCT transition with Figure 3-3 showing this pH titration curve (green). The equivalent titration curve with **GRAND**L12AL16CL26AL30C is shown in blue indicating binding of Cd(II) at lower pH values. The shape of these pH titration curves is consistent with the release of four protons upon binding of two equivalents of Cd(II) and, therefore, it is in agreement with our published model for the simultaneous release of the final two protons upon Cd(II) binding to the remaining two thiolates of either Cys or Pen ($[Cd(pep)(Hpep)_2]^+$ species to give the $[Cd(pep)_3]^-$).[43,44] However, unlike the pH titration curve obtained for the

peptide **GRAND**L16PenL26AL30C (Figure 3-3, red), where two well-defined regions were observed allowing for the independent analysis of both sites,[38] only one inflexion point was obtained for the peptides **GRAND**L16PenL19IL23PenL26I and **GRAND**L12AL16CL26AL30. These results point out that in these **GRAND** peptides both sites bind Cd(II) at similar pH values. Furthermore, based on the ^{113}Cd NMR pH titrations that showed non-selective binding of Cd(II), the experimental data were fit to a model (Scheme 3-1) where binding of two equivalents of Cd(II) occurs simultaneously to both the sites. (The ^{113}Cd NMR pH titrations are described in the results section) The resulting pK_a values extracted from the fitting analysis are reported in Table 3-2.

CD Spectroscopy. The stabilities of the peptides **GRAND**L26AL30C, **GRAND**L12AL16C, **GRAND**L12AL16CL26AL30C and **GRAND**L12AL16CL26AL30CL33I were assessed from their GuHCl-induced unfolding titration curves at pH 6.5. The observed ellipticity in millidegree as a function of denaturant concentration was plotted to generate the titration curves (Figure 3-4). **GRAND**L26AL30C with a denaturation midpoint of 3.7 M (magenta) shows that this is the most stable construct followed by **GRAND**L12AL16C (midpoint 3.3 M, red), . **GRAND**L12AL16CL26AL30CL33I (midpoint 2.9 M, green) and **GRAND**L12AL16CL26AL30C (midpoint 1.7 M, blue). The shape of the denaturation titration curve for **GRAND**L12AL16CL26AL30CL33I is different from the other peptides, which suggests that the Ile group has most likely modified the trimer-monomer two state equilibrium. Fits to the experimental curves were performed as described in the Experimental Section for the peptides **GRAND**L26AL30C, **GRAND**L12AL16C and **GRAND**L12AL16CL26AL30C only (Table 3-4, Figure 3-4).

^{113}Cd NMR Spectroscopy. The ^{113}Cd NMR spectra of **GRAND**L16PenL19IL23PenL26I (Figure 3-5, green) and **GRAND**L12AL16CL26AL30 (Figure 3-5, blue) were recorded in the presence of two equivalents of $^{113}Cd(NO_3)_2$ at pH 9.5 and 8.5, respectively. Two resonances

were observed in each case with chemicals shifts of 686 and 681 ppm for **GRAND**L16PenL19IL23PenL26I, and 589 and 572 ppm for **GRAND**L12AL16CL26AL30. ^{113}Cd NMR pH titrations were carried out for both peptides. The pH of a solution containing 3.0 mM (**GRAND**L16PenL19IL23PenL26I)$_3$ and two equivalents of ^{113}Cd(NO$_3$)$_2$ was varied from pH 7.5 to 9.5 and the ^{113}Cd NMR recorded (Figure 3-6A). Three resonances were observed at pH 7.5 with chemical shifts of 681, 686 and 691 ppm. An increase in pH led to decreased intensity of the resonance at 691 ppm and an increase of the resonance at 681 ppm. The resonance at 686 ppm increased continuously. By pH 9.5, only the resonances at 681 and 686 ppm were observed. The analogous experiment with 3.0 mM (**GRAND**L12AL16CL26AL30)$_3$ and two equivalents of ^{113}Cd(NO$_3$)$_2$ was carried out between pH 6.0 and 8.5 (Figure 3-7). All the spectra showed two resonances with chemical shifts of 589 and 572 ppm. The ^{113}Cd NMR spectrum of 3.3 mM (**GRAND**L12AL16CL26AL30C)$_3$ in the presence of 1.0 eq of ^{113}Cd(NO$_3$)$_2$ at pH 6.0 showed a single resonance at 572 ppm (Figure 3-8). To determine if these peptides show site-selective binding of Cd(II), experiments were carried out where an aliquot of 1.0 equivalents of ^{113}Cd(NO$_3$)$_2$ was added to a solution containing 3.3 mM (**GRAND**L16PenL19IL23PenL26I)$_3$ at pH 9.5 or 3.0 mM (**GRAND**L12AL16CL26AL30)$_3$ at pH 8.5. Three resonances with chemical shifts of 681, 686 and 691 ppm were observed for **GRAND**L16PenL19IL23PenL26I (Figure 3-6B), similar to those observed at low pH under stoichiometric conditions. For **GRAND**L12AL16CL26AL30, the ^{113}Cd NMR spectrum (Figure 3-9) showed two peaks at 589 and 572 ppm.

^{111m}Cd PAC Spectroscopy. ^{111m}Cd PAC spectroscopy was used to determine the Cd(II) coordination geometry of the different complexes [Cd(II)(H$_2$O)](**GRAND**L26AL30C)$_3^-$, [Cd(II)]^{16}[Cd(II)(H$_2$O)]^{30}(**GRAND**L16PenL26AL30C)$_3^{2-}$, [Cd(II)(H$_2$O)]$_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ and [Cd(II)]$_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$. The ^{111m}Cd PAC spectra obtained

are shown in Figure 3-10. The PAC data were analyzed as described in the Experimental Section and the fitted parameters are reported in Table 3-3. The observed values of ω_0, giving information regarding the first coordination sphere ligands, fall into four frequency regions: 1) NQI1: high frequency signal (ω_0 ~ 0.457-0.461 rad/ns) assigned to a trigonal CdS_3 structure;[34,36] 2) NQI2: (ω_0 ~ 0.342-0.349 rad/ns) assigned to a tetrahedral CdS_3X structure (where X may be H_2O);[34,36] 3) NQI3: (ω_0 ~ 0.242-0.255 rad/ns) probably reflecting a coordination geometry with 2-3 Cys and 1-2 other N or O containing ligands; and, 4) NQI4: a low frequency signal (ω_0 ~ 0.105-0.159 rad/ns giving rise to a broad feature around 0.140-0.200 rad/ns in the spectra) which is present in almost all spectra and likely represents one or several Cd(II) non-specifically bound species coordinated to the carboxylate groups at the interface of the helices forming the coiled coils. This NQI4 represents only a minor fraction of the signal and, therefore, it is difficult to obtain a good fit of the PAC parameters. Since it is a small and common feature in the spectra, we will not comment on it further.

The PAC spectrum of **GRAND**L26AL30C with 1/12 Cd(II)/peptide ratio at pH 7.0 shows the signatures corresponding to NQI2 (ω_0 = 0.348 rad/ns) and NQI3 (ω_0 = 0.245 rad/ns). Also, a minor fraction of NQI4 was observed. A very similar spectrum was obtained for the same peptide with 1/12 Cd(II)/peptide ratio at pH 9.1 (Figure 3-10) and the fitting of the experimental data gave the same prominent NQIs (NQI2 and NQI3, Table 3-3). NQI4 was not significantly present in this dataset. The PAC spectra of **GRAND**L16PenL26AL30C with 1/12 Cd(II)/peptide ratio at pH 6.5 and 9.3 are highly similar to those observed for **GRAND**L26AL30C and can be analysed with essentially the same NQIs (NQI2, NQI3, and NQI4) with minor variations in the relative amplitudes. When at pH 9.3 the ratio of Cd(II) to peptide was increased to achieve the loading of both binding sites (1.85/3 Cd(II)/peptide ratio), the high frequency signal (NQI1) appeared in addition to the features corresponding to the three NQIs found at low Cd(II)/peptide ratio (Figure 3-10). Thus, the data indicate that the CdS_3 site is only occupied at high Cd(II)/peptide ratios. The PAC spectra obtained for the

GRANDL12AL16CL26AL30C peptide at pH 8.7 under conditions that favour loading of both binding sites (1.85/3 Cd(II)/peptide ratio) gives a PAC signal with the signature of the L26AL30C site but with a more prominent NQI2 (ω_0 = 0.342 rad/ns). The origin of this increase in amplitude for NQI2 is attributed to the binding of Cd(II) to the L12AL16C site, which is supported by the fact that Cd(II) binds to this site as a single CdS_3O species (as observed for **TRI**L12AL16C).[36] Interestingly, all peptides investigated with PAC spectroscopy in this work and containing the L26AL30C site display this signature: a mixture of NQI2, NQI3, and possibly NQI4 (minor fraction), suggesting structural variability. The PAC spectrum of the peptide **GRAND**L16PenL19IL23PenL26I recorded at pH 9.2 and high Cd(II) concentration (1.85/3 Cd(II)/peptide ratio, Figure 3-10) shows a single prominent signal in the high frequency range. Since this high frequency signal can be fitted with just one NQI, corresponding to NQI1 (trigonal CdS_3), the two binding sites in the triple helix formed by this peptide appear to have highly similar structures.

^{1}H NMR Spectroscopy. To ascertain the site specificity of Cd(II) binding to **GRAND**L12AL16CL26AL30C at pH 6.0 further, we used two-dimensional NMR spectroscopy. The backbone resonances of the peptide were assigned using NOESY and TOCSY spectra, (Figure 3-11 shows the sequential assignment of amide protons (H^N) of Glu7 to Leu33). Using this information, β methylene protons (H^β) of each Cys residues were identified. NOESY spectra were used to monitor the progress of Cd(II) titration into 3.25 mM (**GRAND**L12AL16CL26AL30C)$_3$ at pH 6.0 by observing changes in the H^N-H^β cross peak region of Cys (Figure 3-12). The horizontal axis on Figure 3-12 corresponds to the H^N and the vertical axis corresponds to the H^β of each Cys. Two non-equivalent H^β appear as separate cross peaks corresponding to each H^N. The regions **a** (7.72 ppm) and **c** (8.07 ppm) of Figure 3-12 correspond to cross peaks of H^N_{16}-H^β_{16} and H^N_{30}-H^β_{30}, respectively. Two sets of additional cross peaks at 7.57 ppm and 8.17 ppm correspond to interesidue cross peaks between H^N of Lys17 and H^β of Cys16 and between H^N of Lys31 and H^β of Cys30,

respectively. These peaks contain redundant information as regions **a** and **c** in the figure and will not be considered further. The cross peak at 7.91 ppm can be best assigned as H^N-H^γ cross peak of Glu21. After the addition of 0.5 and 1.0 eq of $^{113}Cd(NO_3)_2$ the major cross peak at 7.81 ppm appears, marked as region **b**. This new peak is broad along the F1 dimension and is an overlap of the two individual H^β resonances of Cys16 which are perturbed due to Cd(II) binding to Cys16. These results display preferential binding of Cd(II) to the Cys16 site at pH 6.0. From 1.0 to 2.0 equivalents of Cd(II), the cross peak at 8.03 ppm appears that corresponds to perturbations of the H^N-H^γ of Glu21 resonance upon addition of Cd(II). From 1.5 to 2.0 equivalents of added Cd(II), a cross peak at 7.86 ppm appears which may represent the perturbation of the Cys30 resonance due to interaction of Cd(II) with the partially deprotonated Cys30 at pH 6.0. To ascertain that region **b** in Figure 3-12 represents the cross peaks due to Cd(II) binding to the Cys16 and not to the Cys30 site, a series of pH dependent NOESY experiments were performed. The pH of a solution containing 3.21 mM (**GRAND**L12AL16CL26AL30C)$_3$ and 2.0 equivalents of Cd(II) was increased from 6.0 to 8.5 (Figure 3-13). With an increase in pH, two new cross peaks appear at 7.6 ppm (region **d**) and 7.93 ppm (region **e**). These resonances are assigned as $H^N{}_{16}$-$H^\beta{}_{16}$ for $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3{}^{2-}$ (region **d**) and $H^N{}_{30}$-$H^\beta{}_{30}$ for $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3{}^{2-}$ (region **e**). The cross peak at 7.81 ppm (region **b**) remains at high pH (8.5) indicating the presence of the species $[Cd(II)(H_2O)]^{16}[apo]^{30}$(**GRAND**L12AL16CL26AL30C)$_3{}^{4-}$ in solution. This result demonstrates that the new cross peaks are associated with Cd(II) binding to the less acidic Cys30 site and that the cross peak at region **b** in Figure 3-12 results from Cd(II) binding to the more acidic Cys16 site.

Discussion

The objective of the study presented in this chapter is to assess whether the position of metal sites along the linear sequence of the peptide influences, possibly fine-tuning, the physical properties of the bound metal ions. A second

objective is to assess whether the presence of a second bound metal influenced the properties of the first coordinated metal. To address these issues, peptides were designed that are capable of binding two equivalents of Cd(II) with identical coordination geometries. Using these peptides I evaluated the physical properties of the bound ions, the selectivity of the metals for the different sites and the influence of pH on the observed properties. While other studies have shown that *de novo* peptides can bind two metal ions in separate mononuclear sites[54,55] or in binuclear centers,[56-58] they did not address the question of selective binding of the same metal ion in different centers having distinct geometries. The model peptides used to test these issues are **GRAND**L12AL16CL26AL30C, designed to have two 4-coordinate sites (CdS_3O) and **GRAND**L16PenL19IL23PenL26I with two 3-coordinate sites (CdS_3). Since these sites are located at different positions in the coiled coil, the final Cd(II) constructs are suitable for evaluating how metal positioning influences their physical properties. Thus, one may assess whether or not Cd(II) displays site-selective binding, whether the binding of one metal in a helix alters the properties of the second metal and whether the spectral properties of a bound Cd(II) are dependent on the proximity to the center or ends of a helical bundle. These important issues have yet to be addressed with designed peptides.

A) Designing the Desired Constructs Containing two Metal Binding Sites

a) Preparing dual 3-coordinate sites

The first challenge was designing the sulfur substituted **GRAND** peptides. The linear correlation between the ^{113}Cd NMR and ^{111m}Cd PAC spectroscopy[44] was used to guide the design of the **GRAND** peptides. The first goal was the preparation of the peptide containing two 3-coordinate sites. Based on previous studies from the group, it was known that positions near the edge of the coiled coil should be avoided in order to obtain a 3-coordinate site.[35,36,38] These positions are prone to coordinate water due to the fraying of the α-helices and thus, both binding sites should be buried in the interior of the coiled coil. The

starting point for the design was the peptide **GRAND**L16PenL26AL30C, known to contain one 3-coordinate site and one 4-coordinate site (Figure 3-5, red).[38] The first strategy was to keep the Pen residue in position 16, since this location is known to generate a 3-coordinate site,[36] and introduce the second Pen site as far as possible from the C-terminal edge to avoid fraying, but separate enough from the first one to avoid the disruption of this site. Based on these requisites, the second Pen could be introduced into position 23 (fourth heptad); however, making a peptide containing two separate binding sites that can encapsulate two Cd(II) ions as 3-coordinate complexes at both the binding sites was not so straightforward as simply adding a Pen to an interior site as a construct previously synthesized in our group, **GRAND**L12AL16CL23Pen (unpublished results) bound Cd(II) in the position 23 as a mixture of species (55% CdS_3O and 45% CdS_3) with a chemical shift of 640 ppm. Since moving the second Pen site towards the interior of the coiled coil could disrupt the first binding site, we decided to substitute Leu 26 for an Ile. Ile is known to confer better packing and should prevent fraying by generating more compact α-helices. The ^{113}Cd NMR spectra of the resulting construct, **GRAND**L16PenL23PenL26I, loaded with 2 equivalents of Cd(II) at pH 9.6 showed two distinct peaks, with chemical shifts of 678 and 666 ppm. Although these chemical shifts could not be assigned to a particular site (16 vs. 23), these values revealed the disruption of the first Pen site (L16Pen) by the introduction of the second site (L23Pen) since a ^{113}Cd chemical shift of 685 ppm was observed for the original fully 3-coordinate 16Pen site.[36,38] While the chemical shifts obtained represented an improvement towards the achievement of fully 3-coordinated sites, they indicate an equilibrium between CdS_3 and CdS_3O (678 ppm ~ 80:20 and 666 ppm ~ 71:29, Table 3-5). Here the ^{113}Cd NMR shifts are used to assess the degree of formation of the CdS_3, but this method could provide a slight underestimate of the amount present as there is not a 100% correspondence between the ^{113}Cd NMR shifts and the ^{111m}Cd PAC assessments of the CdS_3/CdS_3O ratio. Quantum chemical calculations show that a change in Cd-S bond length of 0.01 Å can cause a change in chemical shift of ~20 ppm (Hemmingsen *et.al.* unpublished results). Thus, the

PAC data often indicate a higher percentage of CdS_3 in these samples than one would predict using the strict NMR correlation.

The existence of a single Leu layer (at position 19) between these two sites seemed to be insufficient to retain compact packing and avoid perturbation of the 16Pen site due to the presence of the 23Pen site. To improve these results and minimize the disruption of the 16Pen site, the Leu layer just below (Leu19) was replaced by Ile. This generates an even more compact packing of the hydrophobic core. Indeed, this strategy succeeded as **GRAND**L16PenL19IL23PenL26I bound 2 (±0.14) equivalents of Cd(II) with trigonal planar CdS_3 geometries (686 (CdS_3:CdS_3O ~ 87:13, Table 3-5) and 681 (CdS_3:CdS_3O ~ 83:17, Table 3-5) ppm) as shown in Figure 3-5 (green).

b) **Preparing dual 4-coordinate sites**

Having been successful in preparing constructs capable of binding two Cd(II) ions as 3-coordinate CdS_3 geometry at both the sites the next challenge was to design a construct that can coordinate two Cd(II) ions as a 4-corodinate CdS_3O geometry. The starting point for designing the peptide containing two 4-coordinate sites was again the construct **GRAND**L16PenL26AL30C.[38] In this case, the segment containing the 4-coordinate site (L26AL30C) was kept unaltered and modified the L16Pen center, known to generate a 3-coordinate site,[36] to an L12AL16C center to obtain the second 4-coordinate site. Previous work with **TRI**L12AL16C suggested that this would be a successful approach.[36] This design yields the final peptide **GRAND**L12AL16CL26AL30C. The binding of Cd(II) was characterized by ^{113}Cd NMR spectroscopy and the spectrum of the fully loaded $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ species showed two resonances at 589 (CdS_3:CdS_3O ~ 8:92, Table 3-5) and 572 (CdS_3:CdS_3O ~ 0:100, Table 3-5) ppm, respectively. These chemical shifts are consistent with those expected for a CdS_3O environment,[44] demonstrating that the successful design of a single peptide with two 4-coordinate centers was achieved.

B) Physical Properties of the GRANDL16PenL19IL23PenL26I Peptide

With the two desired **GRAND** peptides in hand, I next investigated the impact of the position of the metal centers on the intrinsic physical properties of the metal ion. Titration of Cd(II) into a solution containing 60 μM (20 μM trimer) **GRAND**L16PenL19IL23PenL26I at pH 9.5 was accompanied by an absorption increase centered at 235 nm and a final extinction coefficient of 54,772 $M^{-1}cm^{-1}$, which is consistent with the formation of two Cd(II) tris thiolate chromophores. The titration curve shown in Figure 3-2 (green color) indicates that the stoichiometry of peptide to Cd(II) is 3 : 2 (±0.14) and, therefore, the best description for the Cd(II) complex formed under these conditions is $[Cd(II)]_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$. As discussed above, based on the ^{113}Cd NMR spectrum (Figure 3-5, green) these two Cd(II) sites are trigonal CdS_3. Further support for this assignment was obtained from the ^{111m}Cd PAC spectrum of $[Cd(II)]_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$ (Figure 3-10F). This dicadmium peptide has a single high frequency signal (NQI1) with parameters ω_0 = 0.461(1) rad/ns and η = 0.02(4) (Table 3-3) that correspond to a trigonal planar CdS_3 geometry with highly similar structures, indicating that the two sites created in (**GRAND**L16PenL19IL23PenL26I)$_3$ yield two metal ion sites with no indication of CdS_3O. The parent **TRI**L16Pen peptide, the first peptide shown to bind Cd(II) exclusively as a 3-corodinate CdS_3 species had similar PAC parameters [ω_0 = 0.4540(9), η = 0.02(10)].[36] Thus, the assignment that Cd(II) in both sites of the **GRAND**L16PenL19IL23PenL26I peptide is bound exclusively as 3-coordinate CdS_3 geometry is consistent with what has been previously observed.

Previous studies have shown that it is possible to discriminate CdS_3 and CdS_3O sites in **TRI** and **GRAND** peptides by examining the pH profiles for the formation of the tris cysteine complexes. In these cases the relevant equilibrium is: $[Cd(pep)(Hpep)_2]^+ \rightarrow [Cd(pep)_3]^- + 2H^+$. It was found that CdS_3 complexes (using Pen or Cys) have effective pK_{a2} values that are far more basic than the corresponding CdS_3O structures (e.g., **TRI**L16Pen, pK_{a2}=15.8 vs. **TRI**L12AL16C,

pK_{a2}=12.2).[44] The pH profile for the binding of 2 equivalents of Cd(II) to the **GRAND**L16PenL19IL23PenL26I peptide (Figure 3-3) is consistent with these results since the pK_{a2} values obtained, 15.2 and 15.9, are very similar, and quite basic, as is expected for 3-coordinate sites. Based on previous observations, we infer that the site with the ^{113}Cd NMR shift of 686 ppm has the more basic of these pK_{a2} values as it has virtually no CdS_3O present. All of these data demonstrate that we have prepared a single peptide capable of binding two equivalents of Cd(II) with trigonal planar CdS_3 coordination. The fact that the physical properties observed for both sites are comparable indicates that the physical properties of the 3-coordinate sites are not dependent on the location of the sites within the coiled coil. Furthermore, at this point it appears that in a well packed helix, the presence of the second CdS_3 center has minimal impact on the starting peptide **GRAND**L16Pen. This latter point will now be addressed in more detail.

C) Does the Metallation State of One Site Perturb the Other Site in **GRAND**L16PenL19IL23PenL26I?

Now I will focus on the two related experiments performed with the **GRAND**L16PenL19IL23PenL26I peptide as shown in Figure 3-6. In the following section I will discuss the possible interpretations of the data presented in Figure 3-6A, and 3-6B in a systematic way. As the assignment of the individual sites is unknown, I will refer to the site with chemical shift of 686 ppm as site A and the site with chemical shift of 681 ppm as site B. In the case of the pH dependent ^{113}Cd NMR measurements presented in Figure 3-6A, the peptide is loaded with two equivalents of Cd(II) so that both sites are filled with the metal. It has been shown previously that the binding model for Cd(II) to these systems is that the metal first coordinates to the site of interest forming a $[Cd(pep)(Hpep)_2]^+$ species at low pH. As the pH is raised, two protons are released simultaneously to give the tris thiolato complex $[Cd(pep)_3]^-$. Thus, the two relevant peptide species under low pH conditions are $[Cd(pep)(Hpep)_2]^+$ and $[Cd(pep)_3]^-$, while no apo

peptide is present. Figure 3-6A shows that at low pH the 691 ppm peak dominates over the 681 ppm peak. With an increase in pH, the 681 ppm peak gains intensity over the 691 ppm peak and at pH 9.5 only the 681 (site B) and 686 (site A) ppm peaks are present. The 686 ppm peak gained intensity continuously during the course of the pH titration. The conversion from the 691 to the 681 ppm peak follows closely to the pK_{a2} determined for this peptide using UV-Vis spectroscopy (Table 3-2).

One interpretation of the data in Figure 3-6A suggests that the 691 ppm peak may arise from site B (681 ppm) when Cd(II) in site B is present as $[Cd(pep)(Hpep)_2]^+$ and disappears when all the Cd(II) is present as only $[Cd(pep)_3]^-$ at site B. However, this model does not withstand the scrutiny of the data in Figure 3-6B. In this experiment all spectra are collected at pH 9.5 which are conditions supporting only the $[Cd(pep)_3]^-$ metal structure. In the presence of only 1.0 equivalent of Cd(II), the simultaneous presence of both the peaks at 681 and 691 ppm is observed. This observation suggests that the 691 ppm peak exists when the $[Cd(pep)(Hpep)_2]^+$ species at site B is absent. Therefore, considering the data in Figure 3-6A and 3-6B together, one can conclude that the 691 ppm peak cannot be an indicator of the presence of $[Cd(pep)(Hpep)_2]^+$ species at site B (681 ppm). Furthermore, the ^{113}Cd NMR resonance for $[Cd(pep)(Hpep)_2]^+$ has not been observed for a **TRI** peptide either as a dual site or mono site derivative,[3] and if observable, would be predicted to be far up field of the $[Cd(pep)_3]^-$ chemical shift region. Thus, a monothiolate Cd(II) species as the origin of the 691 resonance can be eliminated as this peak must be due to a tris thiolato Cd(II) complex.

A second possible interpretation of these data is that the 691 ppm peak is an alternative form of the 686 ppm peak which is a consequence of the partial occupancy of site B that gives rise to the 681 ppm peak. Data in Figure 3-6B are consistent with this idea as it shows that in the presence of 1 equivalent of $^{113}Cd(II)$ at pH 9.5 (where all the Cd(II) is present as $[Cd(pep)_3]^-$) indiscriminate

binding is observed to both the Pen sites (681, 686 and 691 ppm). Subsequent changes upon addition of the second equivalent of the metal are also consistent with this interpretation. However, reexamination of Figure 3-6A, which is collected under conditions where peptide is loaded with 2 equivalents of ^{113}Cd(II) for all the spectra, shows that the 691 peak appears at pH 7.5 through 8.2 even though site B (681 ppm) is fully occupied with Cd(II). Therefore, it can be concluded that the 691 ppm peak cannot be an alternate form of the 686 ppm peak when site B is unoccupied because in these experiments (Figure 3-6A) site B is always filled.

Thus, the model that best explains both the data in Figure 3-6A and 3-6B is that the 691 ppm peak is a form of the 681 ppm peak and exists when site A (686 ppm) is either unfilled (metal titration, apo) or present with only one deprotonated sulfur (pH titration, $[Cd(pep)(Hpep)_2]^+$). Furthermore, the 681 ppm feature only occurs when site A has a fully formed trigonal Cd(II) structure ($[Cd(pep)_3]^-$). Of course, the 681 and 686 ppm signals gain intensity as the pH is raised (Figure 3-6A) as a higher proportion of $[Cd(pep)_3]^-$ is formed in both the sites; however, the 681 specie dominates over the specie giving rise to the 691 ppm signal as all of the Cd(II) in site A becomes $[Cd(pep)_3]^-$ with increasing pH. Thus, the critical factor controlling the 10 ppm upfield shift of the ^{113}Cd NMR resonance is whether the final two thiolate sulfur atoms insert into the coordination sphere of the metal in site A. Hence, the 691 signal can appear either because an apo structure (present at high pH under substoichiometric conditions, Figure 3-6B) or a $[Cd(pep)(Hpep)_2]^+$ (present at lower pH under stoichiometric conditions, Figure 3-6A) structure occurring at site A (686 ppm). Recognizing this duality, it becomes clear that the most likely reason for the switch between the 691 and 681 resonances is the rearrangement of the Pen side chains and subsequent structural/conformational change on going from an apo or partially bound $[Cd(pep)(Hpep)_2]^+$ form to the final $[Cd(pep)_3]^-$ structure. Such an explanation is consistent with X-ray structures of related systems that have recently appeared.[30,32,59] It can be concluded from the discussion of this

section that the two sites in **GRAND**L16PenL19IL23PenL26I peptide are not completely independent of each other as Cd(II) bound to one of the sites (site B, 681 ppm) is perturbed by the metallation state (apo or $[Cd(pep)(Hpep)_2]^+$ or $[Cd(pep)_3]^-$) of the second center (site A, 686 ppm). Site A, on the other hand, seems to be insensitive to the metallation state of site B, based on the data presented in Figure 3-6A, and 3-6B. Of course one can hypothesize that site A is also being perturbed by the metallation state of site B and the resulting ^{113}Cd NMR signal(s) may be hidden under other peaks. However, based on the present data no definite conclusion can be drawn regarding this possibility. Table 3-6 summarizes these conclusions derived from Figures 3-6A and 3-6B. Taking into consideration the previous designs of disubstituted peptides by the group,[37,38,54] two layers of intervening Leu seem to be necessary to keep both metal centers completely independent. In the peptide **GRAND**L16PenL19IL23PenL26I there is only one intervening layer, Ile in this specific case, between both Cd(II) sites.

D) Characterization of the **GRAND**L12AL16CL26AL30C Peptide: Different Physical Properties for the Two Sites

With the **GRAND**L16PenL19IL23PenL26I system well understood, now I will discuss the results obtained for the **GRAND**L12AL16CL26AL30C peptide that should encapsulate two CdS_3O centers. The titration curve (Figure 3-2, blue color) following the binding of Cd(II) to **GRAND**L12AL16CL26AL30C is consistent with the complexation of 2 ($\pm$0.14) equivalents of Cd(II) per trimer. The best description for the Cd(II) complex in this case is $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$. This fully formed Cd(II) species has an extinction coefficient of 39,766 $M^{-1}cm^{-1}$ at 235 nm, which is lower than those of $[Cd(II)]_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$ (Figure 3-2, green color) and $[Cd(II)]^{16}[Cd(II)(H_2O)]^{30}$(**GRAND**L16PenL26AL30C)$_3^{2-}$ (Figure 3-2, red color) and, therefore, consistent with having two 4-coordinate sites.[36,38] Furthermore, this value of extinction coefficient is almost twice than that of $[Cd(II)(H_2O)]^{16}$(**TRI**L12AL16C)$_3^-$ ($\Delta\varepsilon$ = 21, 200 $M^{-1}cm^{-1}$).[36] To verify that both Cd(II)

ions bind to the peptide **GRAND**L12AL16CL26AL30C as a 4-coordinate CdS_3O species, ^{111m}Cd PAC spectroscopic studies were carried out. The ^{111m}Cd PAC spectrum obtained under conditions that favored full loading of the peptide Cd(II), showed two different NQIs, namely NQI2 and NQI3, in addition to NQI4 that is present in almost all the spectra (Figure 3-10G). This result indicates the presence of different Cd(II) species on the ^{111m}Cd PAC timescale. The parameters for NQI2 (ω_0 = 0.342(6) rad/ns and η = 0.148(8), Table 3-3) are very similar to those observed for **TRI**L12AL16C and has been assigned to a distorted tetrahedral CdS_3O species where the metal ion is most likely coordinated above the plane of the β-methylene protons of the Cys in an *exo* configuration, with an exogenous water molecule as a fourth ligand towards the N-terminus.[36,44] The lower frequency NQI3 signal (ω_0 = 0.242(8) rad/ns and η = 0(1), Table 3-3) has parameters similar to those reported for the peptide **TRI**L16CL19A which presents a water-binding pocket below the metal binding site.[44] This peptide contains a 100% CdS_3O species and the best interpretation for this result is that Cd(II) is bound below the thiolate plane in an *endo* configuration with the water towards the C-terminus. This coordination geometry would resemble the one observed in the X-ray structure of As(**CS**L9C)$_3$,[30] (Coil-Ser (CS) is a crystallizable analogue of **TRI** family of peptides)[60] where As(III) is bound to the three Cys sulfurs below the plane defined by their β-methylene protons and the lone pair pointing towards the C-terminus. For Cd(II), the water molecule, as the fourth coordinating ligand, replaces the lone pair of electrons. The *endo* configuration could be favored for **TRI**L16CL19A because the generated hole lies just below the Cys. In contrast, the designed cavities for **GRAND**L12AL16CL26AL30C are located above the Cys sulfur plane; however, the Cys30 is located far towards the C-terminus of the peptide where helical fraying of the coiled coil is significant. This fraying should generate a more promiscuous binding center that would allow the water to enter from the C-terminal direction and coordinate to Cd(II). The NQI3 signal has never been observed for peptides containing the Cys at position 16 and which are capable of binding Cd(II) as a 100% CdS_3O species with an *exo* configuration (**TRI**L12AL16C, **TRI**L12AL16Pen and **TRI**L9AL16C).[44] This

observation suggests that in **GRAND**L12AL16CL26AL30C, NQI3 corresponds to the C-terminal binding site of L26AL30C. The corresponding monosubstituted peptide, **GRAND**L26AL30C, showed essentially the same two NQIs (NQI2 and NQI3) (Figure 3-10B). Indeed, the amplitude corresponding to the NQI2 signal was smaller than that obtained for **GRAND**L12AL16CL26AL30C as would be expected since the peptide **GRAND**L26AL30C is missing the L12AL16C center (Table 3-3). Furthermore, the heterochromic peptide **GRAND**L16PenL26AL30C showed the same three NQI signatures which were observed under conditions where the peptide was in excess compared to Cd(II) (ratio peptide/Cd(II) = 12/1). This heterochromic peptide is known to bind Cd(II) in the L26AL30C center selectively both at low and high pH values (Figure 3-10C and 3-10D, Table 3-3).[38] This result supports the assignment that the L26AL30C center binds Cd(II) as a mixture of CdS_3O species, NQI2 and NQI3, best interpreted as corresponding to an *exo* and an *endo* configuration, respectively.

Once verified that the peptide **GRAND**L12AL16CL26AL30C binds 2 ($\pm$0.14) equivalents of Cd(II) as 4-coordinate CdS_3O structures, the pH dependence associated with Cd(II) binding to this binding was next investigated. The obtained pH titration curve (Figure 3-2, color blue) reveals that the second and third thiolates of this peptide bind to Cd(II) at lower pH values than those of the peptide **GRAND**L16PenL19IL23PenL26I. This observation is consistent with having 4-coordinate sites;[44] however, the pK_{a2} values obtained, 11.8 and 13.9, are not as similar to each other as those observed for Cd(II) binding to **GRAND**L16PenL19IL23PenL26I (15.2 and 15.9). This is an interesting result since both sites are defined by the same type of residues and Cd(II) binds to both as CdS_3O species, as shown by ^{113}Cd NMR and ^{111m}Cd PAC spectroscopy. More importantly, these values are different from those obtained for the coordination of Cd(II) to the peptides **GRAND**L12AL16C (pK_{a2} = 11.3)[34] and **GRAND**L26AL30C (pK_{a2} = 9.9)[38,44], each containing only one binding site (Figure 3-3). Indeed, the simulated pH profile for two equivalents of Cd(II) binding to **GRAND**L12AL16CL26AL30C with the pK_{a2} (11.3 and 9.9) and molar absorption

coefficient values (19247, 18705 M^{-1} cm^{-1} for L12AL16C and L26AL30C sites, respectively) of the individual sites is quite different from the experimentally observed pH profile (Figure 3-14). Upon closer examination of these data, one realizes that one of the pK_{a2} values (11.8) is close to the pK_{a2} value of Cd(II) binding to **GRAND**L12AL16C (11.3). Assigning the value of 11.8 to the pK_{a2} of Cd(II) binding to the site L12AL16C will imply that Cd(II) coordinates to the L26AL30C center with a pK_{a2} of 13.9. This represents a difference of 4 log units between having this binding site in **GRAND**L12AL16CL26AL30C vs. **GRAND**L26AL30C and **GRAND**L16PenL26AL30C. The peptide **GRAND**L16PenL26AL30C is known to coordinate Cd(II) in the 4-coordinate site with a pK_{a2} of 9.6.[38]

E) Assignment of Individual pK_{a2} Values to the Corresponding Sites in the **GRAND**L12AL16CL26AL30C Peptide

At this point, it became essential to use NMR spectroscopy to assign the binding sites corresponding to these different pK_{a2} values. First ^{113}Cd NMR pH titrations were performed. Two important observations can be inferred from the spectra shown in Figure 3-7. First, that the binding site with a ^{113}Cd NMR chemical shift of 572 ppm binds Cd(II) at lower pH values than the site with a ^{113}Cd chemical shift of 589 ppm. Second, in contrast to **GRAND**L16PenL19IL23PenL26I (Figure 3-6A), the two binding centers in **GRAND**L12AL16CL26AL30C are independent of each other since additional ^{113}Cd chemical shifts were not observed for the half-loaded peptides. In this peptide, the two Cys sites are separated by two Leu layers (19 and 23), in addition to the Ala 26 layer, and this can explain this differential behavior. Comparing the resonances at 572 ppm and 589 ppm with those observed for the complexes [Cd(II)(H_2O)](**GRAND**L12AL16C)$_3^-$ (572 ppm)[44] and [Cd(II)(H_2O)](**GRAND**L26AL30C)$_3^-$ (588 ppm)[38], one is tempted to assign the Cd(II) binding pK_{a2} values of 11.9 and 13.8 to L12AL16C and L26AL30C sites of the peptide **GRAND**L12AL16CL26AL30C, respectively.

To confirm these assignments, the ^{1}H chemical shifts of both Cys sites (Cys 16 and Cys 30) were monitored upon addition of Cd(II). Two-dimensional NOESY and TOCSY ^{1}H NMR spectroscopies were employed to assign the backbone resonances of the peptide (Figure 3-11). This information was used to identify the resonances of the *β*-methylene protons (H^{β}) of each Cys. Subsequently, ^{1}H-^{1}H NOESY experiments were carried out to monitor the binding of Cd(II) to Cys 16 and Cys 30 by observing the effects on resonances of the H^{β} protons. If the assignments based on ^{113}Cd NMR spectroscopy are correct, the sequential additions of 0.5 equivalents of $^{113}Cd(NO_3)_2$ up to a total of 2 equivalents at pH 6.0 should cause changes mainly in the region of the spectrum corresponding to the H^{N}-H^{β} cross peak of Cys 16. Figure 3-12 demonstrates that this indeed happens. A new major cross peak appears at 7.81 ppm (region **b**) that corresponds to the perturbation of the two individual H^{β} resonances of Cys16 upon coordination of Cd(II) to this residue. These results indicate that at pH 6.0 Cd(II) binds primarily to Cys 16 forming the species $[Cd(II)(H_2O)]^{16}[apo]^{30}$(**GRAND**L12A L16CL26AL30C)$_3^{-}$.

This assignment was further validated by recording a ^{113}Cd NMR spectrum of a solution containing 3.3 mM (**GRAND**L12AL16CL26AL30C)$_3$ and 1.0 equivalent of $^{113}Cd(NO_3)_2$ at pH 6.0. This spectrum shows a single resonance at 572 ppm (Figure 3-8) that corresponds to Cd(II) bound to the L12AL16C site. Next, pH dependent NOESY experiments were performed in the presence of two equivalents of $^{113}Cd(NO_3)_2$ (Figure 3-13) and used this information for the assignment. An increase in pH of a solution containing 3.21 mM (**GRAND**L12AL16CL26AL30C)$_3$ and two equivalents of $^{113}Cd(NO_3)_2$ from 6.0 to 8.5 (Figure 3-13) generated the appearance of two new cross peaks at 7.6 ppm (region **d**) and 7.93 ppm (region **e**) that were assigned to the $H^{N}{}_{16}$-$H^{\beta}{}_{16}$ and $H^{N}{}_{30}$-$H^{\beta}{}_{30}$ of the new species $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$, respectively. At pH 8.5, where full binding is expected based on the UV-Vis pH titrations, the complete disappearance of the peaks is observed at region **a** and **c**, corresponding to the unmetallated peptide **GRAND**L12AL16CL26AL30C, and the

complete development of the cross peaks at regions **d** and **e** that correspond to the fully metallated species $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$. Interestingly, the cross peak at 7.81 ppm (region **b**) was still present at this high pH values, suggesting that the species $[Cd(II)(H_2O)]^{16}[apo]^{30}$(**GRAND**L12AL16CL26AL30C)$_3^{4-}$ exists in solution at a pH value (8.5) where the peptide is expected to be fully loaded based on the UV-Vis pH titration. The best interpretation for this result is the existence of a dynamic equilibrium regarding the coordination of Cd(II) to the L26AL30C site. This equilibrium between the species $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ and $[Cd(II)(H_2O)]^{16}[apo]^{30}$(**GRAND**L12AL16CL26AL30C)$_3^{4-}$ occurs on the msec time scale. However, the ^{113}Cd chemical shift for the L12AL16C site is the same irrespective of the presence or absence of metal at the L26AL30C site as revealed by the analogous ^{113}Cd NMR experiments (Figure 3-7). As NOEs are observed through space and not through bonds, it is possible that the chemical environment of H_{16}^{N} and H_{16}^{β} of Cys 16 can sense the perturbation of the H_{30}^{N} and H_{30}^{β} of Cys 30 due to the presence or absence of Cd(II) at the L26AL30C site. This perturbation of H_{30}^{N} and H_{30}^{β} at the Cys 30 site can lead to the appearance of different cross peaks of H_{16}^{N} and H_{16}^{β} (Figure 3-13) depending on whether Cd(II) is present at the Cys 30 site or not. In contrast, for the ^{113}Cd NMR experiments we are monitoring the CdS_3O chromophore directly. In this situation, the CdS_3O complex at the L12AL16C site does not have a direct influence on the spectra arising from Cd(II) at the L26AL30C site.

The combination of the UV-Vis, ^{113}Cd NMR and ^{1}H NMR spectroscopic data supports the conclusion that Cd(II) binds to the L12AL16C site with a pK_{a2} value of 11.8 and to the L26AL30C site with a pK_{a2} value of 13.9. In addition, these results have revealed a dynamic behavior of the L26AL30C site. The origin of this dynamic nature is not yet completely clear. However, the close proximity of this site to the edge of the α-helix, where fraying of the helices in the coiled coil is an important factor, could promote this situation. Indeed, ^{111m}Cd PAC spectroscopy shows how all the peptides containing the L26AL30C site display a

mixture of NQI2, NQI3, and NQI4 (minor fraction), suggesting structural variability, whereas the L12AL16C only displays the NQI2. Despite this different behavior, the ^{113}Cd NMR spectrum (Figure 3-9) observed after addition of one equivalent of $^{113}Cd(NO_3)_2$ to a solution containing (**GRAND**L12AL16CL26AL30C)$_3$ at pH 8.5 shows Cd(II) coordination to both sites. Thus, Cd(II) does not show selectivity for either site at pH values where CdS_3O should fully exist in this peptide, as was observed for the CdS_3 centers in $[Cd(II)]_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$.

F) Understanding the Possible Reasons for pK_{a2} Differences in the **GRAND**L12AL16CL26AL30C Peptide

The pK_{a2} value of $[Cd(II)(H_2O)]$(**GRAND**L12AL16C)$_3^-$ is 11.3 and the L12AL16C site of $[Cd(II)(H_2O)]_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ is 11.8. The shift between these two values is relatively minor and requires little additional comment. However, the nearly 4 log unit higher pK_{a2} of the L26AL30C site (13.9) in the **GRAND**L12AL16CL26AL30C peptide as compared to $[Cd(II)(H_2O)]$(**GRAND**L26AL30C)$_3^-$ (9.9) deserves further consideration. **GRAND** peptides, by virtue of the additional heptad, are significantly more stable than the **TRI** peptides. Each substitution of Cys or Ala for Leu causes a marked disruption of this stabilizing hydrophobic layer in either series. The **GRAND** peptides have 10 Leu layers (Table 3-1); however, in **GRAND**L12AL16CL26AL30C, four of these Leu layers have been modified causing as much as a 16-20 kcal/mol destabilization from the parent **GRAND** three stranded coiled coil.[27] Similarly, an 8-10 kcal/mol difference occurs between **GRAND**L12AL16CL26AL30C and a single site analogue (e.g., **GRAND**L12AL16C). Because the packing around the centrally located metal site L12AL16C should be only slightly destabilized by this overall decrease in free energy of association by adding the L26AL30C site, a small shift in pK_{a2} from 11.3 to 11.8 is reasonable. However, the L26AL30C is very close to the C-terminus, which already is expected to suffer from significant fraying.[30,32,59,60] The entire coiled coil is destabilized going from **GRAND**L26AL30C to **GRAND**L12AL16CL26AL30C, especially by disrupting the center of the

assembly. One might expect that this alteration could significantly perturb metal binding in the L26AL30C position, which is the inherently less stable site. One can envision that making a more stable aggregate by conferring improved packing at the C-terminus end of the **GRAND**L12AL16CL26AL30C peptide would shift the pK_{a2} values back towards more acidic range. To test this hypothesis we designed the construct **GRAND**L12AL16CL26AL30CL33I where Leu 33 at the C-terminus end has been replaced by Ile which is known to render better packing. The resulting pK_{a2} values for this new construct are 11.6 and 13.5, respectively (Table 3-2, Figure 3-3 purple). The shift in pK_{a2} for the two sites of this new peptide is minor (0.2 and 0.4) from the parent **GRAND**L12AL16CL26AL30C peptide and is within the error of measurements. To understand the intricate interplay between peptide stabilities and the resulting pK_{a2} values, I analyzed the stability of the unmetallated **GRAND**L12AL16C, **GRAND**L26AL30C, **GRAND**L12AL16CL26AL30C and **GRAND**L12AL16CL26AL30CL33I coiled coils using circular dichroism spectroscopy and guanidinium hydrochloride (GuHCl) denaturation titrations (Figure 3-4).

These experiments illustrate a significant decrease in stability of the **GRAND**L12AL16CL26AL30C construct as compared to the single site peptides **GRAND**L12AL16C, and **GRAND**L26AL30C. Because of only a minor shift in the pK_{a2} value of **GRAND**L12AL16CL26AL30CL33I from **GRAND**L12AL16CL26AL30C and due to the fact that the stability of the **GRAND**L12AL16CL26AL30CL33I construct could not be quantified, the relationship between pK_{a2} and aggregate stability could not be established. Understanding the intricate interplay between coiled coil stability and pK_{a2} would require future studies with many more protein constructs. However, one can qualitatively say that the explanation for the 4 unit shift in pK_{a2} between **GRAND**L26AL30C and **GRAND**L12AL16CL26AL30C is, in part due to the general destabilization of the coiled coil by the addition of the second metal binding site at the middle of the peptide scaffold as it resulted in a significant destabilization of the coiled coil as observed by GuHCl denaturation studies. As a final note, preparing the dual CdS_3 construct

GRANDL16PenL19IL23PenL26I from the parent **GRAND**L16Pen peptide is more tolerant to mutations because in this case, all of the residues replacing Leu (Pen or Ile) confer greater stability to the aggregate; however, one should recognize that to achieve such a dual 3-coordinate peptide still required the addition of a heptad (**TRI** to **GRAND**) and the incorporation of Ile for Leu.

G) Difference in Exchange Behavior of the Two Sites in **GRAND**L12AL16CL26AL30C

A separate study focused on understanding the exchange behavior of Cd(II) into the **GRAND**L12AL16CL26AL30C peptide reveals differences in the exchange properties of the two sites in this peptide. While exchange of Cd(II) is seen to be occurring on the slow exchange regime, the L26AL30C is more susceptible of chemical exchange compared to the L12AL16C site. From NMR line fitting analysis the exchange rates of Cd(II) have been determined to be 5.8 $\pm$ 0.96 ms, and 7.1 $\pm$ 1.5 ms, for the L26AL30C, and L12AL16C sites, in the **GRAND**L12AL16CL26AL30C peptide. On the other hand, the corresponding single site peptide **GRAND**L26AL30C has an exchange rate of 7.9 $\pm$ 1 ms, which is slower than the same site (L26AL30C) in the dual site peptide. This result demonstrates that the introduction of a second metal binding site (L12AL16C), in the middle of the peptide **GRAND**L26AL30C, has fine-tuned the exchange behavior of the L26AL30C site located ~20 Å away in the dual site peptide **GRAND**L12AL16CL26AL30C. Thus introduction of a second metal binding site in the **GRAND**L26AL30C peptide has fine tuned both the thermodynamic (different pk_{a2} values for Cd(II) binding) as well as kinetic properties of the same site in the dual site peptide. Detailed description of the exchange properties of these peptides is presented in chapter 4 of this thesis.

