CRYSTALLOGRAPHIC COMPARISON OF TRIS-THIOLATE SITES IN DESIGNED PROTEINS TO CONTROL METAL GEOMETRIES

by

Leela Ruckthong

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biophysics)
in the University of Michigan
2016

Doctoral Committee:

Professor Vincent L. Pecoraro, Chair
Professor Ari Gafni
Professor James Penner-Hahn
Associate Professor Jeanne A. Stuckey
Assistant Professor Nathaniel Szymczak

To my parents and sister

Acknowledgements

There are a number of people, without whom and their encouragements, my doctoral degree might not have been accomplished. I would like to take this opportunity to express my sincere gratitude to them.

The first words of my deep appreciation goes to my dissertation advisor, Prof. Dr. Vincent L. Pecoraro. I can never thank Vince enough for his trust in a foreign student like me, who had a very hard time speaking English, to join his lab. He has been giving me such a huge encouragement to learn Science in this language and accomplish my degree. Often times through our sciencific discussion, I felt bad that I could not find touch-to-the-point words to explain what I would like to say, but Vince was patient enough to wait and helped finished my sentences as well as teached me English. His dedication and guidance has been mainly responsible for my success. I also would like to sincerely thank him for his understanding on my health situations and a full year that I had to take a break from school to recover. Without his kindness, I could not make it this far. He said one time that he was a "pseudo-father" for all of us in lab (and we were his pseudo-children) and I agree. Vince is like my father in the US.

My second heartfelt gratitude is definitely for Assoc. Prof. Dr. Jeanne Stuckey. When I had a chance to learn a brief introduction on X-ray Protein Crystallography during my first year and heard how painful it was to get a crystal structure, I swear to myself I would never do this technique for my reseacrh. But guess what happened later?, unexpectedly, Vince asked me to try it. Of course, as predicted, all the first sets of trays completely failed. No crystals. No luck. But Jeanne had encouraged me to keep on trying. Until we solved the first GRAND-CoilSer structure, I cried! I was very happy and finally changed my mind toward this challenging technique. Jeanne is a great professor and knowledgable. I really appreciate her time and energy for teaching me on crystallography and letting me ask her whenever I had questions.

I would like to thank all my other committee members (Prof. Ari Gafni, Prof. James Penner-Hahn and Assis. Prof. Natheniel Symzchazk) for their scientific discussion and support. Moreover, I would like to thank the Hemmingsen lab, Penner-Hahn lab and Stuckey lab for their collaborations. My work would not have been completed without the instruments from Department of Chemistry, Biophysics Program and the Center of Structural Biology. Most importantly, I owe my deep sense of gratitude to the Royal Thai Government for the full Ph.D. scholarship that mainly initiated my study here in the US.

I am extremely grateful to work with Dr. Jennifer Meagher upon crystallization setups and data collections. Jen was very helpful and always had a good eye on my crystals. She was very happy every time she saw my crystal diffractions. Other than work, Jen was the person whom I could sit and talk to after having a hard time during the Ph.D. pursuit. Another deepest appreciation goes for Dr. Nicolas Villanueva who has been very nice to spend time helping me on data processing and crystallographic discussion. He was very generous to let me use his experiment time to collect my data, helped solve one of my structures and many more on refinements. Thanks for the help preparing my data meeting presentation. I thank Priya for her kindness while I was working in the Stuckey lab and Dr. Aniruddha Deb for the EXAFS fitting.

To Sara Grosky, thank you for listening and offering me advice through my entire process. Thanks to Rebecca for those smiles to cheer me up. Thanks to Karthik from Saper lab for the 521 Biophysics discussion. Biophysics folks; Andrew, Teppei, Shirley, Josh, Charity, Rohit and Veronica for their friendship. Moreover, thanks the Pecoraro group people who have shared their laugh, fun time and scientific discussion in the lab with me, both old lab members (Ted, Joe, Joe G., Matteo, Evan, Kosh, Alison, Fangting, Ginny, Jeff, Sarah, Corrado, Casey, Feng, Christina and Kitty) and current lab members (Cathy, Karl, Thomas, Tyler, Tu, Jake and Wong). Special thanks to Saumen, my rotation mentor, who taught me to start working on this peptide project and Melissa for her assistance when I first learned crystallography. To Eric, Angie and Thanya from Ragsdale lab, thanks for the valuable advice on my talks.

As my health problems had been unpredictable over the past years, I would like to express my sincere gratitude to all the medical teams in the University of Michigan Hospital and Siriraj Hospital (Thailand), all my MD friends in Michigan and other Thai friends who had been there with me through those difficult times. Thanks Wendy, Fangting, Sunny, Melissa, Gary, Cathy, Alison, Ginny and Mayra for visiting me after I was discharged from the hospital.

Thanks to Ginny, Jeff, Mike, John, Mayra and Taryn for those fun times cooking Thai food together. General help from all friends were all greatly appreciated. Thanks to Thanya, Nate, Jay and Pim for preparing food for me while I was writing my thesis. Thanks Tum, Thian, Pattanapon, Thanapoom and Chayandorn for computer issues. Rattima, Suebpong, Sarawalee, Yuwadee, Anwida, Oat, Panlada, Patima, Chanikarn, Chanipa, Peeraporn, Thitiporn, Radaphat, Nakorn, Supaluk, Janjira, Garn, Art, Nithiwadee, Chanisa, Chanat, Krit, Suwi, Bai, Dee, Aun and so many other Thai folks for the friendship. Bunches of thanks to my best friends (Weekit, Cho, Kanny, Dao, Wattanachai and Tukky) for their 24-7 support.

Lastly, all my success is attributed to my mom, my dad and my wonderful sister. Thanks for having faith in me and not giving up in my dream. My defense was held on the same time as the Father's day in Thailand and that my dad could not sleep until the closed-meeting with my committee was finished. It was so delighted to tell him that I finally passed the last exam. The Ph.D. degree is the best Father's day present I have ever given to him so far.

Table of Contents

page

List of Tables

List of Figures

Abstract

De Novo three-stranded coiled coil proteins formed using **TRI** and **CoilSer** peptides are used to generate a Cys-rich site that arranges thiols to chelate metals into geometries resembling heavy metal sites in metalloregulatory systems. A series of high resolution X-ray diffraction structures was obtained to assess the Cys behavior upon metal binding. In the apo-structures, the three Cys ligands, preferentially orient toward the helical core, revealing a *preorganized* pocket to bind a trigonal pyramidal $Pb(II)S_3^-$ and $As(III)S_3$. This arrangement is predisposed for trigonal planar $Hg(II)S_3^-$ and 4-coordinate $Zn(II)S_3O^-$ structures, requiring significant thiol rotation to bind metals. The structures confirm the control of coordination number is achieved on Cd(II) and Hg(II) complexes through ligand modifications. The *predisposition* of Cys upon metal binding results in a mixture of 3- and 4-coordinate Cd(II), while the bulky Pen restricts thiol rotation, causing a shift in the metal plane that displaces water, generating a pure $Cd(II)S_3^-$. Substitution of Ala for Leu opens a cavity above the Cys site which allows for four waters to enter, forming the $Cd(II)S_3O^-$ center. Additionally, inclusion of D-amino acid ligands compels metals into specific geometry. D-Leu, above the Cys site, reorients the side chain towards the coordination layer diminishing the space for water access yielding $Cd(II)S_3^-$; however, D-Leu below the metal site opens more space, allowing for equal $Cd(II)S_3O^-$ and $Cd(II)S_3O_2^-$ structures. The substitution of D-Leu residues simultaneously above and below the Cys site results in a predominant $Cd(II)S_3O^-$. The chirality perturbation by D-Cys causes a shift of the sulfur plane down, thus making a larger cavity suitable for $Cd(II)S_3O^-$. Moreover, D-Pen binds Hg(II) in a linear fashion, while L-Pen can enforce an unusual trigonal Hg(II) center. These studies provide insights on how to control desired metal geometries in metalloproteins.

Chapter I
Introduction

The purpose of this research is to investigate heavy metal recognition and interactions in metalloregulatory proteins of bacterial systems using a simplified *de novo* 3SCC peptide. The first metal of interest is Hg(II). It is a well-known toxin that interferes with biological processes through the formation of a thermodynamically stable Hg(II)-thiolate bond. Because Hg(II) is a soft Lewis Acid and easily polarizable, it binds well to the soft Lewis Basic sulfhydryl ligands of cysteine residues. It is not surprising that biological systems use the same basic chemistry to detoxify this element.

Detoxification of Hg(II) in humans is achieved using metallomethioneins,[1–5] plants utilize phytochelatins (and metallothioneins)[6,7] whereas in prokaryotic cells an entirely distinct process has been developed. Bacterial mercury resistance is encoded by the detoxification proteins called Mer operon which is controlled by the metalloregulatory proteins MerR.[8–16] MerR is a homodimer bound to DNA in the presence or absence of Hg(II). In the absence of this metal, the MerR/DNA adduct adopts a conformation that represses gene transcription. Metal binding induces a conformational change of the Hg(II)/MerR/DNA ternary complex leading to the 33 degree unwinding of the DNA and subsequent transcription of the Mer genes (MerA, MerB, MerP, and MerT) (**Figure 1-1**).[17] These Mer genes subsequently encode proteins that can sense, transport and reduces a toxic Hg(II) to a nontoxic Hg(0) species.[18] MerR has a strong selectivity for Hg(II) in the presence of nanomolar concentrations, while other ions such as Cd(II), Zn(II) and Au(I) require significantly higher metal concentrations to stimulate this metal sensor protein.[12] Intriguingly, site directed mutagenesis studies indicated that MerR bound Hg(II) as a tridentate $[Hg(II)(SR)_3]^-$ complexes.[11,19] This was surprising as the vast majority of Hg(II) complexes in biochemistry, and also inorganic small molecule studies, preferentially form 2-coordinate, linear species.[20–22]

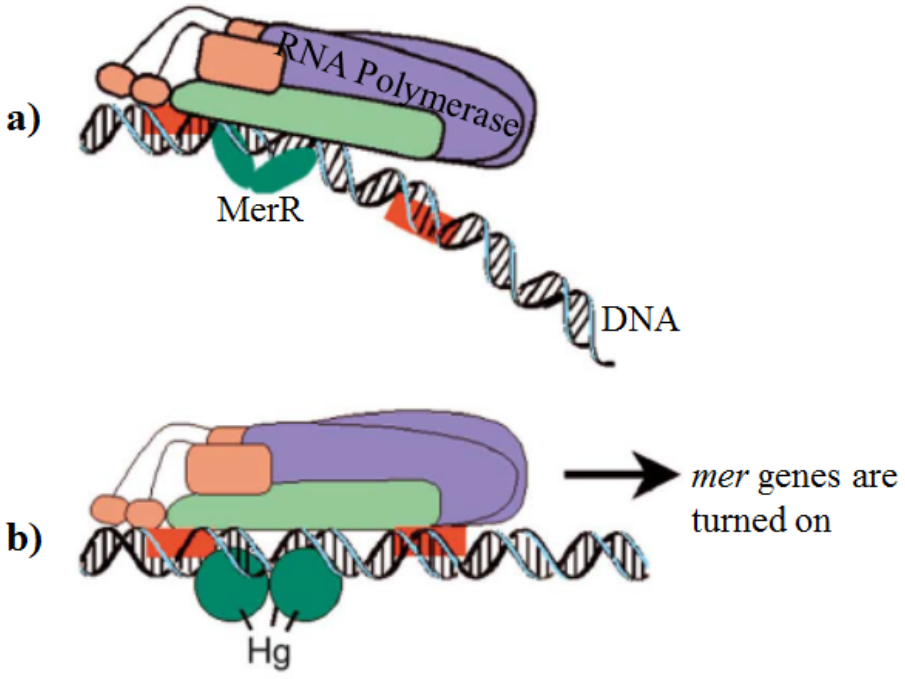

Figure 1-1. **The model for MerR action**. Representing in a) MerR bound to DNA and RNA polymerase recruited but not able to form an open complex. b) The conformational change induced by Hg(II) binding to the ternary complex and initiation of transcription of *mer* genes. The figure is modified from Reference 9.

This unusual three-coordinate Hg(II)-MerR complex, though rare, is thermodynamically stable and it is believed that MerR exploits this geometry to differentiate Hg(II) selectively from other metals, especially other d^{10} metal ions that prefer tetrahedral thiolato coordination. Unfortunately, there is a void in understanding how these metalloregulartory proteins selectively and specifically bind to Hg(II) in this binding mode and the chemistry of this metal site in aqueous solution is not well understood.

While studies of native proteins ultimately yield invaluable information for function they are often extremely complex, causing complications in the interpretation of results. Moreover, often these natural proteins are insoluble or unstable under laboratory conditions. On the other hand, though small molecule mimics are useful as they may be characterized at high structural resolution, it is sometimes difficult to incorporate the exact metal binding functional group into the model. As previously reported in the literature, non-natural bulky thiolate-containing ligands (e.g. aromatic thiolates or aliphatic thiolates) were used in lieu of the Cys ligand in order to enforce this low-coordinate number around the Hg(II) center. The chemistry of tris-thiolate Hg(II) is hard to achieve as it was observed that millimolar concentration ranges of ligands, in non-aqueous solvents, are required to drive the reactions in order to force Hg(II) into this specific geometry. In contrast, O'Halloran *et al*. observed using gel-shift assays that the metalloregulatory protein, merR, possesses nanomolar mercuric ion sensitivity.[10,23] His work and that of Chris Walsh suggested that even at these extremely low concentrations Hg(II) formed a trigonal planar structure.[10,24,25] Moreover, an examination based on August 2015 of the Cambridge Structural Database[26] shows that of over five thousands Hg(II) complexes that have been deposited into the system, only three are reported to be mononuclear $Hg(II)S_3^-$. These compounds are tetralkylammonium salts of $[Hg(SPh)_3]^{1-}$, $[Hg(SBu^t)_3]^{1-}$ and $[Hg(i\text{-}S\text{-}2,4,6\text{-}Pri_3C_6H_2)_3]^{1-}$ **(Figure 1-2)**.[17,23,27] Importantly, none of these were prepared using biologically relevant Cys residues. Usually these complexes are stable in the solid state or non-aqueous solvents, dichloromethane or dimethyformamide. As a consequence, solution spectra are affected by the altered non-protic nature and dielectric of the medium. Therefore, there is a need to look at this binding site in protein environments in order to understand clearly the behavior of this metal center which is believed to be responsible for the regulation of Hg(II) detoxification in bacterial systems.

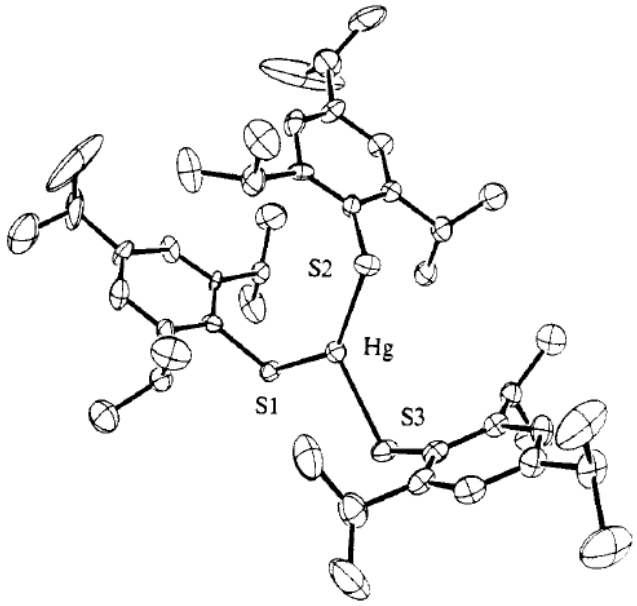

Figure 1-2. ORTEP diagram of $[Hg(S\text{-}2,4,6\text{-}i\text{-}Pr_3C_6H_2)_3]^-$ complex.[27]

Structurally Characterized Tris-thiolate Metal Binding Sites using de Novo Three-stranded Coiled Coil Peptides to Understand Metal Recognition in Metalloregulatory Protein Systems

To avoid the high level of complexity that one encounters in native proteins and to avoid some of the limitations of the small molecule traditional approach, the Pecoraro group has used a *de Novo* metalloprotein strategy, or metalloprotein design from scratch, to study metal-protein interactions. It is an excellent approach that allows one to understand fundamentally the structural and functional properties of metal-protein relationships which may still remain questionable in native systems or might generate novel properties that may not be found previously in natural proteins.[28,29] Though the level of complexity of *de Novo* designed system is much smaller and simpler than those of native counterparts, the minimal model still retains sufficient core features in its primary sequence to study the basic principles of protein folding and metal-protein interactions at a molecular level.[30–32] This strategy can be considered as a bridge between two methods where the native protein studies belongs to one end and the small molecule mimic studies belong to the other.[33] The complicated interpretation caused from the inherent complexity of real protein systems is removed as the alternative strategy of *de Novo* protein design can help isolate factors for studies and simply them to the core elements of the targeted sites. This peptidic framework, moreover, provides a more native-like system, for structural and functional studies, than the small inorganic molecule mimics. Although the inner working system of the small molecule mimics can provide insight into the mechanistic aspect of the studied site, the studies are limited by a variety of factors as previously mentioned. *De Novo* protein design is advantageous because real amino acids can be used to generate the metal binding site and the chemistry can also be performed in water which is relevant to the native systems. Additionally, the secondary and tertiary protein scaffold facilitates the engineering of the secondary coordination sphere of the metal sites. Most significantly, and highly relevant to my research studies, it is easy to incorporate non-natural amino acids into the *de Novo* designed peptides to broaden the knowledge of metalloprotein studies.

The main scaffold that is used to characterize the metal-protein interactions is called the **TRI**-peptide family. The scaffold is an amphipathic α-helix displaying four heptad repeats of seven amino acids, $L_aK_bA_cL_dE_eE_fK_g$ (where subscripts correspond to the position of the substitution in the heptad) in which the N- and C-termini are capped with acetylated and amidated glycine residues, respectively. As shown in **Figure 1-3**, non-polar residues at the *a* and *d* positions make

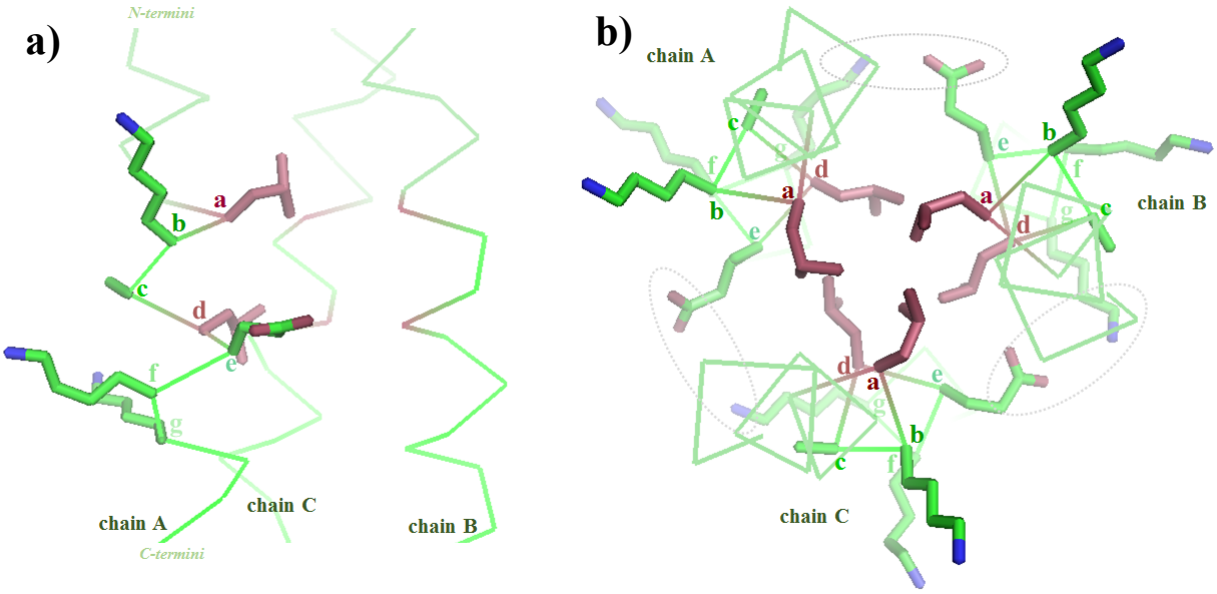

Figure 1-3. Replacement of seven amino acid residues in a heptad of a TRI-family 3SCC peptide. In a) a side view of the arrangements of a heptad in one helix. b) Helical wheel diagram of three helices demonstrating hydrophobic residues occupy the interior face at *a* and *d* positions, while hydrophilic residues are on the exterior face at other positions. The grey circles in b) indicate that the residues can form salt bridge interactions between *e* and *g* positions. Residues at *a* and *d* positions are shown as red sticks, while the remainder are colored green (carbon=green, nitrogen=blue and oxygen=red).

up the hydrophobic core while polar residues at the *e* and *g* positions provide electrostatic interactions assisting the orientation of the tertiary structure. The alanine (Ala) residues at the *c* positions act as a helix inducer. Together, this heptad repeat strategy leads to the parallel assembly of identical helices in aqueous solution to form coiled coils (α-helices that supercoil around a helical axis). The peptide aggregation is pH dependent. At pH lower than 5, these peptides form two stranded coiled coils (2SCCs); however, when the pH is raised Glu residues at the *e* positions are deprotonated and form salt bridge interaction with Lys residues at the *g* positions, which subsequently stabilize the three stranded coiled coil structures (3SCCs).[34,35] A metal binding site can be engineered within the interior of the coiled coil by substitution of a Leu residue at either *a* or *d* position of the **TRI** sequence for an amino acid residue, such as Cys or His, that is capable for metal binding (**Figure 1-4**). Particularly for my dissertation studies where the goals are entirely focused on heavy metal-protein interactions, the Cys ligand is emphasized in the metal ligand designs; however, in a few constructs histidine bound to Zn(II) will also be present. Moreover, the crystallographic analog of **TRI**-family called the **CoilSer** and **GRAND-CoilSer**, in which the sequences, will be introduced are used to perform the structural studies of the designed peptides.

a) Hg(II)-binding Studies in Designed Tris-thiolate Environment using 3SCC Peptides

The *de Novo* **TRI**-family peptide has turned out to be an excellent scaffold because the three sterically hindered thiolate ligands that were used to achieve $[Hg(SR)_3]^-$ complexes in small molecules can be replaced by three α-helical chains containing Cys residues. A trigonal planar Hg(II) site forms because the cysteine residues are placed at the same sequence position of the helix and, since the final aggregate is a parallel 3SCC, this places the thiols in very close proximity in order to bind Hg(II) into the desired geometry (**Figure 1-5**). Thus, we have designed a site that is predisposed to coordinate Hg(II) with three thiolate ligands. The initial work on mercury binding utilized **TRI**L16C (Sequence shown in **Table 1-1**) where the Cys residue replaces the Leu in the sixteenth position of **TRI**.[33,36–41] Three distinct Hg(II) binding behaviors were observed using spectroscopic methods (UV/VIS, EXAFS, ^{199m}Hg PAC and ^{199}Hg NMR) in this system depending on the peptide to Hg(II) ratio and the pH conditions (**Figure 1-6**). It was found that if the peptide solution was titrated into the Hg(II) solution, a linear Hg(II)(**TRI**L16C)$_2$ complex forms when the stoichiometry of peptide to mercury was 2:1, indicating only a two-stranded coiled coil with 2-coordinate Hg(II) was formed. If this third peptide equivalent was added while the pH remains

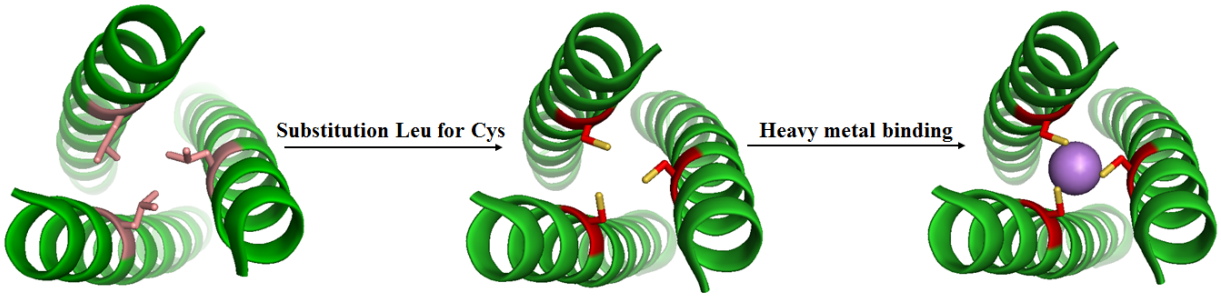

Figure 1-4. PyMOL models demonstrating how to engineer a tris-thiolate binding site into the interior of 3SCC. The substitution of Leu residues (shown as pink sticks) with Cys yields three Cys environment capable of binding heavy metal atoms.

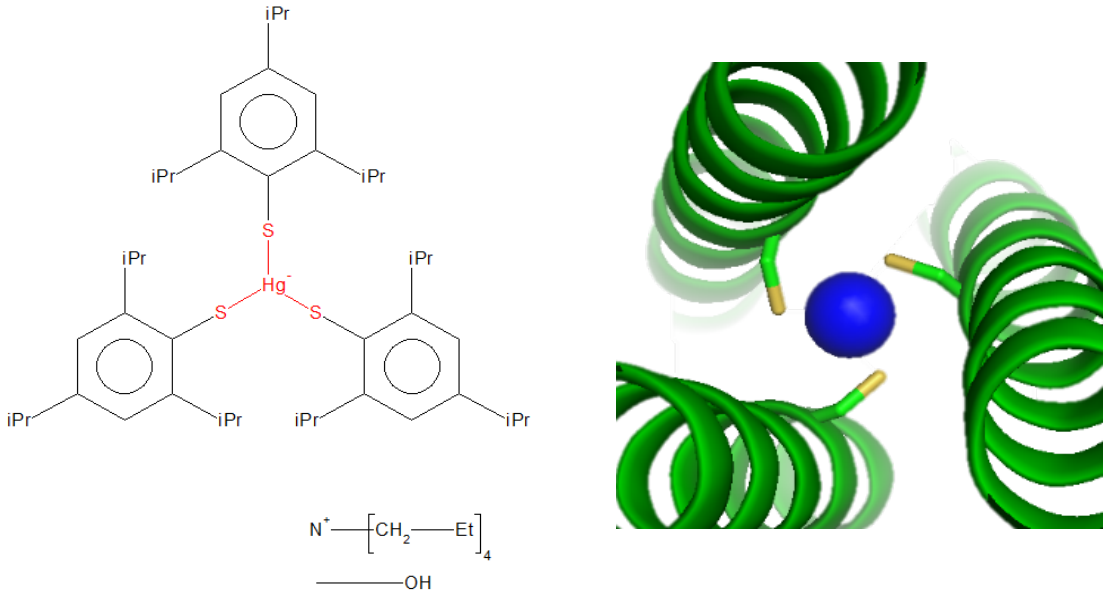

Figure 1-5. Comparison of a tris-thiolate $[Hg(S\text{-}2,4,6\text{-}i\text{-}Pr_3C_6H_2)_3]^-$ small molecule complex[27] with a $Hg(II)S_3^-$ binding site prepared within a 3SCC environment of $Hg(II)_SZn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$ peptide. This figure shows that the bulky thiolate ligands used in the small molecule strategy can be replaced by the peptide backbone in order to force Hg(II) into a trigonal planar geometry.

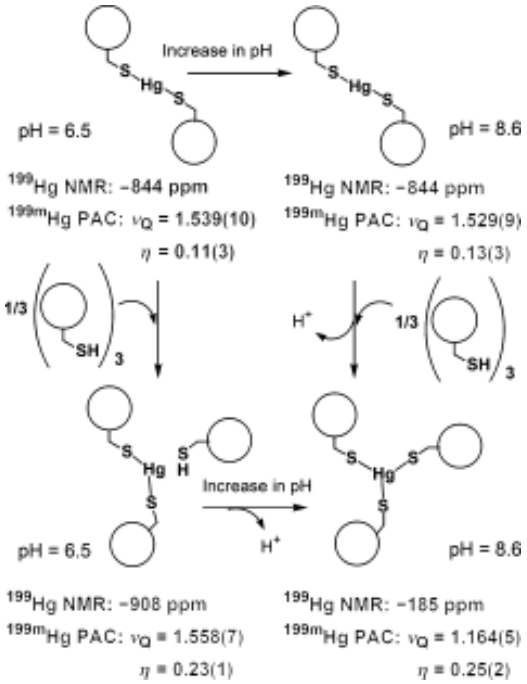

Figure 1-6. Species present at different TRIL9C/Hg(II) ratios and pH values.[36]

Table 1-1. Peptide sequences

Peptides		**a b c d e f g** 2	**a b c d e f g** 9 12	**a b c d e f g** 16 19	**a b c d e f g**	**a b c d e f g** 30
TRI	Ac-G	LKALEEK	LKALEEK	LKALEEK	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	**Pen**KALEEK	LKALEEK	G-NH_2
TRIL12$_D$LL16C	Ac-G	LKALEEK	LKA**$_D$L**EEK	**C**KALEEK	LKALEEK	G-NH_2
TRIL2WL16CL19$_D$L	Ac-G	WKALEEK	LKALEEK	**C**KA**$_D$L**EEK	LKALEEK	G-NH_2
CoilSer (CS)	Ac-E	WEALEKK	LAALESK	LQALEKK	LEALEHG	-NH_2
CSL9C	Ac-E	WEALEKK	**C**AALESK	LQALEKK	LEALEHG	-NH_2
CSL16C	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LEALEHG	-NH_2
CSL16$_D$C	Ac-E	WEALEKK	LAALESK	**$_D$C**QALEKK	LEALEHG	-NH_2
CSL16Pen	Ac-E	WEALEKK	LAALESK	**Pen**QALEKK	LEALEHG	-NH_2
CSL16$_D$Pen	Ac-E	WEALEKK	LAALESK	**$_D$Pen**QALEKK	LEALEHG	-NH_2
GRAND-CoilSer	Ac-E	WEALEKK	LKALESK	LQALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL16CL30H	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LQALEKK	**H**EALEHG -NH_2
GRAND-CSL12AL16C	Ac-E	WEALEKK	LAA**A**ESK	**C**QALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL12$_D$LL16C	Ac-E	WEALEKK	LAA**$_D$L**ESK	**C**QALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL16CL19$_D$L	Ac-E	WEALEKK	LAALESK	**C**QA**$_D$L**EKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL12$_D$LL16CL19$_D$L	Ac-E	WEALEKK	LAA**$_D$L**ESK	**C**QA**$_D$L**EKK	LQALEKK	LEALEHG -NH_2

Bold and underlined residues indicate substitutions.

C-and N-termini are capped by Ac and NH_2 groups, respectively.

around 5.5 to 7, the 2-coordinate bis-thiolato Hg(II) structure encapsulated within the 3SCC was achieved. This observation emphasizes that disrupting the preferred Hg(II) coordination number to make a trigonal planar structure is difficult, even within a pre-folded three-stranded coiled coil that predisposes sulfur binding at this pH range. It is believed that the third thiol remains protonated and uncoordinated to the Hg(II). However, when the pH is raised to 8.5, the third thiol releases its proton, allowing the intrinsic conformational preference of the peptide to enforce the unusual 3-coordinate trigonal planar Hg(II).[36] UV-vis titrations that followed Hg(II) coordination demonstrated that this transition from 2- to 3-coordination occurs with a pK_a =7.7 for *a* site peptides (it is slightly higher when Hg(II) binds to Cys a *d* site).

With a broad range of different spectroscopic techniques having different timescales, the the features between Hg(II)-MerR and Hg(II)(**TRI**L16C)$_3^-$ at pH 8.5 can be compared. The UV/VIS LMCT of Hg(II) to three Cys transitions occurs at 247 nm (ε=19,200 $M^{-1}cm^{-1}$) with shoulders at 265 nm (ε=11,900 $M^{-1}cm^{-1}$) and 295 nm (ε=5,800 $M^{-1}cm^{-1}$)[39] similar to those observed for Hg(II)-MerR as shown in **Table 1-2**.[23] The Hg(II)-coordination environment was probed using a combination of ^{199}Hg Nuclear Magnetic Resonance (NMR) and ^{199m}Hg Perturbed Angular Correlation (PAC) spectroscopies. ^{199}Hg NMR exhibited a chemical shift of -185 ppm[39] for Hg(II)(**TRI**L16C)$_3^-$, which is slightly upfield shifted from the range of mononuclear three-coordinate aliphatic thiolate mercury complexes (-80 to -160 ppm), where Hg(II)-MerR chemical shifts are -106 ppm and -109 ppm (when the protein is bound to DNA).[19,42,43] Moreover, ^{199m}Hg PAC reveals ν_o and η values 1.18 GHz and 0.25.[44] This corresponds to ν_o and η values 1.16 GHz and 0.25 determined from the *a* site **TRI**L9C peptide. EXAFS studies indicated that Hg(II)(**TRI**L16C)$_3^-$ has a Hg(II)-S bond distance of 2.43 Å which is in close agreement with Hg(II)-MerR protein (2.44 Å, EXAFS)[23], the trigonal planar $[Hg(SBu^t)_3]^{1-}$ (2.44 Å), $[Hg(SPh)_3]^{1-}$ (2.44 Å) , $[Hg(i\text{-}S\text{-}2,4,6\text{-}Pri_3C_6H_2)_3]^{1-}$ (2.44 Å) model complexes.[3–5] According to these spectroscopic data, it was clear that by reproduction of the coordination environment of the metal site in the **TRI**L16C construct, the Pecoraro group was the first group to enforce an unusual trigonal thiolate coordination mode of Hg(II) in a simplified protein environment that represented the MerR system in aqueous solution.[39,40,46]

Even though the physical characterizations of **TRI**L16C provided better understanding of protein-Hg(II) interactions, the structural basis of this metal site has not been well-understood because the metallated version of this metalloregulatory protein was only crystallized recently.

Table 1-2: Spectroscopic values for Hg(II)/MerR and Hg(II)/**TRI**-family peptide

sample	λ (Δε) (nm)	Hg-S bond length (Å)	δ (ppm) (^{199}Hg NMR)	ν_Q (GHz), η (^{199m}Hg PAC)
Hg(II)/MerR	242 (19800)[a] 260 (14600)[a] 290 (6450)[a]	2.43[b]	-106[c] -109[c,d]	1.18, 0.25[e]
Trigonal 3-coordinate in *a* site	247 (19200)[f] 265 (11900)[f] 295 (5800)[f]	2.44[f]	-185[g]	1.16, 0.25[g]
Linear 2-coordinate	240 (2700)[h]	2.32[h]	-844[g]	1.53,0.13[g]
Linear 2-coordinate within a 3SCC	247 (2000)[i]	-	-908[g]	1.56,0.23[g]
Tetrahedral 4 coordinate	230 (8100)[j] 289 (7100)	-	-500[j]	

[a] Reference 23 , pH 7.00; [b] Reference 19 , ^{199}Hg NMR chemical shift relative to $Hg(CH_3)_2$; [c] Reference 27 , pH 6.02; [d] Hg/MerR complexed with DNA; [e] Reference 44; [f] Data for **TRI**L16C from reference 39; [g] Data for **TRI**L9C from reference 36; [h] Data for **TRI**L16C from reference 40 [i] Data for **TRI**L12C from reference 47 and [j] Data for **TRI**L12CL16C from reference 50.

In June 2015, Nei-Li Chan and coworkers reported a low resolution (2.5 Å) crystal structure of Hg(II)-bound MerR (**Figure 1-7**).[13] Unfortunately, the resolution of this native structure is not high enough to define the Hg(II)-S bond distance confidently; therefore, one of my dissertation goals in Chapter 2 is to generate a high resolution crystal structure of a trigonal planar Hg(II) in a tris-thiolate environment using the 3SCC peptides.

b) Investigation of Pb(II)-peptide Interactions in 3SCC Peptides

To obtain more insight on the structural requirements for other toxic metals binding to metalloregulatory proteins and metallochaperones, my next goal in Chapter 2 is to obtain and understand structural requirements of Pb(II)-binding to a tris-thiolate environment. Pb(II) is a soft metal toxin that also has high affinity toward thiolate-rich ligands such as cysteines, glutathione and metallothionein.[49–55] It has been found to interfere in several biological processes and causes malfunctions in essential metal ion homeostasis because it can replace and change the metal structures of the cysteine-rich and also oxygen/nitrogen-rich sites of Zn(II) or Ca(II) with alternative coordination numbers upon binding.[56–59] One of the major targets identified for Pb(II) interference is a Zn(II)-containing enzyme known as aminolevulinic acid dehydratase (ALAD), which is responsible for the formation of tetrapyrrole molecules in mammalian heme biosynthetic pathways.[56] Pb(II) inhibits the ALAD active site by replacing the native pseudo-tetrahedral $Zn(II)(Cys)_3(H_2O)$ structure with its preferred $Pb(II)S_3^-$ trigonal pyramidal geometry. This structure has been confirmed by an X-ray crystal structure of Pb(II)-substituted ALAD in yeast (PDB: 1QNV)[60] shown in **Figure 1-8**. It was assumed that the perturbation of the active site structure causes the inhibition of the ALAD enzymatic activity.[61–63]

To survive Pb(II) toxicity, prokaryotic cells utilize the lead resistance operon, PbrR (PbrR2 or PbrR691), which belongs to the MerR metalloregulatory protein class. This protein can be expressed in *Cupriavidus metallidurans* CH34 to regulate Pb(II) level in cells.[9,64,65] It was reported that the PbrR691 can selectively bind Pb(II) almost 1000-fold selectively over other metals.[66] Like MerR, this protein responds to Pb(II) using three Cys residues (Cys78, Cys113 and Cys122) which are also conserved for Hg(II)-binding in the MerR protein as suggested from the sequence alignment between PbrR691 and MerR (Tn 501).[67] However, these highly conserved three Cys residues bind to each of these metals via different coordination geometries. Spectroscopic characterization (above) confirmed that these three Cys sulfur atoms bind Hg(II) as a trigonal

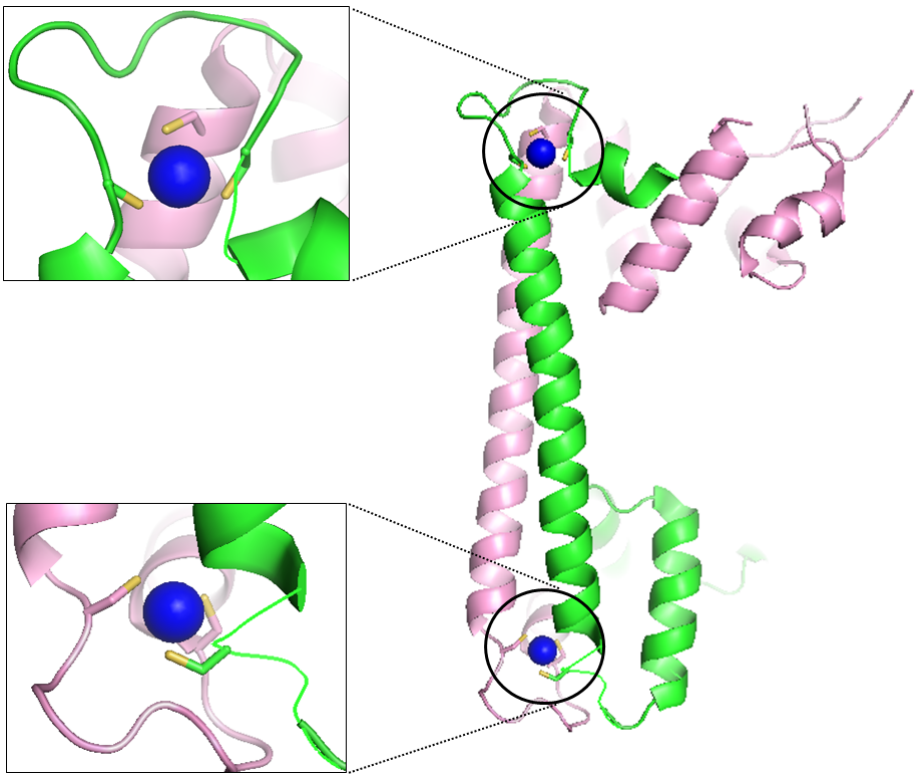

Figure 1-7. PyMOL representation of the Hg(II) bound MerR homodimer crystal structure (PDB code: 4UA1).[13] The Hg(II) binding sites are zoomed-in and shown on the left side panels. The DNA-binding domain, dimerization helix and metal binding site motif of one subunit are colored green and pink for the other. The two Hg(II) ions are shown as blue spheres. Cys ligands are present as sticks where sulfurs are yellow.

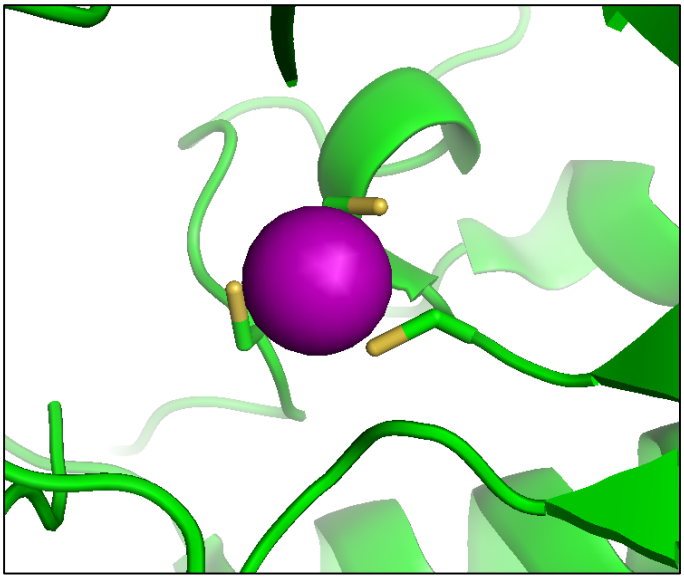

Figure 1-8. PyMOL ribbon diagram of the Pb(II) bound PbrB691 crystal structure (PDB code: 1QNV).[60,61] Pb(II) is bound as a trigonal pyramidal geometry with three Cys ligands. The peptide backbone is green. Cys residues represent as sticks (sulfur=yellow). Pb(II) ion is purple.

planar structure in MerR,[11,19,68] while Pb(II) binds to PbrR in a Pb(II)S_3^- trigonal pyramidal fashion.[46,53] XAS studies showed that the edge spectrum of the low coordination number of Pb(II)-PbrR was thought to be hemidirected where the sulfur ligands are only oriented over one portion of the metal coordination sphere[70] similar to those observed in Pb(II) complexes with thiolate ligands in small molecules.[67,71] The examples of small model tris-thiolate Pb(II)S_3^- complexes are shown in **Figure 1-9**. It has become interesting to understand how the metal recognition is selectively achieved in these metalloregulatory proteins when one realizes that MerR and PbrR use essentially the same highly conserved three Cys residues to bind the metal but are able to differentiate these two ions through their specific geometry. Thus, structural information is needed to identify the metal-ligand organization and distinct properties of these two geometries in order to address this question. Unfortunately, the crystal structure of native Pb(II)-PbrR has not yet been reported. Herein, I attempt to obtain a bound Pb(II) crystal structure prepared in a tris-thiolate binding site using a 3SCC scaffold.

For structural studies, the crystallograprahic analog of the **TRI**-family series of peptides , CoilSer, was used as it is also constructed based on a heptad repeat. **Coil Ser** was first designed based on TM29 and crystallized by Lovejoy and DeGrado *et al*., as an antiparallel triple-stranded coiled coil at low pH.[72] One of the factors proposed to influence the antiparallel orientation was to avoid the steric clash that might occur when the three indole side chains of N-terminal Trp residues are located in close proximity. By substitution of one or two positions containing hydrophobic residues in the **CoilSer** peptides with metal ligands, Cys or His, the Pecoraro group has been able to show that at higher pH the peptides crystallized in a parallel 3SCC manner (**Figure 1-10**) (e.g., As(III)(**CS**L9C)$_3$, apo-(**CS**L9C)$_3$ and apo-(**CS**L19C)$_3$.[73,74] These structures demonstrated that the indole side chains could easily be accommodated within the 3SCC interior and, therefore, was unlikely to be the reason that the original CoilSer structure adopted the antiparallel orientation. In the case of As(III)(**CS**L9C)$_3$ structure, Touw *et al*. reasoned that the pH difference of crystallization between this peptide and the native **CoilSer** could potentially play a significant factor to define the coiled coil orientation because the pH can change the electrostatic interactions of the aggregates.[74] Coil Ser was previously crystallized at pH 5.0, where the hydrogen bonding capabilities of glutamate residues are partially minimized by protonation reducing the amount of interhelical salt bridge interactions (four interactions in total), whereas As(**CS**L9C)$_3$ was crystallized at pH 8. This pH allows for complete deprotonation of Glu residues which causes the

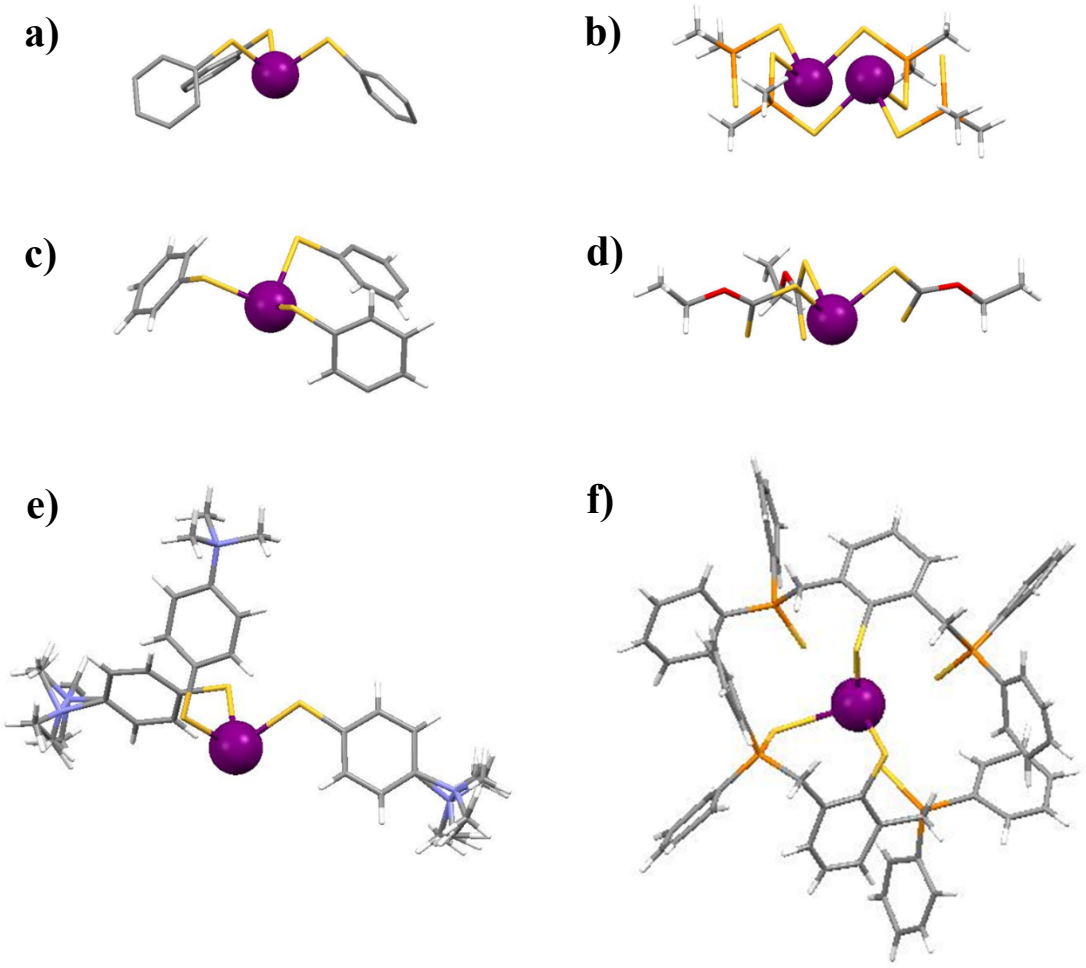

Figure 1-9. Examples of tris-thiolate Pb(II) small molecule complexes. Representing in a) Tetraphenylarsonium tris(phenylthiolato)-lead [$C_{18}H_{15}PbS_3^-$,$C_{24}H_{20}As^+$] *(Note: Only the Pb(II) complex is shown.)*, b) bis(μ$_2$-Dimethyldithiophosphinato-S,S')-bis(dimethyldithio-phosphinat-S,S')-di-lead [$C_8H_{24}P_4Pb_2S_8$], c) Tetra-n-propylammonium tris(phenylthiolato-S)-lead [$C_{18}H_{15}PbS_3^-$,$C_{12}H_{28}N^+$] *(Note: Only the Pb(II) complex is shown.)* d) Tetraethylammonium tris(O-ethylxanthato)-lead [$C_9H_{15}O_3PbS_6^-$,$C_8H_{20}N^+$] *(Note: only the Pb(II) complex is shown)*, e) tris(4-(Trimethylammonio)benzenethiolato-S)-lead(II) diperchlorate [$C_{27}H_{39}N_3PbS_3^{2+}$, $2(ClO_4^-)$] and f) bis(2,6-bis(Diphenylthiophosphinoylmethyl)benzenethiolato)-lead [$C_{64}H_{54}P_4PbS_6$].[17,75–78] Molecules are represent as sticks (sulfur=yellow, carbon=grey, phosphorus=orange, nitrogen=blue, red=oxygen and hydrogen=white). Pb ions are shown as purple spheres.

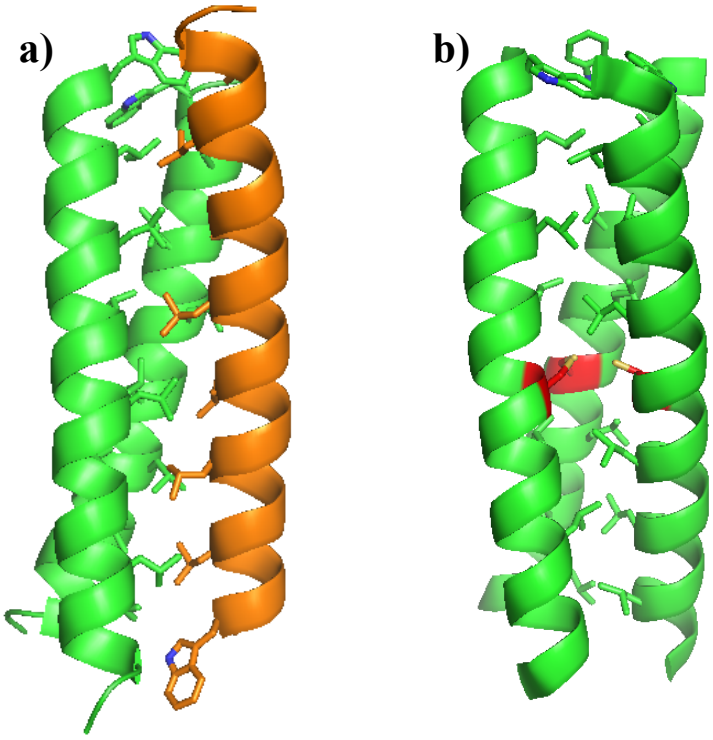

Figure 1-10. Side view of the triple-stranded coiled coils comparing the antiparallel 3SCC orientation observed in the native CoilSer structure (PDB code: 1COS)[72] and the parallel 3SCC fashion observed in a Cys-substituted CoilSer peptides. Shown are a) CoilSer and b) apo-(**CS**L16C)$_3$ crystal structures. The peptide backbones are shown as ribbon diagrams and the interior residues are as sticks. The orange helix in a) emphasizes its opposite orientation from the other two helices in the coiled coil, representing the antiparallel conformation. The Cys residues in b) are red (sulfur = yellow).

latter peptide to have up to ten salt bridge interactions. Moreover, a number of water molecules were found to mediate H-bonding interactions with charged side chain residues, while fewer were observed in the antiparallel **CoilSer** structure. These interactions were thought to strongly help stabilize the orientation of parallel triple-stranded coiled coil in the As(**CS**L9C)$_3$ structure. In addition, 2D NMR experiments also supported that the Cys-substituted peptides were stabilized in a parallel fashion at pH 7.7 (Iranzo, O. and Pecoraro, V. L., unpublished data), even without metal ion coordination linking the peptides. In designs containing a high number of mutations or drastically destabilizing substitutions (e.g. D-amino acid substitution), peptides containing an extra heptad, termed as **GRAND-CoilSer** series (the sequence is given in **Table 1-1**), have been developed to add more stability to the construct. In the Appendix A section, I introduce the overall scaffold features of elongating **GRAND-CoilSer** derivatives which were crystallized, solved and structurally characterized for the first time during these studies. This **GRAND-CoilSer** 3SCC is compared to the shorter **CoilSer** in order to distinguish the distinct structural properties of the elongation.

In the first and second parts of Chapter 2 (*Section I* and *Section II*), I present high resolution X-ray crystal structures of apo-(**CS**L16)$_3$, Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ determined to 1.43, 2.09 and 2.17 Å, respectively. Crystallization and data collection of the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ peptide were performed by Dr. Melissa Zastrow. The metallated crystal structures allow for speculating on the different structural requirements of a trigonal planar Hg(II) compared to a trigonal pyramidal Pb(II) upon binding to a three Cys environment prepared in 3SCC peptides. I discuss how the Cys orientations in the apo-structure change their rotamer organization upon Hg(II) and Pb(II) binding, and address the effect of metal structures on the hydrophobic packing around the metal site. The knowledge gained from this part of the studies should establish general rules for heavy metal binding to Cys-rich sites in designed proteins which may provide insight for understanding how Hg(II) and Pb(II) bind to MerR and PbrR metalloregulatory proteins, respectively.

c) Studies of Cd(II)-protein Interactions in 3SCC Peptides

As with Hg(II) and Pb(II), bacteria also have evolved metalloregulatory proteins for Cd(II) detoxification using Cys coordination chemistry. The class that is responsible for Cd(II) regulation is called the SmtB/ArsR family which functions as transcriptional repressors to respond to the

stress generated by the heavy metal toxicity.[8,79–82] Cd(II) is a transition metal that has a filled *d*-shell similarly to Zn(II). Thus, Cd(II) biology is consequently related to Zn(II) and the toxicity of Cd(II) occurs when Cd(II) interferes with Zn(II)-active site enzymes by replacing the Zn(II) ion because both of the metals prefer to bind into a tetrahedral or pseudo-tetrahedral geometry. However, the increase in crystal ionic radius of Cd(II) (atomic radius for 4-coordinate structure = 0.92 Å) relative to Zn(II) (atomic radius for 4-coordinate structure = 0.72 Å) also allows Cd(II) to adopt higher coordination numbers to five, six, seven and eight.[83] The coordination chemistry of Cd(II) in these cadmium-detecting proteins are interesting as two specific members of SmtB/ArsR proteins; CadC and CmtR were found to coordinate to Cd(II) with different geometries. In the homodimeric repressor CadC, it has been an argument whether the cadmium-detecting transcriptional protein binds to the toxic metal ion as CdS_4 or CdS_3X where X is an oxygen ligand from either an exogenous water or Glu/Asp amino acid side chain when the fourth thiol is weakly associated or not bound.[79,81,84] Mutagenesis studies revealed that the metal ligands for Cd(II) site are Cys58, Cys60 from one monomer, and Cys7 and Cys11 are from the second. This Cys11 residue is the one that has been shown to be non-essential for metal binding. It may contribute to the coordination as the fourth ligand to form a distorted CdS_4 site or weakly bound to the site. It has been proposed that this coordination mode of CdS_4/CdS_3X might be related to the dynamics of the metal exchange.[85] This is a biological example where Cd(II) is bound in its preferred geometry as a 4-coordinate species. In contrast, the structure reported from Banci *et al*. based on 2D NMR of Cd(II)-bound CmtR revealed a CdS_3 coordination geometry (**Figure 1-11**).[80] Such low coordination number mononuclear Cd(II) tris-thiolate complexes are rare and the correlated specific properties are not well-understood because the large complexity of native proteins make the site challenging to be characterized. Moreover, attempts in small molecule mimics studies revealed that among Cd(II) modeled complexes deposited in the Structural Database system[26], only 0.03% of them represent a mononuclear Cd(II)-trithiolate center based on the deposited structures updated to August 2015. As can be seen from an example in **Figure 1-12**, the compound was achieved using ligands not identical to protein ligands and it was prepared in non-aqueous solution which does not always represent the chemistry of a trigonal planar Cd(II) structure in water.[17,87,88] Due to the lack of suitable models using traditional approaches, the Pecoraro group utilized the **TRI**-family peptide, **TRI**L16C, to study Cd(II)-binding.[33,36,37,39–41,46,74,87,88] Intriguingly, unlike other metals in the system, Cd(II) was found to bind to the

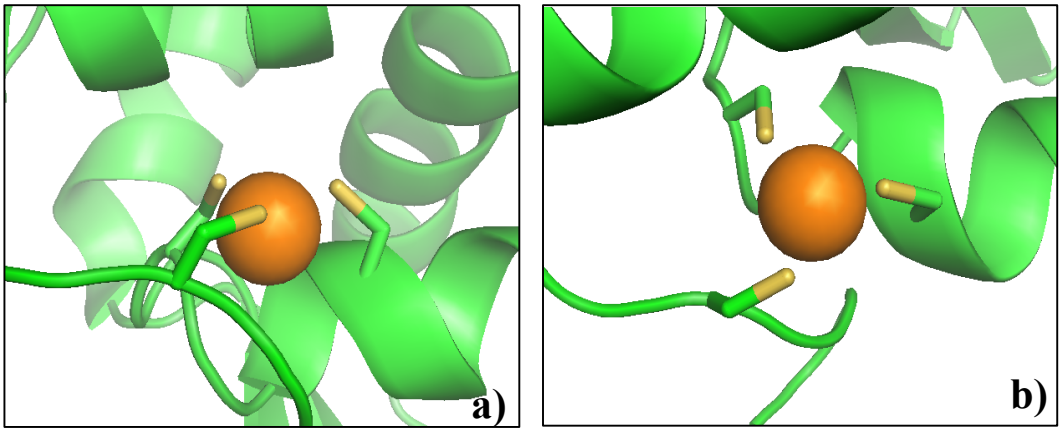

Figure 1-11. PyMOL representation of Cd(II) bound in a tris-thiolate binding site based on the NMR solution structure of CmtR (PDB code: 2JSC).[80] Shown are a) top down view and b) side view of the binding site. Main chain atoms are represented as green helical ribbons, Cys residues as sticks with thiols labeled as yellow and Cd(II) ion as an orange sphere.

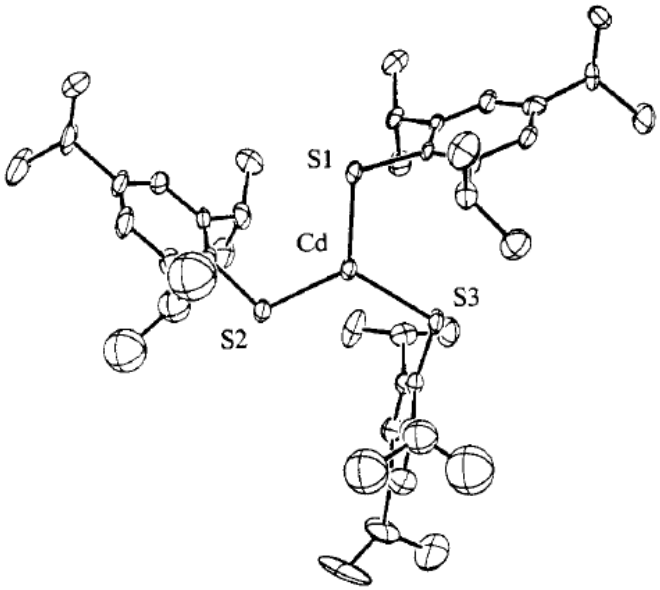

Figure 1-12. ORTEP diagram of $[Cd(S\text{-}2,4,6\text{-}i\text{-}Pr_3C_6H_2)_3]^-$ complex.[27]

TRIL16C peptide in two distinct geometries; CdS_3 and CdS_3O.[39] Much effort has later focused on how to engineer the binding site environment in order to control specific coordination geometries on Cd(II) in the 3SCCs. The knowledge gained is hoped to help clarify Cd(II)-protein interactions that might have been found in nature e.g. CadC and CmtR and to emphasize the strong ability of the "bottom-up approach" of *de Novo* protein design to build a small and stable metalloprotein model to understand basic principles of nature at an unprecedented level.

The main concept to enforce a desired geometry is to engineer the steric environment around the metal pocket either directly through the first coordination sphere ligand modification or indirectly through second sphere ligand modifications. The coordination environment of the designed peptides were probed by the joint use of ^{113}Cd Nuclear Magnetic Resonance (NMR) [89–92] and ^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopies,[93,94] which provide useful information on the coordination environment of cadmium as a result of the different timescales of the two methods. With a timescale of 0.01-10 ms, ^{113}Cd NMR can be used to characterize the number of unique cadmium environments present (number of resonances) as well as help identify the number and type of ligands bound to Cd(II) due to the chemical shift range. Based on the previous published results of the *a* site metal binding peptides, 3- and 4-coordinate thiolate sites have ^{113}Cd NMR chemical shifts that are between 680-700 ppm and 570-600 ppm, respectively. A mixture of 3- and 4-coordinate species falls in 600-680 ppm range.[95] These chemical shift values are now used as a reference for Cd(II) geometry assignment in the system. However, if the ligands bound to the metal exchange faster than this timescale but slower than 0.1-100 ns, ^{111m}Cd PAC can provide more important information about the electronic environment, metal ion symmetry, and dynamics of the site. The PAC spectrum provides information about first coordination sphere ligands (ω_o) and symmetry around the *z*-axis (η).[94,95] The ω_o values gained from previous research of the *a* site-metal binding peptides fall into the frequency regions; 0.450 rad/ns, indicating a trigonal planar Cd(II), and 0.350 rad/ns, indicating a pseudo-tetrahedral Cd(II) geometry.[95]

The parent construct, **TRI**L16C, which contains a Cys residue in an *a* site at the sixteenth position, forms two species in the presence of Cd(II) at pH 8.5. As illustrated in **Figure 1-13** and **Table 1-3**, only one species was observed via ^{113}Cd NMR (δ=625 ppm); however, the ^{111m}Cd PAC spectrum showed that the Cd(II)(**TRI**L16C)$_3^-$ complex contained two species: 40% (ω_o=0.44 rad/ns) of the Cd(II) centers is trigonal planar [Cd(II)S_3^-] and 60% (ω_o=0.34 rad/ns) is pseudo-tetrahedral [Cd(II)S_3O^-], where O is an exogenous water ligand.[39] A **TRI** peptide that would bind

Cd(II) exclusively in a 3- or 4- coordinate geometry was then designed by modifying the steric environment around the metal binding site. This was first achieved by incorporating an Ala residue in place of a Leu at the twelfth position (**TRI**L12AL16C), to give a 4-coordinate species by providing a pocket for an exogenous water molecule above the metal binding plane (with reference to the N-terminus). Cd(II)(**TRI**L12AL16C)$_3^-$ to be a 100% 4-coordinate (Cd(II)S$_3$O$^-$) species, both ^{111m}Cd PAC (ω_o =0.34 rad ns^{-1}) and ^{113}Cd NMR (574 ppm) confirmed.[96,97] The isolation of a fully 3-coordinate species was achieved by increasing the steric bulk of the first sphere coordination ligands around the metal binding site. Here, a non-coded amino acid, Penicillamine (Pen), was incorporated in the sixteenth position.[98,99] Pen can be considered as a bulky derivative of Cys, which has two methyl groups in place of the β-methylene hydrogens in Cys or a thiol-containing derivative of valine (Val) in **Figure 1-14**. Therefore, **TRI**L16Pen was designed to prevent water from binding to Cd(II) to favor a 3-coordinate species. Based on the ^{113}Cd NMR chemical shift of 684 ppm and the ^{111m}Cd PAC data (ω_o = 0.45 rad ns^{-1}), Cd(**TRI**L16Pen)$_3^-$ is fully 3-coordinate Cd(II)S$_3^-$.[99] The published crystal structure of apo-(**CS**L16Pen)$_3$ in **Figure 1-15** proposed that the water exclusion when Cd(II) is bound in the metal site could possibly result from the enhanced packing of Leu hydrophobic layers above and below the Pen plane, generating a dehydrated CdS$_3$ complex. This phenomenon was reported to result from the restricted rotation of the γ-thiols due to the constraint generated from the γ-methylene groups of Pen ligands at the sixteenth position.

Crystallographic Analyses of Steric Modifications of the Coordination Spheres to Control a Metal Coordination Geometry

Due to the lack of Cd(II)-bound crystal structures; however, to still be able to evaluate the control of Cd(II) coordination numbers in structural details, I have used the trigonal planar Hg(II)S$_3^-$ data gained from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ to represent a 3-coordinate Cd(II) structure. The trigonal Hg(II)S$_3^-$ and Cd(II)S$_3^-$ reveal very similar M-S bond lengths based on the X-ray absorption data, therefore, the 3-coordinate Hg(II) structures could be used as analogues to explain the chemistry of Cd(II). In the *Section III* Chapter 2, I distinguish the ligand organization between Cys and Pen upon Hg(II) binding using the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and apo-(**CS**L16C)$_3$ in combination with the

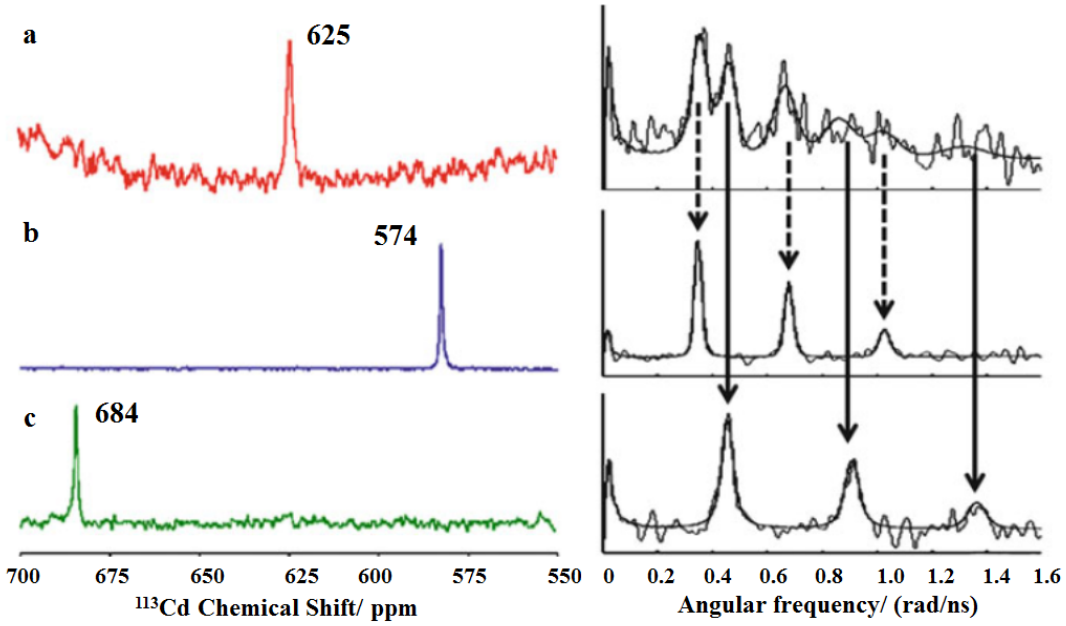

Figure 1-13. ^{113}Cd NMR and corresponding ^{111m}Cd perturbed angular correlation (PAC) spectra showing a) a mixture of $Cd(II)S_3^-$ and $Cd(II)S_3O^-$ for $Cd(TRIL16C)_3^-$, b) 100% $Cd(II)S_3O^-$ for $Cd(II)(TRIL12AL16C)_3^-$, and c) 100% $Cd(II)S_3^-$ for $Cd(II)(TRIL16Pen)_3^-$, respectively. This figure is taken from reference 103.

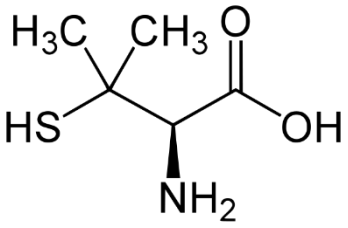

Figure 1-14. A fisher diagram representing an non-coded amino acid Penicillamine (Pen).

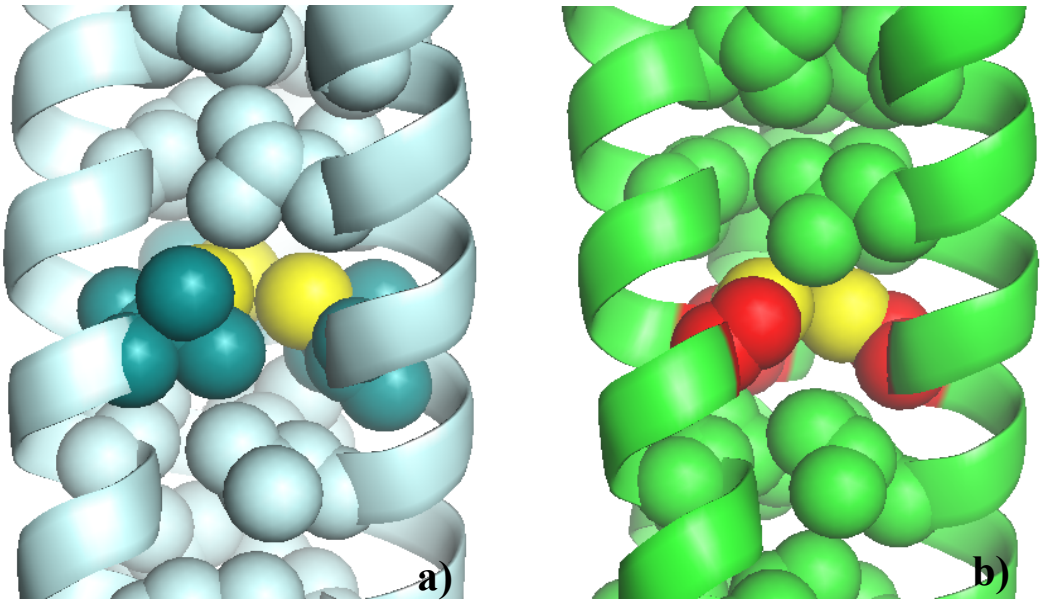

Figure 1-15. Side view of the apo-(CSL16Pen)$_3$ (PDB code: 3H5F)[99] and apo-(CSL16C)$_3$ crystal structures illustrating the steric arrangement of the thiol ligands resulting from two additional γ-methyl groups in penicillamine at the sixteenth position compared to Cys ligands. Shown in a) apo-(**CS**L16Pen)$_3$ and b) apo-(**CS**L16C)$_3$. Main chain atoms are represented as helical ribbon diagrams and internal residues as spheres. Pen and Cys ligands are cyan and red, respectively where sulfurs are color-coded in yellow.

Table 1-3: Spectroscopic values for Cd(II)/**TRI**-family peptides

Peptide	λ (Δε) (nm)	Cd(II)-S bond length (Å)	δ (ppm) (^{113}Cd NMR)	ω_0 (rad/ns), η (^{111m}Cd PAC)
Cd(II)(**TRI**L16C)$_3^-$	232 (22600)[a]	2.49[a]	625[a]	0.337±0.002[a] 0.438±0.004
Cd(II)(**TRI**L12A16C)$_3^-$	231(21000)[b]	2.50[c]	574[b]	0.3405±0.0003[b]
Cd(II)(**TRI**L16Pen)$_3^-$	235(23600)[b]	2.46[d]	684[b]	0.454±0.001[b]

[a] Reference 39 ; [b] Reference 97 ; [c] Data obtained from these studies and [d] Data for Cd(**GRAND**L16Pen)$_3^-$ obtained in these studies.

previously reported structures, apo-(**CS**L16Pen)$_3$ and [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$_2$)]$_N$(**CS**L9PenL23H)$_3^{n+}$.[99,101] In this regard, I address how the *predisposition* of Cys ligands versus the *preorganization* of Pen environment could affect the steric hindrance above the metal site. These structural observations allow me to hypothesize why a Cd(II)S$_3$O$^-$ complex is occasionally formed in Cd(II)(**TRI**L16C)$_3^-$ whereas the water ligand is completely absent in Cd(II)(**TRI**L16Pen)$_3^-$. Moreover, I have generated two more high resolution crystal structures of Hg(II)(**GRAND-CS**L12A16C)$_3^-$ and Zn(II)(**GRAND-CS**L12A16C)$_3^-$ determined to 1.93 and 1.89 Å resolution, respectively, to gain insight on how the substitution of Ala above the metal site influences water access to this peptide region. The pseudo-tetrahedral Zn(II)S$_3$O$^-$ from the Zn(II)(**GRAND-CS**L12A16C)$_3^-$ structure is essential as it reveals basic information of how the designed trithiolate environment could form with a 4-coordinate metal structure. In addition, I provide possible models of Cd(II) centers for Cd(II)S$_3^-$ and Cd(II)S$_3$O$^-$ using the known crystal structures obtained in these studies in order to predict possible Cd(II) structures prepared using the simplified 3SCC designs. In terms of the perspective of protein design, these analyses reveal deeper insight of how ligand engineering is a powerful tool to compel a metal into a specific coordination structures which are useful for broad biophysical applications when one would want to design a desired metal site in a protein to perform a specific function of interest.

Determining Ligand Organization upon Metal Binding into Different Geometries

One of the principles to discover the selectivity and specificity of the metals to their biomacromolecular targets is to understand the orientation of the ligands prior to binding a metal. Among the metalloregulatory proteins, it appears that the core feature of metal binding is the use of three sulfur atoms from Cys ligands that can bind metals in different geometries based on the geometrical preference of the metals. For example, in MerR this tris-thiolate site can form a trigonal planar Hg(II), while ArsR and PbR691 binds As(III) and Pb(II) in a trigonal pyramidal structure, while ZntR binds Zn(II) in a pseudotetrahedral geometry. Unfortunately, not all of these binding sites have been structurally characterized in the native proteins, therefore, structural detail revealing the arrangement of these three Cys ligands has not been achieved for these different geometries. Using 3SCC peptides in *Section IV* of Chapter 2, I can integrate the four structures (trigonal planar Hg(II)S$_3^-$ from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, trigonal pyramidal

$Pb(II)S_3^-$ from $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$, 4-coordinate $Zn(II)S_3O^-$ from Zn(II)(**GRAND-CS**L12A16C)$_3^-$ and a trigonal pyramidal $As(III)S_3$ from the previously published As(III)(**CS**L9C)$_3$[74] to differentiate the requirements for each metal polyhedron to be constructed into the 3SCC environments. These metal structures are not only compared with each other, but also analyzed in consideration with the apo-structures (apo-**CS**L16C)$_3$ and in some cases with the published apo-(**CS**L9C)$_3$[73] in order to illustrate how the ligand orientations are changed upon metal binding and how these resultant metallated structures perturb the hydrophobic packing around the metal site. I will assign the apo-protein as either *predisposed* or *preorganized* for metal binding based upon the change of ligand orientation upon binding a metal. I will describe the system as *predisposed* when there are three Cys side chains available that require significant rotamer reorganization prior to binding a metal. In contrast, the system is said to be *preorganized* when the Cys ligands in the apo-coordination do not have to significantly rotate to bind a metal. The classification of these structures in this way allows for an exploration of metal binding, which could help in unraveling the molecular basis for heavy metal recognition and general rules for metal selectivity that could potentially occur in biological systems.

Using Alternative Chirality of D-amino acids in de Novo Designed Peptides to Enforce Specific Coordination Numbers on Cd(II)

Unlike other protein modification strategies, *de Novo* metalloprotein design is less limited in using non-coded amino acids in the primary sequence.[97,99,102–105] The introduction of artificial amino acids greatly expands the versatility of the protein sequence pool far beyond the limit sets by the common 20 amino acid building blocks. Modification of designed proteins with this broader range of amino acids might reveal some novel physical properties not yet appreciated in native proteins.[104,106–108] D-amino acids constitute one class of non-coded protein building blocks. It has been shown that the mutation of one specific amino acid from a normal L-amino acid to a D-amino analogue does not dramatically affect the overall structure of the whole protein and the replacement site.[101,102,104] Indeed, such changes may engender physical properties that may be useful for future biophysical and pharmaceutical applications.

The stereochemistry of D-amino acids is such an interesting tool to extend the scope of *de Novo* design. One benefit in the design aspect is that they are common non-coded amino acids, thus the side chains remain chemically the same as their L-analogs. In addition, the chirality offers

a geometrically non-superimposable structure compared to its own mirror image. The *de Novo* protein design strategy exploits these "first principles" of the altered configuration to constrain the folding of peptide chains chemically in a specific position because this chirality imposes limited regions of conformational space which results in desired folding.[109,110] Herein, I increase the amount of research addressing the incorporation of D-amino acids as stereochemical directors to generate a number of polypeptide structures. Previous examples of this approach include the use of D-Proline (D-Pro) to force prime turns to nucleate antiparallel β-hairpin formation in LD dipeptide sequences.[106,110] D-Aspatic acid (D-Asp) placed in the N-capped position acts as a strong conformational determinant to induce a type II β-turn conformation and 3_{10} helical structure.[111] Moreover, the stereoisomer of D-amino acids has also been employed to create hetero-chiral fold designs of bracelet and boat-shaped molecules[112,113], a canoe-like construct as an alkali receptor[114], a pi-cup scaffold as an acetylcholine receptor[115] and the Zn-finger hydrolase (αββ fold) structure.[116] A series of alternating LD sequences was also designed to self-assemble into the extended β-helices winding around like a cylinder to mimic the native chiral alternating sequence of gramicidin found in *Bacillus brevis*.[110,117–119] Another prominent advantage of employing D-amino acids is as a helical terminator as evident in D-Alanine and D-Leucine facilitating the Schellman motif to terminate helix formation.[120,121] These examples clearly prove that D-residues strongly provide stereochemical conformational effects that may be used in rational *de Novo* designs. These units considerably expand polypeptide templates that possibly pave ways toward designing novel peptide architectures. Until my work, no studies have been reported that use D-amino acids to engineer metal environments or to control metal geometries.

Following the exploration of the **TRI**-family modifications on controlling Cd(II) structures in Chapter 2 *Section III*, it was shown that the type and position of the second coordination sphere residues can significantly influence Cd(II) coordination geometry as demonstrated from the evidence of steric removal by Ala substitution in **TRI**L12AL16C design that allows 100% CdS_3O. It was consequently proposed that the addition of steric hindrance above the metal site could lead to water exclusion. Peacock *et al.* exploited the reorientation of D-Leu to introduce such steric constraint, while still being able to retain the chemical properties of leucine in the twelfth position, yielding **TRI**L12$_D$LL16C. The ^{113}Cd NMR chemical shift of 697 ppm and the ^{111m}Cd PAC data (ω_o= 0.45 rad ns^{-1}) demonstrated that Cd(II)(**TRI**L12$_D$LL16C)$_3^-$ is a 100% 3-coordinate species. Given the success of these initial designs, I have used the stereochemistry of D-amino acids to

control Cd(II) geometry from the second coordination sphere with the goal of understanding how to fine-tune the design principles of diastereopeptides capable of metal sequestration. To achieve this goal, I needed to learn how the incorporation of D-amino acids affect the folding of α-helical coiled coils. In the next data chapters of my dissertation, I describe using D-Leu to modify the second coordination sphere of the metal site (Chapter 3) and D-Cys, as well as D-Pen, to perturb the first coordination sphere (Chapter 4) to control the coordination numbers of Cd(II) centers. The main goal of these particular studies is to open up new vistas in *de Novo* protein design and the opportunity to control metal active site and metal structure selectivity using an alternative enantiomer of amino acid ligands.

a) Incorporation of D-Leu in the Second Coordination Sphere of the Tris-thiolate Environment to Control Cadmium Coordination Geometry in de Novo Designed Peptides

In Chapter 3, to obtain deeper understanding of the impact of the alternative chirality of D-Leu on designed peptides, a second coordination sphere modification to design a diastereopeptide with a hole below (referenced to the N-terminus) was undertaken. The goal of this work was to alter the Cys layer of **TRI**L16C to yield a higher Cd(II) coordination number complex. To this end, I have designed **TRI**L2WL16CL19$_D$L, in which the Leu side chain is proposed to be reoriented toward the C-terminus. It is hypothesized that this construct would contain a pocket below the metal binding site in order to provide space for an exogenous solvent molecule. Additionally, I have proposed that the double substitution of D-Leu residues in the twelfth position (to block the access of water) at the same time with the nineteenth position (to allow for water inclusion) in the **GRAND-CS**L12$_D$LL16CL19$_D$L design could result in a predominantly 4-coordinate Cd(II) species. A series of physical characterizations (UV/VIS, CD spectroscopy, ^{113}Cd NMR and ^{111m}Cd PAC) are described in order to evaluate the Cd(II) centers that result from these modifications.

Though spectroscopic evaluation has confirmed a high degree of control in the designed Cd(II) binding sites, the assumption of the altered orientation of D-Leu in the second coordination sphere has not been structurally addressed previously. Based on these designed sequences, I have generated a series of high resolution crystal structures of D-Leu containing peptides. The apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and Hg(II)-(**GRAND-CS**L12$_D$LL16C)$_3^-$ structures, determined to 1.34 and 2.11 Å resolution, provides structural information for the published **TRI**L12$_D$LL16C

peptide and verifies how water blockage can occur by the substitution of the 12D-Leu ligands. Moreover, apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and Hg(II)-(**GRAND-CS**L16CL19$_D$L)$_3^-$, determined to 1.83 and 1.93 Å resolution, are for **TRI**L2WL16CL19$_D$L to demonstrate crystallographically how 19D-Leu can open the space below the metal site, resulting in increased water access. In this session, I determine the effect of D-Leu compared to L-Leu using the L-amino acid analogues of apo-(CSL16C)$_3$ and metallated Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ from Chapter 2 to support the results obtained from those physical characterization studies. Finally, models of Cd(II) centers gained from these diastereopeptide designs are provided to investigate Cd(II) binding chemistry in the D-Leu engineered environments fully.

b) Perturbation of the Metal Coordination Sphere by D-amino Acids to Control Cadmium Structures in de Novo Designed 3SCC Peptides

It will be shown in Chapter 4 that the substitution of D-Leu into an L-Leu based 3SCC does not significantly disrupt the native folding of the secondary and tertiary structures of α-helical coiled coil. In addition, this non-coded amino acid has the potential to indirectly perturb the steric environment around the metal site resulting in the structural control of Cd(II) structures. This success has generated another intriguing question which is how the Cd(II)-peptide interactions would be perturbed if the inverted stereochemical configuration is introduced into the metal first coordination sphere. As mentioned previously, when L-Cys was used as a metal ligand and incorporated at the sixteenth position, the resulting Cd(II)(**TRI**L16C)$_3^-$ yielded a mixture of 40% Cd(II)S$_3^-$ and 60% Cd(II)S$_3$O$^-$; while the substitution of L-Pen at the same position is completely able to exclude water from the site giving 100% Cd(II)S$_3^-$ as seen with Cd(II)(**TRI**L16Pen)$_3^-$. At this point, it is instructive to assess whether the Cd(II) polyhedron is perturbed by either D-Cys or D-Pen ligands.

A previous undergraduate student in the Pecoraro laboratory, Cheri Pascoe, designed **TRI**L16$_D$C, where the D-Cys replaced L-Cys at the sixteenth position. Circular dichroism spectra revealed a reasonable initial percent folding for both apo-(**TRI**L16$_D$C)$_3$ and metallated (**TRI**L16$_D$C)$_3^-$ peptides indicating that a single mutation of D-Cys along the **TRI** variant is tolerable. The Cd(II) binding titration performed by UV/Vis and ^{113}Cd NMR (chemical shift of 646 ppm) confirmed that Cd(II) binds to this peptide in a tris-thiolate environment at pH 8.5. When I joined this project, the physical characterization of these systems was partially completed. I have

examined the ^{111m}Cd PAC data for these systems (with Lars Hemmingsen) and performed crystallization studies to address the effect of D-Cys on metal structure. Using **CoilSer** derivatives, I collected X-ray data on a series of apo-(**CS**L16$_D$C)$_3$ and Zn(II)-(**CS**L16$_D$C)$_3^-$ crystal structures, determined to 1.69 and 1.92 Å resolution, respectively. I address the impact of chirality between D-Cys versus L-Cys on the thiol environments by comparing the new apo-(**CS**L16$_D$C)$_3$ with the apo-(**CS**L16C)$_3$ described in Chapter 2. Moreover, I assess the degree of *preorganization* or *predisposition* of apo-chiral ligands using the Zn(II)-(**CS**L16$_D$C)$_3^-$ compared to Zn(II)(**GRAND-CS**L12A16C)$_3^-$ discussed in Chapter 2 to understand Zn(II)-binding. Additionally, revisiting the Cd(II) question, I propose a reasonable possibility for the Cd(II) structure in Cd(**TRI**L16$_D$C)$_3^-$ based on the known Zn(II)-(**CS**L16$_D$C)$_3^-$ structure.

The attempt to incorporate D-Pen into the **TRI**-family system to study Cd(II) interactions was first initiated by Iranzo *et al*.[36] The chemical shift of Cd(**TRI**L16$_D$Pen)$_3^-$ was observed at 557 ppm, which clearly does not fall in the range of chemical shifts observed for CdS_3 and CdS_3O complexes correlated from the chemical shift results of the *a* site **TRI** peptide derivatives. Rather, this chemical shift is more consistent with Cd(II)S_2O_2 or Cd(II)SO_3^+ [95], which is clearly distinguished from a pure Cd(II)S_3^- when the opposite stereochemical configuration, L-Pen, was incorporated at the same position in **TRI**L16Pen. These data primarily suggested that the Cd(II) interactions are very different upon the inversion of the chirality of penicillamine ligands. Peacock *et al*. assessed the structural information of D-Pen ligands by obtaining a crystal structure of apo-(**CS**L16$_D$Pen)$_3$.[99] Show in **Figure 1-16**, the directions of D-Pen versus L-Pen are completely opposite through space (determined based on the major conformers). The D-Pen residues are oriented outward to the helical interface, whereas the L-Pen are inward pointing to the core of the structure. To further identify the structural details of D-Pen non-natural ligands on metal binding in the 3SCC design, I have generated a high resolution crystal structure of Hg(II)(**CS**L16$_D$Pen)$_3^-$, determined to 1.71 Å. Due to the difficulty in achieving Cd(II)(**CS**L16$_D$Pen)$_3^-$, the metallated version of Hg(II) binding to D-Pen may still be beneficial as it can suggest possible D-Pen ligand orientations when bound to a metal center compared to the apo-coordination that was previously reported in the apo-(**CS**L16$_D$Pen)$_3$ structure.[99] To examine the effect of chirality further, I address the difference of ligand orientations between L-Pen and D-Pen in non-metallated forms using the apo-(**CS**L16$_D$Pen)$_3$ and apo-(**CS**L16Pen)$_3$ structures compared to the metallated versions obtained from the new Hg(**CS**L16$_D$Pen)$_3^-$ and the published [Hg(II)]$_S$[Zn(II)(H_2O/

$OH^-)]_N$(**CS**L9PenL23H)$_3^{n+}$structures. The results of these structural studies could allow for a better understanding of how the chirality of penicillamine analogues alter the coordination environment and potentially change the physical location of the metal ion within the construct.

In chapter 5, a summary of the results will be given and the future directions in terms of extended peptide designs and biophysical applications suggested from the knowledge gained from these designs will be provided.

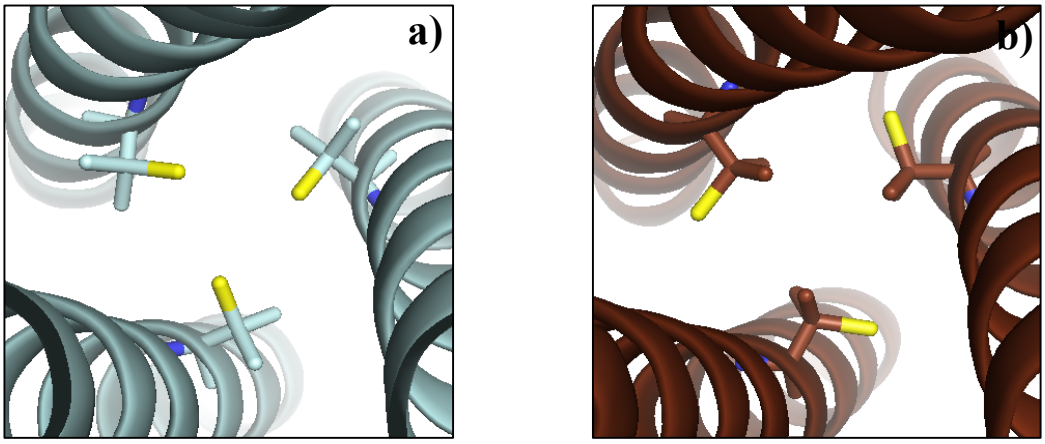

Figure 1-16. Ribbon diagrams shown from a top down view of the 3SCC representing the orientations of L-Pen versus D-Pen ligands. In a) L-Pen environment based on apo-(**CS**L16Pen)$_3$ (PDB code: 3H5F) and b) D-Pen ligands from **apo-(CS**L16$_D$Pen)$_3$ (PDB code: 3H5G).[99] The ligands are shown as sticks with sulfurs labeled as yellow.

References

(1) Zimmermann, M.; Clarke, O.; Gulbis, J. M.; Keizer, D. W.; Jarvis, R. S.; Cobbett, C. S.; Hinds, M. G.; Xiao, Z.; Wedd, A. G. *Biochemistry* **2009**, *48*, 11640–11654.

(2) Rensing, C. *J. Bateriology* **2005**, *187*, 3909–3912.

(3) Brosnan, J.; Brosnan, M. *J. Nutr.* **2006**, *136*, 16365–16405.

(4) Jessop CE. Biochem Soc Trans. 2004, 32, 655 – 8.

(5) Stead, L. M.; Brosnan, J. T.; Brosnan, M. E.; Vance, D. E.; Jacobs, R. L. *Am. J. Clin. Nutr.* **2006**, *83*, 5–10.

(6) Mehra, R. K.; Mulchandani, P. *Biochem. J.* **1995**, *307*, 697–705.

(7) Chekmeneva, E.; Prohens, R.; Díaz-Cruz, J. M.; Ariño, C.; Esteban, M. *Anal. Biochem.* **2008**, *375*, 82–89.

(8) Busenlehner, L. S.; Pennella, M. a.; Giedroc, D. P. *FEMS Microbiol. Rev.* **2003**, *27*, 131–143.

(9) Brown, N. L.; Stoyanov, J. V; Kidd, S. P.; Hobman, J. L. *FEMS Microbiol. Rev.* **2003**, *27*, 145–163.

(10) Shewchuk, L. M.; Verdine, G. L.; Walsh, C. T. *Biochemistry* **1989**, 2331–2339.

(11) Helmann, J. D.; Ballard, B. T.; Walsh, C. T. *Science* **1990**, *247*, 946–948.

(12) Ralston, D. M.; O' Halloran, T. V. *Proc. Natl. Acad. Sci* **1990**, *87*, 3846–3850.

(13) Chang, C.-C.; Lin, L.-Y.; Zou, X.-W.; Huang, C.-C.; Chan, N.-L. *Nucleic Acids Res.* **2015**, *13*, 7612-7623.

(14) Barkay, T.; Wagner-Döbler, I. *Adv. Appl. Microbiol.* **2005**, *57*, 1–52.

(15) Barkay, T.; Miller, S. M.; Summers, A. O. *FEMS Microbiol. Rev.* **2003**, *27*, 355–384.

(16) Barkay, T.; Kritee, K.; Boyd, E.; Geesey, G. *Environ. Microbiol.* **2010**, *12*, 2904–2917.

(17) Christou, G.; Folting, K.; Huffman, J. C. *Polyhedron* **1984**, *3*, 1247-1253.

(18) Freedman, Z.; Zhu, C.; Barkay, T. *Appl. Environ. Microbiol.* **2012**, *78*, 6568–6575.

(19) Wright, J. G.; Tsang, H.; Penner-hahn, J. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, 2434–2435.

(20) Skyllberg, U.; Bloom, P. R.; Qian, J.; Lin, C. M.; Bleam, W. F. *Environ. Sci. Technol.* **2006**, *40*, 4174–4180.

(21) Kaupp, M.; Von Schnering, H. G. *Inorg. Chem.* **1994**, *33*, 2555–2564.

(22) Rulíšek, L.; Vondrášek, J. *J. Inorg. Biochem.* **1998**, *71*, 115–127.

(23) Watton, S. P.; Wright, J. G.; MacDonnell, F. M.; Bryson, J. W.; Sabat, M.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2824.

(24) Helmann, J. D.; Ballard, B. T.; Walsh, C. T. *Science* **1990**, *247*, 946–948.

(25) Shewchuk, L. M.; Verdine, G. L.; Nash, H.; Walsh, C. T. *Biochemistry* **1989**, *28*, 6140–6145.

(26) Allen, F. H. *Acta Crystallogr., Sect. B: Struct. Sci.* **2002**, *58*, 380–388.

(27) Gruff, E.S.; Koch, K. S. *J. Am. Chem. Soc.* **1990**, *112*, 1245–1247.

(28) DeGrado, W. F.; Summa, C. M.; Pavone, V.; Nastri, F.; Lombardi, A. *Annu. Rev. Biochem.* **1999**, *68*, 779–819.

(29) Lu, Y.; Yeung, N.; Sieracki, N.; Marshall, N. M. *Nature* **2009**, *460*, 855–862.

(30) Zastrow, M.; Pecoraro, V. L. *Coord. Chem. Rev.* **2013**, 2565-2588.

(31) Regan, L.; Degrado, W. F. *Science* **1988**, 976–978.

(32) Degrado, W. F.; Wasserman, Z. R.; Lear, J. D. **1989**, *243*, 622–628.

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

(34) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959–9973.

(35) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. *Dalton Trans.* **2009**, *9226*, 2271–2280.

(36) Iranzo, O.; Thulstrup, P. W.; Ryu, S.-B.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry* **2007**, *13*, 9178–9190.

(37) Farrer, B. T.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 3760–3765.

(38) Ghosh, D.; Pecoraro, V. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 97–103.

(39) 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.

(40) 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**, *7863*, 6195–6196.

(41) Pecoraro, V. L.; Peacock, A. F. A.; Iranzo, O.; Marek, Ł. **2009**, 183–197.

(42) Utschig, L. M.; Bryson, J. W.; Halloran, T. V. O. *Science* **1995**, *268*, 380–385.

(43) Utschig, L.M.; Wright J.G.; O'Halloran T. V. In *Methods Enzymology*, Academic Press: San Diego, **1993**, *226*, 71-97.

(44) Tröger, W. *Hyperfine Interact.* **1999**, 117-128.

(45) Gruff, E.S.; Koch, K. S. *J. Am. Chem. Soc.* **1990**, *112*, 1245–1247.

(46) 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.

(47) Dieckmann, G. R. Ph.D. Dissertation, University of Michigan, 1995.

(48) Łuczkowski, M.; Stachura, M.; Schirf, V.; Demeler, B.; Hemmingsen, L.; Pecoraro, V. L. *Inorg. Chem.* **2008**, *47*, 10875–10888.

(49) Delnomdedieu, M.; Basti, M. M.; Otvos, J. D.; Thomas, D. J. *Chem. Res. Toxicol.* **1993**, *6*, 598–602.

(50) Połeć-Pawlak, K.; Ruzik, R.; Lipiec, E. *Talanta.* **2007**, *72*, 1564–1572.

(51) Zampella, G.; Neupane, K. P.; De Gioia, L.; Pecoraro, V. L. *Eur. J. Inorg. Chem.* **2012**, *18*, 2040–2050.

(52) Neupane, K. P.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2010**, *49*, 8177–8180.

(53) Mah, V.; Jalilehvand, F. *Inorg. Chem.* **2012**, *51*, 6285–6298.

(54) Gibney, B. R.; Rabanal, F.; Reddy, K. S.; Dutton, P. L. *Biochemistry* **1998**, *37*, 4635–4643.

(55) Yuan, H.; Zhang, Y.; Huang, S.; Yang, Y.; Gu, C. *Environ. Sci. Pollut. Res.* **2015**, *22*, 2808–2816.

(56) Simons, T. J. *Eur. J. Biochem.* **1995**, *234*, 178–183.

(57) Neal, A. P. *Mol. Neurobiol.* **2011**, *42*, 151–160.

(58) Zawia, N. H.; Crumpton, T.; Brydie, M.; Reddy, G. R.; Razmiafshari, M. *Neuro Toxicol.* **2000**, *21*, 1069–1080.

(59) Simons, T. J. B. *Neuro Toxicol.* **1993**, *14*, 77–85.

(60) Erskine, P. T.; Duke, E. M. H.; Tickle, I. J.; Senior, N. M.; Cooper, J. B. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2000**, 421–430.

(61) Erskine, P. T.; Newbold, R.; Brindley, A. A.; Wood, S. P.; Shoolingin-Jordan, P. M.; Warren, M. J.; Cooper, J. B. *J. Mol. Biol.* **2001**, *312*, 133–141.

(62) Erskine, P. T.; Senior, N.; Awan, S.; Lambert, R.; Lewis, G.; Tickle, I. J.; Sarwar, M.; Spencer, P.; Thomas, P.; Warren, M. J.; Shoolingin-Jordan, P. M.; Wood, S. P.; Cooper, J. B. *Nat. Struct. Biol.* **1997**, *4*, 1025–1031.

(63) Jaffe, E. K.; Martins, J.; Li, J.; Kervinen, J.; Dunbrack, R. L. *J. Biol. Chem.* **2001**, *276*, 1531–1537.

(64) Hobman, J. L.; Julian, D. J.; Brown, N. L. *BMC Microbiol.* **2012**, *12*, 109.

(65) Borremans, B.; Hobman, J. L.; Provoost, A.; Brown, N. L.; Van De Lelie, D. *J Bacteriol.* **2001**, *183*, 5651–5658.

(66) Chen, P.; Greenberg, B.; Taghavi, S.; Romano, C.; Van Der Lelie, D.; He, C. *Angew. Chemie - Int. Ed.* **2005**, *44*, 2715–2719.

(67) Chen, P. R.; He, C. *Curr. Opin. Chem. Biol.* **2008**, *12*, 214–221.

(68) Zeng, Q.; Stålhandske, C.; Anderson, M. C.; Scott, R. A.; Summers, A. O. *Biochemistry* **1998**, *37*, 15885–15895.

(69) Chen, P. R.; Wasinger, E. C.; Zhao, J.; van der Lelie, D.; Chen, L. X.; He, C. *J. Am. Chem. Soc.* **2007**, *129*, 12350–12351.

(70) Morsali, A.; Mahjoub A. R. *HeIv. Chim. Acta* **2004**, *87*, 2717–2722.

(71) Bridgewater, B. M.; Parkin, G. *J. Am. Chem. Soc*. **2000**, *122*, 7140−7141.

(72) Lovejoy, B.; Choe, S.; Cascio, D.; Mcrorie, D. K.; William, F.; Eisenberg, D.; Degrado, W. F. *Science* **1993**, *259*, 1288–1293.

(73) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2010**, *132*, 13240–13250.

(74) Touw, D. S.; Nordman, C. E.; Stuckey, J. a; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11969–11974.

(75) Dean P.A.W.; Vittal, J.J.; Payne, N. C. *Inorg.Chem.* **1984**, *23*, 4232-4236.

(76) P.Perez-Lourido, J.Romero, J.A.Garcia-Vazquez, A.Sousa, Yifan Zheng, J.R.Dilworth, J. Z.; Trans, D. *J.chem.soc.,dalt. Trans.* **2000**, 769.

(77) Zhi-Gang Ren, Xiao-Yan Tang, Li Li, Guang-Fei Liu, Hong-Xi Li, Yang Chen, Yong Zhang, Jian-Ping Lang, I. C. C. *Inorg.Chem.Commun.* **2007**, 1253.

(78) A.J.Rossini, A.W.Macgregor, A.S.Smith, G.Schatte, R.W.Schurko, G. G. B. *Dalt. Trans. (2013),* **2013**, *42*, 9533.

(79) Kandegedara, A.; Thiyagarajan, S.; Kondapalli, K. C.; Stemmler, T. L.; Rosen, B. P. *J. Biol. Chem.* **2009**, *284*, 14958–14965.

(80) Banci, L.; Bertini, I.; Cantini, F.; Ciofi-Baffoni, S.; Cavet, J. S.; Dennison, C.; Graham, A. I.; Harvie, D. R.; Robinson, N. J. *J. Biol. Chem.* **2007**, *282*, 30181–30188.

(81) Ye, J.; Kandegedara, A.; Martin, P.; Rosen, B. P. *J. Bacteriol.* **2005**, *187*, 4214–4221.

(82) Silver, S. *J. Bacteriol.* **1995**, *177*, 4437–4441.

(83) Borsari, M. *Inorg.Chem.* **2011**, 1–16.

(84) Sun, Y.; Wong, M. D.; Rosen, B. P. *J. Biol. Chem.* **2001**, *276*, 14955–14960.

(85) Tebo, A. G.; Hemmingsen, L.; Pecoraro, V. L. *Metallomics* **2015**, 1–7.

(86) Duhme, A.-K.; Strasdeit, H. *Z. Anorg. Allg. Chem.* **1999**, *625*, 6.

(87) Farrer, B. T.; McClure, C. P.; Penner-Hahn, J. E.; Pecoraro, V. L. *Inorg. Chem.* **2000**, *39*, 5422–5423.

(88) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2011**, *133*, 239–251.

(89) Coleman, J. E. *Method Enzym.* **1993**, *227*, 16–43.

(90) Summers, F. *Coord. Chem. Rev.* **1988**, *86*, 43–134.

(91) Oz, G.; Pountney, D. L.; Armitage, I. M. *Biochem. Cell Biol.* **1998**, *76*, 223–234.

(92) Hemmingsen, L.; Olsen, L.; Antony, J.; Sauer, S. P. a. *J. Biol. Inorg. Chem.* **2004**, *9*, 591–599.

(93) Hemmingsen, L.; Stachura, M.; Thulstrup, P. W.; Christensen, N. J.; Johnston, K. *Hyperfine Interact.* **2010**, *197*, 255–267.

(94) Hemmingsen, L.; Sas, K. N.; Danielsen, E. *Chem. Rev.* **2004**, *104*, 4027–4062.

(95) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry* **2009**, *15*, 3761–3772.

(96) Lee, K.-H.; Matzapetakis, M.; Mitra, S.; Marsh, E. N. G.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2004**, *126*, 9178–9179.

(97) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2864–2868.

(98) I have chosen to use the term "non-coded" to refer to any amino acid other than the normally coded 20 amino acids incorporated into mamalian proteins. This terminology is prefered to "non-natural" since an amino acid such as penicillamine is a natural amino acid which is not normally incorporated (hence non-coded) into a protein.

(99) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7371–7374.

(100) Peacock, A. F. A.; Pecoraro, V. L. *Cadmium: From Toxicity to Essentiality*; Sigel, A.; Sigel, H.; Sigel, R. K., Eds.; Metal Ions in Life Sciences; Springer Netherlands: Dordrecht, 2013; Vol. 11.

(101) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Nat. Chem.* **2012**, *4*, 118–

123.

(102) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16566–16571.

(103) Petros, A. K.; Shaner, S. E.; Costello, A. L.; Tierney, D. L.; Gibney, B. R. *Inorg. Chem.* **2004**, *43*, 4793–4795.

(104) Lu, Y. *Curr. Opin. Chem. Biol.* **2005**, *9*, 118–126.

(105) Kumar, A.; Ramakrishnan, V. *ACS Synth. Biol.* **2010**, *4*, 247–256.

(106) Aravinda, S.; Shamala, N.; Roy, R. S.; Balaram, P. *Proc Indian Acad Sci - Chem Sci* **2003**, *115*, 373–400.

(107) Wang, A.; Nairn, N. W.; Marelli, M.; Grabstein, K. In *Protein engineering*; Kaumaya, P. Ed.; **2012**; 253–290.

(108) Privett, H. K.; Reedy, C. J.; Kennedy, M. L.; Gibney, B. R. *J. Am. Chem. Soc.* **2002**, *124*, 6828–6829.

(109) Balaram, P. *J. Pept. Res.* **1999**, *54*, 195–199.

(110) Mahalakshmi, R.; Balaram, P. In *A New Frontier in Amino Acid and Protein Research*; Nova Science Publishers, Inc., **2006**; 416–430.

(111) Dhanasekaran, M.; Fabiola, F.; Pattabhi, V.; Durani, S. *J. Am. Chem. Soc.* **1999**, 5575–5576.

(112) Rana, S.; Kundu, B.; Durani, S. *Chem. Commun. (Camb).* **2004**, 2462–2463.

(113) Rana, S.; Kundu, B.; Durani, S. *Chem. Commun. (Camb).* **2005**, 207–209.

(114) Rana, S.; Kundu, B.; Durani, S. *Bioorg. Med. Chem.* **2007**, *15*, 3874–3882.

(115) Rana, S.; Kundu, B.; Durani, S. *Biopolymer* **2007**, *87*, 231–243.

(116) Petal, K.; Srivastava, K. R.; Durani, S. *Bioinorg. Med. Chem.* **2010**, 8270–8276.

(117) Ramachandran, G. N.; Chandrasekaran, R. *Indian J. Biochem. Biophys.* **1972**, 1–11.

(118) Ghadiri, M. R.; Granja, J. R.; Buehler, L. K. *Nature* **1994**, *369*, 301–304.

(119) Fernandez-Lopez, S.; Kim, H. S.; Choi, E. C.; Delgado, M.; Granja, J. R.; Khasanov, a; Kraehenbuehl, K.; Long, G.; Weinberger, D. a; Wilcoxen, K. M.; Ghadiri, M. R. *Nature* **2001**, *412*, 452–455.

(120) Aravinda, S.; Shamala, N.; Pramanik, a; Das, C.; Balaram, P. *Biochem. Biophys. Res. Commun.* **2000**, *273*, 933–936.

(121) Aravinda, S.; Shamala, N.; Bandyopadhyay, A.; Balaram, P. *J. Am. Chem. Soc.* **2003**, *125*, 15065–15075.

Chapter 2

Understanding Ligand Complexation upon Heavy Metal Binding into Different Geometries within *de Novo* Three-stranded Coiled Coil Proteins

Background

In the introductory chapter I provided background explaining the importance of heavy metal toxicity and methods that organisms from humans to microbes use to combat the effects of these ions. Our approach to understanding heavy metal interactions has been to use *de novo* designed proteins that predispose thiol residues in positions to chelate metals in normally unusual geometries that turn out to be relevant for heavy metal sequestration in natural proteins. Even though significant effort has been expended to characterize spectroscopically protein complexes with Hg(II), Cd(II), Zn(II), Pb(II) and As(III), the structural detail of these metallated designed peptides has remained elusive. To date, X-ray crystal structures have been obtained for apo-**CoilSer** peptides with cysteine substitution in the *a* (L9C) and *d* (L19C) sites, with As(III) complexed in an *a* site (As(III)(**CS**L9C)$_3$)[1] and Hg(II) in peptides containing penicillamine (Hg(II)$_S$Zn(II)$_N$(**CS**L9PenL23H)$_3^+$)[2]. While being major contributions to both heavy metal biochemistry and protein design, these few available structures have been inadequate to assess important features that control metal geometry within α-helical scaffolds. As more complexity in the designed scaffolds has been introduced, peptides containing an additional heptad have also been developed, but no structures were available to assess the impact on elongating the 3SCC on the overall structure. Thus, I have obtained the high resolution crystal structures of **GRAND-CoilSer** as well as **CoilSer** peptides to investigate the structural details of the designed peptides. As the main focus of this dissertation chapter is to understand the heavy metal binding behavior in

the tris-thiolate metal binding sites, readers who are interested in general crystallographic details of the longer **GRAND-CoilSer** sequences in terms of overall scaffold, crystal packing in lattice form and the interactions on the packing interfaces are referred to the Appendix section. Moreover, I will also describe how this additional heptad influences structure compared to the shorter **CoilSer** derivatives

In the first discussion section of this chapter (*Section I*), I will use the crystallographic information of apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ to investigate the trigonal planar Hg(II) complex within the designed peptide that could be a model for the Hg(II)-bound MerR site.

The knowledge that I have learned from the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ will be evaluated to explain the effect of modifications that have been made to the design around the metal site to control Cd(II) structures in the following section (*Section II*). In this regard, I will compare the new structure with Hg(II) bound to cysteine (Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$) with the previously reported structure containing penicillamine. Beyond this, changes can be made in the layer above the metal site (toward the N-terminus) that influences water access to the site. Because 3-coordinate Hg(II) and 3-coordinate Cd(II) have very similar M-S bond lengths, I can use the Hg(II) structures as analogues for clarifying the chemistry with Cd(II). A model will be presented that explains why Cd(II)(**TRI**L16C)$_3^-$ contains a mixture of Cd(II)S$_3^-$ and Cd(II)S$_3$O$^-$ centers, why Cd(II)(**TRI**L16Pen)$_3^-$ shows exclusively Cd(II)S$_3^-$ structures and lastly, how substitution of alanine for leucine in the layer above the binding site of Cd(II)(**TRI**L12AL16C)$_3^-$ exclusively yields Cd(II)S$_3$O$^-$ using Zn(II)(**GRAND-CS**L12A16C)$_3^-$ and Hg(II)(**GRAND-CS**L12A16C)$_3^-$ for structural clarification.

In *Section III*, I will introduce and structurally characterize high resolution Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ crystal structure to not only describe how a stereochemically active lone pair containing ion would bind into the thiolate site in a trigonal pyramidal fashion, but also provides a reasonable model that could represent the PbrR691 active site, which belongs to the SmtB/ArsR family of metalloregulatory proteins.

Through this point, I have shown how we can utilize the simplified 3SCC derivatives to enforce specific coordination geometries on a metal center, whether to increase it into an unusual, but biological-related higher coordination number of Hg(II), decrease it as in the case of controlling Cd(II) into a trigonal planar structure or even achieve the preferred coordination

number of metals using the simplified tris-thiolate environments for trigonal pyramidal Pb(II) and As(III) centers. However, to further distinguish the different metal structures and how the ligand complexations are constrained upon different metal geometries, the Cys environments in the metallated structures described in the previous sections ($Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$, Zn(II)(**GRAND-CS**L12A16C)$_3^-$, As(III)(**CS**L9C)$_3$[1] and $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$) will be analyzed in comparison with the apo-(**CS**L16C)$_3$ in *Section IV*. The results will allow one to assess the extent of *predisposition* (whether metal binding residues are presented to metals in close proximity, but still require rearrangement to bind) and *preorganization* (cases where metal ligands positioned to be predisposed to bind to a metal, but do not require significant atom movement to bind) associated with metal binding to Cys residues in the 3SCC system. Moreover, the knowledge gained from this work will provide deeper insights of heavy metal binding in the protein environments and hopes to reveal some unappreciated structural features of metal selectivity and specificity in the native metalloregulatory proteins.

Materials and Methods

Materials: F-moc-protected amino acids and the MBHA rink amide resin were purchased from Novabiochem; *N*-hydroxybenzotriazole (HOBt) and 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) were bought from Anaspec Inc.; diisopropy-lethylamine (DIEA), acetic anhydride, and pyridine were purchased from Aldrich; piperidine was supplied by Sigma; and *N*-methylpyrrolidinone (NMP) and *N,N*-dimethylformamide (DMF) were obtained from Fisher Scientific.

Peptide synthesis and purification: All peptide variants were synthesized on an Applied Biosystems 433A automated peptide synthesizer with Fmoc-protected amino acids using the standard Fmoc protocol (Applied Biosystems).[3] The C-terminus of the peptides was amidated on the solid support MBHA rink amide resin (0.25 mmole scale) with HBTU/HOBt/DIEPA coupling methods. The N-terminus was acetylated with a solution of 4% (v/v) acetic anhydride, 4.3% (v/v) pyridine, and 91.7% *N,N*-dimethylformamide (DMF). The peptides were cleaved from the resin using a cleavage mixture of 90% trifluoroacetic acid(TFA), 5% anisole, 3% thioanisole, and 2% ethanedithiol for 3.5 hours. The cleaved peptide solution was filtered and evaporated under a dry N_2-flow until a glassy film appeared on the surface. Cold diethyl ether was then added to the thin

film to obtain a precipitated white crude peptide. This crude was redissolved in ddH_2O and lyophilized to get a fluffy white powder which was subsequently dissolved in 10% acetic acid. The peptide was purified by reversed phase HPLC on a Waters 600 Semiprep HPLC peptide C-18 using a linear gradient of 0.1% TFA in water to 0.1% TFA in 9:1 CH_3CN/H_2O program over 30 mins (flow rate 10 mL/min). The purified peptides were identified by electrospray mass spectrometry. A list of the synthesized peptide sequences in this study is given in **Table 2-1**. Concentration of peptide stock solutions was determined by quantitation of the cysteine thiols using Ellman's test, which uses dithionitrobenzoate (DTNB) as an indicator.[4]

Crystallizations: All peptides were crystallized by sitting drop vapor diffusion experiments at 20 ^oC, with drops containing equal volumes of peptide (0.75 μL) and precipitant (0.75 μL) solutions. The peptide solution of apo-**(CS**L16C)$_3$ was prepared from 20 mg/mL apo-**CS**L16C, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5. The precipitant solution to grow apo-**(CS**L16C)$_3$ crystals contained 40% (v/v) PEG-400, sodium acetate buffer pH 4.5 at a final well solution pH 5.4. The Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ solution was prepared from 20 mg/mL **GRAND-CS**L16CL30H peptide, 0.92 eq of $HgCl_2$ per 3SCC peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5. The well solution contained 15% (v/v) PEG-2000 MME and 0.1 M MES pH 6.5. The crystals of Zn(II)(**GRAND-CS**L12A16C)$_3^-$ were grown from a peptide solution (20 mg/mL **GRAND-CS**L12A16C peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5) and a precipitant solution (0.1 M HEPES pH 7.5 and 40% (v/v) 1,2-propanediol). The Hg(II)(**GRAND-CS**L12A16C)$_3^-$ crystals were prepared from a peptide solution (20 mg/mL **GRAND-CS**L12A16C, 0.92 eq of $HgCl_2$ per 3SCC peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5) and a well solution (0.1 M MES pH 6.5 and 25% (w/v) PEG-1000). Lastly, a Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ solution was prepared from 20 mg/mL **GRAND-CS**L16CL30H peptide, 1.0 eq of $Pb(NO_3)_2$ per 3SCC peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5. The well solution contains 0.1 M MES pH 6.5 and 25% (w/v) PEG-1000. Crystals were cryoprotected in a mother liquor containing 20% glycerol prior to supercooling in liquid N_2 for data collection.

Data collections and refinements: Data were collected at the Advanced Photon Source of the Argonne National Laboratory on the LS-CAT Beamline 21-ID-F and 21-ID-D, equipped with a Mar 225 CCD and a Mar 300 CCD detectors, respectively. All data were collected with a 1 oscillation then processed and scaled with HKL2000[5]. All structures presented in this chapter were

solved by molecular replacement using Molrep[6] in the CCP4 suite of programs[7–9], then underwent iterative rounds of electron density fitting and refining in Coot[10] and Buster 2.11.2 program[11], respectively, unless otherwise noted. The crystal of apo-(**CS**L16C)$_3$ diffracted to 1.42 Å resolution in the C2 (#5) space group. The crystals of **GRAND-CoilSer** derivatives ([Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, Zn(II)(**GRAND-CS**L12A16C)$_3^-$, Hg(II)(**GRAND-CS**L12A16C)$_3^-$ and Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$] diffracted to 2.09, 1.89, 1.93 and 2.17 Å resolution, respectively, in the space group R32. The Crystallography Data Collection and Refinement Statistics of these crystal structures are shown in **Tables 2-2** and **2-3**.

The apo-**(CS**L16C)$_3$ crystallized in the space group C2, containing one three stranded coiled coil per asymmetric unit (ASU) with a Matthew's coefficient of 2.21 corresponding to 44.50% solvent content. The structure was solved using the published apo-(**CS**L9C)$_3$ (PDB code:3LJM)[12] as a search model for Molrep in which Cys 9 and Leu 16 of the search model were mutated to Leu and Cys respectively. It was refined to 1.42 Å ($R_{working}$ = 19.81%, R_{free} = 23.57 %). The validity of the models were verified using the MolProbity software.[13] All the residues are present within allowed regions of the Ramachandran plot.

The peptides in the **GRAND-CoilSer** series crystallized in the R32 space group in which the Matthews' coefficient suggests an estimation of one helix of the **GRAND-CoilSer peptide** per ASU. The combination of the three adjacent asymmetric units that are constrained by the threefold axis, results in a well-folded parallel three stranded coiled coil structure. The Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ has Matthews' coefficient is 2.41 consistent with 49.05% solvent content per ASU and was solved using the five heptads of **GRAND-CS**L12$_D$L16C (obtained in chapter 3) as a search model. The side chains of the model were included in which the D-Leu residue at the twelfth position was mutated to L-Leu. Electron difference density maps (F_o-F_c) contoured at 3σ showed two possible metal sites at the sixteenth (16Cys) and the thirtieth (30His) positions corresponding to Hg(II) and Zn(II) ions, respectively. These positions were confirmed by solving the structure by single anomalous dispersion (SAD) using AutoSol in Phenix. [9,14–16] The Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ structure was refined to 2.09 Å ($R_{working}$ = 21.76 %, R_{free} = 23.04 %) in Buster program.

The single strand in an ASU of Zn(II)(**GRAND-CS**L12A16C)$_3^-$ was solved using the **GRAND-CS**L12$_D$L16C model as a search model in Molrep. The top solution result was then used as a starting model in AutoBuild Wizard software (Phenix)[17–20] to estimate iteratively experimental

crystallographic phases and improve the model-based map correlations. During the cycles, the existing model was chopped into pieces and then was rebuilt from the remaining ends based on statistical density distributions. These processes were repeated several times until the overlapping segments covered the entire model. After that, the AutoBuild-solved solution was refined. The solvent content per ASU of this structure is 48.09 % (Matthew's coefficient of 2.37). The structure was refined to 1.89 Å ($R_{working}$ = 19.82%, R_{free} = 23.51 %).

The helix in the ASU of Hg(II)(**GRAND-CS**L12A16C)$_3^-$ was solved by a **GRAND-CS**L12A16C model from the Zn(II)(**GRAND-CS**L12A16C)$_3^-$. The Matthew's coefficient appears to be 2.43 (49.45% solvent content per ASU)). The structure was refined to 1.93 Å ($R_{working}$ = 20.14%, R_{free} = 25.15 %).

The Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ was solved using a single helix of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ as a search model in Molrep. The Matthew's coefficient is 2.47 corresponding to 50.24% per ASU. The structure was refined to 2.17 Å ($R_{working}$ = 19.84%, R_{free} = 24.09 %).