Conclusions

In this chapter I report how *de novo* designed peptides can be used to investigate whether the position of a metal site along a linear sequence that folds into a three stranded α-helical coiled coil defines the physical properties of Cd(II) ions in either CdS_3 or CdS_3O (O-being an exogenous water molecule) coordination environments. Peptides are presented that bind Cd(II) into two identical coordination sites that are located at different topological positions at the interior of these constructs. The peptide **GRAND**L16PenL19IL23PenL26I binds two Cd(II) as trigonal planar 3-coordinate CdS_3 structures whereas **GRAND**L12AL16CL26AL30C sequesters two Cd(II) as pseudotetrahedral 4-coordinate CdS_3O structures. The results obtained from this study demonstrate how for the first peptide, having a more rigid structure, the location of the identical binding sites along the linear sequence does not affect the physical properties of the two bound Cd(II). However, the two sites are not completely independent of each other as Cd(II) bound to one of the sites (^{113}Cd NMR chemical shift of 681 ppm) is perturbed by the metallation state (apo or $[Cd(pep)(Hpep)_2]^+$ or $[Cd(pep)_3]^-$) of the second center (^{113}Cd NMR chemical shift of 686 ppm). **GRAND**L12AL16CL26AL30C shows a completely different behavior. The physical properties of the two bound Cd(II) ions indeed depend on the position of the metal center, having pK_{a2} values for the equilibrium $[Cd(pep)(Hpep)_2]^+ \rightarrow [Cd(pep)_3]^- + 2H^+$ (corresponding to deprotonation and coordination of cysteine thiols) that range from 9.9 to 13.9. In addition, the L26AL30C site shows dynamic behavior, which is not observed for the L12AL16C site. These results indicate that for these systems one cannot simply assign a “4-coordinate structure” and assume certain physical properties for that site since important factors such as packing of the adjacent Leu, size of the intended cavity (*endo* vs *exo*) and location of the metal site play crucial roles in determining the final properties of the bound Cd(II).

References:

(1) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779-819.
(2) Gibney, B.; Rabanal, F.; Dutton, P. *Current Opinion In Chemical Biology* **1997**, *1*, 537-542.
(3) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. *Dalton Transactions* **2009**, 2271-2280.
(4) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. **2009**, *460*, 855-862.
(5) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959-9973.
(6) Gibney, B. R.; Rabanal, F.; Skalicky, J. J.; Wand, A. J.; Dutton, P. L. *J. Am. Chem. Soc.* **1999**, *121*, 4952-4960.
(7) Gibney, B. R.; Dutton, P. L. *Adv. Inorg. Chem, Vol 51* **2001**, *51*, 409-455.
(8) Reedy, C. J.; Gibney, B. R. *Chem. Rev.* **2004**, *104*, 617-649.
(9) Zhuang, J.; Amoroso, J. H.; Kinloch, R.; Dawson, J. H.; Baldwin, M. J.; Gibney, B. R. *Inorg. Chem.* **2004**, *43*, 8218-8220.
(10) Kennedy, M. L.; Petros, A. K.; Gibney, B. R. *J. Inorg. Biochem.* **2004**, *98*, 727-732.
(11) Cochran, F. V.; Wu, S. P.; Wang, W.; Nanda, V.; Saven, J. G.; Therien, M. J.; DeGrado, W. F. *Journal of the American Chemical Society* **2005**, *128*, 663.
(12) McAllister, K. A.; Zou, H.; Cochran, F. V.; Bender, G. M.; Senes, A.; Fry, H. C.; Nanda, V.; Keenan, P. A.; Lear, J. D.; Saven, J. G.; Therien, M. J.; Blasie, J. K.; DeGrado, W. F. *Journal of the American Chemical Society* **2008**, *130*, 11921-11927.
(13) Ghirlanda, G.; Osyczka, A.; Liu, W.; Antolovich, M.; Smith, K. M.; Dutton, P. L.; Wand, A. J.; DeGrado, W. F. *Journal of the American Chemical Society* **2004**, *126*, 8141-8147.
(14) Nastri, F.; Lombardi, A.; D'Andrea, L. D.; Sanseverino, M.; Maglio, O.; Pavone, V. *Peptide Science* **1998**, *47*, 5-22.
(15) Sakamoto, S. O., I.; Ueno, A.; Mihara, H. *J. Chem. Soc., Perkin Trans. 2,* **1999**, 2059–2069.
(16) Obataya, I.; Kotaki, T.; Sakamoto, S.; Ueno, A.; Mihara, H. *Bioorganic & Medicinal Chemistry Letters* **2000**, *10*, 2719-2722.
(17) Calhoun, J. R.; Nastri, F.; Maglio, O.; Pavone, V.; A.;, L.; DeGrado, W. F. *Peptide Sci.* **2005**, *80*, 264-278.
(18) Geremia, S.; DiCostanzo, L.; Randaccio, L.; Engel, D. E.; Lombardi, A.; Nastri, F.; DeGrado, W. F. *J. Am. Chem. Soc.* **2005**, *127*, 17266-17276.
(19) Kharenko, O. A.; Ogawa, M. Y. *J. Inorg. Biochem.* **2004**, *98*, 1971-1974.
(20) Kharenko, O. A.; Kennedy, D. C.; Demeler, B.; Maroney, M. J.; Ogawa, M. Y. *J. Am. Chem. Soc.* **2005**, *127*, 7678-7679.
(21) Li, X. Q.; Suzuki, K.; Kanaori, K.; Tajima, K.; Kashiwada, A.; Hiroaki, H.; Kohda, D.; Tanaka, T. *Protein Sci.* **2000**, *9*, 1327-1333.
(22) Suzuki, K.; Hiroaki, H.; Kohda, D.; Nakamura, H.; Tanaka, T. *J. Am. Chem. Soc.* **1998**, *120*, 13008-13015.

(23) Kiyokawa, T.; Kanaori, K.; Tajima, K.; Koike, M.; Mizuno, T.; Oku, J.; Tanaka, T. *J. Pept. Res.* **2004**, *63*, 347-353.
(24) Shiga, D.; Nakane, D.; Inomata, T.; Masuda, H.; Oda, M.; Noda, M.; Uchiyama, S.; Fukui, K.; Takano, Y.; Nakamura, H.; Mizuno, T.; Tanaka, T. *Biopolymers* **2009**, *91*, 907-916.
(25) Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2004**, *43*, 7902-7915.
(26) Farrer, B.; Pecoraro, V. L. *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3760-3765.
(27) Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. *J. Mol. Biol.* **1998**, *280*, 897-912.
(28) 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.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 6195-6196.
(29) Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. *Biochemistry* **2001**, *40*, 14696-14705.
(30) Touw, D. S.; Nordman, C. E.; Stuckey, J. A.; Pecoraro , V. L. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969-11974.
(31) Luczkowski, M.; Stachura, M.; Schirf, V.; Demeler, B.; Hemmingsen, L.; Pecoraro, V. L. *Inorganic Chemistry* **2008**, *47*, 10875-10888.
(32) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angewandte Chemie International Edition* **2009**, *48*, 7371-7374.
(33) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *Journal of the American Chemical Society* **2010**, *132*, 13240.
(34) Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2002**, *124*, 8042-8054.
(35) Lee, K. H.; Matzapetakis, M.; Mitra, S.; Marsh, E. N. G.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2004**, *126*, 9178-9179.
(36) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2006**, *45*, 2864-2868.
(37) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proceedings of the National Academy of Sciences* **2008**, *105*, 16566-16571.
(38) Iranzo, O.; Cabello, C.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2007**, *46*, 6688-6691.
(39) DeLano, W. L. *The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, California, USA* **2005**, http://www.pymol.org.
(40) **Nomenclature**, *While previously describing in detail the nomenclature for peptide modifications, we have not fully provided descriptors for metal complexation to these peptides. We will provide these definitions here with respect to Cd(II) and the exogenous ligand water; however, the nomenclature is general for any metal or non-protein bound ligand. We refer to an "apo peptide" as one which does not have metal in the designed metal coordination site within the hydrophobic interior of the coiled coils. We do not refer to an "apo peptide" to represent situations without the presence of any cofactors. Metals that may be bound non-specifically to the exterior hydrophilic residues are not considered to be metallated proteins/peptides in this context. Cd(II)(TRIL16C)$_3^-$ indicates*

that Cd(II) is bound to the three stranded coiled coil; however, the metal coordination environment is either unknown or is a mixture of species (e.g., Cd(II)(TRIL16C)$_3^-$ contains a 55:45 mixture of CdS_3O and CdS_3, respectively). When the metal is placed within brackets, as in [Cd(II)(H_2O)](TRIL12AL16C)$_3^-$, we specify the coordination environment (in this case CdS_3O) and the possible exogenous ligand (here water). If the site has only protein ligands, then it would be represented as [Cd(II)](TRIL16Pen)$_3^-$. At low pH, one can encounter the situation where the cysteine sulfurs are protonated and may, or may not, be directly coordinated to the metal. In this case, we indicate this situation as [Cd(II)](TRIL16Pen[S,(SH)$_2$])$_3^-$ in which one cysteine is bound as a thiolate and the remaining two cysteines are protonated. When a peptide contains more than one binding site, we use the above nomenclature, but specify which site is being discussed by adding a superscript outside of the bracket to indicate the amino acid residues coordinating the specific metal (e.g., [Cd(II)]^{16}[Cd(II)(H_2O)]^{30}(GRANDL16PenL26AL30C)$_3^{-2}$). Finally for cases where dual site peptides exhibit specificity for metal complexation, we will utilize [apo] to designate an empty site (e.g., [apo]^{16}[Cd(II)(H_2O)]^{30}(GRANDL16PenL26AL30C)$_3^-$).

(41) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Oxford Univ. Press: New York, 2000.

(42) Mantle, M.; Stewart, G.; Zayas, G.; King, M. *Biochem. J.* **1990**, *266*, 597-604.

(43) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Biol. Inorg. Chem.* **2006**, *11*, 876-890.

(44) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry - A European Journal* **2009**, *15*, 3761-3772.

(45) Pace, C. N.; Scholtz, J. M. *Protein Structure: A Practical Approach*; Oxford University Press: Oxford, 1997.

(46) Boice, J. A.; Dieckmann, G. R.; Degrado, W. F.; Fairman, R. *Biochemistry* **1996**, *35*, 14480-14485.

(47) Cobas, C.; Cruces, J.; Sardina, F. J.; 2.3 ed.; Universidad de Santiago de Compostela, Spain, 2000.

(48) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. In *Journal of Chemical Physics*, 1979; Vol. 71.

(49) Hoult, D. I. In *Journal of Magnetic Resonance*, 1976; Vol. 21.

(50) Rance, M. In *Journal of Magnetic Resonance*, 1987; Vol. 74.

(51) MestReNova; X ed. Santiago de Compostela, Spain, 2009.

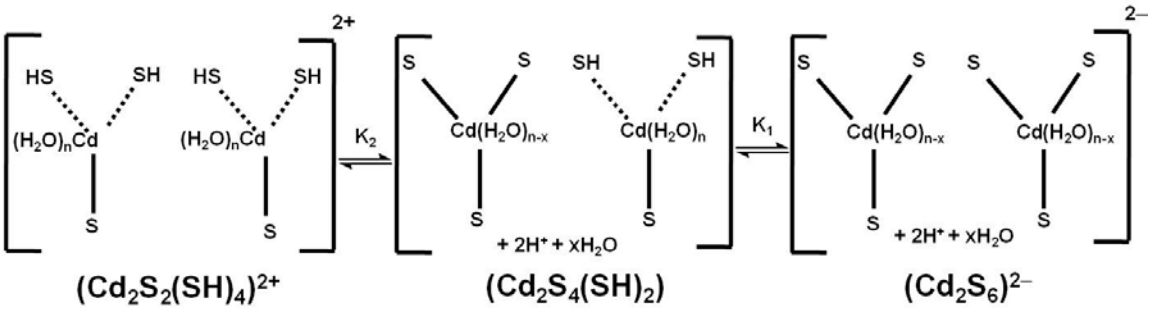
(52) Hemmingsen, L.; Bauer, R.; Bjerrum, M. J.; Zeppezauer, M.; Adolph, H. W.; Formicka, G.; Cedergren-Zeppezauer, E. *Biochemistry* **1995**, *34*, 7145-7153.

(53) Hemmingsen, L. B., R.; Bjerrum, M.; Adolph, H.; Zeppezauer, M.;Cedergren-Zeppezauer, E. *Eur. J. Biochem.* **1996**, *241*, 546-551.

(54) Matzapetakis, M.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2005**, *127*, 18229-18233.

(55) Tanaka, T.; Mizuno, T.; Fukui, S.; Hiroaki, H.; Oku, J.-i.; Kanaori, K.; Tajima, K.; Shirakawa, M. *J. Am. Chem. Soc.* **2004**, *126*, 14023-14028.
(56) Lombardi, A.; Summa, C. M.; Geremia, S.; Randaccio, L.; Pavone, V.; DeGrado, W. F. *Proc. Natl. Acad. Sci., USA* **2000**, *97*, 6298-6305.
(57) Summa, C. M.; Rosenblatt, M. M.; Hong, J.-K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2002**, *321*, 923-938.
(58) Spiegel, K.; De Grado, W. F.; Klein, M. L. *Proteins: Struct., Funct., and Bioinf.* **2006**, *65*, 317-330.
(59) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro , V. L. *Journal of the American Chemical Society* **2010**, *132,* 13240-13250.
(60) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D.; DeGrado, W.; Eisenberg, D. *Science* **1993**, *259*, 1288-1293.

Scheme 3-1. Model used to derive the fitting equations for the experimental UV-Vis pH titration curves of peptides containing two metal binding sites.



NOTES: 1) Solid bonds represent Cd(II)-thiolate bonds. Dotted bonds represent coordination of Cd(II) to thiols

2) n is the number of water molecules bound to Cd(II) initially. x is the number of water molecules released upon formation of Cd(II)-thiolato bond

Derivation of the equation based on the model presented in Scheme 1 for fitting the UV-Vis pH titration curves for the peptides containing two metal binding sites.

$$\left(Cd_2S_2(SH)_4\right)^{2+} \Leftrightarrow \left(Cd_2S_4(SH)_2\right) + 2H^+ \text{.....}(i)$$

$$\left(Cd_2S_4(SH)_2\right) \Leftrightarrow \left(Cd_2S_6\right)^{2-} + 2H^+ \text{.....}(ii)$$

$$(i) \Rightarrow K_2 = \frac{\left[\left(Cd_2S_4(SH)_2\right)\right]\left[H^+\right]^2}{\left[\left(Cd_2S_2(SH)_4\right)^{2+}\right]} \text{.....}(iii)$$

$$(ii) \Rightarrow K_1 = \frac{\left[\left(Cd_2S_6\right)^{2-}\right]\left[H^+\right]^2}{\left[\left(Cd_2S_4(SH)_2\right)\right]} \text{.....}(iv)$$

$$K_1.K_2 = \frac{\left[\left(Cd_2S_6\right)^{2-}\right]\left[H^+\right]^4}{\left[\left(Cd_2S_2(SH)_4\right)^{2+}\right]} \text{.....}(v)$$

$$T_P = \left[\left(Cd_2S_2(SH)_4\right)^{2+}\right] + \left[\left(Cd_2S_4(SH)_2\right)\right] + \left[\left(Cd_2S_6\right)^{2-}\right] \text{.....}(vi)$$

$$(iv) \Rightarrow \left[\left(Cd_2S_4(SH)_2\right)\right] = \left[\left(Cd_2S_6\right)^{2-}\right]\frac{\left[H^+\right]^2}{K_1} \text{.....}(vii)$$

$$(iii)\ \&\ (vii) \Rightarrow \left[\left(Cd_2S_2(SH)_4\right)^{2+}\right] = \left[\left(Cd_2S_6\right)^{2-}\right]\frac{\left[H^+\right]^4}{K_1.K_2} \text{.....}(viii)$$

$$(vi), (vii)\ \&\ (viii) \Rightarrow T_P = \left[\left(Cd_2S_6\right)^{2-}\right]\left(1 + \frac{\left[H^+\right]^2}{K_1} + \frac{\left[H^+\right]^4}{K_1.K_2}\right)$$

$$\Rightarrow \left[\left(Cd_2S_6\right)^{2-}\right] = \frac{T_P}{\left(1 + \frac{\left[H^+\right]^2}{K_1} + \frac{\left[H^+\right]^4}{K_1.K_2}\right)} \text{.....}(ix)$$

$$(vii)\ \&\ (ix) \Rightarrow \left[\left(Cd_2S_4(SH)_2\right)\right] = \frac{T_P.\frac{\left[H^+\right]^2}{K_1}}{\left(1 + \frac{\left[H^+\right]^2}{K_1} + \frac{\left[H^+\right]^4}{K_1.K_2}\right)} \text{.....}(x)$$

$$A = \varepsilon_1.\left[\left(Cd_2S_4(SH)_2\right)\right] + \varepsilon_2.\left[\left(Cd_2S_6\right)^{2-}\right] \text{.....}(xi)$$

$$(ix)\ \&\ (x)\ \&\ (xi) \Rightarrow A = \frac{\left(\varepsilon_1.\frac{\left[H^+\right]^2}{k_1} + \varepsilon_2\right)T_P}{\left(1 + \frac{\left[H^+\right]^2}{K_1} + \frac{\left[H^+\right]^4}{K_1.K_2}\right)}$$

$$\Rightarrow A = \frac{\left(\varepsilon_1.10^{(pK_1 - 2pH)} + \varepsilon_2\right)T_P}{1 + 10^{(pK_1 - 2pH)} + 10^{(pK_1 + pK_2 - 4pH)}}$$

$$\Rightarrow A = \frac{\varepsilon_2.T_P}{1 + 10^{(pK_1 - 2pH)} + 10^{(pK_1 + pK_2 - 4pH)}} + \frac{\varepsilon_1.T_P.10^{(pK_1 - 2pH)}}{1 + 10^{(pK_1 - 2pH)} + 10^{(pK_1 + pK_2 - 4pH)}} \text{.....}(xi)$$

$$\Rightarrow A = \frac{\varepsilon_2.T_P}{1 + 10^{(pK_1 - 2pH)} + 10^{(pK_1 + pK_2 - 4pH)}} + \frac{\varepsilon_1.T_P}{1 + 10^{(2pH - pK_1)} + 10^{(pK_2 - 2pH)}} \text{.....}(xii)$$

NOTES: 1) T_P = Total concentration of metallated peptide

2) It has been assumed that the absorbing species are $[Cd_2S_4(SH)_2]$ and $[Cd_2S_6]^{2-}$

3) Equation (*xii*) is a rearranged version of equation (*xi*) and is obtained by multiplying the numerator and denominator of the second part of equation (*xi*) by $10^{(2pH-pK_1)}$

4) Fitting of the experimental UV-Vis pH titration curves have been performed using equation (*xi*) for **GRAND**L16PenL19IL23PenL26I and equation (*xii*) for **GRAND**L12AL16CL26AL30C

Table 3-1. List of different derivatives of the **TRI** and **GRAND** peptides.

Peptide	***Sequence***
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
GRAND	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
TRIL12AL16C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LKALEEK G-NH_2
TRIL16Pen	Ac-G LKALEEK LKALEEK **X**KALEEK LKALEEK G-NH_2
GRANDL12AL16C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LKALEEK LKALEEK G-NH_2
GRANDL26AL30C	Ac-G LKALEEK LKALEEK LKALEEK LKA**A**EEK **C**KALEEK G-NH_2
GRANDL16Pen	Ac-G LKALEEK LKALEEK **X**KALEEK LKALEEK LKALEEK G-NH_2
GRANDL16PenL26AL30C	Ac-G LKALEEK LKALEEK **X**KALEEK LKA**A**EEK **C**KALEEK G-NH_2
GRANDL12AL16CL26AL30C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LK**A**LEEK **C**KALEEK G-NH_2
GRANDL12AL16CL26AL30CL33I	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LK**A**LEEK **C**KA**I**EEK G-NH_2
GRANDL16PenL23PenL26I	Ac-G LKALEEK LKALEEK **X**KALEEK **X**KA**I**EEK LKALEEK G-NH_2
GRANDL16PenL19IL23PenL26I	Ac-G LKALEEK LKALEEK **X**KA**I**EEK **X**KA**I**EEK LKALEEK G-NH_2

X = Penicillamine. Residues in red indicate modifications.

Table 3-2. Apparent pK_{a2} values for **TRI** and **GRAND** peptides.

Peptide	***Apparent pK_{a2}***	
TRIL12AL16C	12.2 ± 0.2	
TRIL16Pen	15.8 ± 0.2	
GRANDL12AL16C	11.3 ± 0.2	
GRANDL26AL30C	9.9 ± 0.2	
GRANDL16Pen	15.7 ± 0.2	
GRANDL16PenL26AL30C	9.6 ± 0.2	(L26AL30C site)
	16.1 ± 0.2	(L16Pen site)
GRANDL12AL16CL26AL30C	11.8 ± 0.2	(L12AL16C site)
	13.9 ± 0.2	(L26AL30C site)
GrandL12AL16CL26AL30CL33I	11.6± 0.2	(L12AL16C site)
	13.5± 0.2	(L26AL30C site)
GRANDL16PenL19IL23PenL26I	15.2 ± 0.2	(L23Pen site)
	15.9 ± 0.2	(L16Pen site)

Table 3-3. Parameters fitted to the PAC data[a]

Peptide ($c_{Cd(II)}/c_{Peptide\ monomer}$)	pH	ω_0 (rad/ns)	η	$\Delta\omega_0/\omega_0$ (×100)	$1/\tau_c$ (μs^{-1})	A (×100)	Population[c] (%)	χ_r^2
GRANDL26AL30C	7.0	0.348(1)	0.12(2)	1.5(5)	3.6(5)	1.8(2)	21	
(1/12)		0.245(3)	0.30(3)	12(2)	3.6(5)	5.1(7)	61	1.19
		0.107(9)[b]	0.57(9)	15[d]	3.6(5)	1.5(3)	18	
GRANDL26AL30C	9.1	0.347(1)	0.10(2)	1.4(5)	2.2(4)	1.7(2)	19	1.29
(1/12)		0.243(3)	0.31(3)	14(5)	2.2(4)	7.4(6)	81	
GRANDL16PenL26AL30C	6.5	0.347(2)	0.05(8)	2.3(3)	4.2(5)	1.8(3)	17	1.15
(1/12)		0.245(3)	0.32(3)	12(2)	4.2(5)	6.5(8)	61	
		0.105(7)[b]	0.59(7)	15[d]	4.2(5)	2.3(3)	22	
GRANDL16PenL26AL30C	9.3	0.345(1)	0.10(3)	0.5(5)	5.9(7)	1.7(3)	17	1.11
(1/12)		0.247(3)	0.26(3)	12(2)	5.9(7)	7.2(7)	72	
		0.159(8)[b]	0.77(5)	0(4)	5.9(7)	1.1(4)	11	
GRANDL16PenL26AL30C	9.3	0.457(2)	0.0(5)	1.9(6)	7.5(1)	2.3(3)	29	1.23
(1.85/3)		0.349(2)	0.0(3)	0.4(8)	7.5(1)	0.9(2)	11	
		0.255(4)	0.22(5)	9(2)	7.5(1)	3.4(5)	43	
		0.125(9)[b]	0.4(1)	12(6)	7.5(1)	1.3(3)	16	
GRANDL16PenL19IL23PenL26I (1.85/3)[e]	9.2	0.461(1)	0.02(4)	0.7(3)	17(2)	6.8(3)	100	1.44
GRANDL12AL16CL26AL30C	8.7	0.342(6)	0.148(8)	1.8(3)	3.8(6)	5.3(3)	55	1.08
(1.85/3)		0.242(8)	0(1)	16(4)	3.8(6)	3.9(8)	40	
		0.142(2)[b]	0.1(2)	0(2)	3.8(6)	0.5(2)	5	

a: The numbers in parentheses are the standard deviations of the fitted parameters. b: Including this third NQI significantly improves the fit, but at least two different parameter sets give equally good fits – one possible set of parameters is presented here c: The relative population of each species determined using the amplitudes in this Table. d: Fixed in the fit. e: Although the stoichiometry implies that two sites are occupied, the χ_r^2 is not reduced significantly upon inclusion of a second NQI in the fit, and it was not possible to reliably derive parameters for two individual NQIs.

Table 3-4. Thermodynamic parameters determined from GuHCl-induced unfolding curve fitting of the peptides **GRAND**L26AL30C, **GRAND**L12AL16C and **GRAND**L12AL16CL26AL30C.

Peptide	ΔG^0 (kcal mol^{-1})	m (kcal mol^{-1} M_{GuHCl}^{-1})
GRANDL26AL30C	20 ± 0.4	2.17 ± 0.14
GRANDL12AL16C	20.9 ± 0.2	2.43 ± 0.07
GRANDL12AL16CL26AL30C	15.7 ± 0.1	1.84 ± 0.01

Table 3-5. % Cd(II) species (4-coordinate and 3-coordinate) based on ^{113}Cd NMR chemical shift values and the ^{113}Cd NMR -^{111m}Cd PAC spectroscopic correlation.

Peptide	**^{113}Cd NMR δ (ppm)**	**% Cd species 4-coordinate (CdS_3O)**	**% Cd species 3-coordinate CdS_3**
GRANDL12AL16CL23Pen	575 640	100 50	0 50
GRANDL16PenL23PenL26I	678 666	20 29	80 71
GRANDL16PenL19IL23PenL26I	686 681	13 17	87 83
GRANDL12AL16CL26AL30C	572 589	100 92	0 8
GRANDL12AL16C	572	100	0
GRANDL26AL30C	589	92	8

Chemical shifts of 579 and 702 ppm are used as values corresponding to a 100 % 4-coordinate and 3-coordinate Cd(II) species, respectively.[44]

Table 3-6. Metallation states of the two Pen sites of **GRAND**L16PenL19IL23PenL26I and the corresponding ^{113}Cd NMR chemical shifts under different experimental conditions.

Metallation state (site A)	**Metallation state (site B)**	**^{113}Cd δ (ppm) (site A)**	**^{113}Cd δ (ppm) (site B)**
$[Cd(pep)_3]^-$	apo	686	----
$[Cd(pep)_3]^-$	$[Cd(pep)(Hpep)_2]^+$	686	ND[a]
$[Cd(pep)_3]^-$	$[Cd(pep)_3]^-$	686	681
apo	$[Cd(pep)_3]^-$	----	691
$[Cd(pep)(Hpep)_2]^+$	$[Cd(pep)_3]^-$	ND[a]	691

ND=Not Detected.

[a] ^{113}Cd resonances for $[Cd(pep)(Hpep)_2]^+$or $[Cd(pep)(Hpep)_2(OH_2)]^+$ have never been detected for these or single site peptides such as Cd(II)(**TRI**L16C)$_3^-$. [34]

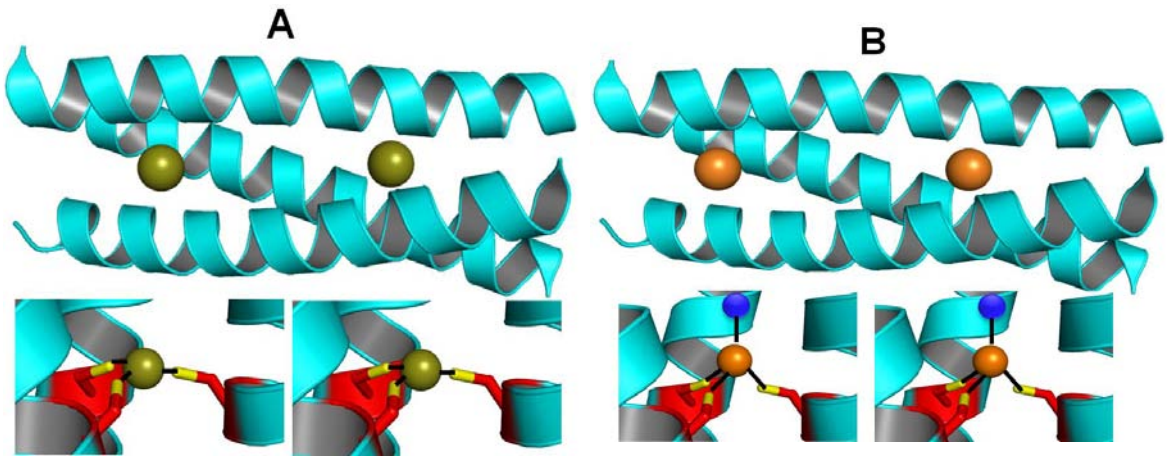


Figure 3-1. PyMol models of the two constructs capable of encapsulating two Cd(II) ions as trigonal planar CdS_3 (A) and pseodotetrahedral CdS_3O (O being water molecule) (B) geometries, respectively. Peptide backbone is shown as cyan. 3-coordinate Cd(II) is shown as green spheres and the 4-coordinate Cd(II) is shown as orange spheres. Exogenous water molecules are shown as blue sphere.

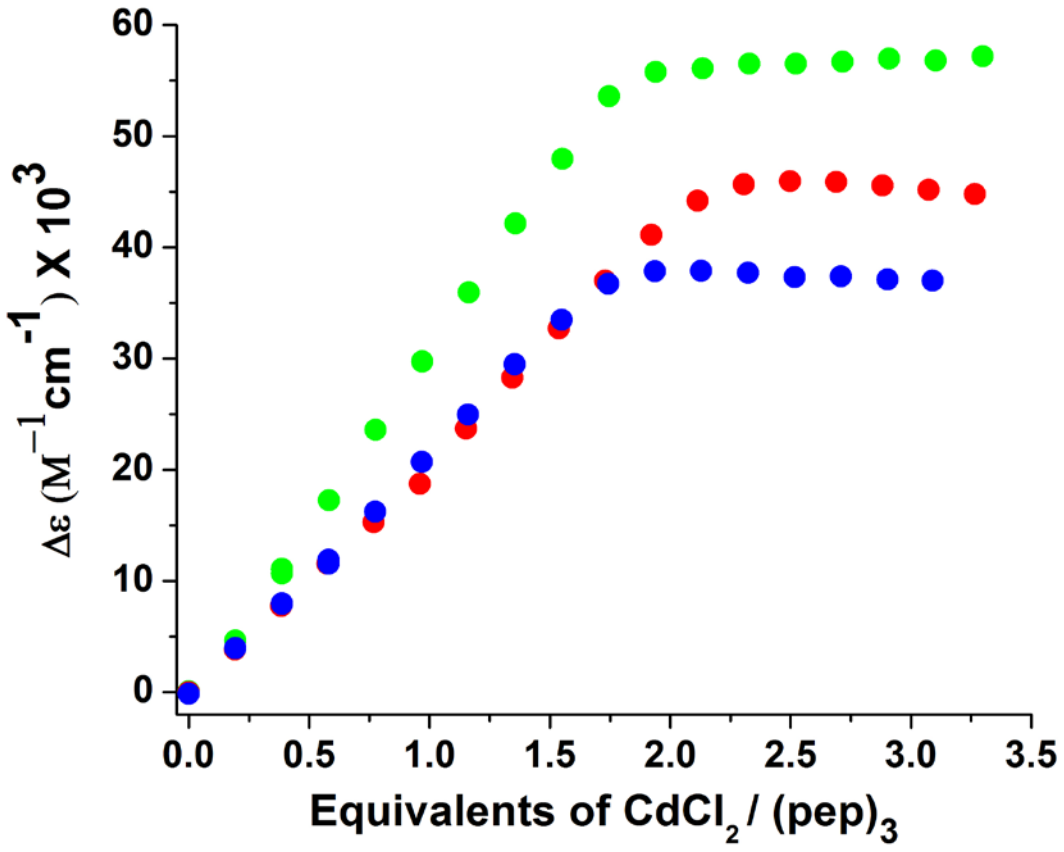


Figure 3-2. Titration curves obtained by plotting the change in the absorbance at 235 nm as a function of the equivalents of $CdCl_2$ added to solutions containing 20 μM (**GRAND**L16PenL26AL30C)$_3$ at pH 9.0 (red), 20 μM (**GRAND**L12AL16CL26AL30C)$_3$ at pH 8.6 (blue) and 20 μM (**GRAND**L16PenL19IL23PenL26I)$_3$ at pH 9.5 (green). All the curves plateau at 2 ± 0.14 equivalents of $CdCl_2$.

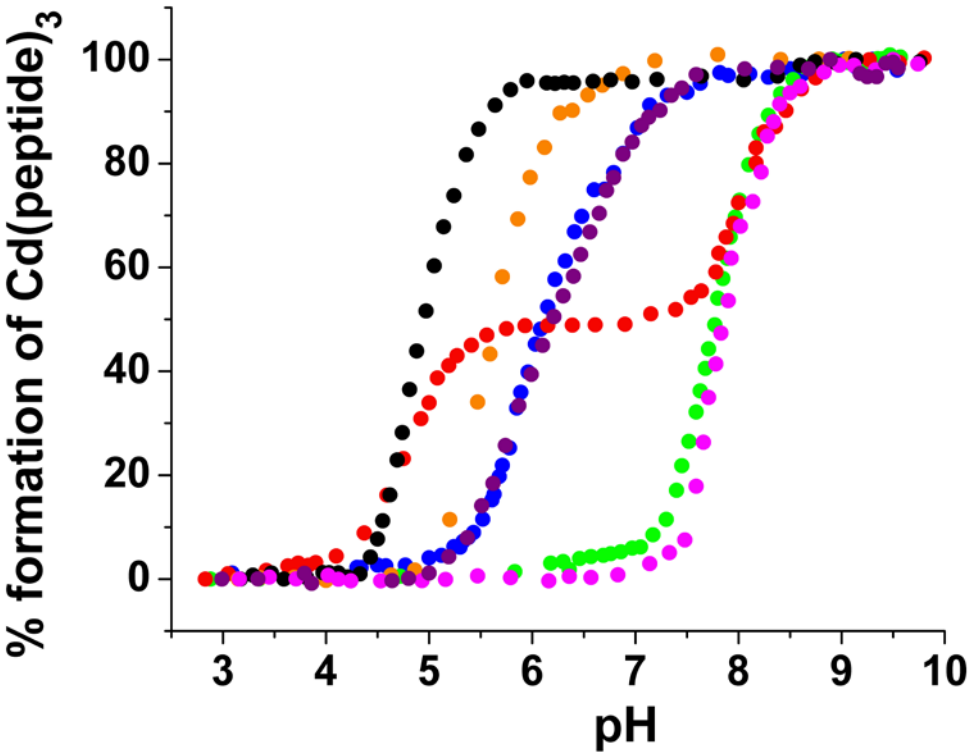


Figure 3-3. pH dependence of the binding of 2 equivalents of $CdCl_2$ to 20 μM (**GRAND**L16PenL26AL30C)$_3$ (red), (**GRAND**L12AL16CL26AL30C)$_3$ (blue), (**GRAND**L16PenL19IL23PenL26I)$_3$ (green), (**GRAND**L12AL16CL26AL30CL33I)$_3$ (purple) and of 1 equivalent of Cd(II) to 20 μM (**GRAND**L26AL30C)$_3$ (black), (**GRAND**L12AL16C)$_3$ (orange) and (**GRAND**L16Pen)$_3$ (magenta). UV/Vis absorbance due to LMCT band at 235 nm was monitored during the course of titration and is plotted as normalized absorbance vs. pH.

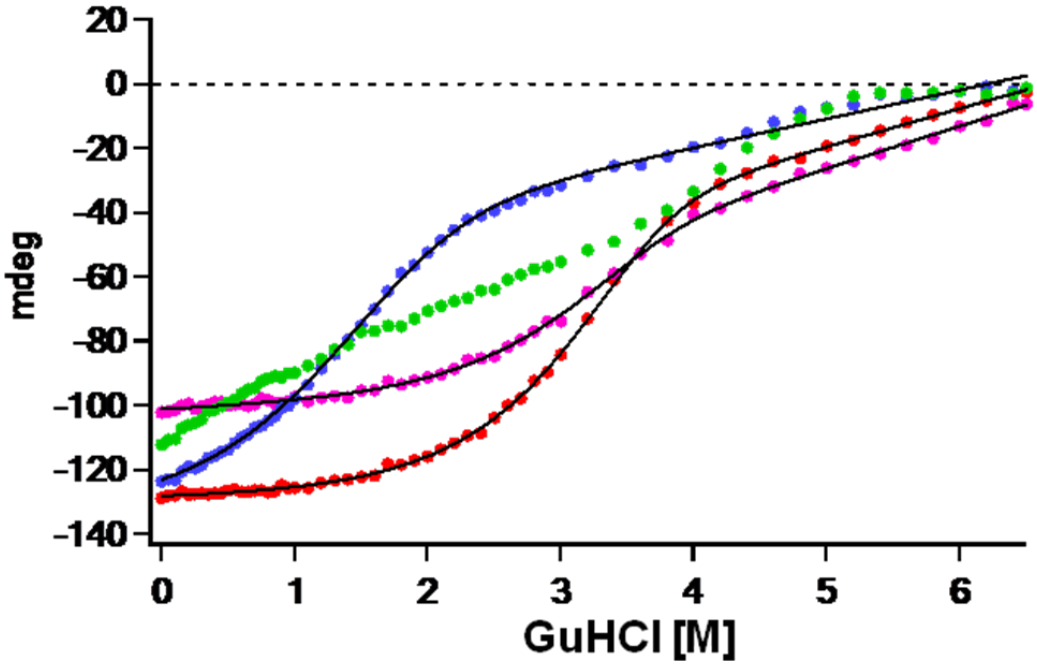


Figure 3-4. GuHCl denaturation titration curves of solutions containing 10 μM monomer peptides **GRAND**L26AL30C (magenta), **GRAND**L12AL16C (red), **GRAND**L12A16CL26AL30C (blue) and **GRAND**L12A16CL26AL30CL33I (green) in 10 mM phosphate buffer at pH 6.5.

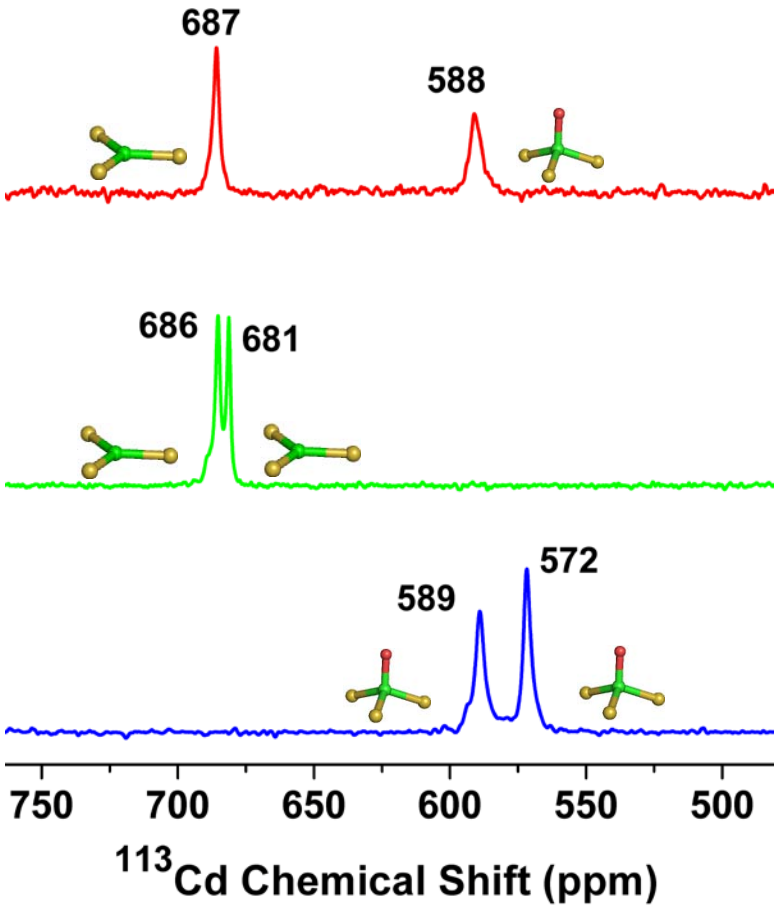


Figure 3-5. ^{113}Cd NMR spectra of 3.0 mM [Cd(II)]^{16}[Cd(II)(H_2O)]^{30}(**GRAND**L16PenL26AL30C)$_3^{2-}$ at pH 9.6 (red), 3.4 mM [Cd(II)]$_2$(**GRAND**L16PenL19IL23PenL26I)$_3^{2-}$ at pH 9.5 (green) and 3.3 mM [Cd(II)(H_2O)]$_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ at pH 8.5 (blue). Ball and stick models show the coordination environments of 3- and 4-coordinate Cd(II) complexes. Yellow spheres represent S atoms of Cys/Pen with the green and red spheres representing Cd(II) ion and water molecule, respectively.

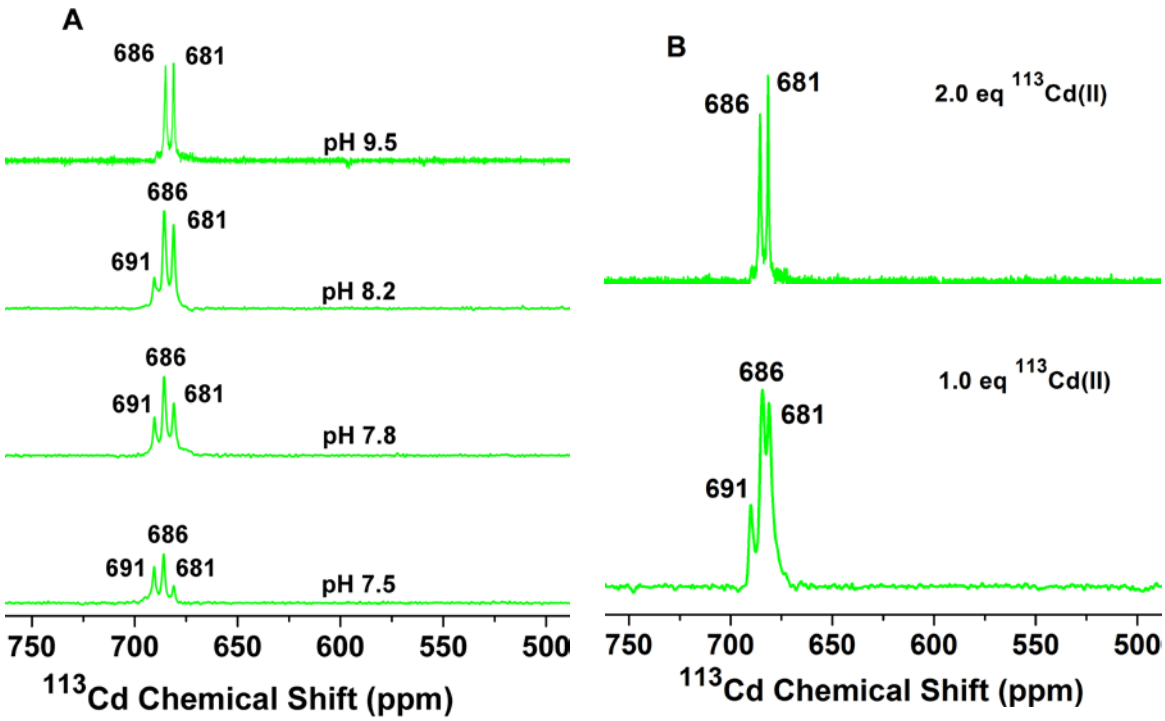


Figure 3-6. ^{113}Cd NMR spectra of solutions containing **(A)** 3.0 mM (**GRAND**L16PenL19IL23PenL26I)$_3$ and 2 equivalents of $^{113}Cd(NO_3)_2$ at different pH values, and (**B**) 3.3 mM (**GRAND**L16PenL19IL23PenL26I)$_3$ loaded with 1 and 2 equivalents of $^{113}Cd(NO_3)_2$ at pH 9.5.

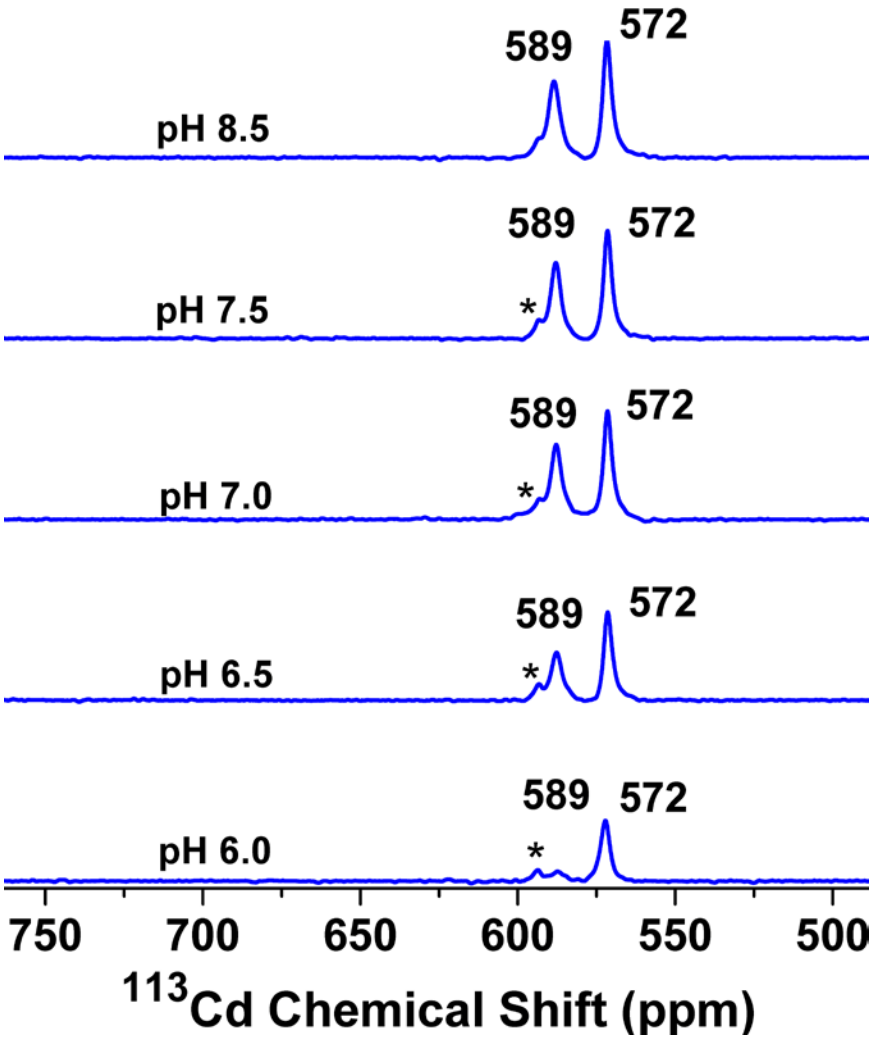


Figure 3-7. ^{113}Cd NMR spectra of solutions containing 3.3 mM (**GRAND**L12AL16CL26AL30C)$_3$ and 2 equivalents of $^{113}Cd(NO_3)_2$ at different pH values. (the most downfield peak marked with star is an impurity. Most likely the impurity is a peptide of shorter length which is produced during automated peptide synthesis and inseparable by HPLC).