Table 2-1. Peptide sequences

Peptides		**a b c d e f g** 2	**a b c d e f g** 9 12	**a b c d e f g** 16 19	**a b c d e f g** 23	**a b c d e f g** 30
TRI	Ac-G	LKALEEK	LKALEEK	LKALEEK	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	**Pen**KALEEK	LKALEEK	G-NH_2
CoilSer (CS)	Ac-E	WEALEKK	LAALESK	LQALEKK	LEALEHG	-NH_2
CSL9C	Ac-E	WEALEKK	**C**AALESK	LQALEKK	LEALEHG	-NH_2
CSL9CL23H	Ac-E	WEALEKK	**C**AALESK	LQALEKK	**H**EALEHG	-NH_2
CSL16C	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LEALEHG	-NH_2
CSL16Pen	Ac-E	WEALEKK	LAALESK	**Pen**QALEKK	LEALEHG	-NH_2
GRAND-CoilSer	Ac-E	WEALEKK	LKALESK	LQALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL16CL30H	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LQALEKK	**H**EALEHG -NH_2
GRAND-CSL12AL16C	Ac-E	WEALEKK	LAA**A**ESK	**C**QALEKK	LQALEKK	LEALEHG -NH_2

Bold and underlined residues indicate substitutions.
C-and N-termini are capped by Ac and NH_2 groups, respectively.

Table 2-2: Crystallography Data Collection and Refinement Statistics

Peptide	apo-(**CS**L16C)$_3$	Hg(II)$_S$Zn(II)$_N$ (**GRANDCS**L16CL30H)$_3^+$	Pb(II)$_S$Zn(II)$_N$ (**GRANDCS**L16CL30H)$_3^+$
Data Collection			
SpaceGroup	C2	R32	R32
Unit Cell a, b, c (Å)	76.119, 28.977, 43.917	38.048, 38.048 143.651	38.199, 38.199, 144.522
Wavelength (Å)	0.97872	0.97856	0.97856
Resolution (Å)[1]	1.42(1.42-1.40)	2.09 (2.13-2.09)	2.17 (2.20-2.17)
Rsym (%)[2]	6.1 (45.7)	9.4 (48.3)	9.1(52.2)
<I/sI>[3]	>50 (2)	>50 (2)	>50 (2)
Completeness (%)[4]	98.2 (97.4)	99.9 (100)	99.6(100)
Redundancy	5.4 (5.0)	18.2 (18.2)	17.6 (18.1)
Refinement			
Resolution (Å)	1.42	2.09	2.17
R-Factor (%)[5]	19.8	21.7	19.8
Rfree (%)[6]	23.6	23.0	24.1
Protein atoms	1549	291	276
Metal ions	3 Zn(II)	1/3 Hg(II) 1/3 Zn(II) 1 Zn(II) on surface	1/3 Pb(II) 1/3 Zn(II) 1 Zn(II) on surface
Water Molecules	52	11	24
Unique Reflections	16518	2584	2421
R.m.s.d.[7]			
Bonds	0.010	0.010	0.010
Angles	1.06	1.15	1.24
MolProbity Score[8]	1.16	1.51	1.30
Clash Score[8]	3.39	4.40	3.83

[1]Statistics for highest resolution bin of reflections in parentheses.

[2]$R_{sym} = \Sigma_h \Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $\langle I_h \rangle$ is the mean intensity for multiply recorded reflections.

[3]Intensity signal-to-noise ratio.

[4]Completeness of the unique diffraction data.

[5]R-factor = $\Sigma_h | IF_oI - IF_cI | / \Sigma_h |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.

[6]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement.

[7]Root mean square deviation of bond lengths and bond angles.

[8]Chen et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica D66:12-21.

Table 2-3: Crystallography Data Collection and Refinement Statistics

Peptide	Zn(II)(**GRAND-CS**L12A16C)$_3^-$	Hg(II)(**GRAND-CS**L12A16C)$_3^-$
Data Collection		
SpaceGroup	R32	R32
Unit Cell a, b, c (Å)	37.857, 37.857, 141.076	38.186, 38.186, 142.385
Wavelength (Å)	0.97856	0.97856
Resolution (Å)[1]	1.89 (1.92-1.89)	1.84 (1.87-1.84)
Rsym (%)[2]	6.5 (45.5)	6.9 (54.4)
<I/sI>[3]	> 50 (2)	> 50 (2)
Completeness (%)[4]	99.9 (100)	98.6 (100)
Redundancy	12.8 (12.0)	15.8 (12.4)
Refinement		
Resolution (Å)	1.89	1.92
R-Factor (%)[5]	19.8	20.1
Rfree (%)[6]	23.2	25.1
Protein atoms	331	318
Metal ions	1/3 Zn(II)	1/3 Zn(II)
	1 Zn(II) on surface	1 Zn(II) on surface
Water Molecules	36	29
Unique Reflections	3341	3266
R.m.s.d.[7]		
Bonds	0.003	0.006
Angles	0.581	0.685
MolProbity Score[8]	1.30	1.25
Clash Score[8]	3.83	3.09

[1]Statistics for highest resolution bin of reflections in parentheses.

[2]$R_{sym} = \Sigma_h \Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h \Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $\langle I_h \rangle$ is the mean intensity for multiply recorded reflections.

[3]Intensity signal-to-noise ratio.

[4]Completeness of the unique diffraction data.

[5]R-factor = $\Sigma_h | IF_oI - IF_cI | / \Sigma_h |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.

[6]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement.

[7]Root mean square deviation of bond lengths and bond angles.

[8]Chen et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica D66:12-21.

Results

Structural Implication of the systems imposed by different space groups

The vast majority of the crystals examined in this dissertation crystallized in the space groups R32 and C2. The basic crystallographic difference between these two lattices is that R32 has crystallographically imposed threefold symmetry along the *z* axis that runs through the center of the three helices of the 3SCC. This symmetry restriction requires the three helices to be identical. The C2 space group, on the other hand, does not have this requirement. Thus, the three helices are independent. The consequence of symmetry is that structures in R32 will have a single reported value for the following crystallographic parameters ($\chi 1$ dihedral angles, atomic distances, interlayer distances, M-S distances, etc.) while three values are usually given for structures crystallizing in C2. These data may be found in **Tables 2-4** and **2-5**.

First coordination sphere: Cysteine Sites at the sixteenth position

*apo-(**CS**L16C)$_3$*

The side chains of the Cys residues at the sixteenth position in chain A and C display two alternate conformations in which the major conformers, at 60% occupancy, are oriented to the coiled coil interior and pointing toward the N-termini. In contrast, the minor conformers (28% occupancy) are directed toward the helical interface and toward the C-termini (**Figure 2-1**). The third thiol comes from chain B that contains only one conformer, which is orientated toward the core of the structure. Only the major 16CysA and 16CysC conformers combined with the 16CysB are reasonable to pre-form a metal binding site because their S_γ atoms are positioned toward the center, resembling the *a* site Cys orientation reported in the published apo-**(CS**L9C)$_3$ structure[12] and the Coil V_aL_d crystal structure.[21] The average separation between the S_γ atoms in this arrangement is 3.29 Å which is too long to represent disulfide formation. The average χ_1 dihedral angle (N-Cα-Cβ-Sγ) for the inner Cys orientation is -69.20^o, conforming to the most common χ_1 value for cysteine rotamers (-65.20^o).[22] Due to the lack of F_o-F_c electron difference at the binding site of the apo-**(CS**L16C**)$_3$**, the structure has been assigned as an apo-structure.

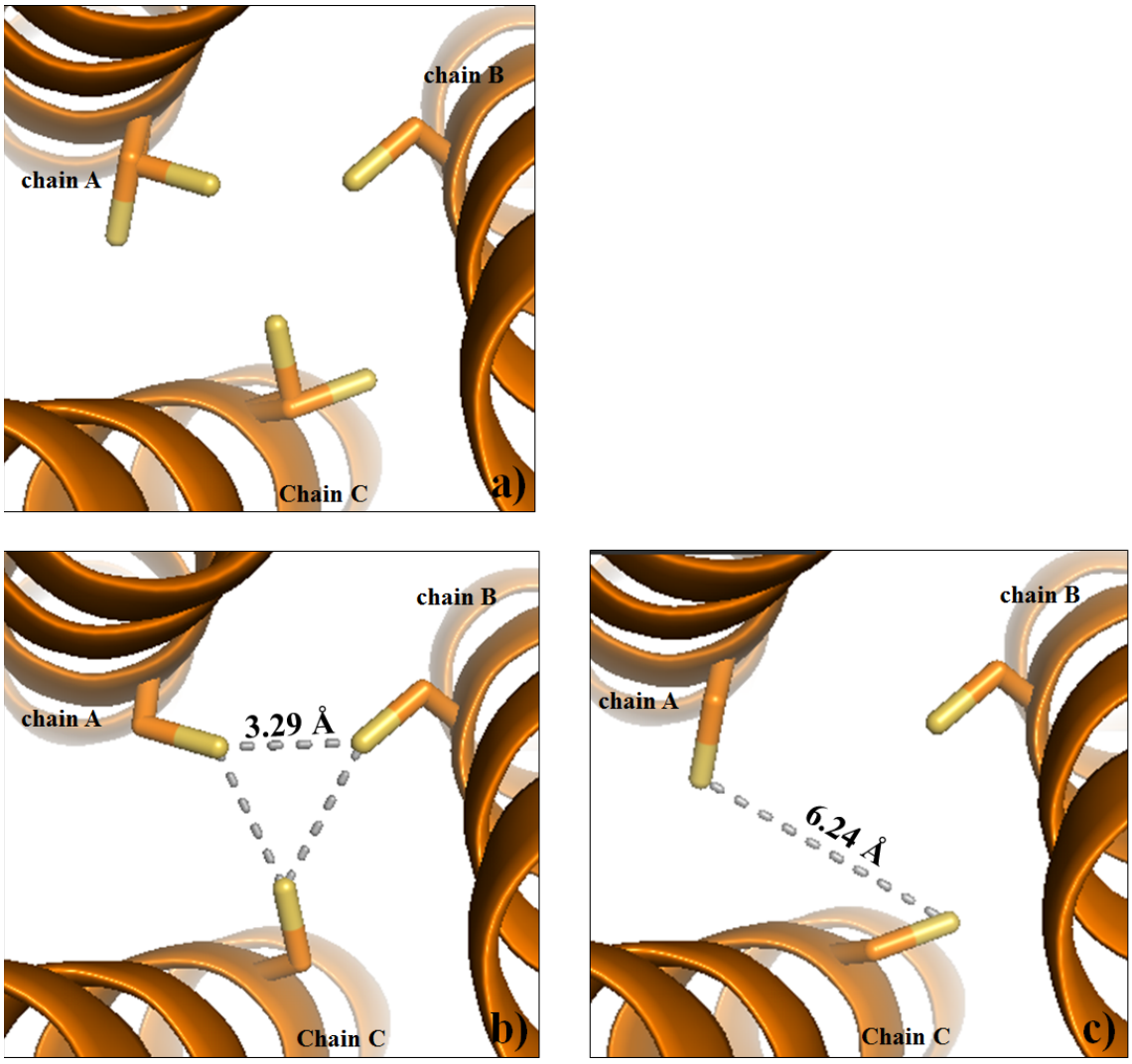

Figure 2-1. PyMOL representation showing the orientations of Cys side chains in the sixteenth layer of apo-(CSL16C)$_3$. The Cys side chains are shown as orange sticks with the thiol groups labeled in yellow. Shown in a Top-down view from the N-termini represents a) the major and minor conformers of Cys residues, b) the major conformers and c) the minor conformers.

Table 2-4: Crystallographic evidence observed from the crystal structures

Peptides	apo-(**CS**L16C)$_3$	Hg(II)$_S$Zn(II)$_N$ (**GRAND-CS**L16CL30H)$_3^+$	Pb(II)$_S$Zn(II)$_N$ (**GRAND-CS**L16CL30H)$_3^+$	Zn(II)(**GRAND-CS**L12A16C)$_3^-$	Hg(II)(**GRAND-CS**L12A16C)$_3^-$
16Cys rotamers					
χ_1 (interior rotamers)[a]	-69.20° (average)	-151.88°	-68.38°	-152.72°	-147.85°
S_γ-S_γ distance[b] (Å)	3.29 (average)	4.16	3.49	3.94	4.20
χ_1 (exterior rotamers)	174.45°,176.96°[j]	173.35°	-	172.42°	-52.83°
S_γ-S_γ distance[c] (Å)	5.36(average)	5.24	-	5.94	3.52
Metal site					
M-S bond length (Å)	-	2.43, Hg(II)-S	2.65, Pb(II)-S	2.27, Zn(II)-S 2.17, Zn(II)-O	2.44, Hg(II)-S
S-M-S angle(average)	-	118.36°	82.37°	119.84°, S-Zn(II)-S 92.13°, S-Zn(II)-O	118.74°
Distance of metal from to the bound Cys plane[d] (Å)	-	-0.3	-1.69	+0.09	-0.26
12Leu rotamers					
χ_1 (chain A, B, C)[e]	-81.68°, -75.68°,-171.24°	-124.20°	179.88°	-	-
χ_2 (chain A, B, C)[f]	-171.45°, 171.07°, 54.60°	-172.02°	57.61°	-	-
Interior C_δ separation[g] (Å)	4.39	4.30	3.83	-	-
Exterior C_δ separation[h] (Å)	6.61	5.57	6.24	-	-
Distance of the layer from the sulfur plane (Å)[i]	5.08	6.07	4.31	-	-
19Leu rotamers					
χ_1 (chain A, B, C)	-78.93°, -65.99°,-62.92°	-69.49°	-69.31°	-71.28°	-72.79°
χ_2 (chain A, B, C)	168.54°, 176.10°,175.32°	-172.87°	-174.92°	178.02°	154.35°
Interior C_δ separation (Å)	4.71	5.95	6.17	5.89	6.64
Exterior C_δ separation (Å)	7.69	8.98	9.13	8.42	9.27
Distance of the layer from the sulfur plane (Å)	4.49	3.49	4.67	3.65	3.49

[a]χ_1 of Cys residue is determined from the dihedral angle of N-C_α-C_β-S_γ; [b]Distance determined between S_γ atoms of the interior Cys conformers of two adjacent chains; [c]Distance determined between S_γ atoms of the exterior Cys conformers of two adjacent chains.; [d]Plus sign (+) indicates the metal is situated above the bound Cys plane; Minus (-) indicates the metal is situated below the bound Cys plane.; [e]χ_1 of Leu residue is determined from the dihedral angle of N-C_α-C_β-C_γ.; [f]χ_2 of Leu residue is the dihedral angle measured from C_α-C_β-C_γ-$C_{\delta 1}$.; [g]Interior C_δ separation defines the average distance between the interior C_δ atoms of Leu residues of two adjacent chains.; [h]Exterior C_δ separation defines the average distance between the exterior C_δ atoms of Leu residues of two adjacent chains.; [I] Distance determined between the interior C_δ (Leu) layer and the inner S_γ (Cys) layer and [j]χ_1 dihedral angles measured from chain A and C, respectively.

Table 2-5: Crystallographic evidence observed from the previous published crystal structures

Peptides	apo-(**CS**L9C)$_3$ (PDB code: 3LJM)[12]	apo-(**CS**L16Pen)$_3$ (PDB code: 3H5F)[31]	[Hg(II)]$_5$[Zn(II) (H_2O/OH^-)]$_N$ (**CS**L9PenL23H)$_3^{n+}$ (PDB code: 3PBJ)[2]	As(III)(**CS**L9C)$_3$ (PDB code: 2JGO)[1]
Metal ligand rotamers				
χ_1 (interior rotamers)[a]	-70.57° (average)	-49.85° (average)	-50.23° (average)	-59.66°(average)
S_γ -S_γ distance[b] (Å)	3.22 (average)	3.71(average)	3.84 (average)	3.25(average)
χ_1 (exterior rotamers)	169.98°(average)	72.96°(average)	-	-
S_γ -S_γ distance[c] (Å)	6.85(average)	8.36(average)	-	-
Metal site				
M-S bond length (Å)	-	-	2.21, Hg(II)-S(average)	2.28, As(III)-S
S-M-S angle(average)	-	-	119.92° (average)	90.0°, S-As(III)-S
Distance of metal from to the bound Cys plane[d] (Å)	-	-	-0.06	-1.34
Leu(above) rotamers		(12Leu)	(5Leu)	(5Leu)
χ_1 (chain A, B, C)[e]	-	-82.66°,-79.23°,-104.59°	-174.14°,-179.20°,-172.74°	-176.72°,-147.90°,-167.01°
χ_2 (chain A, B, C)[f]	-	171.06°, 165.77°,175.34°	52.10°, 78.26°,50.94°	59.51°,-151.49°, 59.33°
Interior C_δ separation[g] (Å)	-	4.94	3.60	3.86
Exterior C_δ separation[h] (Å)	-	6.74	5.84	6.33
Distance of the interior C_δ layer from the sulfur plane[i] (Å)	-	5.76	4.86	4.42
Leu(below) rotamers		(19Leu)	(12Leu)	(12Leu)
χ_1 (chain A, B, C)	-	-128.41°, -80.90°, -65.88°	-128.41°, -80.90°, -65.88°	-64.76°, -61.39°, -68.93°
χ_2 (chain A, B, C)	-	57.85°,-158.84°,176.66°	57.85°,-158.84°,176.66°	169.98°,-158.84°,172.32°
Interior C_δ separation (Å)	-	5.86	5.28	5.26
Exterior C_δ separation (Å)	-	8.66	8.45	7.64
Distance of the interior C_δ layer from the sulfur plane (Å)	-	4.77	4.49	4.66

[a]χ_1 of Cys residue is determined from the dihedral angle of N-C_α-C_β-S_γ; [b]Distance determined between S_γ atoms of the interior Cys conformers of two adjacent chains; [c]Distance determined between S_γ atoms of the exterior Cys conformers of two adjacent chains.; [d]Plus sign (+) indicates the metal is situated above the bound Cys plane; Minus (-) indicates the metal is situated below the bound Cys plane.; [e]χ_1 of Leu residue is determined from the dihedral angle of N-C_α-C_β-C_γ.; [f]χ_2 of Leu residue is the dihedral angle measured from C_α-C_β-C_γ-$C_{\delta 1}$.; [g]Interior C_δ separation defines the average distance between the interior C_δ atoms of Leu residues of two adjacent chains.; [h]Exterior C_δ separation defines the average distance between the exterior C_δ atoms of Leu residues of two adjacent chains. and [I] Distance determined between the interior C_δ (Leu) layer and the inner S_γ (Cys) layer.

*$Hg(II)_S Zn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$*

Constrained by the three fold axis of the R32 space group, the three helices of $Hg(II)_S Zn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ display perfect three-fold symmetry. Difference electron density peak observed on the three fold axis at the sixteenth position is assigned to be a bound Hg(II) with 51% metal occupancy and at the thirtieth position represents a coordinated Zn(II) (100% occupancy) (**Figure 2-2**). The Hg(II) ion is situated in a trigonal plane of three Cys residues, each of which contains two alternate conformations where only the interior conformer accounts for metal coordination (**Figure 2-3**). The C_β atom of this Cys residue is directed toward the N-terminus of the helical scaffold, whereas the S_γ atoms in both conformations are pointing down toward the C-terminus. The inner S_γ conformer (occupancy =54%) is pointing toward the helical interface with a χ_1 side chain torsion angle of -151.88°. The S_γ-S_γ separation is 4.16 Å. The outer S_γ conformer (occupancy = 34%) is oriented even more toward the helical interface with χ_1 angle of -173.35°, leading to a longer S_γ-S_γ separation of 5.24 Å. The Hg(II)-S distance of 2.43 Å (**Figure 2-4**) is in excellent agreement with the 2.44 Å determined from EXAFS as reported in **Table 2-6**. These bond distances also correspond to the Hg(II) to sulfur bond length of a tris-thiolato trigonal planar Hg(II) in native Hg(II)-MerR (EXAFS)[23,24] and small modeled $Hg(II)S_3^-$ inorganic complexes (EXAFS, solid state X-ray crystallography).[24–26] The S-Hg(II)-S angle appears to be 118.36°.

*Zn(II)(**GRAND-CS**L12A16C)$_3^-$*

The binding site of the zinc ion in Zn(II)(**GRAND-CS**L12A16C)$_3^-$ is made of three symmetric Cys ligands which contain two conformations**.** In general, both of them look oriented toward the helical interface with their S_γ atoms pointing toward the C-termini (C_β atoms directs to the opposite ends) (**Figure 2-5)**. Compared with respect to each other, the Zn(II)-bound conformer (with 91% occupancy) is relatively inward with the χ_1 dihedral angle of -152.72°, while the unbound conformer (with 9% occupancy) is more outward with the χ_1 angle of 172.42°. The S_γ-S_γ separations of the bound versus unbound Cys conformers are 3.94 and 5.92 Å,

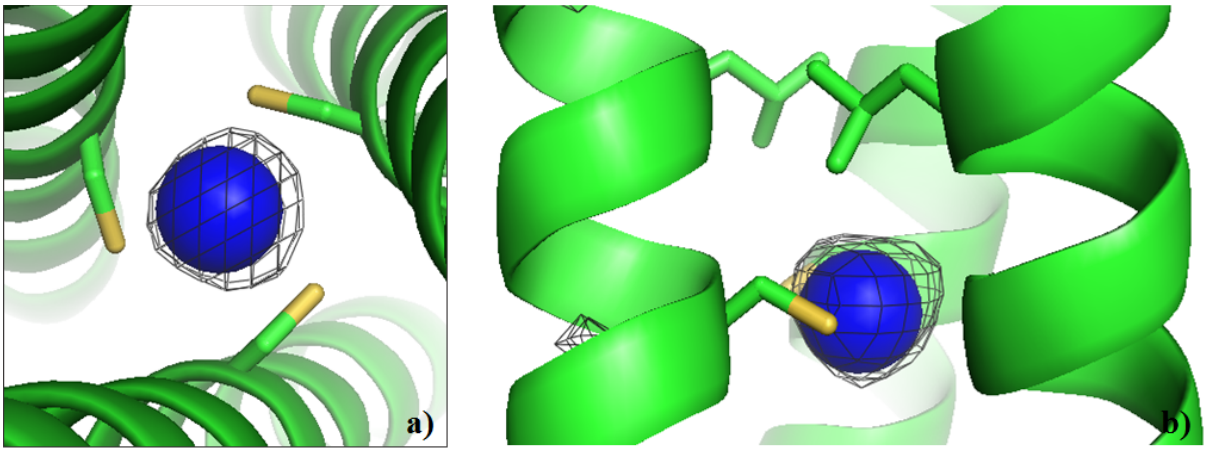

Figure 2-2. Omit map for the trigonal planar coordination of Hg(II)S_3^- site in Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structure. Shown is the ribbon diagram (green) representing the main chains of three helices in which the Cys residues are in stick form (sulfur = yellow). The Hg(II) atom is colored as blue. Electron density maps were generated through refinement with the metal occupancies set to zero. F_o-F_c electron density is shown for Hg(II) contoured at 3σ and colored grey. In a) Top-down view and b) side view of the trigonal Hg(II)S_3^- site.

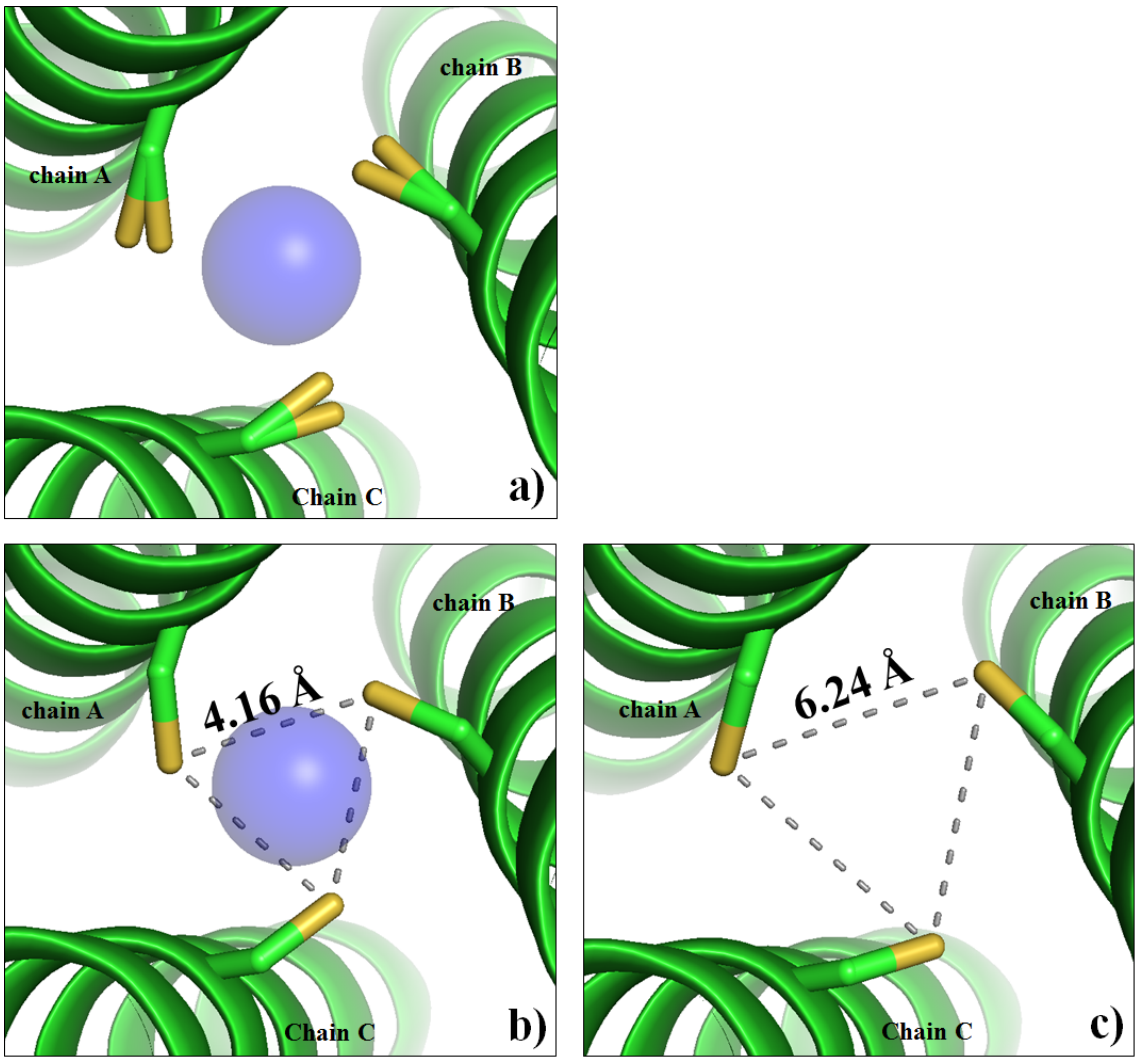

Figure 2-3. Ribbon diagrams of $Hg(II)_S Zn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$ showing the orientation of Cys side chains. Main chain atoms are shown as green ribbon and the Cys side chains as green stick with the thiol groups being labeled as yellow. Top-down view from N-termini showing the exclusive orientations of a) all conformers, b) major conformers and c) minor conformers (unbound).

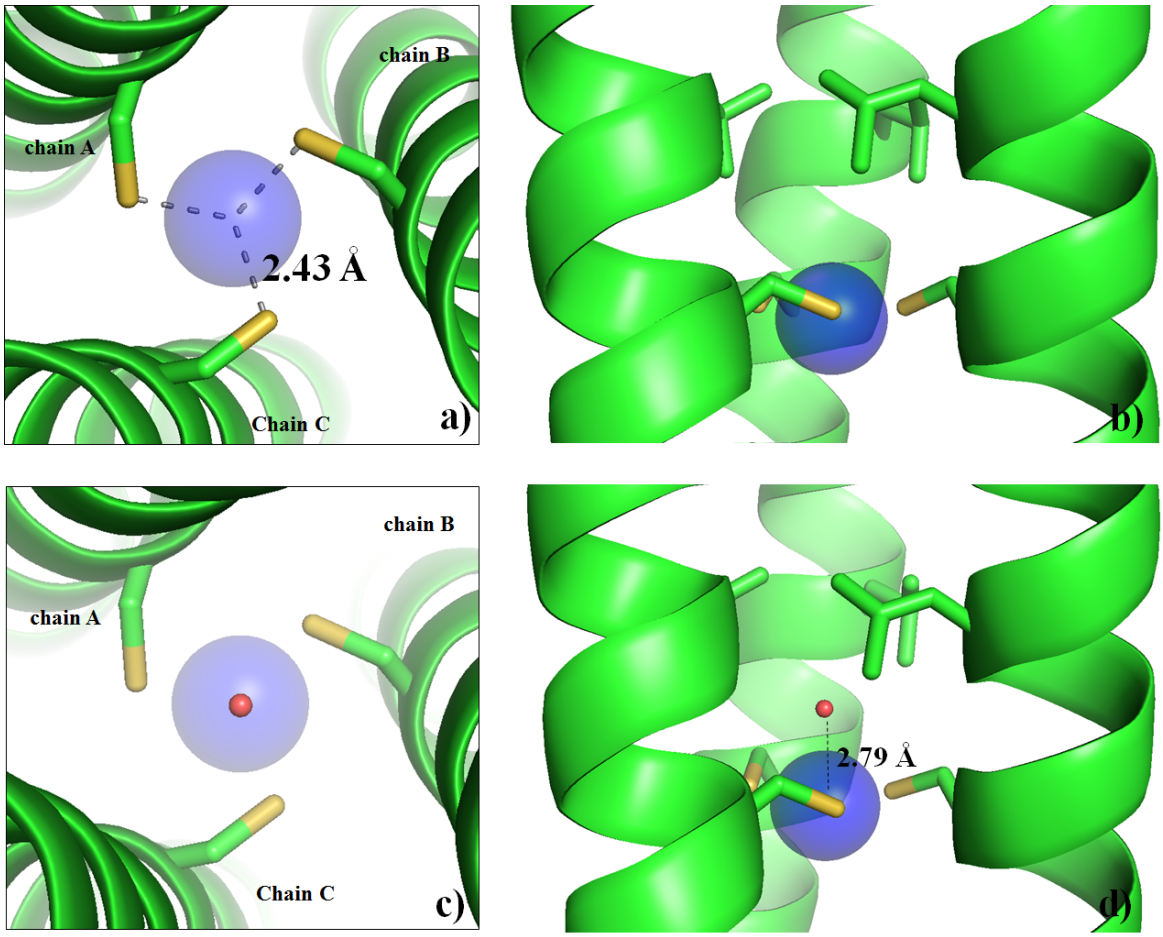

Figure 2-4. PyMOL visualization of the trigonal planar Hg(II)S_3^- in Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structure. Shown in a) top down and b) side view of the first coordination sphere of the binding site. c) and d) representing the uncoordinated water observed above the 16Cys plane. Main chain helices are shown as green ribbons, Cys side chains in the sixteenth position are in stick form (sulfur = yellow), Hg(II) ion is shown as a blue sphere and the uncoordinated water is represented as a small red sphere.

Table 2-6. Spectroscopic data for peptides

Peptide	Cd(II) species	Cd(II)-S bond length (Å)	δ (ppm) (^{113}Cd NMR)	ω_0 (rad/ns), η (^{111m}Cd PAC)
Cd(II)(**TRI**L16C)$_3^-$	mixtures between 3- and 4-coordinate	2.49[a]	625[a]	0.337±0.002[a] 0.438±0.004
Cd(II)(**TRI**L12A16C)$_3^-$	pure 4-coordinate	2.50[c]	574[b]	0.3405±0.0003[b]
Cd(II)(**TRI**L16Pen)$_3^-$	pure 3-coordinate	2.46[d]	684[b]	0.454±0.001[b]

[a] Reference 27 ; [b] Reference 28 and 29 ; [c] Data obtained from these studies and [d] Data for Cd(**GRAND**L16Pen)$_3^-$ obtained in these studies. EXAFS data were analyzed by Dr. Aniruddha Deb, in collaboration with Prof. James Penner Hahn.

respectively. **Figure 2-6** shows a 3σ-contoured F_O-F_C electron density map in the center of the Cys plane (located on the three fold symmetry axis) indicating a metal difference electron density which is consistent with the scattering power of Zn(II). This Zn(II) is in a distorted 4-coordinate structure with three Cys ligands (interior conformers) and a water molecule (**Figure 2-7**). This water is also sitting on the three-fold symmetry axis and it is 2.17 Å directly above the Zn(II) ion. The Zn(II)-S distance appears to be 2.27 Å. The S-Zn(II)-S and S-Zn(II)-O angles are 119.84° and 92.13°, respectively. Moreover, there are three water molecules observed above the metal plane and appear to be H-bonded to the coordinated water. These three waters are crystallographically imposed and at distances of 3.65 and 2.43 Å from the Zn(II) center and the bound water, respectively.

*Hg(II)(**GRAND-CS**L12A16C)$_3^-$*

The Cys side chains at the sixteenth position are symmetrically equivalent in all three helices and contain two conformations each (**Figure 2-8**). The major conformer (with 88% occupancy) is pointing toward the C-termini and outward to the helical interface with the χ_1 dihedral angle of -147.85°, while the minor conformer (with 12% occupancy) is oriented at the helical core and upward to the N-termini with the χ_1 angle of -52.83°. The $S\gamma$-$S\gamma$ separation between two major Cys residues is 4.20 Å and it is 3.52 Å for the minor Cys. In fact, only the major conformers are reasonable to bind to Hg(II) at the center of the 16Cys plane as the Hg(II) to $S\gamma$ (major) distance of 2.44 Å (**Figure 2-9, top panels**) is in excellent agreement with the Hg(II)-S bond distance in a trigonal planar $Hg(II)S_3^-$ of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure (while Hg(II)-$S\gamma$(minor) distance is 2.90 Å). The S-Hg(II)-S angle is 118.74°.

*$Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$*

Like the other **GRAND-CoilSer** variants, the $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ is constrained by the crystallographic three-fold symmetry of the R32 space group. The Cys residue points the $C\beta$ and $S\gamma$ atoms at the core of the helices and also toward the N-terminus **(Figure 2-10)**. The $S\gamma$-$S\gamma$ separation is 3.49 Å. The side chain χ_1 angle is -68.38°. The difference electron

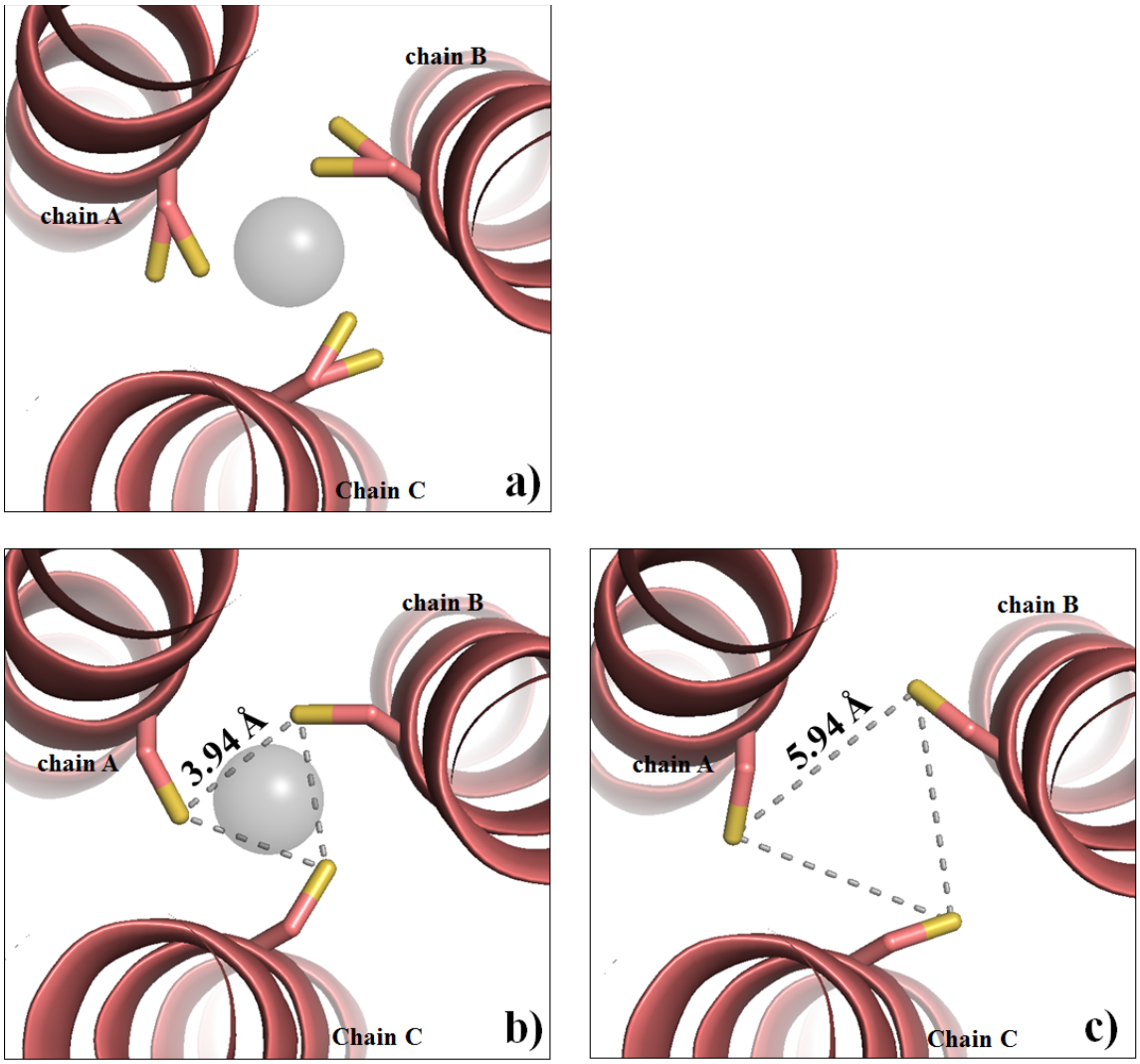

Figure 2-5. PyMOL representation showing the orientations of Cys side chains in the sixteenth layer of Zn(II)(GRAND-CSL12AL16C)$_3^-$ structure. The Cys side chains are shown as light red sticks with the thiol groups labeled in yellow. The Zn(II) ion is in present as a grey sphere. Shown in a Top-down view from the N-termini represents a) both major and minor conformers of Cys residues, b) the major conformers (bound rotamers) and c) the minor conformers.

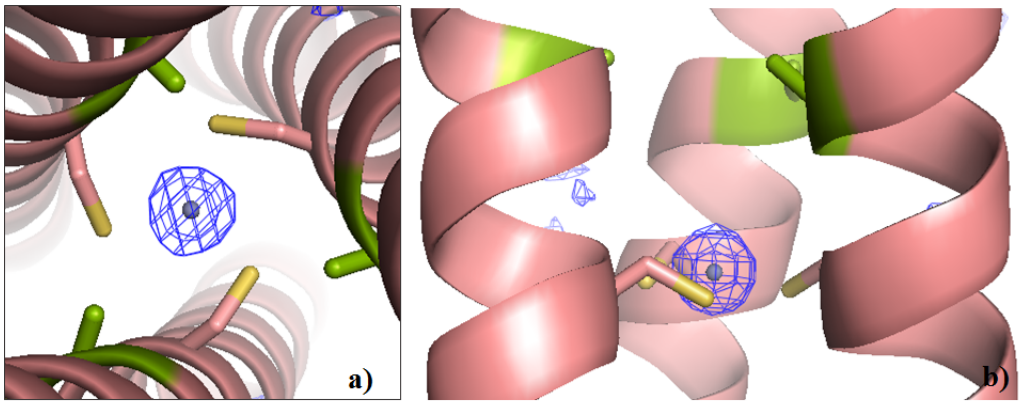

Figure 2-6. Omit map for the trigonal planar coordination of $Zn(II)S_3(H_2O)^-$ in $Zn(II)(GRAND\text{-}CSL12AL16C)_3^-$ structure. Shown is the ribbon diagram (pink) representing the main chains of three helices in which the Cys residues are in stick form (sulfur = yellow). The Zn(II) atom is colored as grey. Electron density maps were generated through refinement with the metal occupancies set to zero. F_O-F_C electron density is shown for Zn(II) contoured at 3σ and colored blue. In a) Top-down view and b) side view of the 4-coordinate $Zn(II)S_3(H_2O)^-$ site.

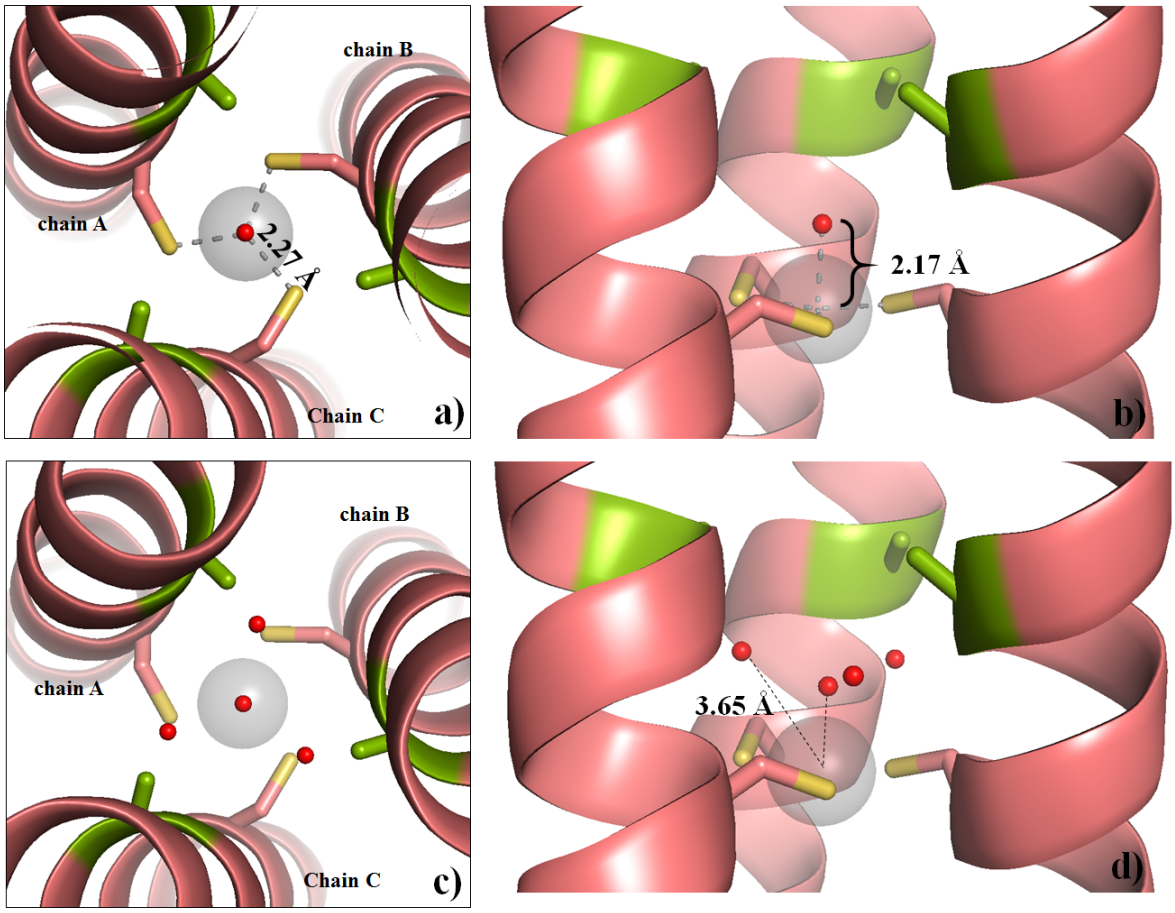

Figure 2-7. PyMOL visualization of the pseudo-tetrahedral $Zn(II)S_3(H_2O)^-$ in $Zn(II)(GRAND\text{-}CSL12AL16C)_3^-$ structure. Top panels: a) Top down view from the N-termini and b) side view of the first coordination of the 4-coordinate binding site. (NOTE: The other three uncoordinated waters observed above the binding site are omitted for clarity.) c) and d) represents the environment above site when those waters are included. Main chain helices are shown as light red ribbons, Cys side chains in the sixteenth position are present as sticks (sulfur = yellow). Ala residues at the twelfth position are colored in green. Zn(II) ion is shown as a grey sphere and the coordinated water is represented as a small red sphere.

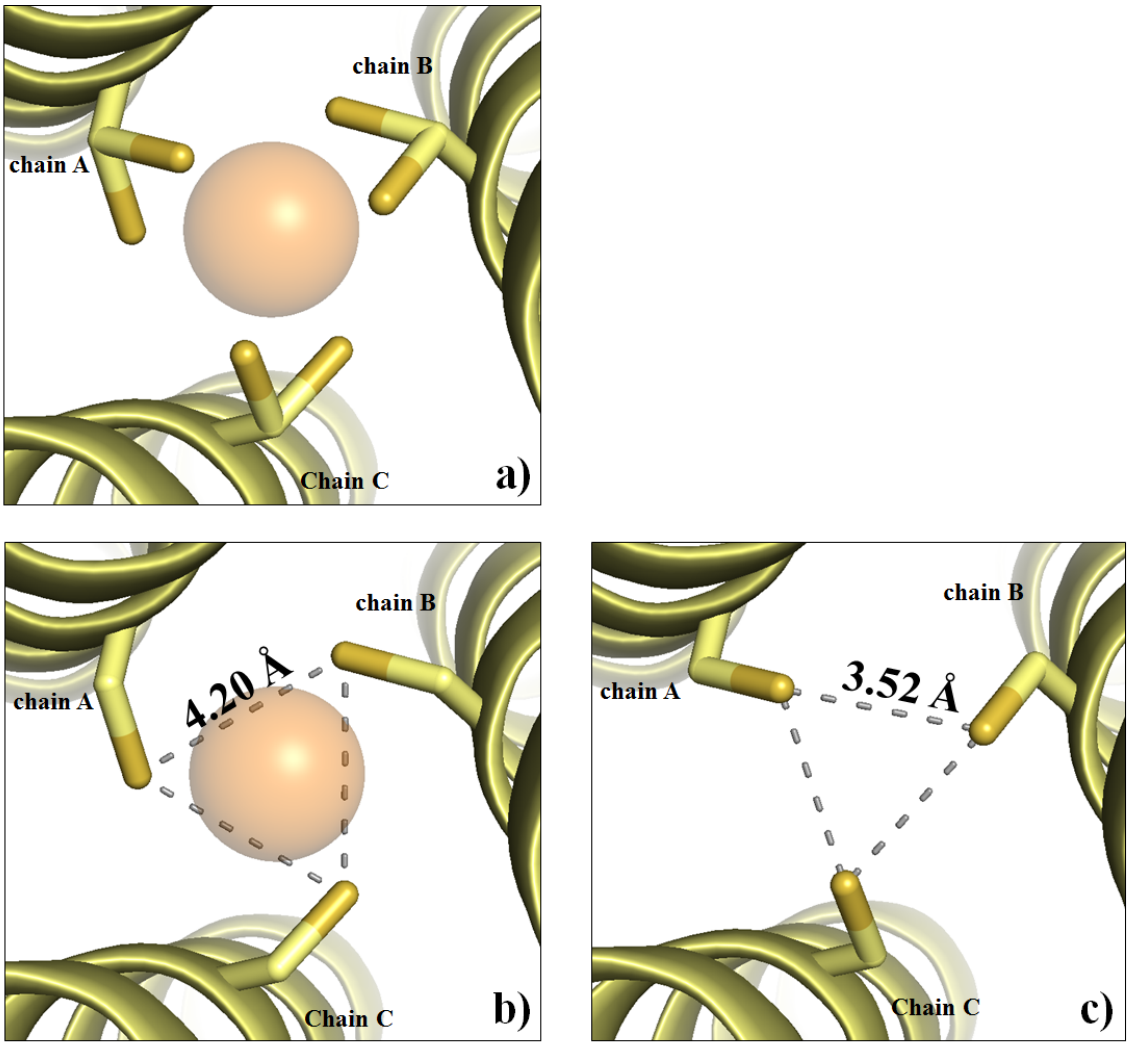

Figure 2-8. Cys side chain arrangements in the sixteenth position of Hg(II)(GRAND-CSL12AL16C)$_3^-$. The Cys side chains are shown as pale yellow sticks in which the thiol groups are labeled in bright yellow. Shown from a top-down view of the structure represents a) both alternate conformers of Cys residues, b) the major conformers. c) the minor (unbound) conformers.

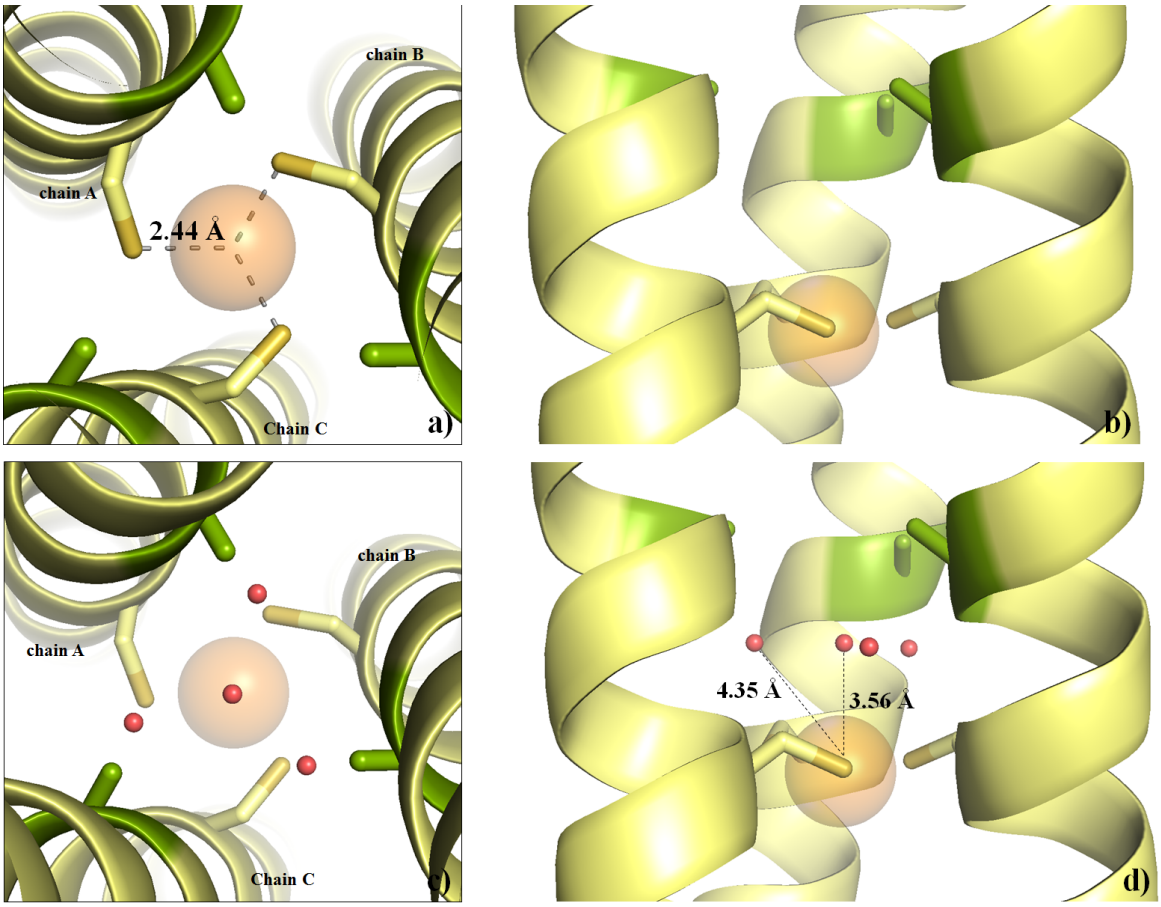

Figure 2-9. PyMOL visualization of the trigonal Hg(II)S_3^- in Hg(II)(GRAND-CSL12AL16C)$_3^-$ structure. a) Top down view from the N-termini and b) side view of the trigonal Hg(II) binding site. (NOTE: The other three uncoordinated waters observed above the binding site are omitted for clarity.) c) and d) represents the environment above site when the other three waters are included. Main chain helices are shown as pale yellow ribbon diagrams, Cys side chains in the sixteenth position are present as sticks (sulfur = bright yellow). Ala residues at the twelfth position are colored in green. The Hg(II) ion is shown as an orange sphere.

density (F_o-F_c) contoured at 3σ from the beginning round of refinements confirms that a coordinated Pb(II) is bound in a trigonal pyramidal geometry at a distance of 1.69 Å below the tris-thiolate plane with respect to the N-termini. The Pb(II)-S distance is 2.65 Å in good agreement with a previous determination using EXAFS.[27] The S-Pb(II)-S angle is 82.37°.

Second coordination sphere: the twelfth and nineteenth positions

*apo-(**CS**L16C)$_3$*

The packing of the Leu layer at the twelfth position is shown in **Figure 2-11a**. The side chains of these residues orient toward the helical interface by having the Cβ atoms pointed toward the N-termini. Only one of the two δ-methyl groups of the Leu residues are pointed toward the core of the structure, whereas the other atoms point toward the outer face. The size of hydrophobic cavity above the metal site is determined from the distance between the plane made by the core Cδ atoms of 12Leu to the interior Cys conformer plane, which is ~5.08 Å (**Figure 2-12a**). The 19Leu side chains are oriented even more toward the helical interface compared to the 12Leu (**Figure 2-11b**). The hydrophobic interlayer space below the metal site is ~4.49 Å determined from the separation between the interior Cys conformer plane to the plane of the core Cδ atoms of 19Leu.

*Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$*

Examination of the second coordination sphere of metals bound to these 3SCCs provides additional insight into the processes that control metal coordination number and geometry. The side chain of 12Leu is oriented away from the center of the coiled coil as illustrated in **Figure 2-13**. Only one of the Cδ atoms is directed to the core while the other is pointing outward to the helical interface. The distance between the core Cδ atoms of 12Leu to the major Cys plane is 6.07 Å, which defines the size of the hydrophobic cavity above the metal site (**Figure 2-12b**). In the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ structure, there is a water sitting on the three-fold axis, 2.79 Å above the Hg(II) ion. This particular water is situated in a hydrophobic cavity between the twelfth and sixteenth positions. The water is 4.35 Å away from the interior Cδ atom plane,

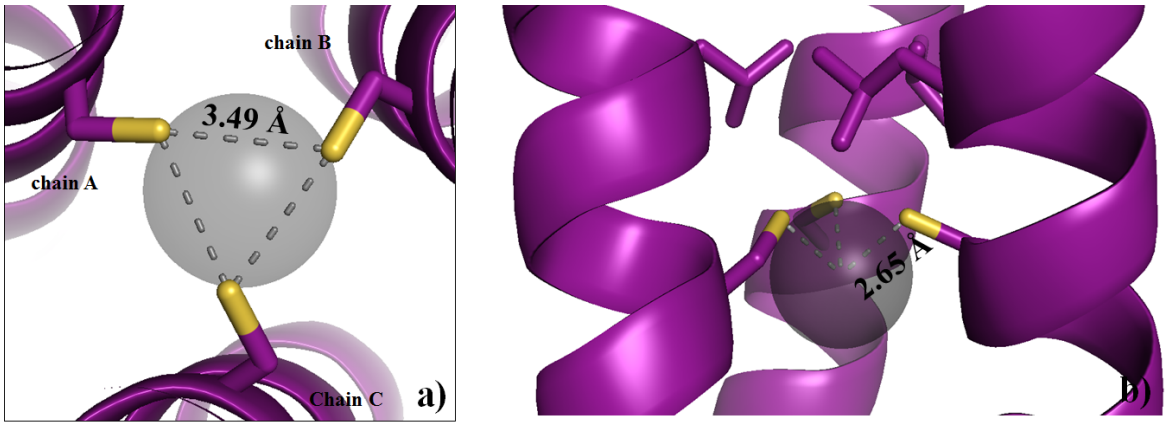

Figure 2-10. PyMOL visualization of the trigonal pyramidal Pb(II)S_3^- in Pb(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structure. a) Top down view from the N-termini and b) side view of the binding site. Main chain atoms are shown as ribbon diagrams in purple, 12Leu and 16Cys side chains in the sixteenth position are present as sticks (sulfur = bright yellow). The Pb(II) ion is shown as a grey sphere.

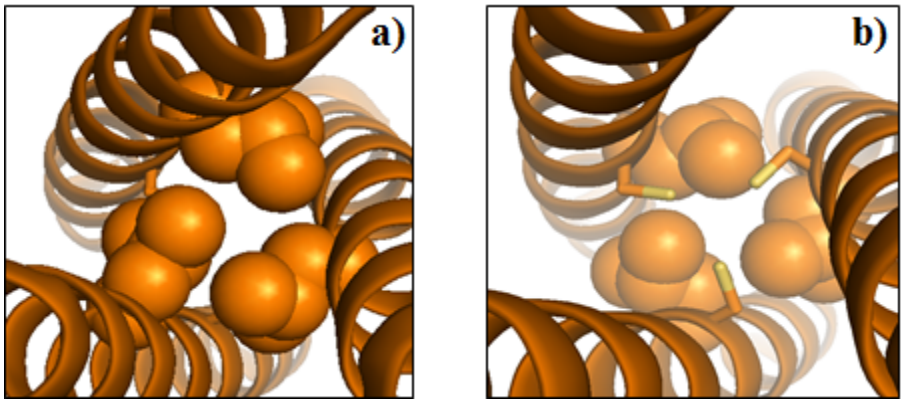

Figure 2-11. Packing of the hydrophobic residues around the 16Cys site of apo-(CSL16C)$_3$. Shown from top down view of the N-termini, representing Leu residues as spheres in a) the twelfth and b) the nineteenth positions. The other amino acid side chains have been omitted for clarity.

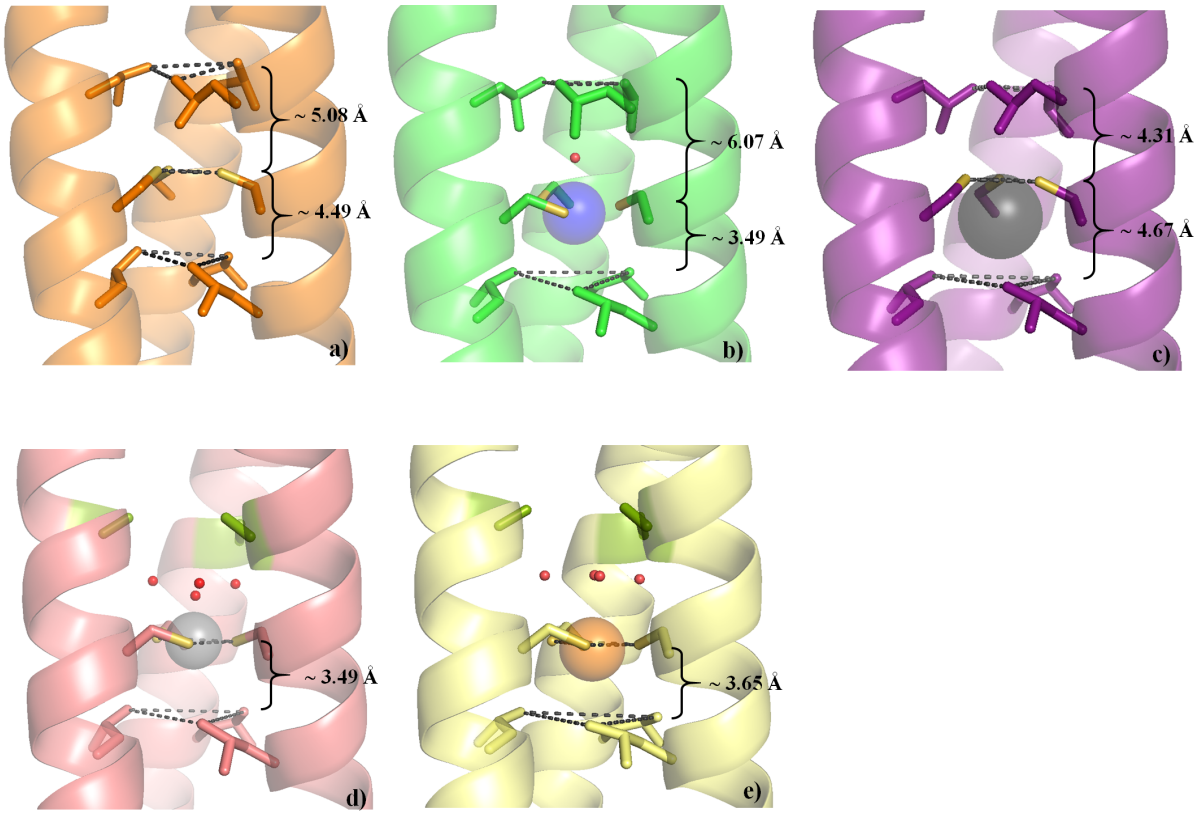

Figure 2-12. Interlayer spaces around the 16Cys layer. Shown in a) apo-(**CS**L16C)$_3$, b) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, c) Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, d) Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ and e) Hg(II)(**GRAND-CS**L12AL16C)$_3^-$. Main chain helices are shown as ribbon diagrams, Residue side chains are present as sticks. Metal centers are shown as spheres.

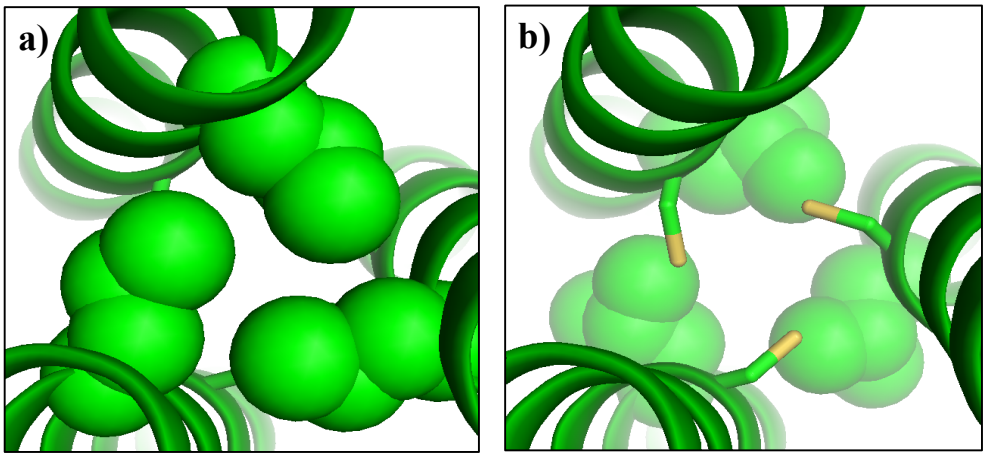

Figure 2-13. PyMOL representation of the hydrophobic packing around the 16Cys layer of $Hg(II)_S Zn(II)_N$(GRAND-CSL16CL30H)$_3^+$. The packing of Leu residues is shown as spheres from the top down view of a) the twelfth layer and b) the nineteenth layer. The Cys side chains are shown as sticks where sulfur atoms are colored yellow. The Hg(II) and observed water above the 16Cys site are omitted for clarity.

and 3.80 Å when considered from the plane made by these Cδ atoms (**Figure 2-14**). For the 19Leu layer, the residue side chain is facing more outward toward the helical interface. This layer is at a distance of 3.49 Å below the metal site (**Figure 2-12b**).

*Zn(II)(**GRAND-CS**L12A16C)$_3^-$*

Alanine is a smaller amino acid than leucine (containing a methyl group rather than an isopropyl group) so the side chain does not contain the four atoms (N-Cα-Cβ-Cγ) that are used to define the side chain torsion angle. Here, only the Cβ-Cβ separation (5.94 Å) between two residues is reported in order to provide information on the cavity size when Ala is used at the twelfth position in Zn(II)(**GRAND-CS**L12A16C)$_3^-$ (**Figure 2-15a**). Similar to other constructs, Leu is placed at the nineteenth position, which is one layer below the metal site. Obviously, the side chain orientation points toward the outer face. The nineteenth layer is separated from the sulfur plane at a distance of 3.65 Å (**Figure 2-12d**).

*Hg(II)(**GRAND-CS**L12A16C)$_3^-$*

Similar to Zn(II)(**GRAND-CS**L12A16C)$_3^-$, the Ala residue is incorporated in the hydrophobic layer above the metal site. The Cβ-Cβ of Ala residues from two neighboring chains is 5.86 Å separated from each other (**Figure 2-15b**). The overall site is illustrated in **Figure 2-9.** It is obvious that once Hg(II) is bound there are four water molecules observed within the hydrophobic cavity between the 12Ala and 16Cys layers (**Figure 2-9c** and **2-9d**). One water molecule (with 33% occupancy per ASU) is situated on the threefold axis directly above the Hg(II) at a distance of 3.56 Å. The rest of the waters (4.35 Å away from the Hg(II)) are threefold related and located close to the helical interface between two neighboring chains. Their positions are almost at the same height as the central water (with respect to the N-termini) with a separation of 2.63 Å between these water molecules, suggesting that they are hydrogen bonded. Unlike the twelfth position, Leu is placed below the metal site at the nineteenth position. The residue displays a single conformation, which is oriented to the helical interface. The nineteenth layer is at a distance of 3.49 Å below the Cys plane (**Figure 2-12e**).

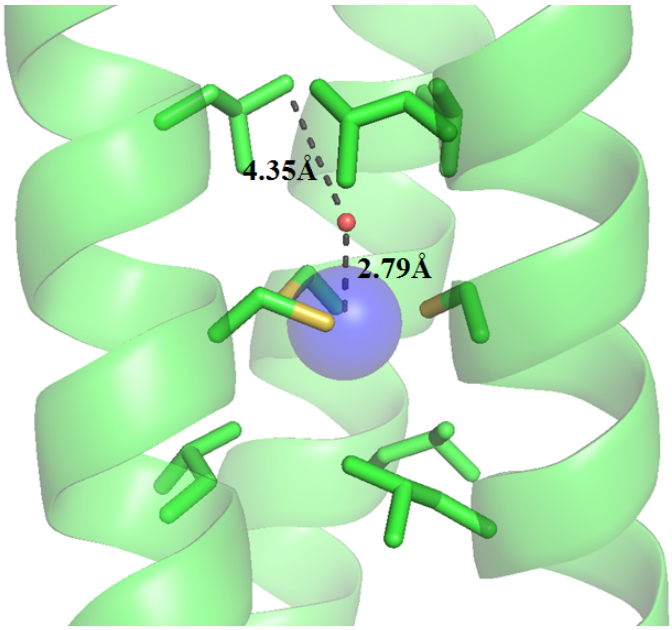

Figure 2-14. PyMOL visualization representing the distances of the uncoordinated water observed in $Hg(II)_S Zn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$ from the Hg(II) and the core δ-methyl atoms of the 12Leu residues. Main chain helices are shown as green ribbons, Cys side chains in the sixteenth position are in stick form (sulfur = yellow), The Hg(II) ion is shown as a blue sphere and the uncoordinated water is represented as a small red sphere.

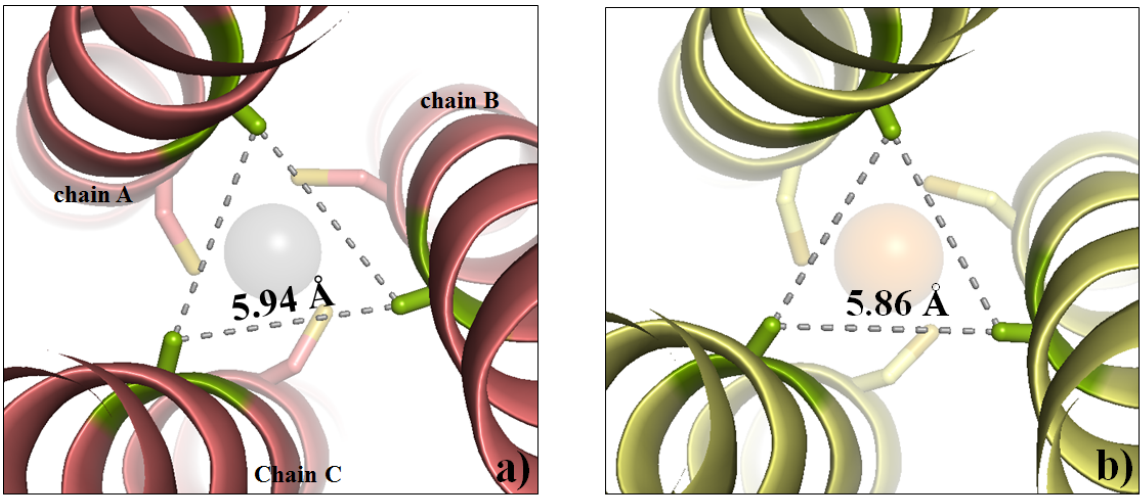

Figure 2-15. **Representation of the distances between the Cβ atoms of 12Ala residues above the 16Cys site of a) Zn(II)(GRAND-CSL12AL16C)$_3^-$ and b) Hg(II)(GRAND-CSL12AL16C)$_3^-$.** Main chain helices are shown as ribbon diagrams, Cys side chains in the sixteenth position are present as sticks (sulfur = bright yellow). Ala residues at the twelfth position are colored in green. Zn(II) and Hg(II) ions are shown as grey and orange spheres, respectively. All waters are omitted for clarity.

*Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$*

The overall packing around the 16Cys site is shown in **Figure 2-16**.The side chains in the layer at the twelfth position are directed toward the center of the structure with a distance of 4.31 Å above the Cys plane (**Figure 2-12c**). In the nineteenth position, the orientation of 19Leu side chains is points toward the C-terminus and the helical interface. This layer is 4.67 Å below the Cys plane.

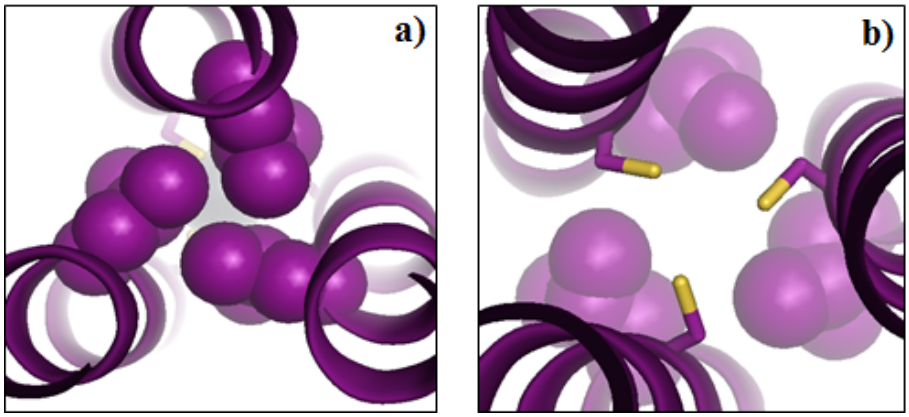

Figure 2-16. **PyMOL representation of the hydrophobic packing around the 16Cys layer of $Pb(II)_S Zn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$**. The packing of Leu residues is shown as spheres from the top down view of a) the twelfth layer and b) the nineteenth layer. The Cys side chains are shown as sticks where sulfur atoms are colored yellow. The Pb(II) ion is omitted for clarity.

Discussion

<u>Section I</u>: Structural studies of mercury binding to a tris thiolate site within a simplified three stranded coiled coil peptide

A combined approach using UV-VIS, EXAFS, ^{199}Hg NMR and ^{199m}Hg PAC spectroscopies confirmed that at an appropriate pH and with a stoichiometric 1:1 ratio of 3SCC to Hg(II), **TRI**L16C can form a tris-(thiolato)Hg(II) complex. It was reported that Hg(II)(**TRI**L16C)$_3^-$ was the first water soluble spectroscopic model prepared in a simplified protein environment for Hg(II) bound to the metalloregulatory protein MerR.[27,30–32] One of the important questions in metalloprotein studies is to gain better understanding of how protein ligands in an active site are arranged prior to metal binding and how those ligands change their geometries when the metal is bound. That is, to what extent are these designed sites *preorganized* for metal binding? Native protein ligands are frequently required to be *preorganized* before metal binding to impart structural constraints that control the properties (e.g., reduction potential) of a metals.[33] From this perspective, one should ask whether the 16Cys sites engineered within the hydrophobic core of these 3SCCs are truly *preorganized* or simply *predisposed* for heavy ion complexation.

In the apo-(**CS**L16C)$_3$, the Cys residues in chain A and C display two alternate conformations as shown in **Figure 2-1**. The major conformation for all cysteine groups orients the thiol sulfurs toward the center of the core and in the direction of the N-termini of the peptides. The major conformation for the three cysteine side chains considered as a single unit is based on the major Sγ conformers (60% occupancy) of these chains combined with the third Cys in chain B (70% occupancy). Thus, approximately 25% of the time the orientation of these three Cys conformations might, at first inspection suggest a *preorganized* metal binding because it represents a reasonable metal chelating environment with an averaged Sγ-Sγ distance of 3.29 Å. The remaining possibilities constitute combinations of cysteines with the thiol pointed toward the helical interface, with the minor conformation (accounting for ~ 3%) where all three thiols are supposed to arrange toward the helical interface. It is clear that these mixed cysteine orientations are at best partially *preorganized* and, in the case of the minor isomer fully *predisposed*, but not *preorganized* for metal binding. The same Cys arrangements are also observed in the inner Sγ conformers of the published apo-(**CS**L9C)$_3$ where the Cys residue are in the ninth position (**Figure 2-17**). The Sγ-Sγ separation of the 9Cys is at a distance of 3.32 Å, whereas the

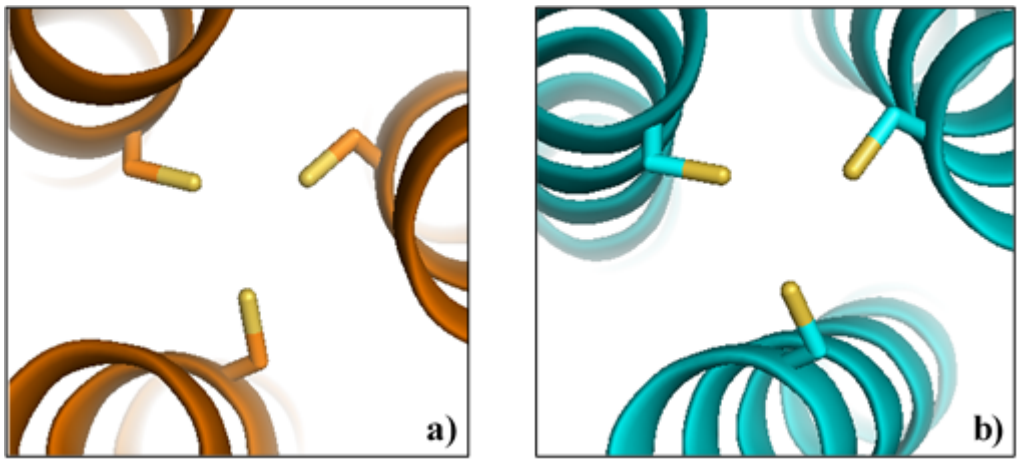

Figure 2-17. Ribbon diagrams shown from the top down view of 3SCC represents the similarity of the major Cys conformer orientation between a) apo-(CSL16C)$_3$ and b) the published apo-(CSL9C)$_3$[12] structures. The Cys side chains are shown as sticks with sulfurs labeled as yellow.

$S\gamma$-$S\gamma$ separation of the *d* site, nineteenth layer in the apo-(**CS**L19C)$_3$ gets longer to 3.85 Å.[12] These observations clearly suggest that Cys layers at the *a* site generate the metal binding sites that are comparable in size, which is slightly smaller than the *d* site. Again, in both previous structures, disorder was observed between two conformations for each cysteine side chain.

The averaged χ_1 torsion angle of these inner Cys rotamers in apo-(**CS**L16C)$_3$ is -69.20^o (the range is approximately 25%), which is within the distribution of value observed in apo-(**CS**L9C)$_3$ (-70.52 ^{o}).[12] This analysis indicates that the *a* site Cys residues behave roughly similarly regardless of the sequence position and are predisposed similarly for metal complexation. Importantly, both χ_1 values correspond to the most common χ_1 reported for unbound Cys residues in protein structures demonstrating that these simple constructs mimic natural systems well.[22]

The minor (exterior) conformations in chain A and C point their $S\gamma$ atoms toward the helical interface and outward to the C-termini and the $S\gamma$-$S\gamma$ separation lengthens to 6.24 Å (**Figure 2-1, c**). Combining the exterior 16CysA and 16CysC conformations with the single rotamer 16CysB, the overall binding pocket is bigger and does not look *preorganized* for metal binding as only one of the Cys residues is directed at the core while the others are facing out. As a consequence, the averaged $S\gamma$-$S\gamma$ separation becomes longer (5.36 Å). Based on these crystallographic observations, only the interior $S\gamma$ arrangements appear to have the potential for a *preorganized* metal binding environment.

As a consequence of crystallizing in the R32 space group, the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ structure has an imposed three fold axis (along *z* directed from N- to C- termini) that requires each helix in the 3SCC to be identical. The Cys side chains in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ display two alternate conformations (**Figure 2-3**), both of which are pointing to the helical interface and downward to the C-terminal end of the protein, in marked contrast to the inner Cys conformers observed in the apo-(**CS**L16C)$_3$ (**Figure 2-1**). This orientation of the Cys residues causes a large separation between adjacent thiols ($S\gamma$-$S\gamma$) at 4.16 Å for the interior $S\gamma$ conformers (χ_1 torsion angle = -151.88^o) and 5.24 Å for the exterior $S\gamma$ conformers (χ_1 torsion angle = -173.35^o). As the outer Cys side chains are directed even further outward to the solvent region, only the inner Cys side chains are considered to be the bound Hg(II)

rotamers with an excellent agreement for the Hg(II)-S bond distance of 2.43 Å (**Figure 2-4**). (The Hg(II) to outer S_γ distance is 3.22 Å, which is too long to be a Hg(II)-S bond length for a trigonal thiolate Hg(II) complex.) This value (2.43 Å) is consistent with the Hg(II)-S bond length reported from the tris-(thiolato) trigonal planar Hg(II) complexes in native Hg(II)-MerR (EXAFS)[23,24] and small model $Hg(II)S_3^-$ complexes (EXAFS, solid state X-ray crystallography).[24–26] The bound Hg(II) ion is situated at a distance of 0.30 Å below a trigonal plane of the three interior Cys residues with a S-Hg(II)-S angle of 118.36°.

Surprisingly, there is a water molecule present in the hydrophobic cavity between the 12Leu and bound-16Cys layers as shown in **Figures 2-4 and 2-14**. This water is situated on the threefold axis at a distance of 2.79 Å above (with respect to the N-termini) the Hg(II) atom. I assign this as an unbound water because the observed Hg(II)-O distance is too long to represent a Hg(II)-O bond (typically 2.20-2.30 Å).[34,35] Moreover, the metal structure certainly is trigonal with a S-Hg(II)-S angle close to the perfect 120° and the Hg(II) ion does not look as if it is distorted out toward the water. Thus, the possibility of a pseudo-tetrahedral $Hg(II)S_3O$ structure is ruled out. This water, on the other hand, H-bonds to the p-orbitals of sulfurs stabilizing the overall charge of the metal site (O-S distance of 3.41 Å). In addition, no steric clashing between the 12Leu residues and water was observed as validated by the MolProbity program.

To determine the extent to which the metal site organization in the first coordination sphere is changed from an unbound to a Hg(II) bound state and how this affects the hydrophobic environments around the metal site, the interior Cys layer of the apo-(**CS**L16C)$_3$ is overlaid on the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ as illustrated in **Figure 2-18**. The Cys residues in both structures similarly point their C_β carbons toward the N-termini; however, the directions through space of the S_γ thiols are completely different. The S_γ atoms in the apo-protein are directed to the helical core with the χ_1 torsion angle of -69.20° and S_γ-S_γ separation of 3.29 Å. In contrast, longer S_γ-S_γ (4.16 Å) and larger negative χ_1 side chain torsion angle (-151.88°) are observed in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure. These observations suggest that metal binding induces significant rotation of the interior Cys conformations by moving the cysteine sulfur atoms from directing at the core toward the N-termini (in the absence of the metal) to facing outward to the helical interface and downward to the C-termini (in the presence of the metal). The

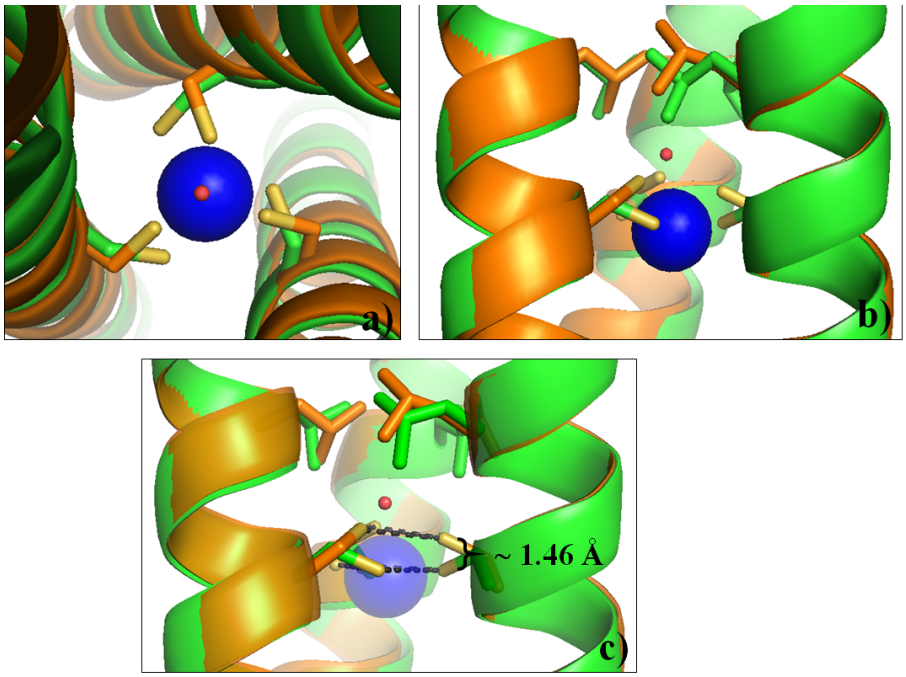

Figure 2-18. Ribbon diagrams demonstrating an overlay of the sixteenth layer between Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ and apo-(CSL16C)$_3$. a) Top down view from the N-termini and b) Side on view. Main chains of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are colored in green and apo-(**CS**L16C)$_3$ in orange. Cysteine side chains are shown as sticks. Hg(II) is present as a blue sphere where the observed water is shown in red.

sulfur planes in these two situations are different by 1.46 Å (**Figure 2-18, c**). Also, the shift in sulfur orientations causes a longer S_γ-S_γ separation by 0.87 Å from the apo-structure, implying more space of the metal pocket is required in order to fit the large ionic radius (1.16 Å) of Hg(II) into the trigonal plane with a Hg(II)-S distance of 2.43 Å.

After assessing the impact of metal binding on the first coordination sphere ligands, it is imperative to examine how the bound metal influences the hydrophobic environment around the 16Cys layer. This is especially important when one considers the water molecule anchored above the bound metal site in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. This observation raises the question of how a polar molecule like water can be housed within the core of the hydrophobic region. To address this question, comparison of the 12Leu and 19Leu layers of the two structures are taken into consideration.

The overlay structure reveals that the overall packing of the 12Leu layers of both structures look similar to each other but are not exactly identical as shown in **Figure 2-19**. All 12Leu residues are oriented away from the core center and directed toward the helical interface, loosening the packing of the hydrophobes above the metal site. Each leucine side chain points one of the two δ-methyl atoms inward to the core (higher in plane with respect to the N-termini) compared to the second δ-methyl that is outward toward the interface (lower in height, pointing toward the C-termini). However, in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ the $C_{\delta 2}$ atoms are directed to the center of the helix and are higher in plane, while the other $C_{\delta 1}$ atoms are oriented outward. The related distances in this metal bound structure become slightly shorter indicating the hydrophobic leucine residues are packing slightly tighter than the apo-peptide (**Figure 2-19, c**). Moreover, the interior $C_{\delta 2}$ plane of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ is slightly lower than the interior δ-methyl plane of apo-(**CS**L16C)$_3$ with respect to the N-termini. Based on the crystallographic evidence mentioned thus far, it can be concluded that the rotation of the Cys residues upon metal binding causes the metal plane to lower toward the C-terminus. Even though the thermodynamically favorable Hg(II)-thiolate bonds add more stability to hold the three helices together[36–38], in order to help maintain a reasonable degree of hydrophobic packing of the 3SCC scaffold the 12Leu layer in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ also has to move downward toward the 16Cys layer in order to keep the separations between the twelfth and sixteenth layers

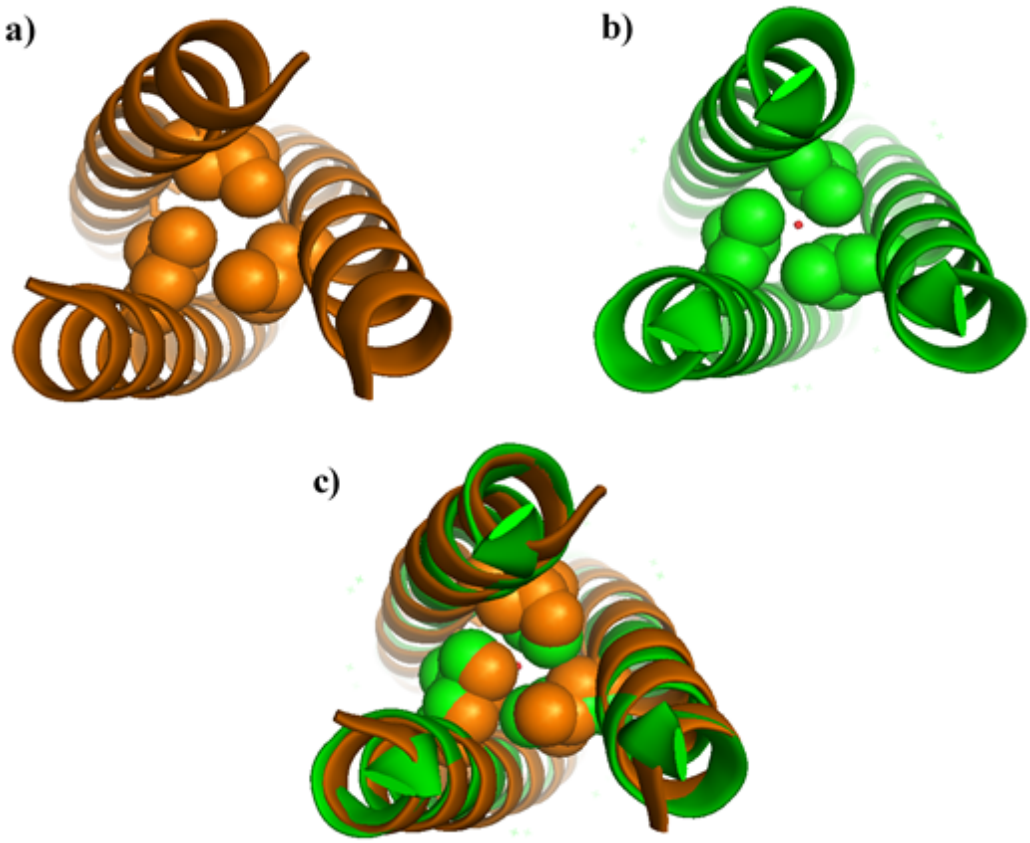

Figure 2-19. Packing of hydrophobic layers at the twelfth position shown as spheres above the 16Cys site. From top down view of the N-termini, representing the 12Leu layer of a) apo-(**CS**L16C)$_3$, and b) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. In c) represents the packing comparison in the twelfth layers between a) and b). Main chains of apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in orange and green respectively. The observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is shown as a small red sphere.

comparable. The slight shift of 12Leu in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ does not lead to a steric clash with the observed water molecule. The distance between the interior $C_{\delta 2}$ atom to the oxygen of water is 4.35 Å (**Figure 2-14**). In the nineteenth layer, as shown by the distance and angles in **table 2-4** and the PyMOL visualization in **Figure 2-20**, in order to avoid the steric clashing between the sulfur layer and the 19Leu layer, the Leu side chains must reorient far out to the helical interface and shift toward the C-termini.

This structural analysis demonstrates that the Hg(II) binding not only directly perturbs the metal organization in the primary coordination sphere, but also indirectly affects the packing around the metal site in the second coordination. In order to fit a Hg(II) ion into the metal binding plane, the thiols in the apo-protein must adapt to the geometry and size requirement of the trigonal planar structure by reorienting the torsion angles of the Cys residues from -69.20^o (upward to the N-termini) to -151.88^o (downward to the C-termini). This shift causes the sulfur plane of the metallated structure lower than the apo-structure by 1.46 Å, which consequently affects both of the packing of the leucine residues above and below the metal site. It is reasoned that the lowering bound-16Cys residues would clash with the 19Leu residues if they were to position in the same orientations as observed in the apo-structure. Thus, in order to avoid the steric clashing between these adjacent layers, the 19Leu residues have to reorient far out to the helical interface and lower farther toward the C-termini. Similarly, the lowering in height of the 12Leu layer occurs in the metallated versus apo-structure; however, the degree of hydrophobic packing is changed in the opposite direction from the 19Leu. As the bound-16Cys layer is repositioned toward the C-termini, the 12Leu residues in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ become more tightly packed as indicated by the closer distances of δ-methyl atoms of 12Leu above the metal site encouraging the Van Der Waal contacts of the hydrophobic groups. The combination of the favorable enthalpy for the Hg(II)-thiolate formation and the tightened 12Leu layer is believed to make up for the reorganization energy penalty cost by the rotation of thiols to bind Hg(II) in the 16Cys site and allows retention of the native folding of the 3SCC structure.

In sum, the crystal structure of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ clearly confirms the previous published physical characterizations of Hg(II)(**TRI**L16C)$_3^-$ [27,30–32] that at a specific pH 8.5, Hg(II) is bound to the hydrophobic interior of 3SCC as a trigonal thiolate Hg(II) complex which provides a good crystallographic model of the active site of the MerR metalloregulatory protein when it is bound to a toxic Hg(II) ion. One can now establish general rules for mercury

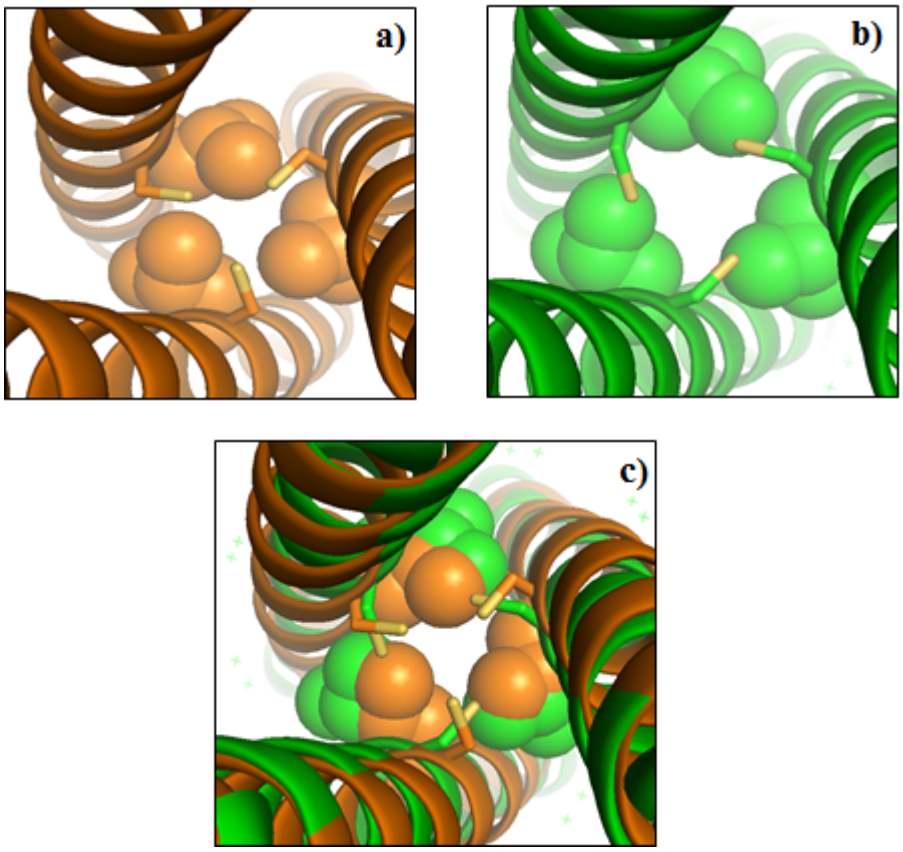

Figure 2-20. Packing of hydrophobic layers at the nineteenth position shown as spheres below the 16Cys site (sticks). From top down view of the N-termini, representing the 19Leu layer of a) apo-(**CS**L16C)$_3$, and b) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. In c) represents the packing comparison in the twelfth layers between a) and b). Main chains of apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in orange and green respectively.

binding to a Cys site in a designed protein and apply this understanding to the real biology of MerR regulatory proteins.

<u>Section II</u>**: **Structural analysis of factors controlling Cd(II) coordination geometry

From the perspective of metalloprotein design it is essential to understand how to construct a peptidic environment that constrains a metal within a specific coordination geometry and a defined oxidation state. Defining the structural principles that control structure, and subsequently function, are core objectives of protein design. Having been spectroscopically characterized for over a decade, the studies of tris-thiolato Cd(II) interactions within a three stranded coiled coil has been well-established.[12,27,30,39–46] While the parent metallopeptide Cd(II)(**TRI**L16C)$_3^-$ yielded a mixture of 3- and 4-coordination[27], a combination of UV-vis, EXAFS, ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies confirmed that an exclusively 3-coordinate Cd(II)S_3^- center could be achieved by either direct modification of the first coordination sphere ligand (using penicillamine) or indirectly through perturbation of the second coordination sphere of the metal (L12$_D$L modification).[40,47] Similarly, an exclusive 4-coordinate Cd(II)S_3O^- [28,29] could be isolated by decreasing the steric constraints of the layer above (L12A) the metal binding site. Despite unequivocal assignment of these trigonal planar and pseudo tetrahedral structures, an explanation of precisely how these sequence modifications influenced the obtained metal structures has been elusive in the absence of crystallographic data.

Unfortunately, structural details of such modifications have not yet been revealed even though great effort has been spent on optimizing the crystal growth conditions of Cd(II)-bound designed peptides. To date, no crystals have been isolated of **CoilSer** or **GRAND-CoilSer** constructs with a bound Cd(II). Promising candidate crystals revealed the overall packing of three-helices was obtained; however, both of the resulting electron density maps ($2F_o$-F_c and F_o-F_c) clearly demonstrated that Cd(II) was not bound in the Cys site. It is reasoned that Cd(II) was not able to bind to the protein under the crystal growth conditions due to the high affinity of this metal ion toward the oxygen-containing precipitants that are present (e.g. polyethylene glycol, glycerol and ethoxyethanol). Usually these materials are at concentrations much higher than the protein, so it is possible that the Cd(II)-thiolate center might not be able to compete thermodynamically in the presence of those concentrated oxygen ligands. Thus, direct observation of the Cd(II) with these 3SCCs has been unsuccessful.

In order to determine the impact of changing the internal hydrophobic residues on controlling Cd(II) coordination numbers within 3SCC constructs, the known trigonal planar $Hg(II)S_3^-$ of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ described in the previous section could provide a reasonable model for tris-thiolato Cd(II) complexes or at least for a trigonal planar $Cd(II)S_3^-$ structure. The Hg(II)-S bond distance of the trigonal planar $HgS(II)_3^-$ from the X-ray crystal structure of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ (2.43 Å) is in excellent agreement with the X-ray absorption data (2.44 Å) of Hg(II)(**TRI**L16C)$_3^-$. At the same time, the EXAFS Cd(II)-S bond distance for the trigonal planar Cd(II)(**TRI**L16Pen)$_3^-$ is 2.46 Å (see **Table 2-6)**. Obviously, the bond distances of the trigonal planar $Hg(II)S_3^-$ and $Cd(II)S_3^-$ structures are within experimental errors. Regardless of the metal size difference, the coordinated $Hg(II)S_3^-$ in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ crystal structure, thus, can be used to explain general characteristics of $Cd(II)S_3^-$ and model other Cd(II) complexes in the studies.

In this section, structural analyses of various crystal structures will be discussed to assess possible models of Cd(II)-binding studies of three peptide designs (**TRI**L16C, **TRI**L16Pen and **TRI**L12AL16C). First, the inherent structural difference between the Cys site [apo-(**CS**L16C)$_3$] and Pen site [published apo-(**CS**L16Pen)$_3$][30] will be compared to the metallated forms of $Hg(II)(S_{Cys})_3^-$ in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and the $Hg(II)(S_{Pen})_3^-$ in the reported $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N$(**CS**L9PenL23H)$_3^{n+}$.[63] The perturbation caused by the Pen ligand on water exclusion in Cd(II)(**TRI**L16Pen)$_3^-$ will be discussed. Second, the coordinated Hg(II) in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ crystal structure will be used to model the mixture of $Cd(II)S_3^-$ and $Cd(II)S_3O^-$ in Cd(II)(**TRI**L16C)$_3^-$. Finally, the impact of 12Ala on water access to the metal site will be demonstrated by the metallated structures Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ and Zn(II)(**GRAND-CS**L12AL16C)$_3^-$.

Lee *et al*. demonstrated that the incorporation of the Pen ligand in lieu of Cys at the sixteenth position, generating **TRI**L16Pen, led to $Cd(II)S_3^-$ coordination as confirmed by ^{113}Cd NMR and ^{111m}Cd PAC spectroscopies.[29] There are three possible hypotheses of how Pen can perturb the metal coordination environments of the $Cd(II)S_3^-$ center in Cd(II)(**TRI**L16Pen)$_3^-$. The first hypothesis is that the Pen ligands position their γ-methyls toward the space above the metal plane. When enough space is not available, this may lead to water exclusion. The second hypothesis is that Pen ligands at the sixteenth position cause a tighter hydrophobic packing of 12Leu plane reorienting the sidechains toward the metal layer which does not allow for water

access. The third hypothesis is that the Pen ligands perturb the primary coordination sphere of metal in a specific way that encourages a $Cd(II)S_3^-$ formation. To test the first and second assumptions, the apo-(**CS**L16Pen)$_3$ structure is aligned with apo-(**CS**L16C)$_3$. As shown in **Figure 2-21**, the Pen-containing construct is unperturbed and the helical backbones of the two structures clearly demonstrate their similarity in secondary structure. The Pen ligands observed in the apo-(**CS**L16Pen)$_3$ also display two alternative conformations (**Figure 2-22**). Only the major conformers (with 95% occupancy) are considered to be *preorganized* metal ligands resembling the major Cys residues in apo-(**CS**L16C)$_3$, where the $S\gamma$ atoms point at the central core and toward the N-termini. The average $S\gamma$-$S\gamma$ separation in apo-(**CS**L16Pen)$_3$ is 3.71 Å which is longer than one observed in apo-(**CS**L16C)$_3$ (3.29 Å) indicating that Pen makes a slightly larger metal pocket than Cys. The average χ_1 torsion angle for the γ-thiols in apo-(**CS**L16Pen)$_3$ is -49.85^o and it is -69.20^o in apo-(**CS**L16C)$_3$. In contrast, the γ-thiols of the minor conformers (with 5% occupancy) in apo-(**CS**L16Pen)$_3$ are oriented to the helical interface, generating a much longer $S\gamma$-$S\gamma$ separation of 8.36 Å (average). These minor $S\gamma$ atoms of Pen, though they are facing outward from the core like the minor Cys conformers, are directed upward to the N-termini which is in the opposite direction of those observed in apo-(**CS**L16C)$_3$. However, only the major ligand conformers of both structures (**Figure 2-23**) are overlaid and used to discuss the extent of ligand pre-organization.

The thiol orientations achieved by Cys and Pen ligands are almost identical as illustrated in **Figure 2-23**, **a**. While the $S\gamma$ atoms of Pen in apo-(**CS**L16Pen)$_3$ are oriented to the interior of the coiled coil, the γ-methyl groups are pointing to the exterior. The fact that the Pen rotamers are almost at full occupancy suggests that these side chain conformations are geometrically preferred when Pen is placed at the sixteenth position; however, the similar orientations of γ-thiols observed in Cys structure exhibits only 60% occupancy, implying that there is more free rotation of γ-thiols in Cys rather than Pen. The rigidity of these $S\gamma$ atoms in apo-(**CS**L16Pen)$_3$ likely results from the restricted thiol rotation around $C\beta$ atom due to the steric constraint imposed by the two γ-methyl groups of Pen. Based on these crystallographic observations, the γ-methyl groups of Pen, which are oriented toward the helical interface, are not positioned to block the space above the

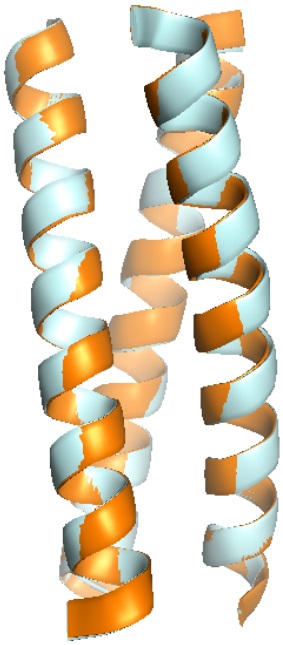

Figure 2-21. Side view overlay of apo-(CSL16C)$_3$ (orange) with the published apo-(CSL16Pen)$_3$ [47] (grey) demonstrates their structural similarity. Both structures are well-aligned indicating that the incorporation of a non-natural amino acid, Pen does not perturb the secondary structure of the three-stranded coiled coil.

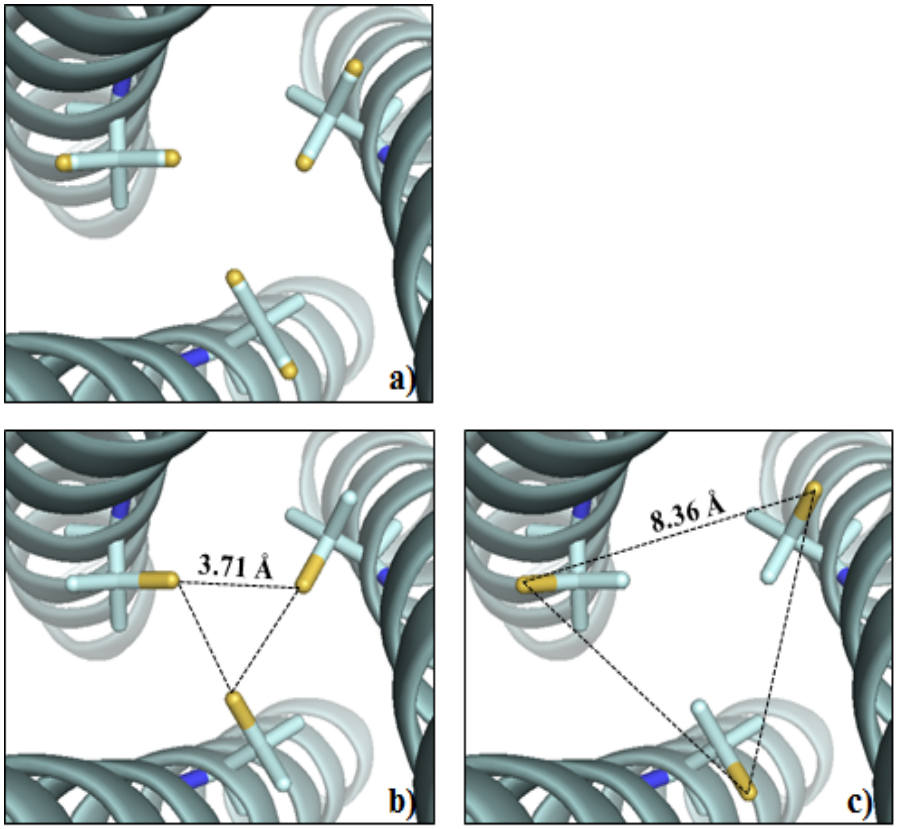

Figure 2-22. Ribbon diagrams of apo-(CSL16Pen)$_3$ (PDB code: 3H5F)[47] showing the orientation of Pen side chains. Main chain atoms are shown in grey and the Cys side chains are in sticks with the thiol groups being yellow colored coded. Top-down view from N-termini showing the exclusive orientations of a) all conformers, b) major conformers and c) minor conformers.

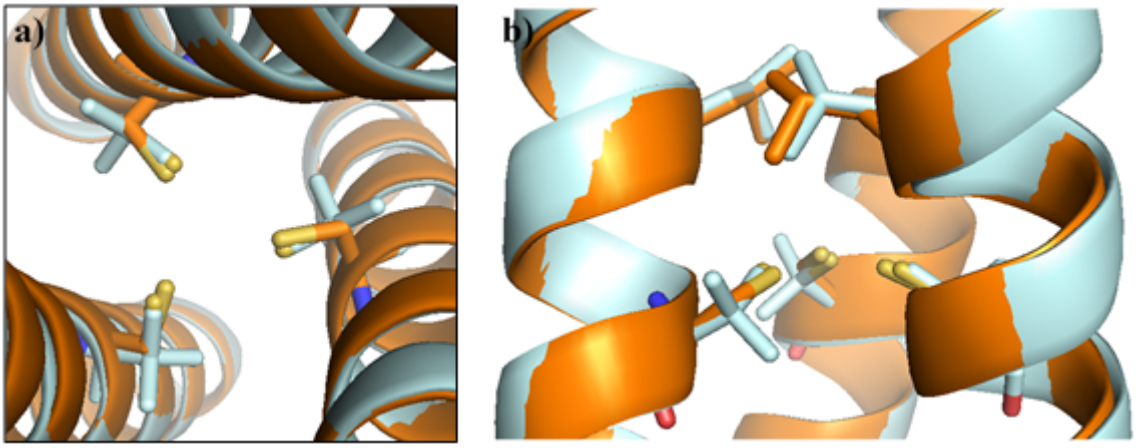

Figure 2-23. Comparison of the sixteenth layer between the Cys ligands from apo-(CSL16C)$_3$ and Pen ligands from apo-(CSL16Pen)$_3$ (PDB code: 3H5F)[47]**.** a) Top down view from the N-termini and b) Side on view. Main chains of apo-(**CS**L16C)$_3$ are colored in orange and apo-(**CS**L16Pen)$_3$ in grey. Cysteine and Pen side chains are shown as sticks.

metal site (between the 12Leu and 16Pen layers) that can exclude water from binding to Cd(II). Thus, the first hypothesis is ruled out.

The second hypothesis can be addressed by comparing the hydrophobic layer above the metal site between apo-(**CS**L16Pen)$_3$ and apo-(**CS**L16C)$_3$. In **Figure 2-24**, regardless of which ligands are used to produce the metal site, the overall hydrophobic packing between both layers are equivalent (or even slightly tighter in the apo-(**CS**L16C)$_3$ structure because of the 12LeuB that is oriented more at the helical core). All the 12Leu side chains are directed toward the helical interface with the Cβ carbons positioned toward the N-termini of the helical scaffolds. Moreover, their δ-methyl atoms are also positioned comparatively in height. These observations indicate that comparable size and side chain orientation of the hydrophobic pockets above the metal layer exist in both structures. A comparison of the apo structures would suggest that if the cavity above the L16C layer (which is slightly tighter than in Pen-structure) allows for water coordination to the Cd(II) site resulting in ~60% Cd(II)S$_3$O$^-$, then, the hydrophobic cavity between the 12Leu and 16Pen should potentially do so as well. Thus, based on the apo-structures alone, there is no *preorganization* of the 12Leu layers to support hypothesis 2; however, it is possible that differential packing in the 12Leu layer might occur solely between the two metallated structures.

To make this comparison the two crystal structures of metallated Hg(II)(S_{Pen})$_3^-$ from Zastrow *et al.*[63] and Hg(II)(S_{Cys})$_3^-$ in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ will be compared. Although the structure of metallated-(**CS**L16Pen)$_3$ is not available, the published Hg(II)(**CS**L9Pen)$_3^-$ site is known and can reveal structural information required for *a* site metal binding. The [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$^-$)]$_N$(**CS**L9PenL23H)$_3^{n+}$ design contains a trigonal planar Hg(II)(S_{Pen})$_3^-$.[2,47,48] The orientations of the bound Pen ligands are shown in **Figure 2-25** demonstrating that their Sγ atoms are directed at the interior core and have the γ-methyl groups positioned out toward the helical interface. The Cγ_1 atoms are pointing up toward the N-termini and the Cγ_2 atoms are downward to the C-termini. These configurations are similar to what was observed in the apo-(**CS**L16Pen)$_3$ (**Figure 2-22, b**).[47] The Hg(II) ion is situated in the center of the trigonal plane generated from the three penicillamine-γ-thiols. In fact, this metal center is quite distorted with S-Hg(II)-S angles of 135.86^o, 115.65^o and 108.26^o. These angles deviate significantly from 120^o of a perfect trigonal planar structure; however, the average angle of 119.92^o

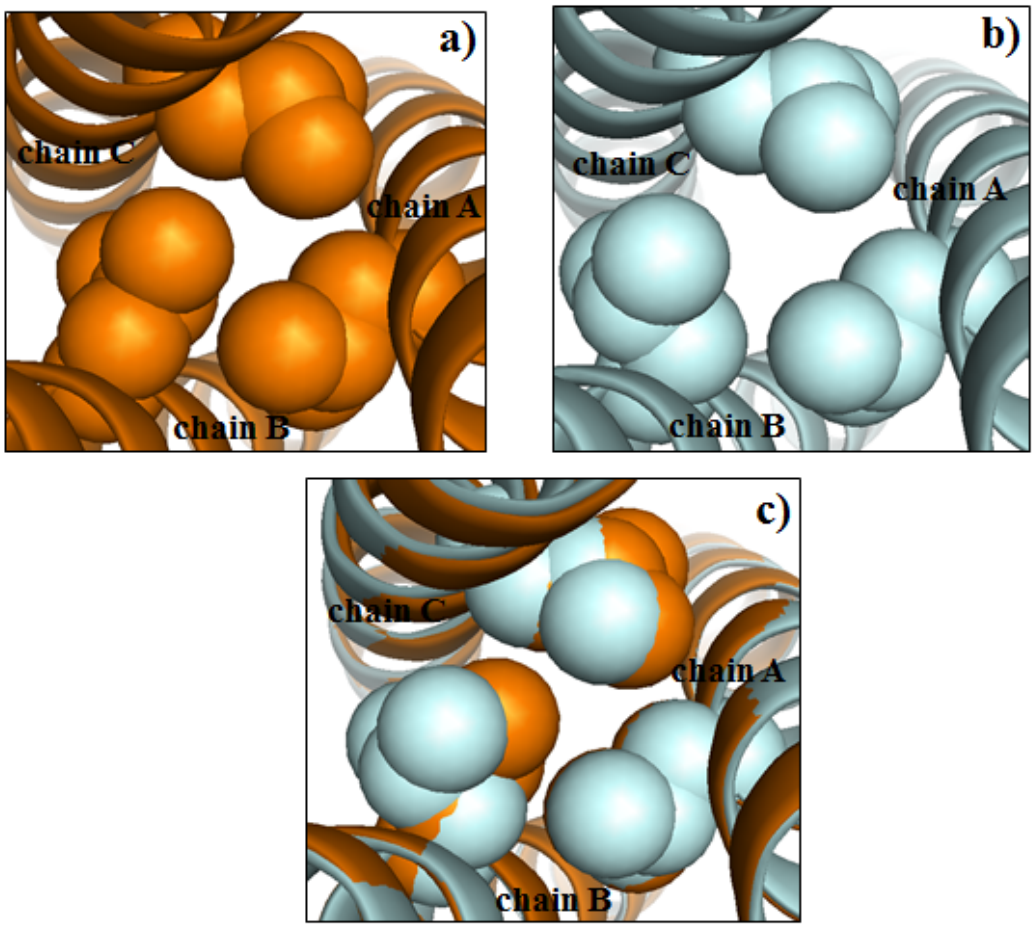

Figure 2-24. Packing of hydrophobic layers at the twelfth position shown as spheres above the metal binding layers. From top down view of the N-termini, representing the 12Leu layer of a) apo-(**CS**L16C)$_3$, and b) apo-(**CS**L16Pen)$_3$ (PDB code: 3H5F)[47]. In c) represents the packing comparison in the twelfth layers between a) and b). Main chains of apo-(**CS**L16C)$_3$ and apo-(**CS**L16Pen)$_3$ are shown in orange and grey, respectively.

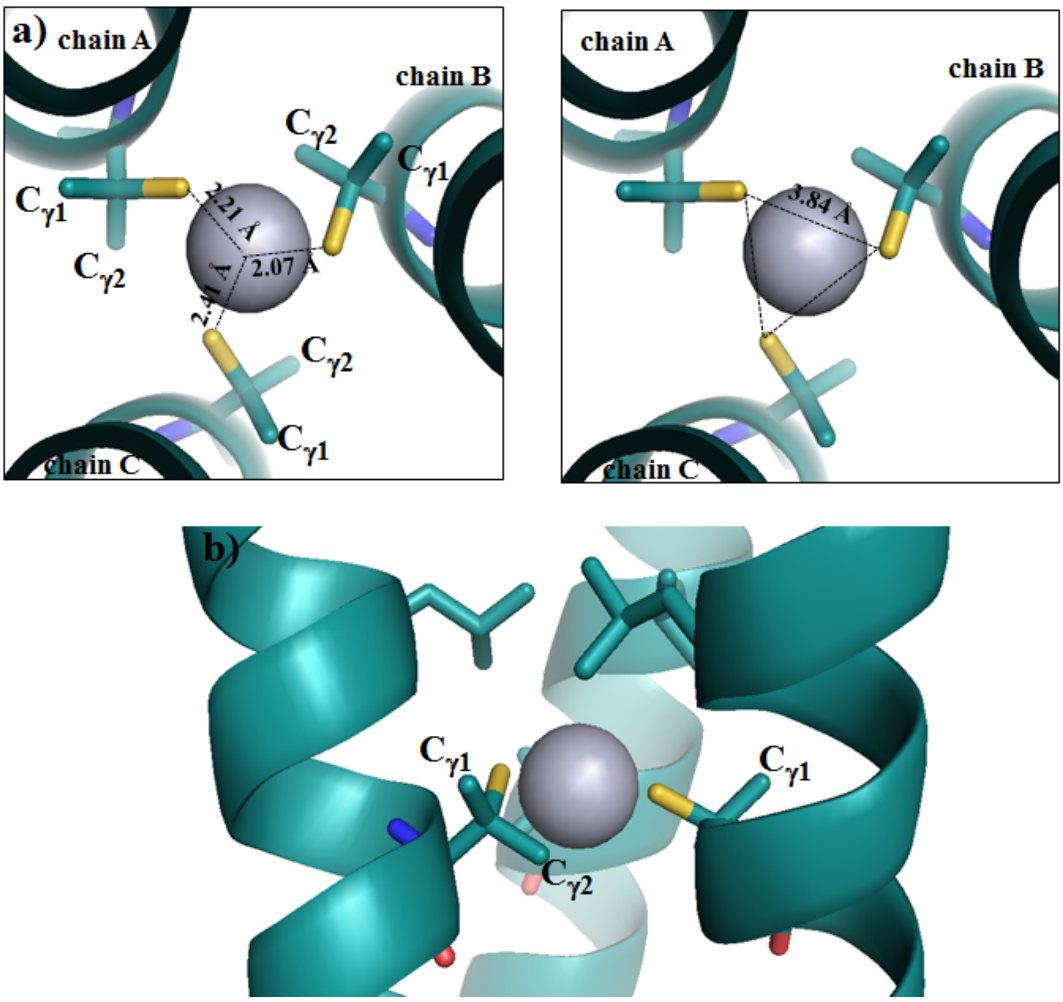

Figure 2-25. Ribbon diagram of a trigonal thiolate site, $Hg(II)S_3^-$ in the published $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$ crystal structure (PDB code: 3PBJ)[2]. The binding site shown in a) top down view from the N-termini and b) side view. Main chains of the helices are colored in cyan. Cys and 12Leu residues are shown as sticks. Hg(II) atom is present in grey.

suggests that this Hg(II) ion is still in the plane of the three sulfurs. The Hg(II)-S bond lengths are 2.21 (chain A), 2.07(chain B) and 2.41(chain C) Å. The first two distances are relatively short compared to the distance of 2.43 Å observed in $Hg(II)(S_{Cys})_3^-$ and also to the most common Hg(II)-S distance reported from the tris-thiolate Hg(II) complexes (2.43 Å).[24–26] These resulting short Hg(II)-S bond distances in $Hg(II)(S_{Pen})_3^-$ structure come from the smaller sulfur pocket. It is reasoned that the thiols of the Pen ligands are not open wide enough (average $S\gamma$-$S\gamma$ distance of 3.84 Å) to allow for the Hg(II) atom to fit into the center with an agreement of 2.43 Å Hg(II) to sulfur separation. This ideal bond distance could be made if the Pen-γ-thiols of chain B and chain C move upward slightly and at the same time rotate to the side to orient away from the core center. With these modeled orientations (**Figure 2-26**), the metal site would subsequently enlarge to an average $S\gamma$-$S\gamma$ distance of 4.17 Å. These distances would then appear to be close to the $S\gamma$-$S\gamma$ distances observed in $Hg(II)(S_{Cys})_3^-$ structure (4.16 Å) and small molecule complexes.[24–26]

An overlay of the apo-Pen and bound-Pen structures shown in **Figure 2-27**, **a** strongly suggests that the S_γ atoms are almost identically positioned in space toward the core, while the $C_{\gamma 1}$, $C_{\gamma 2}$ atoms are oriented outward. The invariance of the metal ligands between metallated and apo structures is further underscored by comparison of χ_1 torsion angles (average) of unbound and bound forms (**Table 2-5**). This very tiny torsion angle shift results in almost equal $S\gamma$-$S\gamma$ separations of the two structures (3.71 versus 3.84 Å for apo-$(Pen)_3^-$ and $Hg(II)(S_{Pen})_3^-$ respectively). The highly similar sulfur planes emphasizes that apo-Pen ligands exhibit a high degree of *preorganization* for metal binding, which could be due to the rigidity caused by the bulky γ-methyl substitution that prevents the γ-thiol from moving freely through space. This subsequently limits the size of metal pocket, leading to the relatively short Hg(II)-S bond distances in $Hg(II)(S_{Pen})_3^-$.