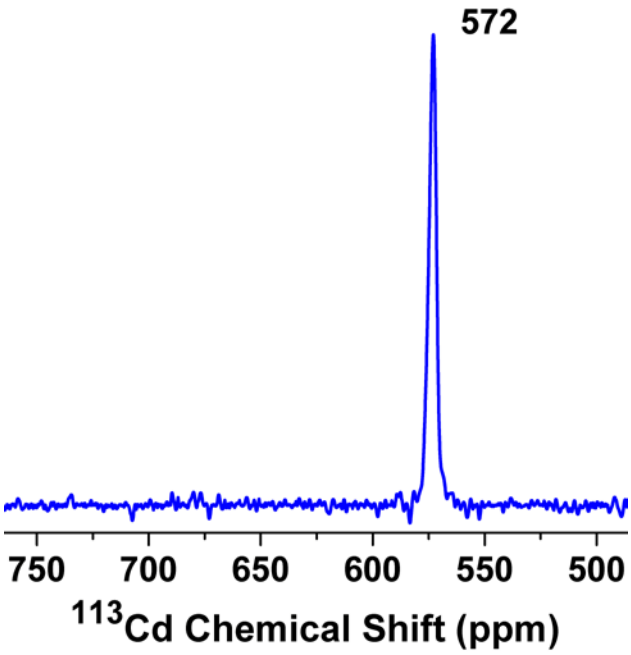


Figure 3-8. ^{113}Cd NMR spectrum of 3.3 mM (**GRAND**L12AL16CL26AL30C)$_3$ in the presence of 1.0 eq of $^{113}Cd(NO_3)_2$, at pH 6.0.

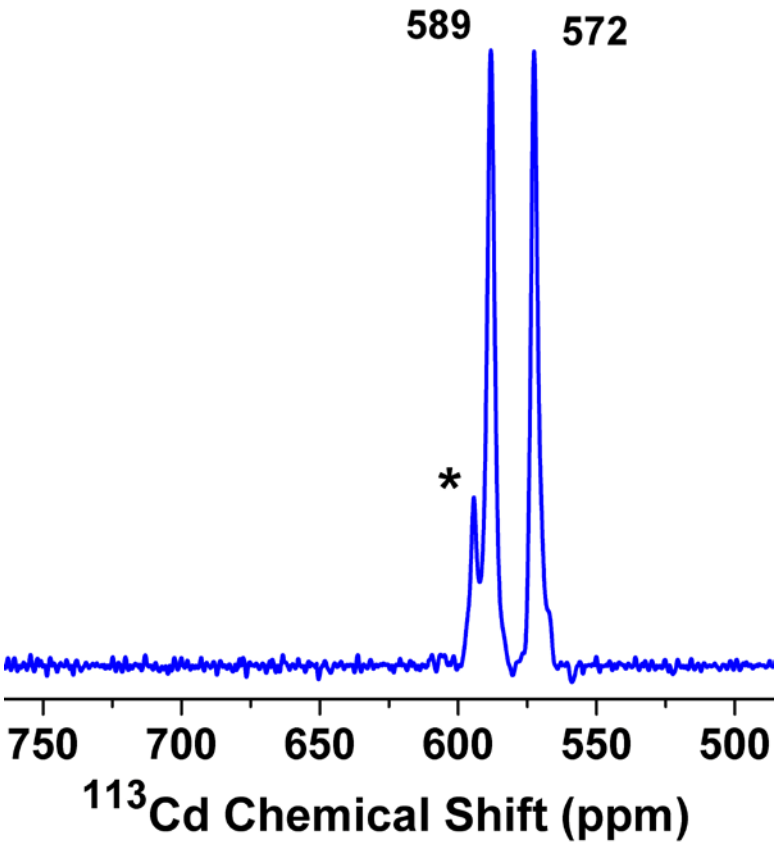


Figure 3-9. ^{113}Cd NMR spectra of solutions containing 1.0 equivalent of $^{113}Cd(NO_3)_2$ and 3.0 mM (**GRAND**L12AL16CL26AL30C)$_3$ at pH 8.5 (the most downfield peak marked with star is an impurity. Most likely the impurity is a peptide of shorter length which is produced during automated peptide synthesis and inseparable by HPLC).

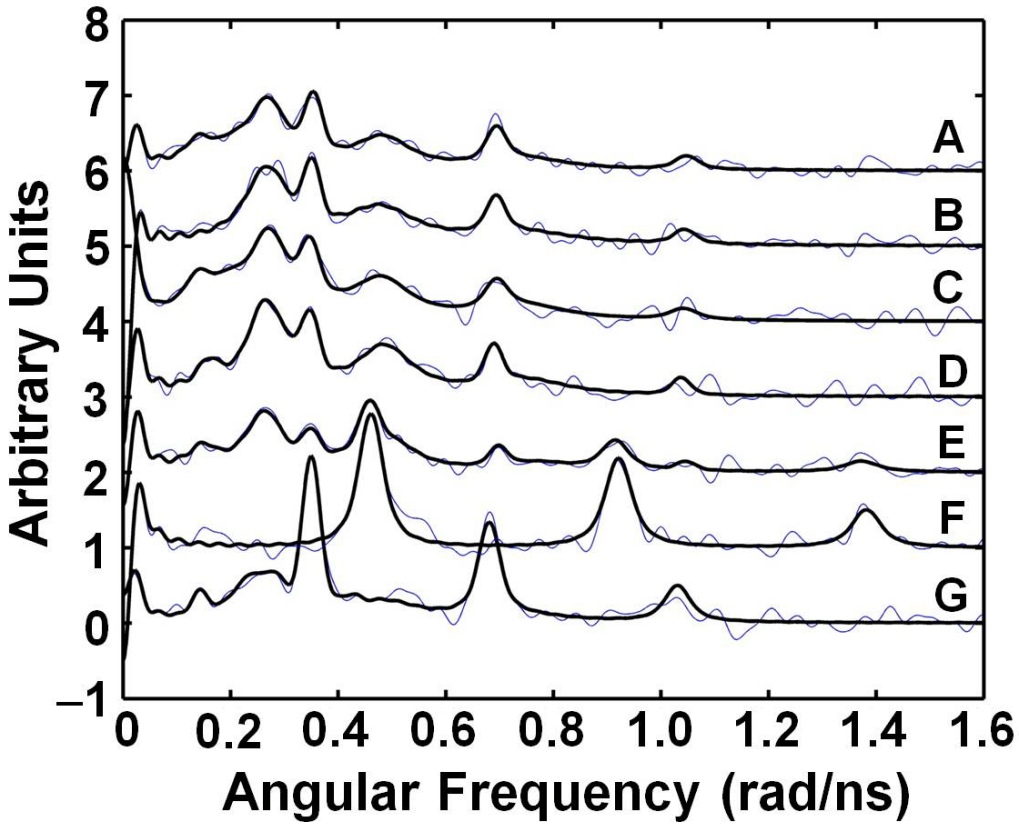


Figure 3-10. ^{111m}Cd PAC spectra of the different **GRAND** peptides [Fourier transforms: experimental data (thin line) and fits (bold faced line) are shown overlaid]. All the samples contained 20 mM appropriate buffer and 250 - 300 µM peptide. **A**) **GRAND**L26AL30C, 1/12 eq Cd(II), pH 7.0; **B**) **GRAND**L26AL30C, 1/12 eq Cd(II), pH 9.1; **C**) **GRAND**L16PenL26AL30C, 1/12 eq Cd(II), pH 6.5; **D**) **GRAND**L16PenL26AL30C, 1/12 eq Cd(II), pH 9.3; **E**) **GRAND**L16PenL26AL30C, 1.85/3 eq Cd(II), pH 9.3; **F**) **GRAND**L16PenL19IL23PenL26I, 1.85/3 eq Cd(II), pH 9.2; **G**) **GRAND**L12AL16CL26AL30C, 1.85/3 eq Cd(II), pH 8.7.

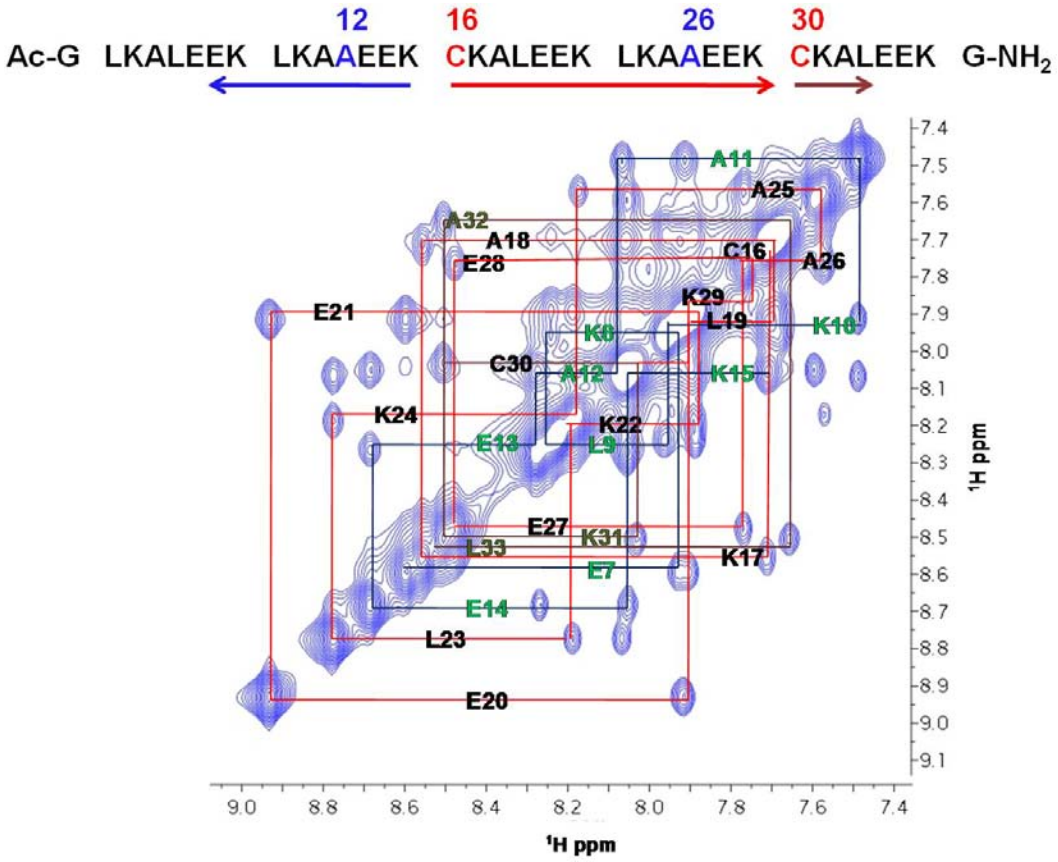


Figure 3-11. Amide region of NOESY spectrum of 3.25 mM (**GRAND**L12AL16CL26AL30C)$_3$ solution in 10% D_2O at pH 6.0 showing sequential assignments of amide protons of E7 through L33.

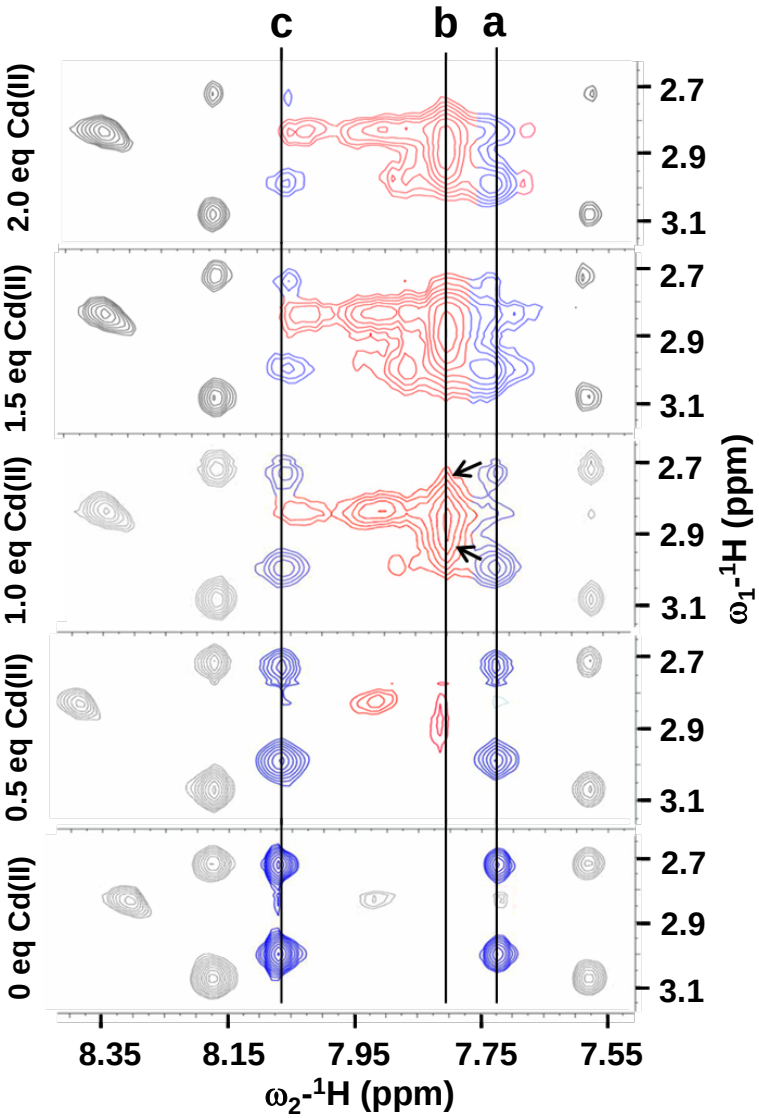


Figure 3-12. Sections of ^{1}H-^{1}H NOESY spectra of 3.25 mM (**GRAND**L12AL16CL26AL30C)$_3$ as a function of added equivalents of Cd(II) at pH 6.0. Peak at 7.91 ppm corresponds to H^N of E21 and other peaks displayed in gray correspond to interesidue NOEs (H^N_i-H^β_{i+1}) (see text for full description). Colored peaks correspond to intraresidue NOEs of the amide protons (H^N, vertical axis) and the β methylene protons (H^β, horizontal axis) of Cys: a) H^N_{16}-H^β_{16} for **GRAND**L12AL16CL26AL30C; b) H^N_{16}-H^β_{16} for [Cd(II)(H_2O)]^{16}[apo]^{30}(**GRAND**L12A16L16CL26AL30C)$^-$; and c) H^N_{30}-H^β_{30} for **GRAND**L12AL16CL26AL30C.

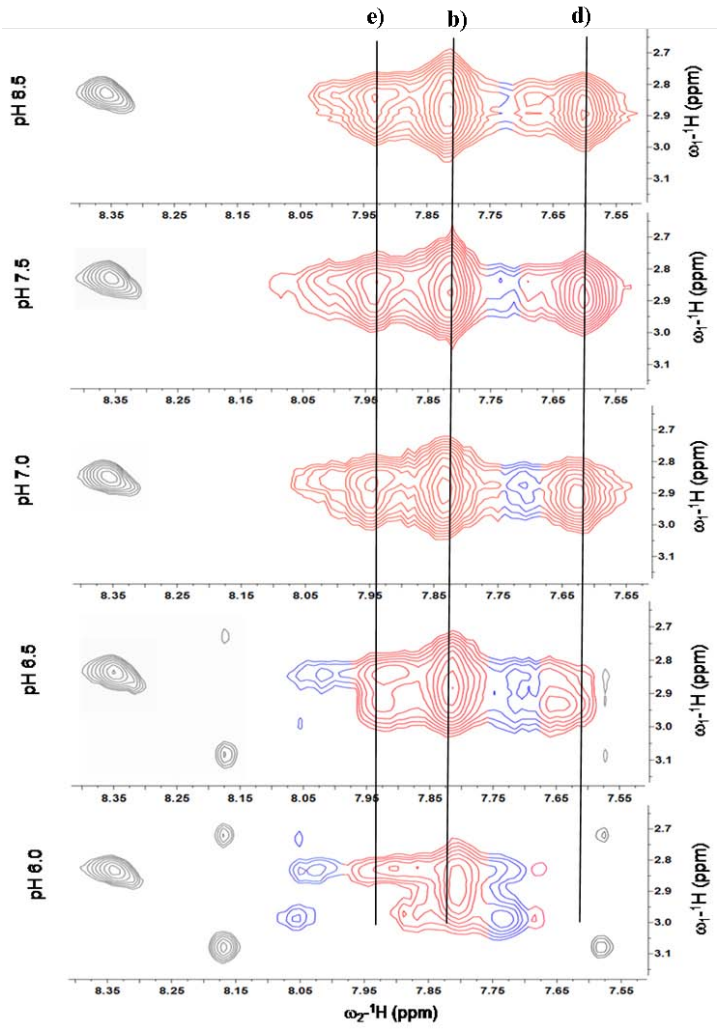


Figure 3-13. Sections of ^{1}H-^{1}H NOESY spectra of 3.21 mM (**GRAND**L12AL16CL26AL30C)$_3$ as a function of pH in the presence of 2.0 equivalents of Cd(II). With increase in pH two new cross peaks appear at 7.6 ppm (region d) and 7.93 ppm (region e) respectively. The best assignments to these resonances are d) H^{N}_{16}-H^{β}_{16} for [Cd(II)(H$_2$O)]$_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$ and e) H^{N}_{30}-H^{β}_{30} for [Cd(II)(H$_2$O)]$_2$(**GRAND**L12AL16CL26AL30C)$_3^{2-}$. The peak at 7.81 ppm (region b) indicates the presence of the [Cd(II)(H$_2$O)]^{16}[apo]^{30}(**GRAND**L12A16L16CL26AL30C)$^{4-}$ species at high pH (8.5).

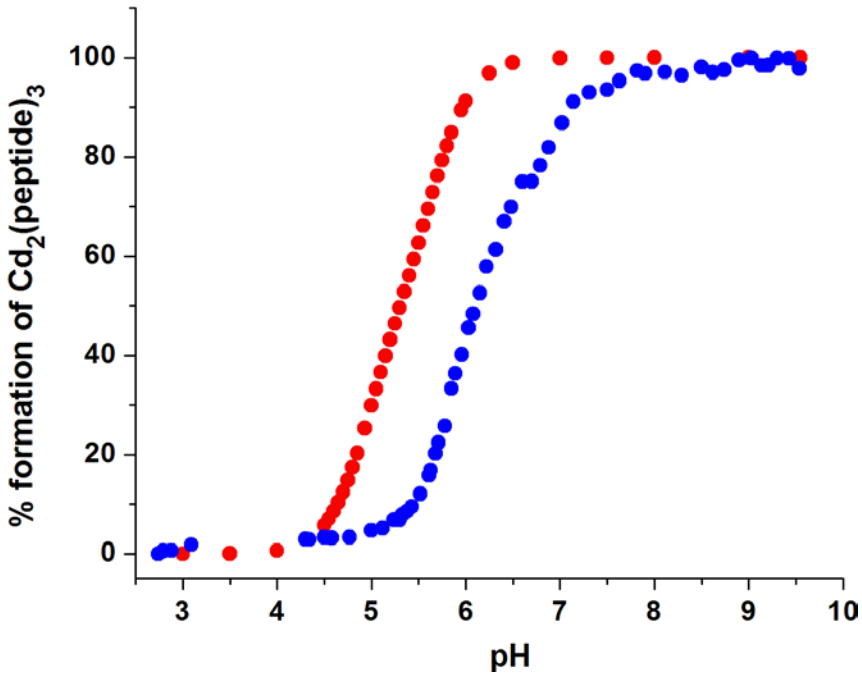


Figure 3-14. Simulated (red) and experimental (blue) UV-Vis pH titration curves for **GRAND**L12AL16CL26AL30C. pK_{a2} and extinction coefficient values of the corresponding mono substituted peptides **GRAND**L12AL16C and **GRAND**L26AL30C were used to generate the simulated curve, the values being 11.3, 9.9 and 19247 M^{-1} cm^{-1}, 18705 M^{-1} cm^{-1}, respectively.

Chapter 4

Investigation of Cd(II) Exchange in Designed Three-Stranded Coiled Coil Peptides

Introduction

Metal ion transport and homeostasis are critical cellular processes required for the maintenance of different biochemical activities. There are two important classes of metal ion control within this rubric. Firstly, one may consider the proper handling of essential elements (e.g., Fe, Ni, Cu) in order to localize these metals in the appropriate protein, while minimizing the potential deleterious effects of the free ion. Alternatively, purely toxic ions (e.g., Hg(II), Cd(II) or Pb(II)) which a cell encounters must be sequestered and detoxified before they are released back into the cellular environment.

Nature often employs different strategies for uptake and transport of different essential metal ions. Some proteins such as albumin (binds Ca^{2+}, Na^{+}, K^{+}), calmodulin (binds Ca^{2+}) sequester metals directly without any assistance by other proteins.[1,2] On the other hand, intracellular trafficking of transition metals such as copper and nickel occurs by the assistance of metallochaperone proteins, that are part of the cellular machinery which ensure that the right metal ion is delivered to appropriate targets in a safe manner.[3-7] Remarkably, little is

known about the fundamental rates and exchange pathways that control both the uptake and transport of metal ions by proteins into different cellular environments or different protein structure types.

Lower organisms such as bacteria, have evolved a detoxification mechanism to cope with heavy metals such as Hg(II), Cd(II), As(III) and Pb(II) by using different repressor proteins such as MerR,[8,9] CadC/CmtR,[10-13] ArsR/SmtB,[14] and pbrR.[15,16] In the absence of heavy metals these proteins bind to DNA, blocking the transcription of the genes that code for proteins that extrude heavy metals from cells. Upon metal binding the repressor protein either twists the DNA (MerR) or dissociates from DNA (CadC/CmtR, ArsR/SmtB) after undergoing a conformational change which then allow transcription to happen. In all of these cases, the proteins display exclusive Cys ligation around the metal centers. Using *de novo* designed three-stranded coiled coil peptides as model systems, research in the Pecoraro group has focused on generating structural and spectroscopic models of different metalloregulatory proteins. Additionally, the group has been interested in understanding the fundamentals of heavy metal biochemistry which requires a complete description of the metal's first coordination sphere as well as the general knowledge of the surrounding polypeptide.

De novo protein design provides constructs that are intermediate in size and complexity between small molecule inorganic complexes and native metalloproteins. These constructs are much simpler than the natural proteins, yet they retain sufficient complexity and information in their sequences required for understanding various aspects of structure-function relationships as well as metallobiochemistry. From previous studies the group has reported numerous successes for heavy metal chemistry using *de novo* peptide systems designed based on the heptad repeat, a seven amino acids repeat unit designated as **abcdefg.** These peptides self assemble in solution to form three-stranded coiled coils above pH 5.5. Studies from the designed TRI and Coil Ser series of

peptides have provided us with better understanding of heavy metal biochemistry including the first water soluble spectroscopic model[17,18] of the MerR protein, the first trigonal pyramidal As(III) bound to three cysteines[19] as a model of ArsR, and the first example of a trigonal thiolate Cd(II).[20]

One important factor that defines the exchange processes between metal ions and biological macromolecules is the dynamics involving metal exchange. In almost all of the proteins mentioned above, the rates of metal incorporation and removal are essential parts for proper functioning of these proteins. However, complete understanding of the mechanistic pathways in which metals insert in and out of the proteins has been elusive and is of fundamental interest to the bioinorganic community. My goal in this chapter is to understand the exchange of Cd(II) using designed peptides as models for more complicated native proteins. Clarification of the mechanisms of encapsulation of heavy metals such as Cd(II), and the factors that influence these rates, should serve as the basis for understanding heavy metal complexation and exchange processes in more complex systems such as metallochaperones, and metalloregulatory proteins such as CmtR which belongs to the ArsR/SmtB metal sensing family and senses Cd(II) and Pb(II).[11]

Previous kinetic studies have been performed by the Pecoraro group on the insertion of Hg(II) into the smaller, unstructured peptide **Baby**L9C, which showed that the three-stranded coiled coil assembly was unstructured and weakly associated in the absence of metal.[21] The mechanistic model for the insertion of Hg(II) to **Baby**L9C was termed as “StepAD” mechanism,[21] where the first step involved rapid collapse of the two peptide strands in the presence of Hg(II) leading to the formation of linear Hg-dithiolato complex. Addition of the third peptide strand leading to the formation of the three-stranded coiled coil was the rate determining state and resulted in an intermediate state involving two separate species, one of them being the properly folded state and another was a misfolded peptide. The properly folded species then underwent rapid

deprotonation of the third Cys thiol leading to the formation of the trithiolato-Hg(II) complex. Our subsequent studies have shown that the longer constructs such as TRI, Coil Ser and GRAND series of peptides exist as well-structured three-stranded coiled coils even in the absence of metals.[22-24] Mechanistic investigation of Hg(II) encapsulation to the associated peptide **TRI**L9C showed that Hg(II) insertion was consistent with the "Dissociation Mechanism" where one peptide strand dissociates from the three-stranded coiled coil giving the less-stable two-stranded form.[25] The two-stranded coiled coil would then encapsulate Hg(II) to form the $Hg(pep)_2$ complex as a linear Hg(II) coordination geometry. The subsequent steps were similar to the "StepAD" mechanism that involved the addition of the third peptide strand to form the three-stranded coiled coil followed by the deprotonation of the third Cys thiol leading to the formation of the trigonal $Hg(pep)_3$ complex.

The group has also performed kinetic studies involving the encapsulation of Cd(II) to **TRI**L9C.[25] Cd(II) binding to **TRI**L9C follows a different pathway than that of Hg(II) binding. In the case of Cd(II) encapsulation, the three-stranded coiled coil first opens up in a slow and partially rate-determining step to allow insertion of the metal inside the three-stranded coiled coil followed by binding of the Cys residues at the interior of the coiled coil as a fast step and consequently possible loss of exogenous water molecule(s) bound to Cd(II) ion. This mechanism is known as the "Breathing Mechanism". Although these studies were significant, the following questions need to be answered to obtain a better molecular level understanding of the metal insertion process into helical scaffolds: do the metals first interact with amino acids located at the surface of the peptide before being inserted into the coiled coil interior? Do the cysteines at the designed metal binding site play any role in assisting the insertion of the metal from the surface to the hydrophobic interior once the metal is bound to the surface residues? What are the rates of metal exchange? Do the rates depend on the location of the metal binding site at the interior of the coiled coils (middle of the helix vs. towards the helical terminus)? Is there any impact of the metal's

coordination number on the exchange dynamics (4-coordinate CdS_3O, vs. 3-coordinate CdS_3)? To address these questions, in this chapter I have investigated the exchange of Cd(II) using the GRAND series of peptides containing both dual and single metal binding sites. In the following section I will give a brief introduction about chemical exchange.

Chemical Exchange: Chemical exchange refers to a process where a nucleus exchanges between two or more chemical environments in which its NMR parameters such as chemical shift, scalar coupling and relaxation, differ. There are two types of chemical exchange, in general. The first one is intramolecular exchange that includes processes such as motions of protein sidechains, helix-coil transition, unfolding of proteins and conformational equilibria. The second type of exchange is intermolecular exchange which refers to processes such as ligand or metal binding to macromolecules, protonation/deprotonation equibria, isotope exchange processes and enzyme catalyzed reactions. NMR spectroscopy can be used to study chemical and biological processes occurring in a wide range of time scales starting from tens of picoseconds to hundreds of millisecond. A simple chemical exchange process involving ligand binding to a protein is illustrated in Figure 4-1 where the protein A and the protein-ligand complex B are present in equal populations. When the ligand is not bound, the free protein has an NMR signal at the resonance frequency ν_A. Once the ligand is bound to the protein the resulting protein-ligand complex B has an NMR signal at the resonance frequency ν_B. When there is no exchange the two species have equally intense, narrow peaks at the respective resonance frequencies of the two species, ν_A and ν_B, respectively. When the two species undergo slow exchange ($k_{ex} << \nu_A - \nu_B$) the two peaks broaden and decrease in intensity while the respective resonance frequencies do not shift their positions. With an increase in exchange rate, the two peaks start shifting towards one another and at some point the two individual peaks coalesce into a single, broad resonance. This point is called the coalescence point where the exchange rate constant is almost equal to the difference of the individual resonance

frequencies of the two exchanging species ($k_{ex} \sim \nu_A - \nu_B$). Further increase of the exchange rate produces a sharp resonance. Under fast exchange conditions the magnitude of the exchange rate constant is much greater than the difference in the individual resonance frequencies ($k_{ex} >> \nu_A - \nu_B$) of the two exchanging species.

In this chapter I present a specific case of metal binding to protein and investigate the chemical exchange behavior of the metal. Exchange behavior Cd(II) into parallel three-stranded coiled coil peptides is investigated using NMR spectroscopy. Exchange of Cd(II) is shown to be occurring on the slow exchange reagime of the NMR time scale. A possible mechanism of insertion of Cd(II) into and out of the helical bundles is presented.

Materials and Methods

Peptide Synthesis and Purification. The GRAND peptides (see Table 1 for sequence nomenclature) were synthesized on an Applied Biosystems 433A peptide synthesizer using standard Fmoc protocols,[26] and purified and characterized as described previously.[27]

UV-Vis Spectroscopy. UV-Vis spectroscopy was used to determine the stoichiometry and pH dependence of Cd(II) binding to the peptide **Grand**L26AL28QL30C.[28] All solutions were purged with argon before use to minimize the chances of oxidation of the peptides and formation of disulfide bonds. Fresh stock solutions of the purified peptides were prepared for each experiment in doubly distilled water and their concentrations determined by quantization of the Cys thiol groups using a known assay with 4,4'-dipyridyl disulphide.[29]

a) Metal Binding Titrations. Cd(II) into peptide titrations were performed at room temperature on a Cary 100 Bio UV-Vis spectrometer using a 1-cm quartz cuvette. Aliquots of 8.932 mM $CdCl_2$ stock solution were added into a 3-mL solution containing 60 µM peptide and 50 mM TRIS buffer at pH 8.5. In each

case, the difference spectra were obtained by subtracting the background spectrum of the peptide in the absence of metal (60 µM peptide and 50 mM TRIS buffer at pH 8.5). The difference molar extinction coefficients ($\Delta\varepsilon$) are calculated based on the total metal concentration.

b) pH Titrations. UV-Vis pH titrations were carried out at room temperature on an Ocean Optics SD 2000 fiber optic spectrometer and the pH measured using a mini-glass combination pH electrode (Hamilton Biotrode) coupled to a Fisher Accumet digital pH meter model 805 MP. pH titrations were performed by adding small aliquots of concentrated solution of KOH to unbuffered solutions containing $CdCl_2$ (20 µM) and peptide (60 µM), and monitoring the change in absorbance at 235 nm as a function of pH. Equilibration time was always allowed before reading the final pH. In all cases, reverse titrations were carried out by adding small aliquots of concentrated solution of HCl to verify the reversibility of the process. The UV-Vis pH titration curve of the peptide **Grand**L26AL28QL30C was analyzed using the model and the fitting equation published previously for the simultaneous release of two Cys thiol protons upon formation of tristhiolato Cd(II) species $(CdS_3)^-$ at high pH from a monothiolato Cd(II) complex $(CdS(SH)_2)^+$ that exists at low pH .[30,31]

^{113}Cd NMR Spectroscopy. All the spectra were collected at room temperature on a Varian Inova 500 spectrometer (110.92 MHz for ^{113}Cd) equipped with a 5 mm broadband probe. ^{113}Cd NMR spectra were externally referenced to a 0.1 M $Cd(ClO_4)_2$ solution in D_2O. A spectral width of 847 ppm (93,897 Hz) was sampled using a 5.0 µs 90^o pulse and 0.05 second acquisition time with no delay between scans. Samples were prepared under a flow of argon by dissolving 30 – 35 mg of the lyophilized and degassed peptides in 450 – 500 µL 15% D_2O/H_2O solution. The peptide concentrations were determined by using the assay with 4,4'-dipyridyl disulphide,[29] and the concentrations range from 9 to 18 mM peptide, which corresponds to 3 – 6 mM three-stranded coiled coil. The final samples were prepared by the addition of the appropriate amount of 250

mM $^{113}Cd(NO_3)_2$ solution (prepared from 95% isotopically enriched ^{113}CdO obtained from Oak Ridge National Laboratory) and the adjustment of the pH with KOH or HCl solutions. Whenever necessary, 500 mM stock solution of $Mg(NO_3)_2.6H_2O$ (Mallinckrodt Chemicals) and 250 mM stock solution of $Ca(NO_3)_2.4H_2O$ (Aldrich) prepared in 15% D_2O were added. The pH value was measured both before and after the experiment. An argon atmosphere was maintained when possible but the samples came in contact with air during addition of $^{113}Cd(NO_3)_2$, pH adjustment and acquisition. The data were analyzed using the software MestRe-C.[32] All free induction decays (FIDs) were zero filled to double the original points and were processed by application of 100 Hz line broadening prior to Fourier transformation.

Line Fitting Analysis of the ^{113}Cd NMR Spectra. Line fitting analysis of the ^{113}Cd NMR peaks was performed using the software MestRe-C by first defining the region of the spectra to be analyzed followed by selecting the peak of interest. Fittings were performed by keeping the chemical shift and the intensity of the peaks fixed while varying the line width (LW) of the peak during the iterative fitting process. The ratio of Lorentzian/Gaussian functions for each peak were kept constant at 1 (L/G = 1) during the fitting analysis. The line widths of the NMR peaks obtained from the analysis were used to calculate the exchange rate constants. Due to slow exchange, exchange rate constants (k_{ex}) were calculated using the equation for slow exchange regime k_{ex} = increase in LW (Hz) $\times$ 3.14. Whenever necessary an average increase in LW was calculated from the sets of spectra obtained during ^{113}Cd NMR titration and used to obtain the exchange rate constants.

^{1}H NMR Spectroscopy. 2D ^{1}H NMR experiments were performed at room temperature on a 500 MHz Varian Inova spectrometer equipped with an inverse detection probe. Samples were prepared by dissolving 30-35 mg of lyophilized and degassed peptide in 500-600 uL 10% D_2O/H_2O under a flow of argon. Final solutions contained 3.0 – 4.0 mM of peptide trimer, as determined by the assay

with 4,4'-dipyridyl disulphide.[29] Whenever necessary $^{113}Cd(II)$ was added in the form of 250 mM $^{113}Cd(NO_3)_2$ stock solution and pH was adjusted with concentrated KOH or HCl solutions as described in the previous section. Attempts were made to keep an argon atmosphere throughout, but the samples did come into contact of air during addition of $^{113}Cd(NO_3)_2$, pH adjustment and data acquisition. Nuclear Overhauser Effect Spectroscopy (NOESY) experiments were performed using standard Varian pulse sequences with mixing time of 100 ms.[33] For water suppression, pre-saturation was used for 1.5 s with saturation power of 5 dB.[34] An 80 ms spin lock was used during the experiment. The data were processed either by NMR Pipe,[35] or MestRe Nova.[36]

Results

UV-Vis Spectroscopy.

a) Stoichiometry. Stoichiometry of Cd(II) binding to the peptide **Grand**L26AL28QL30C was probed by monitoring the characteristic ligand-to-metal charge transfer transition (LMCT) at 235 nm due to the formation of Cd(II)-thiolate bonds. Addition of aliquots of 8.932 mM $CdCl_2$ stock solution to peptide solution containing 60 μM **Grand**L26AL28QL30C at pH 8.5, where the metal fully bound to the peptide based on the pH titrations, generated the characteristic LMCT bands. Figure 4-2 shows the UV-Vis spectra of **Grand**L26AL28QL30C obtained during the course of the titration. Inset of Figure 4-2 shows a plot of molar extinction coefficient ($\Delta\varepsilon$) at 235 nm vs. equivalents of Cd(II) added. The curve plateaus at 1.0 equivalent of Cd(II).

b) pH Dependence. Increase in pH of solutions containing 60 μM of **Grand**L26AL28QL30C peptide in the presence of 20 μM Cd(II) resulted in an increase of the characteristic LMCT bands. The resulting pH titration curve was obtained by plotting absorbance at 235 nm vs. pH (Figure 4-3) and was fit to a previously published model, simultaneous deprotonation of two Cys thiols leading

to the formation of Cd-trithiolato complex.[30,31] The resulting pK_{a2} value extracted from this analysis is 10.1 ± 0.2.

^{113}Cd NMR Spectroscopy.

Peptides Containing two Metal Binding Sites. First, ^{113}Cd NMR titrations were carried out with the dual site peptide **Grand**L12AL16CL26AL30C that is known to bind two Cd(II) ions as 4-coordinate pseudotetrahedral CdS_3O geometry at both the metal binding sites 12A16C and 26A30.[28] Gradual addition of 0.5 equivalent of ^{113}Cd(II) to a solution containing 3.2 mM (**Grand**L12AL16CL26AL30C)$_3$ at pH 8.5 led to an increase in intensity of the ^{113}Cd NMR resonances at 572 and 589 ppm up to 2.0 equivalents of ^{113}Cd(II) (Figure 4-4) (the most downfield peak marked with a star is an impurity from a peptide of shorter length that is produced during automated peptide synthesis and is inseparable by HPLC). The signals at 572 ppm and 589 ppm correspond to chemical shifts of the L12A16C and 26A30C sites, respectively.[28] In the presence of 2.0 equivalents (stoichiometric) of ^{113}Cd(II), the two peaks gained full intensity. Adding an additional 0.5 equivalent of ^{113}Cd(II) (2.5 equivalents total) resulted in a significant decrease in intensity and broadening of the peak at 589 ppm, indicating chemical exchange of Cd(II). The peak at 572 ppm underwent only slight decrease in intensity and broadening in the presence of 2.5 equivalents of ^{113}Cd(II). None of the peaks however, shifted their respective resonance frequencies, indicating that the process is occurring on the slow exchange regime. When 3.0 equivalents of ^{113}Cd(II) were added, the peak at 589 ppm had broadened beyond detection. The 572 ppm peak, on the other hand, was still present and did not broaden or lose further intensity. This observation suggests that the 26AL30C site of **Grand**L12AL16CL26AL30C peptide is undergoing site-selective chemical exchange of Cd(II). The increase in line width of the 589 ppm peak from 2.0 to 2.5 equivalents of ^{113}Cd(II) was 55 Hz, whereas for the peak at 572 the increase was 45 Hz., yielding rate constants of 173 ± 28.63 s^{-1} (5.8 ± 0.96 ms) (Table 2) and 140 ± 29.57 s^{-1} (7.1 ± 1.5 ms) (Table 2)

for the 589 and 572 ppm sites, respectively. Next, exchange studies were performed with the peptide **Grand**L16PenL19IL23PenL26I, which contains two metal binding sites and sequesters two of Cd(II) ions in trigonal planar CdS_3 geometry at both the binding sites (16 Pen and 23 Pen).[28] ^{113}Cd NMR titrations were carried out where 2.0, 2.5 and 3.0 equivalents of ^{113}Cd(II) were added to a solution containing 3.25 mM (**Grand**L16PenL19IL23PenL26I)$_3$ at pH 9.5. Two NMR resonances at 681 and 686 ppm were observed (Figure 4-5) during the course of the titration where no broadening of the peaks occurred when excess (2.5, 3.0 equivalents) ^{113}Cd(II) were added, indicating no chemical exchange was taking place for this peptide. These results show that the **Grand**L16PenL19IL23PenL26I peptide that encapsulates two Cd(II) ions with a 3-coordinate CdS_3 geometry has a slow exchange process, unlike the **Grand**L12AL16CL26AL30C peptide that sequesters two Cd(II) ions as 4-coordinate CdS_3O geometry.

Peptides Containing one Metal Binding Site. For further understanding of the exchange phenomenon and get more insight into the process, experiments were performed with simpler constructs containing a single metal binding site that sequester one equivalent of Cd(II) as a 4-corodinate CdS_3O geometry. The exchange behavior of the peptide **Grand**L26AL30C, which has been shown to bind one Cd(II) ion with a 4-coordinate CdS_3O geometry, was investigated.[37] First, ^{113}Cd NMR experiments were performed to determine whether any exchange could be observed under sub-stoichiometric amounts of ^{113}Cd(II). Addition of 0.2 to 0.8 equivalents of ^{113}Cd(II) to a solution containing 4 mM (**Grand**L26AL30C)$_3$ at pH 8.5 led to a gradual increase in the intensity of the ^{113}Cd NMR resonance with chemical shift of 587 ppm, while no line broadening of the peak was observed (Figure 4-6). These results show that under sub-stoichiometric amounts of ^{113}Cd(II), no measurable exchange of Cd(II) was taking place. Next, ^{113}Cd NMR experiments were carried out where 1.0 to 1.7 equivalents of ^{113}Cd(II) were added to a solution containing 3.71 mM (**Grand**L26AL30C)$_3$ at pH 8.5 (Figure 4-7). In the presence of a stoichiometric

(1.0 equivalent) amount of ^{113}Cd(II) this peptide has a single sharp resonance at 589 ppm. With the gradual addition of more than a stoichiometric amount of ^{113}Cd(II), the resonance at 589 ppm decreased in intensity and subsequently broadened. In the presence of 1.7 equivalents of ^{113}Cd(II), the resonance had broadened beyond detection. Furthermore, the resonance at 589 ppm did not shift during the course of ^{113}Cd(II) addition, showing that exchange is occurring on the slow exchange regime. Moreover, in the presence of excess Cd(II) no additional resonance was observed which could be attributed to the free ^{113}Cd(II), which is expected to appear in the up-field region between 0-100 ppm. This observation suggests that the exchange under investigation is not a simple two-site exchange process involving free Cd(II) in solution and Cd(II) bound at the interior of the coiled coil as CdS_3O, rather it is a multi-site exchange phenomenon. As the peptides under study are rich in glutamates at the surface, one can hypothesize that the surface Glu residue(s) can potentially be the other possible site(s) for Cd(II) in the multi-site exchange process. Fittings of the ^{113}Cd NMR spectra are shown in Figure 4-8. From the line fitting analysis the average increase in line width of the 589 ppm peak was determined to be 40.6 Hz, yielding an exchange rate constant of 127 ± 16 s^{-1} (7.9 ± 1 ms) (Table 2).

To test the hypothesis regarding the role of surface Glu residues as potential sites involved in the multi-site exchange process, Glu 28, located closest to the metal binding residue Cys 30 in the sequence of the **Grand**L26AL30C peptide, was mutated to non-metal binding residue Gln and exchange studies were performed with the resulting mutant **Grand**L26AE28QL30C. Aliquots of ^{113}Cd(II) were added to a solution containing 3.7 mM (**Grand**L26AE28QL30C)$_3$ at pH 8.5. In the presence of 1.0 equivalent of ^{113}Cd(II), a single sharp resonance with a chemical shift of 587 ppm is observed (Figure 4-9). The ^{113}Cd NMR resonance at 587 ppm broadened slowly with increasing amounts of ^{113}Cd(II) and is broadened beyond detection in the presence of 2.1 equivalents of ^{113}Cd(II), which is much higher than what was observed for the **Grand**L26AL30C peptide. The resonance for the latter peptide

was completely broadened beyond detection in the presence of 1.7 equivalents of ^{113}Cd(II). From line fitting analysis (Figure 4-10) the increase in line width of the 587 ppm peak was determined to be 29.3 Hz, resulting in a rate constant of 92 ± 28 s^{-1} (10.9 ± 3.3 ms) (Table 2).

To probe the effect of other divalent redox-inactive metal ions on exchange, experiments were performed where excess of $Mg(NO_3)_2$ and $(CaNO_3)_2$ were added to **Grand**L26AL30C under exchanging conditions of 1.1 equivalents of ^{113}Cd at pH 8.5 and the ^{113}Cd NMR spectra were recorded. Addition of up to 5 equivalents of Mg(II), to 3.3 mM (**Grand**L26AL30C)$_3$ in the presence of 1.1 equivalents ^{113}Cd(II) did not cause any broadening or shifting of the resonance at 588 ppm (Figure 4-11), indicating that excess Mg(II) has no effect on exchange. Similar experiments were performed where aliquots of Ca(II) were added up to 100 equivalents in the presence of 3.8 mM (**Grand**L26AL30C)$_3$ and 1.1 equivalents of ^{113}Cd(II). Figure 4-12 shows that during the course of titration the ^{113}Cd NMR resonance at 587 ppm is unperturbed, suggesting that excess Ca(II) also has no effect on the exchange process.

^{1}H NMR Spectroscopy.

To investigate whether the excess Cd(II) is interacting with other external sites of the protein, ^{1}H-^{1}H NOESY experiments were performed with **Grand**L26AL30C peptide in the presence of different equivalents of Cd(II). Addition of up to 3.0 equivalents of ^{113}Cd(II) to 3.4 mM (**Grand**L26AL30C)$_3$ at pH 8.5 did not disrupt the overall secondary structure of the peptide as shown by the well-dispersed chemical shifts in the 2D ^{1}H-^{1}H NOESY spectra (Figure 4-13 to 4-15) characteristic of a well-folded α-helical structure. When 1.0 equivalent Cd(II) was added new cross peaks appeared at 7.8, 7.88, 8.05, 8.14, 8.18, 8.21 and 8.71 ppm (Figure 4-13) along the F2 dimension of the NOESY spectrum indicating metal binding and subsequent structural/conformational changes to the protein. In the presence of 1.5 equivalents of Cd(II) cross peak movements were observed in the regions of 8.15, 7.88, 7.79 ppm of the NOESY spectra (Figure 4-

14). Subsequent addition of up to 3.0 equivalents of Cd(II) resulted in movement of cross peaks at 7.81, 7.85, 7.87, 8.13, and 8.47 ppm region of the NOESY spectra (Figure 4-15). As movements of the cross peaks were observed under more than stoichiometric amounts of Cd(II), this observation indicates that the excess Cd(II) is coordinating to external sites of the protein other than the Cys site.

Discussion

The objective of the study presented in this chapter is to understand how metal ions insert into and out of helical scaffolds at a molecular level using *de novo* designed three-stranded coiled coil peptides as model systems. Additionally, determination of the rates of metal ion exchange and the factors that influence these rates would help us better understand the metal insertion processes in more complex and native metalloproteins. I have investigated the exchange of the heavy metal Cd(II) into the designed peptides containing both two and single metal binding sites and exist as parallel three-stranded coiled coils in solution at pH higher than 5.5. The main focus of my study has been to address fundamental issues of metal exchange such as: does the Cd(II) initially coordinate to any amino acids located at the surface of the peptide? Do the cysteine residues at the designed metal binding site play any role in internalizing the surface bound Cd(II)? What are the rates of exchange? Does the rate depend on the location of the metal binding site within the three-stranded coiled coil interior (middle of the helix vs. towards the helical terminus)? Can one site fine-tune the rate of a distant second metal site within the same peptide? Is there any different in the exchange behavior of the 4-coordinate CdS_3O vs. 3-coordinate CdS_3 structure? While previous kinetic studies have been carried out by the Pecoraro group regarding the insertion of Hg(II) and Cd(II) into designed systems, these studies did not address the issues identified above.

Exchange Studies with Dual Site Peptides: Site-Selective Exchange of Cd(II).

The starting peptide used as the model to address these issues is **Grand**L12AL16CL26AL30C containing two metal binding sites that encapsulates two Cd(II) ions as 4-coordinate CdS_3O geometry at both the binding sites. The 12A16C site is located almost in the middle of the helix whereas the 26A30C site is located at the C-terminal end of the coiled coil assembly. Although detailed understanding of the exchange phenomenon might be complicated using this peptide as it contains two metal binding sites, this peptide would allow us to evaluate whether the two sites located at different positions along the linear sequence would show any difference in the exchange behavior of Cd(II). Said differently, would the position of the designed metal binding site within the interior of the coiled coil impact the kinetic properties of Cd(II) at the two sites? With addition of 0.5 equivalent ^{113}Cd(II) the resonances at 572 and 589 ppm increased in intensity. In the presence of 2.0 equivalents (stoichiometric) of ^{113}Cd(II) the two peaks had gained maximum intensity (Figure 4-4). Figure 4-4 also shows that under excess conditions of ^{113}Cd(II) the resonance with a chemical shift of 589 ppm is much more prone to exchange broadening, compared to the peak with chemical shift of 572 ppm. In the presence of 2.5 equivalents of ^{113}Cd(II) both the 572 and 589 ppm peaks decreased in intensity and broadened. The 589 ppm peak underwent a line broadening of 55 Hz whereas the corresponding 572 ppm peak broadened by 45 Hz. The corresponding rate constants were determined to be $173 \pm 28.63\ s^{-1}$ (5.8 ± 0.96 ms) and $140 \pm 29.57\ s^{-1}$ (7.1 ± 1.5 ms) (Table 2) for the 589 and 572 ppm sites, respectively, indicating a higher exchange rate constant for the 589 ppm peak compared to the 572 ppm peak. However, the resonance frequencies of none of the peaks shifted, indicating the occurrence of exchange on the slow exchange regime. In the presence of 3.0 equivalents of ^{113}Cd(II), the 589 ppm peak was broadened beyond detection while the 572 ppm peak was still present with almost similar intensity to that of 2.5 equivalents ^{113}Cd(II). These results tell us that the 26A30C site, having a chemical shift of

589 ppm,[28] and located close to the C-terminal end of the three-stranded coiled coil is more susceptible to chemical exchange of Cd(II) as compared to the 12A16C site (chemical shift 572 ppm) which is located almost in the middle of the helical scaffold. The 26A30C site, being located towards the C-terminal end of the three-stranded coiled coil is more exposed to the bulk solvent where the α helices are subject to significant fraying.[19,23,24,38] Thus, the Cd(II) bound at the 26A30C site being in a more solvent-exposed environment undergoes facile exchange with the Cd(II) present in bulk solvent compared to the Cd(II) bound at the 12A16C site, located in the middle of the coiled coil assembly in a more hydrophobic environment. These studies demonstrate that even though both the 12A16C and 26A30C sites in the **Grand**L12AL16CL26AL30C peptide bind Cd(II) in "identical" first coordination sphere geometries (4-coordinate CdS_3O), the 26A30C site is kinetically more labile and displays site-selective chemical exchange of Cd(II). This observation conveys the importance of the location of metal binding sites along the linear sequence and how the resulting dynamical properties of the bound metal ions can be fine-tuned depending on their position within the helical scaffold. Having understood how the position of metal binding sites impact the exchange dynamics of 4-coordinate Cd(II) (CdS_3O) structures in the dual site peptide **Grand**L12AL16CL26AL30C, exchange studies were performed with the peptide **Grand**L16PenL19IL23PenL26I. This construct also contains two metal binding sites and sequesters two Cd(II) ions as 3-coordinate CdS_3 geometry in both the Pen (Penicillamine) sites.[28] This peptide would allow us to address whether the 3-coordinate CdS_3 sites are also subject to chemical exchange of Cd(II) and whether there is any impact of the location of the metal sites on the resulting exchange properties of the metal ion. Figure 4-5 shows that the addition of more than stoichiometric amount of ^{113}Cd(II)) to this peptide does not cause any line broadening of either of the resonances with chemical shifts of 681 and 686 ppm, respectively, indicating that no exchange is occurring. These results demonstrate that the 3-coordinate CdS_3 sites are inert to chemical exchange unlike the 4-coordinate CdS_3O sites which are exchange labile. The

Pen and Ile residues in the **Grand**L16PenL19IL23PenL26I peptide provide improved packing conferring higher stability to the coiled coil assembly,[28] making the 3-coordinate Cd(II) sites more rigid and compact. In the **Grand**L12AL16CL26AL30C peptide, on the other hand, each Leu substitution to Cys or Ala results in a loss of free energy of ~ 4 kcal/mole,[28] making this an inherently less stable aggregate compared to the **Grand**L16PenL19IL23PenL26I peptide. For this peptide to do the exchange the Pen side chains must rotate within the helical interior. It would be much harder for Pen to rotate with the bulkier methyl groups which would slow down the exchange. Thus, this may be a kinetic effect, not necessarily related to the thermodynamic stability of the construct. This would mean that the on rate with Pen derivative is much slower than for Cys derivative which.

Investigation of Exchange in Single Site Peptides.

Exchange studies performed with the dual site peptide **Grand**L12AL16CL26AL30C have revealed that the 26A30C site of this peptide is more susceptible to chemical exchange of Cd(II) compared to the 12A16C site. To further elucidate and obtain a better understanding of the exchange phenomenon, studies were performed with simpler constructs containing one metal binding site that bind one Cd(II) ion in 4-coordinate CdS_3O geometry. Exchange experiments were carried out with the single site peptide **Grand**L26AL30C that contains the metal binding site in the same position in the sequence to that of the more dynamic 26A30C site in the corresponding dual site analogue **Grand**L12AL16CL26AL30C. First, studies were performed to see whether any chemical exchange could be observed under sub-stoichiometric concentrations of ^{113}Cd(II). Figure 4-6 shows that during the addition of 0.2 to 0.8 equivalents of ^{113}Cd(II) to the **Grand**L26AL30C peptide the peak with chemical shift of 587 ppm gained intensity. No broadening of the resonance was observed, indicating that no exchange was occurring under these conditions. Next, ^{113}Cd(II) NMR titrations were performed where 0.1 equivalents of excess ^{113}Cd(II) were

added at each point during the course of the experiment to follow the exchange behavior closely, and the resulting ^{113}Cd NMR spectra were recorded (Figure 4-7). In the presence of 1.0 equivalent ^{113}Cd(II) (stoichiometric) a single resonance with chemical shift of 589 ppm is observed. With gradual addition of excess ^{113}Cd(II), the resonance decreased in intensity with concomitant increase in line width. At 1.7 equivalents of ^{113}Cd(II) the peak had broadened beyond detection. During the course of the titration the position of the resonance did not shift, again showing that the exchange is occurring on the slow exchange regime, similar to the dual site peptide. From line fitting analysis (Figure 4-8) the increase in line width was determined to be 40.6 Hz, corresponding to an exchange rate constant of 127 $\pm$ 16 s^{-1} (7.9 $\pm$ 1 ms) (Table 2). This is an interesting result. While the same site (26AL30C) in the dual site peptide **Grand**L12AL16CL26AL30C had an exchange rate constant of 173 $\pm$ 28.63 s^{-1} (5.8 $\pm$ 0.96 ms) (Table 2), the corresponding single site peptide **Grand**L26AL30C has an exchange rate constant of 127 $\pm$ 16 s^{-1} (7.9 $\pm$ 1 ms) (Table 2) which is slower than the same site (26A30C) in the dual site peptide. These results demonstrate that the introduction of a second metal binding site (12A16C site) in the middle of the **Grand**L26AL30C peptide, has fine-tuned the dynamics of the 26A30C site located ~20 Å away in the resulting dual site construct **GrandL12AL16C**L26AL30C, leading to a faster Cd(II) exchange rate of the 26A30C site in the dual site peptide. These observations show how one metal site can influence the physical properties such as dynamics of metal encapsulation of a distant second metal site within the same helical scaffold.

Further examination of Figure 4-7 shows that during the course of excess Cd(II) addition, no peak corresponding to the free Cd(II) appeared, expected in the up field region of 0-100 ppm in the ^{113}Cd NMR spectrum. This observation suggests that the exchange under investigation is not a simple two-site exchange involving Cd(II) bound at the coiled coil interior as CdS_3O and free Cd(II) in solution, but it is a multi-site chemical exchange involving other potential sites of the polypeptide.

To examine if other sites of the protein are indeed involved in the exchange, 2D ^{1}H-^{1}H NOESY experiments were performed with **Grand**L26AL30C peptide with the addition of different equivalents of Cd(II). The NOESY spectra (Figure 4-13 to 4-15) of this peptide retained well dispersed cross peaks characteristic of a well-folded α-helical structure up to the addition of 3.0 equivalents of Cd(II), indicating the secondary structure of the peptide was not disrupted by the addition of excess Cd(II). When 1.0 equivalent of Cd(II) was added, new cross peaks appeared in the NOESY spectrum at 7.8, 7.88, 8.05, 8.14, 8.18, 8.21 and 8.71 ppm along the F2 dimension indicating metal binding to the protein and subsequent structural/conformational changes in the peptide (Figure 4-13). Several regions (7.81, 7.85, 7.87, 8.13, and 8.47 ppm) of the NOESY spectra showed movements of cross peaks during the course of addition of up to 3.0 equivalents of Cd(II) (Figure 4-14, 4-15). From the UV-Vis studies it has been proven that the **Grand**L26AL30C peptide binds one equivalent of Cd(II) as 4-corodinate CdS_3O at the coiled coil interior. The observation of peak movements in the NOESY spectra in the presence of more than 1.0 equivalent of Cd(II) would indicate that the excess Cd(II) is coordinating to other external sites of the protein.

What could be the other possible sites that Cd(II) can interact with and thus be involved in the multi-site exchange process? Other than the metal binding Cys residues the sequence of the peptide **Grand**L26AL30C is rich in the amino acids Glu, Lys, Ala and Leu (Table 1). Out of these different types of amino acid residues, Glu residues at the surface of the peptide are the most probable candidates that can potentially coordinate to the excess Cd(II) before the metal ion is inserted into the designed metal binding site at the protein interior. To test this hypothesis regarding the involvement of surface Glu residue(s) as other potential site(s) for Cd(II), Glu 28 located closest in sequence to the metal binding Cys residue at position 30, was mutated to a non-metal binding residue Gln. There are two reasons for choosing Glu 28 as the residue to be mutated. Firstly, it was assumed that as the Glu 28 is located closest in

sequence to the metal binding Cys 30, mutation of this Glu residue would have the most pronounced effect on the ^{113}Cd NMR chemical shifts, if it is involved in exchange. It is known that the Glu residues located at the **e** position of the heptad (5th) are involved in forming interhelical electrostatic interactions with the Lys residues located at **g** position (7th) of the heptad that are critical to the formation of parallel three-stranded coiled coils in solution. Glu 28 is located in the 6th position of the heptad (**f**) and residues located at **f** position of the heptads are not known to form interhelical salt bridge interaction with Lys that are essential for the formation of parallel three-stranded coiled coils.[22,39,40] Mutation of Glu 28 to Gln was not expected to disrupt the formation of parallel three-stranded coiled coils. However, Glu 28 can potentially be involved in intrahelical salt bridge interactions with other Lys residues. Thus, no additional Glu from either **e** or **f** position was mutated to Gln [22,39,40] so as to minimize the effect of mutation on the stabilizing electrostatic interaction.

E28Q Mutant: Possible Role of Glu 28 on Exchange.

First, the resulting mutant **Grand**L26AE28QL30C was characterized by UV-Vis spectroscopy to obtain the stoichiometry and pK_{a2} of Cd(II) binding to this peptide. Figure 4-2 shows that the UV-Vis spectral features (λ_{max} = 235 nm) and the molar extinction coefficients ($\Delta\varepsilon_{235\ nm}$ =) due to LMCT are similar to what has been previously observed for Cd(II)-tristhiolato complexes of other peptides studied by the group.[28,30,41] The inset of the figure shows that **Grand**L26AE28QL30C binds 1.0 equivalent of Cd(II) ion per trimer. Furthermore, the pK_{a2} of 10.1 $\pm$ 0.2 extracted from the fitting of the UV-Vis pH titration curve (Figure 4-3), corresponding to the binding of Cd(II) to the Cys thiols is consistent with the formation of a 4-coordinate CdS_3O complex,[31] and is similar to the pK_{a2} of 9.9 $\pm$ 0.2 for the **Grand**L26AL30C peptide. Thus the E28Q mutation did not perturb the stoichiometry and pK_{a2} of Cd(II) binding to the resulting mutant. Next, exchange studies were performed with the mutant peptide **Grand**L26AE28QL30C where excess of ^{113}Cd(II) were added and the resulting ^{113}Cd NMR spectra were

recorded throughout the course of titration. Under stoichiometric condition of ^{113}Cd(II) a single sharp resonance with a chemical shift of 587 ppm is observed (Figure 4-9). Under excess conditions of ^{113}Cd(II), Figure 4-9 shows that the exchange behavior of this peptide is different than what was observed for the **Grand**L26AL30C peptide. With gradual addition of excess ^{113}Cd(II), the ^{113}Cd NMR resonance at 587 ppm slowly decreased in intensity and broadened at higher excess of ^{113}Cd(II) compared to **Grand**L26AL30C. The peak did not shift in position, showing that the exchange is occurring on the slow exchange regime, similar to the **Grand**L26AL30C peptide. When 2.1 equivalents of ^{113}Cd(II) were added the resonance broadened beyond detection, as compared to the 1.7 equivalents of ^{113}Cd(II) needed for complete broadening of the NMR signal for the **Grand**L26AL30C peptide. The average increase in line width of the resonance is 29.3 Hz, resulting in a rate constant of 92 $\pm$ 28 s^{-1} (10.9 $\pm$ 3.3 ms) (Table 2) for the **Grand**L26AE28QL30C peptide. The fact that higher equivalents of ^{113}Cd(II) are required for complete broadening of the NMR signal for the **Grand**L26AE28QL30C peptide compared to that of **Grand**L26AL30C indicates that the surface Glu 28 residue plays an important role in the exchange of Cd(II) and is most likely one of the sites that the free Cd(II) interacts with prior to being inserted at the peptide interior. The slower exchange rate for the E28Q mutant compared to the **Grand**L26AL30C peptide would suggest that Glu 28 facilitates the insertion of Cd(II) into the interior of the three-stranded coiled coil by initially coordinating to the free Cd(II). Mutation of Glu 28 to the non metal binding residue Gln thus limits the interaction of Glu 28 with free Cd(II) making the insertion of the metal ion into the interior of the peptide less assisted, resulting in a slower exchange in **Grand**L26AE28QL30C.

Effect of Other Metal Ions on Exchange.

Studies were performed to investigate whether coordination of any other divalent redox-inactive metal ions such as Mg(II) and Ca(II) to the surface Glu residues can cause exchange broadening of the ^{113}Cd NMR signal in the

presence of exchanging conditions of ^{113}Cd(II). These metals have low polarizability (hard acids) and are not expected to interact with the soft (high polarizability) thiolate ligands of Cys residues at the interior of the coiled coils. Thus, experiments were performed under exchanging conditions of 1.1 equivalents of ^{113}Cd(II) per trimer of **Grand**L26AL30C and excess of Mg(II) and Ca(II) were added and the resulting ^{113}Cd NMR spectra were recorded. These experiments would allow us to test the effect of only the Mg(II) and Ca(II) ions on exchange as the amount of ^{113}Cd(II) is kept constant at throughout the experiments. Addition of up to 5 equivalents of Mg(II) did not cause any exchange broadening of the ^{113}Cd NMR signal at 588 ppm (Figure 4-11), indicating Mg(II) has no effect on exchange. However, as Mg(II) tends to exist in solution as hexahydrate $Mg(H_2O)_6$, [42] no definitive conclusion could be reached regarding the incorporation of the surface Glu carboxylate side chain into the first coordination sphere of Mg(II) and the resulting effect on exchange. Similar experiments with Ca(II) showed that the presence of as high as 100 equivalents of Ca(II) (Figure 4-12) also did not cause any line broadening of the ^{113}Cd NMR resonance at 587 ppm. Ca(II) is known to bind to Glu rich proteins with milimolar affinity,[43] and is expected to coordinate to the surface Glu residues of the peptide. Thus, of the fact that no exchange broadening of the^{113}Cd NMR signal is observed even in the presence of 100 equivalents of Ca(II) shows that binding of Ca(II) to the surface Glu residues does not affect the exchange behavior of the Cd(II) ion, invoking a Cd(II) ion specific exchange behavior of the peptide and shows that binding of other divalent metal ions such as Ca(II) to the surface Glu residues has no effect on the exchange of Cd(II).

Proposed Exchange Scheme.

Based on the results obtained from this study a plausible exchange scheme of Cd(II) is presented in Figure 4-16. From our previous studies we know that the peptides containing 4 heptads or more (TRI, GRAND and Coil Ser) exist as parallel three-stranded coiled coils even in the absence of metals at any pH

higher than 5.5.[19,22-24] **A** in Figure 4-16 represents situations where Cd(II) is bound to three cysteines at the interior of the three-stranded coiled coils under stoichiometric conditions. When free Cd(II) is present, it is initially coordinated by the Glu residue(s) at the surface located closest on sequence to the metal binding Cys residues in a fast process (**B**). This step of the scheme is supported by the exchange experiments performed with the E28Q mutant peptide as described before. Next, a slow step is envisioned that involves breaking of a Cd(II)-thiolate bond followed by reorientation of the Cys side chain from the coiled coil interior towards the helical interface which then binds to the Cd(II) ion initially coordinated to the surface Glu residue(s) (**C**). Such orientation of the Cys side chains towards the helical interface is supported by the X-ray structures of related polypeptide systems that have recently been reported in literature,[23,24] and would suggest that the alternate conformations of Cys side chains facilitate the movement of an internally bound Cd(II) ion to the helical interface and vice-versa. At this step the Cd(II) at the interior of the helices is coordinated to two Cys thiolates. This intermediate is closely represented by X-ray structure of a related peptide system Coil Ser where one Hg(II) ion is bound to two thiols of Pen (Penicillamine) as linear HgS_2 at the interior of the coiled coil, whereas another Hg(II) is present at the helical interface coordinated to a Glu and a Pen side chain oriented towards the helical interface (Figure 4-17) (Zastrow *et al.* unpublished results). The next step in the scheme involves the transfer of the Cd(II) ion bound at the interior (to two cysteines) to the helical interface where it would now be coordinated to one Cys and a Glu residue (**D**). As this step involves breaking of a Cd(II)-thiolate bond, it is expected to be slow. At this point no Cd(II) is present at the interior of the coiled coil. Next, we envision insertion of the Cd(II) ion, that was initially coordinated to the surface Glu and Cys residues at the helical interface (**C**), to the interior of the coiled coil assisted by the reorientation of the Cys side chain from the helical interface to the interior with concomitant formation of two Cd(II)-thiolate bonds. This step is thus expected to be a fast process. Finally, the reorientation of the third Cys side chain from the

helical interface to the interior would complete the exchange process and at this step the external Cd(II) is fully internalized and is bound to all three Cys thiolates. The proposed exchange scheme shows how the exchange of an internally bound Cd(II) ion can occur with an external Cd(II) ion. The scheme illustrates the possible roles of the surface Glu residue(s) which initially coordinate(s) to the free Cd(II) followed by anchoring of the metal ion by an alternate conformation of a Cys side chain at the helical interface. Movement of the internally bound Cd(II) to the exterior followed by the insertion of the external Cd(II) to the interior of the coiled coil assisted by the orientation of Cys side chain to the helical interior would then complete the exchange process. Because of the millisecond timescale of ^{113}Cd NMR spectroscopy that is being used to monitor the exchange of Cd(II), the most probable step of the exchange scheme that is reporting the life time broadening of the ^{113}Cd NMR peaks, would be the slow step involving the inter conversion of the species **B** and **C**.

Prof. Erik Zuiderweg has developed a mathematical model that addresses the exchange process. The model involves a high-affinity Cd(II) binding in the Cys site, followed by low-affinity binding in the Glu site. A five-site NMR exchange scheme with potentially different chemical shifts of Cd(II) has been developed involving free Cd(II), Cd(II), bound at the Cys site, Cd(II) bound at the Glu site and Cd(II) bound to both the Cys and Glu site. Simulations of NMR spectra have been performed involving this five-site chemical exchange broadening. The proposed slow-exchange model and the simulations are capable of predicting the experimentally observed behavior of Cd(II) binding to the Cys site. The mathematical model predicts that more NMR lines must be observed when excess Cd(II) is added. However, as described before the NMR signal for free Cd(II) in H_2O cannot be experimentally observed. It is therefore possible that Cd(II) in the other binding sites is experimentally invisible as well.

Conclusions

In this chapter I have investigated the exchange of Cd(II) into *de novo* designed three-stranded coiled coil peptides that contain both dual and single metal binding sites. Chemical exchange of Cd(II) is observed when the metal ion is bound as 4-corodinate CdS_3O structure in pseudotetrahedral geometry. The corresponding 3-coordinate trigonal planar CdS_3 structures are inert to chemical exchange as the exchange involving the Pen side chains is kinetically slow. The exchange of Cd(II) from the CdS_3O site is observed to be occurring on the slow exchange regime with the exchange rates determined to be in the range of 5.8 – 10.9 ms. Exchange studies performed with the peptide **Grand**L12AL16CL26AL30C that encapsulates two Cd(II) ions as 4-corodinate CdS_3O geometry at both the binding sites, show that the 26AL30C site, located proximal to the helical terminus is more dynamic and susceptible to chemical exchange of Cd(II), compared to the 12A16C site located at the middle of the helical scaffold. This demonstrates that the dynamics of the same metal bound in "identical" coordination environments (CdS_3O) can be different depending on the location of the metal sites along the linear sequence. Studies with the corresponding single site peptide **Grand**L26AL30C show that the exchange rate of Cd(II) in this peptide is slower than that of the 26A30C site in the dual site peptide **Grand**L12AL16CL26AL30C. This observation suggests that introduction of a metal binding site (12A16C) in the middle of the helix has fine tuned the dynamics of a second metal site located ~20 Å away, close to the helical terminus. A multi site exchange scheme is proposed where the free Cd(II) is believed to be interacting with the surface Glu residues prior to being inserted to the interior of the coiled coil. The initial interaction of the metal ion with the Glu residues and the alternate conformations of Cys side chains are thought to be facilitating the exchange process.

References

(1) Masuoka, J.; Saltman, P. *Journal of Biological Chemistry* **1994**, *269*, 25557-25561.
(2) Means, A. R.; Dedman, J. R. *Nature* **1980**, *285*, 73-7.
(3) Rosenzweig, A. C.; O'Halloran, T. V. *Curr. Opin. Chem. Biol.* **2000**, *4*, 140-147.
(4) Xia, W.; Li, H.; Sze, K.-H.; Sun, H. *Journal of the American Chemical Society* **2009**, *131*, 10031-10040.
(5) Rosenzweig, A. C. *Accounts of Chemical Research* **2000**, *34*, 119-128.
(6) O'Halloran, T. V.; Culotta, V. C. *J. Biol. Chem.* **2000**, *275*, 25057-25060.
(7) Wernimont, A. K.; Huffman, D. L.; Lamb, A. L.; O'Halloran, T. V.; Rosenzweig, A. C. *Nat. Struct. Biol.* **2000**, *7*, 766-771.
(8) Wright, J. G.; Natan, M. J.; MacDonnell, F. M.; Ralston, D. M.; O'Halloran, T. V. *Progress in Inorganic Chemistry:Bioinorganic Chemistry* **1990**, *38*, 323-412.
(9) Wright, J. G.; Tsang, H.-T.; Penner-Hahn, J. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2434-2435.
(10) Ye, J.; Kandegedara, A.; Martin, P.; Rosen, B. P. *J. Bacteriol.* **2005**, *187*, 4214-4221.
(11) Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Cavet, J. S.; Dennison, C.; Graham, A. I.; Harvie, D. R.; Robinson, N. J. *Journal of Biological Chemistry* **2007** *282* 30181-30188
(12) Busenlehner, L. S.; Weng, T. C.; Penner-Hahn, J. E.; Giedroc, D. P. *J. Mol. Biol.* **2002**, *319*, 685-701.
(13) Busenlehner, L. S.; Cosper, N. J.; Scott, R. A.; Rosen, B. P.; Wong, M. D.; Giedroc, D. P. *Biochemistry* **2001**, *40*, 4426-2236.
(14) Shi, W.; Dong, J.; Scott, R. A.; Ksenzenko, M. Y.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 9291-9297.
(15) Sun, Y.; Wong, M.; Stalhandske, C.; Scott, R. A.; Rosen, B. P. *FASEB J.* **1999**, *13*, A1464-A1464.
(16) Borremans, B.; Hobman, J. L.; Provoost, A.; Brown, N. L.; van der Lelie, D. *J. Bacteriol.* **2001**, *183*, 5651-5658.
(17) Dieckmann, G. R., Ph.D. Thesis, University of Michigan, 1995.
(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.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 6195-6196.
(19) Touw, D. S.; Nordman, C. E.; Stuckey, J. A.; Pecoraro , V. L. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969-11974.
(20) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2006**, *45*, 2864-2868.
(21) Farrer, B.; Pecoraro, V. L. *Proc. Natl. Acad. Sci., U.S.A.* **2003**, *100*, 3760-3765.
(22) Dieckmann, G. R.; McRorie, D. K.; Lear, J. D.; Sharp, K. A.; DeGrado, W. F.; Pecoraro, V. L. *J. Mol. Biol.* **1998**, *280*, 897-912.

(23) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angewandte Chemie International Edition* **2009**, *48*, 7371-7374.
(24) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *Journal of the American Chemical Society* **2010**, *132*, 13240.
(25) Ghosh, D. Ph.D., University of Michigan, 2006.
(26) Chan, W. C.; White, P. D. *Fmoc Solid Phase Peptide Synthesis: A Practical Approach*; Oxford Univ. Press: New York, 2000.
(27) Farrer, B. T.; Harris, N. P.; Balchus, K. E.; Pecoraro, V. L. *Biochemistry* **2001**, *40*, 14696-14705.
(28) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro , V. L. *Journal of the American Chemical Society* **2010**, *Submitted*.
(29) Mantle, M.; Stewart, G.; Zayas, G.; King, M. *Biochem. J.* **1990**, *266*, 597-604.
(30) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Biol. Inorg. Chem.* **2006**, *11*, 876-890.
(31) Iranzo, O., Lee, K.H., Jakusch, T., Hemmingsen, L., Pecoraro, V. L. *Chem. Eur. J.,* **2009**, *15*, 3761-3772.
(32) Cobas, C.; Cruces, J.; Sardina, F. J.; 2.3 ed.; Universidad de Santiago de Compostela, Spain, 2000.
(33) Jeener, J.; Meier, B. H.; Bachmann, P.; Ernst, R. R. *The Journal of Chemical Physics* **1979**, *71*, 4546-4553.
(34) Hoult, D. I. *Journal of Magnetic Resonance* **1976**, *21*, 337-347.
(35) Delaglio, F.; Grzesiek, S.; Vuister, G. W.; Zhu, G.; Pfeifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277-293.
(36) MestReNova, v. X., Mestrelab Research S.L., Santiago de Compostela,; Spain *www.mestrelab.com* **2009**.
(37) Iranzo, O.; Cabello, C.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2007**, *46*, 6688-6691.
(38) Lovejoy, B.; Choe, S.; Cascio, D.; McRorie, D.; DeGrado, W.; Eisenberg, D. *Science* **1993**, *259*, 1288-1293.
(39) Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2004**, *43*, 7902-7915.
(40) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959-9973.
(41) Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2002**, *124*, 8042-8054.
(42) Maguire, M. E.; Cowan, J. A. In *BioMetals*; Springer Netherlands, 2002; Vol. 15.
(43) Haber-Pohlmeier;, S.; Abarca-Heidemann;, K.; Ko¨rschen;, H. G.; Dhiman;, H. K.; Heberle;, J.; Schwalbe;, H.; Klein-Seetharaman;, J.; Kaupp;, U. B.; Pohlmeier, A. *BIOPHYS J* **2007**, *92*, 3207-3214.
(44) http://131.104.156.23/Lectures/CHEM_207/CHM_207_NMR.htm.

Table 4-1. Peptide sequences used in this study.

Peptide	Sequence
Baby	Ac-G LKALEEK LKALEEK LKALEEK G-NH_2
BabyL9C	Ac-G LKALEEK **C**KALEEK LKALEEK G-NH_2
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
TRIL9C	Ac-G LKALEEK **C**KALEEK LKALEEK LKALEEK G-NH_2
GRAND	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK LKALEEK G-NH_2
GRANDL26AL30C	Ac-G LKALEEK LKALEEK LKALEEK LKA**A**EEK **C**KALEEK G-NH_2
GRANDL26AE28QL30C	Ac-G LKALEEK LKALEEK LKALEEK LKA**A**E**Q**K **C**KALEEK G-NH_2
GRANDL12AL16CL26AL30C	Ac-G LKALEEK LKA**A**EEK **C**KALEEK LK**A**LEEK **C**KALEEK G-NH_2
GRANDL16PenL19IL23PenL26I	Ac-G LKALEEK LKALEEK **X**KA**I**EEK **X**KA**I**EEK LKALEEK G-NH_2

X = Penicillamine (Pen). Residues in red represent changes from the parent sequence.

Table 4-2. Line width of ^{113}Cd NMR peaks for different peptides in the presence of different equivalents of added ^{113}Cd(II).

Peptide	Eq. of Cd(II)	^{113}Cd NMR δ (ppm)	LW (Hz)	Exchange Rate (k_{ex}) s^{-1}
GRANDL26AL30C	1.0	589	432	
	1.1	589	470	
	1.2	589	538	
	1.3	589	570	126 ± 16
	1.4	589	606	
	1.5	589	634	
GRANDL26AE28QL30C	1.0	587	378	
	1.1	587	407	
	1.2	587	401	
	1.3	587	413	92 ± 28
	1.4	587	518	
	1.5	587	524	
	2.0	587	737	
GRANDL12AL16CL26AL30C	2.0	572, 589	303 (572 ppm)	
			392 (589 ppm)	140 ± 30 (572 ppm)
	2.5	572, 589	348 (572 ppm)	
			447 (589 ppm)	173 ± 29 (589 ppm)
	3.0	572	350 (572 ppm)	

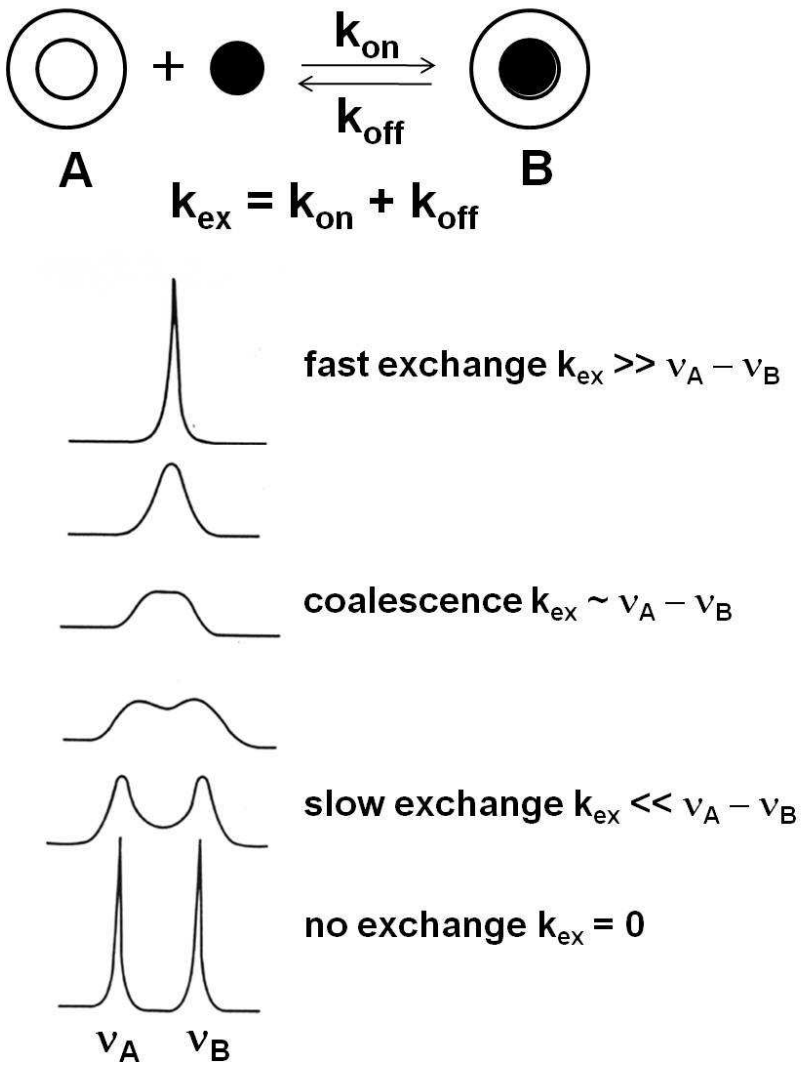


Figure 4-1. Illustration of a chemical exchange process involving ligand binding to a protein. A is represented as a protein, black sphere as a ligand and B is the resulting protein-ligand complex. Shown are the changes in the NMR spectra from slow to fast exchange between the species A and B. The NMR spectra are taken from ref [44].

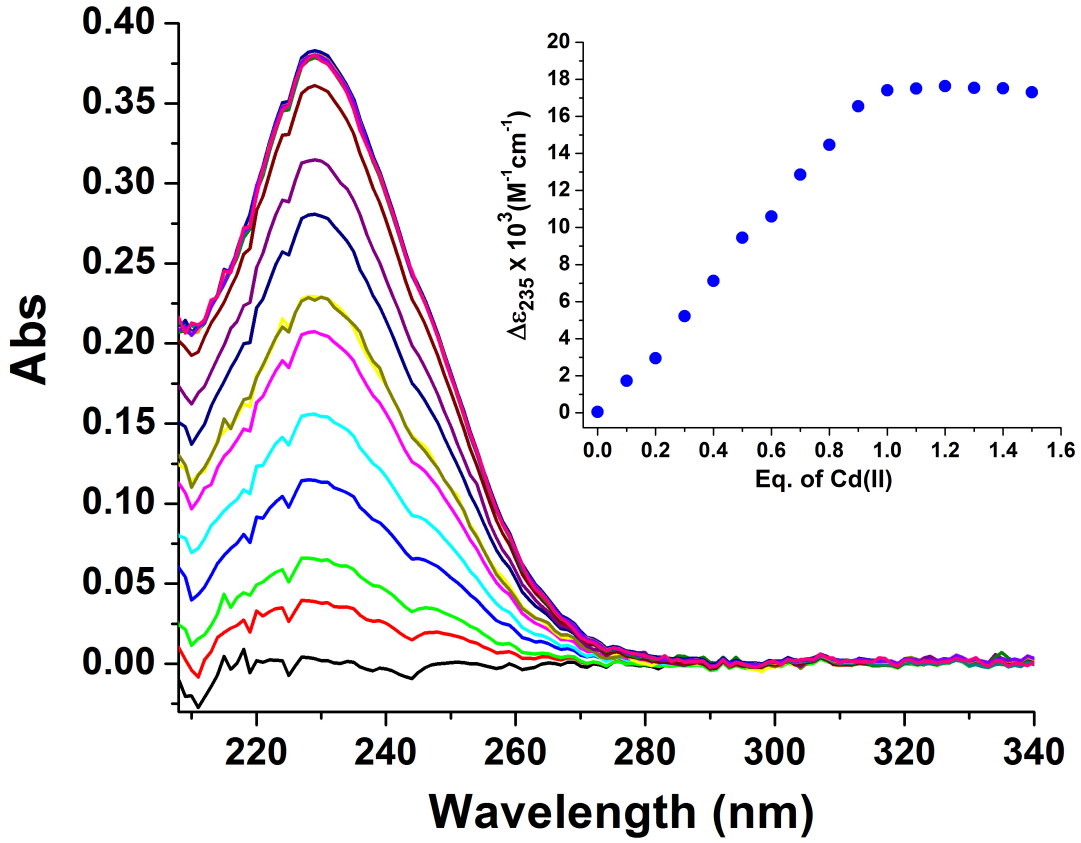


Figure 4-2. UV-Vis titration of $CdCl_2$ to a solution containing 60 μM GrandL26AE28QL30C in 50 mM TRIS buffer at pH 8.5. Shown is a plot of Abs. vs. wavelength (nm). The inset of the figure shows a plot of $\Delta\varepsilon_{235nm}$ as a function of Cd(II).

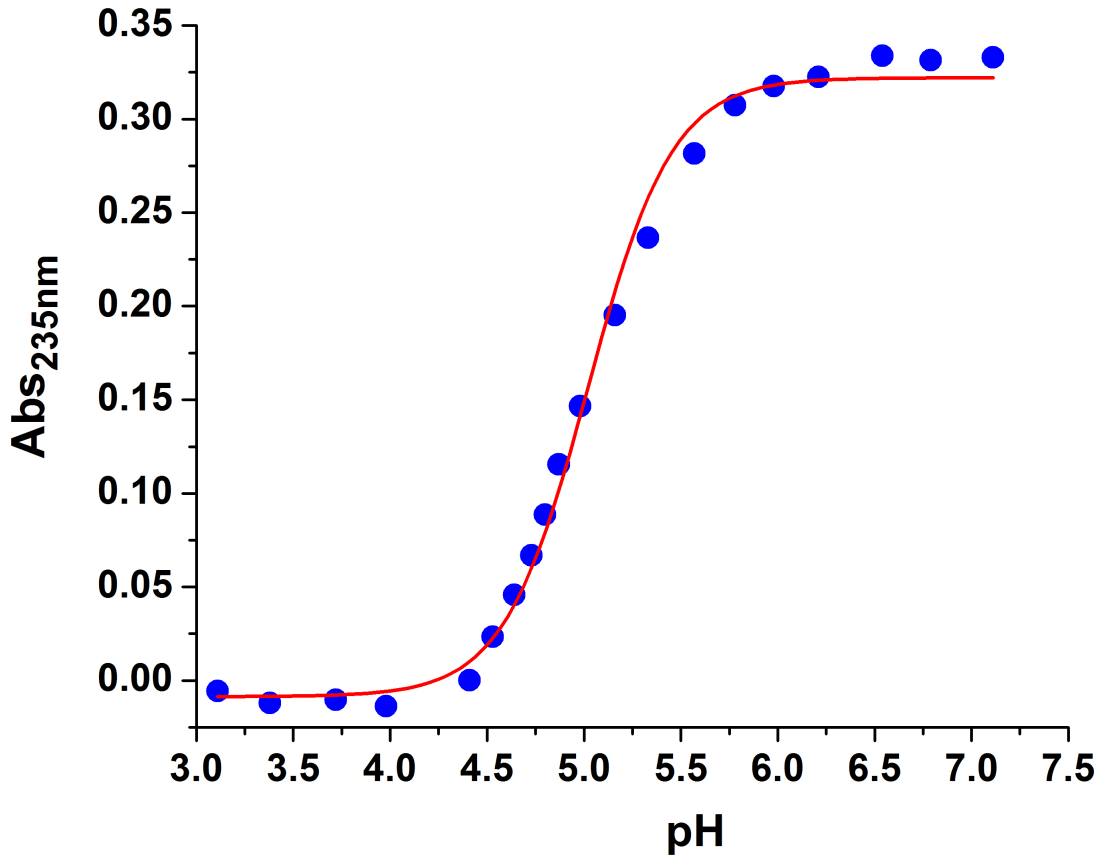


Figure 4-3. pH dependence of the binding of 1 equivalent $CdCl_2$ to 60 μM GrandL26AL28QL30C along with the fit of the experimental data. Experiment was followed by monitoring UV-Vis absorbance at 235 nm. Titration data were fit to the model simultaneous deprotonation of two Cys thiols from $CdS(SH)_2$ to CdS_3.

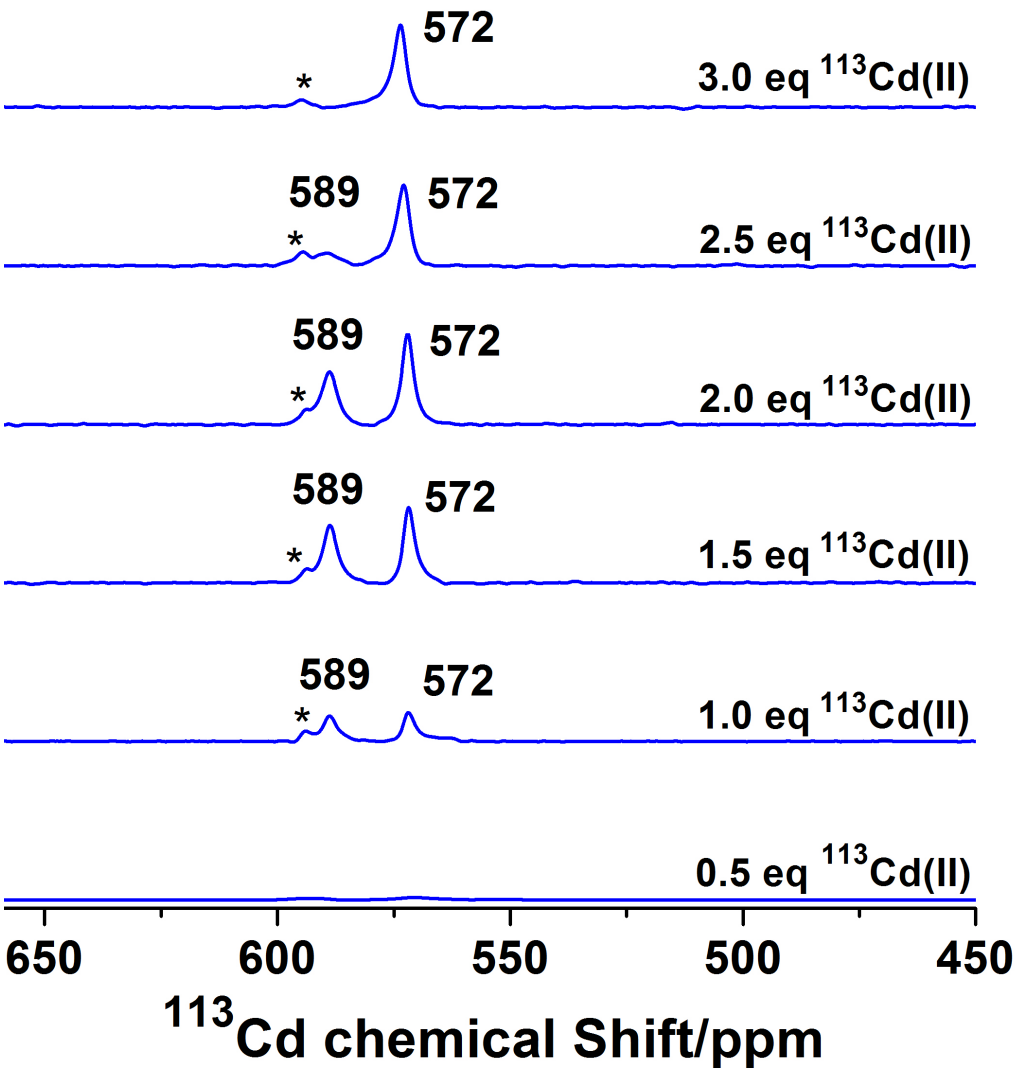


Figure 4-4. ^{113}Cd NMR spectra of a solution containing 3.2 mM (**Grand**L12AL16CL26AL30C)$_3$ at pH 8.5 with different equivalents of added $^{113}Cd(NO_3)_2$. In the presence of more than stoichiometric amount of ^{113}Cd(II), 589 ppm peak broadened significantly more compared to the 572 ppm peak, showing site-selective chemical exchange of Cd(II) for the 26A30C site. The peak marked with a star is an impurity most likely in the form of a peptide of shorter length that is produced during automated peptide synthesis and is inseparable by HPLC.

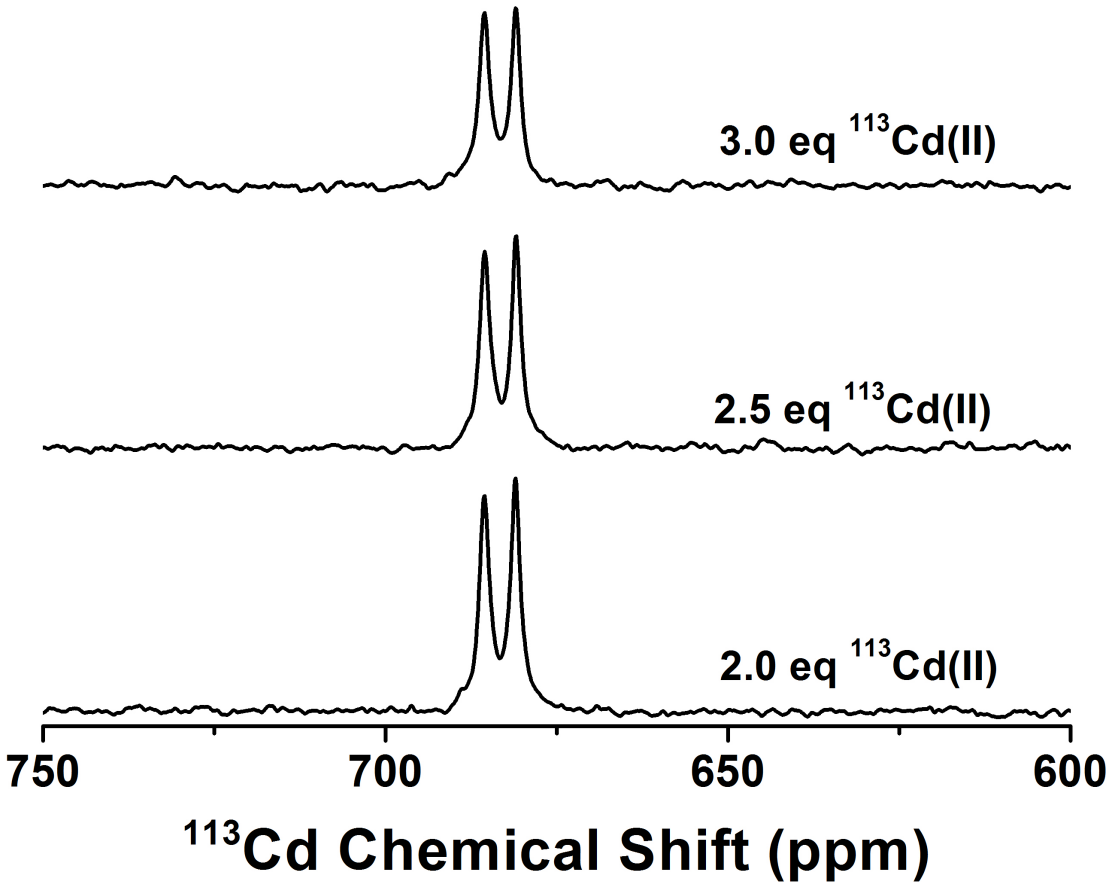


Figure 4-5. ^{113}Cd NMR spectra of a solution containing 3.25 mM (**Grand**L16PenL19IL23PenL26I)$_3$ at pH 9.5 with different equivalents of added ^{113}Cd(NO$_3$)$_2$. None of the peaks broaden during the course of addition of ^{113}Cd(II), indicating no exchange.