Based on the actual observations, it is obvious that the thiols of the Pen ligands do not form a symmetric trigonal plane. To see the effect of steric hindrance more clearly, the binding site of $Hg(II)(S_{Pen})_3^-$ is overlaid onto $Hg(II)(S_{Cys})_3^-$ as shown in **Figure 2-28**. Apparently, the thiol orientations of the bound-Cys and bound-Pen point in completely opposite directions. While the Pen ligands are oriented to the interior, the Cys residues instead are directed further out toward the

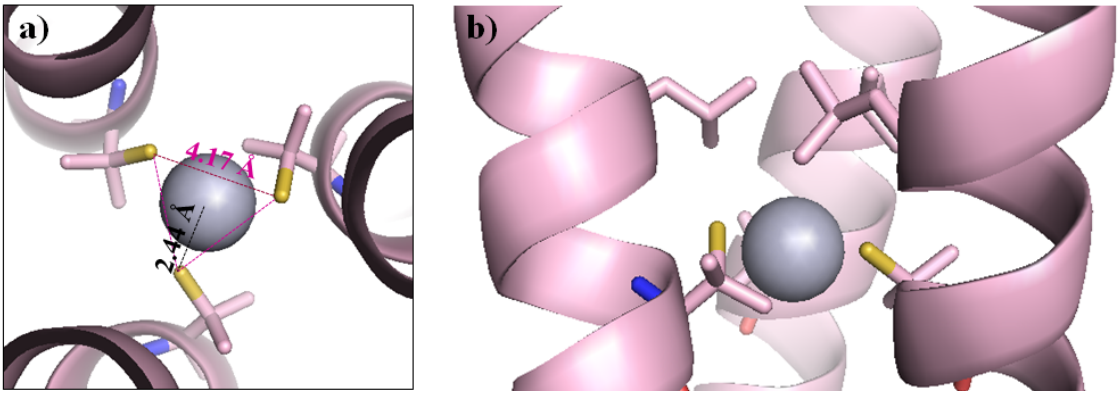

Figure 2-26. Model of Hg(II)(S-Pen)$_3$ using the published [Hg(II)]$_S$[Zn(II)(H$_2$O/ OH$^-$)]$_N$(CSL9PenL23H)$_3^{n+}$ crystal structure (PDB code: 3PBJ)[2] representing the binding site when the Hg(II) to the sulfur bond distances are modeled to be 2.44 Å. Shown in a) Side view and b) top down view from the N-termini of the model. Main chains of the helices are in pink. Hg(II) is grey . Cys and Leu side chains are shown in pink as sticks where sulfurs are labeled in yellow.

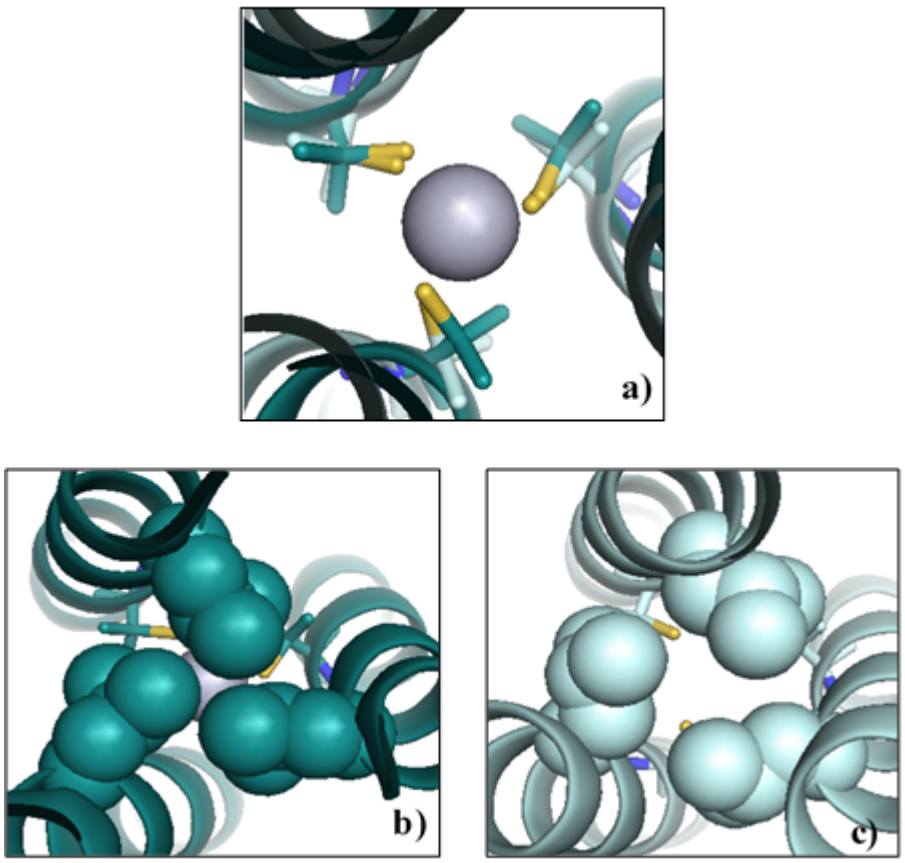

Figure 2-27. Structure comparison between the apo-$(CSL16Pen)_3$ and $[Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(CSL9PenL23H)_3^{n+}$(PDB code: 3PBJ)[2] structures shown from the top down view of the helices. Shown in a) an overlay of the Pen layer, b) 5Leu residues demonstrating the packing above the metallated 9Pen layer and c) 12Leu packing above the apo-16Pen layer. Main atoms are shown as ribbon diagrams in which the apo-structure is in grey and metallated-structure is in cyan. Pen residues are shown as sticks. Hg(II) in the metallated $Hg(II)]_S[Zn(II)(H_2O/OH^-)]_N(\mathbf{CS}L9PenL23H)_3^{n+}$ is represent as a grey sphere.

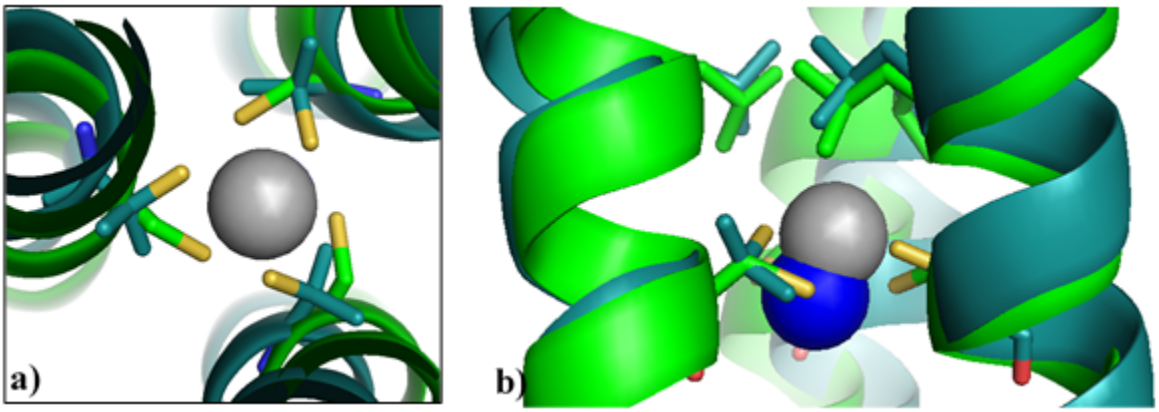

Figure 2-28. Comparison of the trigonal structures of Hg(II)(S-Cys)$_3^-$ from Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ and Hg(II)(S-Pen)$_3^-$ from the published [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$^-$)]$_N$(CSL9PenL23H)$_3^{n+}$ (PDB code: 3PBJ)[2]. Representing in a) side view and b) top down view from the N-termini of the overlay. Main atoms of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are colored in green and [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$^-$)]$_N$(**CS**L9PenL23H)$_3^{n+}$ in cyan. Cys and Pen side chains are shown as sticks where sulfurs are yellow. Hg(II) atoms in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$^-$)]$_N$(**CS**L9PenL23H)$_3^{n+}$ are present as blue and grey spheres, respectively.

helical interface as confirmed by the different χ_1 (N-Cα-Cβ-Sγ) torsion angles of the two ligands [-49.85^o observed in Hg(II)$(S_{Pen})_3^-$ and -151.88^o in Hg(II)$(S_{Cys})_3^-$]. The average γ-thiol separation in Pen is subsequently shorter than Cys by 0.32 Å, verifying that Cys can make a larger triangular metal plane compared to Pen. These observations again suggest that the γ-methyl groups of Pen inhibit the expansion of the three atom sulfur plane to the requisite distances that are optimal for trigonal Hg(II) species. It should be noted that this effect should work against forming a trigonal planar species for either Hg(II) or Cd(II). Therefore, another factor must be operative for Cd(II) to adopt the trigonal planar structure observed in Cd(II)(**CS**L16Pen)$_3^-$.

In contrast, totally different behavior is observed with Cys ligand construct. As discussed in *Section I* (**Figure 2-28**), the apo-Cys ligands (directed upward to the N-termini and to the core of the structure) have to reorient their γ-thiols downward to the C-termini and face out to the helical interface in order to accept Hg(II) into the metal plane. Quantitatively, the change of χ_1 torsion angles was shown to make a huge difference in S_γ-S_γ separations; however, equally important, is the change in position relative to the N-termini between the thiol planes. The Hg(II)-Cys plane is shifted 1.46 Å down toward the C-termini plane with respect to the apo-Cys structure. Moreover, a PyMOL visualization of the side on view of the aligned Hg(II)$(S_{Cys})_3^-$ and Hg(II)$(S_{Pen})_3^-$ binding sites (**Figure 2-28, b**) demonstrates that the heights of the metallated planes are also shifted. The Hg(II)$(S_{Pen})_3^-$ is positioned about 1.80 Å toward the N termini relative to Hg(II)$(S_{Cys})_3^-$. I propose that the absence of the γ-methyl groups around the β-carbon allows for more S_γ free rotation through space, allowing for a huge change in the Cys rotamers between un-metallated and metallated coordination environments. In contrast, the inclusion of the bulky groups in Pen restricts the S_γ rotation upon metal binding yielding an opposite effect as indicated by the similar thiol orientations between apo-$(Pen)_3$ and Hg(II)$(S_{Pen})_3^-$ structures. Most important, the immobility of the Pen side chains requires the metal to bind further toward the N-termini and closer to the 12Leu layer. It is likely that this effect restricts water access to the hydrophobic cavity.

It is proposed that if the packing of 12Leu layers between the Hg(II)$(S_{Cys})_3^-$ and Hg(II)$(S_{Pen})_3^-$ remain unchanged while the metallated-Pen plane is higher than in the bound-Cys form, then a smaller space for water above the Pen site will be available. As illustrated in **Figure 2-29**, the packing of the 12Leu layer in Hg(II)$(S_{Pen})_3^-$ is in fact slightly tighter as the inner Cδ-Cδ

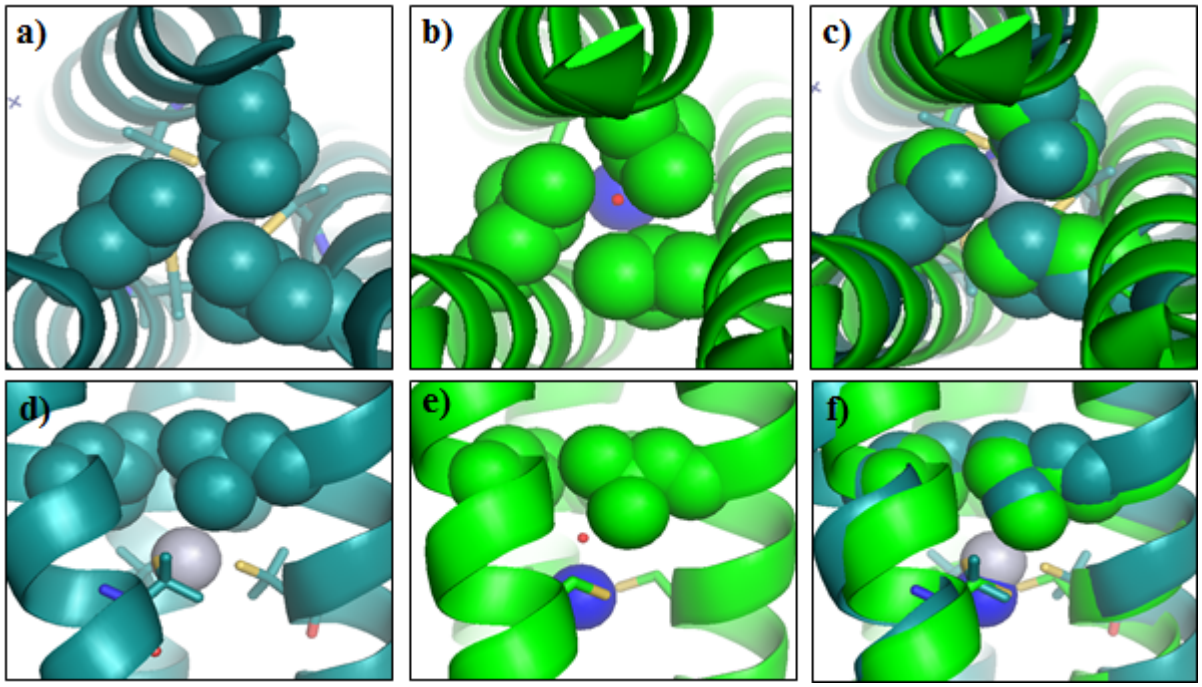

Figure 2-29. Packing comparison of Leu residues above the trigonal Hg(II) sites of Hg(II)(S_{Cys})$_3^-$ from Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ and Hg(II)(S_{Pen})$_3^-$ from the published [Hg(II)]$_S$[Zn(II)(H_2O/OH^-)]$_N$(CSL9PenL23H)$_3^{n+}$ (PDB code: 3PBJ)[2]. Top panels: Representing a top down view from the N-termini of a) 5Leu residues above the Hg(II)(S_{Pen})$_3^-$, b) 12Leu residues above the Hg(II)(S_{Pen})$_3^-$ and c) an overlay of a) and b). Bottom panels: Representing a side view of d) 5Leu residues above the Hg(II)(S_{Pen})$_3^-$, e) 12Leu residues above the Hg(II)(S_{Pen})$_3^-$ and f) an alignment between d) and e). Main chains of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are colored in green and [Hg(II)]$_S$[Zn(II)(H_2O/OH^-)]$_N$(**CS**L9PenL23H)$_3^{n+}$ in cyan. Cys and Pen side chains are shown as sticks where sulfurs are yellow. Hg(II) atoms in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and [Hg(II)]$_S$[Zn(II)(H_2O/OH^-)]$_N$(**CS**L9PenL23H)$_3^{n+}$ are present as blue and grey spheres, respectively.

separation is only 3.60 Å. This indicates stronger Van Der Waal interactions in the helical core occur in $Hg(II)(S_{Pen})_3^-$ rather than in $Hg(II)(S_{Cys})_3^-$ where the inner Cδ-Cδ separation is 4.30 Å. The most significant observation to support hypothesis 3 is that the 12Leu layer of $Hg(II)(S_{Pen})_3^-$ is only separated by 4.86 Å from the bound-16Pen plane, whereas the related distance determined from the $Hg(II)(S_{Cys})_3^-$ lengthens to 6.07 Å. The impact of this difference in interlayer spacing on water access is supported by the absence of a water molecule in the area between 12Leu and 16Pen in the $Hg(Pen)_3^-$ structure.[2,47,48] The aligned binding sites (**Figure 2-29, bottom panels)** also shows that the coordinated Hg(II) in $Hg(II)(S_{Pen})_3^-$ occupies the space where the water was previously observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$.

I will now address whether the rules of Hg(II) binding to **CS**L16C derivatives can be transferred to complexation of Cd(II). When Cd(II)(**TRI**L16C)$_3^-$ is formed, one observes 60% of the centers with bound water [$Cd(II)S_3O^-$] and the remainder is $Cd(II)S_3^-$ [27,30], while Cd(II)(**TRI**L16Pen)$_3^-$ has an exclusive $Cd(II)S_3^-$ structure and Cd(II)(**TRI**L12AL16C)$_3^-$ yields solely $Cd(II)S_3O^-$.[28,29] I will first address the formation of $Cd(II)S_3O^-$and then consider $Cd(II)S_3^-$.

Based on studies by Alber who examined water entrapment upon replacement of hydrophobes in 4-stranded α-helical bundles[49], Lee *et al*. substituted Ala for the more bulky Leu at the twelfth layer to provide a water pocket above the metal site in (**TRI**L12AL16C)$_3$.[28,29] While a structure of Cd(II)(**TRI**L12AL16C)$_3^-$ has not been obtained, inferences about this system can be obtained using two related metallated structures, Hg(II)(**GRAND-CS**L12A16C)$_3^-$ and Zn(II)(**GRAND-CS**L12A16C)$_3^-$. The Hg(II)(**GRAND-CS**L12A16C)$_3^-$ structure is used to assess how a smaller amino acid perturbs the steric hindrance of Leu at the twelfth position compared to $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ using a metal that has nearly comparable M-S bond distances that are expected for Cd(II)-S four coordinate complexes. Next, Zn(II)(**GRAND-CS**L12A16C)$_3^-$is examined to reveal the general requirements for a sulfur pocket prepared with a 4-coordinate metal to provide some structural restrictions for $Cd(II)S_3O^-$.

To see the effect of 12Ala compared to 12Leu, the metallated structure of Hg(II)(**GRAND-CS**L12A16C)$_3^-$ is overlaid onto the previous $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ (**Figure 2-30**). In both cases Hg(II) is found as a trigonal planar structure. The three Cys residues in Hg(II)(**GRAND-CS**L12A16C)$_3^-$ are symmetric and apparently only the major conformer (**Figure 2-8b**) can bind the metal with an orientation toward the helical interface ($\chi_1 = -147.85^o$). This χ_1

-

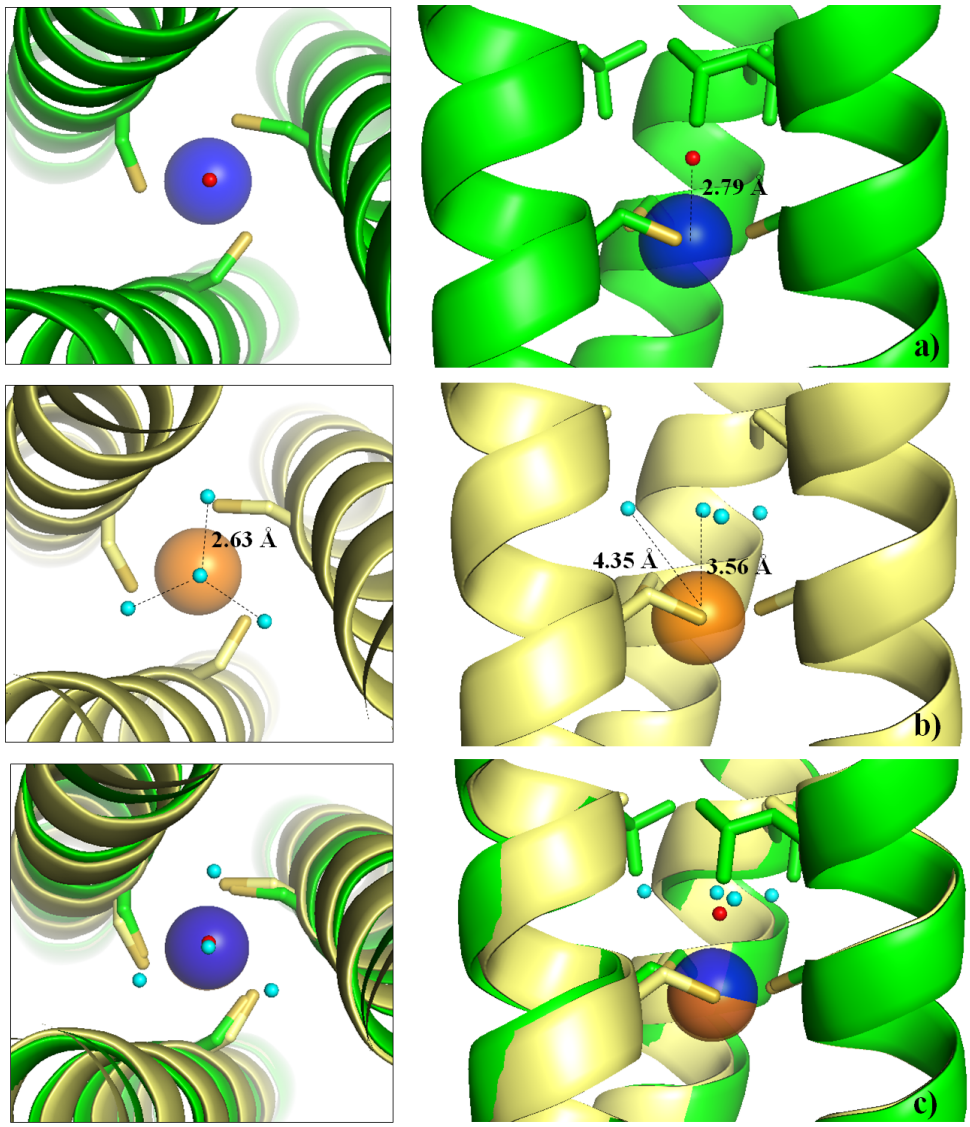

Figure 2-30. Comparison of the trigonal Hg(II)$(S)_3^-$ binding sites and the numbers of waters observed between Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ and Hg(II)(GRAND-CSL12AL16C)$_3^-$. Shown in a) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, b) Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ and c) the overlays between top-down and side views of the two structures. Main chains of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are colored in green and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ in pale yellow. Cys, Leu and Ala side chains are shown as sticks. Sulfurs atoms are bright yellow. The Hg(II) atoms in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are present as blue and orange spheres, respectively. The observed waters in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown as red and cyan small spheres.

value is close to the -151.88º observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ indicating that the orientations of the bound Cys ligands in both structures are similar. Such arrangements make the metal pocket sizes comparable as the $S\gamma$-$S\gamma$ separations of the two structures are in the same range, 4.20 Å for Hg(II)(**GRAND-CS**L12A16C)$_3^-$ and 4.16 Å for $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. Hg(II) is situated at a distance of 0.26 Å below the 16Cys plane with a Hg(II)-S distance of 2.44 Å and S-Hg(II)-S angle of 118.74º , values that correspond closely to those observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. Taken together these parameters confirm that both designs show essentially identical first coordination environments for $Hg(II)S_3^-$, therefore, we can expect that the change in the $Cd(II)S_3O^-$ to $Cd(II)S_3^-$ ratio is not a consequence of the first coordination sphere, but rather depends on factors associated with the second coordination sphere that surrounds the metal pocket of Hg(II)(**GRAND-CS**L12A16C)$_3^-$.

Unlike leucine, alanine contains a single methyl group (Cβ carbon) attached to the α-carbon. For this reason, alanine fills much less space in the position above the metal. Alber had shown that conversion of leucine to alanine allowed for the addition of four waters into the cavity generated in 4-helix bundles by removing the leucine isopropyl groups.[49] Though a narrower construct, 3SCCs might be expect to behave similarly. The PyMOL representation in **Figure 2-31** emphasizes the steric hindrance generated from the aliphatic isobutyl side chain of Leu compared to the methyl group of Ala. It is obvious that Ala opens a huge space generating a hole above the metal site confirming the proposed impact of the modification.[29] As a consequence, the larger space in Hg(II)(**GRAND-CS**L12A16C)$_3^-$ allows for up to four waters to access the metal binding site (**Figure 2-9**), an observation consistent with Alber's previous studies. As discussed below, a similar conclusion may be drawn when Zn(II) coordinates to an L12AL16C derivative (**Figure 2-7**). In contrast, in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$, only a single, unbound water that sits on the threefold axis directly above the metal at a 2.79 Å distance is observed (**Figure 2-30**). In Hg(II)(**GRAND-CS**L12A16C)$_3^-$, one of the waters behaves in the same way as the observed water in the 12Leu-structure. Indeed, it is again located on a threefold axis at a non-bonding distance of 3.56 Å from Hg(II). Moreover, the other three waters are threefold related, but located close to the helical interface between two neighboring strands with a distance of 4.35 Å from the Hg(II). These waters form a hydrogen bonding network and are separated by a distance of 2.63 Å from the first water. Each solvent is found within the same plane (with respect to the N-termini).

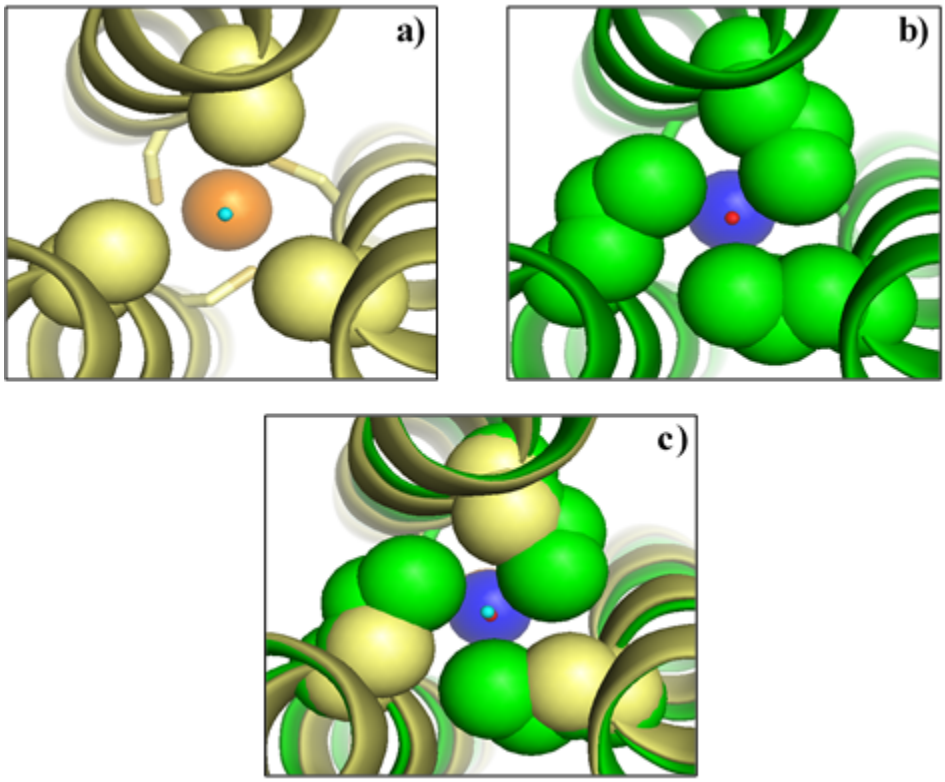

Figure 2-31. Packing comparison of hydrophobic layers at the twelfth position shown as spheres above the metal binding sites of Hg(II)(GRAND-CSL12AL16C)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$. From top down view of the N-termini, representing the 12Leu residues of a) Hg(II)(**GRAND-CS**L12AL16C)$_3^-$, and b) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. In c) showing as overlay between a) and b). Main chains of Hg(II)(**GRAND-CS**L12AL16C)$_3^-$and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in pale yellow and green, respectively. Cys residues are shown as sticks with sulfurs atoms being bright yellow. The Hg(II) atoms in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are present as blue and orange spheres, respectively. The observed waters in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown as red and cyan small spheres.

Such distances of Hg(II) to water are too long to be a Hg(II)-O bond (predicted to be ~2.2 Å), therefore, all of the waters found within the cavity are considered to be uncoordinated and stabilized through H-bonding interactions between each other. Another compelling point to support the large size of the cavity formed with 12Ala is the observation of a longer Hg(II) to the water distance in Hg(II)(**GRAND-CS**L12A16C)$_3^-$ (3.56 Å) as compared to the related distance of 2.79 Å in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. This 0.77 Å increase in Hg(II)-O separation clearly demonstrates that more space is available for the water to move upward in the 12Ala-containing structure.

Generally, to anchor waters within the hydrophobic environment, the sulfur *p*-orbitals in the sixteenth position are exploited to mediate H-bonding interactions. This is confirmed crystallographically as the Cys residues in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ structure use both major and minor conformers (both directed to the helical interface) to form H-bonds with the observed water (S-O distances of 3.39 and 3.97 Å, respectively) (**Figure 2-32**). However, when the water shifts upward to the N-termini in Hg(II)(**GRAND-CS**L12A16C)$_3^-$, this situation causes the water to move 4.08 Å farther away from the major conformer of Cys. The three coordinated Cys sulfur atoms, at a distance of 4.08 Å, are too far away to form good H-bond contacts. Instead, the additional three waters that occupy the cavity stabilize the core water as each is only 2.63 Å away. Such distance promises to make a strong H-bonding network among waters. In addition, each one is also 2.58 Å away from the minor and 3.25 Å from the major Cys residues indicating that both Cys rotamers are capable of stabilizing the H-bonding network within the hydrophobic region above the metal site of Hg(II)(**GRAND-CS**L12A16C)$_3^-$. The combination of these stabilizing interactions and the availability of more space generated by 12Ala rationally explain why the Hg(II) to the central water distance is longer in Hg(II)(**GRAND-CS**L12A16C)$_3^-$ compared to Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. Therefore, this particular analysis additionally emphasizes that the larger space generated by 12Ala does not only provide more possibility of water access, but it also allows for Cys rotation (minor conformers) by facilitating the H-bonding interactions to stabilize waters above the metal site.

The different number of water molecules observed between Hg(II)(**GRAND-CS**L12A16C)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ can explain the different degrees of solvation of Cd(II) in L12A16C and L16C designs. The metallated Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ contains a small hydrophobic pocket that is able to house a single water molecule

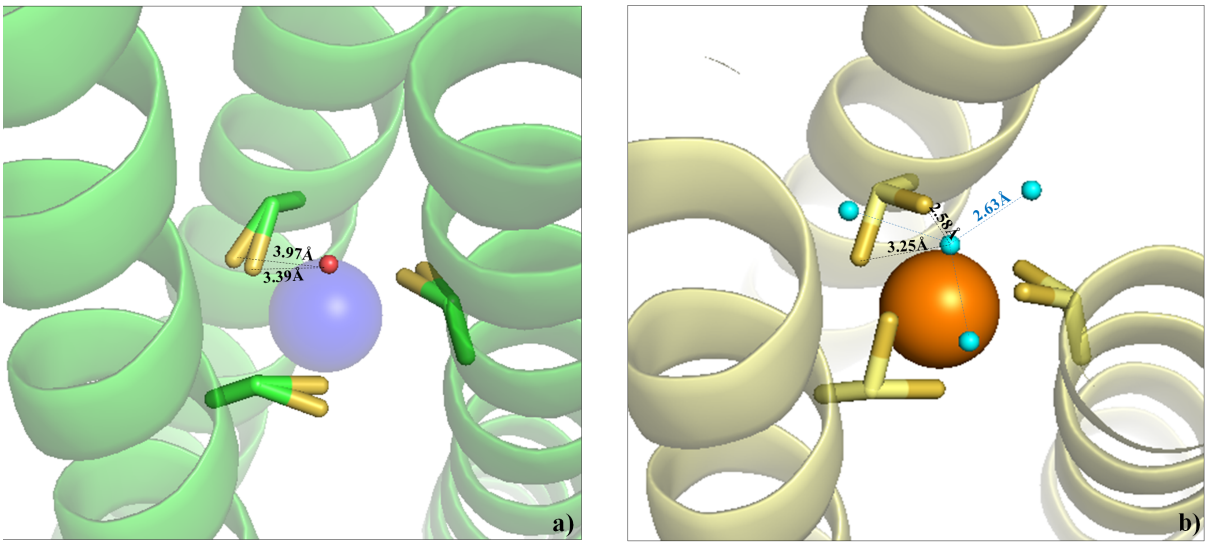

Figure 2-32. Distances of the core water(uncoordinated) from the Cys conformers in a) Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ and b) Hg(II)(GRAND-CSL12AL16C)$_3^-$ structures. Main chains of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown in green and pale yellow, respectively. Cys residues are shown as sticks with sulfurs atoms being bright yellow. The Hg(II) atoms in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are present as blue and orange spheres, respectively. The observed waters in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown as red and cyan small spheres.

at a relatively short Hg(II) to O distance of 2.79 Å. This observation proves that a cavity for solvent exists and it may allow water access when Cd(II) is bound to the metal site. Furthermore, EXAFS data indicate that a Cd(II)-O bond in a $Cd(II)S_3O^-$ structure is 2.32 Å, a distance that would appear achievable based on the Hg(II) water separation. However, one must remember that Hg(II) forms a trigonal planar structure whereas Cd(II) would have a 4-coordinate pseudo tetrahedral polyhedron. This means that the Cd(II) would need to be displaced above the three S atom plane toward the solvent ligand. As the spectroscopic data indicated that only 60% of $Cd(II)S_3O^-$ is present in Cd(II)(**TRI**L16C)$_3^-$, this suggests that in a four coordinate structure the cavity may not be capable of stabilizing water in the hydrophobic core. In the L12AL16C design, this steric restriction is no longer operative even for a four coordinate complex. Thus, $Cd(II)S_3O^-$ is the only species that forms in Cd(II)(**TRI**L12AL16C)$_3^-$. I will now discuss the binding of the four coordinate metal Zn(II) in order to understand the constraints on water binding in Cd(II)(**TRI**L16C)$_3^-$.

Unfortunately, a crystal structure of $Cd(II)S_3O^-$ in (**GRAND-CS**L12AL16C)$_3^-$ has not been achieved; however, I have been successful in solving Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ which contains the 4-coordinate Zn(II) ion. While Zn(II) is smaller than Cd(II), it is still hoped that some general information can be obtained for a 4-coordinate metal structure when the twelfth layer is perturbed by 12Ala. The overall binding environment is shown in **Figure 2-7.** The bound Zn(II) is situated on the threefold axis. The metal is coordinated to four ligands, three of which are from the major S_γ conformer of three 16Cys residues and the fourth ligand comes from a water molecule. This observed water sits on top of the structure and is also constrained by the threefold at a distance of Zn(II)-O 2.17 Å directly above the Zn(II) center. The Zn(II)-S distance is 2.27 Å. The S-Zn(II)-S and S-Zn(II)-O angles are 119.84^o and 92.13^o, respectively. These values actually do not follow the perfect tetrahedral geometry angle of 109.5^o indicating that the first coordination sphere of this $Zn(II)(Cys)_3(H_2O)^-$ structure is quite distorted. One factor is because the bound Zn(II) is situated relatively close to the sulfur plane (at a distance of 0.09 Å above the plane) which affects the S-Zn(II)-O and S-Zn(II)-S angles. It is found that to make an agreement of these particular angles and metal-ligand distances, the χ_1 dihedral angle of the bound Cys rotamer has to be -146.96^o, which makes the S_γ positioned toward the helical interface. The separation between the two S_γ atoms is at a distance of 3.94 Å in order to bind Zn(II). Interestingly, the effect of 12Ala is also shown in the pocket above the bound 16Cys layer of Zn(II)(**GRAND-CS**L12A16C)$_3^-$ because

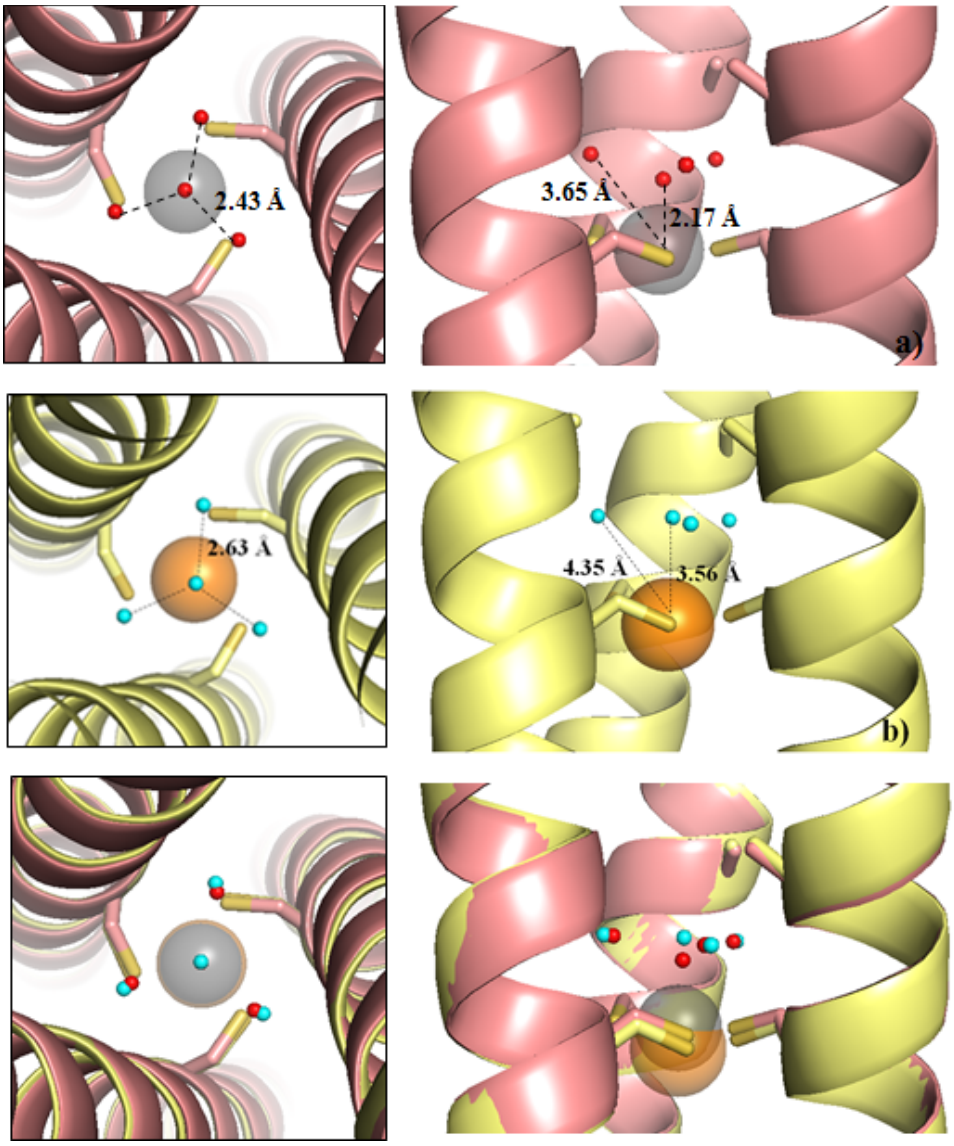

Figure 2-33. Aligned structures demonstrating the observed waters above the metal sites between the Zn(II)(Cys)$_3$(H$_2$O)$^-$ from Zn(II)(GRAND-CSL12AL16C)$_3^-$ and Hg(II)(S)$_3^-$ from Hg(II)(GRAND-CSL12AL16C)$_3^-$. a) Zn(II)(**GRAND-CS**L12AL16C)$_3^-$, b) Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ and c) the overlays of the top-down and side views of the two structures. Main chains of Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ are colored in pink and Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ in pale yellow. Sulfurs atoms are bright yellow. The Zn(II) and Hg(II) are grey and orange, respectively. The observed waters in Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ are red and in Hg(II)(**GRAND-CS**L12AL16C)$_3^-$ are cyan.

there are again three waters filling out the cavity and mediating H-bonding interactions apart from the bound-Zn(II) water that is located on the threefold axis (total number of waters is four). All of the side waters are observed in the same manner as Hg(II)(**GRAND-CS**L12A16C)$_3^-$. Three of them are again crystallographically threefold related and unbound to Zn(II) at a distance of 3.65 Å. Their locations are close to the helical interface (**Figure 2-33, b**) between two adjacent helices with a 2.43 Å separation from the bound water. The plane generated from these side waters is actually higher than the central water with respect to the N-termini. The overlay of Zn(II)(**GRAND-CS**L12A16C)$_3^-$ with the Hg(II)(**GRAND-CS**L12A16C)$_3^-$ structure illustrates that the side waters are comparable in height (**Figure 2-33**) (a little higher and further out to the helical interface in the Hg(II)-structure) but that the central water coordinates are not overlaying. The lower height of the central water observed in the Zn(II)-bound structure compared to Hg(II)(**GRAND-CS**L12A16C)$_3^-$ strongly emphasizes the requirement of the Zn(II)-O distance needed to achieve for the 4-coordinate Zn(II) structure.

The consistent number of waters determined between the two metallated structures of L12AL16C derivatives demonstrates that the cavities generated above the sixteenth layer are, independent of metal, comparable in size. The following observations confirm this conclusion. First, the directions of the C_β atoms of 12Ala between the two structures are similar as shown in **Figure 2-15** with close values of C_β-C_β separation (5.94 versus 5.86 Å for Zn(II)(**GRAND-CS**L12A16C)$_3^-$ and Hg(II)(**GRAND-CS**L12A16C)$_3^-$, respectively). Second, the major conformer of Cys residues of the two structures are positioned through space almost with nearly identical χ_1 angles indicating that the height of the sulfur planes with respect to the N-termini of the two structure are very close. These two observations, combined with the presence of four waters in the cavity in both structures, strongly supports the contention that the designed protein defines the extent of water access rather than the chosen metal.

With the Zn(II)(**GRAND-CS**L12A16C)$_3^-$ structure available, I can now attempt to model the coordination environment for Cd(II) in an Cd(II)S_3O^- environment. Combining the Zn(II)(**GRAND-CS**L12A16C)$_3^-$ structural motif with the Cd(II)-S and Cd(II)-O distances of Cd(II)(**TRI**L12AL16C)$_3^-$ from the X-ray absorption data (**Table 2-6**), I have generated the proposed structure shown in **Figure 2-34**. Cd(II) is bound to the metal site with three sulfurs (2.50 Å) and a water (2.37 Å). To achieve these distances, the S_γ atoms have to rotate slightly outward to the helical interface because the Cd(II)-S distance (2.50 Å) is longer than Zn(II)-S (2.27 Å).

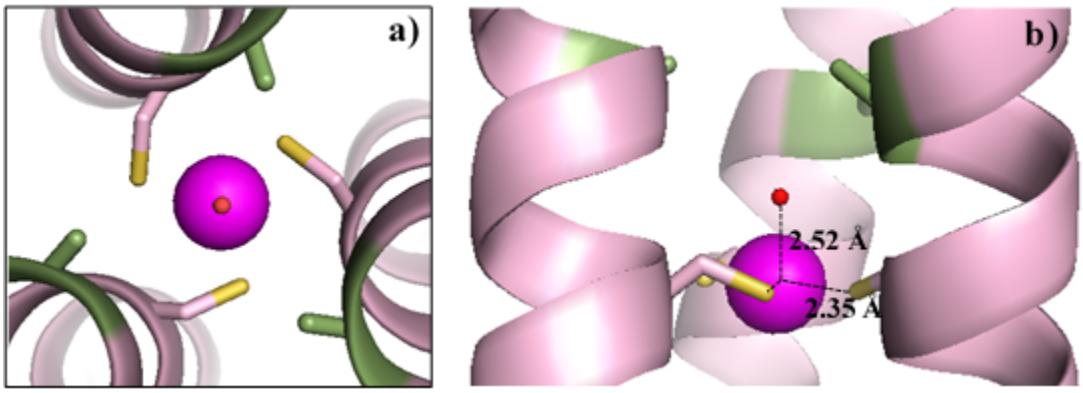

Figure 2-34. Model of $Cd(II)S_3O^-$ using $Zn(II)(GRAND\text{-}CSL12AL16C)_3^-$ crystal structure. a) Top down view from the N-termini and b) side view of the binding site representing Cd(II)-O and Cd(II)-S distances of 2.35 and 2.52 Å, respectively. Cd(II) is modeled as a pink sphere. 12Ala and 16Cys residues are shown in as sticks (sulfurs = yellow). The modeled water is in red. Backbones of the peptide helices are shown as pink ribbon diagrams.

This rotation causes the χ_1 dihedral angle of Cys to change from -146.96° to -154.26°. It should be noted that the χ_1 dihedral angle of Cys in the Hg(II) trigonal plane is -151.88° so that this expansion is within reasonable tolerance for the cysteine conformers. The central water is moved upward by 0.2 Å. The final model shows the S-Cd(II)-S and S-Cd(II)-O angles of 119.24° and 93.29°.

An alternative strategy to model a $Cd(II)S_3O^-$ structure is to begin with the $Hg(II)S_3^-$ structure and perturb it to $Cd(II)S_3O^-$ structural constraints. Presumably if the water coordinates were fixed to where it is located in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure, the Cd(II) would likely move out of the trigonal Cys plane toward the water. This would require a shift of approximately 0.5 Å of the Cd(II) ion toward the N-termini. In this case, the modeled Cd(II) would be 2.35 Å from the coordinated water, consistent with the Cd(II)-O bond distance determined from the EXAFS of a fully $Cd(II)S_3O^-$ formation in **TRI**L12AL16C (**Table 2-6**). Now that Cd(II) is above the Cys plane, the χ_1 torsion angle of Cys has to shift from -151.88° to -156.07° in order to keep the EXAFS derived Cd(II)-S bond length for 4-coordination. This results in a change of S_γ orientation from pointing to the helical interface, downward to the C-termini to directing slightly upward to the N-termini. The model requires a S_γ-S_γ separation of 4.39 Å, while 4.16 Å is required for the trigonal planar Hg(II). This suggests that a tetrahedral structure may require a slightly bigger sulfur pocket than a trigonal structure in the 3SCC environments. The final model of the first species is shown in **Figure 2-35** shows a pseudo-tetrahedral $Cd(II)S_3O^-$ structure with S-Cd(II)-S and S-Cd(II)-O angles of 119.84° and 91.84°, respectively. The most important conclusion of this analysis is that a reasonable and consistent model for $Cd(II)S_3O^-$ can be obtained starting either from the 4-coordinate Zn(II) or 3-coordinate Hg(II).

Using these models, we can come to the following conclusions. First, solvent can gain access to the Cd(II) in L16C, but at a potentially significant price. First, the only factor stabilizing the water in this environment is the coordination bond to Cd(II). This is in marked contrast to the situation with L12AL16C where the additional three waters can H-bond both to the bound water and the coordinated sulfur atoms. Thus, the presence of leucine significantly destabilizes the water in the core of the 3SCC by inhibiting favorable H-bonding contacts. Second, the movement of the Cd(II) toward the hydrophobic layer at the 12Leu as it forms a four coordinate complex, decreases the volume in which water can bind to the metal. Thus, there are direct steric constraints that inhibit solvent interactions in the core. Moreover, movement of the leucine methyl groups to

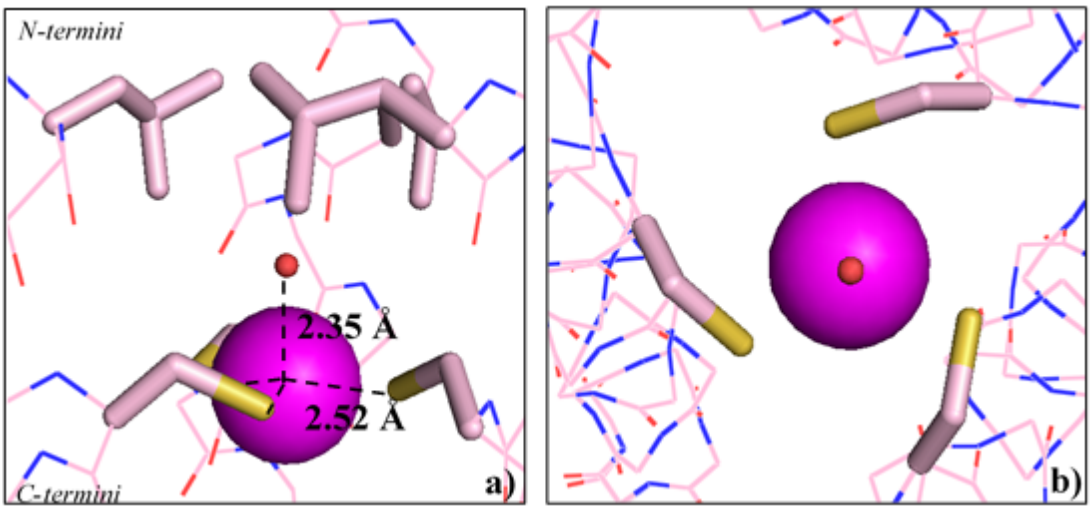

Figure 2-35. Model of $Cd(II)S_3O^-$ using $Hg(II)_SZn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$ crystal structure. a) Side view of the binding site representing Cd(II)-O and Cd(II)-S distances of 2.35 and 2.52 Å, respectively. b) Top down view from the N-termini of the model. Cd(II) is modeled as a pink sphere. Cys side chains are shown in purple as sticks where sulfurs are labeled in yellow. The modeled water is in red. Backbones of the peptide helices are shown as lines (the red color refers to oxygen atoms and blue refers to nitrogen atoms).

accommodate the coordinated water will destabilize the 3SCC by diminishing hydrophobic contacts in the core. Finally, complexation of Cd(II) in a 4-coordinate structure requires that the three atom sulfur plane expand significantly from the original apo-conformation of the cysteine thiols. By retaining a three coordinate structure, less conformational rearrangement is required to bind the metal. All three factors work in conjunction to bias the Cd(II) ion toward the normally less favorable 3-coordinate geometry. The final model shown in **Figure 2-36** displays a trigonal $Cd(II)S_3^-$. Here, the modeled Cd(II) can remain in the position observed for trigonal Hg(II) in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$; however, in order to reach an agreement of 2.46 Å Cd(II)-S bond distance based on EXAFS experiments for 100% 3-coordinate, the S_γ atoms need to open up toward the helical interface ~0.3 Å. The χ_1 dihedral angle changes to -154.03° with a longer S_γ-S_γ separation of 4.27 Å and the S-Cd(II)-S angles of 118.10°. As a result, energy penalty to internalize the water molecule into the hydrophobic cavity between the 12Leu and 16Cys layers is too great leading to the observed 40% trigonal planer $Cd(II)S_3^-$ formation and 60% $Cd(II)S_3O^-$.

At this juncture, we can now understand the structural influences that control the binding of water to Cd(II) in L16C and L12AL16C type peptides; however, equally interesting are the factors that exclude water from the cadmium coordination environment to yield $Cd(II)S_3^-$ systems. We have learned that in L16C and L12AL16C type peptides, Hg(II) induces a high degree of S_γ rotation of the Cys residues causing a significant lowering of the metal plane toward the C-termini, which expands the hydrophobic cavity above the sulfur plane sufficiently to accommodate a water molecule. This model may also be operative for Cd(II) binding to these peptides. To test this, a water was modeled at the same distance (2.79 Å) above the coordinated Hg(II) in $Hg(II)(S_{Pen})_3^-$ as seen for Hg(II)-water distance observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. To attain this distance, the modeled water must be placed almost at the same height as the interior C_δ plane of 5Leu (**Figure 2-37**). The MolProbity clash score (6.20) confirms that these distances are too close to prevent steric clashing between the 5Leu residues and the water. However, if this water were to sit at this modeled coordinate, the 5Leu residues would have to open wider by orienting their side chains toward the helical interface. This situation is unlikely, as the 5Leu layer would lose too many hydrophobic contacts, destabilizing the entire 3SCC. In $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ the water is 3.80 Å below the 12Leu plane . It is obvious that Pen cannot generate enough space to accommodate a polar molecule within the hydrophobic core above the metal site. Therefore, all these crystallographic observations support the third hypothesis that had been

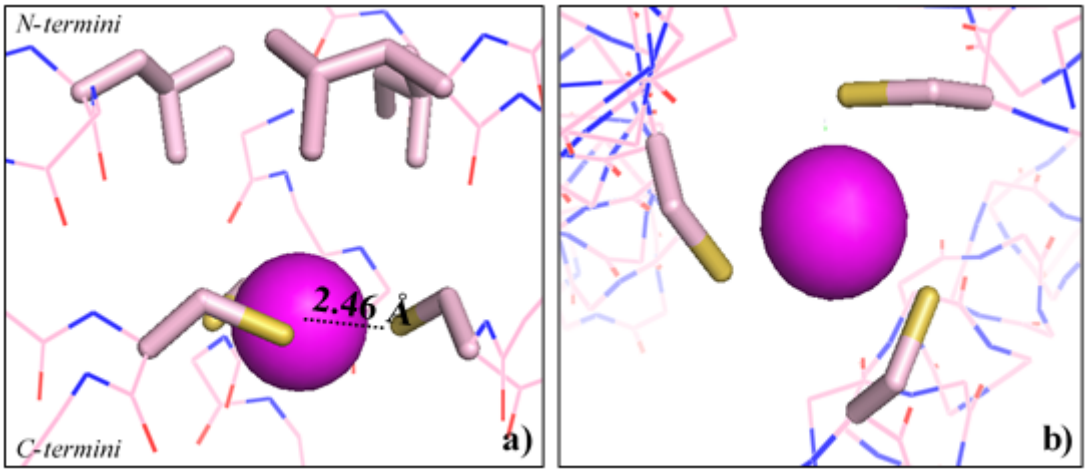

Figure 2-36. Model of $Cd(II)S_3^-$ using the $Hg(II)_S Zn(II)_N(GRAND\text{-}CSL16CL30H)_3^+$ crystal structure. a) Side view of the binding site representing a Cd(II)-S distance of 2.46 Å. b) Top down view from the N-termini of the model. Cd(II) is modeled as a pink sphere. Cys side chains are shown in purple as sticks where sulfurs are labeled in yellow. Backbones of the helices are shown as lines (red color referred to oxygen atoms and blue referred to nitrogen atoms).

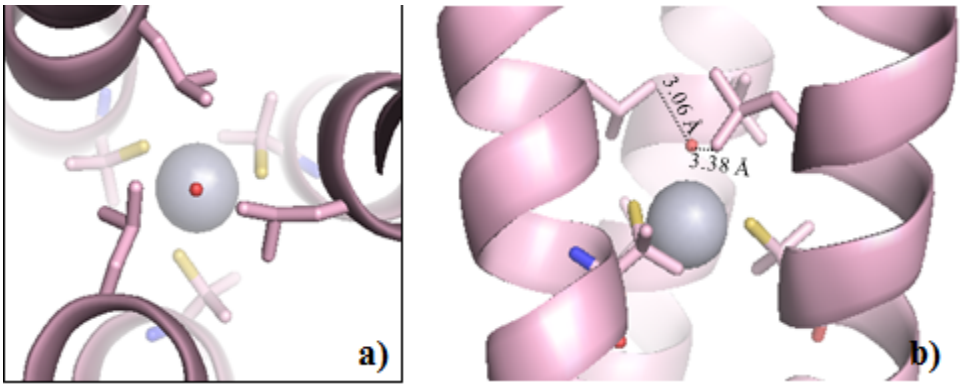

Figure 2-37. Model of a water placed above the Hg(II)(S-Pen)$_3$ binding site of the published [Hg(II)]$_S$[Zn(II)(H$_2$O/OH$^-$)]$_N$(CSL9PenL23H)$_3^{n+}$ structure (PDB code: 3PBJ)[2] representing possibility of steric clashes that could happen between 5Leu residues and the modeled water. a) Top down view from the N-termini showing the distances of the $C_{\delta 2}$-water = 3.06 Å and $C_{\delta 1}$-water = 3.38 Å. b) Side view of the binding site. Backbones of the helices are shown as ribbons where Pen and 5Leu residues are represented as sticks (sulfur = yellow). Hg(II) is a grey sphere. The modeled water is in red.

formulated. The restriction of metallated-Pen ligands shows an impact directly on the metal plane and influences tighter hydrophobic packing above the metal site, subsequently generating less space for solvents. It is believed that the Pen ligands will similarly impose these constraints on the Cd(II) center and strongly supports a model for complete water exclusion (fully $Cd(II)S_3^-$ formation) in Cd(II)(**TRI**L16Pen)$_3^-$. This hypothesis has been confirmed by the corresponding ^{113}Cd NMR and ^{111m}Cd PAC results.[29]

<u>Section III</u>: Structural comparisons between trigonal planar Hg(II) and trigonal pyramidal Pb(II) and As(III) using designed three-stranded coiled coils

I will now discuss the trigonal pyramidal $Pb(II)S_3^-$ center observed in $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ to develop further insight on the structural requirements for toxic metal binding to metalloregulatory proteins and metallochaperones such as PbPbr691.

The binding site of $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ shown in **Figure 2-10** represents a homodirected trigonal pyramidal geometry of $Pb(II)S_3^-$. Unlike the bound Cys orientation of the trigonal $Hg(II)S_3^-$ in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure that points outward to the helical interface, the major Cys rotamer binds to the metal by directing the side chain at the core of the structure and toward the N-termini with the observed χ_1 value of -68.38°. This χ_1 angle is almost identical to the value of -69.20° observed in the major conformation of apo-(**CS**L16C)$_3$ (**Figure 2-38**), implying that the orientation of Cys in the apo-peptide is highly *preorganized* for a trigonal pyramidal structure rather than the trigonal planar site observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. Subsequently, it results in close values of $S\gamma$-$S\gamma$ separation between the apo-(**CS**L16C)$_3$ (3.29 Å) and the metallated $Pb(II)S_3^-$ structure (3.49 Å). Due to the crystallographic symmetry constraints imposed by the R32 space group, Pb(II) sits on the threefold axis yielding a Pb(II)-S bond distance of 2.65 Å. This separation is within the error of the values determined from X-ray absorption spectroscopy (2.63 Å)[27] and also consistent with the Pb(II)-S bond distance observed in published reports.[25,50–53] The S-Pb(II)-S angle is 82.37°. The Pb(II) ion is located at distances of 1.69 Å below the S_γ plane toward the C-terminus of the 3SCC, forming an *endo* configuration. Actually this type of conformation was previously observed in the $As(III)S_3$ site of As(III)(**CS**L9C)$_3$.[1] Touw *et al.* discovered that the lone pair of As(III) may play an important role in influencing the As(III) coordination sphere within the 3SCC

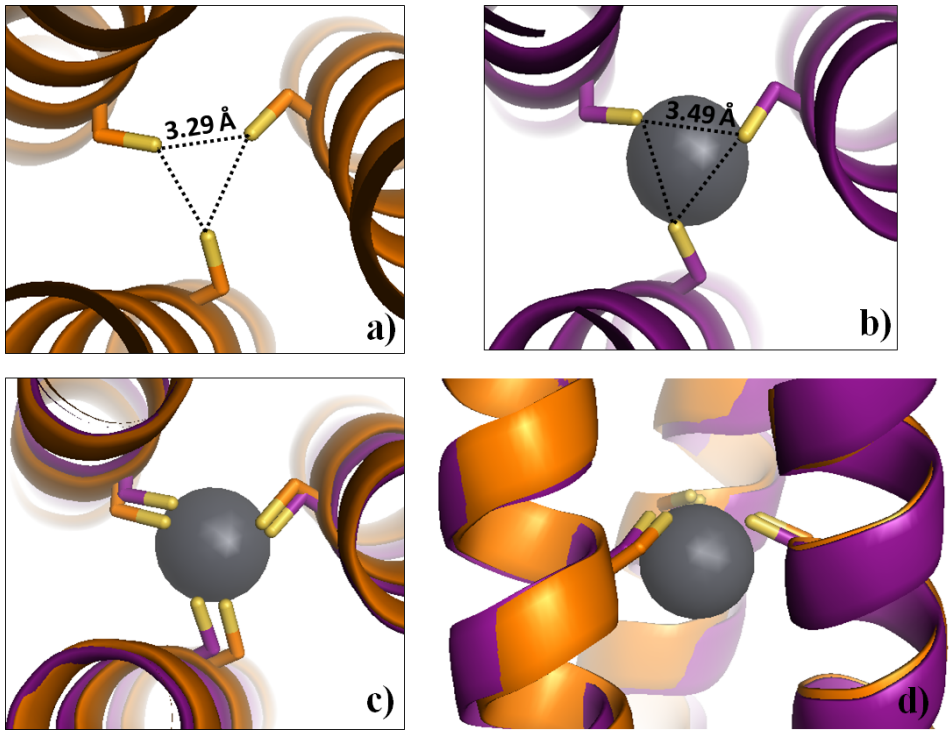

Figure 2-38. An overlay demonstrating the trigonal pyramidal structure comparison between $Pb(II)(S)_3^-$ from $Pb(II)_SZn(II)_N$(GRAND-CSL16CL30H)$_3^+$and the apo-$(CSL16C)_3$ structure. In a) representing top down view from the N-termini and b) side view of the binding sites. Main chains of $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$are colored in purple and apo-(**CS**L16C)$_3$ in orange. Sulfurs atoms are bright yellow.

environment.[1] It was reasoned that the poorer packing of the 12Leu residues (one hydrophobic layer below the 9Cys site) compared to the tighter packing of 5Leu residues that are more accommodated this lone pair toward the C-termini. In the *endo* conformation the As(III) can orient the lone pair to the C-terminal side of the metal ion so that the lone pair will not suffer from steric clashes with the 5Leu side chains. Comparably to As(III), the alignment of the trigonal pyramidal As(III)S_3 and Pb(II)S_3^- binding sites (**Figure 2-39**) suggests that Pb(II) binds in a similar fashion to a tris-thiolate site. This may be due to the stereochemically active lone pair of Pb(II) that may direct the metal to form preferentially in this conformation. An *exo* configuration for Pb(II)S_3^- had been predicted based on structures of small molecule models[25,50–53], a Pb(II)-bound ALAD protein structure[54,55] and a computational study further corroborated this structural preference. However, the observed *endo* conformation contradicts these previous precedents and suggests that the dominant factor controlling the metal structure in the **CoilSer** and **GRAND-CoilSer** peptides is the side chain rotamer orientation within the 3SCC scaffold.