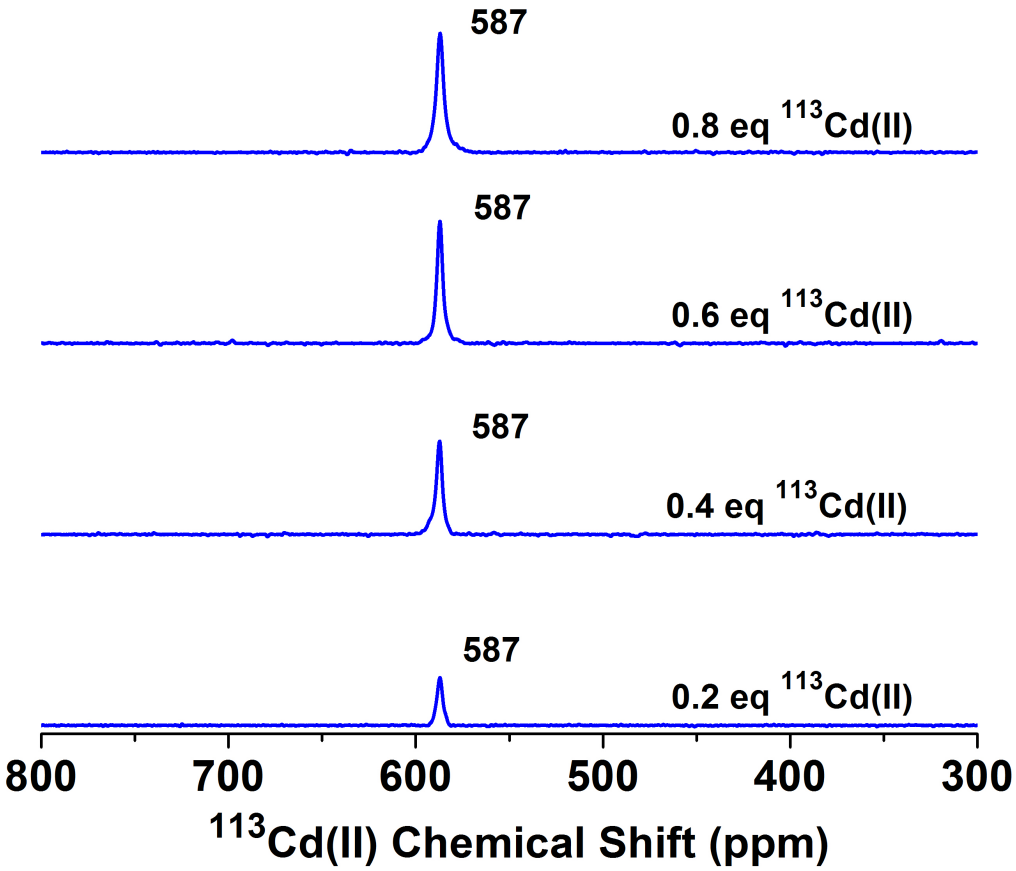


Figure 4-6. ^{113}Cd NMR spectra of a solution containing 4.0 mM (**Grand**L26AL30C)$_3$ at pH 8.5 in the presence of 0.2 to 0.8 equivalents of added ^{113}Cd(NO$_3$)$_2$. The resonancne with chemical shift of 587 ppm gained intensity with increase in the amount of added Cd(II), while no broadening of the peak was observed, indicating chemical exchange was not occurring under sub-stoichiometric amounts of Cd(II).

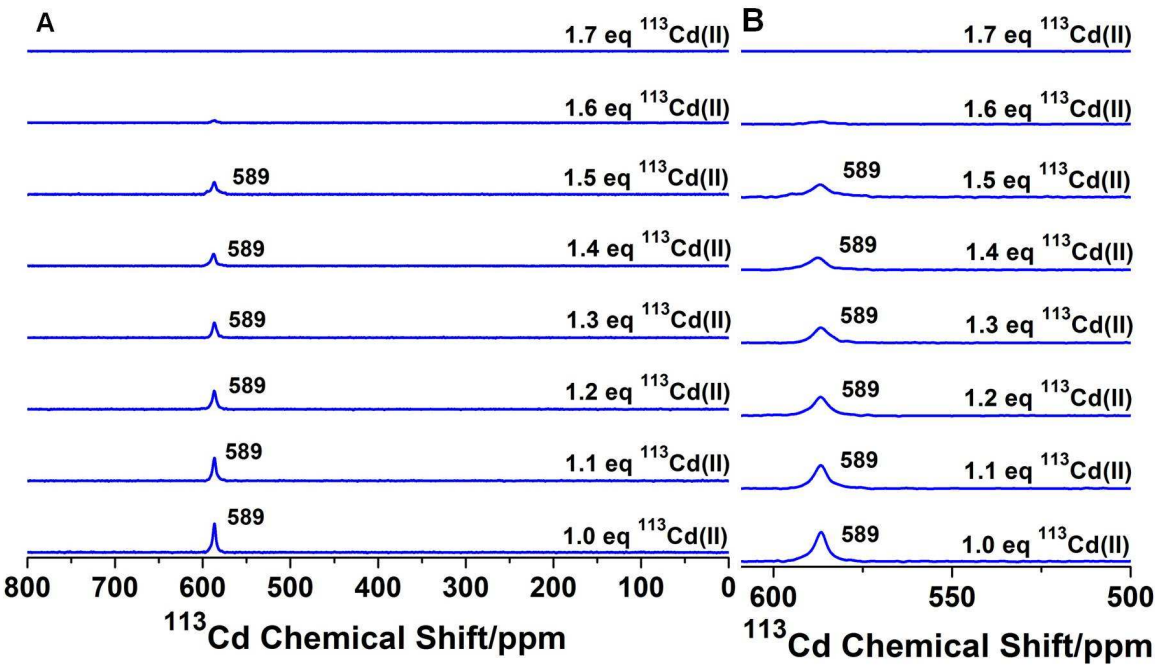


Figure 4-7. **A)** ^{113}Cd NMR spectra of a solution containing 3.7 mM (**Grand**L26AL30C)$_3$ at pH 8.5 with different equivalents of added $^{113}Cd(NO_3)_2$. The critical region of the spectra is enlarged and shown in **B**.

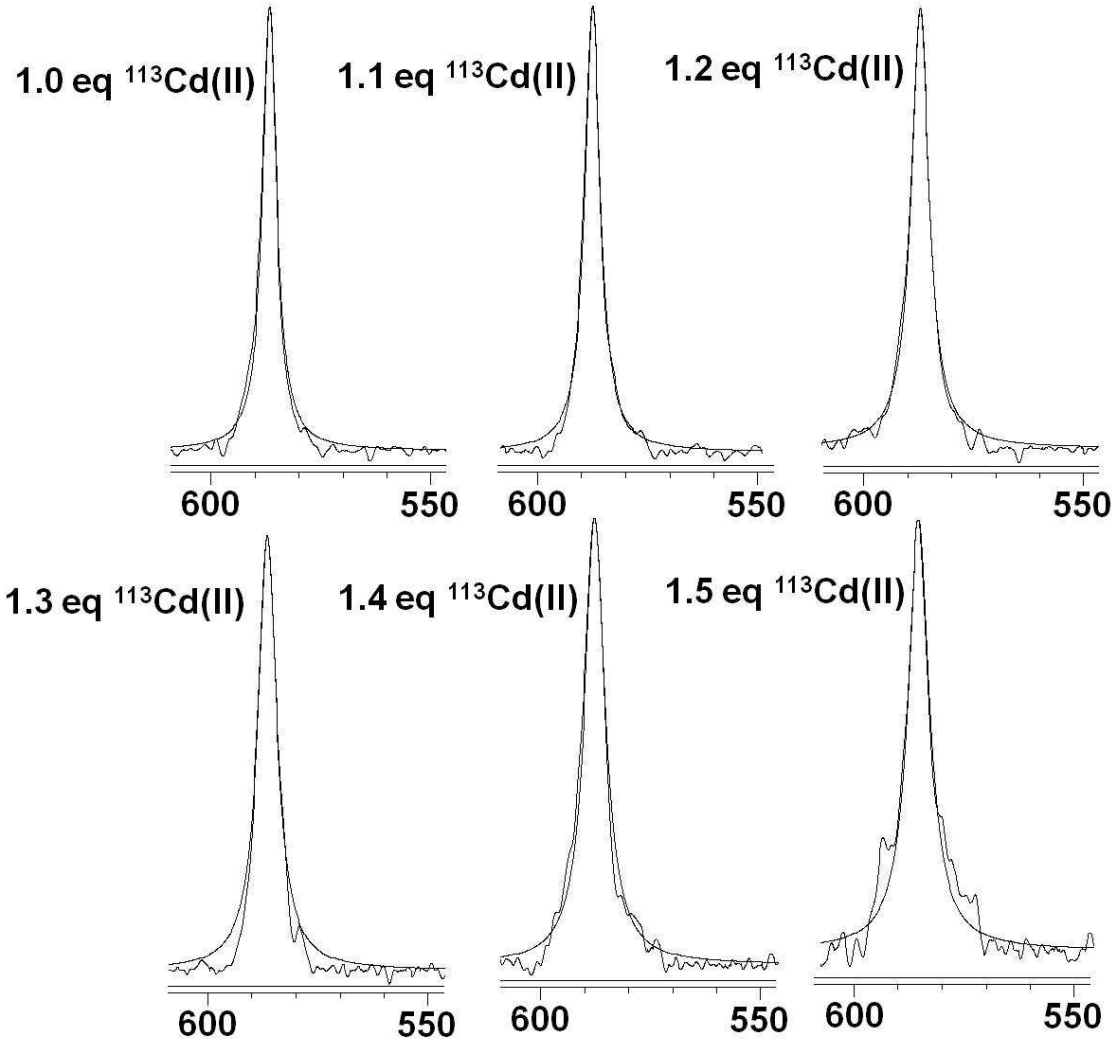


Figure 4-8. Fittings (smooth lines) of the ^{113}Cd NMR spectra of **Grand**L26AL30C obtained under different equivalents of added $^{113}Cd(NO_3)_2$. Shown are the absolute intensities of the spectra along with the fits performed with a constant Lorentzian/Gaussian ratio of 1.

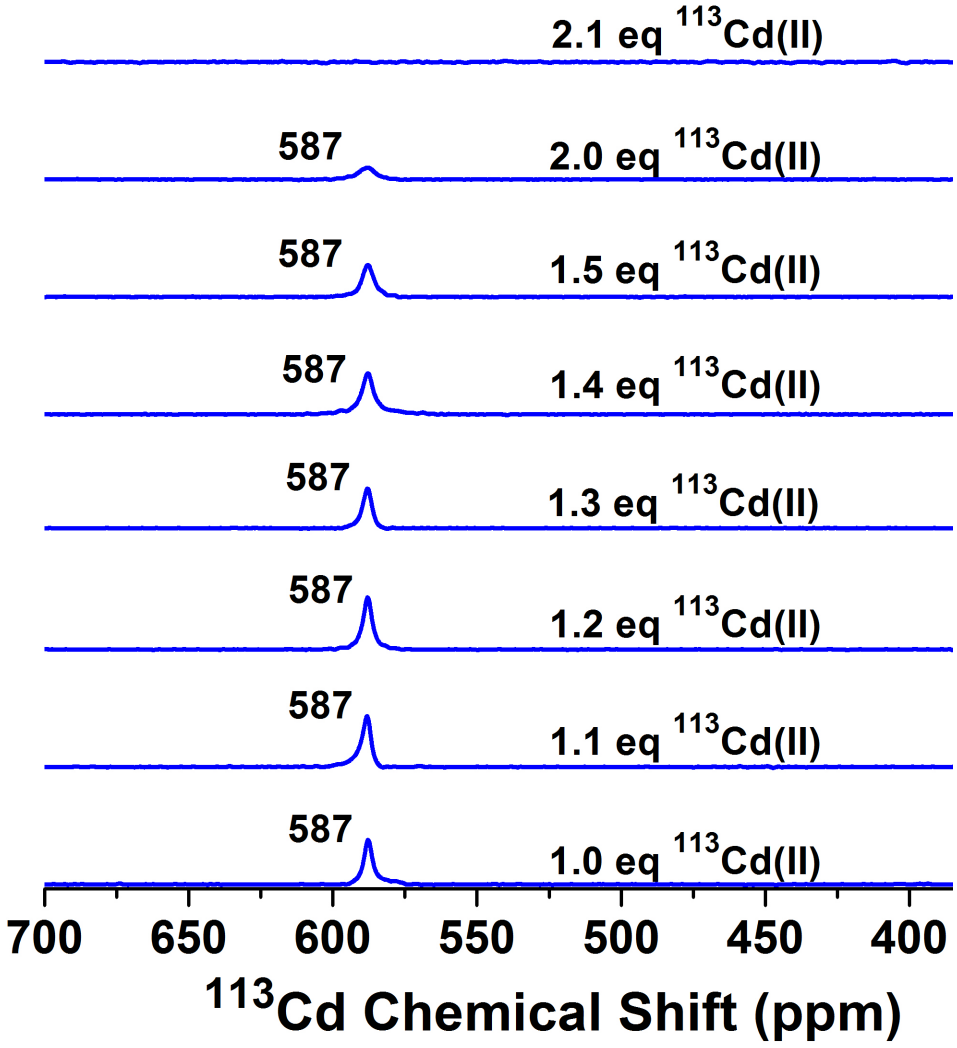


Figure 4-9. ^{113}Cd NMR spectra of a solution containing 3.7 mM (**Grand**L26AE28QL30C)$_3$ at pH 8.5 with different equivalents of added $^{113}Cd(NO_3)_2$. The ^{113}Cd NMR peak slowly decreased in intensity and broadened. In the presence of 2.1 equivalents of added ^{113}Cd(II), the resonance broadened beyond detection.

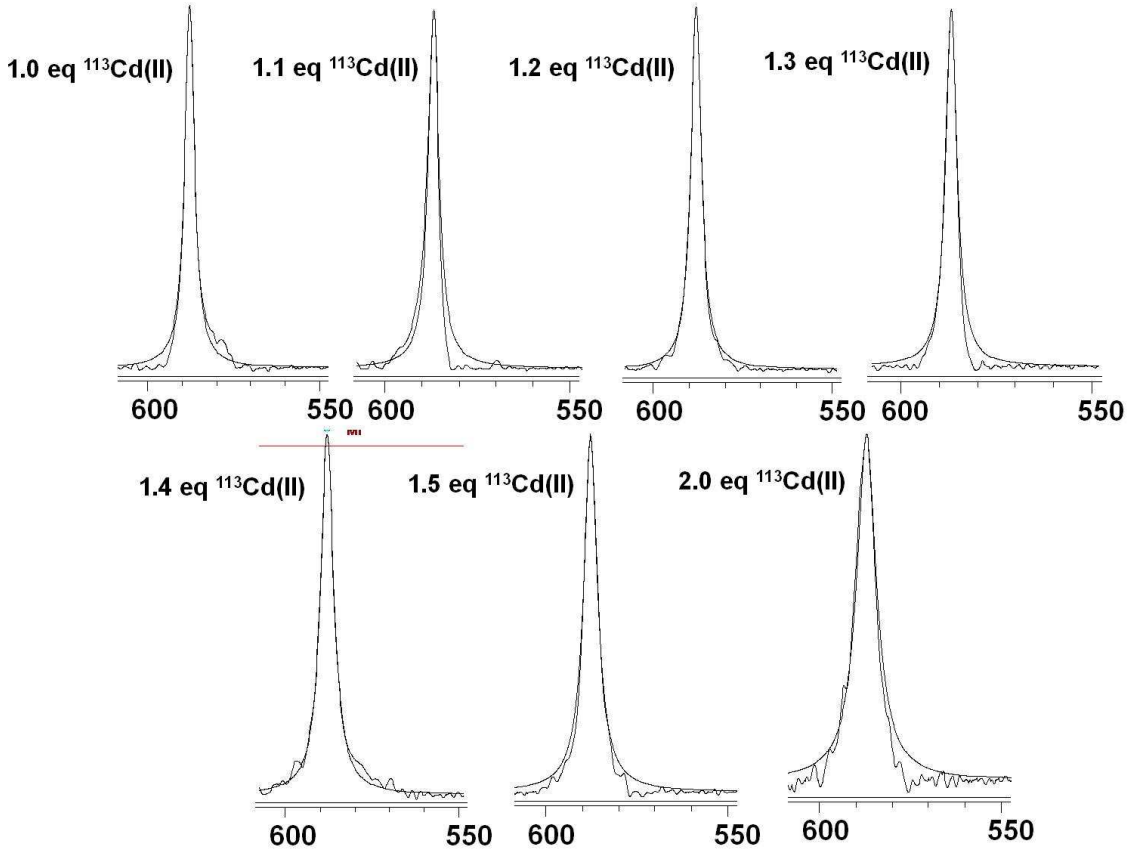


Figure 4-10. Fittings (smooth lines) of the ^{113}Cd NMR spectra of **Grand**L26AE28QL30C obtained under different equivalents of added $^{113}Cd(NO_3)_2$. Shown are the absolute intensities of the spectra along with the fits performed with a constant Lorentzian/Gaussian ratio of 1.

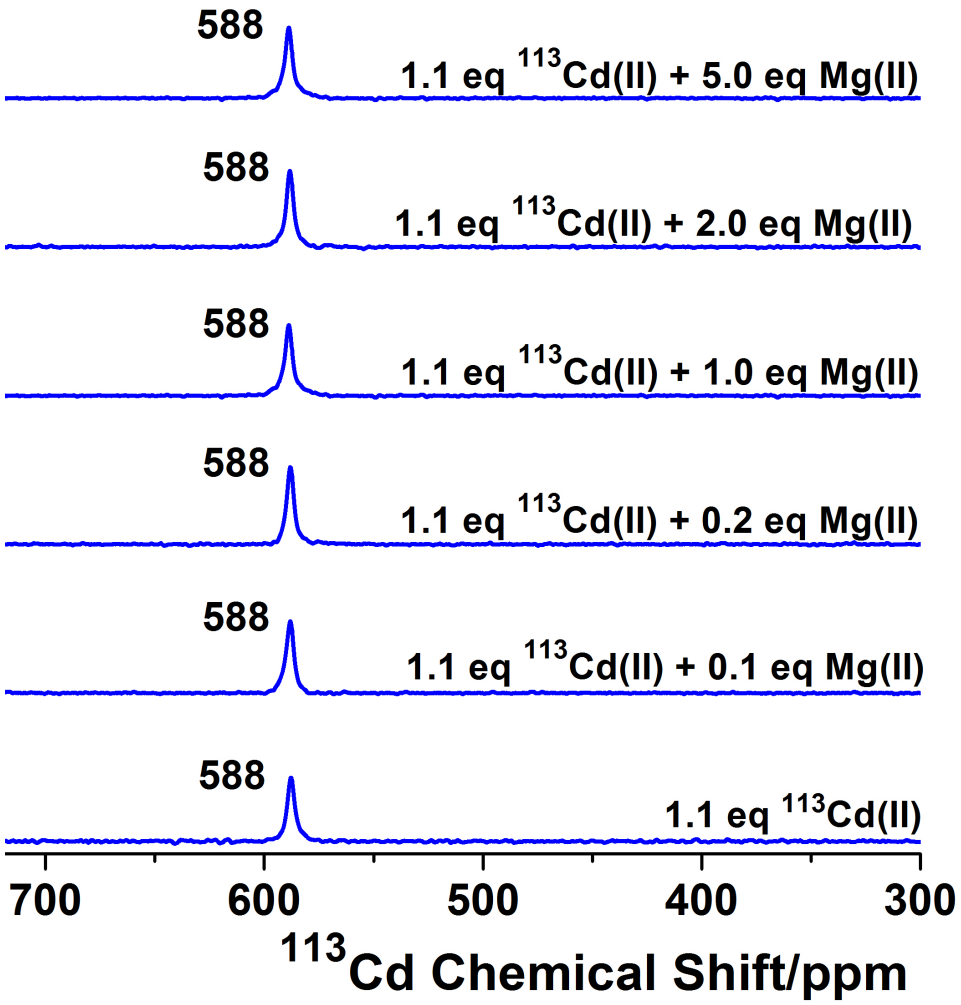


Figure 4-11. ^{113}Cd NMR spectra of a solution containing 3.3 mM (**Grand**L26AL30C)$_3$ and 1.1 equivalents of ^{113}Cd(II) at pH 8.5 with different equivalents of added $Mg(NO_3)_2$. The peak at 588 ppm did not broaden in the presence of excess Mg(II), indicating excess Mg(II) has no effect on exchange.

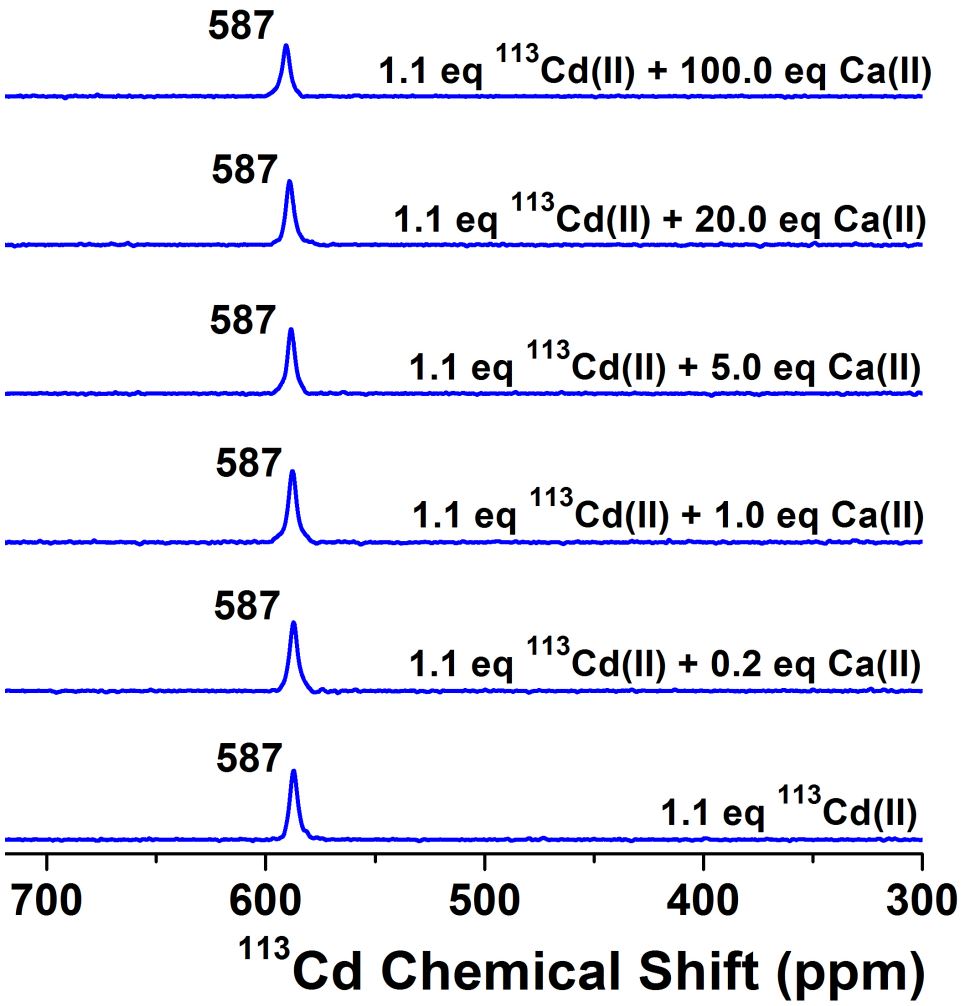


Figure 4-12. ^{113}Cd NMR spectra of a solution containing 3.8 mM (**Grand**L26AL30C)$_3$ and 1.1 equivalents of ^{113}Cd(II) at pH 8.5 with different equivalents of added $Ca(NO_3)_2$. The peak at 587 ppm did not broaden in the presence of excess Ca(II), indicating excess Ca(II) has no effect on exchange.

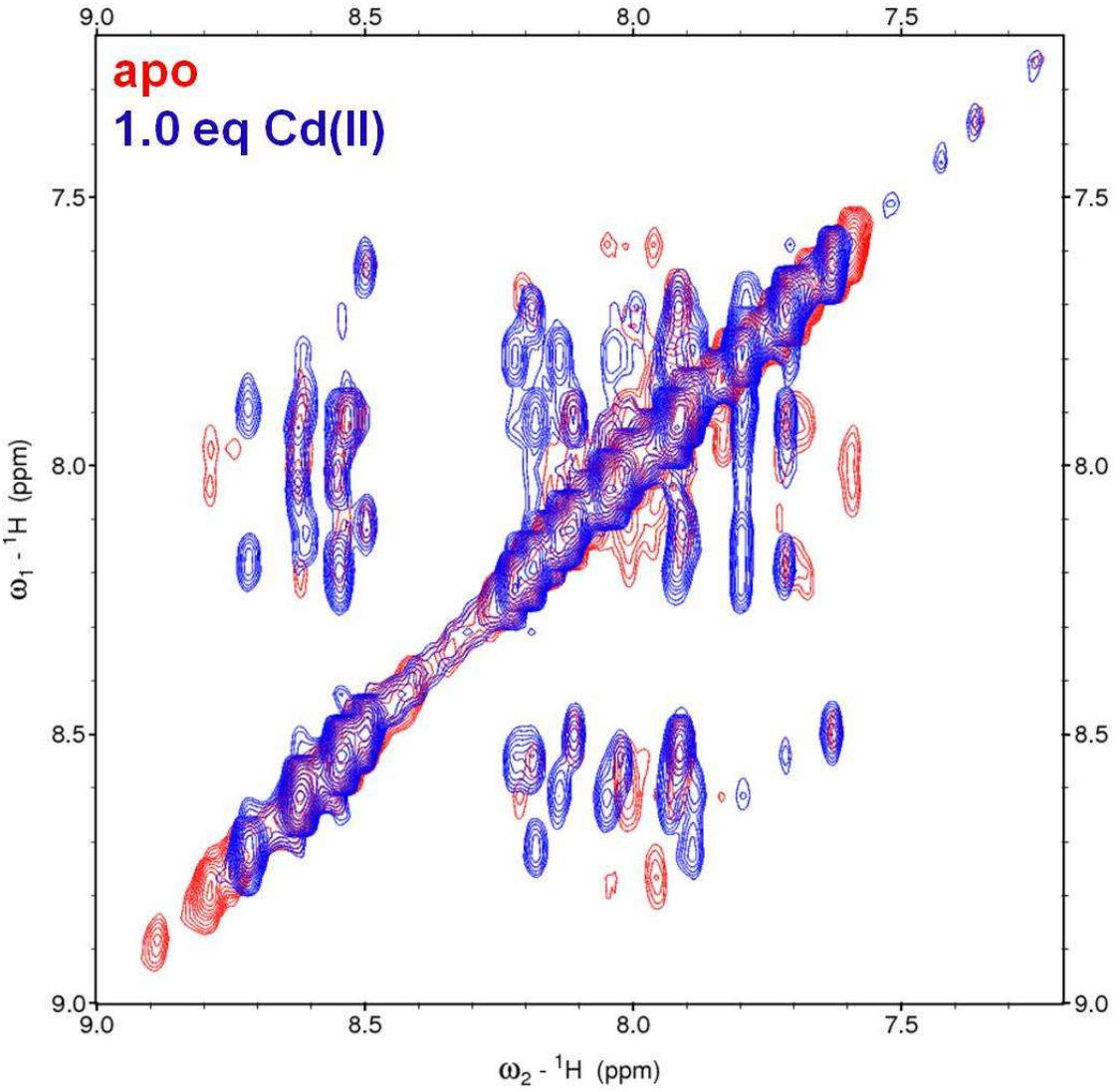


Figure 4-13. Overlay of NOESY spectra of 3.4 mM apo (**Grand**L26AL30C)$_3$ (red) and in the presence of 1.0 equivalent of Cd(II) (blue) at pH 8.5. New cross peaks appeared at 7.8, 7.88, 8.05, 8.14, 8.18, 8.21 and 8.71 ppm in the presence of 1.0 equivalent Cd(II) indicating metal binding and subsequent structural/conformational changes.

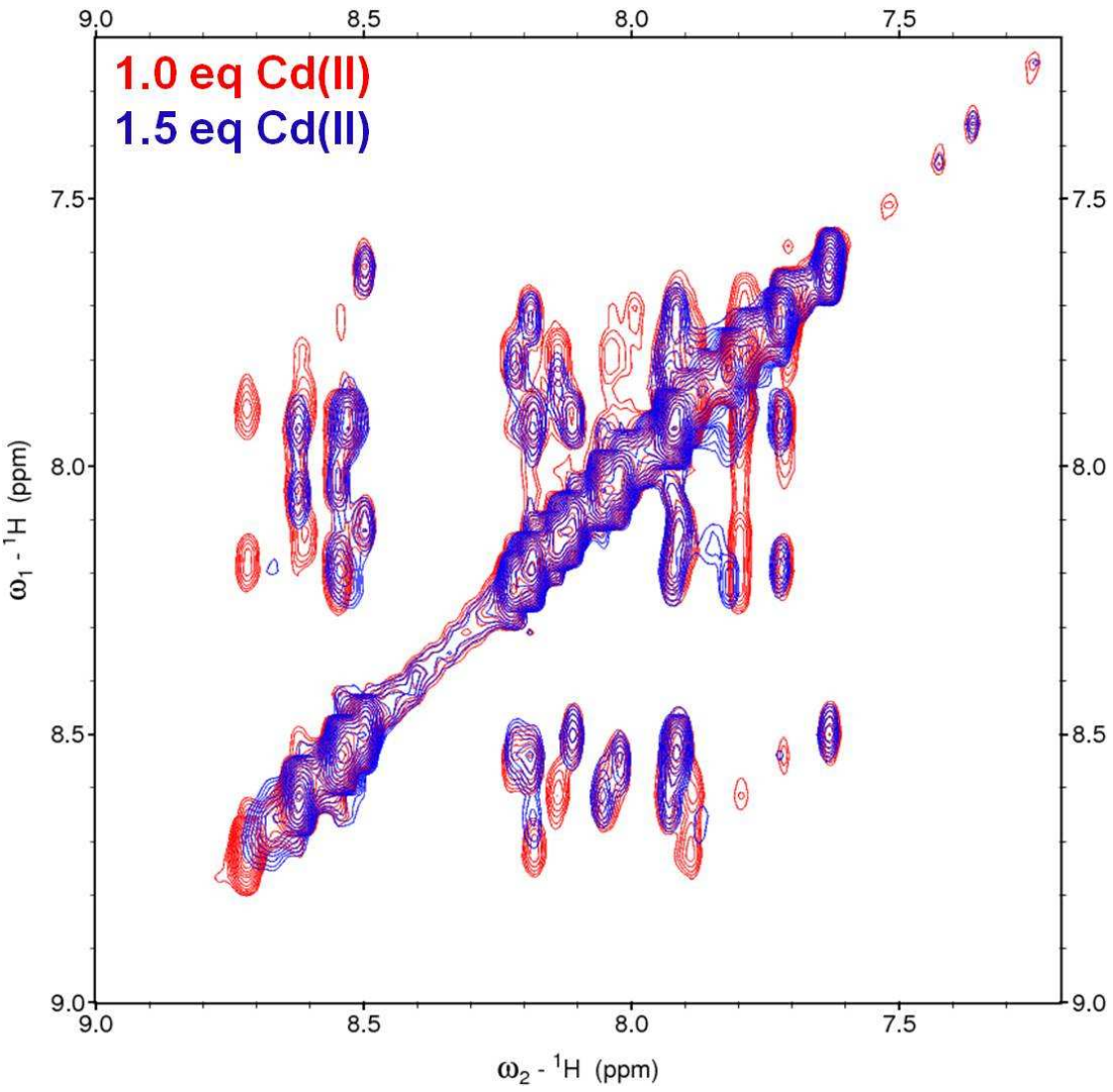


Figure 4-14. Overlay of NOESY spectra of 3.4 mM (**Grand**L26AL30C)$_3$ in the presence of 1.0 (red) and 1.5 equivalent Cd(II) (blue) at pH 8.5. Cross peak movements are observed at 8.15, 7.88, 7.79 ppm region.

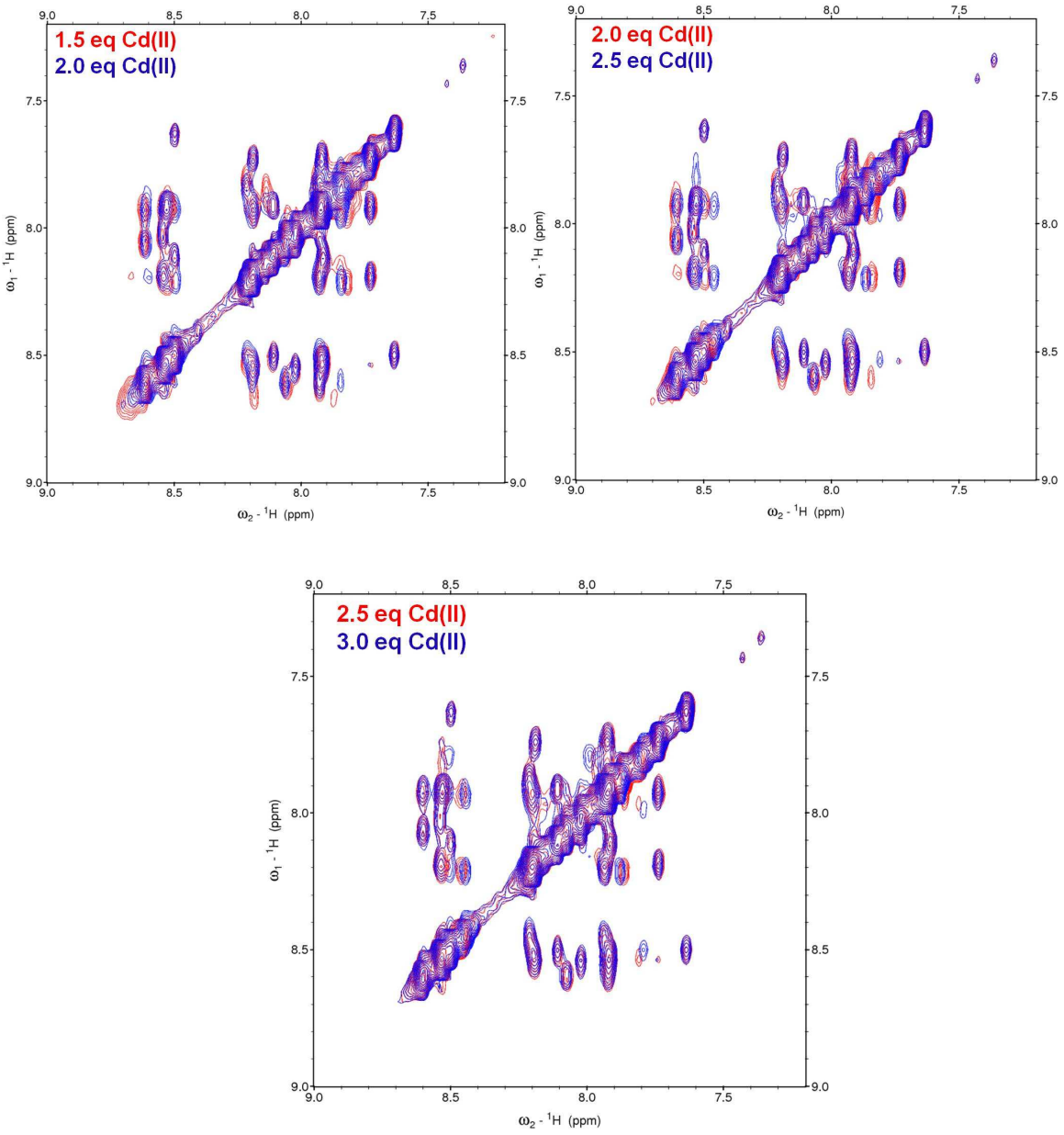


Figure 4-15. NOESY spectra of 3.4 mM (**Grand**L26AL30C)$_3$ in the presence of 1.5 to 3.0 equivalents Cd(II) at pH 8.5. Cross peak movements are observed at 7.81, 7.85, 7.87, 8.13, and 8.47 indicating that excess Cd(II) binds to other external sites of the peptide.

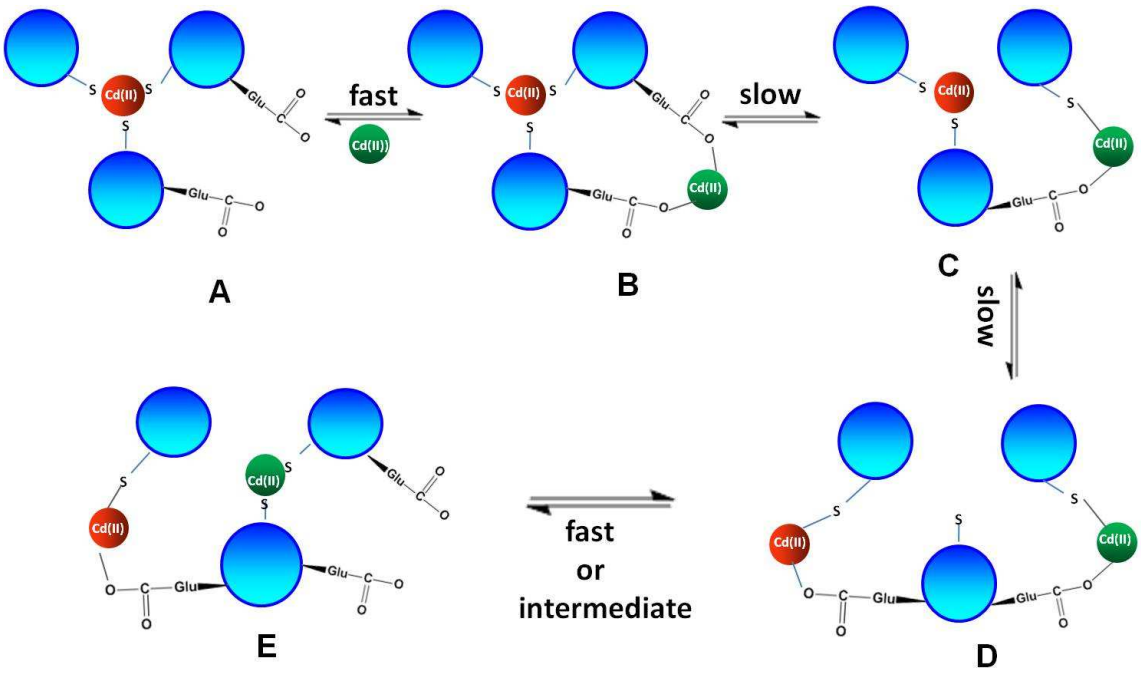


Figure 4-16. Proposed exchange scheme of insertion of Cd(II) into and out of helical scaffolds. **A** represents Cd(II) bound at the interior of three-stranded coiled coils under stoichiometric conditions. Free Cd(II) initially coordinates to the surface Glu residues (**B)**. A Cys side chain orients towards the helical interface and coordinates to the Cd(II) initially bound to Glu residues (**C**). Internally bound Cd(II) is then transferred to the helical interface and coordinates to Glu residue(s) (**D**). Insertion of the externally bound Cd(II) assisted by the reorientation of the Cys side chain towards helical interior then internalizes the external Cd(II) ion (**E**). At this step exchange of an internal Cd(II) ion occurs with an external Cd(II) ion. Reorientation of the third Cys side chain towards the interior would then lead to the formation of Cd-tristhiolato complex at the coiled coil interior.

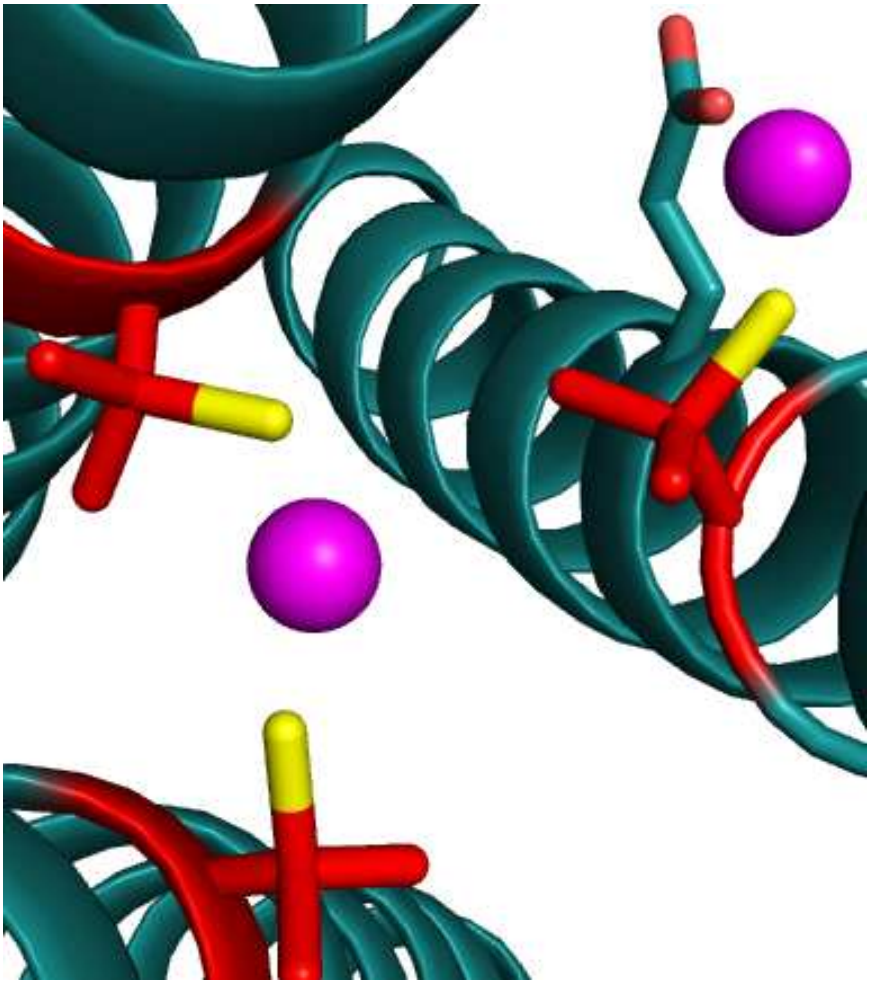


Figure 4-17. X-ray structure of **CS**L9PenL23H where one Hg(II) ion is bound to two thiols of Pen (Penicillamine) as linear HgS_2 at the interior of the coiled coil, whereas another Hg(II) is present at the helical interface coordinated to a Glu and a Pen side chain oriented towards the helical interface. This figure represents intermediate C of the exchange scheme shown in Fig 4-16. Protein backbone is shown as blue, Pen side chains as red/yellow sticks, Hg(II) ions as purple. The structure is solved and refined by Melissa Zastrow.

Chapter 5

Realization of a Designed Three-Helix Bundle Capable of Binding Heavy Metals in a Tris(Cysteine) Environment

Introduction

An important objective of *de novo* protein design has been the preparation of metalloproteins, as many natural systems contain metals which play crucial roles for the function and/or structural integrity of the biopolymer.[1,2] Metalloproteins catalyze some of the most important processes in nature ranging from energy generation and signal transduction to complex chemical transformations. At the same time, metals in excess can be deleterious to cells and some ions are purely toxic, having no known beneficial effects (e.g., Hg^{II}, Cd^{II} or Pb^{II}). One hopes to use a first principles approach for realizing both known metallocenters and also to prepare novel sites which may lead to exciting new catalytic transformations. However, designing novel metalloproteins is a challenging and complex task, especially if one desires to prepare asymmetric metal environments in the first coordination sphere of metals containing more than one type of amino acid ligands.

Numerous metalloprotein systems have been designed over the past fifteen years. These proteins typically use unassociated peptides that aggregate into parallel three stranded coiled coils or helix-loop-helix motifs that assemble to

form anti-parallel 4 stranded bundles capable of binding hemes,[3,4] and non-heme mononuclear,[5]and binuclear centers.[6,7] Because of the symmetry of these systems relying on homo-oligomerization, it is often difficult to prepare non-symmetrical metal sites through these strategies. Preparing metalloproteins with asymmetric metal binding sites would allow us to mimic active sites of many natural proteins such as the blue copper proteins plastocyanin,[8,9] azurin,[10-12] introducing a single specific hydrogen bonding interaction to the metal site such as in carbonic anhydrase,[13,14] as well as making catalytic metalloproteins. There has been an attempt to design a heterotrimeric coiled coil intended to form a two-His, one-Cys trigonal site by DeGrado *et. al*.[15] Even though the intended heterotrimer was formed, the construct was unable to incorporate metal ions. The only successful incorporation of metal ions in a heterooligomeric system is the heterotetrameric assembly of DF2tet $A_aA_bB_2$ designed by DeGrado *et. al.* which in the presence of Fe(II) and dioxygen formed a peroxo-bridged diferric species.[16] Due to very limited success in preparing heterooligomeric systems involving single self-associating peptides, preparing a single polypeptide chain capable of binding metals in predefined coordination geometries has been a highly sought-after objective by which to impart proteins with novel properties and activities.[17-22] Thus, the preparation of a single polypeptide chain, capable of controlling a metal coordination environment in a defined manner, has become a sought after objective. However, preparing novel metal binding sites into *de novo* designed single-chain helical bundles have proven to be challenging.

From our research over the last decade using TRI and GRAND peptides we have learned about designing soft, thiol-rich metal binding sites into the interior of parallel, three-stranded α-helical coiled coils involving cysteine and/or penicillamine as the ligating amino acid residues.[23,24] These systems, designed based on the heptad repeat (**a → g**), have served as hallmarks for understanding metallobiochemistry of different heavy metals such as Cd^{II}, Hg^{II}, As^{III}, and Pb^{II},[23-26] and delineated salient features of the interaction of these heavy metals with different metalloregulatory proteins such as MerR, CmtR, ArsR/SmtB.[27-29] From

the protein design perspective we have shown how to control the geometry and coordination number of metals such as Cd^{II} and Hg^{II} at the protein interior as well as fine-tuning the physical properties of the metals bound within the hydrophobic core of the polypeptide and achieving site-selective molecular recognition of Cd^{II}.[30-32] Unfortunately, the production of heterotrimeric systems that are capable of fine tuning metal environments controllably has been elusive.[33] For example, we have not designed asymmetric metal binding environments in the first coordination sphere of the metals involving different types of amino acids as ligating residues nor has it been possible to introduce a single secondary shell interaction to orient a single specific hydrogen bond within the first coordination sphere of the metal because when the α helices self assemble to form parallel, homomeric three-stranded coiled coils, three such interactions are introduced. Therefore, we have chosen an alternative strategy to satisfy this objective using a single polypeptide chain instead of multiple self associating peptides.

Existing heteromeric helical bundles and coiled-coils show energetic preferences of several kcal/mol for the desired heteromeric versus homomeric assemblies.[34,35] However, the energy gap between hetero and homomeric assemblies often depends critically on ionic strength, pH, and other environmental parameters. Moreover, the objective of many studies in *de novo* protein design is to impart an energetically sub-optimal coordination geometry on the metal ion, and the degree to which this strategy will be successful depends on the size of this energy gap between the desired heteromeric assembly and other homomeric or mis-folded states. Even when heterooligomeric bundles have been used to successfully identify specific environmental effects that influence the reactivity of metal ion cofactor capable of binding substrates,[36] the non-covalently assembled complexes have often been difficult to characterize structurally, possibly due to small populations of alternately assembled species. In this case, cassetting the active site residues into a construct with linked helices greatly facilitated structural analysis and catalytic characterization.[37]

An attractive starting scaffold with a single polypeptide chain is the *de novo* designed three-helix bundle α_3D. The NMR structure of this protein is known (Figure 5-1A) and it has been proven that the helices are oriented in a counter-clockwise topology (Figure 5-1B).[38] Although the α_3D protein originated from a coiled-coil, its helices were shortened to the point where it might be better considered as a globular protein rather than a coiled-coil, whose repetitive structure make each of the heptads very similar to one another. α_3D has a diverse set of hydrophobic amino acids that constitute a well packed hydrophobic core (Figure 5-2A) and many hydrophilic amino acids (Figure 5-2B) that form salt bridge interactions. The stability of α_3D was similar to that of native proteins even though it shares no sequence homology with the proteins in the data base. Because of its high stability, α_3D should be tolerant to mutations allowing this protein to serve as an excellent framework to engineer specific metal binding sites. Additionally, with this protein scaffold we can study the effect of the ligating residue located on the second helix which is antiparallel to the first and third helices of the bundle.

Studies have been performed to test the malleability and adaptability of α_3D by mutating Ala 60 in its core to larger, hydrophobic side chains Leu and Ile.[39] Such changes introduce strain into the structure of natural proteins, and therefore generally destabilize the native state. By contrast, these mutations were slightly stabilizing (~1.5 kcal mole^{-1}) to the tertiary structure of α_3D. The value of ΔC_p for unfolding of these mutants was not greatly affected relative to the wild type protein, indicating that the change in solvent accessibility for unfolding was similar. However, two-dimensional heteronuclear single quantum coherence spectra indicated that the protein adjusts to the introduction of steric bulk in different ways. A60L mutant showed serious erosion in the dispersion of both the amide backbone as well as the side-chain methyl chemical shifts. By contrast, A60I showed excellent dispersion of the backbone resonances, and selective changes in dispersion of the aliphatic side-chains proximal to the site of

mutation. Together, these data suggested that the α_3D, although folded into a unique three-dimensional structure, is nevertheless more malleable and flexible than most natural, native proteins. The backbone and side-chain dynamics of α_3D were investigated using ^{15}N, ^{13}C, and ^{2}H nuclear magnetic resonance relaxation methods with the aim of assessing the character of the internal motions of this native-like protein of *de novo* design.[40] At the backbone level, both ^{15}N and $^{13}C_\alpha$ relaxation studies indicate highly restrictive motion on the picosecond to nanosecond time scale in the α-helical regions of α_3D, with increasing mobility at the ends of the α-helices and in the two loop regions. This is largely consistent with what is seen in proteins of natural origin. Overall, the view provided by both ^{2}H and ^{13}C methyl relaxation methods suggest that the side chains of α_3D are more dynamic compared to natural proteins. Regions of relative flexibility bound clusters of rigid methyl-bearing side-chain groups that are interspersed with aromatic and β-branched amino acids. The time scale of motions associated with methyl-bearing side chains of α_3D are significantly longer than that seen in natural proteins. These results indicate that the strategies underlying the design of α_3D have largely, but not completely, captured both the structural and dynamic character of natural proteins. Hydration of amides of α_3D was examined by molecular dynamics simulation, IR, and NMR.[41] Molecular dynamics calculations show that the amide carbonyls on the surface of the protein tilt away from the helical axis to interact with solvent water, resulting in a lengthening of the hydrogen bonds on this face of the helix. Water molecules are bonded to these carbonyl groups with partial occupancy (~ 50%–70%), and their interaction geometries show a large variation in their hydrogen bond lengths and angles on the nsec time scale. This heterogeneity is reflected in the carbonyl stretching vibration (amide I' band) of a group of surface Ala residues. The surface-exposed amides are broad, and shift to lower frequency (reflecting strengthening of the hydrogen bonds) as the temperature is decreased. By contrast, the amide I' bands of the buried ^{13}C-labeled Leu residues are significantly sharper and their frequencies are consistent with the

formation of strong hydrogen bonds, independent of temperature. The rates of hydrogen-deuterium exchange and the proton NMR chemical shifts of the helical amide groups also depend on environment. The partial occupancy of the hydration sites on the surface of helices suggests that the interaction is relatively weak, on the order of thermal energy at room temperature. One unexpected feature that emerged from the dynamics calculations was that a Thr side chain subtly disrupted the helical geometry 4–7 residues N-terminal in sequence, which was reflected in the proton chemical shifts and the rates of amide proton exchange for several amides that engage in a mixed $3_{10}/\alpha/\pi$-helical conformation. The folding/unfolding kinetics of α_3D, has been described.[42] Both IR temperature jump and ultrafast fluorescence mixing methods reveal a single-exponential process consistent with a minimal folding time of 3.2 ± 1.2 μs (at ~ 50℃), indicating that a protein can fold on the 1- to 5-μs time scale. Furthermore, the single-exponential nature of the relaxation indicates that the prefactor for transition state (TS)-folding models is probably ≥1 $(\mu s)^{-1}$ for a protein of this size and topology. Molecular dynamics simulations and IR spectroscopy provide a molecular rationale for the rapid, single-exponential folding of this protein. α_3D shows a significant bias toward local helical structure in the thermally denatured state. The molecular dynamics-simulated transition state ensemble is highly heterogeneous and dynamic, allowing access to the transition state via multiple pathways.

Before attempting to prepare asymmetric metal coordination environments or site specific H-bonds, we felt it was important to first redesign α_3D so as to introduce symmetric metal binding sites involving cysteine residues. This approach would allow us to exploit the extensive body of work defining heavy metal complexation using the TRI and Coil Ser peptides to assess whether a specific metal structure could be achieved in the modified α_3D construct. It would also allow us to probe the physical properties of metals such as Cd^{II}, Hg^{II} and Pb^{II} in a more natural antiparallel helical system. A significant amount of literature

exists where metal binding sites have been designed in existing four-helix bundles and in a mixture of α/β protein structural frameworks.[22,43] However, there are no such examples where novel metal binding site(s) have been engineered within an antiparallel single chain three-helix bundle. Lessons learned from this exercise can then be applied to the subsequent design of more complex and challenging systems such as fully asymmetric metalloproteins.

Positioning of the Metal Binding Site:

Based on visual inspection of the α_3D structure, four potential sites along the bundle were identified where three cysteines, one from each helix, could be introduced, designated as α_3D**I**, α_3D**II**, α_3D**III**, and α_3D**IV**, starting from the N-terminal end o the helical bundle to the C-terminal end (Figure 5-3). Each Cys was contributed by each of the helices of the three-helix bundle. Out of these four mutants, the sites in one such mutant, α_3D**IV** (Figure 5-4), located at the C-terminal end of the bundle, seemed to be optimal based on the properties of the starting protein as described earlier. [38-40,44] Previous NMR structural and dynamic investigations showed a gradient in the dynamic behavior and malleability of the protein, with the C-terminal end of the bundle being most amenable to amino acid substitutions. The selected location has a well-ordered backbone conformation, but the side chains of the mutated residues are less well ordered than residues in other locations of the bundle. The 3-Cys site, which is largely sequestered from solvent, occupies a “box”; the sides are formed by the backbone of the helices and the bottom by the apolar side chains of Phe31, Ile14, and Ile63. The aromatic residue of Phe31 lies directly over the predicted metal-binding site and lines most of the bottom of the box. The top is formed by the main-chain and side chains of residues in the non-helical loops including Leu21, as well as the apolar portion of Tyr 70 at the terminus of helix 3. His72, which was entirely disordered in the NMR solution structure, also lies proximal to the site. Moreover, after introduction of the Cys side chains in one of the two

preferred rotamers for Cys, the thiol S^G atoms formed a nearly equilateral triangle with the side chains well oriented to form the desired site (inter S^G distances = 3.5 - 4.8 Å). Overall, the location is ideal to explore the effects of hydrophobic sequestration in the present work. Sequence of α_3D**IV** is shown in Table 5-1.

Materials and Methods

Expression and Purification: A synthetic DNA of α_3D**IV** was cloned into pET-15b (Celtek Genes) vector and expressed in *E. coli* BL21(DE3) competent cells (Stratagene) and grown in M9 media. After sonication and heat denaturation at 55°C the lyophilized powder was purified on a C18 preparative reverse phase HPLC column over a linear gradient of 100% H_2O/0.1% TFA to 10% H_2O/90% Acetonitrile/0.1% TFA over 50 minutes. A chromatogram is shown in Figure 5-5A. The peak eluted with a retention time of 28 min corresponds to the fraction of pure α_3D**IV**. The molecular weight of the pure peptide was determined to be 7945.1 Da by Electrospray Ionization Mass Spectrometry (Figure 5-5B) which corresponds to α_3D**IV** with the deletion of the first Met (calc MW 7946.9 Da). The yield of the pure protein was 17 mg/L. The concentration of the protein was determined by A_{280nm} and using known extinction coefficient ε_{280} = 8.61 mM^{-1} corresponding to 1 Trp, 2 Tyr and 3 Cys.[45]

CD Spectroscopy: CD spectra of 5 μM α_3DIV were collected on an Aviv model 202 CD spectrometer using rectangular open top quartz cuvettes at 25°C at different pH values (pH 3-9). GuHCl titration experiments were carried out using a Microlab 500 series syringe pump automatic titrator controlled by Aviv software. Titrations were carried out by mixing two separate solutions of 10 μM peptide containing 0.0 and 7.63 M GuHCl in 10 mM phosphate buffer at pH 8.

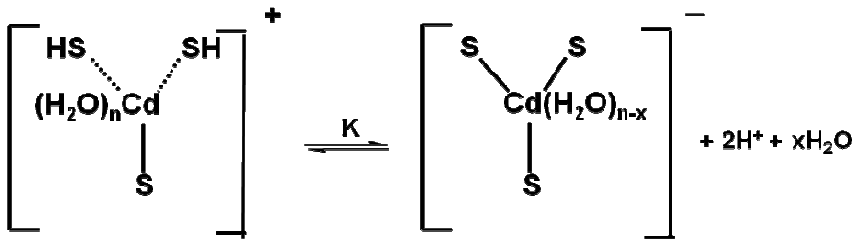
The observed ellipticity in millidegree was converted to molar ellipticity (deg cm^2 $dmole^{-1}$ res^{-1}) as described previously,[46] using 59 amino acids that are present in the helical region.[38] GuHCl titration data were plotted as the concentration of folded protein vs. the concentration of GuHCl and was fit to a previously published equation, derived based on a two-state model.[47-49]

^{1}H NMR Spectroscopy: ^{1}H-^{1}H NOESY experiments were performed at room temperature on a 500 MHz Varian Inova spectrometer equipped with an inverse detection probe using standard procedures using 100 ms of mixing time.[50] The sample was prepared by dissolving 15.5 mg of lyophilized and degassed peptide in 509 uL 10% D_2O/H_2O under a flow of argon which resulted in a protein concentration of 3 mM. The pH was adjusted to 6 by adding concentrated solutions of KOH and HCl.

Metal Binding Titrations: Cd^{II}, Hg^{II} and Pb^{II} binding titrations were performed at room temperature on a Cary 100 Bio UV/Vis spectrometer using anaerobic cuvettes (Starna Inc.) of 1-cm path length by adding aliquots of stock solutions of different metals. Peptide samples of 20-30 μM were prepared in 50 mM of appropriate buffers (TRIS for pH 8 and CHES for pH 8.6) and 40-60 μM *tris*(2-carboxyethyl)phosphine (TCEP) inside an inert atmosphere box (Vacuum Atmospheres Co., model OMNI-LAB). Stock solutions of 8 mM $CdCl_2$, 7.37 mM $HgCl_2$ and 5.16 mM $PbCl_2$ were also prepared inside the inert atmosphere box. In each case, difference spectra were obtained by subtracting the background spectra of samples containing peptide, buffer and TCEP in the absence of metals. Direct titration data were analyzed by non-linear least squares fits to the equation described previously.[51] The difference molar extinction coefficients ($\Delta\varepsilon$) were determined based on the total metal concentrations after subtracting the blank spectra in the absence of metals.

pH Titrations: pH titrations were performed as described previously,[52,53] by adding small aliquots of KOH to solutions containing 20-30 μM α_3D**IV** and 1 equivalent of $CdCl_2$, $HgCl_2$ and $PbCl_2$. In the case of Cd^{II} and Pb^{II} the titration

data were analyzed using the model: simultaneous dissociation of the two protons of Cys thiols.[52,53] The corresponding equilibrium for this process is shown below:

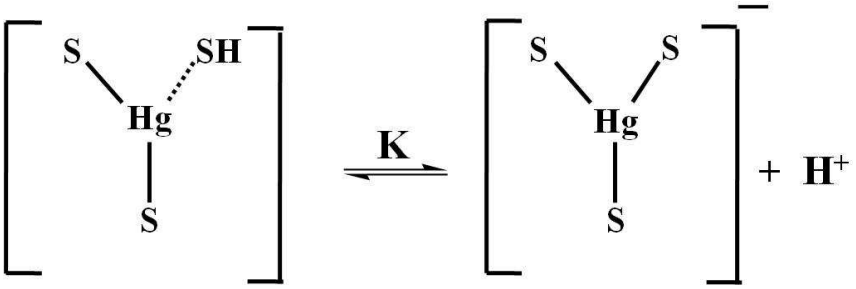


Based on this equilibrium, the following equation can be derived as previously described[52,53]

$$A = (M*P1)\ (10^{\ (P2-2*x)}+1)+P3$$

where, A is absorbance of the Cd(II)-trithiolate complex being formed during the course of titration, M is the molar concentration of total Cd(II) used in the titration, P1 is the molar extinction coefficient of the Cd(II)-trithiolate complex, P2 is the pK_{a2} corresponding to the release of two thiol protons of Cys, x is the pH at each point during titration and P3 is the baseline absorbance.

For Hg^{II}, the data were analyzed using the model: dissociation of one thiol proton of Cys as shown below by the following equilibrium:[33]



The following equation can be derived based on the above equilibrium, as previously described[33]

$$A = ((M*P1)\ (10^{\ (P2-x)}+1))+P3$$

where A is the absorbance of the Hg(II)-trithiolate complex, M is the molar concentration of total Cd(II) used in the titration, P1 is the molar extinction coefficient of the Hg(II)-trithiolate complex, P2 is the pK_a corresponding to the release of one thiol proton of Cys from the linear HgS_2 complex, x is the pH at each point during titration and P3 is the baseline absorbance.

^{113}Cd and ^{199}Hg NMR: ^{113}Cd NMR and ^{199}Hg NMR experiments were performed according to standard procedures.[54] Samples contained 10-16 mg of lyophilized peptide. Final protein concentrations were 2.8 – 3 mM. pH of the solutions were adjusted by adding concentrated solutions of KOH and HCl. The pH values of the solutions were also checked at the end of each experiment. An exponential line broadening of 200 Hz was applied prior to Fourier transformation while processing ^{199}Hg NMR data.

^{111m}Cd and ^{199m}Hg PAC: Samples for ^{111m}Cd PAC measurement contained 300 μM α_3D**IV**, 1/12 equivalent of Cd^{II} and 55% sucrose (w/w) in 20 mM TRIS buffer at pH 8.1. Sample preparation and data collection were performed at the University of Copenhagen as previously described.[53,54] Samples for ^{199m}Hg PAC experiments contained 200 μM α_3D**IV**, 80 – 400 μM Hg^{II} and 55% sucrose in 100 mM of appropriate buffer (phosphate for pH 5.8, 7.4 and CHES for pH 8.6). Sample preparation and data collection were performed at CERN as described previously.[55]

Protein Gel. A 4-12% Bis-Tris polyacrylamide gel (NuPage) was run in 1X MES running buffer at pH 6.4. The following samples were prepared: apo α_3D**IV**, α_3D**IV** + 0.8 equivalent Cd(II) pH 8.0, α_3D**IV** + 0.8 equivalent Hg(II) pH 5.8, α_3D**IV** + 0.8 equivalent Hg(II) pH 8.6, α_3D**IV** + 2.0 equivalent Hg(II) pH 8.6, α_3D**IV** + 0.8 equivalent Pb(II) pH 8.0, α_3D**IV** + 2.0 equivalent Pb(II) pH 8.0. Each well contained ~ 10 μg of protein. The gel was stained by coomassie blue and destained overnight in a solution of 50% methanol (HPLC grade), 10% acetic acid, 40% water.

Results

Characterization of the apo Peptide. After expression and purification of α_3D**IV** by HPLC, the folding behavior of α_3D**IV** was studied in solution by CD and NMR spectroscopy. The CD spectrum of 5 μM α_3D**IV** has double minima at 222 and 208 nm at pH 8 with the molar ellipticity [Θ] values at these wavelengths being –34726 and –33000 deg cm^2 dmole^{-1} res^{-1}, respectively (Figure 5-6A). These values are characteristic of a well-folded α-helical construct (97% α-helical based on $[\Theta]_{222}$). α_3D**IV** remained well-folded between pH 3 and 9 (Figure 5-6B). The ^{1}H-^{1}H NOESY spectrum of 3 mM α_3D**IV** at pH 6, shows chemical shift dispersions characteristic of a well-folded α-helical structure (Figure 5-7). GuHCl-induced unfolding of α_3D**IV** was monitored by the change in ellipticity (in millidegree) at 222 nm (local minimum in the CD spectrum of α helices) as a function of GuHCl concentration at pH 8. The resulting titration curve was plotted as concentration of folded protein vs. GuHCl concentration (Figure 5-8) and was fit by a non-linear least square method to a two-state equilibrium, yielding the free energy of unfolding (ΔG_u) associated with the transition from a well folded α–helical structure to an unfolded random coil to be 2.5 kcal mole^{-1}, with the degree of cooperativity (m) being 1.4 kcal mole^{-1} M_{GuHCl}^{-1}. The midpoint of the transition (C_m) occurred at a 1.8 M GuHCl concentration. These results indicate that replacing three Leu residues of the WT α_3D with cysteines resulted in a loss of unfolding free energy of ~ 2.5 kcal mole^{-1}.[38]

Characterization of Hg^{II} Complexes of α_3D**IV:**

UV/Vis Studies. Having established that α_3D**IV** is well folded and stable in solution, first the complexation of Hg^{II} to α_3D**IV** was investigated by UV/Vis spectroscopy. Direct metal binding titrations were performed under anaerobic conditions where aliquots of Hg^{II} stock solution were added to 30 μM of peptide solution at pH 8.6 in the presence of 60 μM TCEP. The progress of the titration was monitored by the appearance of characteristic absorption bands due to

Ligand to Metal Charge Transfer (LMCT) transitions at 247 nm, upon formation of Hg^{II}-S bonds in the complex $Hg^{II}(\alpha_3D\mathbf{IV})^-$. The resulting UV/Vis absorption spectra are shown in Figure 5-9. Molar extinction coefficients ($\Delta\varepsilon$), are 12 500 (λ_{247nm}), 8 400 (λ_{265nm}) and 3 900 (λ_{295nm}) $M^{-1}cm^{-1}$ for the complex $Hg^{II}(\alpha_3D\mathbf{IV})^-$ (Table 5-2). Due to the high affinity binding, stability constants could not be determined directly. Hg(II) formed 1:1 complex with $\alpha_3D\mathbf{IV}$. In the presence of excess Hg(II), overall UV-Vis spectra changed and the λ_{max} shifted from 247 nm towards 240 nm, representing similar spectral features as observed for a linear HgS_2 complex. Next, I examined pH dependent complexation of Hg^{II} to $\alpha_3D\mathbf{IV}$ monitoring changes in the LMCT band as a function of pH. As Hg^{II} binds to the Cys thiols as a linear $HgS_2(SH)$ complex at low pH to TRI peptides, such as in the complex Hg(**TRI**L16C)$_3$, titration data of Hg^{II} were fit to the release of one thiol proton upon forming a HgS_3 complex for $\alpha_3D\mathbf{IV}$,[33] resulting in a pK_a of 7.1±0.1 (Table 5-2, Figure 5-10). This pK_a for $\alpha_3D\mathbf{IV}$ is slightly more acidic than observed previously for TRI peptides.[33,46]

^{199}Hg NMR and ^{199m}Hg PAC Studies. ^{199}Hg NMR and ^{199m}PAC (Perturbed Angular Correlation) spectroscopy were used to probe the coordination environment around Hg^{II} bound to $\alpha_3D\mathbf{IV}$. ^{199}Hg NMR experiments were performed with $\alpha_3D\mathbf{IV}$ at pH 5.8, 8.6 and 7.4. Based on the pK_a value Hg^{II} is expected to form a linear HgS_2 complex at pH 5.8 and a trigonal HgS_3 complex at pH 8.8, with a mixture of linear and trigonal complexes at the intermediate pH. ^{199}Hg NMR spectra of 2.9 mM $\alpha_3D\mathbf{IV}$ loaded with 0.8 equivalents of Hg^{II} has resonances at δ = –938 and –240 ppm at pH 5.8 and 8.6, respectively (Figure 5-11). The chemical shift of –938 ppm is within the range (–760 to –1200 ppm) of what is expected for a dithiolate-Hg^{II} complex.[27,56] ^{199m}Hg PAC spectrum of $\alpha_3D\mathbf{IV}$ loaded with 0.4 equivalents of Hg^{II} at pH 5.8 (Figure 5-12) has one set of signals that corresponds to an NQI with an angular frequency ν_Q of 1.48 GHz and η of 0.15 (Table 5-3), also consistent with the formation of a linear HgS_2 complex with distorted geometry.[55,57,58] The chemical shift of –244 ppm somewhat up field of

what has been previously observed for trithiolate-Hg^{II} complexes of TRI family of peptides.[25,55] The ^{199m}Hg PAC spectrum (Figure 5-12) at pH 8.6 has one set of resonances with ν_Q of 1.11 GHz and η of 0.4 (Table 5-3) which agree well with the formation of a distorted trigonal HgS_3 complex.[55,59,60] At pH 7.4, two ^{199}Hg NMR resonances with δ = –926 and –240 ppm are observed (Figure 5-11), indicating the presence of both the dithiolate-Hg^{II} and trithiolate-Hg^{II} complexes of α_3D**IV** at this pH. These conclusions are confirmed with the PAC data (Figure 5-12) which was fit for two NQIs with ν_Q of 1.48 GHz (η = 0.15) and 1.11 GHz (η = 0.4) (Table 5-3) indicating the presence of two Hg^{II} species, HgS_2 and HgS_3 present as distorted geometries.[55,57-60] ^{199}Hg NMR of 2.2 mM α_3DIV bound to 2 equivalents of Hg^{II} at pH 8.6 has a resonance at δ = –927 ppm, indicating the formation of linear HgS_2 complex under these conditions (Figure 5-11). Same chemical shift is observed under similar stoichiometric conditions at pH 5.8 indicating that the formation of this complex is independent of pH when Hg^{II} is present in excess.

Characterization of Cd^{II} Complexes of α_3DIV**:**

UV/Vis Spectroscopy. The Cd^{II} complexes of α_3D**IV** were characterized by UV/Vis spectroscopy. Titration of Cd^{II} to 20 μM α_3D**IV** at pH 8 in the presence of 40 μM TCEP was performed under anaerobic conditions and the complexation was monitored by the appearance of UV bands due to the LMCT in the complex $Cd^{II}(\alpha_3$D**IV**$)^-$ (Figure 5-13). The λ_{max} and $\Delta\varepsilon$ values are as follows: (λ_{max} = 232 nm and $\Delta\varepsilon$= 18 200 $M^{-1}cm^{-1}$) (Table 5-2). The UV-Vis spectral features and the molar extinction coefficient values and consistent with the formation of Cd^{II}-tristhiolate complex of α_3D**IV**. A binding constant of 2.0×10^7 M^{-1} (Table 5-2) was extracted from a direct metal binding titration using known fitting equations.[51]The pH-dependent complexation of Cd^{II} to α_3D**IV** was monitored by UV/Vis spectroscopy. The resulting pH titration curve (Figure 5-10) was fit to the

previously published model, [33,52,53] simultaneous dissociation of two Cys thiols, yielding a pK_{a2} of 10.6 ± 0.1 (Table 5-2).

^{113}Cd NMR and ^{111m}Cd PAC Studies. ^{113}Cd NMR spectrum of 0.8 equivalents of Cd^{II} bound to 3 mM α_3D**IV** has two resonances at δ = 595 and 583 ppm at pH 8 (Figure 5-14) indicating the presence of two Cd^{II} species with chemical shift values similar to what has been observed for 4-coordinate pseudotetrahedral CdS_3O species.[53] However, definitive assignments of these peaks to the corresponding Cd^{II} complexes were not possible based only on the ^{113}Cd NMR data. To confirm the coordination environment and geometry of Cd^{II} complexes of α_3D**IV**, ^{111m}Cd PAC spectroscopic studies were performed. ^{111m}Cd PAC spectrum of α_3D**IV** loaded with 1/12 equivalents of Cd(II) has three NQIs, at ω_0 = 0.35 (η = 0), 0.268 (η = 0.18) and 0.17 (η = 0.5) rad/ns (Figure 5-15, Table 5-4) at pH 8.1.

Characterization of Pb^{II} complexes of α_3DIV**:**

UV/Vis Studies with Pb^{II}. Pb^{II} binding studies with α_3D**IV** were performed by UV/Vis. Spectroscopy. The Pb^{II} titration to 20 μM α_3D**IV** was performed at pH 8.0 in the presence of 40 μM TCEP under anaerobic conditions and followed the LMCT transitions due to the formation of Pb-S bonds. From the resulting UV/Vis spectra (Figure 5-16) molar extinction coefficients of 18 000 (λ_{236nm}), 14 400 (λ_{260nm}), 9 100 (λ_{278nm}), and 3 150 (λ_{346nm}) $M^{-1}cm^{-1}$ (Table 5-2) are determined. A Pb^{II} binding constant of 3.1×10^7 M^{-1} (Table 5-2) was determined from the analysis of the metal binding data using a known equation.[51] In the presence of more than stoichiometric amount of Pb(II), the intensity of the peak at λ_{346nm} decreased, while the intensity of the peak at λ_{236nm} increased continuously. Intensity of the UV signal between 300 and 330 nm also increased. Other spectral features of the UV-Vis spectra in the presence of excess Pb(II) are similar to those of the stoichiometric conditions. Three isosbestic points are observed at λ_{330nm}, λ_{280nm} and λ_{260nm} under excess conditions of Pb(II). To

examine pH-dependent complexation of Pb^{II} to α_3D**IV**, pH titrations were performed and the progress of the experiment was followed by LMCT transition at 236 nm. The resulting pH titration curve (Figure 5-10) was fit to the simultaneous dissociation of two Cys thiols,[52], yielding a pK_{a2} of 10.2 $\pm$ 0.1 (Table 5-2).

Protein Gel. Figure 5-18 shows a 4-12% NuPage Bis-Tris gel run in 1X MES buffer pH 6.4 loaded with samples of apo α_3D**IV** as well as in the presence of different equivalents of metals. Samples 1, 2, 3, 4, 5, 6, and 7 correspond to apo α_3D**IV**, α_3D**IV** + 0.8 equivalent Cd(II) pH 8.0, α_3D**IV** + 0.8 equivalent Hg(II) pH 5.8, α_3D**IV** + 0.8 equivalent Hg(II) pH 8.6, α_3D**IV** + 2.0 equivalent Hg(II) pH 8.6, α_3D**IV** + 0.8 equivalent Pb(II) pH 8.0, α_3D**IV** + 2.0 equivalent Pb(II) pH 8.0. Samples 1, 2, 3, 4, 6 and 7 show a low molecular weight band with similar molecular weight for all of these samples at ~ 5 kDa. Sample 5, loaded with 2.0 equivalents of Hg(II) shows a high MW band ~ 13.5 kDa. All the bands appear at lower molecular weights than expected for these samples.

Discussion

The objective of the study in this chapter is to engineer a tris(cysteine) metal binding site on a existing three-helix bundle α_3D by rational design. If successful in expressing the desired mutant then we can address the following questions: 1) How well folded the apo protein would be? 2) Would a stable protein be achieved? 3) What would the stability of the mutant protein be relative to the wild type protein? 4) Would the mutant protein be able to incorporate metal ions such as Hg(II), Cd(II), Pb(II)? 5) If yes, then what would be affinity and coordination geometry of the metal ions binding to the protein? 6) Would the observed physical properties of the metallated complexes of the reengineered three-helix bundle protein be comparable to the three-stranded coiled coil TRI peptides? Based on visual inspection of the structure of α_3D, four potential sites

were identified along the helical bundle where the triple-Cys site could be formed (Figure 5-3). Each Cys was contributed by replacing one amino acid from each of the three helices. These desired mutants were designated as α_3D**I** (F7CL42CL56C), α_3D**II** (L11CF38CA60C), α_3D**III** (I14CI35CI63C), and α_3D**IV** (L18CL28CL67C), starting from the N-terminal end of the bundle to the C-terminal end. Inspection of the mutated residues shows that in α_3D**I**, one Phe and two Leu residues are replaced by cysteines. In α_3D**II** one Leu, one Phe and one Ala are mutated to cysteines. In α_3D**III** all the cysteines are obtained by mutating three Ile residues. In α_3D**IV** three Leu residues are mutated to cysteines. Figure 5-3 shows that in α_3D**II** and α_3D**III** the desired metal binding sites are located almost in the middle of the bundle and thus these two mutants might not be able to access metal ions. The metal binding sites in α_3D**I** and α_3D**IV** are located at each end of the helix and hence would be potential candidates for metal binding studies as these two mutants would allow easy access of metals to the designed sites. Additionally, as more hydrophobic residues such as Phe, and Ile confer better hydrophobic packing than Leu, α_3D**IV** as the preferred mutant for our studies as α_3D**IV** has all Leu substitutions and thus would minimize the effect of loss in hydrophobic interaction in the targeted mutant. Furthermore, α_3D**IV** was the most suitable mutant based on the properties of the wild type protein as described earlier in the chapter.

Comparison of Cys site in α_3DIV** with CS**L9C. PyMol model of α_3D**IV** was generated by mutating the Leu 18, Leu 28, and Leu 67 to one of the preferred Cys rotamers. A view from perpendicular to the helical axis (Figure 5-17A) shows that the β-methylenes of Cys 18 and Cys 67 Cys in α_3D**IV** are pointing towards the C-terminus while that of Cys 28 is pointing towards the N-terminus of the bundle. In comparison, all the β-methylenes of the Cysteines in the X-ray structure of parallel three-stranded coiled coil **CS**L9C are shown to be oriented towards the N-terminus if the helices.[51] Cys 28 in α_3D**IV** is contributed from the second helix which is oriented in a mutually anti-parallel orientation to

the first and third helices, respectively. Thus, the observed orientation of the β-methylenes of Cys 28 is towards the C-terminus of the bundle is predicted based on the parallel three-stranded coiled coil structure of **CS**L9C. Such orientation of β-methylenes of Cys 28 is similar to what has been observed for the β-methylenes of D-amino acid D-Peniclillamine (Pen) in parallel three-stranded coiled coil structure of **CS**L16D-Pen.[61] Similar orientation of D-Leu has also been predicted in **TRI**L12L$_D$L16C studied previously by the group.[32] Thus, similar types of structural orientations of amino acids side chains and the implications of metal binding can be studied as those of the non-coding D-amino acids, but using the naturally occurring L-amino acids in a three-helix bundle structure. A view from the N-terminus (Figure 5-17B) of α_3D**IV** shows that the S^G atoms of Cys 18 and Cys 28 are oriented towards the interior of the bundle whereas for Cys 67 it is directing almost towards the helical interface. In **CS**L9C all three Cys side chains had alternate conformations where each Cys orientation was partitioned between the coiled coil interior and the helical interface in almost equal proportions. In **CS**L9C the S^G – S^G separation for the interior Cys conformers is 3.32 Å. The S^G – S^G separations of α_3D**IV** are 3.5, 4.0, and 4.8 Å. These distances form an almost equilateral triangle which is ideal for investigating the binding of metals such as Hg(II), Cd(II), and Pb(II) to this construct.

Characterization of the apo Peptide. α_3D**IV** was successfully expressed *E. coli* and grown in M9 media. After purification by HPLC the desired peptide was obtained as confirmed by ESI mass spectroscopy (MW = 7945.1 Da, calc = 7946.9 Da) with the deletion of Met 1. The yield of the pure protein was ~17 mg/L. First, the folding of α_3D**IV** was studied by CD and ^{1}H NMR spectroscopy. CD spectrum of 5 μM α_3D**IV** at pH 8.0 shows double minima at 208 and 222 nm characteristic for α-helices (Figure 5-6). The moalr ellipticity value of –34726 and –33000 deg cm^2 dmole^{-1} res^{-1}, at these wavelengths are characteristic of a well-folded α-helical structure (97% α-helical). Furthermore, α_3D**IV** remained well-folded between pH 3 and 9. ^{1}H-^{1}H NOESY spectrum (Figure 5-7) of α_3D**IV** at pH

6.0 shows well-dispersed chemical shift values characteristic of a well folded protein. Thus, the desired mutant was not only expressed successfully, but also a well-folded protein was obtained in solution.

Stability of α_3D**IV** was measured by performing a GuHCl-induced unfolding of the apo peptide at pH 8. The titration curve shows a single cooperative transition (Figure 5-8). The titration data was fit to a two state equilibrium, yielding the free energy of unfolding (ΔG_u) of 2.5 kcal mole^{-1}, with the degree of cooperativity (m) being 1.4 kcal mole^{-1} M_{GuHCl}^{-1}. The midpoint of the transition (C_m) occurred at a 1.8 M GuHCl concentration. The degree of cooperativity (m) determined from GuHCl denaturation has been attributed to an increase in the accessible surface area upon unfolding.[39] For α_3D, the m value was 2.41 kcal mole^{-1} M_{GuHCl}^{-1}.[38] A lower m value for α_3D**IV** suggests that upon folding the increase in accessible surface area for α_3D**IV** is less than that of the WT protein. This consistent with the face that three Leu residues, each having large surface area have been mutated to Cys which have smaller surface area than Leu.[62] The ΔG_u of 2.5 kcal mole^{-1} is less than what is found for native proteins, the stability for which are reported to be ~5-10 kcal mole^{-1}.[63] The wild type protein α_3D had free energy of unfolding (5.1 kal mole^{-1})[38], similar to those of native proteins. Thus, mutating three Leu of α_3D to three Cys residues resulted in a loss of free energy of unfolding of ~2.5 kcal mole^{-1} in the mutant protein α_3D**IV**. Even though the stability of α_3D**IV** is almost half than that of α_3D, it was a stable enough construct so that further characterization of different metallated complexes of α_3D**IV** could be performed.

Characterization of Hg(II) Complexes of α_3DIV.** Having established that α_3D**IV** was well-folded and stable enough in solution, Hg(II) binding studies were performed with this protein. First, direct metal binding studies were performed under anaerobic and reducing conditions (TCEP), and monitored by appearance of characteristic LMCT transitions in the UV-Vis spectra of the metallated

complex of α_3D**IV** (Figure 5-9). Gradual addition of Hg(II) to α_3D**IV** resulted in an increase of the characteristic absorbance bands. Difference extinction coefficients ($\Delta\varepsilon$) at the characteristic wavelengths are as follows: 12 500 (λ_{247nm}), 8 400 (λ_{265nm}) and 3 900 (λ_{295nm}) $M^{-1}cm^{-1}$ (Table 5-2). These values of extinction coefficients are lower than that observed for Hg(**TRI**L16C)$_3^-$ that values for which were $\Delta\varepsilon_{247nm}$ = 19, 200 $M^{-1}cm^{-1}$, $\Delta\varepsilon_{265nm}$ = 11, 900 $M^{-1}cm^{-1}$, and $\Delta\varepsilon_{295nm}$ = 5, 800 $M^{-1}cm^{-1}$.[46] Nonetheless, the appearance of the UV bands at these wavelengths and the respective extinction coefficients are consistent with the fact that all three Cys thiolates of α_3D**IV** are incorporated into the first coordination sphere of Hg^{II} leading to the formation of the complex $Hg^{II}(\alpha_3D\mathbf{IV})^-$. Due to strong binding, Hg(II) binding affinity could not be obtained from the direct metal binding titration data. α_3D**IV** formed the trigonal complex with Hg(II) with a stoichiometry of 1:1.

Next, pH dependence for the complexation of Hg(II) to α_3D**IV** was examined by UV-Vis spectroscopy by monitoring the LMCT band due to the formation of Hg(II)-thiolato bonds. With an increase in pH of a solution containing 30 μM α_3D**IV** in the presence of 1.0 equivalent Hg(II), led to an increase in the intensity of the UV bands. For the three-stranded coiled coil systems, Hg(II) is known to first form a T-shaped HgS_2SH complex within the three-stranded coiled coil at low pH.[33,46,55] With an increase in pH the proton of the third Cys thiol is released leading to the formation of tris-thiolato HgS_3 complex. Thus analysis of the titration data is performed by fitting the titration curve to an equation derived based on the equilibrium $[HgS_2SH] \rightarrow [HgS_3]^- + H^+$. Similarly, the pH titration data for Hg(II) binding to α_3D**IV** (Figure 5-10) was fit to the release of a single thiol proton of Cys leading to the formation of HgS_3 complex of α_3D**IV** from the linear HgS_2 complex. A pK_a of 7.1 ± 0.1 was obtained from this analysis. This pK_a is marginally more acidic than that of 7.6 ± 0.2 observed for TRI peptides.[33]

Having established that all three Cys thiolates are incorporated into the first coordination sphere of Hg(II) and that the $Hg^{II}(\alpha_3D\mathbf{IV})^-$ complex is formed,

coordination geometry around the Hg(II) ion was probed by using ^{199}Hg NMR and ^{199m}Hg PAC spectroscopies. These two techniques are complimentary to each other and provide important information about the nature of the coordinating ligands to the metal ion, coordination number as well as metal ion geometry. Based on the pK_a of Hg(II) binding to α_3D**IV**, at pH 5.8 only the linear HgS_2 complex is expected to form. At pH 7.4, a mixture of linear and trigonal HgS_3 complexes are expected to form, while at pH 8.6 only the trigonal species is expected. ^{199}Hg NMR spectra of 2.9 mM α_3D**IV** in the presence of 0.8 equivalent of ^{199}Hg(II) has resonances at δ = –938 and –240 ppm at pH 5.8 and 8.6, respectively (Figure 5-11). The chemical shift of –938 ppm is within the range (–760 to –1200 ppm) of what is expected for a dithiolate linear HgS_2 complex.[27,56] However, this chemical shift is slightly up filed compared to those of linear Hg(II) complexes of TRI peptides (–834 to –844 ppm).[25,55] This difference in chemical shift of ~100 ppm is most likely due to a difference in Hg(II)-S bond distances in the TRI peptides compared to α_3D**IV**, as ^{199}Hg NMR chemical shifts are very sensitive to the perturbations of the Hg(II)-S distances. Quantum chemical calculations indicate a change in Hg(II)-S bond distance of ~0.05 Å will give rise to a change of about 100 ppm, but other structural changes may also affect the chemical shift (Hemmingsen *et. al.* unpublished results. They are also exploring various structural changes as well theoretically). ^{199m}Hg PAC spectrum of α_3D**IV** in the presence of 0.4 equivalent Hg(II) at pH 5.8 has only one set of resonances corresponding to a Nuclear Quadrupole Inetraction (NQI) with an angular frequency ν_Q 1.48 GHz and the asymmetry parameter η of 0.15 (Figure 5-12, Table 5-3). These parameters are consistent with the formation of HgS_2 complex in distorted geometry.[55,57,58] The corresponding values for the TRI peptides are ν_Q of 1.53 GHz and η of 0.11,[55] which are very similar to that of α_3D**IV**. Thus, both the ^{199}Hg NMR and ^{199m}Hg PAC data confirm that indeed linear HgS_2 complex of α_3D**IV** is formed in distorted geometry under low pH conditions. The chemical shift of –240 ppm at pH 8.6 is slightly up field from what is expected for a trigonal thiolato HgS_3 complex, the range for which has been observed to be

–80 to –160 ppm for mononuclear 3-coordinate aliphatic thiolate complexes,[27,56] as well as those of the **a** site TRI peptides (–179 to –185 ppm).[25,33,55] Nonetheless, this chemical shift is far from the bisthiolato Hg(II) complexes. The chemical shift of **d** substituted TRI peptides have been reported to be –313 ppm. Interestingly, the chemical shift of –240 ppm is almost the average of the **a** and **d** site TRI peptides. Thus, it may be possible that the Cys site of α_3D**IV** has properties of both **a** and **d** substituted TRI peptides. However, as previously mentioned that a small change in Hg(II)-S bond distance as well as structural changes may cause a significant change in chemical shift. To further confirm the coordination environment around the Hg(II) ion, ^{199m}Hg PAC studies were performed. ^{199m}Hg PAC spectrum of α_3D**IV** in the presence of 0.4 equivalent Hg(II) at pH 8.6 has one set of resonances with ν_Q of 1.11 GHz and η of 0.4. These values agree strikingly well with the formation of distorted trigonal HgS_3 complexes reported in literature.[57,59,60] Thus the PAC parameters conclusively suggest that at pH 8.6, indeed $Hg^{II}(\alpha_3$D**IV**$)^-$ complex is formed where Hg(II) coordinated in distorted trigonal geometry. At pH 7.4 two ^{199}Hg NMR resonances are observed with chemical shifts of δ = –926 and –240 ppm. (Figure 5-11), Thus, the presence of both the dithiolate-Hg^{II} and trithiolate-Hg^{II} complexes of α_3D**IV** is indicated at this pH. The PAC data (Figure 5-12) was fit for two NQIs with ν_Q of 1.48 GHz (η = 0.15) and 1.11 GHz (η = 0.4) (Table 5-3). These parameters are consistent with the presence of two Hg^{II} species, HgS_2 and HgS_3 as distorted geometries at the intermediate pH 7.4. In the presence of excess Hg^{II}, it is known that trigonal HgS_3 complex bound to three-stranded coiled coil, breaks down and converts to linear complex within a two-stranded coiled coil.[55] ^{199}Hg NMR experiments were performed with α_3D**IV** under excess Hg^{II} conditions to examine if indeed the linear complex is formed. ^{199}Hg NMR of 2.2 mM α_3DIV in the presence of 2 equivalents of Hg^{II} at pH 8.6 has a resonance at δ = –927 ppm, indicating the presence of linear HgS_2 complex under these conditions (Figure 5-11). At pH 8.5 in the presence of 2 equivalents of Hg^{II}, one set of PAC signals is

observed with ν_Q of 1.42 GHz (η = 0.19) (Figure 5-12, Table 5-3). These parameters are consistent with the formation of a linear HgS_2 complex under these conditions.

As mentioned, in the case of the three-stranded coiled coil systems (TRI), in the presence of excess Hg(II) the three-stranded coiled coil converts to two stranded coiled coil due to very high affinity for the formation of linear HgS_2 complex. However, such conversion is not possible for the three-helix bundle α_3D**IV** as all the helices are tether together. The best possible way in which it can form a linear HgS_2 complex under excess conditions of Hg(II) is if α_3D**IV** forms a dimeric complex encapsulating three equivalents of Hg(II), having a molecular formula $Hg_3(\alpha_3D\mathbf{IV})_2$. Two Hg(II) ions would form intramolecular linear HgS_2 complexes within each monomeric unit, while one Hg(II) ion would form a cross-linked HgS_2 complex bridging two molecules of α_3D**IV** using the remaining third Cys from each monomer as its ligating residues. To confirm whether the cross-linked dimer is indeed formed under excess conditions of Hg(II), a protein gel was run. Figure 5-18 shows that the samples 1, 2, 3, 4, 6, and 7 containing apo α_3D**IV**, α_3D**IV** + 0.8 equivalent Cd(II) pH 8.0, α_3D**IV** + 0.8 equivalent Hg(II) pH 5.8, α_3D**IV** + 0.8 equivalent Hg(II) pH 8.6, α_3D**IV** + 0.8 equivalent Pb(II) pH 8.0, α_3D**IV** + 2.0 equivalent Pb(II) pH 8.0 have a single low molecular weight band ~5 kDa. The ESI of apo α_3D**IV** yields a molecular weight of ~ 8 kDa. Thus, the other samples having the similar low molecular weight band in the gel like that of α_3D**IV** are also expected to have a MW of ~ 8 kDa. It is intriguing that all of these samples show a lower molecular weight in the gel than expected. The most likely interpretation for this observation is that there is error associated with the pre-stained molecular weight marker or the fact that the shape of the protein in these samples is not perfectly globular. However, only one of all the samples, sample 5 containing 2.0 equivalents of Hg(II) shows a higher molecular weight band at ~13.5 kDa. The best interpretation for this species is a dimer of α_3D**IV** cross-linked by Hg(II) in linear HgS_2 geometry (expected MW = 16 kDa). Based on the

PyMol model of α_3D**IV**, Cys 67 is separated at a longer distance from Cys 18, and Cys 28 at 4.0, and 4.8 Å, respectively. Thus it is most likely that Cys 67 from two molecules of α_3D**IV** are involved in forming the intermolecular cross-linked dimer with the Hg(II) ion, while Cys 18 and Cys 28 from each protein molecule form intramolecular HgS_2 bonds (Figure 5-18).

Characterization of the Cd(II) Complexes of α_3D**IV**. Having characterized the Hg(II) complexes of α_3D**IV**, characterization of the Cd(II) complexes of α_3D**IV** was performed next. First, direct metal binding titrations were performed under anaerobic conditions in the presence of TCEP. Progress of the titration was monitored by appearance of characteristic LMCT transitions in the UV-Vis spectra of the Cd(II) complex of α_3D**IV** (Figure 5-13). The spectral features as well as the molar extinction coefficient $\Delta\varepsilon_{232nm}$ of 18, 200 $M^{-1}cm^{-1}$ are consistent with the formation of Cd(II)-tristhiolato complex of α_3D**IV**. Stoichiometry of α_3D**IV** : Cd(II) is 1:1. The value of extinction coefficient is lower than that of the Cd(**TRI**L16C)$_3^-$ ($\Delta\varepsilon_{232nm}$ = 22, 600 $M^{-1}cm^{-1}$)[46], Cd(**TRI**L12AL16C)$_3^-$ ($\Delta\varepsilon_{231nm}$ = 21, 200 $M^{-1}cm^{-1}$), and Cd(**TRI**L12C)$_3^-$ ($\Delta\varepsilon_{231nm}$ = 20, 600 $M^{-1}cm^{-1}$).[30] Interestingly, this value is closer to the Coil Ser complex Cd(**CS**L9C)$_3^-$ ($\Delta\varepsilon_{231nm}$ = 19, 000 $M^{-1}cm^{-1}$).[33] Even though the molar extinction coefficient values are smaller than those of TRI peptides, these values are consistent with the fact that all three thiolates of are α_3D**IV** incorporated into the first coordination sphere of Cd(II). From the direct metal binding titration the Cd(II) binding constant to α_3D**IV** is determined to be 2.0×10^7 M^{-1}. This value is very similar to the affinity of Cd(II) binding to the **d** substituted peptides **TRI**L12C (2.6×10^7 M^{-1}),[46] and **CS**L19C (1.7×10^7 M^{-1})[33] as well as the **a** site Coil Ser peptide **CS**L19C (2.7×10^7 M^{-1}),[33] but an order of magnitude lower than the **a** site peptide **TRI**L16C (1.6×10^8 M^{-1})[52]. However, due to very strong binding of Cd(II), these are the lower limits of the binding constants that have been determined from the fitting analysis. Thus, one can conclude that the affinity of Cd(II) binding to α_3D**IV** is very similar to those observed for the TRI and Coil Ser peptides.