We can examine this hypothesis using the crystallographic parameters obtained in this study. The arrangement of the bound thiols in the first coordination sphere of the As(III) and Pb(II) structures are almost identical. The Cys ligands are directed toward the core of the structure and upward to the N-termini. In the As(III)-structure, the Cys rotamer has the χ_1 dihedral angle of -59.66°, which is within 10 degrees of that reported for the χ_1 of $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ (-68.38°). As a consequence, the Sγ-Sγ separation of Pb(II)S_3^- is 0.24 Å longer than the As(III)S_3 structure (3.25 Å, average) implying that the larger size of Pb(II) ion requires a larger metal site to sequester the metal ion. Moreover, it is observed that to reach an agreement of a Pb(II)-S distance of 2.65 Å, the *endo* Pb(II) has to be located at a distance of 1.69 Å lower than the bound 16Cys layer, while the smaller As(III) exhibits Sγ-Sγ separation of 3.25 Å. This value is even slightly shorter than the Sγ-Sγ separation observed in the apo-(**CS**L16C)$_3$ (3.29 Å), indicating that in order to bind As(III), the apo-thiols have to compress marginally toward the core resulting in a smaller metal binding pocket. In contrast, for Pb(II) the thiols must spread out to 3.49 Å(Sγ-Sγ) because Pb(II) is bigger. Additionally, the *endo* As(III) ion, does not have to move significantly below the 9Cys plane (only 1.34 Å) because the As(III)-S (2.28 Å) is much shorter than the Pb(II)-S distance. Therefore, for these trigonal pyramidal metalloids, the observations of the lower

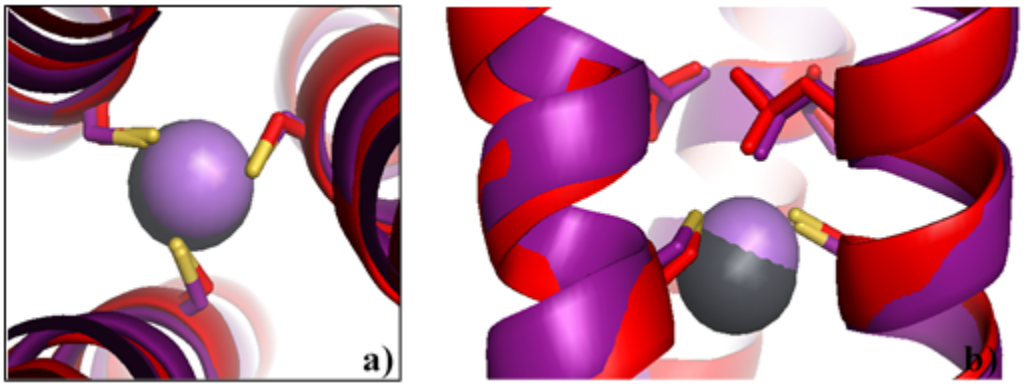

Figure 2-39. An overlay demonstrating the trigonal pyramidal structure comparison between $Pb(II)(S)_3^-$ from $Pb(II)_SZn(II)_N$(GRAND-CSL16CL30H)$_3^+$and $As(III)(S)_3$ from the published $As(III)(CSL9C)_3$ structure (PDB code: 2JGO)[1]. In a) representing top down view from the N-termini and b) side view of the binding sites. Main chains of $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$are colored in purple and As(III)(**CS**L9C)$_3$ in red. The sulfurs atoms are yellow. The Pb(II) and As(III) ions are shown as grey and light purple spheres.

position of the bound metal with respect to the thiol plane indicates that the lone pair of both Pb(II) and As(III) point toward the C-termini with the lone pair well accommodated by the hydrophobic packing around the metal site. The recognition from these crystallographic studies that Pb(II) and As(III) form such similar trigonal pyramidal geometries that causes little perturbation to the core of the 3SCC allows me to now address important fundamental questions in protein design and heavy metal trafficking. Is the cysteine containing layer of the **CoilSer** peptide generally *preorganized* or *predisposed* for metal binding? Is pre-organization dependent on the type of metal? Is pre-organization dependent on the oxidation state of the metal? Is pre-organization dependent on the coordination geometry of the metal? The answers to these questions that are provided in the following section inform future *de novo* protein designs and provides deeper insight for metal recognition in native proteins (e.g., Pb(II) binding to the metalloregulatory protein PbrR691).

Section IV: ***On the Role of Ligand Pre-organization in de Novo Designed Metalloproteins***

Metalloprotein design is challenging because it requires a stable protein scaffold into which an additional level of complexity, an inorganic active site, may be constructed.[48] Further complicating design criteria are a variety of factors (e.g., number of ligands, coordination geometries and oxidation states and appropriate ligand donor sets) that have to be taken into consideration. In the previous four sections, I have described how the **CoilSer** 3SCC peptides are versatile and stable scaffolds that may be modified to generate a tris-thiolate environment capable of enforcing a specific metal geometry. Examples were given that increase the common coordination number of an ion (e.g., **Figure 2-40**, $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ in *Section I*), that lower the usual coordination number (e.g., $Cd(II)S_3^-$) or to satisfy the ion's preferred geometries (e.g., $Cd(II)S_3O^-$, $Zn(II)S_3O^-$, $Pb(II)S_3^-$ and $As(III)S_3$). At this point it is useful to assess the complementary between the protein "host" and the metal ion guest, specifically answering the question whether these 3SCCs are simply *predisposed* to bind metals or actually *preorganized* for specific metal structures.

Given the S_γ-S_γ separation of 3.29 Å in the apo-(**CS**L16C)$_3$, it is essentially impossible to bind a metal in a trigonal planar geometry without imparting some degree of cysteine side chain rearrangement. This is because the central ion would be required to have a M-S distance of 1.90 Å to fit within the 3 sulfur atom plane[56], which is unrealistically short for ions of Hg(II), Cd(II),

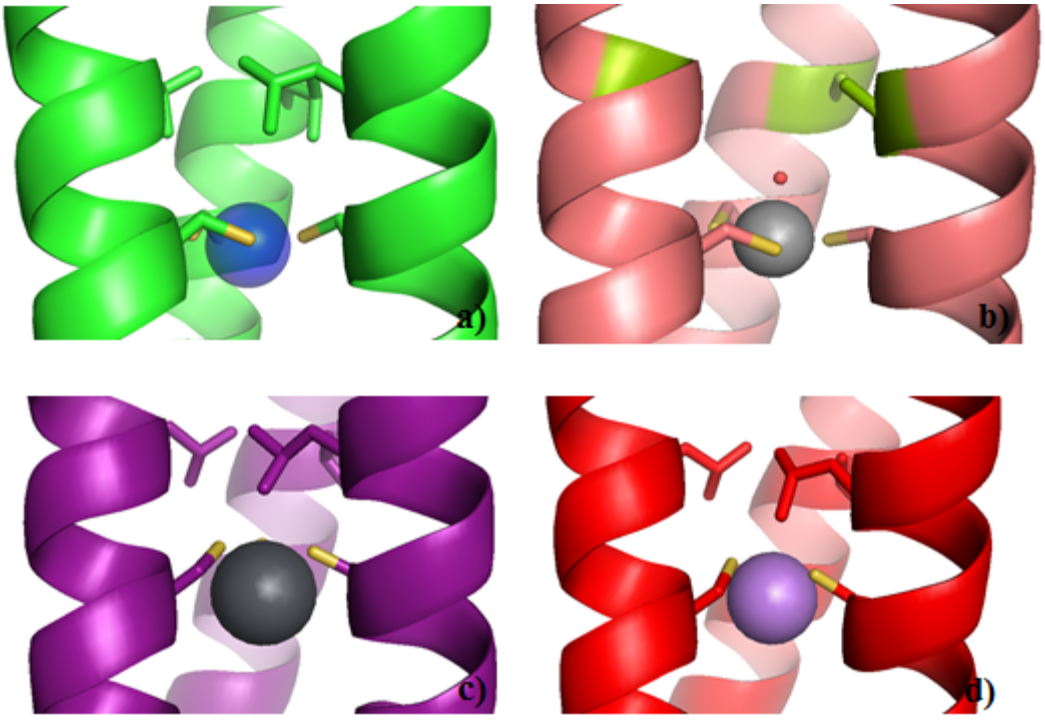

Figure 2-40. PyMOL visualization of the metal binding sites prepared in the 3SCC peptides. Shown, side view of a) trigonal planar $Hg(II)S_3^-$ from $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$, b) pseudo-tetrahedral $Zn(II)S_3O^-$ from Zn(II)(**GRAND-CS**L12AL16C)$_3^-$, c) trigonal pyramidal $Pb(II)S_3^-$ from $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and d) trigonal pyramidal $As(III)S_3$ from As(III)(**CS**L9C)$_3$.

Pb(II) and As(III). Therefore, simply by assessing the apo structure, we can conclude that the **CoilSer** peptides containing a sulfur layer are *predisposed*, but not *preorganized* toward trigonal planar metal complexes. Examination of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12A16C)$_3^-$ confirms this conclusion with S_γ-S_γ separations of over 4 Å and significant perturbations of the χ_1 Cys torsion angle. Similarly, we can conclude that these peptides are also *predisposed* to trigonal planar Cd(II) binding when sufficient steric constraints are imposed to exclude water from the metal first coordination sphere (**Figure 2-41**).

Examination of the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure reveals that the Hg(II) ion is actually situated at a distance of 0.30 Å below the sulfur plane, thus resulting in a slightly distorted trigonal structure with a S-Hg(II)-S of 118.36°. Interestingly, when the first coordination sphere of the trigonal planar $Hg(II)S_3^-$ is aligned to the trigonal pyramidal $Pb(II)S_3^-$ structure it shows that the coordinates of Hg(II) and Pb(II) are within 0.10 Å, with Pb(II) lower in position with respect to the N-termini (**Figure 2-42**). This observation is important as it demonstrates that when the two metals sit at almost identical positions in the 3SCC scaffolds, the geometric preference of the metal defines the observed polyhedron.

The next polyhedron to analyze is the pseudo-tetrahedral $Zn(II)S_3O^-$ ion found in Zn(II)(**GRAND-CS**L12A16C)$_3^-$. Because this is an L12A variant, the second coordination sphere residues above the Zn(II) provide additional space for solvent molecules. None-the-less, the Zn(II) polyhedron may still reveal some useful information regarding the basic requirements necessary for a 4-coordinate metal binding site in 3SCCs. As shown in **Figure 2-7**, the peptide scaffold provides three Cys ligands coordinating to the Zn(II) center and an exogenous water serves as the fourth ligand to complete the 4-coordinate structure. Interestingly, the conformations of the threefold related Cys are similar, but not identical, to that seen in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. This means that the formation of 4-coordinate environment also requires the Cys rotamers to adopt positions pointing up toward the N-termini (apo-protein) and directing out toward the helical interface (**Figure 2-43**).

Unlike Hg(II) that forms a trigonal planar structure which requires the metal to sit roughly within the 3 atom sulfur plane, Zn(II) being 4 coordinate is likely displaced toward the fourth ligand. Despite this difference, the χ_1 torsion angle of the Zn(II) is -146.96°, whereas the $Hg(II)S_3^-$ shows χ_1 of -151.88°. None-the-less, this diminished dihedral angle in Zn(II)(**GRAND-**

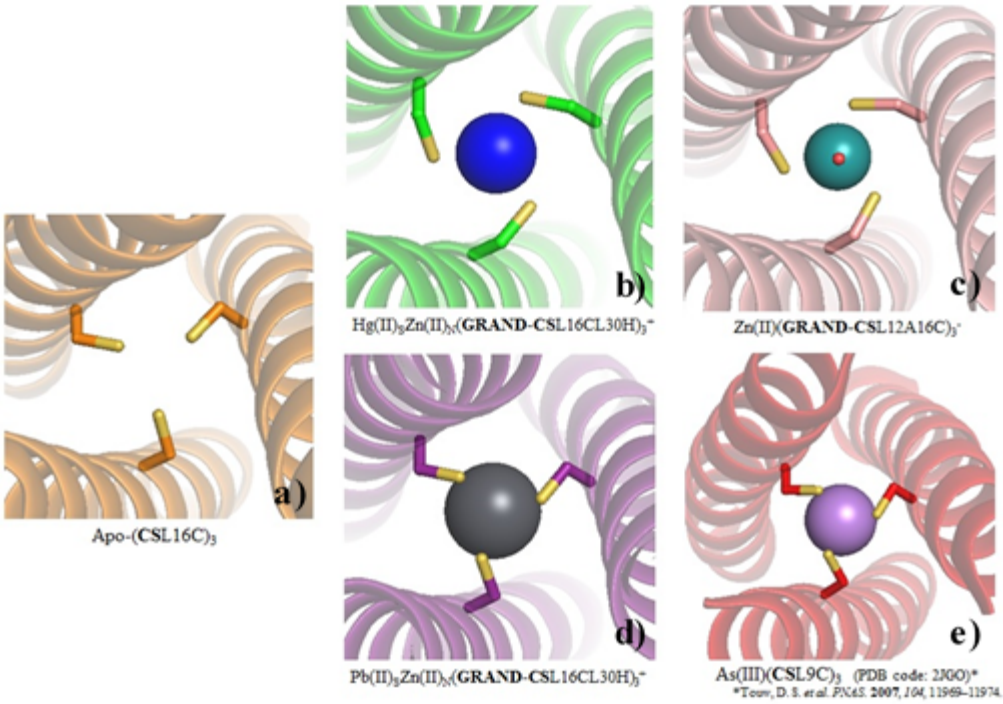

Figure 2-41. Ligand organizations of the first coordination sphere of the designed metal sites. In a) apo-coordination from apo-(**CS**L16C)$_3$, b) trigonal Hg(II)S_3^- from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ (The water observed in the structure is unbound to Hg(II).), c) 4-coodinate Zn(II)S_3O^- from Zn(II)(**GRAND-CS**L12AL16C)$_3^-$, d) trigonal pyramidal Pb(II)(S)$_3$ from Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and e) trigonal pyramidal As(III)S_3 from As(III)(**CS**L9C)$_3$ (PDB code: 2JGO)[1].

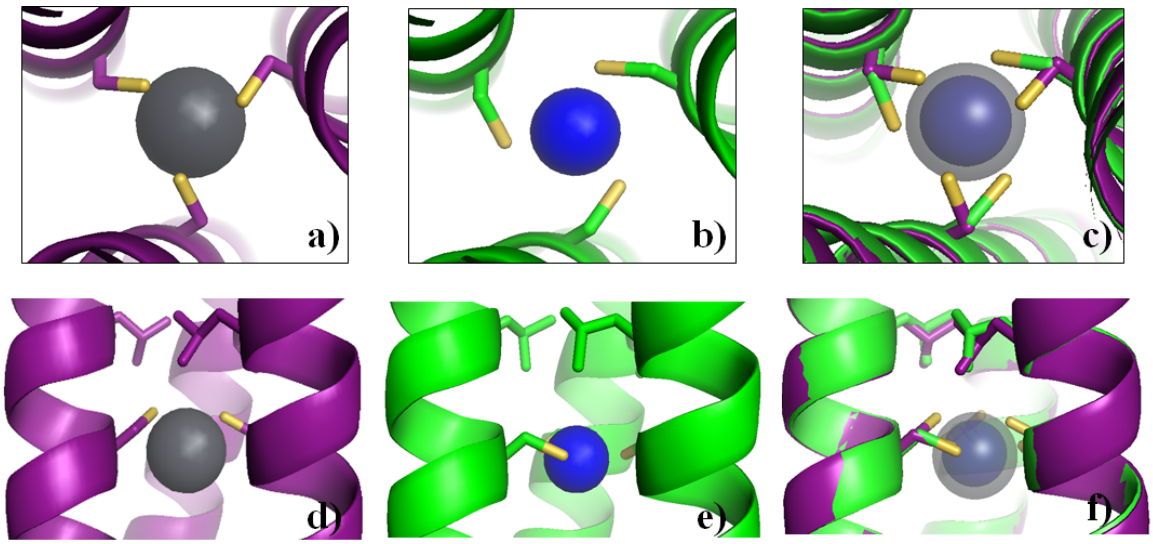

Figure 2-42. PyMOL visualization representing the comparison of the metal coordination between the trigonal pyramidal Pb(II)S_3^- to the trigonal planar Hg(II)S_3^-. Top panels: Top down view from the N-termini representing a) trigonal pyramidal Pb(II)S_3^- from Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, b) trigonal planar Hg(II)S_3^- from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ (The unbound water observed in the structure is excluded for clarity.) and c) an overlay between a) and b). **Bottom panels**: Side view of d) Pb(II)S_3^-, e) Hg(II)S_3^- and f) an overlay between d) and e). The 16Cys side chains in the sixteenth position are present as sticks (sulfur = bright yellow). The Pb(II) and Hg(II) are shown as grey, blue and purple spheres.

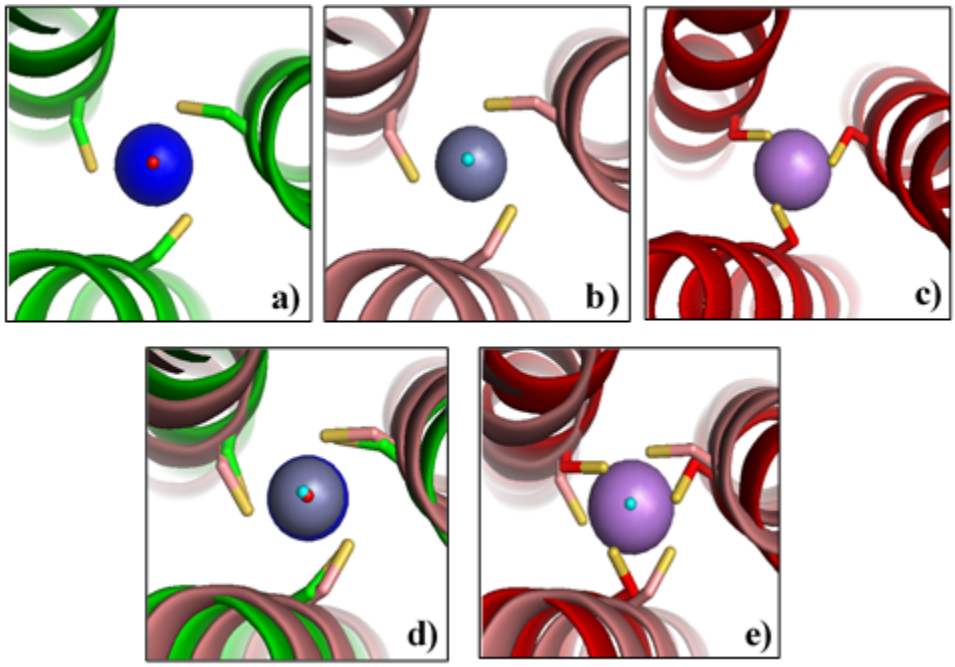

Figure 2-43. PyMOL visualization representing the comparison of the metal coordination between the trigonal planar Hg(II)S_3^-, trigonal pyramidal As(III)S_3 to the pseudo-tetrahedral Zn(II)$S_3(H_2O)^-$ from Zn(II)(GRAND-CSL12AL16C)$_3^-$. Top panels: Top down view from the N-termini representing a) trigonal planar Hg(II)S_3^- from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ (The water observed in the structure is uncoordinated to Hg(II).), b) Zn(II)S_3O^- from Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ and c) trigonal pyramidal As(III)S_3 from As(III)(**CS**L9C)$_3$. **Bottom panels**: Representing an overlay between two of the metal layers,d) Hg(II)S_3^- (green) versus Zn(II)S_3O^- (pink) and e) As(III)S_3 (red) versus Zn(II)S_3O^- (pink). Main chain atoms are shown as ribbon diagrams. 16Cys side chains are present as sticks (sulfur = yellow). The Hg(II), Zn(II) and As(III) ions are shown as blue, grey and purple spheres. The waters observed in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown as red and cyan small spheres, respectively. PDB code for As(III)(**CS**L9C)$_3$ is 2JGO.[1]

CSL12A16C)$_3^-$ causes a shorter Sγ-Sγ separation (3.94 Å) verifying that the sulfur plane of the Zn(II) structure is smaller than the trigonal planar Hg(II)S_3^- (4.16 Å). As previously described in *Section II*, the Zn(II) structure is distorted as the angles of S-Zn(II)-S and S-Zn(II)-O are not perfectly at 109.5°. This is mainly because Zn(II) actually sits relatively close to the sulfur plane, with a surprising small out-of-plane distance of 0.09 Å toward the water ligand. Thus, the Zn(II) ion is displaced less than the trigonal planar Hg(II)S_3^- (0.30 Å below the plane). This may be a consequence of the shorter Zn(II)-S bonds (2.27 Å), which do not require the thiols to open as much as for the larger Hg(II).

Another factor that makes Zn(II) unique from Hg(II) is its preference for a 4-coordinate geometry. The Zn(II)-O distance is 2.17 Å, reflective of a real coordinative bond; while the water that is present above the Hg(II)S_3^- in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Hg(II)(**GRAND-CS**L12A16C)$_3^-$ is found to be 2.79 and 3.56 Å, respectively. A very important observation for both the Hg(II) and Zn(II) structures is that the solvent is ONLY found in the core of the 3SCC toward the N-termini. In none of these structures is water observed in the region between the 16Cys and 19Leu. This observation may be critical for understanding the difference between pseudotetrahedral and trigonal pyramidal geometries.

To compare the pseudo-tetrahedral and trigonal pyramidal geometries, the Zn(II)$(S)_3O^-$ is overlaid onto the As(III)$(S)_3$ in **Figures 2-43** and **2-44**. Again, both of the structures share the same basic three Cys ligand set to anchor the ions in the protein; however, while their radii are comparable [As(III), 0.72 Å and Zn(II), 0.74 Å], the specific preference in geometry causes the bound-thiols to orient differently. To bind As(III), the thiols exhibit a χ_1 torsion angle of -59.66°, resulting in a sulfur plane similar to the apo-structure. As previously described, Zn(II) is bound when the thiols are reoriented outward to the helical interface, making the plane of Cys shifted lower by ~1.30 Å than for the plane in the unmetallated, and As(III) structures.

Intriguingly, even though the shift of sulfur planes are observed between these two geometries, the alignment of the two structures (**Figure 2-44, e**) clearly shows that the positions of As(III) and Zn(II) are relatively close in the binding sites. Their positions are separated by only 0.1 Å (where As(III) is higher). This implies that the rotation of the Cys required to make the Zn(II)$(S_3O^-$ site allows the Zn(II) to fit in the same position as the As(III) despite the fact the cysteine rotamer angles and the three sulfur planes are very different. Moreover, this Zn(II)S_3O^-

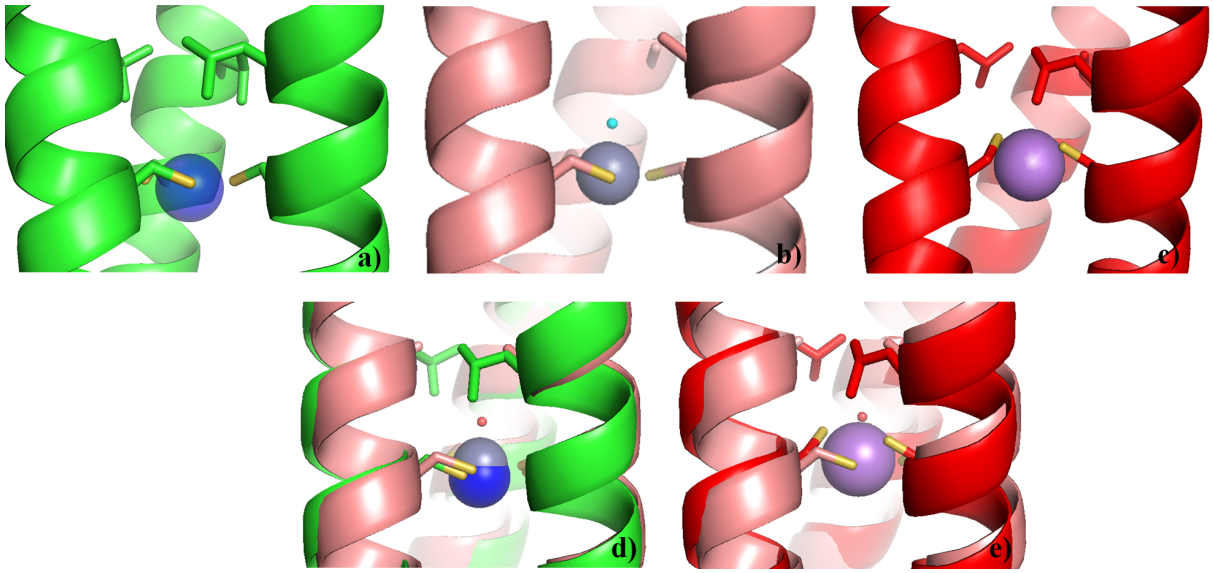

Figure 2-44. Side view visualization representing the comparison between the trigonal planar Hg(II)S_3^-, trigonal pyramidal As(III)S_3 to the pseudo-tetrahedral Zn(II)$S_3(H_2O)^-$ from Zn(II)(GRAND-CSL12AL16C)$_3^-$. Top panels: Representing a) trigonal planar Hg(II)S_3^- from Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, b) Zn(II)S_3O^- from Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ and c) trigonal pyramidal As(III)S_3 from As(III)(**CS**L9C)$_3$. **Bottom panels**: Representing an overlay between two of the metal layers, d) Hg(II)S_3^- (green) versus Zn(II)S_3O^- (pink) and e) As(III)S_3 (red) versus Zn(II)S_3O^- (pink). Main chain atoms are shown as ribbon diagrams. 16Cys side chains are present as sticks (sulfur = yellow). The Hg(II), Zn(II) and As(III) ions are shown as blue, grey and purple spheres. The waters observed in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ are shown as red and cyan small spheres, respectively. The PDB code for As(III)(**CS**L9C)$_3$ is 2JGO.[1]

bound Cys orientation results in a longer $S\gamma$-$S\gamma$ separation than for the As(III) site (3.94 Å vs 3.25 Å). This is because the As(III) is bound as an *endo*-conformation that allows the metal to sit 1.34 Å below the sulfur plane. A simple minded model might suggest that with respect to the metal sulfur bonding, trigonal pyramidal versus pseudo tetrahedral polyhedral would be closely similar. Yet we see here that there are marked differences in $S\gamma$-$S\gamma$ separations, χ_1 Cys torsion angles and the relative position of the three atom sulfur plane for the two geometries. To understand these observations, we must consider the differences in packing at the Leu19 position and how these hydrophobes can accommodate a stereochemically active lone pair versus a coordinated solvent.

The packing of the hydrophobic layer below the metal site is compared for the $Hg(II)S_3^-$, $Zn(II)S_3O^-$, $As(III)S_3$ and $Pb(II)S_3^-$ structures in **Figure 2-45**. Based on the PyMOL visualization, the packing of the 19Leu residues below the Zn(II) site appears to be similar to that seen for $Hg(II)S_3^-$, although the inner $C\delta$-$C\delta$ (5.95 versus 5.89 Å) and exterior $C\delta$-$C\delta$ (8.98 versus 8.42 Å) separations observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ are slightly longer than Zn(II)(**GRAND-CS**L12A16C)$_3^-$, the effect is not strikingly pronounced. However, the same $C\delta$-$C\delta$ separations lengthen by 1-2 Å in the $Pb(II)S_3^-$ structure, which is likely a consequence of the lower position and larger size of Pb(II) when it is bound to the site. The packing below the trigonal pyramidal $As(III)S_3$ site [inner $C\delta$-$C\delta$ separation (5.26 Å) and exterior $C\delta$-$C\delta$ separation (7.64 Å)] is similar to, but slightly shorter, than the $Zn(II)(S)_3O^-$ separations in Zn(II)(**GRAND-CS**L12A16C)$_3^-$. The fact that Zn(II) and As(III) have similar size and are relatively close in positions through space may explain this similarity. The most important factor is that the leucine layer comes closer to the three atom sulfur plane in the $Zn(II)S_3O^-$. When Zn(II) is bound, the rotation of Cys causes the ligand pocket to get closer to the 19Leu plane. Given the small distances, there is not sufficient space to add a solvent molecule at a 2.17 Å distance from the Zn towards the 19Leu layer. One should remember at this point that the peptide sequence from the 12Leu position to the 16cys position has three intervening residues whereas the 16cys to 19Leu has only two residues. Thus, there is less space available for solvent between 16cys to 19Leu than for 12Leu to the 16cys position. This is consistent with observation that we have never observed a solvent below the metal plane when Leu is in the nineteenth position. Thus, there are two possibilities for the

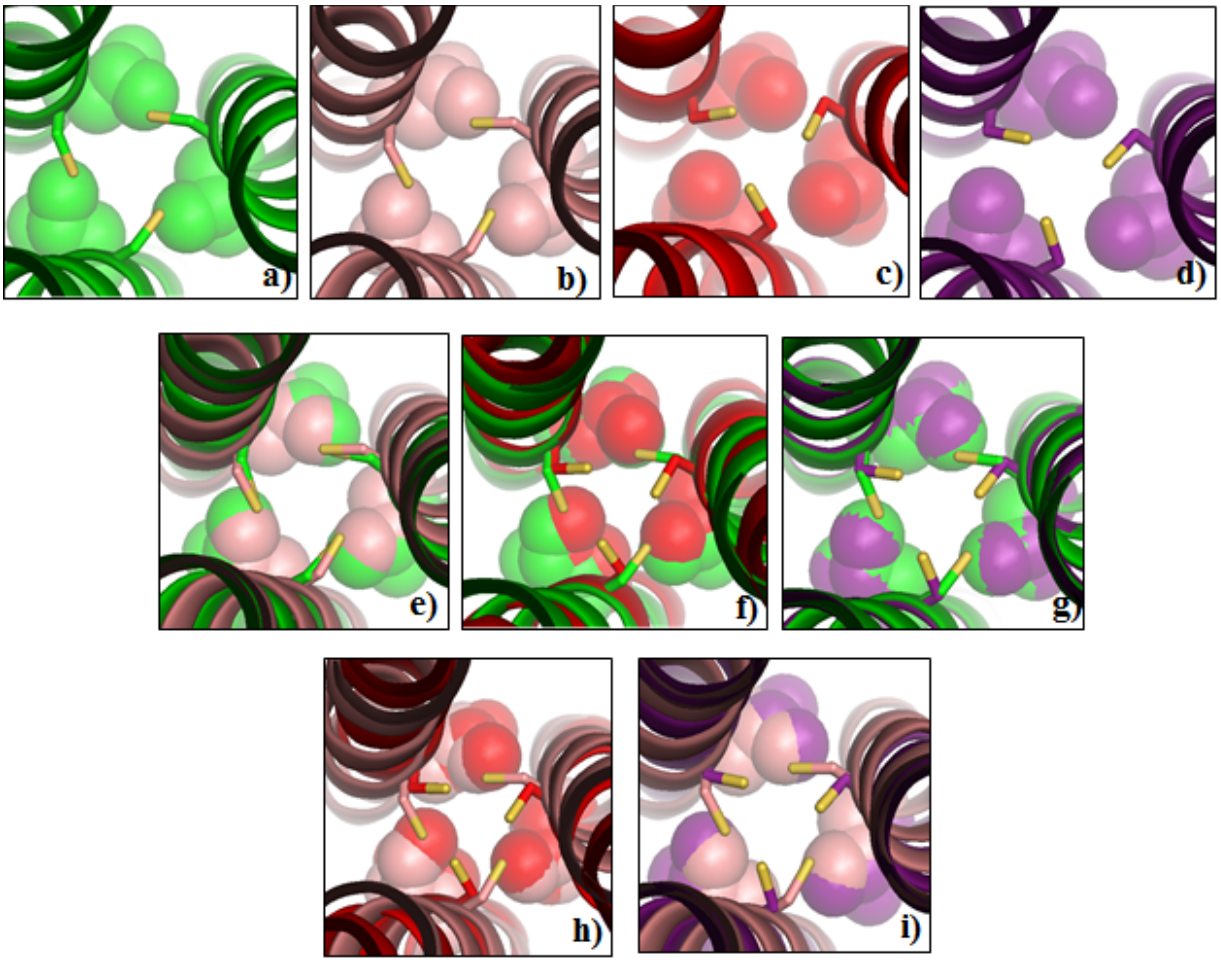

Figure 2-45. Packing comparison of Leu residues below the metal site. Top panels: Top down view from the N-termini representing the Leu packing (shown as spheres) of a) 19Leu in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ (green), b) 19Leu in Zn(II)(**GRAND-CS**L12AL16C)$_3^-$ (pink), c) 12Leu in As(III)(**CS**L9C)$_3$ (red) and d) 19Leu in $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ (purple). **Middle and bottom panels**: Representing an overlay between the two structures, e) $Hg(II)S_3^-$ versus $Zn(II)S_3O^-$, f) $Hg(II)S_3^-$ versus $As(III)S_3$, g) $Hg(II)S_3^-$ versus $Pb(II)S_3^-$, h) $As(III)S_3$ versus $Zn(II)S_3O^-$ and i) $Pb(II)S_3^-$ (green) versus $Zn(II)S_3O^-$ (pink). Main chain atoms are shown as ribbon diagrams. The 16Cys side chains are present as sticks (sulfur = yellow). The metal ions are omitted for clarity. The PDB code for As(III)(**CS**L9C)$_3$ is 2JGO.[1]

formation of a pseudotetrahedral center. The first is that the Zn(II) adopts the cysteine rotamers of the Pb(II) or As(III) ions. This is unlikely because it would require that the bound water point toward the C-termini. The water simply cannot fit in the resultant cavity whereasthere is enough space to accommodate even the largest lone pairs. The second possibility is that which is observed. The cysteines rotate as in the Hg(II) structure in order to allow the water to occupy the fourth coordination site in the cavity towards the N-termini. A similar phenomenon is expected for Cd(II), which is larger than Zn(II) and has a Cd(II)-O distance of 2.35 Å based on EXAFS studies. For these reasons, I conclude that the **CoilSer** peptide is *predisposed*, but not *preorganized* for the sequestration of pseudotetrahedral polyhedra, regardless of the metal.

From this analysis, I conclude that the only two cases where the peptide scaffold is *pre-organized* is for Pb(II) and As(III) because the geometrical orientation of their bound Cys rotamers are almost identical to the apo-peptide (**Figure 2-51**). It should be emphasized that this is not a size effect as Pb(II) is much larger than Hg(II) or Cd(II) and As(III) is comparable in size to Zn(II). The primary basis for Pb(II) and As(III) accepting the apo-structure conformations is the trigonal pyramidal polyhedron that these ions form. This is not to say that size is unimportant in defining the structure. The apo-protein shows an average χ_1 Cys torsion angle of -69.20° with a $S\gamma$-$S\gamma$ separation of 3.29 Å while for Pb(II) binding, the χ_1 angle observed in $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ is -68.38° which makes a $S\gamma$-$S\gamma$ separation to be 3.49 Å and for As(III)(**CS**L9C)$_3$ one observes an average χ_1 of -59.66° with a $S\gamma$-$S\gamma$ separation of 3.25 Å. The close values of $S\gamma$-$S\gamma$ separations indicate that the ligand spheres of As(III) and Pb(II) are similar, yet they are not identical. The size of the bound metal defines the necessary size of the thiol plane for each metal. As(III) with an atomic radius of 0.72 Å requires the thiol ligands to rotate slightly more inward to the core by 10 degree compared to the apo-proteins (both apo-(**CS**L16C)$_3$ and apo-(**CS**L9C)$_3$), compressing the diameter of the sulfur plane. On the other hand, the binding of the larger Pb(II), atomic radius of 1.33 Å, causes the thiols to slightly move outward from the core as indicated by the longer $S\gamma$-$S\gamma$ observation in $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. Both the metallated As(III)S_3 and Pb(II)S_3^- crystal structures demonstrate that the main reason that the bound-Cys rotamers of the trigonal pyramidal geometry does not drastically change is because the metal is bound to the site in an *endo*-configuration. The smaller As(III) ion achieves a 2.28 Å As(III)-S

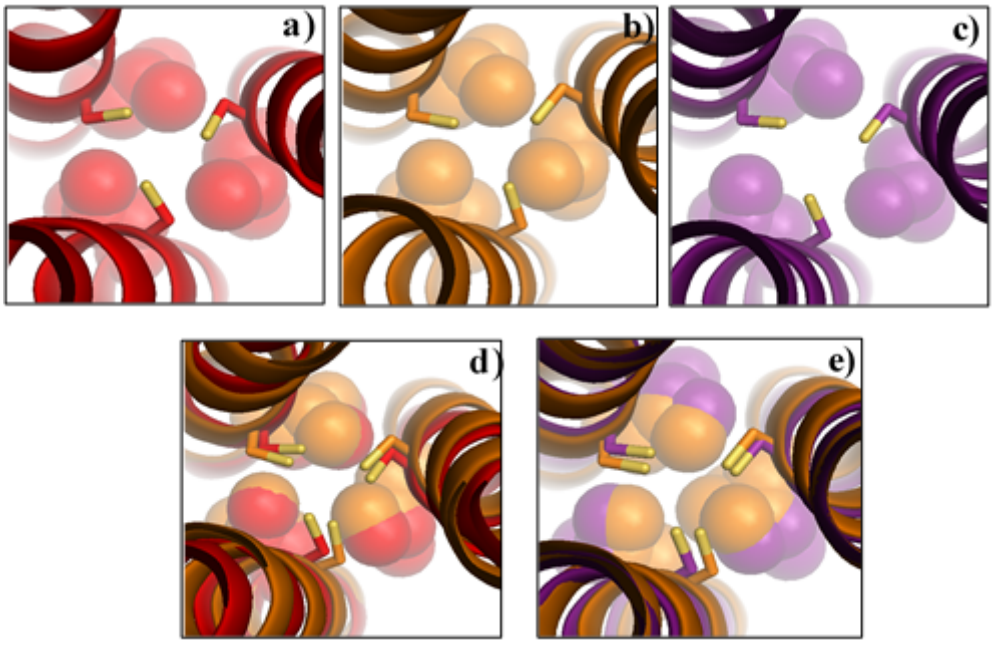

Figure 2-46. Packing of Leu residues (shown as spheres) demonstrating the hydrophobic packing below the trigonal pyramidal binding site compared to the apo-structure. Top panels: From top down view of the N-termini, representing the packing of a) 12Leu below the 9Cys layer in As(III)(**CS**L9C)$_3$, b) 19Leu below the 16Cys in apo-(**CS**L16C)$_3$ and c) 19Leu below the 16Cys in Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. **Bottom panels**: Representing in f) the packing comparison between a) and b). In e) an overlay between b) and c).Main atoms of Pb(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$,apo-(**CS**L16C)$_3$ and As(III)(**CS**L9C)$_3$ are represent as purple, orange and red ribbon diagrams, respectively. Cys side chains are present as sticks in which the thiols are yellow. Leu residues are shown as spheres indicating the packing. Pb(II) and As(III) ions are omitted for clarity. The PDB code for As(III)(**CS**L9C)$_3$ is 2JGO.[1]

bond distance when the metal is situated at a distance of 1.34 Å below the sulfur plane. On the other hand, Pb(II), with a longer bond length (2.65 Å), is located more downward to the C-termini at a distance of 1.69 Å below the sulfur plane.

To gain more structural insight of how the secondary coordination spheres around the metal site are affected when the trigonal pyramidal structure is formed, the metallated $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and As(III)(**CS**L9C)$_3$ structures are aligned with their analogous apo-proteins. It is obvious in **Figure 2-46** that when the larger Pb(II) ion is bound to the metal site, it causes the 19Leu residues to be directed further out to the helical interface compared to the apo-(**CS**L16C)$_3$. This statement is confirmed by the longer δ-methyl carbon separations determined in the $Pb(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure. The distances of interior $C\delta1$-$C\delta1$ (higher in position) and exterior $C\delta2$-$C\delta2$ (lower in position) are 6.17 and 9.13 Å, respectively in the Pb(II) structure, whereas the related distances of 4.65 and 7.28 Å observed in the non-metallated form. Similarly, the binding of As(III) also causes the hydrophobic residues below the metal site to reorient toward the outer face; however, the effect of this smaller ion is not significantly pronounced as compared to the case of Pb(II) ion. These observations suggest that though the space below the metal site is already available to facilitate the *endo*-formation of these metals, the Leu residues below the metal plane have to move slightly toward the helical interface then they were in the apo-form to accommodate the metal lone pairs.

In sum, the structural observations imply that when a lone-pair containing metal binds to a tris thiolate Cys site in a trigonal pyramidal *endo*-configuration, the metal sits below the sulfur layer. This means that it does not require the Cys residues to change their orientations from the apo-peptide so they can remain oriented in the same manner prior to metal binding.

Conclusion

In sum, these crystallographic studies have shown that the 3SCC peptide are excellent scaffolds that can provide a three Cys environment to adopt a variety of different polyhedra. Simply starting from a threefold symmetric Cys ligands, one can generate a trigonal planar structure as observed in the case of $Hg(II)S_3^-$. Moreover, a trigonal pyramidal mode can be made when a lone pair containing metal like As(III) or Pb(II) is placed into the binding site. Finally, if the lone pair is replaced with an exogenous ligand such as water in the Zn(II) or Cd(II) systems a

pseudo-tetrahedral geometry is obtained. Moreover, these studies demonstrate that the apo-protein is only *pre-rganized* for a trigonal pyramidal polyhedron because this type of geometry does not require the Cys residues to expand for metal incorporation or for a fourth ligand to be accommodated in the hydrophobic cavity below the sulfur plane. On the other hand, both the trigonal planar and pseudo-tetrahedral geometries must incur significant cysteine side chain re-orientation to bind metals. For trigonal planar structures, this is because the intrinsic packing of the cysteines in the apo structure places the thiol groups to close to accommodate a metal with the three sulfur plane. In the case of pseudotetrahedral compounds, the only position that allows a fourth ligand to bind to the metal is when the ion is oriented toward the N-termini and the solvent approaches the 12Leu or 12Ala hydrophobic plane. The 19Leu plane below the sulfur plane can only accept a lone pair, but not a fourth ligand. Therefore, I conclude that irrespective of metal size, the **TRI** family/**CoilSer**/**GRAND-CoilSer** of peptides are only *pre-rganized* to sequester trigonal pyramidal ions and are predisposed for binding other coordination geometries. Based on these conclusions, one would predict that these 3SCC peptides would also be *predisposed* but not *preorganized* to bind trigonal bipyramidal metals. This issue is addressed in Chapter 3. Finally, it was found that not only the first coordination sphere of Cys residues is perturbed but also the secondary coordination spheres around the metal are indirectly affected upon metal binding. The degree of perturbation is dependent on the position of the bound metal, size of the metal and the location of the sulfur plane.

References

(1) Touw, D. S.; Nordman, C. E.; Stuckey, J. a; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11969–11974.

(2) Zastrow, M. L.; Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Nat. Chem.* **2012**, *4*, 118–123.

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

(4) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70–77.

(5) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307-326.

(6) Vagin, A.; Teplyakov, A. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2010**, 22–27.

(7) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242.

(8) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2003**, *59*, 1131–1137.

(9) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(10) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2004**, *60*, 2126-2132.

(11) Roversi, P.; Sharff, A.; Smart, O. S.; Vonrhein, C.; Womack, T.O. (**2011**). BUSTER version 2.11.2 Cambridge, United Kingdom: Global Phasing Ltd.

(12) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2010**, *132*, 13240–13250.

(13) Chen, V. B.; Arendall III, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D.C. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2010***, 66*, 12-21.

(14) Terwilliger, T. C.; Adams, P. D.;Read, R. J.; McCoy, A. J.; Moriarty, N. W.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Zwart, P. H.; Hung, L. W. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2009**, *65*, 582-601.

(15) Grosse-Kunstleve, R.W; Adams, P.D. *Acta Cryst.* **2003**, *59*, 1966-1973.

(16) Terwilliger, T. C. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2000**, *56*, 965-972.

(17) Terwilliger, T. C.;Grosse-Kunstleve, R. W.; Afonine, P. V.; Moriarty, N. W.; Zwart, P. H.; Hung, L. W.; Read, R. J.; Adams., P. D. *Acta Cryst.* **2008**, *64*, 61-69.

(18) Terwilliger, T. C.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Moriarty, N. W.; Adams, P. D.; Read, R. J.; Zwart, P. H.; Hung, L. W. *Acta Cryst.* **2008**, *64*, 515-524.

(19) Terwilliger, T. C. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2004**, *60*, 2144-9.

(20) Afonine, P. V.; Grosse-Kunstleve, R.W.; Echols, N.; Headd, J. J.; Moriarty, N. W.; Mustyakimov, M.; Terwilliger, T. C.; Urzhumtsev, A.; Zwart, P. H.; Adams, P. D. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2012**, *68*, 352-67.

(21) Ogihara, N. L.; Weiss, M. S.; DeGrado, W. F.; Eisenberg, D. *Protein Sci.* **1997**, *6*, 78–86.

(22) Ponder, J. W.; Richard, F. W. *J. Mol. Biol.* **1987**, *193*, 775–791.

(23) Wright, J. G.; Tsang, H.; Penner-hahn, J. E.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, 2434–2435.

(24) Watton, S. P.; Wright, J. G.; MacDonnell, F. M.; Bryson, J. W.; Sabat, M.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2824.

(25) Christou, G.; Folting, K.; Huffman, J. C. *Polyhedron* **1984**, *3*, 1247-1253.

(26) Gruff, E. S.; Koch, S. A. *J. Am. Chem. Soc.* **1990**, *112*, 1245-1247.

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

(28) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2864–2868.

(29) Lee, K.-H.; Matzapetakis, M.; Mitra, S.; Marsh, E. N. G.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2004**, *126*, 9178–9179.

(30) 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.

(31) 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**, *7863*, 6195–6196.

(32) Iranzo, O.; Ghosh, D.; Pecoraro, V. L. *Inorg. Chem.* **2006**, *45*, 9959–9973.

(33) Pasternak, A.; Kaplan, J.; Lear, J. D.; Degrado, W. F. *Protein Sci.* **2001**, *10*, 958–969.

(34) Shannon, R. D. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1976**, *32*, 751–767.

(35) Cremer, D.; Kraka, E.; Filatov, M. *ChemPhysChem.* **2008**, *9*, 2510–2521.

(36) Vallee, B. L.; Ulmer, D. D. *Annu. Rev. Biochem.* **1972**, *41*, 91–128.

(37) Onyido, I.; Noriis, A. R.; Buncel, E. B. *Chem. Rev.* **2004**, *104*, 5911–5929.

(38) Natan, M. J.; Millikan, C. F.; Wright, J. G.; O' Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 3255–3257.

(39) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry* **2009**, *15*, 3761–3772.

(40) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16566–16571.

(41) Peacock, A. F. A.; Pecoraro, V. L. *Cadmium: From Toxicity to Essentiality*; Sigel, A.; Sigel, H.; Sigel, R. K., Eds.; Metal Ions in Life Sciences; Springer Netherlands: Dordrecht, 2013; Vol. 11.

(42) Peacock, A. F. A.; Iranzo, O.; Pecoraro, V. L. *Dalton Trans.* **2009**, *9226*, 2271–2280.

(43) Ghosh, D.; Lee, K.-H.; Demeler, B.; Pecoraro, V. L. *Biochemistry* **2005**, *44*, 10732–10740.

(44) Chakraborty, S.; Iranzo, O.; Zuiderweg, E. R. P.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2012**, *134*, 6191–6203.

(45) Iranzo, O.; Cabello, C.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 6688–6691.

(46) Iranzo, O.; Chakraborty, S.; Hemmingsen, L.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2011**, *133*, 239–251.

(47) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7371–7374.

(48) Zastrow, M.; Pecoraro, V. L. *Coord. Chem. Rev.* **2013**, 2565-2588.

(49) Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. *Science* **1993**, *262*, 1401–1407.

(50) Dean, P. A. W.; Vittal, J. J.; Payne, N. C. *Inorg.Chem.* **1984**, *23*, 4232.

(51) Perez-Lourido, P.; Romero, J.; Garcia-Vazquez, J. A.; Sousa, A.;Zheng, Y.; Dilworth, J. R.; Zubeita, J.; Trans, D. *J. Chem. Soc., Dalton Trans.* **2000**, *5*, 769-774.

(52) Ren, Z.-G.; Tang, X. -Y.; Li, L.; Liu, G.-F.; Li, H.-X.; Chen, Y.; Zhang, Y.;Lang, J. P. *Inorg.Chem.Commun.* **2007**, 1253.

(53) Rossini, A. J.; Macgregor, A. W.; Smith, A. S.; Schatte, G.; Schurko, R. W.; Briand, G.G.; *Dalt. Trans.* **2013**, *42*, 9533-9546.

(54) Erskine, P. T.; Duke, E. M. H.; Tickle, I. J.; Senior, N. M.; Cooper, J. B. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2000**, 421–430.

(55) Erskine, P. T.; Norton, E.; Cooper, J. B.; Lambert, R.; Coker, A.; Lewis, G.; Spencer, P.; Sarwar, M.; Wood, S. P.; Warren, M. J.; Shoolingin-jordan, P. M. *Biochemistry* **1999**, 4266–4276.

(56) To calculate the circumradius that gives this number go to http://mathworld.wolfram.com/EquilateralTriangle.html.

Chapter 3

Using Alternative Chirality of D-amino acid in the Second Coordination Sphere of *De Novo* Designed Proteins to Control the Coordination Environment of Cd(II)

Background

The alternative chirality of D-amino acids has been used specifically incorporated into the second coordination sphere to control a specific geometry of Cd(II) or to modify steric environments around the metal site in the **TRI**-family sequences.[1,2] As previously described in Chapter 2, the parent construct, **TRI**L16C (Sequence shown in **Table 3-1**), which contains a Cys residue in an *a* site at the sixteenth position, forms two species in the presence of Cd(II) at pH 8.5. Only one species was observed via ^{113}Cd NMR (δ=625 ppm); however, the ^{111m}Cd PAC spectrum showed that the Cd(II)(**TRI**L16C)$_3^-$ complex contains two species: 40% (ω_o=0.44 rad/ns) of the Cd(II) centers were trigonal planar [Cd(II)S_3^-] and 60% (ω_o=0.34 rad/ns) were pseudo-tetrahedral [Cd(II)S_3O^-], where O is an exogenous water ligand.[3] To enforce an unfavorable, but biologically relevant, Cd(II)S_3^- environment within the 3SCC exclusively, the use of alternative chirality was employed. It was reasoned that by reorienting the 12L-Leu side chain in the twelfth layer toward the metal site, steric interference could occur that would exclude the water from the coordination site. Successfully, the introduction of D-Leu in the twelfth position of **TRI**L16C, resulting in **TRI**L12$_D$LL16C, yielded a 100 % 3-coordinate species as confirmed by the ^{113}Cd NMR chemical shift of 697 ppm and the single ^{111m}Cd PAC angular frequency (ω_o) of = 0.45 rad/ns.[4] This observation revealed a milestone towards the *de Novo* protein design since based on our knowledge, there were no previous reports for the incorporation of D-amino acid into an L-amino acid coiled coil structure to harness geometries of a metal in a binding site and still retain its native

folded structure. This work has identified an exciting strategy to achieve the goal of controlling metal coordination geometry and served as the inspiration for this thesis. Alternating the chirality from L-Leu to D-Leu allows for modification of the sterics above the metal coordination site, subsequently obtaining the desired Cd(II)S_3^- environment. The subsequent hypothesis is that if the stereochemical configuration is inverted below rather than above the Cys layer, the reorientation of D-Leu would produce 4- or 5-coordinate Cd(II) complexes. To this end, D-Leu was incorporated below the Cys layer of **TRI**L16C at the nineteenth position (**TRI**L2WL16CL19$_D$L) to presumably open another hole to provide a space below the metal binding site for an exogenous solvent molecule. Second, to investigate the effect of alternative chirality in the second coordination sphere further, L-Leu were replaced by D-Leu both above and below the L16C layer to make a D-Leu sandwich binding site in **GRAND-CS**L12$_D$LL16CL19$_D$L. It was proposed that re-directing the alkyl groups of D-Leu in the layer at the twelfth position toward the metal plane in a combination with shifting the alkyl groups down in the nineteenth position below the metal plane could result in predominantly Cd(II)S_3O^-, where O is from a water that enters the site from the C-terminal side.

The goal of this particular work is to understand how the variation of D-Leu positions around the metal site has an impact on metal-protein interactions especially in the perspectives of how the coordination chemistry of the metal site is controlled, which is essential to its function. This chapter will first discuss the published physical characterizations of **TRI**L12$_D$LL16C and the context of the new designs presented in this work, **TRI**L2WL16CL19$_D$L and **GRAND-CS**L12$_D$LL16CL19$_D$L. In the second part, to visualize the effect of D-Leu, the crystal structures of apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and mercurated-(**GRAND-CS**L12$_D$LL16C)$_3$ are used to represent the structural details of **TRI**L12$_D$LL16C peptide, while apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and mercurated-(**GRAND-CS**L16CL19$_D$L)$_3$ are models for **TRI**L2WL16CL19$_D$L. These crystal structures provide important insights into explaining how the sterics around the metal site are perturbed by D-Leu, thus resulting in the control of metal geometries. Furthermore, the obtained metallated structures allow for possible Cd(II) complex predictions for each peptide design based on the physical characterizations.

Table 3-1. Peptide sequences

Peptides		**a b c d e f g** 2	**a b c d e f g** 9 12	**a b c d e f g** 16 19	**a b c d e f g**	**a b c d e f g** 30
TRI	Ac-G	LKALEEK	LKALEEK	LKALEEK	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
TRIL12$_D$LL16C	Ac-G	LKALEEK	LKA$_{\mathbf{D}}$**L**EEK	**C**KALEEK	LKALEEK	G-NH_2
TRIL2WL16CL19$_D$L	Ac-G	WKALEEK	LKALEEK	**C**KA$_{\mathbf{D}}$**L**EEK	LKALEEK	G-NH_2
Coil-Ser (CS)	Ac-E	WEALEKK	LAALESK	LQALEKK	LEALEHG	-NH_2
CSL16C	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LEALEHG	-NH_2
GRAND-CoilSer	Ac-E	WEALEKK	LAALESK	LQALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL16CL30H	Ac-E	WEALEKK	LAALESK	**C**QALEKK	LQALEKK	**H**EALEHG -NH_2
GRAND-CSL12$_D$LL16C	Ac-E	WEALEKK	LAA$_{\mathbf{D}}$**L**ESK	**C**QALEKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL16CL19$_D$L	Ac-E	WEALEKK	LAALESK	**C**QA$_{\mathbf{D}}$**L**EKK	LQALEKK	LEALEHG -NH_2
GRAND-CSL12$_D$LL16CL19$_D$L	Ac-E	WEALEKK	LAA$_{\mathbf{D}}$**L**ESK	**C**QA$_{\mathbf{D}}$**L**EKK	LQALEKK	LEALEHG -NH_2

Bold and underlined residues indicate substitutions.

C-and N-termini are capped by Ac and NH_2 groups, respectively.

Experimental section

Materials

Lists of chemicals used for peptide synthesis and characterizations were noted in the experimental section of Chapter 2.

Peptide synthesis and purification

The peptides (sequences as given in **Table 3-1**) were synthesized and purified based on the methods described in Chapter 2.

Physical Characterizations

Circular dichroism (CD) spectroscopy: CD spectral measurements were performed at room temperature on an AVIV 62DS spectrometer using 1-mm strain-free quartz cuvette. CD spectra for 20 uM(monomer concentration) apo-peptide in a solution containing 10mM phosphate buffer pH 8.0 and 10:1 TCEP to monomer ratio(to prevent Cys oxidation), were collected from 280 nm to 190 nm every 1 nm with 1s signal averaging. In the case of holo-peptide, $CdCl_2$ solution was added into the peptide solution to make 1:1 Cd(II):trimer peptide ratio. The CD signal observed in milli-degree of ellipticity was converted to molar ellipticity, [Θ], and is reported in units of $deg.dmol^{-1}cm^2$.

UV/Vis Cd(II) binding titrations: UV/Vis spectra were monitored on a Cary 100 UV/Vis spectrophotometer in 1-cm quartz cuvette. All solutions in the experiments were purged under Ar atmosphere before titrations to reduce chances of oxidations. The titrations were performed by titrating small aliquots (0.1 eq) of 10 mM $Cd(II)Cl_2$ into a 3-mL solution containing 60 uM peptide monomer, 50 mM TrisHCl buffer pH 8.5 and 10eq of TCEP solution per monomer ratio. Upon titrations, the solutions were stirred by flat rectangular magnet for 5 minutes and allowed to sit for 3 minutes before being scanned. The spectra between 350 to 190 nm were collected to observe the increase in LMCT bands with a scan rate of 240 nm min^{-1}. Difference spectra were obtained by subtracting the background spectrum of apo- peptide solution (60 uM peptide monomer in 50 mM TrisHCl and TCEP, pH 8.5) in the absence of metal.

pH titrations: pH titrations were done at room temperature by the addition of small aliquots of concentrated KOH to unbuffered solutions of 20 uM $CdCl_2$ and 60 uM peptide monomer on an Ocean Optics SD 2000 fiber-optic spectrometer. The scans were collected from 350 to 190 nm and the absorbance at λmax was measured as a function of pH using a gel-filled combination electrode (Accumet) attached to a Fisher Accumet model 805 pH meter. All the

solutions were constantly purged with argon to avoid disulfide-linked formations in the metal binding site.

^{113}Cd NMR spectroscopy: All the spectra were recorded at room temperature on a Varian Inova 700 spectrometer (110.92 MHz for ^{113}Cd) equipped with a 5 mm broadband probe and externally referenced to a 0.1 M $Cd(ClO_4)_2$ solution in D_2O. A spectral width was sampled using a 5 us 90° pulse and 0.05s acquisition time, with no delay between scans. Samples were prepared by dissolving ~30 mg of lyophilized and degassed peptide in 500 uL of 10% D_2O/ddH_2O under a flow of argon. Peptide concentration was determined using the Ellman's test which quantifies the aliphatic thiol products from the reaction of Cys thiol of the peptide to cleave the disulfide bond in 5,5'-dithiobis-(2-nitrobenzoic acid) (DTNB), yielding colormetric 2-nitro-5-thiobenzoate (NTB^-).[5] The sample concentrations usually range from 4-6 mM 3SCC. An appropriate amount of a 250 mM $^{113}Cd(NO_3)_2$ solution (prepared from 95% isotopically enriched ^{113}CdO obtained from Oak Ridge National Laboratory) was then added to make a 1:1 Cd(II):3SCC ratio. The final pH was adjusted with concentrated KOH or HCl to reach pH 8.5.The data were processed by using MestRe-C.[6] The free induction decays were zero-filled and processed with an exponential function with a line-broadening value of 100 Hz before Fourier transformation.

^{111m}Cd PAC: All perturbed angular correlation (PAC) experiments were performed with a setup using six detectors at a temperature of 1 (2 °C, which was controlled by a Peltier element. The radioactive cadmium was produced on the day of the experiment at the University Hospital cyclotron in Copenhagen and extracted as described previously,[7] except for the HPLC separation of zinc and cadmium. This procedure may lead to zinc contamination of the sample, but the level of contamination should not interfere with the experiment. The ^{111m}Cd solution (10-40µL)was mixed with nonradioactive cadmium acetate and TRIS buffer. The peptide was then added (dissolved in ion-exchanged water), and the sample was left to equilibrate for 10 min to allow for metal binding. Finally, sucrose was added to produce a 55% w/w solution, and the pH of the solution was adjusted with H_2SO_4 or KOH. To measure the pH, a small volume of sample was removed from the solution to avoid chloride contamination of the sample. The pH reported in was measured at room temperature the following day and corrected to the pH at 1 °C. The pH of solutions buffered by TRIS is temperature dependent. Because of the pH dependence on the temperature of TRIS solutions, the pH of the solution at 1 °C was calculated using pH (1 °C) 0.964 [pH(25 °C)] + 0.86.26. The samples were either used immediately after preparation or left on ice

for up to 2 h until the measurement was started. All buffers were purged with Ar and treated so as to lower metal contamination. The final volume of the samples ranged between 0.05 and 0.5 mL with concentrations of 300 μM peptide and 20 mM TRIS. All fits were carried out with 300-400 data points, disregarding the 3-5 first points due to systematic errors in these. The analytical expression for the perturbation function is known, and five parameters are fitted to the data: ω_0 is a measure of the nuclear quadrupole interaction strength; the asymmetry parameter (η) refers to the symmetry of the structure at the site of the PAC probe. In an axially symmetric structure η is 0, and the largest value of this parameter, namely 1, is found, for example, in a tetrahedral complex with 2A and 2B ligands; the amplitude of the signal, A. If more than one structure at the site of the PAC probe is present in the sample, the relative amplitudes will in general directly reflect the relative population of the different sites. The line shape of the Fourier transformed data can be affected by structural variations at the site of the PAC probe from one molecule to the next, reflected in the relative peak width $\Delta\omega_0/\omega_0$; dynamics similarly affects the PAC signal, and is reflected in the rotational correlation time τ_c. In summary, five parameters (A, η, τ_Q (or ω_0), δ and τ_c) are determined for each structure present at the site of the PAC probe in the sample.

EXAFS: XAS Data Collection and Analysis. The samples containing 2 mM Cd(II) and 8 mM each peptide at pH 8.5 were independently prepared for EXAFS measurements. XAS data were collected at Stanford Synchrotron Radiation Laboratory (SSRL) on beam lines 7-3. A Si(220) double crystal monochromator was detuned to 50% of the maximum intensity for harmonic rejection on beam line 7-3. The beam energy was 3 GeV, and the beam size was 1×10 mm^2. The X-ray energies were calibrated by collecting the absorption spectrum of a Cd foil or $HgCl_2$ references at the same time as the fluorescence data, with the first inflection points assigned as 26,714 eV. The sample temperature was maintained between 9.8 and 12.6 K during data collection using an Oxford liquid helium flow cryostat. XAS data were collected as fluorescence excitation spectra using a 13-element (7-3) Ge solid-state detector array. Spectra were measured with 10 eV increments in the pre-edge region (26,485-26,695 eV), 1.0 eV increments in the edge region (26,695-26,745 eV), and 0.05 Å-1 increments in the EXAFS region (2.75-14.00 Å-1), with integration times of 1s in the pre-edge and edge and 1-20s (k^3-weighted) in the EXAFS region, for a total scan time of 35 min. For all samples, the first and last spectra were compared to confirm the lack of radiation damage during the measurements. Each channel of each scan was examined for glitches. The fluorescence from the good channels was averaged for each sample to give the

raw data. The raw data were normalized to the tabulated X-ray absorption coefficients by using MBACK.[8] Data were converted to k space, k) $[2me(E - E_0)/h^2]^{1/2}$, using a threshold energy (Cd, E_0 = 26 716.3 eV) that was calibrated by fitting the EXAFS data for the crystallographically characterized reference compounds $Cd(ClO_4)_2$ and $Cd(SC(^iPr)_3)_3^-$.[9] Background in EXAFS data was further removed by using a least-squares spline with k^3 weighting. The EXAFS data, $\varnothing(k)$, were fit to eq 2 where Ns is the number of scatterers with atom type s at a distance of R_{as}, $A_{as}(k)$ is the effective backscattering amplitude function, σ_{as}^2 is the Debye-Waller factor (disorder in R_{as}), and $\phi_{as}(k)$ is the phase-shift of the photoelectron wave traveling between the potentials of the absorbing and scattering atoms. The amplitude and phase functions, $A_{as}(k)$ and $\phi_{as}(k)$, were calculated using FEFF 7.02, with an amplitude scale factor (0.9) and the threshold energy E_o calibrated by fitting crystallographically defined model complexes. Raw data were k^3-weighted, and the individual fits were performed in k space (1.5-13.0 $Å^{-1}$), allowing R_{as} and σ_{as}^2 to vary for each shell.[10]

Crystallizations: All peptides were crystallized by sitting drop vapor diffusion experiments at 20 ºC with drops containing equal volumes of peptide (0.75 μL) and precipitant (0.75 μL) solutions. The peptide solution of apo-**GRAND-CS**L12$_D$LL16C was prepared from 20 mg/mL **GRAND-CS**L12$_D$LL16C, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5. The precipitant solution contains 40% (v/v) PEG-400, sodium acetate buffer pH 4.5 at a final well solution pH 5.4. The apo-**GRAND-CS**L16CL19$_D$L solution was prepared from 20 mg/mL peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5. The well solution contains 25% (v/v) PEG-2000 MME and 0.1 M MES pH 6.5. The crystals of Hg(II)**GRAND-CS**L12$_D$LL16C were grown from a peptide solution (20 mg/mL **GRAND-CS**L12$_D$LL16C, 0.92 eq of $HgCl_2$ per 3SCC peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5) and a precipitant solution (0.1 M MES pH 6.5 and 25% (v/v) PEG-2000 MME).The crystals of Hg(II)**GRAND-CS**L16CL19$_D$L were crystallized from a peptide solution (20 mg/mL **GRAND-CS**L16CL19$_D$L, 0.92 eq of $HgCl_2$ per 3SCC peptide, 15 mM $Zn(OAc)_2$ and 0.5 mM Tris buffer pH 8.5) and a well solution (0.2 M Lithium acetate and 20% (v/v) PEG-3350). Crystals were cryoprotected in a mother liquor containing 20% glycerol prior to supercooling in liquid N_2 for data collection.

Data collections and refinements: Data were collected at the Advanced Photon Source of the Argonne National Laboratory on the LS-CAT Beamline 21-ID-F, equipped with a Mar 225 CCD detector, respectively. All data were collected with a 1º oscillation then processed and scaled

with HKL2000[11]. All structures presented in this chapter were solved by molecular replacement using Molrep[12] in the CCP4 suite of programs[13–15], then underwent iterative rounds of electron density fitting and refining in Coot[16] and Buster 2.11.2 program[17], respectively.

The X-ray crystal structures of well-folded, three-stranded parallel coiled coil peptides of apo-(**GRAND-CS**L12$_D$LL16C)$_3$, apo-(**GRAND-CS**L16CL19$_D$L)$_3$, Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$, and Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ were determined to 1.34, 1.83, 2.11 and 1.93 Å resolution, respectively. The crystallographic data of each structure is shown in **Table 3-2**. The apo-(**GRAND-CS**L12$_D$LL16C)$_3$ crystallized in the space group R32, (#155) contains one single strand of peptide per ASU with a Matthew's coefficient of 2.38 corresponding to 47.67% solvent content. The three stranded coiled coil is obtained by the combination of three adjacent symmetric units that are crystallographic imposed by the three fold axis. The structure was solved using a four-heptad chain A of apo-(**CS**L9C)$_3$ (PDB code: 3LJM)[18] to serve as a search model in MolRep in which Cys 9 and Leu 16 were mutated to Leu and Cys, respectively. The phrases calculated from the molecular replacement solution were then used in autobuilding Arp/warp software[19,20] to generate a starting model that has five heptads in length corresponding to the **GRAND-CoilSer** structure. After the first round of rigid body refinement, the D-Leu residue was built at the twelfth position using the difference density present in F_o-F_c map. The structure was refined to 1.42 Å ($R_{working}$ = 19.6%, R_{free} = 20.3 %).

Sharing similar lattice packing of the R32 space group, the refined apo-(**GRAND-CS**L12$_D$LL16C) was subsequently employed to be a search model for apo-(**GRAND-CS**L16CL19$_D$L) by mutating the 12 D-Leu to L-Leu. 19 D-Leu was replaced after the first round of refinement. The solvent content per ASU of this structure is 48.60 % (Matthew's coefficient of 2.37). The structure was refined to 1.83 Å ($R_{working}$ = 20.0%, R_{free} = 20.6 %).

The Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$, assigned to space group $P2_12_12_1$ (#19), was solved using AutoSol Wizard in Phenix.[21–24] To solve the structure, the anomalous difference of heavy atoms, Hg(II) and Zn(II), was determined to generate the experimental phases. The obtained solution reveals a possible three-stranded coiled coil packing per asymmetric unit, yet the third strand was broken in the middle, missing the residues 19Leu, 20Glu, 21Lys and 22Lys.

Table 3-2: Crystallography Data Collection and Refinement Statistics.

Peptide	apo-(**GRAND-CS**L12$_D$LL16C)$_3$	apo-(**GRAND-CS**L16CL19$_D$L)$_3$	Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$	Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$
Data Collection				
Space Group	R32	R32	C2	$P2_12_12_1$
Unit Cell a, b, c (Å)	38.213, 38.213, 140.655	37.898, 37.898, 140.667	70.602, 40.808, 52.438	32.636, 80.508, 88.730
Wavelength (Å)	0.97872	0.97872	0.97856	0.97856
Resolution (Å)[1]	1.42(1.42-1.40)	1.83 (1.87-1.83)	2.11(2.15-2.11)	1.84 (1.87-1.84)
Rsym (%)[2]	5.6 (43.4)	9.4 (48.3)	5.4 (42.1)	12.9 (60.8)
<I/sI>[3]	>50 (2)	>50 (2)	>50 (2)	> 50 (2)
Completeness (%)[4]	99.3 (100)	99.4 (100)	98.6 (100)	97.6 (99.6)
Redundancy	5.6 (5.5)	35.6 (39.8)	4.6 (4.9)	8.3 (7.6)
Refinement				
Resolution (Å)	1.42	1.83	-	1.84
R-Factor (%)[5]	19.6	20.0	-	21.1
Rfree (%)[6]	20.3	20.6	-	22.6
Protein atoms	302	273	-	870
Metal ions	3 Zn(II)	1 Zn(II) on surface	-	1 Hg(II) 3 Zn(II)
Water Molecules	52	44	-	189
Unique Reflections	8093	2584	-	20219
R.m.s.d.[7]			-	
Bonds	0.010	0.010	-	0.010
Angles	1.15	1.01	-	1.08
MolProbity Score[8]	1.11	0.50	-	1.45
Clash Score[8]	3.17	0.00	-	4.2

[1]Statistics for highest resolution bin of reflections in parentheses.

[2]$R_{sym} = \Sigma_h\Sigma_j | I_{hj} - \langle I_h \rangle | / \Sigma_h\Sigma_j I_{hj}$, where I_{hj} is the intensity of observation j of reflection h and $\langle I_h \rangle$ is the mean intensity for multiply recorded reflections.

[3]Intensity signal-to-noise ratio.

[4]Completeness of the unique diffraction data.

[5]R-factor = $\Sigma_h | IF_oI - IF_cI | / \Sigma_h |F_o|$, where F_o and F_c are the observed and calculated structure factor amplitudes for reflection h.

[6]R_{free} is calculated against a 10% random sampling of the reflections that were removed before structure refinement.

[7]Root mean square deviation of bond lengths and bond angles.

[8]Chen et al. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallographica D66:12-21.

By using the $2F_o$-F_c electron density as a guide, all missing residues were built back into the chain to generate the final starting model which consequently served as a search model in MolRep. D-Leu at the nineteenth position was replaced with L-Leu after one round of refinement according to the difference density shown in F_o-F_c map. The Matthew's coefficient is 4.68 corresponding to 73.74% per ASU. The structure was refined to 1.84 Å ($R_{working}$ = 21.1%, R_{free} = 22.6 %). The validity of the models were verified using the MolProbity software.[25] All non-glycine residues of these structures fall in the preferred right handed α-helical region of the Ramanchandran plot. Every side chain is present in the preferred rotameric conformation.

Results

Physical characterizations

Circular Dichromism Spectroscopy (CD)

CD spectra were obtained to determine the effects of introducing a D-amino acid on the secondary and tertiary structure of the coiled coils. The coiled coil and alpha helix have CD spectral signatures in the far-UV region at 208 and 222 nm respectively, as a negatively double-well band. At pH 8.5, the molar ellipticity values at 222 nm designated folded coiled-coil structures reveals an average of -25,468 deg.dmol^{-1}cm^2 for apo-**TRI**L2WL16CL19$_D$L and of -26,212 deg.dmol^{-1}cm^2 for metalated-**TRI**L2WL16CL19$_D$L which corresponds to 71.7±0.1% and 73.8±1.2 % initial folded coiled coil structures, respectively. Representative spectra are shown in **Figure 3-1**. **GRAND-CS**L12$_D$LL16CL19$_D$L has a negative peak at 222 nm with an average of -25,052 deg.dmol^{-1}cm^2 for apo-peptide and of -26,055 deg.dmol^{-1}cm^2 for metalated-peptide corresponding to 70.6±0.9% and 73.4±0.8% initial folding, respectively (**Figure 3-2**).

UV/VIS metal binding titrations

This characterization was performed to investigate Cd(II)-protein interactions in aqueous solution by following the LMCT transition of Cd(II)-trithiolate complex formation. Upon the addition of aliquots of Cd(II) to the **TRI**L2WL16CL19$_D$L peptide solution at pH 8.5, a well-resolved peak at 232 nm was observed. The absorbance of each titration point was plotted as a function of equivalence of Cd(II), showing a steep slope before 1 equivalence of Cd(II) per 3SCC, following by a suddenly sharp turn at 1 equivalence of Cd(II) per 3SCC and leveling off

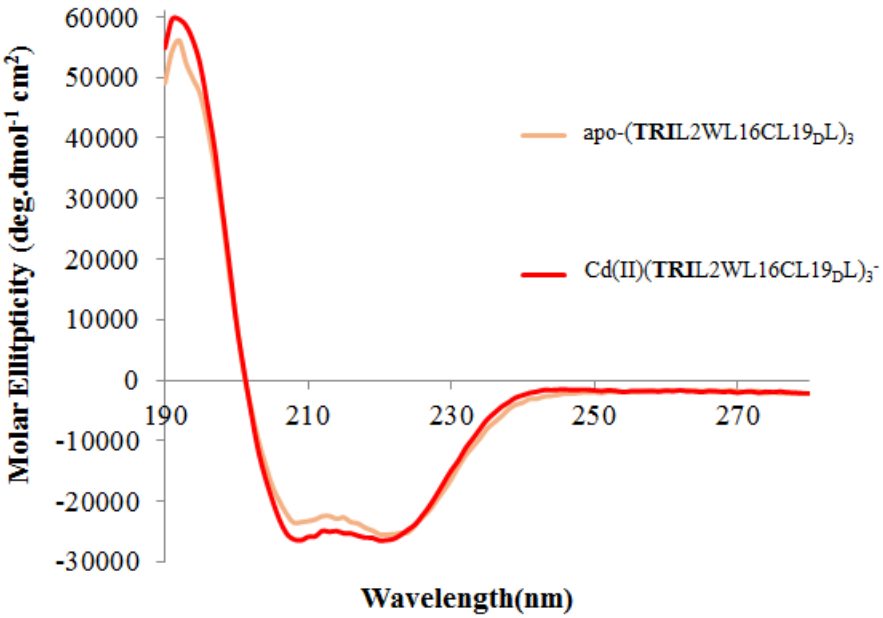

Figure 3-1. Far UV circular dichroism spectra of apo-diastereopeptide and metalated- TRIL2WL16CL19$_D$L monitored by CD. The negatively doubled well at 208 and 222 nm reveals the α-helical coiled coil structure. The light and dark red lines represent the spectra of apo- and Cd(II)-TRIL2WL16CL19$_D$L (1 eq. of Cd(II) per 3SCC), respectively, at micromolar concentration of monomer. The spectra were recorded at pH 8.5 and ambient temperature.

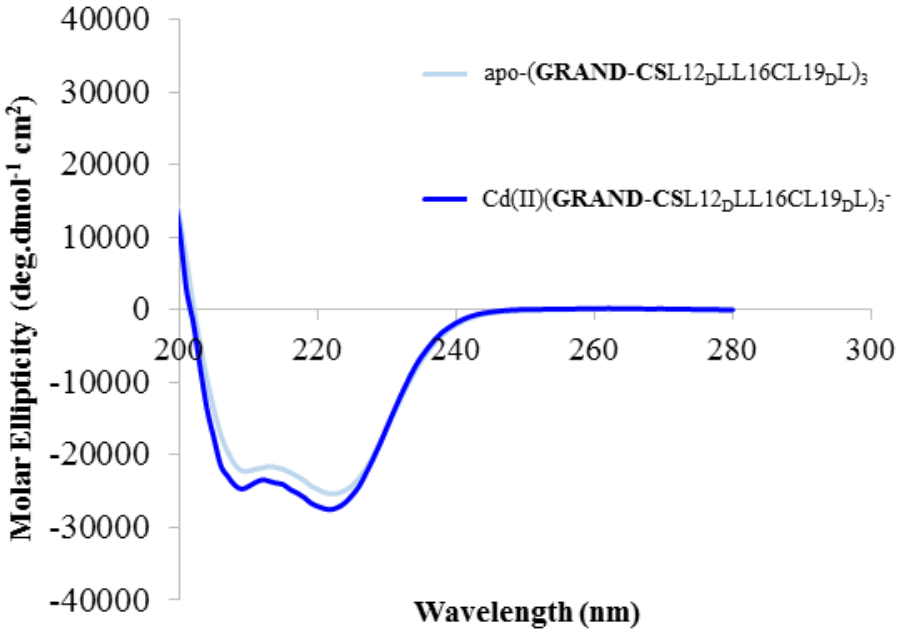

Figure 3-2. Far UV circular dichroism spectra of GRAND-CSL12DL L16CL19DL in the absence and presence of Cd(II). The light and dark blue lines represent the spectra of apo- and Cd(II)-peptide (1 eq. of Cd(II) per 3SCC), respectively, at micromolar concentration of monomer. The spectra were recorded at pH 8.5 and ambient temperature.

afterwards (**Figure 3-3**). This confirmed the expected stoichiometry of Cd(II) incorporated into the design peptide was 1:3. The total concentration of Cd(II) was used to calculate an extinction coefficient of λ_{232} 19,900 $M^{-1}cm^{-1}$. A similar observation was found with **GRAND-CS**L12$_D$LL16CL19$_D$L where the λ_{max} of LMCT band occurs at 232 nm with a final extinction coefficient of 19,700 $M^{-1}cm^{-1}$ at pH 8.5 with 1:1 Cd(II):trimer (**Figure 3-4**).

pH Dependence studies

pH dependent titrations for Cd(II) binding to the various peptides were used to determine pK_{aS} (S is abbreviated from the number of thiol deprotonation events in the equilibrium) of the thiol residues in the 3SCC following the Ligand-to-Metal-Charge-Transfer (LMCT) of three thiolates to Cd(II) observed at 235 nm. When small amounts of concentrated KOH solution were added to a solution containing 1:1 $CdCl_2$:3SCC ratio, the increase in the bands of 235 nm as a function of pH was observed until reaching a flat plateau slope around pH 8.5 as shown in **Figure 3-5**. The pH value curves were fitted using the non-linear regression models for the release of two protons upon Cd(II) binding to three-thiolates[3], assuming $Cd[(peptide)(peptide\text{-}H)_2]^+$ is a starter species which is initially formed at low pH. The equilibria are shown as follows.

$$(peptide\text{-}H)_3 + Cd(II) \leftrightarrow Cd(II)(peptide)(peptide\text{-}H)_2^+ + H^+$$

$$Cd(II)(peptide)(peptide\text{-}H)_2^+ \leftrightarrow Cd(II)(peptide)_2(peptide) + H^+$$

$$Cd(II)(peptide)_2(peptide) \leftrightarrow Cd(II)(peptide)_3^- + H^+$$

The titration curve of **TRI**L2WL16CL19$_D$L yielded pK_{a2} of 13.20 (One at pK_a of 6.50 and the other at pK_a of 6.70) whereas in **GRAND-CS**L12$_D$LL16CL19$_D$L, the peptides yielded pK_{a2} of 13.74 (One at pK_a of 6.70 and the other at pK_a of 7.04).

^{113}Cd NMR

The ^{113}Cd NMR spectra of the different Cd(II) peptide complexes were recorded at pH 8.5 under conditions where the metal ion was fully bound based on pH titration studies. A sharp single resonance was observed at chemical shifts of 600 ppm and 610 ppm for **TRI**L2WL16CL19$_D$L (**Figure 3-6**) and **GRAND-CS**L12$_D$LL16CL19$_D$L (**Figure 3-7**), respectively.

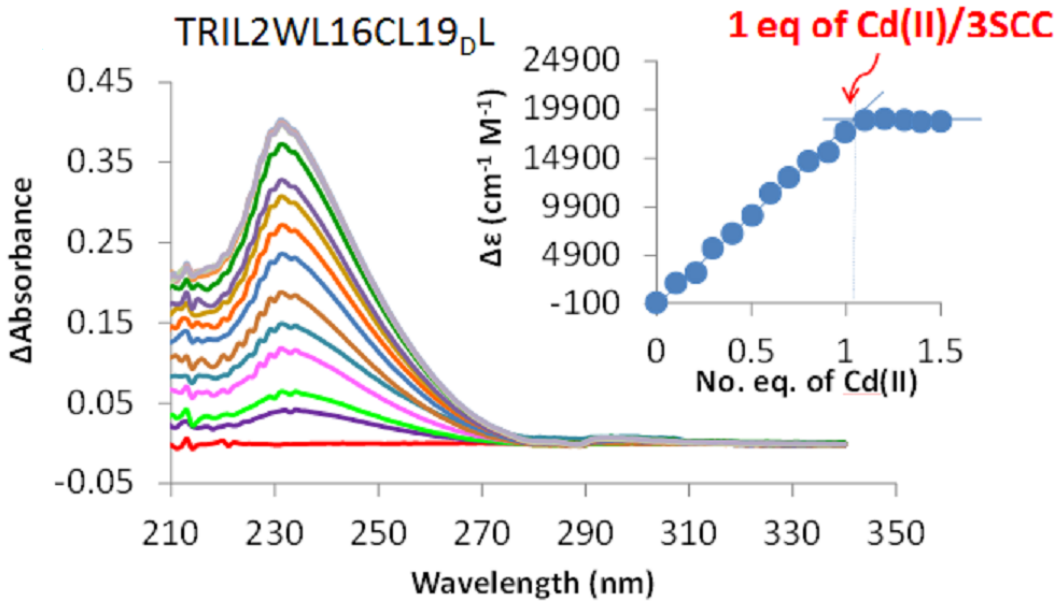

Figure 3-3. UV/VIS Cd(II) binding titrations to $(TRIL2WL16CL19_DL)_3$ peptide. UV/VIS titration of Cd(II) into a solution of $TRIL2WL16CL19_DL$ at pH 8.5. The data were plotted as ΔAbsorbance (M^{-1} cm^{-1}) vs wavelength (nm). The inset to each titration profile shows the titration curve plotted as Δε at λ_{max} vs equivalents of Cd(II) added.

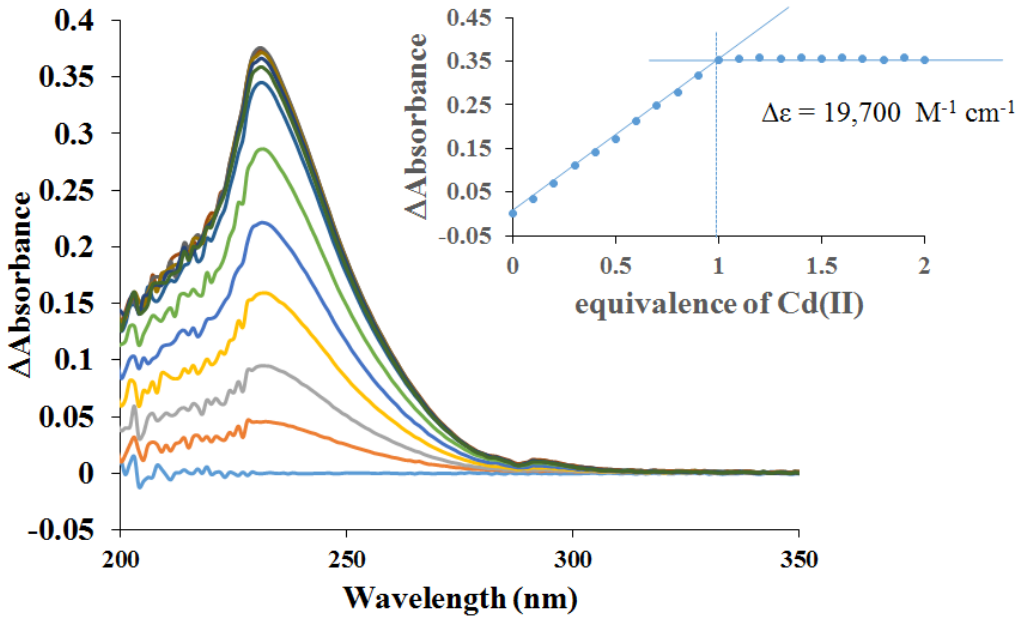

Figure 3-4. UV/VIS Cd(II) binding titration studies of (GRAND-CSL12$_D$L L16CL19$_D$L)$_3$ peptide. UV/VIS titration of Cd(II) into a solution of (GRAND-CSL12$_D$L L16CL19$_D$L)$_3$ at pH 8.5. The data were plotted as ΔAbsorbance (M^{-1} cm^{-1}) vs wavelength (nm). The inset to each titration profile shows the titration curve plotted as Δε at λ_{max} vs. equivalents of Cd(II) added.

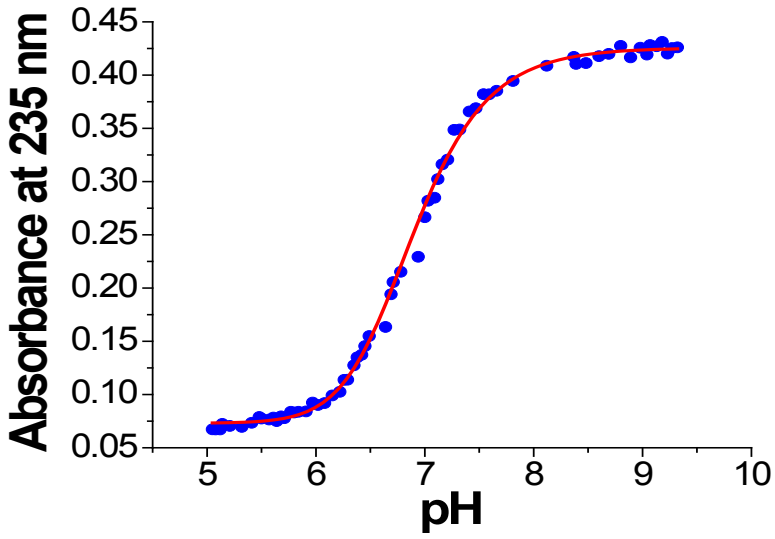

Figure 3-5. pH dependence for the formation of Cd(II)(TRIL2WL16CL19$_D$L)$_3^-$. The red line represents the fit for the pH titration of Cd(II) binding to TRIL2WL16CL19$_D$L to a one proton stepwise dissociation (two total protons) according to a published model.[26]

^{111m}Cd Perturbed Angular Correlation Spectroscopy

^{111m}Cd Perturbed Angular Correlation Spectroscopy was used to assign the Cd(II) coordination geometry of the peptides. The Fourier transform of ^{111m}Cd and **TRI**L2WL16CL19$_D$L monomer in a 1:12 ratio at pH 8.6, 1 °C shows two clear NQIs at ω_o = 0.316 rad/ns (η = 0.17) and ω_o = 0.159 rad/ns (η = 0.33) with equal amplitudes indicating 50:50 mixtures of two species (**Figure 3-8**). **GRAND-CS**L12$_D$LL16CL19$_D$L also reveals two values of angular frequencies, both of which are similar to **TRI**L2WL16CL19$_D$L. The first NQIs at ω_o = 0.330 rad/ns (η = 0.26) and ω_o = 0.170 rad/ns (η = 0.39) with a 70:30 ratio suggests the major species belongs to the higher angular frequency at pH 8.6, 1°C (**Figure 3-9**). The ^{111m}Cd PAC parameters are given in Table **3-3**.

EXASF

XAS technique was used to obtain precise metrical parameters for the Cd(II) coordination environment. The EXAFS data for Cd(II)(**TRI**L2WL16CL19$_D$L)$_3^-$ is shown in **Figure 3-10**. The Fourier transform of the data gave a single peak. The best fit to the k^3-weighted EXAFS (shown as the red line) was obtained using three sulfur atoms at a Cd(II)-S distance of 2.48 Å and a Cd(II)-O distance of 2.37 Å. The EXAFS data for Cd(II)(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3^-$ is shown in **Figure 3-11**. The best fit to the k^3- weighted EXAFS (shown as the red line) was obtained using three sulfur atoms a Cd(II)-S distance of 2.48 Å and a Cd(II)-O distance of 2.28 Å.

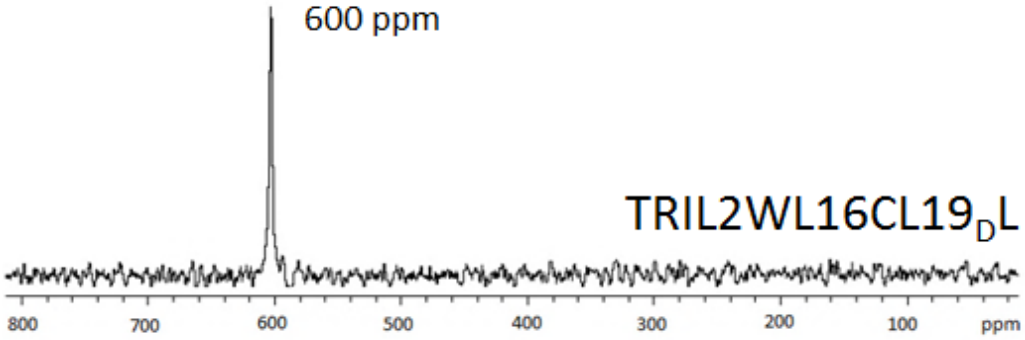

Figure 3-6. ^{113}Cd NMR of Cd(II)(TRIL2WL16CL19$_D$L)$_3^-$. The spectrum was recorded at pH 8.5.

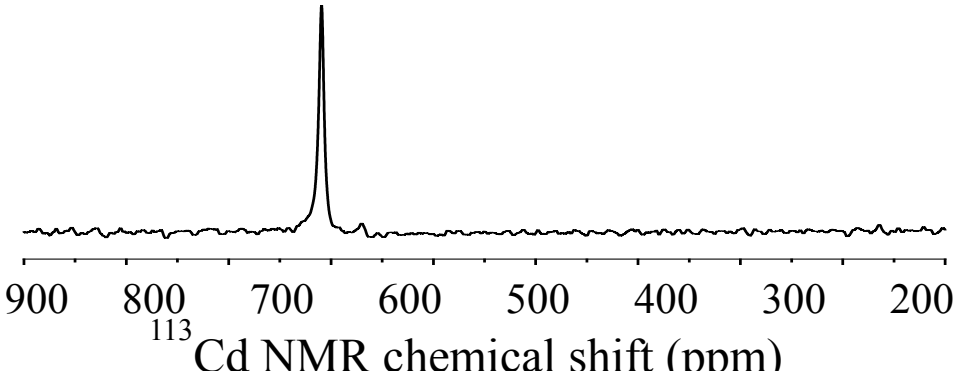

Figure 3-7. ^{113}Cd NMR of Cd(II)(GRAND-CSL12$_D$LL16CL19$_D$L)$_3^-$. The spectrum was recorded at pH 8.5.

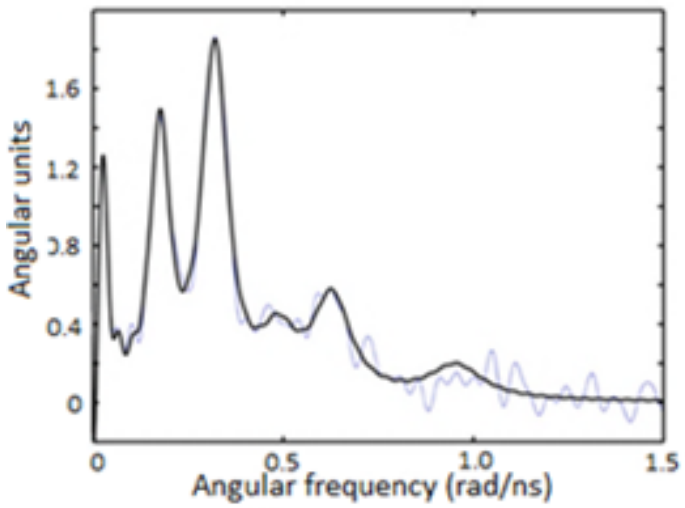

Figure 3-8. ^{111m}Cd PAC spectrum of Cd(II)(TRIL2WL16CL19$_D$L)$_3^-$. The Fourier transforms (the thin line is the Fourier transform of the experimental data and the bold line is the Fourier transform of the fit) of Cd(II)(TRIL2WL16CL19$_D$L)$_3^-$. Data was collected and fit by Prof. Lars Hemmingsen, University of Copenhagen.

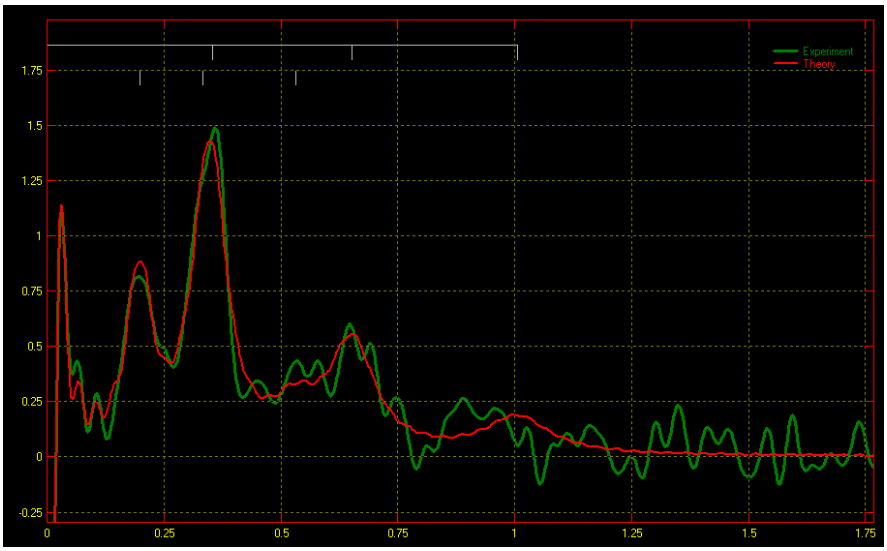

Figure 3-9. ^{111m}Cd PAC spectrum of Cd(II)(GRAND-CSL12$_D$LL16CL19$_D$L)$_3^-$. The Fourier transforms (the thin line is the Fourier transform of the experimental data and the bold line is the Fourier transform of the fit) of Cd(II)(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3^-$. Data was collected and fit by Prof. Lars Hemmingsen, University of Copenhagen.

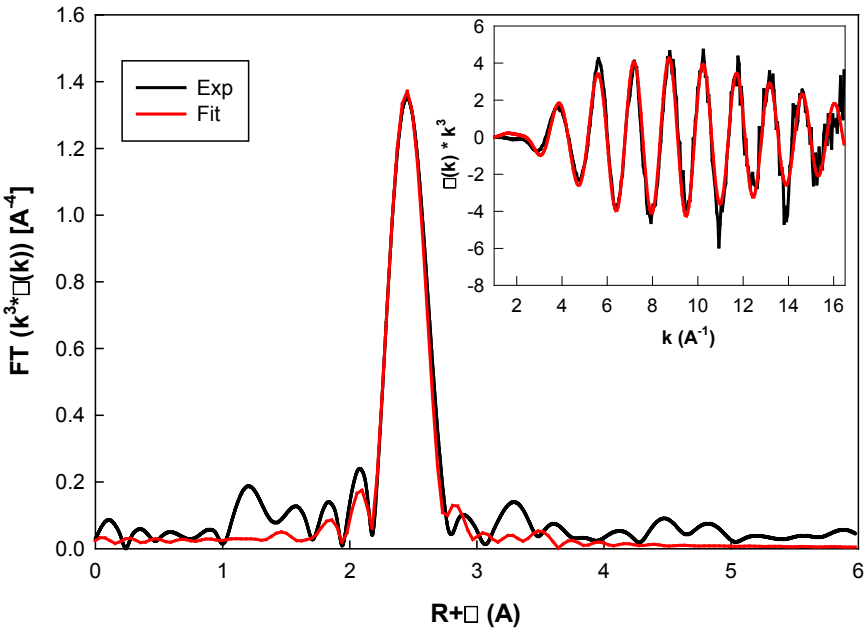

Figure 3-10. Extended X-Ray Absorption Fine Structure data for Cd(II)(TRIL2WL16CL19$_D$L)$_3^-$. In each graph, the black line corresponds to the experimental data and the red line to the fit. The insets contain the data in the k space before Fourier transformation. The data were fit by Dr. Aniruddha Deb, in collaboration with Professor James Penner-Hahn.

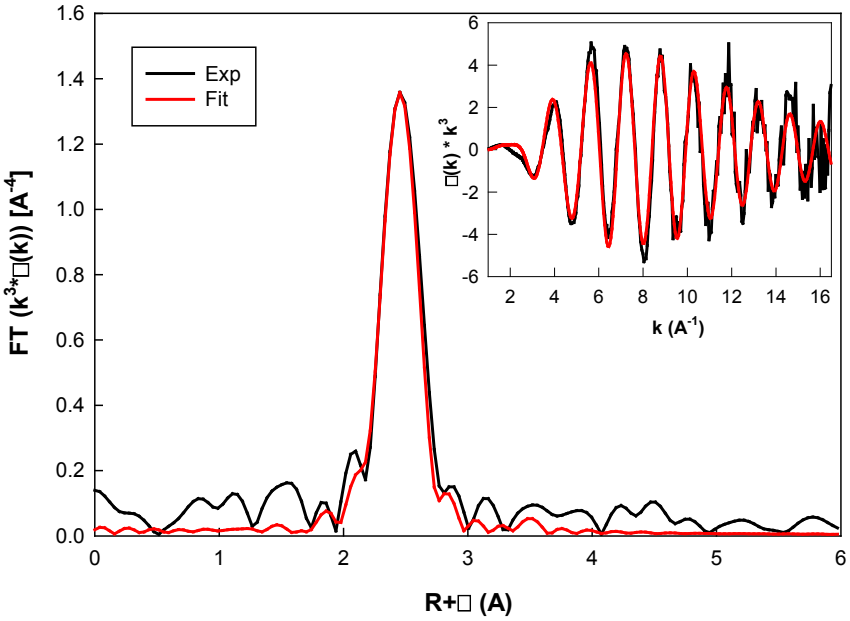

Figure 3-11. Extended X-Ray Absorption Fine Structure data for Cd(II)(GRAND-CSL12$_D$LL16CL19$_D$L)$_3^-$. In each graph, the black line corresponds to the experimental data and the red line to the fit. The insets contain the data in the k space before Fourier transformation. The data were fit by Dr. Aniruddha Deb, in collaboration with Professor James Penner-Hahn.

Table 3-3: ^{113}Cd NMR chemical shifts and parameters fitted to the ^{111m}Cd PAC data

Peptide	^{113}Cd NMR (ppm)	^{111m}Cd PAC parameters[a]							
		pH	Temperature (^{o}C)	ω_o (rad/ns)	η	$\Delta\omega_o/\omega_o$ (x100)	$1/\tau_c$ [μs^{-1}]	A (x100)	χ_r^2
TRIL16C[b]	625	8.7	1	0.337(8)	0.23(2)	5.1(7)	8(5)	5.1(6)	1.10
				0.438(6)	0.20(3)	5.4(7)	3(3)	3.6(4)	
TRIL12$_D$LL16C[c]	697	9.1	1	0.456(4)	0.10(1)	0.7(2)	14(1)	8.2(2)	1.18
TRIL2WL16CL19$_D$L	600	8.6	1	0.316(2)	0.17(4)	8(1)	7.3(7)	5.9(4)	1.12
				0.159(2)	0.33(3)	9(1)	7.3(7)	5.9(3)	
GRANDCSL12$_D$LL16CL19$_D$L	610	8.6	1	0.330	0.26(4)	7(2)	9.7(2)	7(2)	0.68
				0.170	0.39(6)	8(3)	9.7(2)	3(5)	

[a] The numbers in parentheses are the standard deviations of the fitted parameters. [b] Data from reference 3. and [c]Data from reference 4. ^{113m}Cd PAC data of the peptide samples were collected and fitted by Professor Lars Hemmingsen.

Structural Characterizations

The X-ray crystal structures of well-folded, 3SCC peptides of apo-(**GRAND-CS**L12$_D$LL16C)$_3$, apo-(**GRAND-CS**L16CL19$_D$L)$_3$, Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$, and Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ were determined to 1.34, 1.83, 2.11 and 1.93 Å resolution, respectively **(Figure 3-12)**. The results of the following crystallographic evidence (χ1 dihedral angles, atomic distances, interlayer distances, M-S distances, etc.) are referred to **Table 3-4**.

First Coordination Sphere: the sixteenth position

*apo-(**GRAND-CS**L12$_D$LL16C)$_3$*

According to crystallographic constraints imposed by the three-fold axis of R32, the structure of apo-(**GRAND-CS**L12$_D$LL16C)$_3$, the three helical strands display perfect three-fold symmetry. The Cys residues at the sixteenth position are composed of two alternate conformations **(Figure 3-13)**. The Sγ of the major conformer, at 56% occupancy, is oriented to the coiled coil interior and directed toward the N-terminus, corresponding to the *a* site Cys orientation found in the Coil V$_a$L$_d$ crystal structure.[27] The Sγ-Sγ separation is 3.45 Å. This Sγ-Sγ distance of thiols is too long to represent disulfide formation (unlikely at pH 6.5 where the crystal was grown). Moreover, the cysteines should also be protonated at this pH. The other Sγ conformer, with 30% occupancy, is pointing toward the helical interface and at the same time downward toward the C-terminus with the Sγ-Sγ distance of 5.38 Å. The Cβ carbons of both Cys orientations are pointing similarly toward the N-terminus of the helical scaffold. The χ_1 dihedral angle for the inner Cys orientation is -66.77^o corresponding to the most common χ_1 value for cysteine rotamers of -65.20^o.[28] In contrast, the χ_1 torsion angle for the outer Cys orientation is -166.66^o indicating that the thiols in both conformers are positioned differently through space. The apo-(**GRAND-CS**L12$_D$LL16C)$_3$ structure is assigned as an apo-structure due to the lack of 3σ contour density at the center of the metal binding site in the difference map.

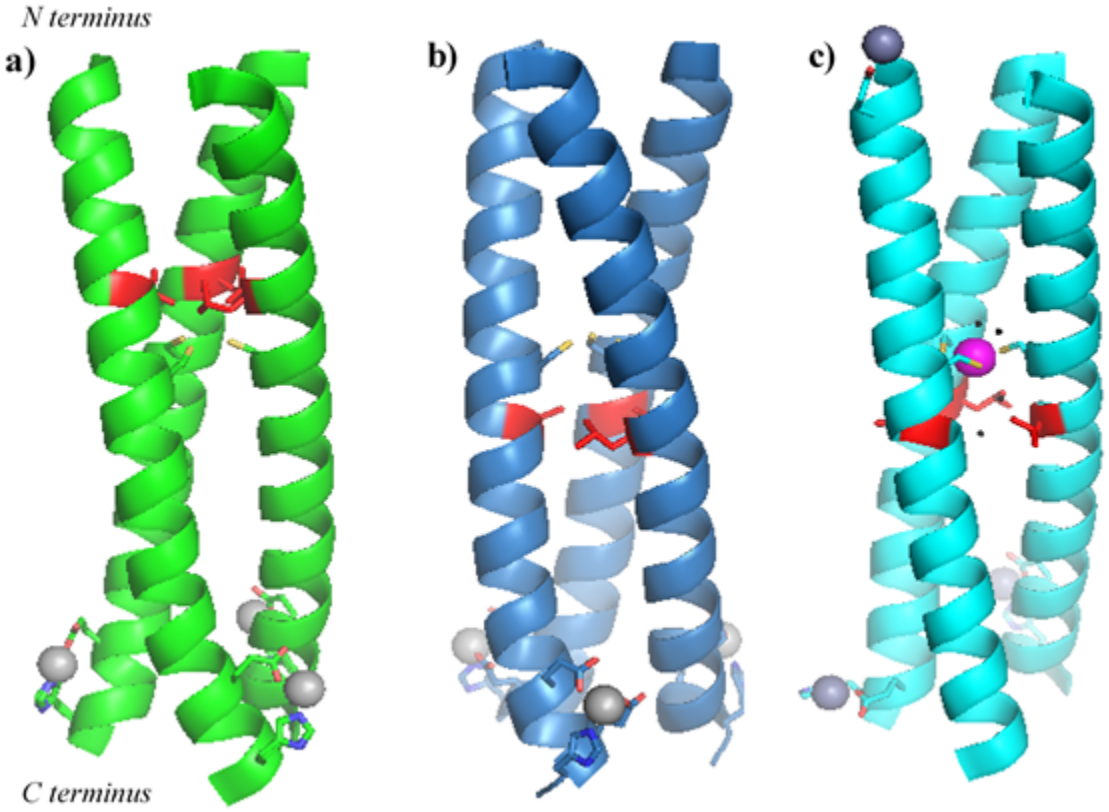

Figure 3-12. Side view of trimeric diastereopeptide structures. Shown in a) apo-(**GRAND-CS**L12$_D$LL16C)$_3$ (green), b) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ (blue) and c) Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ (cyan). D-Leu(red), Cys and Zn(II) coordinating ligands side are shown as sticks. Hg(II) and Zn(II) are shown as spheres in pink and gray respectively. Waters are shown in gray small spheres.

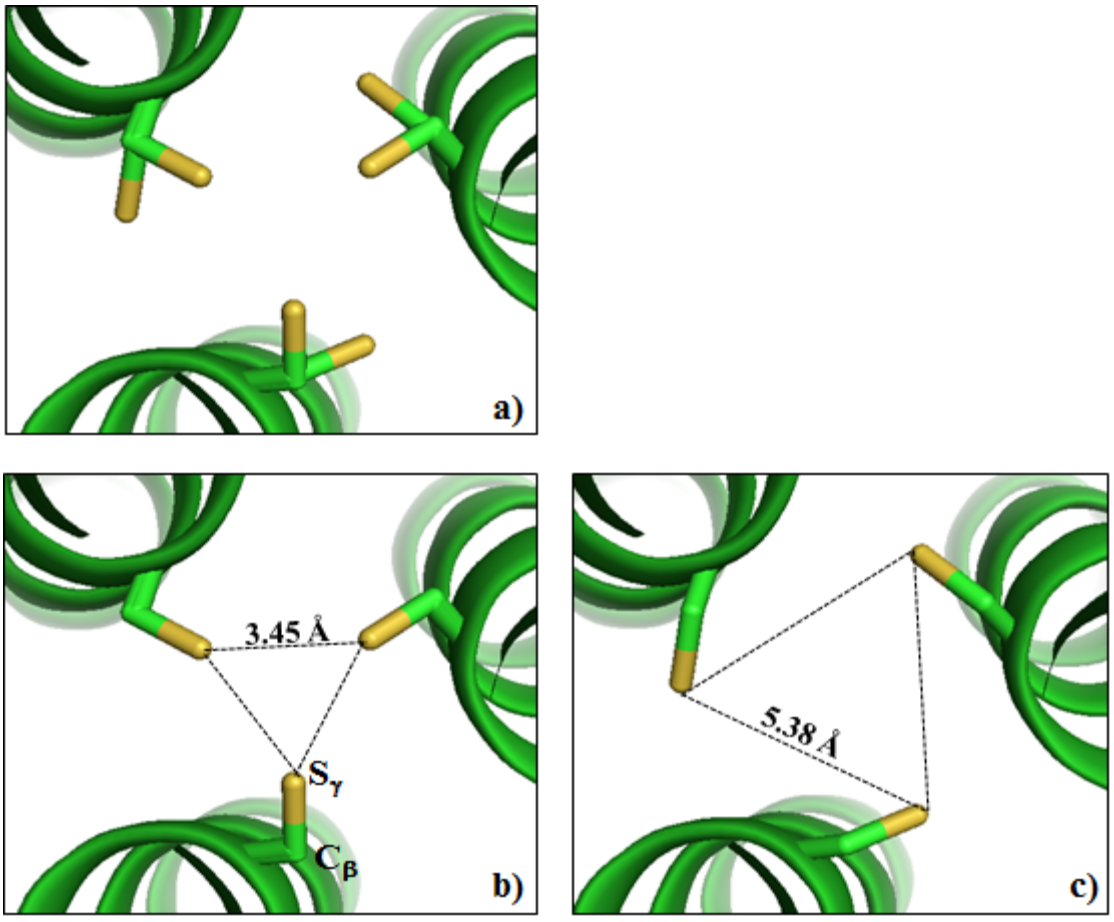

Figure 3-13. PyMOL representation showing the orientations of Cys side chains in the sixteenth layer of apo-(GRAND-CSL12$_D$LL16C)$_3$. The Cys side chains are shown as green sticks with the thiol groups labeled in yellow. a) Top-down view from the N-termini represents the major and minor conformers of Cys residues. b) Top-down view of the major conformers. c) Top-down view of the minor conformers.

Table 3-4: Crystallographic evidence observed from the crystal structures

Peptides	apo-(**CS**L16C)$_3$	Hg(II)$_S$Zn(II)$_N$ (**GRAND-CS**L16CL30H)$_3^+$	apo-(**GRAND-CS**L12$_D$LL16C)$_3$	apo-(**GRAND-CS**L16CL19$_D$L)	Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$
16Cys rotamers					
χ_1 (interior rotamers)[a]	-69.20° (average)	-151.88°	-66.77°	-61.17°	-152.49° (average)
S_γ -S_γ distance[b] (Å)	3.29 (average)	4.16	3.45	3.31	4.21(average)
χ_1 (exterior rotamers)	174.45°,176.96°[j]	173.35°	-166.66°	-	-51.32°, -89.72°[#]
S_γ -S_γ distance[c] (Å)	5.36 (average)	5.24	5.38	-	3.65(average)
Metal site					
M-S bond length (Å)	-	2.43, Hg(II)-S	-	-	2.43(average), Hg(II)-S
S-M-S angle(average)	-	118.36°	-	-	119.75°
Distance of metal from to the bound Cys plane[d] (Å)	-	-0.3	-	-	-0.12
12Leu rotamers			(**D-Leu**)	(**L-Leu**)	(**L-Leu**)
χ_1 (chain A, B, C)[e]	-81.68°, -75.68°,-171.24°	-124.20°	119.78°	-154.78°	-159.63°, -67.37°,-134.35°
χ_2 (chain A, B, C)[f]	-171.45°, 171.07°, 54.60°	-172.02°	-40.73°	48.20°	151.55°,177.78°,-163.59°
Interior C_δ separation[g] (Å)	4.39	4.30	3.80	4.04	4.22
Exterior C_δ separation[h] (Å)	6.61	5.57	5.82	6.45	6.22
Distance of the layer from the interior sulfur plane[i] (Å)	5.08	6.07	2.32	4.72	~6.17
19Leu rotamers			(**L-Leu**)	(**D-Leu**)	(**D-Leu**)
χ_1 (chain A, B, C)	-78.93°, -65.99°,-62.92°	-69.49°	-71.34°	66.04°	82.94°, 58.16°,60.87°
χ_2 (chain A, B, C)	168.54°, 176.10°,175.32°	-172.87°	-178.68°	75.20°	-117.23°, 73.66°,-158.89°
Interior C_δ separation (Å)	4.71	5.95	5.42	5.71	9.35(average)
Exterior C_δ separation (Å)	7.69	8.98	8.21	8.96	10.42(average)
Distance of the layer from the interior sulfur plane (Å)	4.49	3.49	4.04	5.18	3.76

[a] χ_1 of Cys residue is determined from the dihedral angle of N-C_α-C_β-S_γ; [b] The distance determined between S_γ atoms of the interior Cys conformers.; [c] The distance determined between S_γ atoms of the exterior Cys conformers.; [d] Plus sign (+) indicates the metal is situated above the bound Cys plane. Minus (-) indicates the metal is situated below the bound Cys plane.; [e] χ_1 of Leu residue is determined from the dihedral angle of N-C_α-C_β-C_γ.; [f] χ_2 of Leu residue is the dihedral angle measured from C_α-C_β-C_γ-$C_{\delta 1}$.; [g] Interior C_δ separation define the average distance between the interior C_δ atoms of Leu residues of two adjacent chains.; [h] Exterior C_δ separation definesthe average distance between the exterior C_δ atoms of Leu residues of two adjacent chains.; [I] Distance determined between the interior C_δ (Leu) layer and the inner S_γ (Cys) layer and [j] χ_1 dihedral angles measured from chain A and C, respectively.

*apo-(**GRAND-CS**L16CL19$_D$L)$_3$*

Likewise, apo-(**GRAND-CS**L16CL19$_D$L)$_3$ is restricted by the crystallographic three-fold symmetry of R32 space group. The Cys side chain presents as a single rotamer where both the Cβ and Sγ atoms are oriented at the core of the helices and also toward the N-terminus (**Figure 3-14**). The Sγ-Sγ separation is 3.31 Å. The side chain χ_1 angle is -61.17°.

*Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$*

The crystallographic model for the Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ peptide clearly demonstrates a coordinated Hg(II) is found in a center of trigonal planar plane of three Cys residues (major conformers) as indicated by the difference electron density (F_o-F_c) contoured at 3σ from the beginning round of refinements (**Figure 3-15**). The resulting Cys ligands in chain A and chain B contain two conformations in which only the dominant orientation, pointing toward the center of the core, plays a role in metal binding (**Figure 3-16**). The third ligand comes from a single Cys rotamer of chain C. The average χ_1 angle of the Cys coordinating ligand is -152.49°, analogous to the value of -151.88° for the Cys-bound conformation in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ reported in the previous chapter (**Table 3-4**). The average Sγ-Sγ distance is 4.20 Å. The Hg(II) atom is located in the trigonal plane with an average S-Hg(II)-S angle of 119.75°. The refined Hg(II)-S distances are 2.41 Å, 2.43 Å and 2.46 Å, resulting in a mean Hg(II)-S distance of 2.43 Å. These bond angles and bond length observations are not only consistent with the previous Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$structure but also with published trigonal Hg(II)-thiolate small molecule complexes.[29–31]

Second Coordination Sphere: the twelfth and nineteen positions

*apo-(**GRAND-CS**L12$_D$LL16C)$_3$*

The packing of the D-Leu layer situated above the metal site is shown in **Figure 3-17**,**a**. The whole side chain of this residue orients toward the center of the coiled coil. Here, the β-methyl carbon of D-Leu is deviated from its L-counterpart due to the D-configuration effect. Instead of directing toward the N-terminus, Cβ atoms of D-Leu orient oppositely through space, toward the other end of the helices. This change in chirality leads to a twist of the δ-methyl groups, which as

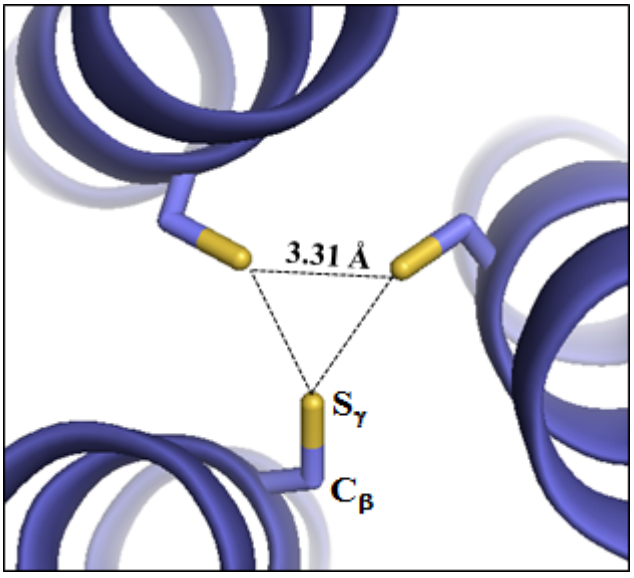

Figure 3-14. Top-down view from the N-termini showing the orientation of Cys side chains in the sixteenth layer of apo-(GRAND-CSL16CL19DL)3. The Cys side chains are shown as blue sticks with the thiol groups color-coded in yellow.

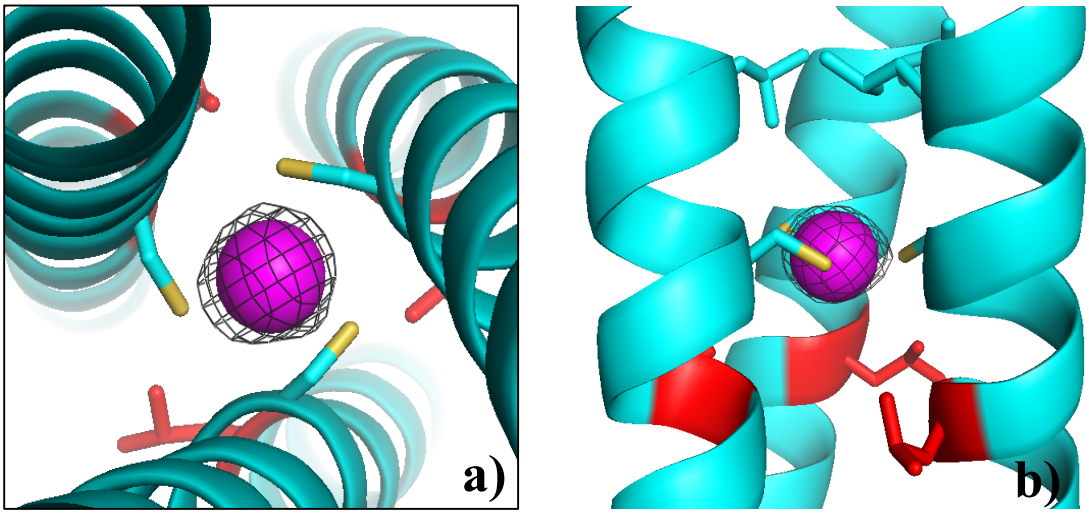

Figure 3-15. Top-down view from the N-termini representing the Hg(II) ion coordinated to the three major cysteine residues in Hg(II)(GRAND-CSL16CL19$_{D}$L)$_3^-$ structure. The metal-binding site with an F_O-F_C omit map contoured at 3σ demonstrating a trigonal planar Hg(II)S_3^- structure.