Next, pH dependence of the complexation of Cd(II) to α_3D**IV** was examined by UV-Vis spectroscopy and followed by the resulting LMCT transitions of Cd(II)-tristhiolato chromophore at 232 nm. The resulting pH titration curve (Figure 5-10) has a single transition. Fitting of the titration curve to a known model, simultaneous deprotonation of two Cys thiols, yielded a pK_{a2} of 10.6 ± 0.1. This value is ~3 log unit lower than those of TRI peptides,[33,52,53], ~1 log unit lower than the **Grand**L12AL16C.[53] There two possible explanations for the observed lowering of the pK_{a2} in α_3D**IV** compared to the TRI or Grand peptides. According to our model, at low pH the Cd(II) complex that is present is $[CdS(SH)_2]^+$.[53] As α_3D**IV** is a single chain three-helix bundle, this complex can be formed more easily compared to the TRI or Grand peptides as for these peptides to form the derised Cd(II) complex the self assembly of the individual helices would be required to form the three-stranded coiled coil. Secondly, it is possible that the His 72 residue located at the N-terminal end of α_3D**IV**, proximal to the three-Cys site, can potentially lower the pK_{a2} of Cd(II) binding to cysteines. Nonetheless, the pK_{a2} value of 10.6 ± 0.1 is consistent with the coordination mode of Cd^{II} to α_3D**IV** as a 4-coordinate CdS_3O complex (O-being an exogenous water molecule), as for a 3-coordinate CdS_3 complex, a much higher pK_{a2} value is expected, based on the pK_{a2} values observed for 3-coordinate TRI and GRAND peptides.[53]

To confirm the coordination environment around Cd(II), ^{113}Cd NMR and ^{111m}Cd PAC studies were performed. ^{113}Cd NMR spectrum of α_3D**IV** loaded with 0.8 equivalent of ^{113}Cd(II) has two resonances with chemical shifts of 595, and 583 ppm at pH 8.0, indicating the presence of two Cd(II) species (Figure 5-14). These chemical shift values are similar to what has previously been observed for TRI peptides where Cd(II) is bound as a 4-coordinate pseudotetrahedral CdS_3O species.[30,53] The two chemical shifts might arise where the Cd(II) is possibly coordinated to α_3D**IV** as *exo* and *endo* conformers. It is also possible that one of the signals is arising due to coordination of a nearby amino acid residue to the

Cd(II) ion proximal to the Cys site. However, a definitive conclusion regarding the Cd(II) species could not be reached based only on the ^{113}Cd NMR chemical shift values. To confirm the coordination environment and geometry of Cd(II) complex of α_3D**IV**, ^{111m}Cd PAC studies were performed. ^{111m}Cd PAC spectrum of α_3D**IV** bound to Cd(II) shows three NQIs with angular frequencies ω_0 = 0.35 (η = 0), 0.268 (η = 0.18) and 0.17 (η = 0.5) rad/ns at pH 8.1 (Figure 5-15, Table 5-4). The first two peaks at 0.35 rad/ns and 0.268 rad/ns agree strikingly well with a typical CdS_3O signal in an *exo* and *endo* conformation, respectively. The PAC parameters for Cd(**TRI**L12AL16C)$_3^-$ where Cd(II) is bound as 100% CdS_3O in an *exo* conformation are ω_0 = 0.3405, η = 0.141.[30] Thus the first peak for α_3D**IV** at ω_0 = 0.35 rad/ns agree well with a pure 4-coordinate CdS_3O species where Cd(II) is bound as an *exo* conformer. This is also supported by PAC data from many other TRI peptides.[53] The corresponding PAC parameters for Cd(**TRI**L12AL16C)$_3^-$, where Cd(II) is bound as 100% CdS_3O as an *endo* conformation, are ω_0 = 0.274 rad/ns, η = 0.19 (Suh *et. al.* unpublished results). Thus, the second peak for α_3D**IV** at ω_0 = 0.268 rad/ns (η = 0.18) is very similar to a pure 4-coordinate CdS_3O species where Cd(II) is bound as an *endo* conformer. The lowest frequency NQI at 0.17 rad/ns contributes to ~52% of the total amplitude and falls in a range that several structures may give. Interestingly, a high value of the asymmetry parameter (η = 0.5) is observed indicating asymmetric nature of the ligands to Cd(II). As thiolates are present, most likely they will be among the ligands, but almost certainly there are also other ligands as a pure trigonal CdS_3 species would give a higher frequency signal.[53] This signal can be best assigned to a CdS_3N species where N is His 72 located at the C-terminal end of the bundle (Figure 5-15). Even though the chemical shift of 595 ppm in the ^{113}Cd NMR spectrum is lower than what has been reported for a CdS_3N species,[64] PAC studies show that a change in Cd-S bond length of 0.01 Å can cause a change in chemical shift of ~20 ppm (Hemmingsen *et.al.* unpublished results). Tentatively, the ^{113}Cd NMR resonance at δ = 595 ppm can be best assigned to the CdS_3N species (N= His 72).

Characterization of the Pb(II) Complexes of α_3D**IV**. Finally Pb(II) complexes of α_3D**IV** were characterized by UV-Vis spectroscopy. First, Pb(II) titration to α_3D**IV** was performed under anaerobic conditions in the presence of the reducing agent TCEP at pH 8.0 (Figure 5-16). The overall UV-Vis spectral features are consistent with the formation of Pb(II)-tristhiolato complex of α_3D**IV** (Figure 5-16). Extinction coefficients at different wavelengths are determined to be 18 000 (λ_{236nm}), 14 400 (λ_{260nm}), 9 100 (λ_{278nm}), and 3 150 (λ_{346nm}) $M^{-1}cm^{-1}$. For Pb(**TRI**L16C)$_3^-$ the extinction coefficients at these wavelengths were 18, 500 (λ_{236nm}), 16, 500 (λ_{260nm}), 14, 500 (λ_{278nm}), and 3, 400 (λ_{346nm}) $M^{-1}cm^{-1}$.[52] At λ_{236nm} and λ_{346nm} the values of the extinction coefficients of α_3D**IV** are very similar to **TRI**L16C and are consistent with the formation of Pb(α_3D**IV**)$_3^-$ complex in trigonal pyramidal geometry. α_3D**IV** formed a 1:1 complex with Pb(II). A Pb(II) binding constant of 3.1×10^7 M^{-1} was determined from the analysis of the titration data. This value is very similar to the Pb(II) binding affinity for **TRI**L16C (1.2×10^7 M^{-1}).[52] In the presence of more than stoichiometric amount of Pb(II), the intensity of the peak at λ_{346nm} decreased, indicating that the PbS_3 chromophore is dissipating. The intensity of the peak at λ_{236nm} increased continuously as well as the UV signal between 300 and 330 nm. Other spectral features of the UV-Vis spectra in the presence of excess Pb(II) are similar to those of the stoichiometric conditions. The appearance of three isosbestic points at λ_{330nm}, λ_{280nm} and λ_{260nm} under excess conditions of Pb(II) suggests that other chomophore(s) are being formed at the expense of PbS_3 chromophore. Excess Pb(II) may coordinate to other amino acid side chains of the protein. However, the identity of the new Pb(II) chromophore could not be established based on the UV-Vis data. Next, pH dependent complexation of Pb(II) to α_3D**IV** was performed by UV-Vis spectroscopy by monitoring the LMCT transition at 236 nm. The resulting titration curve (Figure 5-10) has a single inflexion point and was fit to a known model: simultaneous deprotonation of two thiol protons of Cys, yielding a pK_{a2} value of 10.2 ± 0.1. This pK_{a2} value is ~2 log unit lower than that of **TRI**L16C.[52]

Nonetheless, the pK_{a2} value of 10.2 ± 0.1 supports the formation of $Pb^{II}(\alpha_3D\mathbf{IV})^-$ complex as trigonal pyramidal geometry with the incorporation of all three Cys thiolates to the metal.

It may be noted that in order to achieve 1:1 stoichiometry of $\alpha_3D\mathbf{IV}$ with Cd(II), Hg(II), and Pb(II) it was necessary to maintain strictly anaerobic conditions in the presence of ~2 fold excess of the reducing agent TCEP. Samples and stock solutions of metals were prepared inside an inert atmosphere box, and the titrations were performed in anaerobic cuvettes. Titrations under aerobic conditions in the presence and absence of ~2 fold excess of TCEP gave a stoichiometry of $\alpha_3D\mathbf{IV}$: M(II) of 1: 0.8. The best interpretation of this fact is that $\alpha_3D\mathbf{IV}$ is susceptible to oxidation (~20%). For our studies with TRI peptides strict anaerobic conditions were not required to be maintained to achieve proper stoichiometry of metal binding to these peptides.

Conclusions: In conclusion, in this chapter I have described successful engineering of metal binding sites containing thiol-rich cysteine residues in an existing *de novo* antiparallel three-helix bundle, yielding a well folded protein $\alpha_3D\mathbf{IV}$ in solution capable of binding heavy metals with high affinity ($>10^7$ M^{-1}). Characterization of different metallo derivatives of $\alpha_3D\mathbf{IV}$ shows that it binds Cd^{II} as a 4-coordinate pseudotetrahedral CdS_3O/CdS_3N species, Hg^{II} as distorted linear HgS_2 and distorted trigonal HgS_3 complexes depending on experimental conditions and Pb^{II} as trigonal pyramidal PbS_3 geometries. The physical and spectroscopic properties of Cd^{II}, Hg^{II} and Pb^{II} complexes of $\alpha_3D\mathbf{IV}$ are very similar to the parallel three-stranded coiled coils. Thus the objective of preparing a single polypeptide chain capable of binding metals with high affinity and predefined coordination geometry is achieved. The current study is the first such example where a Cys_3 metal binding site has been redesigned in a versatile protein structural motif, three-helix bundle, which is capable of binding different heavy metals with high affinity. Future studies will explore the possibilities of

preparing similar constructs containing asymmetric metal binding sites as mimics of many natural proteins and making catalytic metalloproteins.

References:

(1) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chemical Reviews* **1996**, *96*, 2239-2314.
(2) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855-862.
(3) Gibney, B. R.; Dutton, P. L. *Adv. Inorg. Chem, Vol 51* **2001**, *51*, 409-455.
(4) Reedy, C. J.; Gibney, B. R. *Chem. Rev.* **2004**, *104*, 617-649.
(5) Kennedy, M. L.; Petros, A. K.; Gibney, B. R. *J. Inorg. Biochem.* **2004**, *98*, 727-732.
(6) Cochran, F. V.; Wu, S. P.; Wang, W.; Nanda, V.; Saven, J. G.; Therien, M. J.; DeGrado, W. F. *Journal of the American Chemical Society* **2005**, *127*, 1346-1347.
(7) Ghirlanda, G.; Osyczka, A.; Liu, W.; Antolovich, M.; Smith, K. M.; Dutton, P. L.; Wand, A. J.; DeGrado, W. F. *Journal of the American Chemical Society* **2004**, *126*, 8141-8147.
(8) Guss, J. M.; Harrowell, P. R.; Murata, M.; Norris, V. A.; Freeman, H. C. *Journal of Molecular Biology* **1986**, *192*, 361-387.
(9) Mitchell Guss, J.; Freeman, H. C. *Journal of Molecular Biology* **1983**, *169*, 521-563.
(10) Adman, E. T.; Stenkamp, R. E.; Sieker, L. C.; Jensen, L. H. *Journal of Molecular Biology* **1978**, *123*, 35-47.
(11) Baker, E. N. *Journal of Molecular Biology* **1988**, *203*, 1071-1095.
(12) Korszun, Z. R. *Journal of Molecular Biology* **1987**, *196*, 413-419.
(13) Sly, W. S.; Hu, P. Y. *Annual Review of Biochemistry* **1995**, *64*, 375-401.
(14) Deutsch, H. F. *International Journal of Biochemistry* **1987**, *19*, 101-113.
(15) Lombardi, A.; Bryson, J. W.; DeGrado, W. F. *Biopolymers* **1997**, *40*, 495-504.
(16) Marsh, E. N. G.; DeGrado, W. F. *Proc. Natl. Acad. Sci.* **2002**, *99*, 5150-5154.
(17) Regan, L. *Trends in Biochemical Sciences* **1995**, *20*, 280-285.
(18) Regan, L. *Annual Review of Biophysics and Biomolecular Structure* **1993**, *22*, 257-281.
(19) Tainer, J. A.; Roberts, V. A.; Getzoff, E. D. *Current Opinion in Biotechnology* **1992**, *3*, 378-387.
(20) Hellinga, H. W. *Design of metalloproteins*; Wiley-Liss: New York, 1996.
(21) Hellinga, H. W. *Current Opinion in Biotechnology* **1996**, *7*, 437-441.
(22) Hellinga, H. W. *Folding & Design* **1998**, *3*, R1-R8.
(23) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. *Dalton Transactions* **2009**, 2271-2280.
(24) Pecoraro, V. L.; Peacock, A. F. A.; Iranzo, O.; Luczkowski, M. In *Bioinorganic Chemistry*; American Chemical Society, 2009.
(25) 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.; Pecoraro, V. L. *J. Am. Chem. Soc.* **1997**, *119*, 6195-6196.

(26) Touw, D. S.; Nordman, C. E.; Stuckey, J. A.; Pecoraro , V. L. *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969-11974.
(27) Wright, J. G.; Natan, M. J.; MacDonnell, F. M.; Ralston, D. M.; O'Halloran, T. V. *Progress in Inorganic Chemistry:Bioinorganic Chemistry* **1990**, *38*, 323-412.
(28) Shi, W.; Dong, J.; Scott, R. A.; Ksenzenko, M. Y.; Rosen, B. P. *J. Biol. Chem.* **1996**, *271*, 9291-9297.
(29) Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Cavet, J. S.; Dennison, C.; Graham, A. I.; Harvie, D. R.; Robinson, N. J. *Journal of Biological Chemistry* **2007** *282* 30181-30188
(30) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2006**, *45*, 2864-2868.
(31) Iranzo, O.; Cabello, C.; Pecoraro, V. L. *Angew. Chem., Int. Ed.* **2007**, *46*, 6688-6691.
(32) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proceedings of the National Academy of Sciences* **2008**, *105*, 16566-16571.
(33) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959-9973.
(34) Marsh, E. N. G.; DeGrado, W. F. *Proc. Natl. Acad. Sci., U.S.A.* **2002**, *99*, 5150-5154.
(35) Summa, C. M.; Rosenblatt, M. M.; Hong, J.-K.; Lear, J. D.; DeGrado, W. F. *J. Mol. Biol.* **2002**, *321*, 923-938.
(36) Kaplan, J.; DeGrado, W. F. *Proc. Natl. Acad. Sci., U.S.A.* **2004**, *101*, 11566-11570.
(37) Faiella, M.; Andreozzi, C.; de Rosales, R. T. M.; Pavone, V.; Maglio, O.; Nastri, F.; DeGrado, W. F.; Lombardi, A. *Nat. Chem. Biol* **2009**, *5*, 882-884.
(38) Walsh, S. T. R.; Cheng, H.; Bryson, J. W.; Roder, H.; DeGrado, W. F. *Proc. Natl. Acad. Sci., U.S.A.* **1999**, *96*, 5486-5491.
(39) Walsh, S. T. R.; Sukharev, V. I.; Betz, S. F.; Vekshin, N. L.; DeGrado, W. F. *Journal of Molecular Biology* **2001**, *305*, 361-373.
(40) Walsh, S. T. R.; Lee, A. L.; DeGrado, W. F.; Wand, A. J. *Biochemistry* **2001**, *40*, 9560-9569.
(41) Walsh, S. T. R.; Cheng, R. P.; Wright, W. W.; Alonso, D. O. V.; Daggett, V.; Vanderkooi, J. M.; DeGrado, W. F. *Protein Science* **2003**, *12*, 520-531.
(42) Zhu, Y.; Alonso, D. O. V.; Maki, K.; Huang, C.-Y.; Lahr, S. J.; Daggett, V.; Roder, H.; DeGrado, W. F.; Gai, F. *Proceedings of the National Academy of Sciences of the United States of America* **2003** *100* 15486-15491
(43) Regan, L. *Advances in Molecular and Cell Biology* **1997**, *22A*, 51-80.
(44) Walsh, S. T.; Cheng, R. P.; Wright, W. W.; Alonso, D. O.; Daggett, V.; Vanderkooi, J. M.; DeGrado, W. F. *Protein Science* **2003**, *12*, 520-531.
(45) Aitken, A.; Learmonth, M. l. In *The Protein Protocols Handbook*; Walker, J. M., Ed.; Humana Press, 1996.
(46) Matzapetakis, M.; Farrer, B. T.; Weng, T.-C.; Hemmingsen, L.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2002**, *124*, 8042-8054.
(47) Agashe, V. R.; Udgaonkar, J. B. *Biochemistry* **1995**, *34*, 3286-3299.

(48) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1988**, *27*, 8063-8068.
(49) Santoro, M. M.; Bolen, D. W. *Biochemistry* **1992**, *31*, 4901-4907.
(50) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro , V. L. *Journal of the American Chemical Society* **2010**, *Submitted*.
(51) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *Journal of the American Chemical Society* **2010**, *132*, 13240.
(52) Matzapetakis, M.; Ghosh, D.; Weng, T.-C.; Penner-Hahn, J. E.; Pecoraro, V. L. *J. Biol. Inorg. Chem.* **2006**, *11*, 876-890.
(53) Iranzo, O., Lee, K.H., Jakusch, T., Hemmingsen, L., Pecoraro, V. L. *Chem. Eur. J.,* **2009**, *15*, 3761-3772.
(54) Luczkowski, M.; Stachura, M.; Schirf, V.; Demeler, B.; Hemmingsen, L.; Pecoraro, V. L. *Inorganic Chemistry* **2008**, *47*, 10875-10888.
(55) Iranzo, O.; Thulstrup, P. V.; Ryu, S.-B.; Hemmingsen, L.; Pecoraro, V. L. *Chem. Eur. J.* **2007**, *13*, 9178-9190.
(56) Huffman, D. L.; Utschig, L. M.; O'Halloran, T. In *Metal Ions in Biological Systems*; Sigel, H., Ed.; Marcel Dekker, Inc.: New York, 1997; Vol. 34.
(57) Butz, T.; Troger, W.; Pohlmann, T.; Nuyken, O. *Z. Naturforsch* **1992**, *47a*, 85-88.
(58) Dieckmann, G. R., Ph.D. Thesis, University of Michigan, 1995.
(59) Lippert, C.; Troger, W.; Butz, T. 30th Zakopane School of Physics, Krakow, Poland, 1995.
(60) Troger, W. *Hyperfine Interactions* **1999**, *120/121*, 117-128.
(61) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angewandte Chemie International Edition* **2009**, *48*, 7371-7374.
(62) *Protein Structure: A Practical Approach*; Creighton, T. E., Ed.; Oxford University Press: Oxford, 1997.
(63) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annual Review of Biochemistry* **1999**, *68*, 779-819.
(64) Coleman, J. E. In *Metallobiochemistry Part D: Physical and Spectroscopic Methods for Probing Metal Ion Environment in Metalloproteins*; Volume 227 ed.; Vallee, J. F. R. a. B. L., Ed.; Academic Press, 1993.

Table 5-1. Sequences of α_3D**I** to α_3D**IV**.[a]

Peptide	**Sequence**			
α_3D**I**	MGS WAEC K QR	LAAIKTR	LQAL	GG
F7C-L42C-L56C	SEAE LAAF E KE	IAAFESE	CQAY	KGKGNPE
	V E AC R KE	AAAIRDE	LQAY	RHN
α_3D**II**	MGS WAEF K QR	CAAIKTR	LQAL	GG
L11C-F38C-A60C	SEAE LAAF E KE	IAACESE	LQAY	KGKGNPE
	V E AL R KE	CAAIRDE	LQAY	RHN
α_3D**III**	MGS WAEF K QR	LAACKTR	LQAL	GG
I14C-I35C-I63C	SEAE LAAF E KE	CAAFESE	LQAY	KGKGNPE
	V E AL R KE	AAACRDE	LQAY	RHN
α_3D**IV**	MGS WAEF K QR	LAAIKTR	CQAL	GG
L18C-L28C-L67C	SEAE CAAF E KE	IAAFESE	LQAY	KGKGNPE
	V E AL R KE	AAAIRDE	CQAY	RHN

[a] Residues in red represent mutations from the WT α_3D.

Table 5-2. Physical Parameters of Cd^{II}, Hg^{II} and Pb^{II} Complexes of α_3D**IV**.

Complex	**UV/Vis** λ_{nm} **(**$\Delta\varepsilon$ **M** $^{-1}$**cm**$^{-1}$**)**	**NMR** δ **(ppm)**		**Apparent** pK_a	**Binding Constants** (K_b) (M^{-1})[a]
		^{113}**Cd**	^{199}**Hg**		
$Cd(\alpha_3D\mathbf{IV})^-$	232 (18, 200)	583 595		10.6 ± 0.1[b]	2.0×10^7
$Hg(\alpha_3D\mathbf{IV})^-$	247 (12,500) 265 (8,400) 295 (3,900)		–244	7.1 ± 0.1[c]	
$Hg(\alpha_3D\mathbf{IV})$[d]	240 (850)		–938[e] –926[f]		
$Hg_3(\alpha_3D\mathbf{IV})_2$	240		–927[g]		
$Pb(\alpha_3D\mathbf{IV})^-$	236 (18,000) 260 (14,400) 278 (9,100) 346 (3,150)			10.2 ± 0.1[b]	3.1×10^7

[a] Model used to obtain binding constants is $M^{II} + (\alpha_3D\mathbf{IV})^{3-} \rightleftharpoons M^{II}(\alpha_3D\mathbf{IV})^- (K_b)$. These values represent the lower limit of K_b. [b] Model used to obtain pK_{a2} values for Cd^{II} and Pb^{II} is $M^{II}(\alpha_3D\mathbf{IV}S(SH)_2)^+ \rightleftharpoons M^{II}(\alpha_3D\mathbf{IV})^- + 2H^+ (K_{a2})$. [c] Model used to obtain the pK_a for Hg^{II} is $Hg^{II}(\alpha_3D\mathbf{IV}S_2(SH)) \rightleftharpoons Hg^{II}(\alpha_3D\mathbf{IV})^- + H^+ (K_a)$. [d] Linear HgS_2 complex of $\alpha_3D\mathbf{IV}$. [e] ^{199}Hg NMR δ of linear HgS_2 complex at pH 5.8. [f] ^{199}Hg NMR δ of linear HgS_2 complex at pH 7.4. [g] ^{199}Hg NMR δ of linear HgS_2 complex of $Hg_3(\alpha_3D\mathbf{IV})_2$ pH 8.6 and 5.8.

Table 5-3. Parameters fitted to the ^{199m}Hg PAC data.[a]

α_3DIV (μM)	Hg^{II} (μM)	pH	ν_Q (GHz)	η	$\Delta\omega_0/\omega_0\times100$	$1/\tau_c$ (μs^{-1})	A×100	χ_r^2
200	80	5.8	1.48(2)	0.15(5)	2(2)	18(21)	12(1)	0.63
200	80	7.5	1.48[b]	0.15[b]	2[b]	10(18)	4(1)	0.67
			1.11[b]	0.40[b]	6[b]		11(1)	
200	80	8.6	1.11(2)	0.40(3)	6(2)	0(20)	10(1)	0.62
200	400	8.5	1.42(2)	0.19(4)	2(2)	17(23)	11(1)	0.58

[a] Numbers in the parenthesis are the standard deviations of the fitted parameters. [b] Fixed in the fit.

Table 5-4. Parameters fitted to the ^{111m}Cd PAC data.[a]

Peptide ($C_{CdII}/C_{peptide}$)	**ω_0 (rad/ns)**	**η**	**$\Delta\omega_0/\omega_0 \times 100$**	**$1/\tau_c$ (μs^{-1})**	**A ×100**	**χ_r^2**
α_3D**IV** (1/12)	0.350(6)	0.0(1)	14(3)	4.5(6)	5(1)	1.36
	0.268(4)	0.18(7)	1(2)	4.5(6)	0.3(3)	
	0.17(2)	0.5(2)	24(3)	4.5(6)	5.8(8)	

[a] Numbers in the parenthesis are the standard deviations of the fitted parameters.

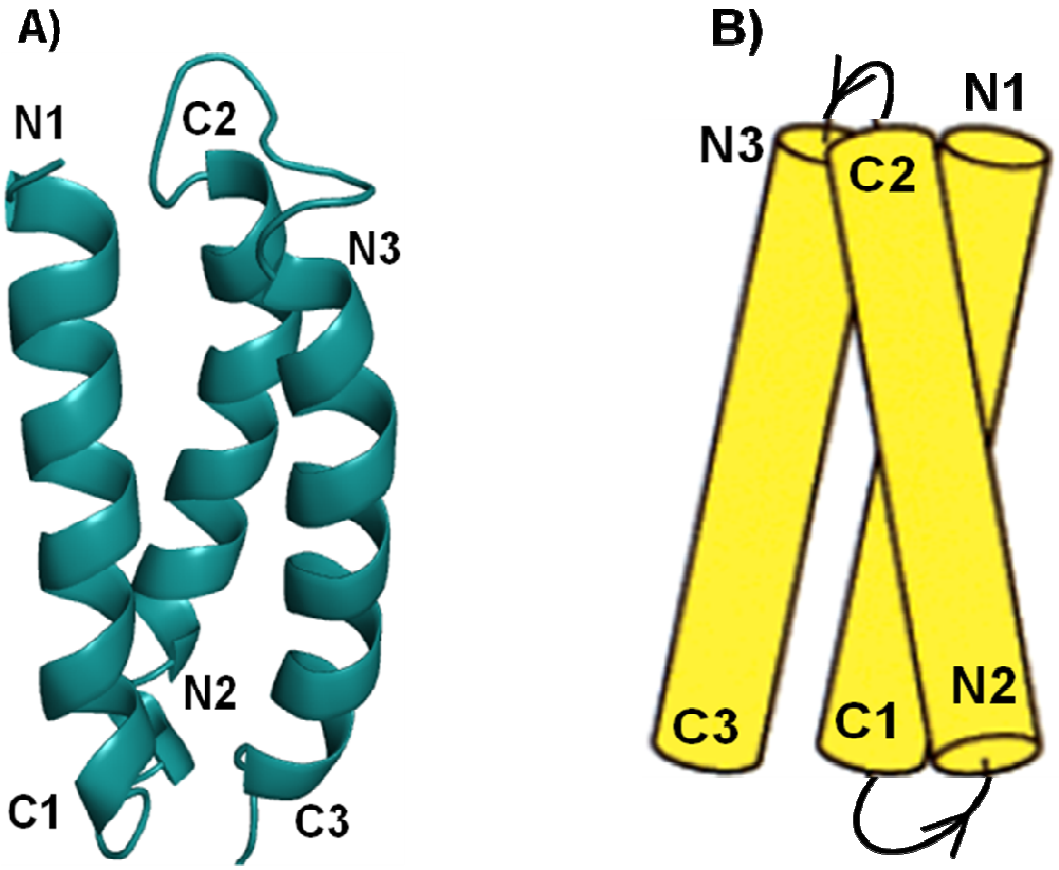


Figure 5-1. **A)** Solution structure of the three-helix bundle protein α_3D (PDB ID 2A3D, ref 19). Figure was generated in PyMol. N1 - C1, N2 - C2, N3 - C3 represent N and C terminus of the 1st, 2nd and 3rd helix, respectively. **B)** Cartoon representation of an anticlockwise three-helix bundle. Figure is reprinted in part from the reference "*De Novo* Design of ProteinssWhat Are the Rules?", Lars Baltzer, Helena Nilsson and Jonas Nilsson, *Chem. Rev.,* **2001**, 101, 3153.

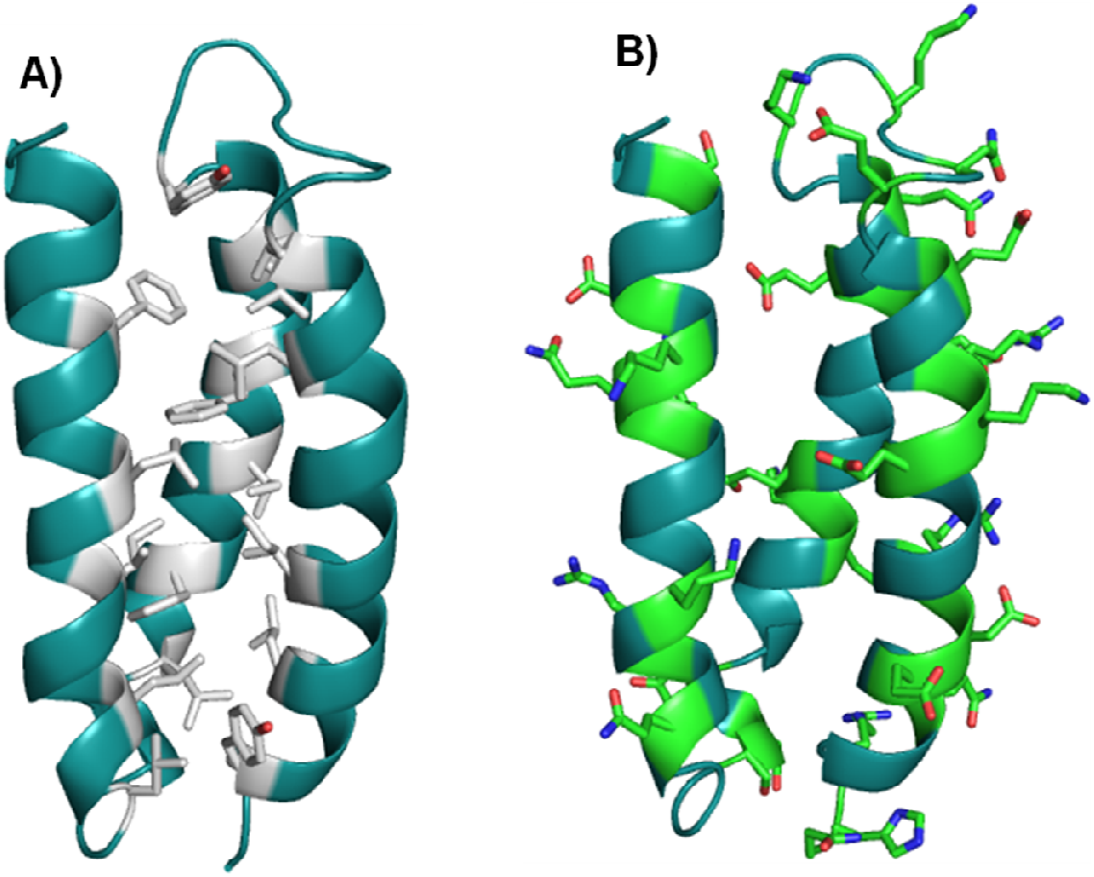


Figure 5-2. **A)** A PyMol figure of α_3D showing the hydrophobic amino acids as sticks (color white) that form the core of the bundle. Protein backbone is shown as blue. **B)** Hydrophilic amino acids of α_3D are shown as green. Protein backbone is shown as blue.

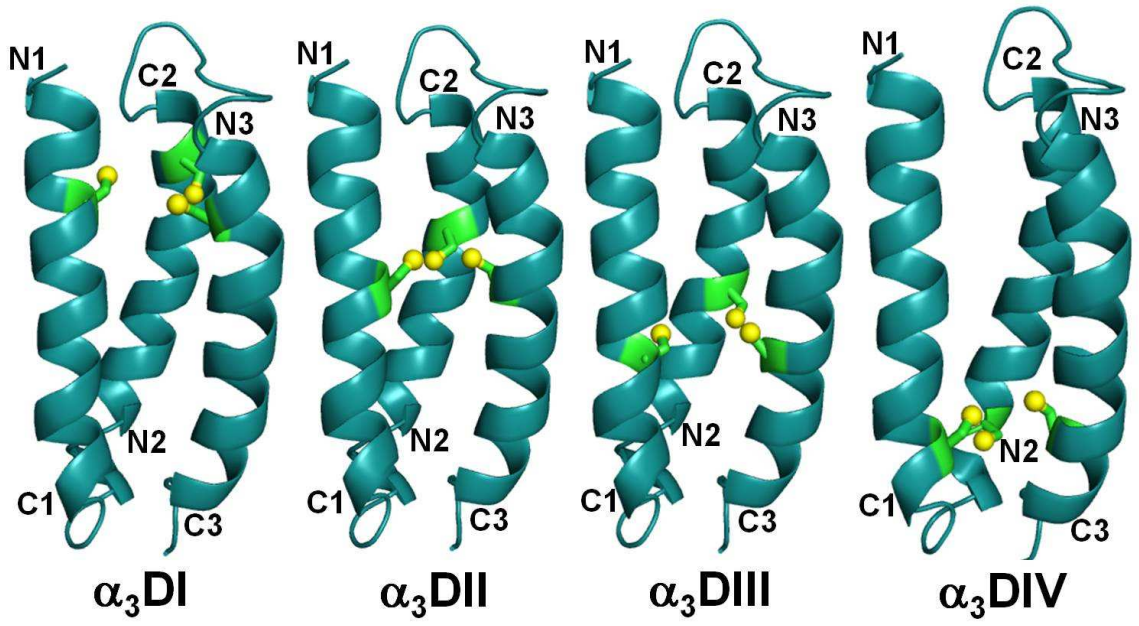


Figure 5-3. Models of α_3D**I**, α_3D**II**, α_3D**III** and α_3D**IV** generated from the NMR structure of α_3D. Respective amino acids from α_3D were mutated to Cys using the program PyMol to generate each mutant. Shown are the most preferred rotamers of cysteines in each mutant and represented as ball and sticks in green and yellow colors. Protein backbone is shown in blue. N1 - C1, N2 - C2, N3 - C3 represent N and C terminus of 1st, 2nd and 3rd helix, respectively.

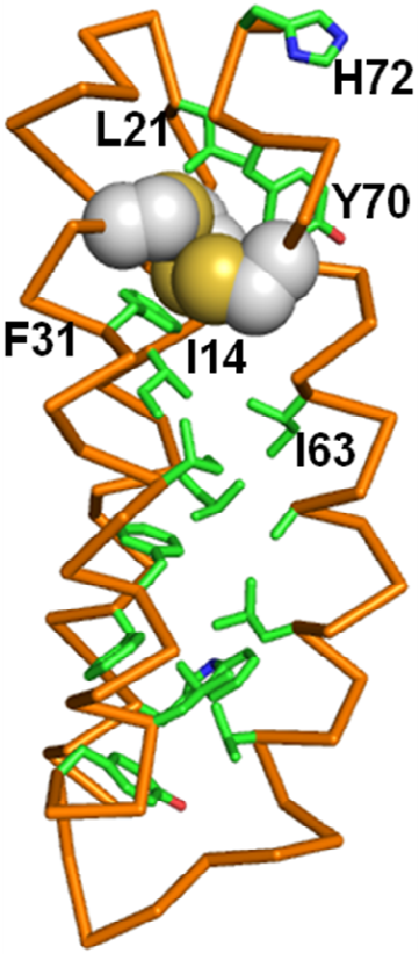


Figure 5-4. PyMol model of α_3D**IV** generated from the NMR structure of α_3D. Cys residues, located at the C-terminal end of the bundle are shown as spheres. Protein backbone is shown in orange. The Cys site can be considered to be located in a hydrophobic “box” formed by the hydrophobic residues F31, I14, I63, L21 and Y70, shown as sticks.

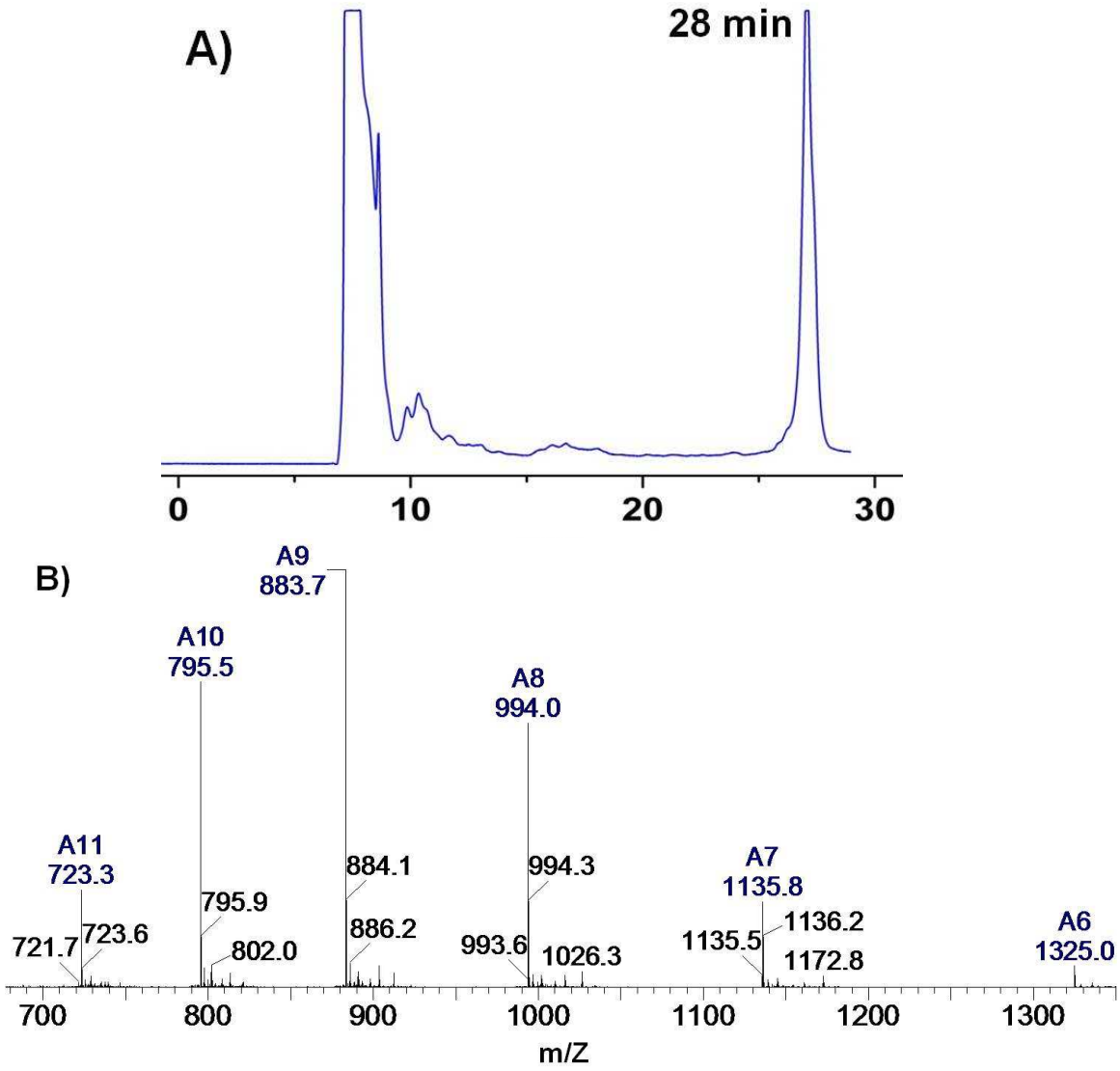


Figure 5-5. **A)** HPLC trace of pure α_3D**IV** eluted with retention time of 28 min. **B)** Electro Spray Ionization mass spectrum of α_3D**IV.** The MW of 7945.1 Da determined from the m/Z distribution profile corresponds to α_3D**IV** with the deletion of the first Met.

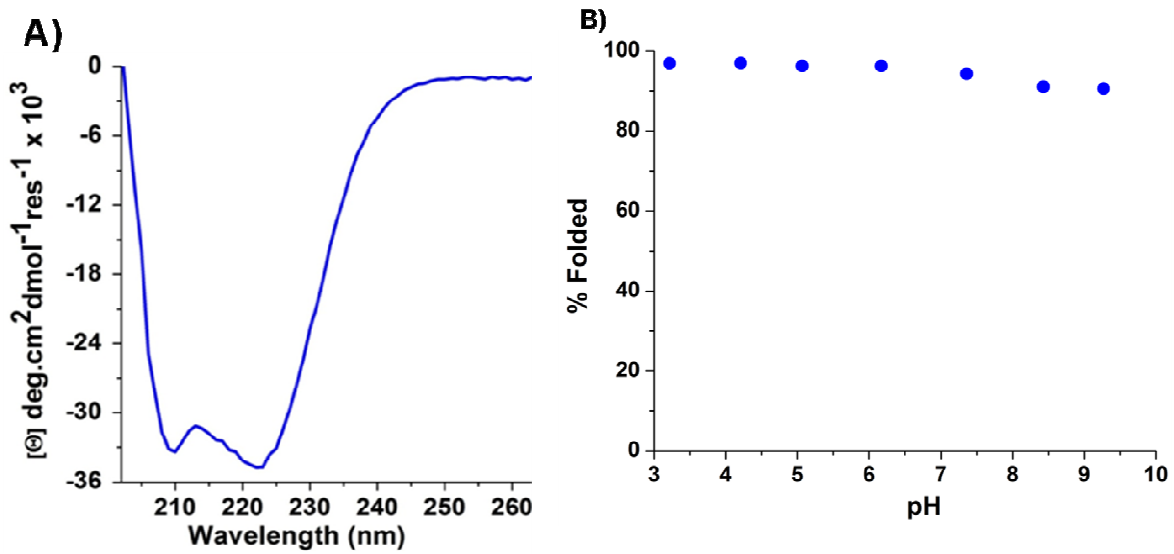


Figure 5-6. **A)** CD spectrum of α_3D**IV** at pH 8.0 recorded at 25°C. The double minima at 222 and 208 nm are representative of α helical structure. Molar ellipticity of –34726 deg cm^2 dmol^{-1} res^{-1} at 222 nm corresponds to a 97% helical structure. **B)** Folding of α_3D**IV** as a function of pH shows that α_3D**IV** is well folded between pH 3 and 9.

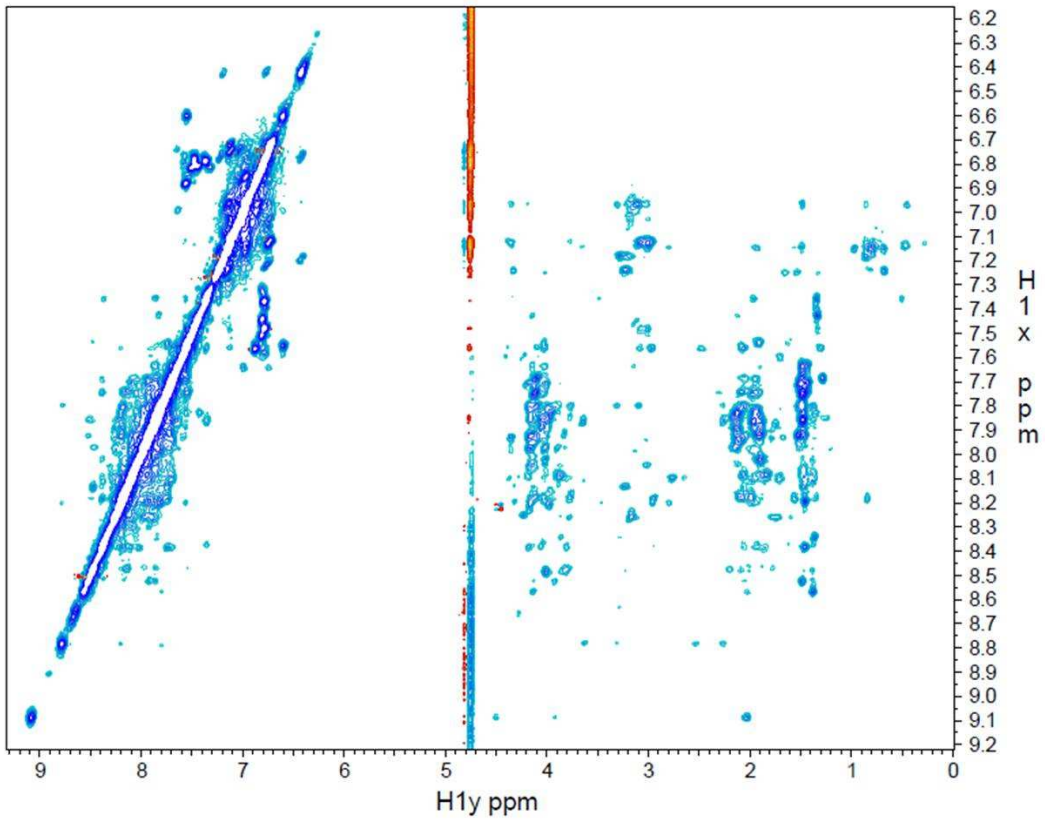


Figure 5-7. ^{1}H-^{1}H NOESY spectrum of α_3D**IV** at pH 6 shows chemical shift dispersion characteristic of a well folded α– helical structure.

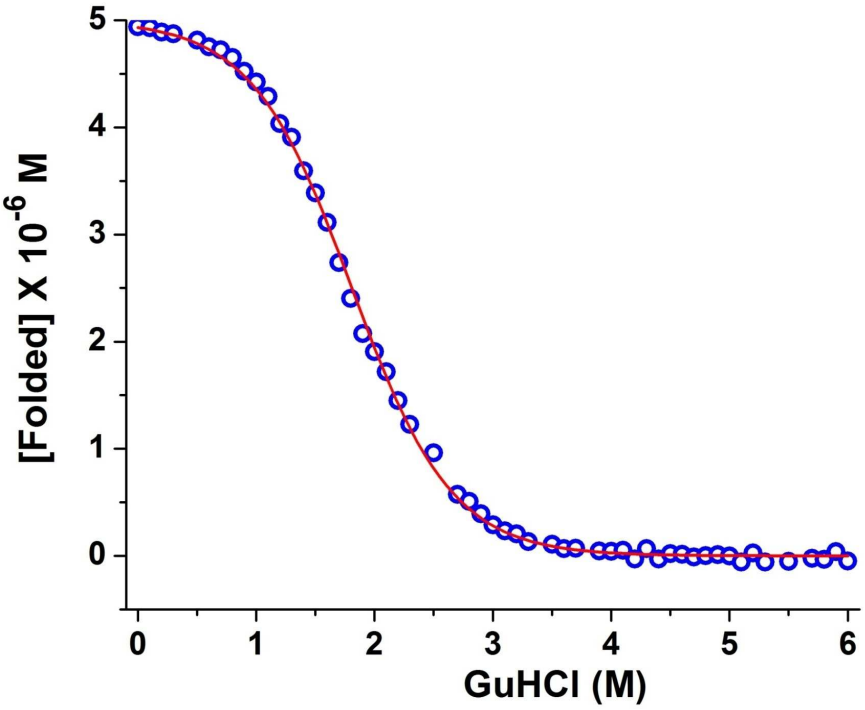


Figure 5-8. GuHCl denaturation titration curve of α_3D**IV** plotted as a function of concentration of folded protein vs. concentration of GuHCl at pH 8.0. The solid line represents fit to the experimental data.

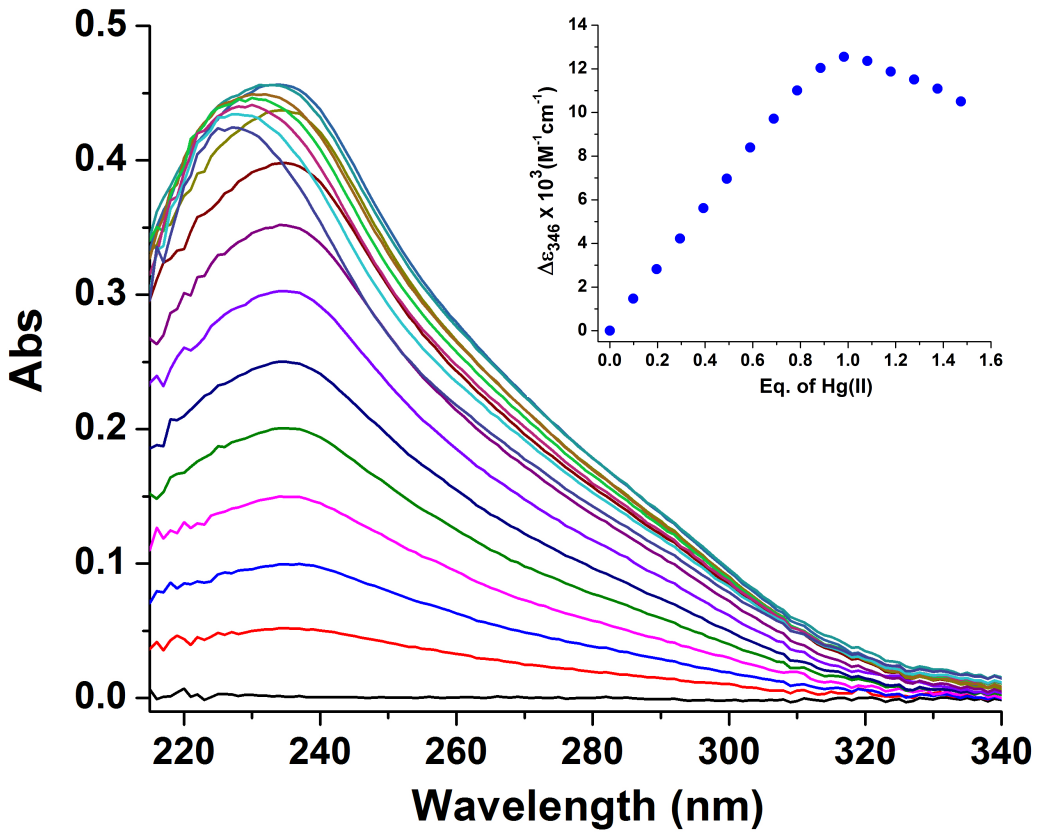


Figure 5-9. UV/Vis titration of $HgCl_2$ to a solution containing 30 μM α_3D**IV** and 60 μM TCEP in 50 mM CHES buffer at pH 8.6. Data are plotted as Abs vs. wavelength. The inset of the figure shows the titration curve plotted as $\Delta\varepsilon_{247nm}$ vs. equivalents of Hg^{II} added.

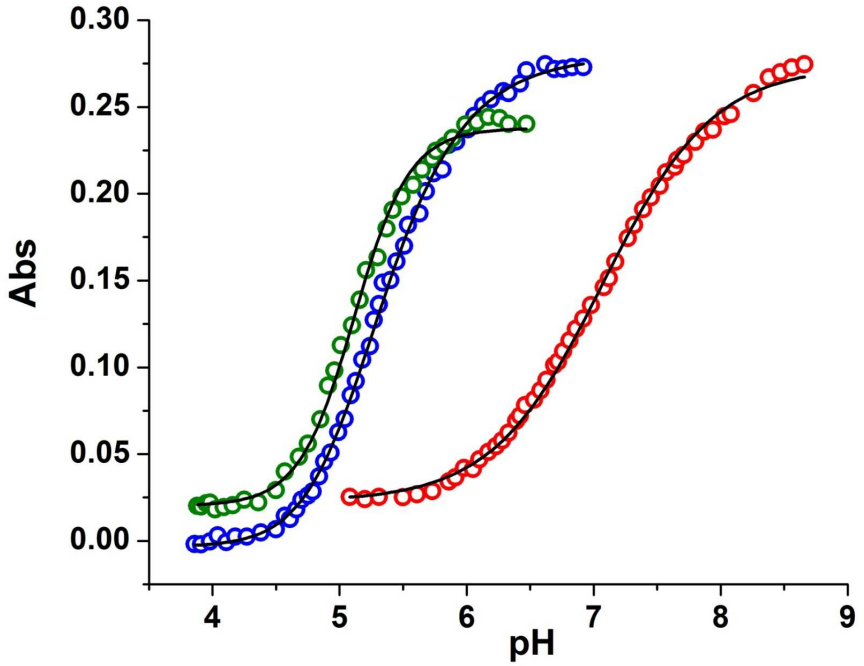


Figure 5-10. pH dependence of the binding of 1 equivalent of Cd^{II} (blue circles), Hg^{II} (red circles) and Pb^{II} (green circles) to 20 μM, 30 μM and 20 μM α_3D**IV**, respectively, along with the fits of the experimental data. Experiments were followed by monitoring UV/Vis absorbance due to LMCT bands at 232, 247 and 236 nm for Cd^{II}, Hg^{II} and Pb^{II}, respectively. Cd^{II} and Pb^{II} titration data were fit to the model simultaneous dissociation of two Cys thiol protons; Hg^{II} was fit to a single 1-H step corresponding to the formation of HgS_3 from $HgS_2(SH)$.

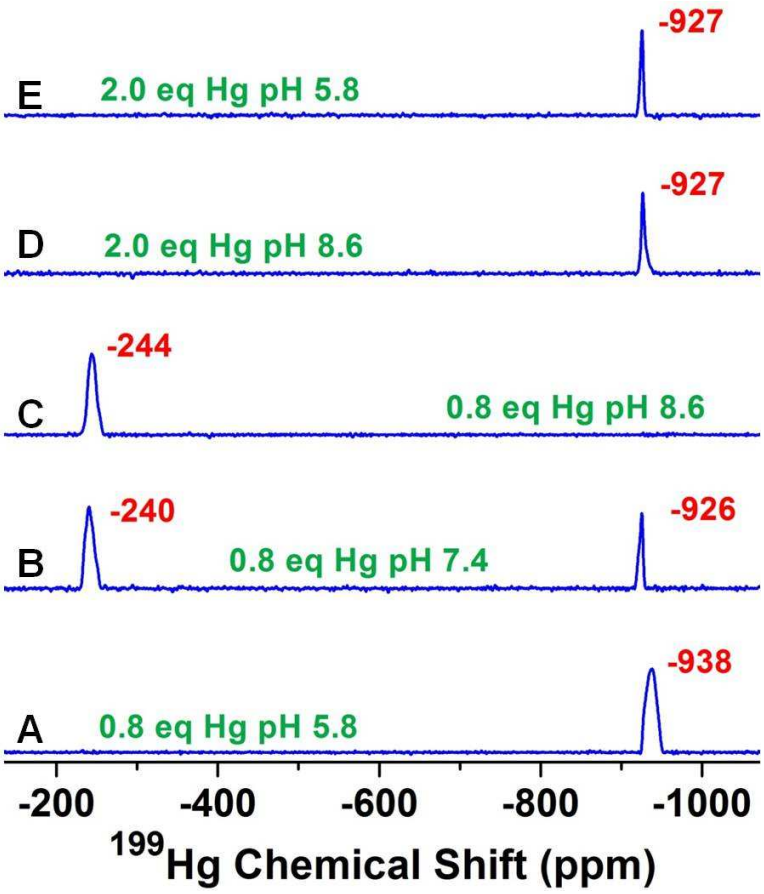


Figure 5-11. ^{199}Hg NMR spectra of solutions containing 2.93 mM α_3DIV and 0.8 equivalent of $^{199}Hg(NO_3)_2$ at pH 5.8 (A), 7.4 (B), 8.6 (C); and 2.95 mM α_3DIV and 2 equivalents of $^{199}Hg(NO_3)_2$ at pH 8.6 (D) and 5.8 (E).

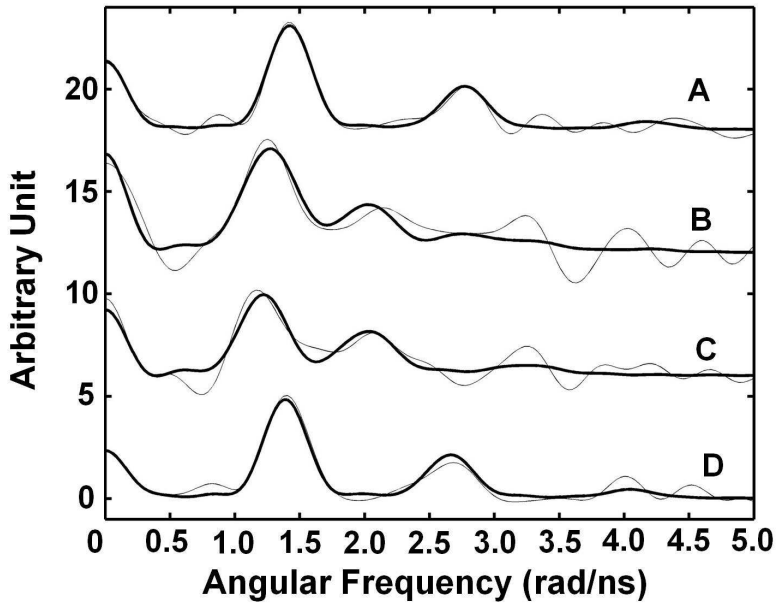


Figure 5-12. ^{199m}Hg PAC spectra of α_3DIV (faint lines) under different conditions along with the fits of the Fourier transformed data (dark lines). Sample conditions were A) 200 μM α_3DIV, 80 μM $HgCl_2$, 100 mM phosphate buffer at pH 5.8, B) 200 μM α_3DIV, 80 μM $HgCl_2$, 100 mM phosphate buffer at pH 7.4, C) 200 μM α_3DIV, 80 μM $HgCl_2$, 100 mM CHES buffer at pH 8.6 and D) 200 μM α_3DIV, 400 μM $HgCl_2$, 100 mM CHES buffer at pH 8.5.

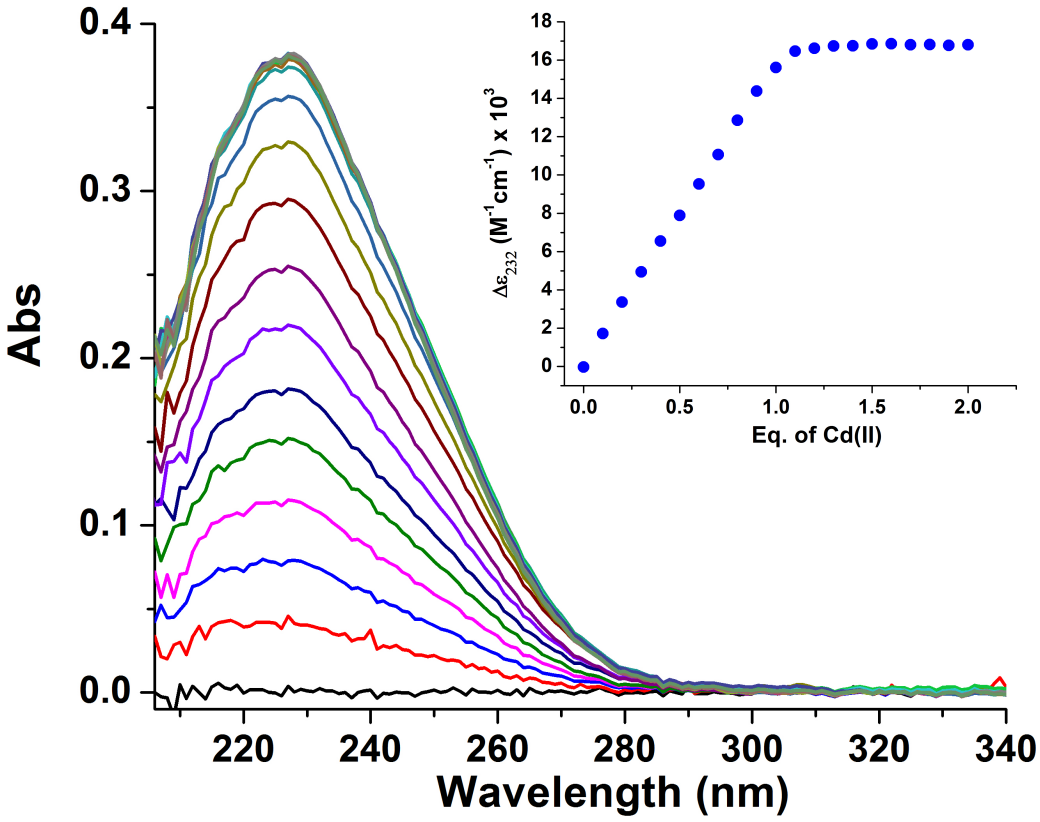


Figure 5-13. UV/Vis titration of $CdCl_2$ to a solution containing 20 μM α_3D**IV** and 40 μM TCEP in 50 mM TRIS buffer at pH 8.0. Data are plotted as Abs vs. wavelength. The inset of the figure shows the titration curve plotted as $\Delta\varepsilon_{232nm}$ vs. equivalents of Cd^{II} added.

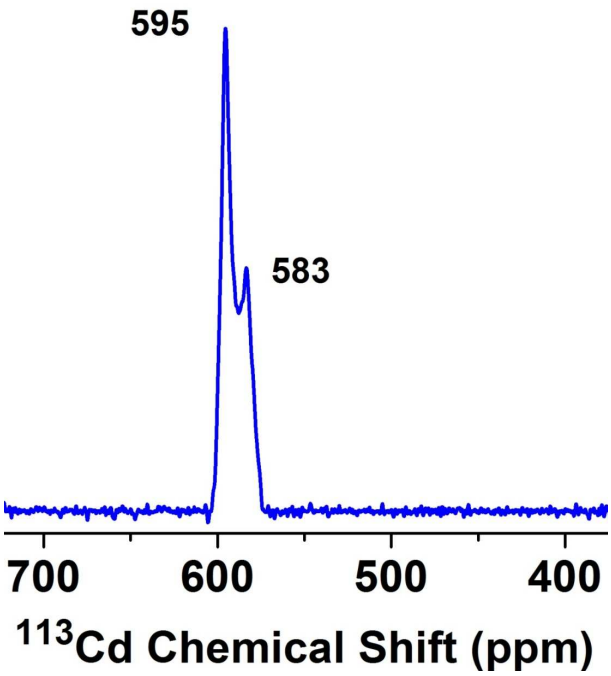


Figure 5-14. ^{113}Cd NMR spectrum of 3 mM α_3D**IV** with 0.8 equivalents of $^{113}Cd(NO_3)_2$ recorded at pH 8.

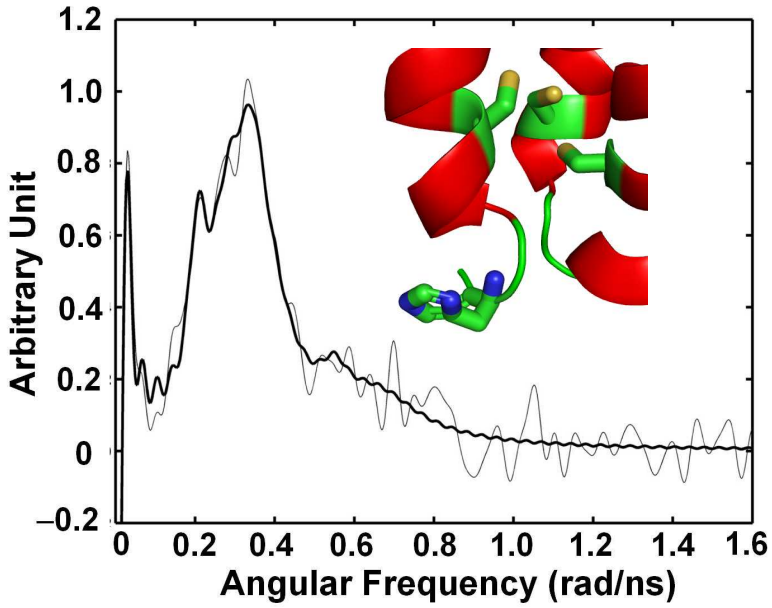


Figure 5-15. ^{111m}Cd PAC spectrum of 300 μM α_3D**IV** loaded with 1/12 equivalents of Cd^{II} (faint line) along with the fits (dark line) to the Fourier transformed data. Sample was prepared in 20 mM TRIS buffer at pH 8.1. PyMol model at the inset shows the His 72 at the end of the bundle which can potentially coordinate to Cd^{II} along with three Cys-thiolates.

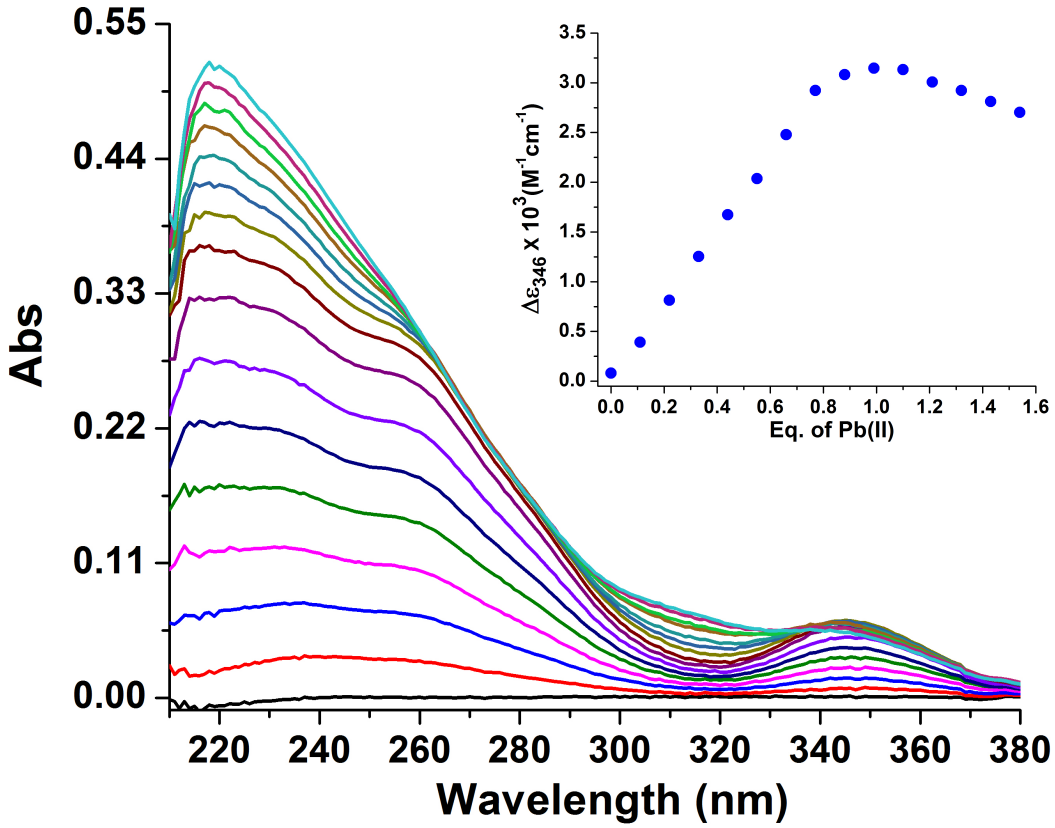


Figure 5-16. UV/Vis titration of $PbCl_2$ to a solution containing 20 μM α_3D**IV** and 40 μM TCEP in 50 mM TRIS buffer at pH 8. Data are plotted as Abs vs. wavelength. The inset of the figure shows the titration curve plotted as $\Delta\varepsilon_{346nm}$ vs. equivalents of Pb(II) added. In the presence of more than stoichiometric amount of Pb(II), the intensity of the peak at λ_{346nm} decreased, while the intensity of the peak at λ_{236nm} increased continuously as well as the UV signal between 300 and 330 nm. Three isosbestic points at λ_{330nm}, λ_{280nm} and λ_{260nm} under excess conditions of Pb(II) are observed, suggesting that other chomophore(s) are being formed at the expense of PbS_3 chromophore.

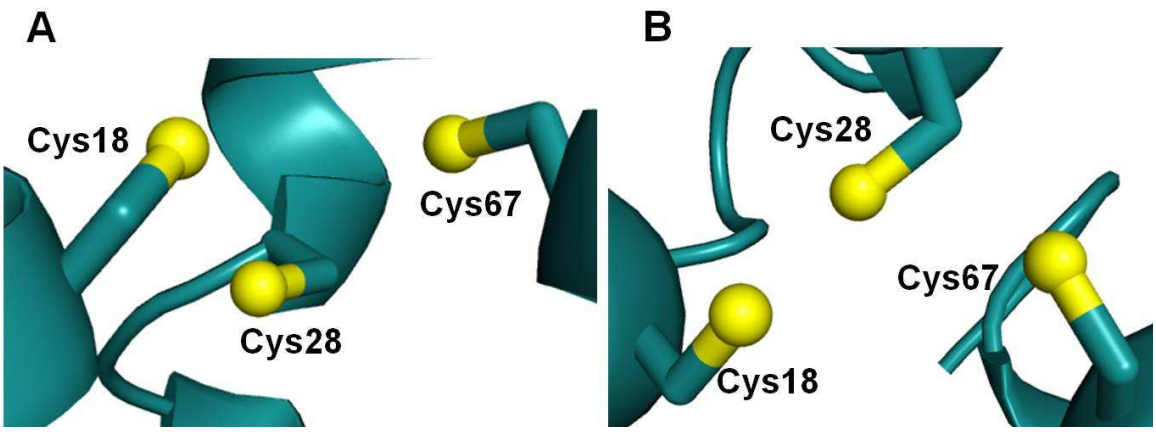


Figure 5-17. PyMol model of α_3D**IV** showing views of Cys residues A) perpendicular to the helical axis, and B) from N-terminal end of the bundle. A shows that the β-methylenes of Cys 18 and Cys 67 Cys in α_3D**IV** are pointing towards the C-terminus while that of Cys 28 is pointing towards the N-terminus of the bundle. B shows that the S^G atoms of Cys 18 and Cys 28 are oriented towards the interior of the bundle whereas for Cys 67 it is directing almost towards the helical interface.

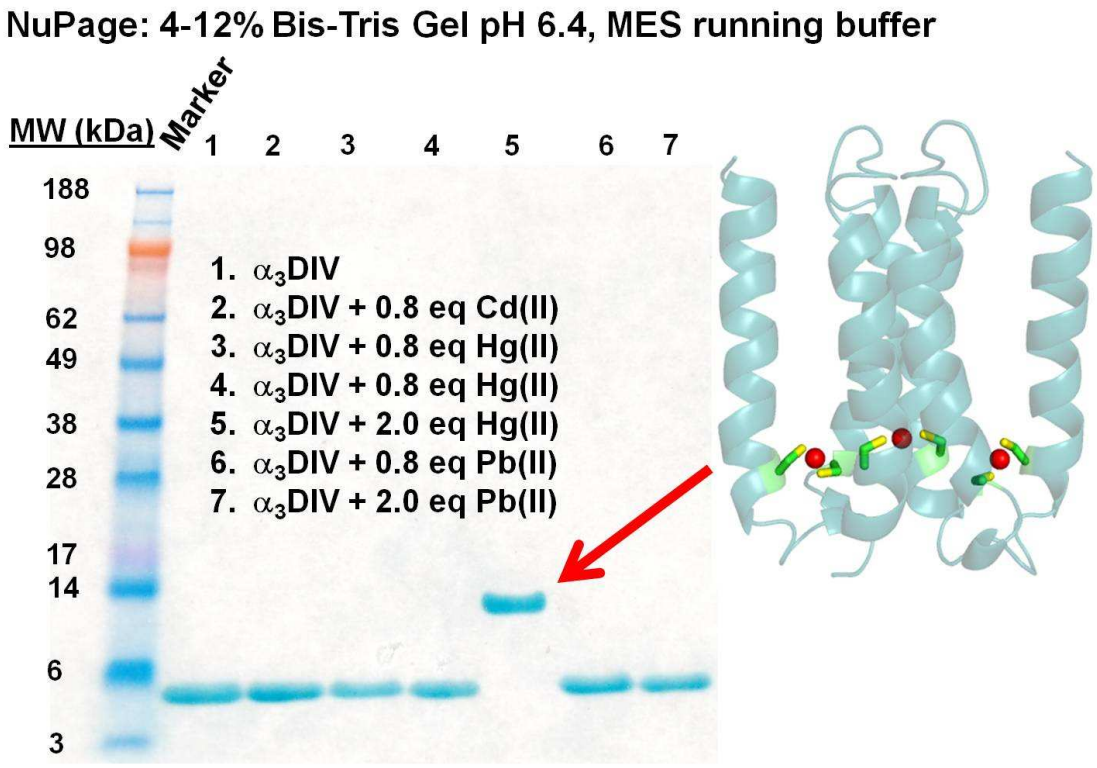


Figure 5-18. A 4-12% NuPage Bis-Tris gel run in 1X MES buffer pH 6.4. Lanes are loaded with samples of apo α_3D**IV** as well as in the presence of different equivalents of metals. Samples 1, 2, 3, 4, 6, and 7 show low molecular weight bands at ~ 5 kDa. Sample 5, loaded with 2.0 equivalents of Hg(II) shows a high MW band ~ 13.5 kDa. All the bands appear at lower molecular weights than expected for these samples. A model of the proposed high molecular weight species, the cross-linked dimer of α_3D**IV**, $Hg_3(\alpha_3D\mathbf{IV})_2$ is shown on the right.

Chapter 6

Conclusions

Metalloproteins account for nearly one half of all proteins in nature.[1] Protein-metal binding sites play crucial roles in important biology processes starting from photosynthesis, respiration, signal transduction to complex chemical transformations. Therefore, designing metalloproteins have been an in exciting area of research over the last two decades to understand structural and functional features of metalloproteins. Due to the complexity of naturally occurring metalloproteins rigorous studies are often precluded. One approach to simplify this problem is the *de novo* design of metalloproteins that are much simpler than the natural counterparts, yet which retain sufficient complexity and information in their sequences that help us understand the structure-function relationships of the naturally occurring proteins. However, *de novo* design of metalloproteins is not just an intellectual exercise. Equipped with the lessons learned from this 'bottom up' approach will help us design novel metalloproteins with improved properties and 'state-of-the art' functionalities. My research has focused on fundamental studies using *de novo* designed α-helical three-stranded coiled coils and single-chain three-helix bundle protein to elucidate the biochemistry of heavy metal ions such as Cd(II), Hg(II), and Pb(II).

In the second chapter of my thesis I have described the structural analysis of two *de novo* designed parallel three-stranded coiled coils **CS**L9C and **CS**L19C in their non-metallated forms, determined to 1.36 and 2.15 Å resolutions, respectively. Previously the group has previously characterized the metallated peptide[2] As(**CS**L9C)$_3$ which served as the first structural model of the metalloregulatory protein ArsR,[3] where As(III) ion was bound to the Cys residues as an *endo* conformation below the Cys thiol plane within the protein environment. From previous studies the group has demonstrated that the affinities and selectivities of metal binding to the **a** vs. **d** site substituted peptides are different in solution.[4, 5] Cd(II) has binds selectively to the **a** site and has a ~10-fold higher preference for the **a** site over **d** site.[6] Pb(II) on the other hand has ~2-4 fold preference for the **d** site.[7, 8] The structures of apo **CS**L9C and **CS**L19C are used to understand the inherent structural differences between **a** and **d** sites to clarify the basis of the observed differential behavior of metal binding in these types of peptides. With **CS**L9C peptide, I have discussed local/global structural differences with the metallated peptide As(**CS**L9C)$_3$, effect of As(III) binding on the overall structure of the peptide, and address whether As(III) binding changes the conformations and the extent of the degree of preorganization of the metal binding cysteine residues. With the apo **CS**L19C structure I have addressed whether there are differences in the orientation of the Cys residues in the **a** vs. **d** site structures, size of the thiol pocket generated by the two types of structures, packing of the adjacent hydrophobic Leu residues. On the basis of the observed structural difference between the two sites, I also hypothesize the possible binding modes in which Cd(II), Hg(II), and Pb(II) would coordinate in the two types of structures which would help us understand the observed differences in physical properties for metal binding to **a** vs. **d** substituted three-stranded coiled coil peptides.

Cys side chains of (**CS**L9C)$_3$ show alternate conformations and are oriented essentially in equal proportion between the coiled coil interior and helical interface. Thus the Cys side chains are partially preorganized (15%) for metal

binding. The β-methylenes and Cys S_γ atoms are pointing towards the N-termini of the helices. Cysteines in (**CS**L19C)$_3$ are present as a single conformer. Thiol groups of two Cys side chains are oriented toward the interior of the helices and toward C-termini, while one Cys is oriented toward the helical interface and almost perpendicular to the helical axis. With S-S separations of 3.32 Å for the interior Cys conformers, cysteines in **CS**L9C generate a smaller thiol pocket than that of **CS**L19C for which the S-S separations are 4.64, 3.40, 4.59 Å. The secondary structure of both the apo and metallated **CS**L9C overlay very well showing that metal binding does not perturb the overall secondary structure of the peptide. Intriguingly, the Leu layers above and below the Cys 9 site are oriented in a similar way in both the apo and metallated (**CS**L9C)$_3$ structures. The Leu 5 layer is tucked in towards the center of the helical axis and the Leu 12 layer is oriented away from the center, thus opening up space below the metal binding site. It is believed that As(III) binds in a hemidirected **endo** conformation to (**CS**L9C)$_3$ so that the stereochemically active lone pair can be housed in the open space generated by the outward orientation of the Leu 12 layer below Cys 9. The Leu residues at the 16 layer of **CS**L19C are positioned away from the center of the helical axis whereas those at layer 23 are directing towards the center of the helical axis. Thus, the Leu 16 layer generates more open space compared to that of the Leu 23 layer. The orientation of these Leu layers above and below the thiol site is exactly opposite compared to those of Leu 5 and Leu 12 of **CS**L9C as mentioned above. From the structural analysis (orientation of Cys S_γ atoms, size of the thiol pockets, packing of the adjacent hydrophobes) of the two types of structures, it is concluded that ions with smaller ionic radii such as Cd(II), Hg(II), and As(III) are better fitted within the **a** site, while ions with larger ionic radii such as Pb(II) can be well accommodated within the larger size of **d** site.

In the third chapter of my thesis, I have investigated the fine tuning of the physical properties of two Cd(II) ions bound in “identical” first coordination spheres at different locations along the linear sequence of three-stranded coiled

coils. The group has previously designed heterochromic peptides that contain two binding sites in close proximity (~ 20 Å) within the same three-stranded coiled coil, and are capable of binding two Cd(II) ions with different coordination geometries (three coordinate, trigonal planar as CdS_3 and four coordinate, pseudotetrahedral as CdS_3O) and thus display different physical properties of the bound ions.[9, 10] Furthermore, these peptides showed site-selective Cd(II) recognition, where Cd(II) binds selectively to the 4-coordinate site (CdS_3O) regardless of the pH conditions.

While the heterochromic peptides allowed us to discriminate between 3- and 4-coordination sites that were at a first glance nearly identical, this success raised an important new question: Are all 3-coordinate or 4-coordinate sites in the same peptide identical? Said differently, can we refer to, as an example, an environment simply as a 4-coordinate site or we must specify whether this sulfur rich environment is in the center of the helix versus towards the helical terminus or whether the Cys resides are in **a** vs. **d** heptad positions? If there is a difference, will the physical properties of the center be modified, will there be any selectivity for binding of the same metal in two sites or the dynamics of the structures be different? Ultimately, we can address whether the position of these sites in the coiled coil play any role in fine tuning the physical properties of the bound Cd(II). I have addressed these issues by sets of designed single peptides capable of binding two equivalents of Cd(II) with the same coordination geometry. A construct such as this has two 3-coordinate sites (CdS_3) or two 4-coordinate sites (CdS_3O) at different positions in the coiled coil. These peptides are used to assess the selectivity of metal binding between two equivalent positions within the peptide, whether binding of one metal would influence the physical properties of the second bound metal and if different physical properties can be observed for the same metal in two similar first coordination environments but in different topological positions within the coiled coil.

The peptide **GRAND**L16PenL19IL23PenL26I, binds two Cd(II) ions as trigonal planar 3-coordinate CdS_3 structures as observed from ^{113}Cd NMR chemical shifts and ^{111m}Cd PAC spectroscopic data. The results of the studies performed with this peptide show that the physical properties (pK_{a2} for Cd(II) binding, UV-Vis spectra, ^{113}Cd NMR chemical shifts) of the two sites in this peptide are similar; however these two sites are not completely independent of each other. The metallation state (apo or $[Cd(pep)(Hpep)_2]^+$ or $[Cd(pep)_3]^-$) of one site perturbed the Cd(II) bound to the second site. A minimum of two Leu layers were found to be required to keep both metal centers completely independent of each other. The **GRAND**L12AL16CL26AL30C peptide, binds two Cd(II) ions as pseudotetrahedral 4-coordinate CdS_3O structures and displays dramatically different physical properties for the two metal sites. The pK_{a2} values, corresponding to the equilibrium $[Cd(pep)(Hpep)_2]^+ \rightarrow [Cd(pep)_3]^- + 2H^+$, range from 9.9 to 13.9, and the L26AL30C site shows dynamic behavior, which is not observed for the L12AL16C site. It is concluded that this is the combined effect of several factors playing a crucial role in determining the final properties of the bound Cd(II), among them, are the packing of the adjacent Leu residues, size of the intended cavity (*endo* vs. *exo*) and location of the binding centers along the coiled coils. Thus, one can not define a location simply as a “4-coordinate” site without being cognizant of the numerous factors that determine the final physical properties of these types of structures.

Metal ion homeostasis and transport are important cellular processes for maintenance of different biochemical activities. Essential metal ions are safely delivered to their partners by assistance with metallochaperones.[11-15] Toxic metal ions such as Hg(II), Cd(II), As(III), and Pb(II) that a cell encounters must be detoxified before they are released back to cellular environment. One important factor that defines the exchange processes between metal ions and biological macromolecules is the dynamics involving metal exchange. The rates of metal incorporation and removal are essential parts for proper functioning of these proteins. However, little is known about the fundamental rates of metal ion

exchange and a complete understanding of the mechanistic pathways in which metals insert in and out of the proteins has been elusive and is of fundamental interest to the bioinorganic community. In the fourth chapter of my thesis I have attempted to understand the exchange of Cd(II) using designed GRAND peptides as models for more complicated native proteins such as metallochaperones, and metalloregulatory proteins such as CmtR which belongs to the ArsR/SmtB[3] metal sensing family and senses Cd(II) and Pb(II). The group has previously investigated the insertion of Hg(II), and Cd(II) into well-structures TRI peptides that exist as parallel three-stranded coiled coils even in the absence of metals. Studies with insertion of Hg(II) **TRI**L9C was shown to initially follow the "Dissociation Mechanism" followed by the "StepAD" mechanism.[16] In case of Cd(II) the insertion mechanism followed the "Breathing" mechanism.[16] Although these studies were significant, the following questions need to be answered to obtain a better molecular level understanding of the metal insertion process into helical scaffolds: Do the metals first interact with amino acids located at the surface of the peptide before being inserted into the coiled coil interior? Do the cysteines at the designed metal binding site play any role in assisting the insertion of the metal from the surface to the hydrophobic interior once the metal is bound to the surface residues? What are the rates of metal exchange? Do the rates depend on the location of the metal binding site at the interior of the coiled coils (middle of the helix vs. towards the helical terminus)? Is there any impact of the metal's coordination number on the exchange dynamics (4-coordinate CdS_3O, vs. 3-coordinate CdS_3)? To address these questions, in the fourth chapter I have investigated the exchange of Cd(II) using the GRAND series of peptides containing both dual and single metal binding sites. ^{113}Cd NMR spectroscopy has been used to probe the chemical exchange of Cd(II).

The results show that chemical exchange of Cd(II) is observed when the metal ion is bound as 4-corodinate CdS_3O structure in pseudotetrahedral geometry, whereas the corresponding 3-coordinate trigonal planar CdS_3 structures are inert to chemical exchange on the NMR time scale. The exchange

of Cd(II) is observed to be occurring on the slow exchange regime with the exchange rates determined to be in the range of 5.8 – 10.9 ms. Exchange studies performed with the peptide **Grand**L12AL16CL26AL30C that encapsulates two Cd(II) ions as 4-corodinate CdS_3O geometry at both the binding sites, show that the L26AL30C site, located proximal to the helical terminus is more dynamic and susceptible to chemical exchange of Cd(II), compared to the L12A16C site located at the middle of the helical scaffold. This demonstrates that the dynamics of the same metal bound in "identical" coordination environments (CdS_3O) can be different depending on the location of the metal sites along the linear sequence. Studies with the corresponding single site peptide **Grand**L26AL30C show that the exchange rate of Cd(II) in this peptide is slower than that of the L26AL30C site in the dual site peptide **Grand**L12AL16CL26AL30C. This observation suggests that introduction of a metal binding site (L12AL16C) in the middle of the helix has fine tuned the dynamics of a second metal site located ~20 Å away, close to the helical terminus. A multi site exchange scheme is proposed where the free Cd(II) is believed to be interacting with the surface Glu residues prior to being inserted to the interior of the coiled coil. The initial interaction of the metal ion with the Glu residues and the alternate conformations of Cys side chains are thought to be facilitating the exchange process.

In the fifth chapter of my thesis I have discussed successful engineering of a tris(cysteine) metal binding site at the C-terminal end of an existing *de novo* designed antiparallel single-chain three-helix bundle α_3D.[17] The NMR structure of α_3D was solved by DeGrado *et. al.* Properties of α_3D are similar to those of native proteins and thus this protein should be tolerant to mutations and serve as an excellent framework to evolve functions. The redesigned protein is designated as α_3D**IV** [α_3D (L18CL28CL67C)]. Previous NMR structural and dynamic investigations showed a gradient in the dynamic behavior and malleability of the protein, with the C-terminal end of the bundle being most amenable to amino acid substitutions.[18-20]

α_3D**IV** was successfully expressed in ~17 mg/ L yield. Circular dichroism and NOESY spectroscopy showed that the protein was very well folded with an α-helicity of ~97%. Furthermore, α_3D**IV** remained well folded between pH 3 and 9. The stability of α_3D**IV** was determined to be 2.5 kcal mole^{-1} from GuHCl denaturation titrations. Even though the substitution of three Leu residues to Cys residues resulted in a loss of free energy of ~2.5 kcal mole^{-1} of the mutant protein, it was stable enough for further characterization in the presence of different metal ions such as Hg(II), Cd(II), and Pb(II). The pK_a for Hg(II) binding leading to the formation of the HgS_3 complex from linear $HgS_2(SH)$ complex of α_3D**IV** is 7.1 $\pm$ 0.1. The pK_{a2} for Cd(II) and Pb(II) binding to α_3D**IV** corresponding to the equilibrium $[M^{II}S(SH)_2]^+ \rightarrow [M^{II}S_3]^- + 2H^+$ were determined to be 10.6 $\pm$ 0.1 and 10.2 $\pm$ 0.1. These pK_a values are slightly more acidic than the corresponding TRI peptides studied by the group.[21-23] This designed protein displays high affinity (>10^7 M^{-1}) for binding metals such as Hg^{II}, Cd^{II} and Pb^{II} in predefined coordination geometries as observed by a combination of spectroscopic techniques such as UV-Vis, ^{113}Cd NMR, ^{199}Hg NMR, ^{111m}Cd PAC, and ^{199m}Hg PAC. The Hg^{II} ion is complexed as a linear HgS_2 center at low pH and as a trigonal planar HgS_3 complex at pH 8.6. The corresponding $Pb^{II}\alpha_3$D**IV** forms a trigonal pyramidal PbS_3 chromophore using the same three cysteine thiolate ligands. In contrast, Cd^{II} forms four coordinate species either as CdS_3O (where O is a coordinated water) or CdS_3N (where N is a bound histidine). Despite the fact that numerous metalloprotein systems have been designed over the last fifteen years, α_3D**IV** represents the first example where metal binding sites have been engineered within an anti-parallel three-helix bundle scaffold.

Successful preparation of the three-Cys mutant of α_3D now opens the possibilities of engineering asymmetric metal binding sites into this scaffold. There are many different proteins that contain asymmetric metal binding environments. Type I copper centers (T1Cu), often called ‘cupredoxins’, are characterized by a single copper atom coordinated by two histidine residues and

a cysteine residue in a trigonal planar structure, and a variable axial ligand.[24-28] In class I T1Cu proteins (e.g. amicyanin, plastocyanin and pseudoazurin) the axial ligand is a methionine, whereas amino acids other than methionine (e.g. glutamine) give rise to class II T1Cu copper proteins. Azurins contain the third type of T1Cu centres: besides a methionine in one axial position, they contain a second axial ligand (a carbonyl group of a glycine residue). Cupredoxins show similar three-dimensional structures, relatively high reduction potentials (> 250 mV), and strong absorption near 600 nm (due to S→Cu charge transfer), which usually gives rise to a blue color.

As a starting point towards our future endeavors in designing asymmetric metal coordination sites, our short term goal could be to engineer Type I copper center into α_3D without the axial ligand. Once successful in forming the trigonal planar site with two histidines and one cysteine and fully characterizing this structure with various spectroscopic techniques, we would focus in introducing the axial ligand to the copper ion. It has been shown that the axial ligand is not necessary to structurally form the Type I center; however the axial ligands have been shown to be critical to achieve the high reduction potentials.[24, 28] Thus it would be important for us to reengineer α_3D so as to introduce the fourth axial ligand to the copper ion in order to prepare a Type I copper center in a *de novo* designed three-helix bundle. Attempts have been made to prepare the His_2Cys asymmetric site on *de novo* designed coiled coil systems.[29] However, metal incorporation was not achieved by this system. α_3D, being a single polypeptide chain should serve as a better template to engineer asymmetric metal binding sites into this protein. Furthermore, we should be able to introduce a single specific secondary shell interaction, such as hydrogen bonding, and making more efficient catalytic metalloproteins using this construct. Additionally, with this system we can study a mixture of **a** and **d** sites substituted peptides, which is often precluded using three-stranded coiled coil systems. Thus, in future we envision evolving a tremendous amount of exciting new functionalities and structure types using α_3D as a template.

References

[1] Y. Lu, N. Yeung, N. Sieracki, N. M. Marshall, *Nature* **2009**, *460*, 855.
[2] D. S. Touw, C. E. Nordman, J. A. Stuckey, V. L. Pecoraro *Proc. Natl. Acad. Sci., U.S.A.* **2007**, *104*, 11969.
[3] W. Shi, J. Dong, R. A. Scott, M. Y. Ksenzenko, B. P. Rosen, *J. Biol. Chem.* **1996**, *271*, 9291.
[4] M. Matzapetakis, V. L. Pecoraro, *J. Am. Chem. Soc.* **2005**, *127*, 18229.
[5] M. Matzapetakis, University of Michigan (Ann Arbor), **2005**.
[6] M. Matzapetakis, B. T. Farrer, T.-C. Weng, L. Hemmingsen, J. E. Penner-Hahn, V. L. Pecoraro, *J. Am. Chem. Soc.* **2002**, *124*, 8042.
[7] D. S. Touw, *Ph.D Thesis* **2007**, *University of Michigan*.
[8] K. P. Neupane, V. L. Pecoraro, *Angewandte Chemie International Edition* **2010**, *49*, 8177.
[9] O. Iranzo, C. Cabello, V. L. Pecoraro, *Angew. Chem., Int. Ed.* **2007**, *46*, 6688.
[10] A. F. A. Peacock, L. Hemmingsen, V. L. Pecoraro, *Proceedings of the National Academy of Sciences* **2008**, *105*, 16566.
[11] A. C. Rosenzweig, *Accounts of Chemical Research* **2000**, *34*, 119.
[12] A. C. Rosenzweig, T. V. O'Halloran, *Curr. Opin. Chem. Biol.* **2000**, *4*, 140.
[13] A. K. Wernimont, D. L. Huffman, A. L. Lamb, T. V. O'Halloran, A. C. Rosenzweig, *Nat. Struct. Biol.* **2000**, *7*, 766.
[14] T. V. O'Halloran, V. C. Culotta, *J. Biol. Chem.* **2000**, *275*, 25057.
[15] W. Xia, H. Li, K.-H. Sze, H. Sun, *Journal of the American Chemical Society* **2009**, *131*, 10031.
[16] D. Ghosh, V. L. Pecoraro, *Inorg. Chem.* **2004**, *43*, 7902.
[17] S. T. R. Walsh, H. Cheng, J. W. Bryson, H. Roder, W. F. DeGrado, *Proc. Natl. Acad. Sci., U.S.A.* **1999**, *96*, 5486.
[18] S. T. Walsh, R. P. Cheng, W. W. Wright, D. O. Alonso, V. Daggett, J. M. Vanderkooi, W. F. DeGrado, *Protein Science* **2003**, *12*, 520.
[19] S. T. R. Walsh, A. L. Lee, W. F. DeGrado, A. J. Wand, *Biochemistry* **2001**, *40*, 9560.
[20] S. T. R. Walsh, V. I. Sukharev, S. F. Betz, N. L. Vekshin, W. F. DeGrado, *Journal of Molecular Biology* **2001**, *305*, 361.
[21] M. Matzapetakis, D. Ghosh, T.-C. Weng, J. E. Penner-Hahn, V. L. Pecoraro, *J. Biol. Inorg. Chem.* **2006**, *11*, 876.
[22] O. Iranzo, Lee, K.H., Jakusch, T., Hemmingsen, L., Pecoraro, V. L., *Chem. Eur. J.,* **2009**, *15*, 3761.
[23] O. Iranzo, D. Ghosh, V. L. Pecoraro, *Inorg. Chem.* **2006**, *45*, 9959.
[24] J. P. Klinman, *Chemical Reviews* **1996**, *96*, 2541.
[25] E. I. Solomon, R. K. Szilagyi, S. DeBeer George, L. Basumallick, *Chemical Reviews* **2004**, *104*, 419.
[26] H. B. Gray, B. G. MalmstrÃ¶m, R. J. P. Williams, in *Journal of Biological Inorganic Chemistry, Vol. 5*, Springer Berlin / Heidelberg, **2000**, pp. 551.
[27] C. Dennison, *Coordination Chemistry Reviews* **2005**, *249*, 3025.

[28] R. H. Holm, P. Kennepohl, E. I. Solomon, *Chemical Reviews* **1996**, *96*, 2239.
[29] A. Lombardi, J. W. Bryson, W. F. DeGrado, *Biopolymers* **1997**, *40*, 495.