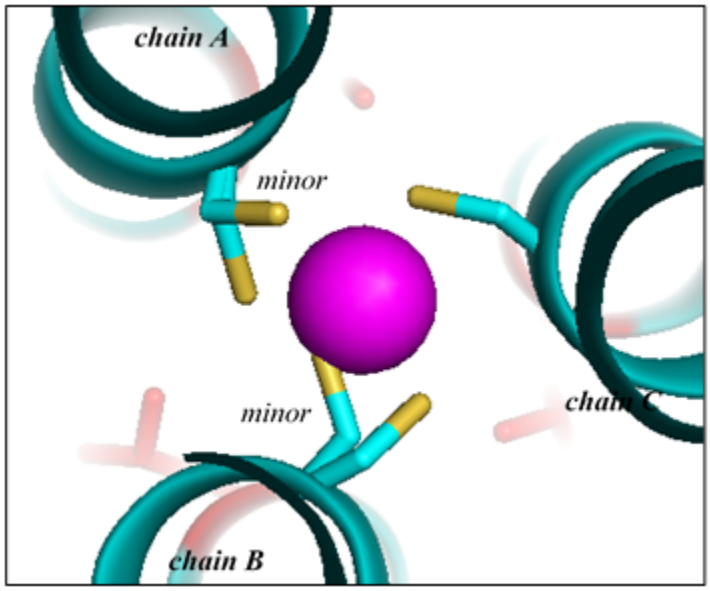

Figure 3-16. PyMOL representation showing the 16Cys conformers in the metal binding layer of Hg(II)(GRAND-CSL16CL19DL)$_3^-$. The Cys side chains are shown as sticks with the thiol groups labeled in yellow. Hg(II) ion is shown as a pink sphere.

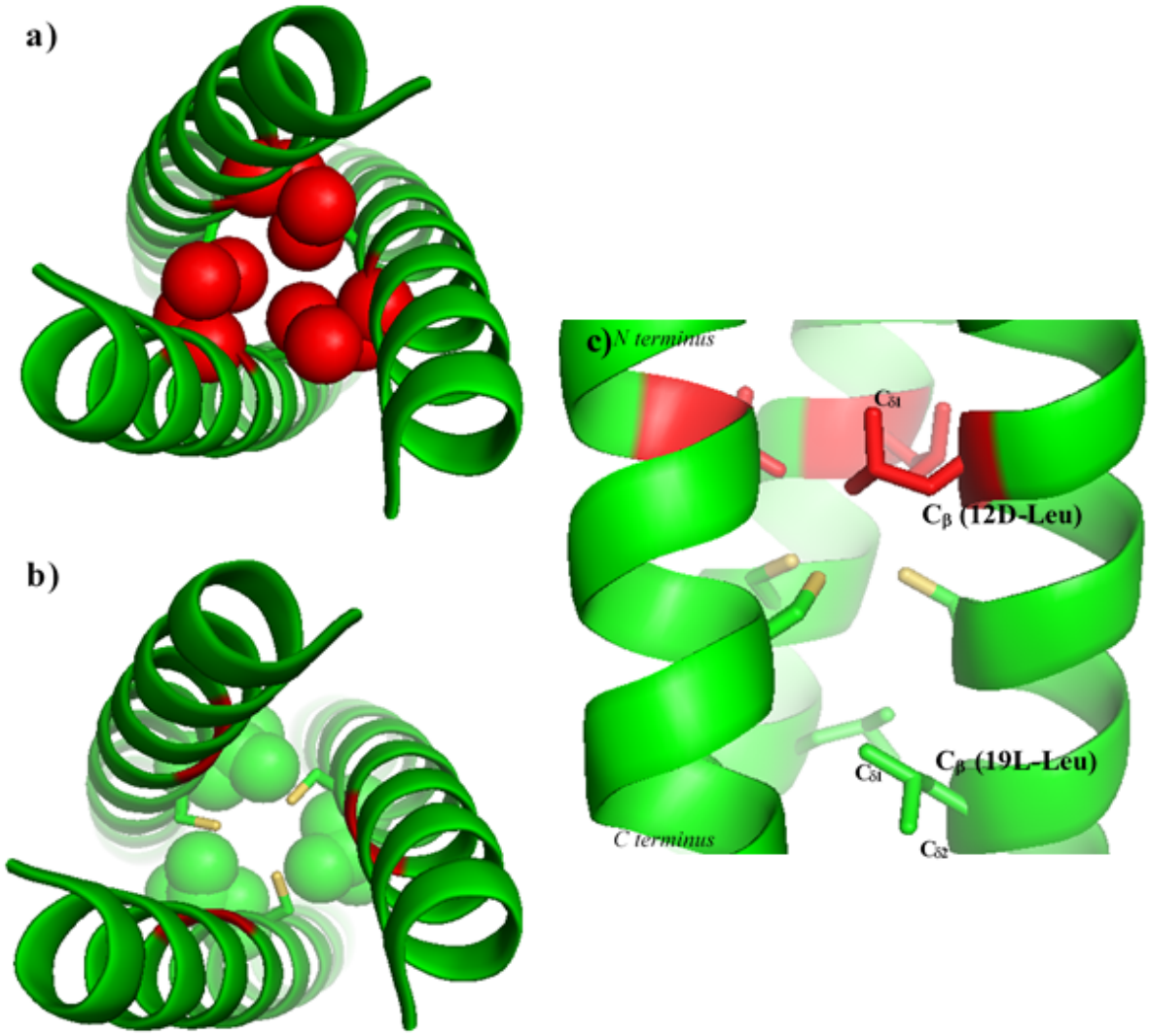

Figure 3-17. PyMOL Representation of the 12D-Leu and 19L-Leu layers of apo-(**GRAND-CSL12DLL16C)3.** Top down view from the N-termini demonstrating a) the 12D-Leu packing shown as red spheres, b) the 19L-Leu packing shown as green spheres. c) Side view of the 12D-Leu layer (red sticks) and the 19L-Leu layer(green sticks) above and below the 16Cys site (green sticks with thiols colored coded in yellow).

a consequence causes the $C_{\delta 2}$ atoms to become the lowest atom of this D-Leu residue, positioning it down toward the Cys plane **(Figures 3-17, b** and **c)**. The hydrophobic interlayer space above the metal site is determined from the distance between the planes formed by the interior $C_{\delta 2}$ atom of 12D-Leu to the interior S_{γ} conformer of 16Cys (2.32 Å) (**Figure 3-18**). The 19L-Leu residues are oriented differently from 12D-Leu as expected due to the change in chirality, with the C_{β} atoms positioned toward the N-terminus. The side chains of 19L-Leu are directed slightly toward the helical interface compared to the 12D-Leu in **Figures 3-17**,**b**. With this arrangement, the $C_{\delta 1}$ plane is the closest plane below the Cys layer (**Figure 3-17**,**c**). This plane is far from the interior Cys conformer plane (4.04 Å) indicating that the space below the metal site is larger than the one above.

*apo-(**GRAND-CS**L16CL19$_D$L)$_3$*

In the layer at the nineteenth position, the deviation of C_{β} atoms causes the reorientation of D-Leu side chains to the C-terminus and more toward the helical interface, subsequently resulting in a large separation between the two layers. The $C_{\delta 1}$ in 19D-Leu **(Figure 3-19**,**a)** forms the closest plane below the metal site and the S_{γ} atom in the 16 Cys planes with a separation of 5.18 Å. With the L-configuration, the β-methyl carbons of 12L-Leu layer point toward the N-terminus. Thus, the side chain atoms are directed toward the center of the structure as shown in **Figure 3-19**,**b**. The $C_{\delta 1}$ atom of 12 L-Leu is interior and becomes the lowest atom in the layer above the metal binding site. The $C_{\delta 1}$ plane is separated by 4.72 Å far from the interior Cys plane.

*Hg(II)-(**GRAND-CS**L16CL19$_D$L)$_3$*

Similarly, the 19D-Leu side chains are oriented away from the center of the coiled coil according to the C_{β} perturbation **(figure 3-20)**.The separation between the interior $C_{\delta 1}$ of 19D-Leu to the major S_{γ} planes is 3.76 Å. The 12L-Leu packing can be seen in **Figure 3-20**,**b**. This layer is at a distance of 6.17 Å above the Cys plane. The interlayer spacing is shown in **Figure 3-18**.

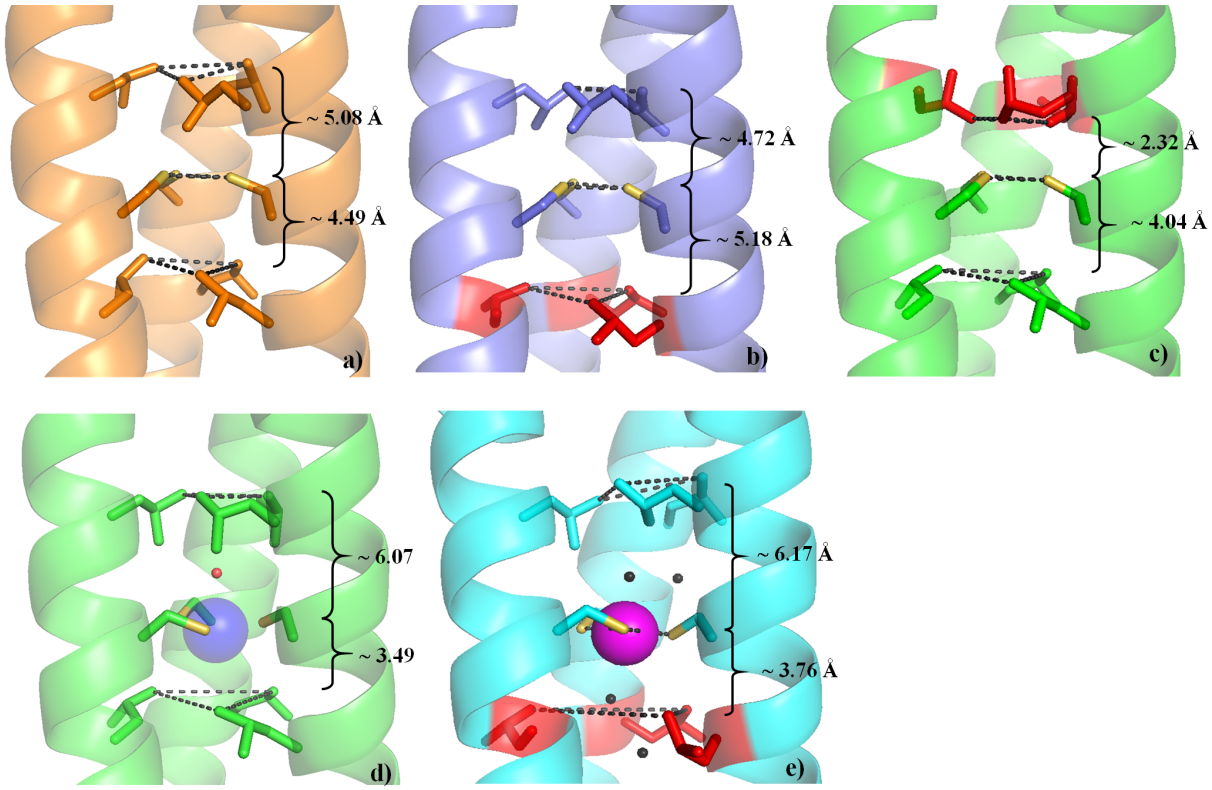

Figure 3-18. Interlayer spaces around the 16Cys layer of diastereopeptides. a) apo-(**CS**L16C)$_3$, b) apo-(**GRAND-CS**L16CL19$_D$L)$_3$, c) apo-(**GRAND-CS**L12$_D$LL16C)$_3$, d) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and e) Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. Main atoms are shown as ribbon diagrams, Residue side chains are present as sticks. Metal centers are shown as spheres.

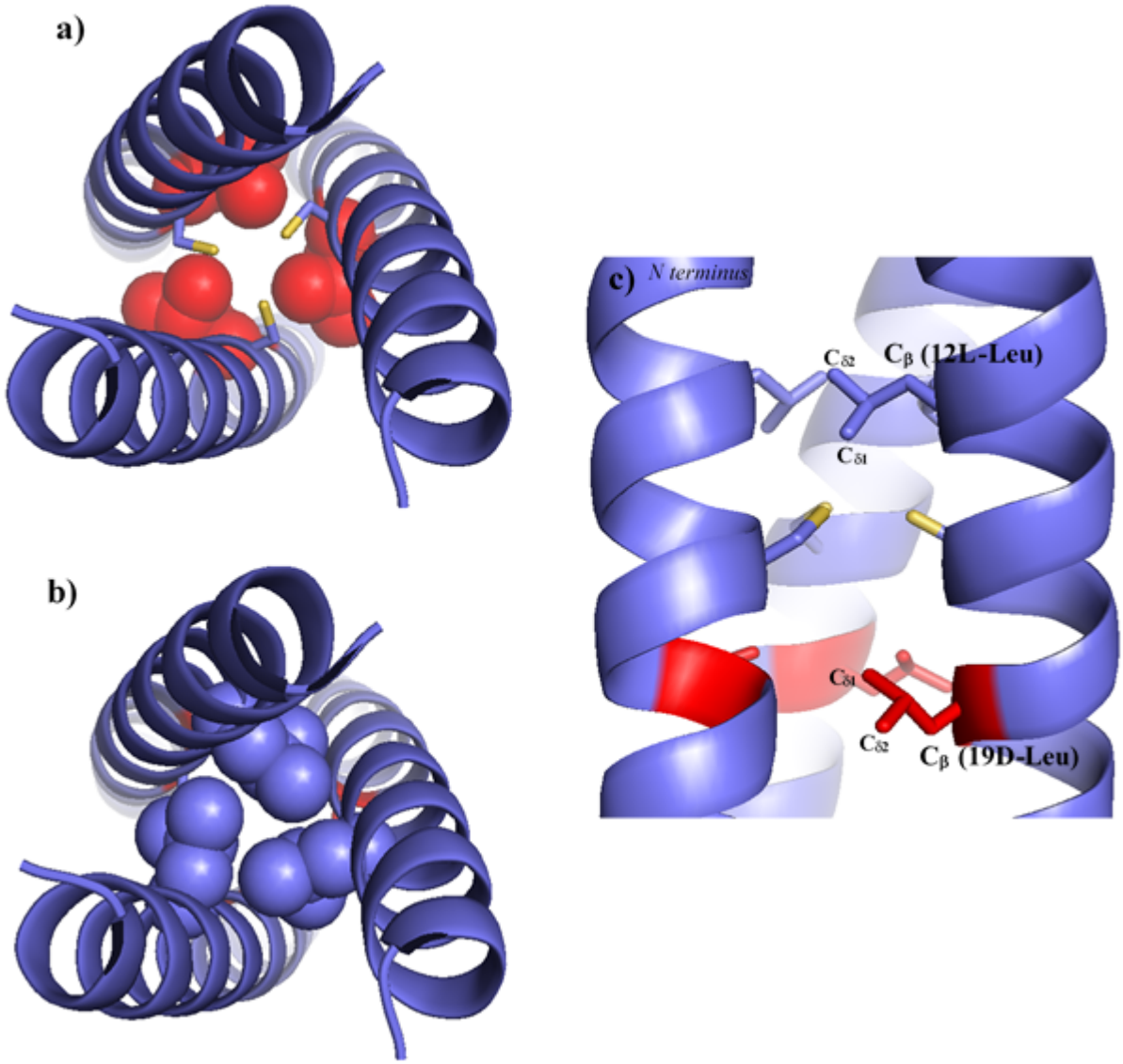

Figure 3-19. PyMOL Representation of the 19D-Leu and 12L-Leu layers of apo-(GRAND-CSL16CL19DL)3. Top down view from the N-termini demonstrating a) the 19D-Leu packing shown as red spheres, b) the 12L-Leu packing shown as blue spheres. c) Side view of the 19D-Leu layer (red sticks) and the 12L-Leu layer(blue sticks) below and above the 16Cys site (blue sticks with thiols labeled in yellow color).

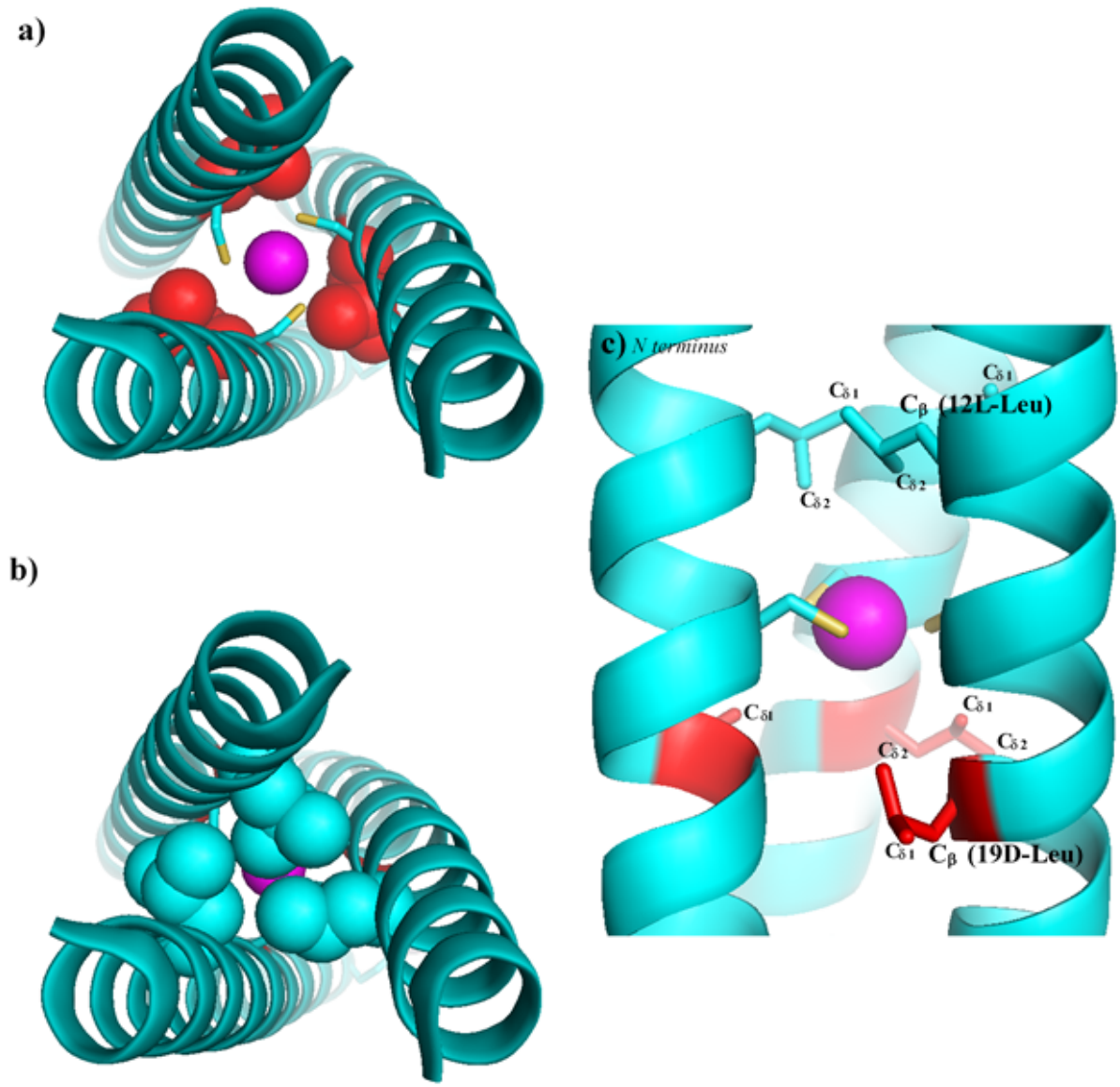

Figure 3-20. PyMOL Representation of the 19D-Leu and 12L-Leu layers of Hg(II)(GRAND-CSL16CL19DL)$_3^-$. Top down view from the N-termini demonstrating a) the 19D-Leu packing shown as red spheres, b) the 12L-Leu packing shown as cyan spheres. c) Side view of the 19D-Leu layer (red sticks) and the 12L-Leu layer(cyan sticks) below and above the 16Cys site (cyan sticks with thiols being in yellow color). (Note: The experimentally observed waters are not shown in this figure.)

Discussion

Physical characterization

One of the true tests for the complete understanding of metalloprotein design is the capability to force specific coordination structures on the metal site. This objective is desired in order to optimize stability and control reactivity. One approach to achieve this control with *de Novo* metalloprotein design is to exploit changes in amino acid stereochemistry. This objective is explored here for Cd(II).[4,32] While not used in nature, the advanced concept of using the alternative configuration of D-amino acids as an approach to control metal site properties is now shown to be a viable strategy.

Encouraged by the success from the previous physical characterization results for CdS_3 formation with the **TRI**L12$_D$LL16C peptide[4], I attempted to obtain higher coordination numbers for Cd(II), either Cd(II)S_3O^- or Cd(II)$S_3O_2^-$, using alternative chirality of D-Leu. It was hypothesized that when D-Leu was placed in the nineteenth position, one layer below the metal site (the sixteenth position), the altered configuration of D-Leu would make the side chain point down toward the C-terminus. Such a modification would provide room for a water ligand to access Cd(II)S_3O^- or Cd(II)$S_3O_2^-$ complexes within this **TRI**L2WL16CL19$_D$L diastereopeptide design (the Tryptophan residue (W) at the second position is for spectroscopic tag).

The effect of introducing a D-amino acid on the secondary and tertiary structure of the coiled coils in aqueous solution at pH 8.5 was first determined by CD spectroscopy. In **Figure 3-1**, the coexistence of the two bands at 208 and 222 nm designated $n\rightarrow\pi^*$ and $\pi\rightarrow\pi^*$ transitions between the backbone amide and carbonyl groups, is the signature of α-helical and coiled coil structures, respectively.[33] Based on the molar ellipticity at 222 nm, Cd(II)(**TRI**L2WL16CL19$_D$L)$_3^-$ is 73.8±1.2% folded, which is less than the L-amino acid sequence Cd(II)(**TRI**L16C)$_3^-$.[34] indicating that the inclusion of a D-amino acid destabilizes the peptide. A UV/VIS metal binding titration of Cd(II) with the peptide shows that after one equivalent of Cd(II) per 3SCC was added, a Cd(**TRI**L2WL16CL19$_D$L)$_3^-$ complex formed (**Figure 3-3**). This result suggests that even though the peptide is destabilized by the introduction of a D-amino acid, it is still capable of sequestering Cd(II). A comparison of UV titrations with the analogous L-amino acid parent peptide, TRIL16C, yields similar results. First, the LMCT bands of both peptides are centered at 232 nm. It is noted that both of the peptides bind in a 3:1 monomer peptide to metal ratio with strong metal binding as revealed from the sharp curve of Δε versus equivalents of metal plotted for both **TRI**L16C and

TRIL2WL16CL19$_D$L. However, the final extinction coefficients per metal of the peptides are different (**Table 3-4**). There is an increase in Δε from 19,900 $M^{-1}cm^{-1}$ to 22,600 $M^{-1}cm^{-1}$ going from **TRI**L2WL16CL19$_D$L to **TRI**L16C. When these Cd peptides are compared with the published **TRI**L12$_D$LL16C result, which gives a fully Cd(II)S_3^- complex, a trend is observed from **TRI**L2WL16CL19$_D$L (having D-Leu below) 19,900 $M^{-1}cm^{-1}$, **TRI**L16C (mixture of 3- to 4-coordinate 40:60) 22,600 $M^{-1}cm^{-1}$ and **TRI**L12$_D$LL16C (D-Leu above, 100 % Cd(II)S_3^- complex) 23,600 $M^{-1}cm^{-1}$.[3,4] There is a trend in these observations suggesting that the higher extinction coefficients correlate with lower Cd(II) coordination numbers. At one end **TRI**L2WL16CL19$_D$L is assumed to bind Cd(II) in higher coordination numbers probably Cd(II)S_3O^- and Cd(II)$S_3O_2^-$ has the lowest Δε, while **TRI**L12$_D$LL16C published with a fully trigonal site has the highest Δε among the three. The **TRI**L16C, containing mixtures of the two species, has Δε falling in between the two diastereopeptides; however, the UV spectral characteristics do not quantitatively differentiate geometries of Cd(II) complexes.

According to the previous report, pK_as of the Cys residues for Cd(II)-complexes can be determined for the release of two protons from the initial Cd[(peptide)(H-peptide)$_2$]$^+$ complex.[26] Following the LMCT at 235 nm, the UV/VIS spectrum pattern for TRIL2WL16CL19$_D$L is similar to those for other Cd(II)S_3^- and Cd(II)S_3O^- species found in the previous research. This, is a further indication that the desired structure was obtained. The pH curve shown in **Figure 3-5** fits well with a non-linear regression model for the stepwise release of one proton (two protons in total) upon Cd(II) binding to tris-thiolates yielding the first pK_a for one Cys of 6.5 and 6.7 for the other (pK_{a2} = 13.2). This model indicates a pathway that represents the loss of one proton from Cd[(peptide)(H-peptide)$_2$]$^+$ to form Cd[(peptide)$_2$(H-peptide)] and then, the second proton is released at slightly more basic conditions to finally form Cd(II)(**TRI**L2WL16CL19$_D$L)$_3^-$. Moreover, this experiment also suggests that Cd(II) is fully bound to **TRI**L2WL16CL19$_D$L with three thiolate ligands at pH 8.5.

It has been expected that the pK_{a2} values for the formation of different Cd(II) geometries in the diastereopeptides might also reflect the speciation present in solution as observed in our L-analogue peptide system. It was previously determined that by having a space adjacent to the metal layer to encourage formation of a 4-coordinate site [Cd(II)S_3O^-] the thiols deprotonate at lower pH than the ones cysteines in a trigonal site.[3] The pK_{a2} value of 13.2 in **TRI**L2WL16CL19$_D$L actually falls in the range of pK_{a2} values observed for Cd(II)S_3O^- species, which demonstrated that by

changing the chirality of the second coordination sphere using D-Leu the site acidity behaves the same as when the coordination increases due to smaller L-amino acid residues.[35] The pK_{a2} of **TRI**L16C (13.4) with a high proportion of Cd(II)S_3O^- is also consistent with this value, while the pK_{a2} of **TRI**L12$_D$LL16C (fully three coordinate) is 15.1.[4,26]

^{113}Cd NMR spectrum of Cd(**TRI**L2WL16CL19$_D$L)$_3^-$ shows a single resonance at 600 ppm revealing almost 100% of Cd(II)S_3O^- (**Figure 3-6**); however, the PAC Fourier Transform exhibits two angular frequency values of 0.316 rad/ns and 0.159 rad/ns displaying approximately equal amplitudes (A) (**Figure 3-8**). The sharp pattern of these two NQI values suggests well-defined binding sites in the interior of the coiled coil. The parameter at 0.316 rad/ns clearly confirms the presence of a Cd(II)S_3O^- species as predicted. The low value of $\eta = 0.17$ is indicative of a small perturbation of the geometry from the axial symmetry. Intriguingly, the uncommon angular frequency at 0.159 rad/ns is also present. Such a low angular frequency is different from the values that have usually been observed which strongly suggests that the NQIs around the Cd(II) site are distinct from Cd(II)S_3O^- and Cd(II)S_3^- species. This could be due to different types of coordination ligands or different numbers of ligands. Moreover, the closer to zero in angular frequency indicates that the NQIs around the metal site is relatively symmetrical (where the prefect tetrahedral and octahedral geometries ideally show ω_o of 0 rad/ns because of the cancellation of equal charge distribution of extranuclear fields of ligands to the nuclear site).[7] Based on the theoretical calculations by Professor Lars Hemmingsen, this ω_o value could indicate a high symmetry Cd(II) with a higher coordination number, potentially Cd(II)$S_3O_2^-$. The number of thiols bound to Cd(II); however, is still certainly considered as tris-thiolates as reflective of ^{113}Cd NMR chemical shift, but the axial ligands could come from two water molecules (one above and the other below the thiolate plane). This assignment is possible since **TRI**L2WL16CL19$_D$L designs have two potential spaces both at the twelfth and nineteenth positions that are above and below the metal site. This structure possibly allows for the simultaneous access of two water molecules resulting in Cd(II)$S_3O_2^-$. The 0.159 rad/ns, therefore, possibly suggests a trigonal bi-pyramidal Cd(II) structure and the heterogeneity observed from the PAC spectra strongly demonstrates that the open space by 19D-Leu in **TRI**L2WL16CL19$_D$L allows for more waters around the metal site.

Moreover, the X-ray absorption data in **Table 3-5** shows that the best fit for Cd(II)-S and Cd(II)-O distances of Cd(**TRI**L2WL16CL19$_D$L)$_3^-$ (**Figure 3-10**) are 2.48 and 2.37 Å corresponding to a mixed 4- and 5-coordinate system; while the pure Cd(II)S_3O^- formed in

Cd(II)(**TRI**L12AL16C)$_3^-$ shows Cd(II)-S and Cd(II)-O distances of 2.50 and 2.35 Å, and the Cd(II)-S bond distance of the pure Cd(II)S_3^- in the Cd(II)(**TRI**L12$_D$LL16C)$_3^-$ or Cd(II)(**TRI**L16Pen)$_3^-$ is shortened to 2.46 and 2.45 Å, respectively. As the coordination number increases, the average bond distances should increase by 0.07A.[36] Theoretically, a mixture of 4 and 5 coordinate in Cd(**TRI**L2WL16CL19$_D$L)$_3^-$ should have longer average bond lengths than the pure Cd(II)S_3O^-; however, based on these X-ray absorption data it suggests that the Cd(II)$S_3O_2^-$ species that is really much like a Cd(II)S_3^- with two more weakly bound oxygen atoms.

So far, the orientation of D-Leu appears to introduce a space below (referenced to the N-terminus) the Cys-plane allowing for solvent water molecule(s) to coordinate to Cd(II) resulting in a mixture of 4- and 5-coordinate Cd(II) species. Based on these spectroscopic results, it is concluded that we efficiently achieved control of the Cd(II) geometries in the *a* site Cys binding system by using D-Leu in the second coordination sphere. Going from the parent TRIL16C design which was found to sequester Cd(II) non-selectively in a mixture of 40:60 Cd(II)S_3^- : Cd(II)S_3O^- species, the perturbation at the outer sphere one layer (twelfth position) above the metal site by the reorientation of D-Leu achieved an exclusively Cd(II)S_3^- complex. On the other hand, when positioned one layer (the nineteenth position) below the metal site of **TRI**L16C, **TRI**L2WL16CL19$_D$L a mixture of 4- and 5-coordinate Cd(II) complexes results.

As it was successfully shown that a single mutation of D-Leu either above or below the metal site may compel a metal into a specific geometry in our simplified construct, it was next interesting to ask whether D-Leu residues placed simultaneous at both the twelfth and nineteenth positions lead to pure 4-coordination. The hypothesis is that the side chain reorientation would add steric hindrance above the metal site (blocking water) while at the same time remove the steric restrictions below the metal site, resulting in predominantly a 4-coordinate complex Cd(II) species. The knowledge gained from this work will not only provide a better understanding on the benefits of using stereochemistry in metalloprotein design, but also extend the insight into investigating how well a three-stranded coiled coil design can handle multiple D-amino acid substitutions in its sequence.

The longer length scaffold of **GRAND-CoilSer** was used to avoid a drastic decrease in folding free energy due to the disruption of hydrophobic interactions with the double D-Leu substitutions within the helical core. Here D-Leu residues are introduced at the twelfth and the nineteenth positions of the sequence where the coordination site is at the sixteenth position

Table 3-5: Spectroscopic data for Cd(II) binding to the different **TRI** peptides

Peptide	Cadmium species(%)			Apparent	UV/VIS		EXAFS
	$Cd(II)S_3^-$	$Cd(II)S_3O^-$	$CdS_3O_2^-$	pK_{a2}	λ_{max} (nm)	$\Delta\varepsilon$ ($M^{-1}cm^{-1}$)	Cd(II)-S bond distance (Å)
TRIL16C	40	60	-	13.40[a]	232[b]	22,600[b]	2.49[b]
TRIL12$_D$LL16C	100	-	-	15.10[c]	232[c]	23,600[c]	2.46
TRIL2WL16CL19$_D$L	-	50	50	13.20	232	19,900	2.48 (Cd(II)-O = 2.37 Å)
GRAND-CSL12$_D$LL16CL19$_D$L	-	70	30	13.74	232	19,700	2.48 (Cd(II)-O = 2.28 Å)
TRIL12AL16C	-	100	-	12.20[d]	231[e]	21,200[e]	2.50 (Cd(II)-O = 2.35 Å)

[a]Data from reference 26; [b] Data from reference 3; [c] Data from reference 4; [d] Data from reference 37; and [e] Data from reference 38.

generating **GRAND-CS**L12$_D$LL16CL19$_D$L. The physical properties of the peptide were experimentally examined in the same manner as done with **TRI**L2WL16CL19$_D$L. According to the CD band at 222 nm, the apo-peptide is ~70% folded. The incorporation of two D-Leu residues has a strong effect on the folding of the coiled coil structure when one realizes that the same level of folding is achieved for **GRAND-CS**L12$_D$LL16CL19$_D$L containing double D-Leu mutation as was observed for **TRI**L2WL16CL19$_D$L that had a single mutation. This suggests that the destabilization associated with the D-Leu substitution roughly cancels the added stability of an extra heptad (generally thought to be ~ 1.6 kcal/mol per monomer).[39] This observation suggests that substituting two D-Leu residues into the **TRI**-family peptides would yield variants with an insufficient level of folding for the apo protein. Moreover, the percent folding of the constructs did not improve when Cd(II) was added into the system. This could be because Cd(II) does not help with the folding of the peptide or because the peptide is more dynamic allowing the large excess of chloride access to extract more easily the Cd(II) under the high concentrations of denaturant.

To further determine the effect of the double D-Leu mutation on the stability of **GRAND-CS**L12$_D$LL16CL19$_D$L, the GuaHCl denaturation titration studies were performed to compare the results with other single D-Leu containing **GRAND-CoilSer** variants (sequences shown in **Table 3-1**). The denaturation curves are shown in **Figure 3-21** and the statistic of the fitting is shown in **Table 3-6**. Obviously, the **GRAND-CS**L12$_D$LL16CL19$_D$L is not highly stable with the $\Delta G_{folding}$ only 0.77 kcal/mole and a C_m= 0.21 M for the apo-peptide and 0.95 kcal/mole with C_m= 0.27 M for the Cd(II)-peptide. On the other hand, a single D-Leu substitution in **GRAND-CS**L12$_D$LL16C (This peptide, which has the same mutation as **TRI**L12$_D$LL16C, is prepared for crystallization.) has $\Delta G_{folding}$ of 3.02 kcal/mole with C_m= 1.86 M (in the absence of Cd(II)) and 3.07 kcal/mole with C_m= 1.93 M (in the presence of Cd(II)) The **GRAND-CS**L16CL19$_D$L (analogous to the **TRI**L2WL16CL19$_D$L) shows less stability compared to **GRAND-CS**L12$_D$LL16C. The apo-peptide has $\Delta G_{folding}$ of 2.12 kcal/mole with C_m= 1.46 M and the Cd(II)-peptide of 2.6 kcal/mole with C_m= 1.83 M. From these observations, the replacement of D-Leu in the twelfth position leads to the most stable peptide among the three in the series. This is because the substitution is close to the N-termini where the hydrophobes are more tightly-packed. Consequently, D-Leu residues in the twelfth position could tuck the whole side chain into the center of the core. This conclusion is confirmed crystallographically and will be described in detail below. In summary, it appears that in **TRI**, D-Leu substitution towards the N-terminus leads to less destabilization than when

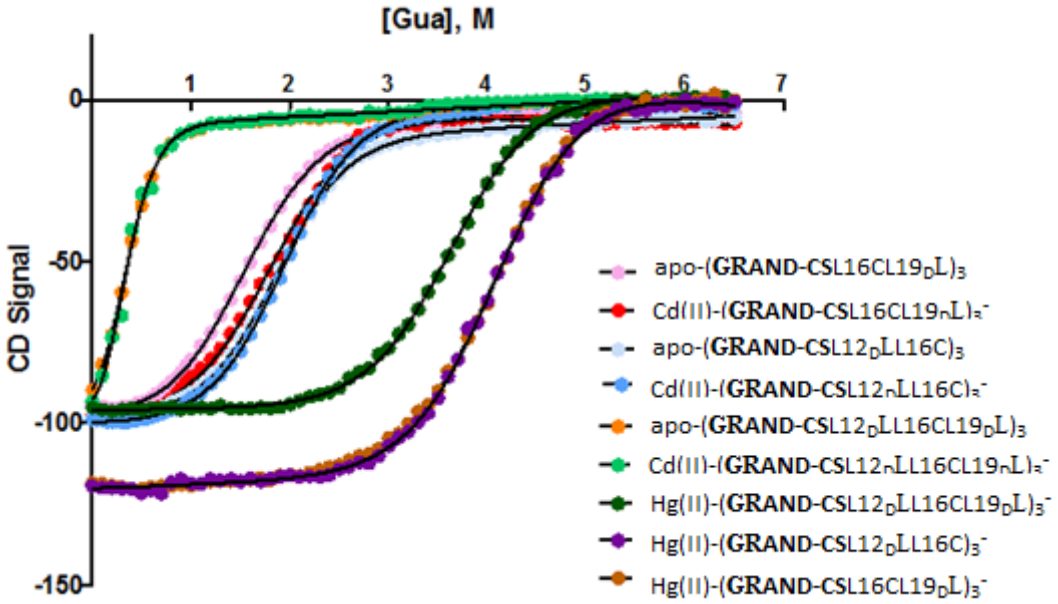

Figure 3-21. Fits of GuaHCl denaturation curves for D-Leu containing peptides in the absence of metal and presence of metal (Cd(II) or Hg(II)) at pH 8.5. Fit results are reported in the table below. The fitting results are shown in **Table 3-6**.

Table 3-6: Results of fits to GuaHCl denaturation curves of the diastereopeptides.

Peptide	**C_m [M]**	**R^2**	**ΔG(kcal mol^{-1})**
apo-(**GRAND-CS**L16CL19$_D$L)$_3$	1.463	0.9999	2.115
Cd(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$	1.829	0.9991	2.650
Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$	4.134	0.9996	5.544
apo-(**GRAND-CS**L12$_D$LL16C)$_3$	1.856	0.9998	3.023
Cd(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$	1.938	0.9998	3.066
Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$	4.110	0.9995	5.561
apo-(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3$	0.2731	0.9981	0.948
Cd(II)(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3^-$	0.2191	0.9942	0.769
Hg(II)(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3^-$	3.641	0.9996	5.068

incorporated near the C-terminus. The destabilization of these two peptides is not as great as seen with **GRAND-CS**L12$_D$LL16CL19$_D$L where the amount of the folding free energy is dramatically diminished both in the absence and presence of Cd(II). This strongly emphasizes the negative influence of having double D-Leu substitution in helical assemblies.

Even though the coiled coil construct of **GRAND-CS**L12$_D$LL16CL19$_D$L is less stably folded, the peptide could form a Cd(II)-(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3^-$ complex when a precise 1:1 stoichiometry of Cd(II) was added to the peptide solution. As demonstrated by UV/VIS titrations, the growth of bands having λ_{max} of 232 nm was observed at pH 8.5. As shown in **Figure 3-4**, the peaks increase until 1 equivalent of Cd(II) per peptide trimer and then they slowly level off at a final extinction coefficient of 19,700 M^{-1} cm^{-1}. Thus, it is confirmed that the three helices of **GRAND-CS**L12$_D$LL16CL19$_D$L bind to one Cd(II) as a coiled coil with the best description for the complexes to be [Cd(II)(**GRAND-CS**L12$_D$LL16CL19$_D$L)$_3$]$^-$ (**Figure 3-4**). The final extinction coefficient is close to what was reported to **TRI**L2WL16CL19$_D$L (**Table 3-5**). From the UV-vis data, it is highly likely that **GRAND-CS**L12$_D$LL16CL19$_D$L peptide shares a similar type of metal coordination geometry (a mixture of 4- and 5-coordinate Cd(II)) as **TRI**L2WL16CL19$_D$L.

^{113}Cd NMR experiments were performed to assign the Cd(II) binding environment. Cd(II) bound to **GRAND-CS**L12$_D$LL16CL19$_D$L results in a chemical shift of 610 ppm **(Figure 3-7)** which falls in an acceptable range of values that are reported for Cd(II)(SR)$_3^-$ complexes.[40] If constrained to a published correlation model comparing chemical shifts for Cd(II)S_3^- and Cd(II)$S_3O_2^-$ species for the *a* site peptide system from Iranzo et al.[35], this observed shift would be consistent with predominantly Cd(II)S_3O^- present (~75%); however, the possible existence of Cd(II)$S_3O_2^-$ in **GRAND-CS**L16CL19$_D$L may render this assignment incorrect. However, the upfield shift does indicate the presence of water ligand(s) on the Cd(II) center in **GRAND-CS**L12$_D$LL16CL19$_D$L. Thus, the angular information was collected from ^{111m}Cd PAC spectroscopy to assign the Cd(II) geometries. The ^{111m}Cd PAC spectra of **GRAND-CS**L12$_D$LL16CL19$_D$L in **Figure 3-10** show a sharp line with two clear NQIs for a minor species (30%) at ω_o = 0.17 rad/ns (η=0.39) indicating the 3Cys trigonal bi-pyramidal Cd(II)$S_3O_2^-$ structure where O is from a solvent water and a major species (70%) at ω_o = 0.33 rad/ns (η=0.26) suggesting a pseudo-tetrahedral Cd(II)S_3O^- . Given the fact that **GRAND-CS**L12$_D$LL16CL19$_D$L contains a heterogeneous metal center with two species (Cd(II)$S_3O_2^-$ and Cd(II)S_3O^-) based on the PAC result. Because a single resonance was obtained from ^{113}Cd NMR, this clearly illustrates these two

coordination geometries have faster dynamic exchange than the ^{113}Cd NMR timescale. The ^{111m}Cd PAC integrations demonstrate the coexistence of ca. 70% Cd(II)S_3O^- and 30% Cd(II)$S_3O_2^-$. It appears that when adding steric interference into the layer at the twelfth above the L16CL19$_D$L design, the **GRAND-CS**L12$_D$LL16CL19$_D$L shifts the coordinates of Cd(II) from 50% 5-coordination formation by less than 20% from that observed for **TRI**L2WL16CL19$_D$L. A rational explanation to this observation is while the reorientation of 19D-Leu definitely removes the steric hindrance below the 16Cys site encouraging the water access from the bottom, the 12D-Leu placed one layer on top of the sixteenth position can only control the environment toward the metal site to a limited extent. This could be due to the drastic effect of having two D-Leu residues at the same time along the sequence that might greatly disrupt the hydrophobic stacking in the core causing the helices to fray apart and have a hard time folding into a desired structure as indicated by the destability of the peptide. I conclude that the steric packing made by 12D-Leu in the **GRAND-CS**L12$_D$LL16CL19$_D$L construct is not as efficient in excluding a water ligand as found in TRIL12$_D$LL16C. While not completely blocking the access of the fifth water molecule, these data clearly support the conclusion that the design succeeded to some degree.

In summary, I conclude that the replacement of D-Leu in the second coordination sphere around the metal binding site displays different effects on Cd(II) structures. The orientation of D-Leu when positioned at the twelfth position (close to the N-terminus) above the 16Cys site follows the hypothesis that this residue can block the access of water giving pure Cd(II)S_3^-. In contrast, the opposite effect of steric removal was found when D-Leu is placed below the metal site resulting in an equal mixture of Cd(II)S_3O^- and a new species Cd(II)$S_3O_2^-$. Moreover, the combination of 12D-Leu and 19D-Leu placed in a peptide sequence also produces a mixture in which the major product is Cd(II)S_3O^-. Even though all of these physical properties are confirmed, there has been no structural details to support how the D-Leu substitution affects the secondary structure of the entire coiled coil scaffold and how the reconfiguration of the D-chirality impacts the sterics in the outer sphere around a metal binding site of the engineered peptides. To address these questions, it is very important to perform crystallization studies to advance the complete understanding of my diastereopeptide designs. In the following section, the structural analysis of apo-(**GRAND-CS**L12$_D$LL16C)$_3$ Hg(II)(**GRAND-CS**L12$_D$LL16C)$_3^-$ (preliminary electron density map result), apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ are described in combination with the previous apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$

structures reported in Chapter 2, to evaluate the structural implications on how a change in the chirality can drastically alter the coordination environment of the metal within a designed protein.

Crystal structure characterization

The Cys substitution in **CoilSer** variants appears to crystalize into a parallel 3SCC structure.[18,32,41,42] However, the first attempt to produce crystals of **CS**L12$_D$LL16C, containing two mutations of an unnatural D-Leu and Cys, was unsuccessful. It was previously found that **CoilSer** can, to some extent, handle the destabilization from the mutation of the Cys residue placed in one of the hydrophobic layers; however, with a combination of D-Leu the instability of the **CS**L12$_D$LL16C became worse, subsequently resulting in poor peptide folding. Obviously, D-Leu causes more drastic destabilization to the peptide. This statement was supported by the result of GuaHCl denaturation titration curve shown in **Figure 3-22**. The denaturation titration curve of **CS**L12$_D$LL16C does not follow a two state model changing from the completely folded 3SCC to the unfolded monomer states. This observation demonstrates low propensity for the peptide to fold into a native state. Similarly, the peptide does not show a reasonable percent folding, even in the absence of denaturant where only 30% folding was observed at 10 μM peptide concentration. This value is about 50% less than the percent folding determined for **TRI**L12$_D$LL16C.[4] Therefore, in order to increase the stability of the construct and enhance the folding, the variant is prepared using the elongating **GRAND-CoilSer**. As predicted, the resulting **GRAND-CS**L12$_D$LL16C peptide (sequence shown in **Table 3-1**) has significantly greater stability. The initial folding of this new design increased to 78± 0.5%. The denaturation titration curve of the apo-peptide shows a significant improved stability with $\Delta G_{folding}$ of 3.02 kcal/mole and the midpoint (C_m) of 1.85 M of [GuaHCl]. This evidence suggests that the addition of a D-amino acid combined with a Cys residue in **GRAND-CS**L12$_D$LL16C is well-tolerated. Moreover, when Cd(II) was added to the peptide, ^{113}Cd NMR showed the exact chemical shift as was observed for **TRI**L12$_D$LL16C (697 ppm) which clearly demonstrates that the metal environment between these two metal binding sites are identical**(Figure 3-23)**. Thus, **GRAND-CS**L12$_D$LL16C could reasonably serve as a crystallographic analog for **TRI**L12$_D$LL16C.

To illustrate the effect of alternate chirality on the internal hydrophobic residues in the second coordination sphere, the apo-(**GRAND-CS**L12$_D$LL16C)$_3$ is overlaid onto the apo-(**CS**L16C)$_3$ from the previous chapter in **Figure 3-24**. This figure compares a parent peptide that

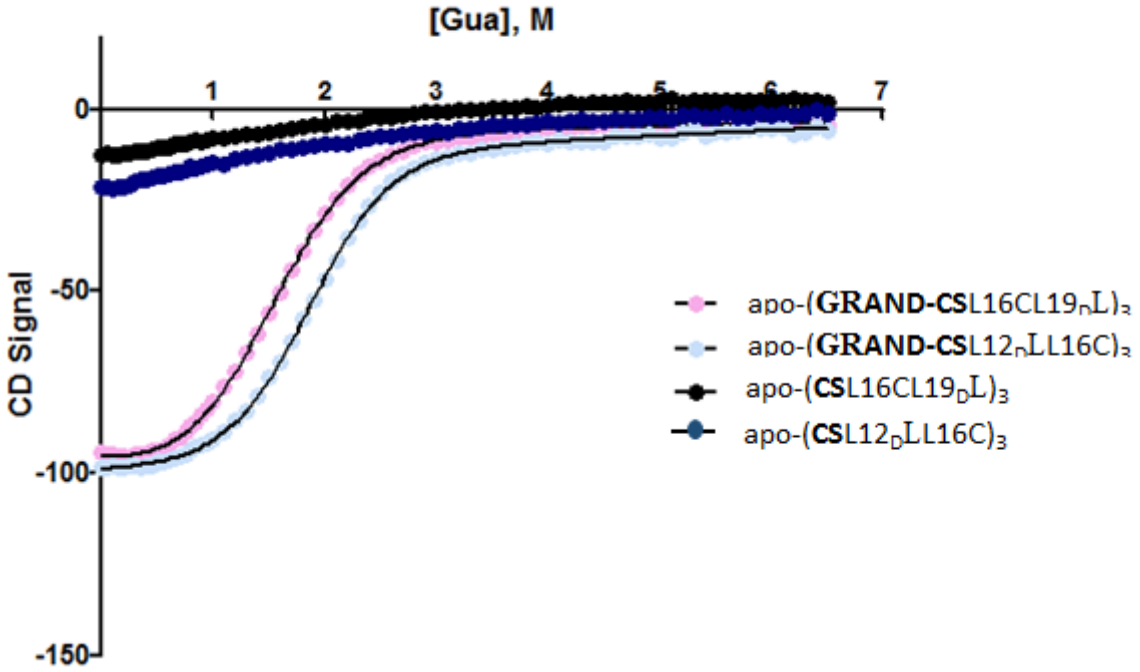

Figure 3-22. GuaHCl denaturation curves showing the effect of D-Leu on the GRAND-CoilSer and CoilSer peptides. The titrations were performed at pH 8.5, ambient temperature.

contains solely L-amino acids in the sequence to one with a single D-Leu substitution. Both of the peptides fold into parallel three stranded coiled coils as predicted. Although they are different in length by one heptad, the α-helical backbones of the two structures are extremely similar except that the last heptad in the chain A of apo-(**CS**L16C)$_3$ is slightly frayed because both 24GluA and 28HisA are constrained to bind to two Zn(II) ions. In contrast, in the other two chains only His28 is responsible for complexing a single Zn(II) while 24Glu is freely facing the helical interface. Intriguingly, there are no kinks in the helical backbones observed in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ suggesting that the incorporation of a D-Leu does not disturb the secondary structure of the coiled coil. As shown, the Cβ carbons of the Cys residues in both structures point toward the N-termini of the helices and the Sγ atoms adopt two conformations. The major Cys rotamers have the thiol positioned toward the metal binding core of the peptide as if prepared for a *preorganized* metal binding **(Figure 3-24**,**b)**, corresponding to the *a* site Cys orientation reported in apo-(**CS**L9C)$_3$.[18] These thiol ligands of apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and apo-(**CS**L16C)$_3$ not only point in the same direction, but exhibit a similar range of torsion angles. The inner thiol rotamers in the D-Leu containing construct have χ_1 values of -66.77^o resembling the -69.20^o in apo-(**CS**L16C)$_3$. Furthermore, the Sγ-Sγ distances are comparable between both structures; 3.45 Å in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and 3.29 Å(averaged) in apo-(**CS**L16C)$_3$ as expected for a good chelating sulfur environment. The minor Cys orientations of both structures (**Figure 3-24**,**c)** point their thiol groups to the outer interface, subsequently causing a long Sγ-Sγ separation between minor Cys conformers, which are not suitable for metal binding. According to this first structural analysis in the layer at the sixteenth position, it appears that the apo-structures of (**GRAND-CS**L12$_D$LL16C)$_3$ and (**CS**L16C)$_3$ present a relatively similar metal binding ligand environment.

I next attempted to explain how the change in chirality above the metal site can compel Cd(II) into a specific 3-coordinate Cd(II)S_3^- structure whereas L16C produces a mixture of Cd species. The packing of 12D-Leu layer in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and 12L-Leu in apo-(**CS**L16C)$_3$ are compared in **Figure 3-25**. It is obvious that D-Leu residues are tightly packed as opposed to L-Leu, causing the steric hindrance above the layer at the sixteenth. This perturbation occurs because the D-configuration is inverted leading to the reorientation of Cβ atoms from directing toward the N-termini (in L-Leu) to the C-termini (**Figure 3-26**). This Cβ deviation

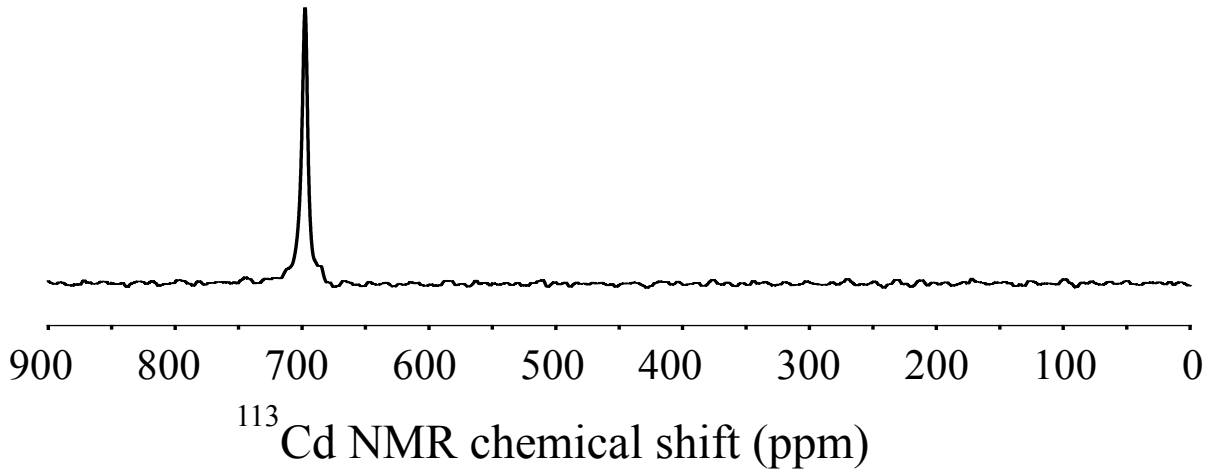

Figure 3-23. ^{113}Cd NMR of Cd(GRAND-CSL12$_D$LL16C)$_3^-$. The spectrum was recorded at pH 8.5.

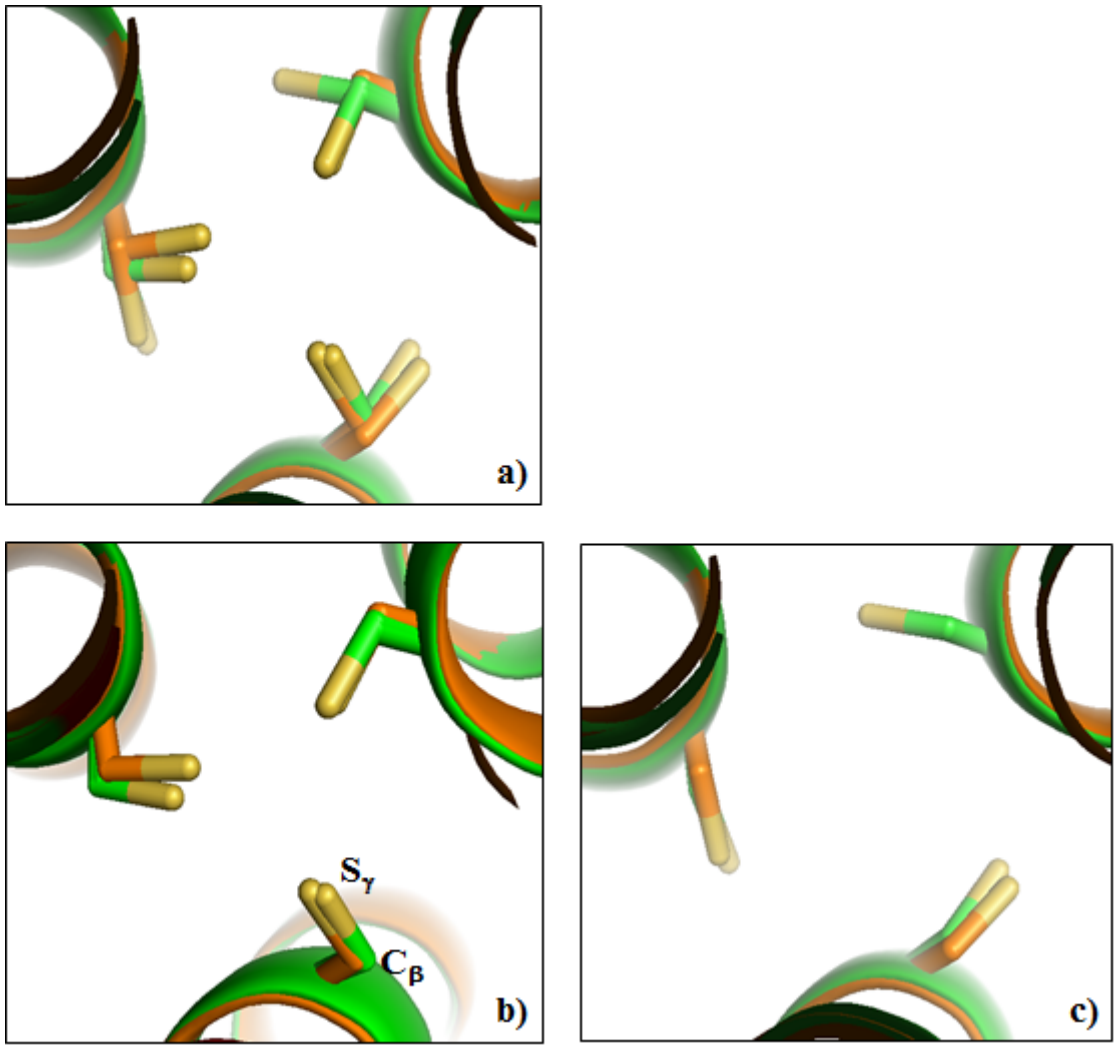

Figure 3-24. An overlay of apo-(GRAND-CSL12$_D$LL16C)$_3$ and apo-(CSL16C)$_3$ structures showing the orientations of Cys side chains. a) Top-down view from the N-termini represents the major and minor conformers of Cys residues of the two structures. b) Top-down view of the major conformers. c) Top-down view of the minor conformers. (note: In apo-(**CS**L16C)$_3$ only chain A and B contain minor conformation. Main atoms are shown in green for apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and in orange for apo-(**CS**L16C)$_3$. The Cys side chains are shown as stick with the thiol groups labeled in yellow.

twists the positions of δ-methyl groups ($C\delta_1$, $C\delta_2$) toward the center of the coiled coil (**Figure 3-27**). In the apo-(**CS**L16C)$_3$ structure, only one of the two δ-methyl atoms of each L-Leu residue is pointed toward the center, while the other points to the helical interface, thus opening up more space above the metal binding site and potentially making it less-packed compared to 12D-Leu (**Figure 3-25**). This D-Leu effect shortens the separation between 12D-Leu and 16Cys layer to be (2.32 Å), while it is 5.08 Å in apo-(**CS**L16C)$_3$ (**Figure 3-18**). Additional support for the steric interference model comes from how close the inner δ-methyl atoms are located above the metal site with 12D-Leu. The shorter $C\delta_2$ - $C\delta_2$ distance of 3.80 Å implies that the $C\delta_2$ groups of D-Leu are close-packed, while this separation becomes much longer (average of 6.60 Å) using 12L-Leu. With all these apo-peptide crystallographic observations, it is obvious that the space above the layer at the sixteenth position made by L-Leu is bigger and less well packed compared to D-Leu. The differential orientations of leucine layers in the second coordination sphere, therefore, could represent an important effect of the amino acid side chains chirality on metal structures and binding mode preferences in the metallated-forms. To analyze this hypothesis further, it was useful to build Cd(II) models based on our known crystallographic metallated structures.

Unfortunately, Cd(II)-peptide crystals have not yet been obtained, however, $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ from the previous chapter may again be used to represent a metallated form of a trigonal planar Cd(II) structure as described in Chapter 2 *Section I*. Structural overlays between $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and apo-(**CS**L16C)$_3$ are shown in **Figure 3-28**. In brief, only the dominant thiol conformers from the apo-(**CS**L16C)$_3$ could form a *preorganized* metal binding site which is able to internalize the metal in the protein's hydrophobic core. This is because these ligands present a good metal chelating environment by pointing toward the helical core with an averaged $S\gamma$-$S\gamma$ separation of 3.29 Å. In addition, the averaged $C\beta$-$S\gamma$ torsion angle of -69.20^o is close to the value of -65.20^o reported for common unbound Cys rotamers.[28] However, one observes that when aligned with the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure, the metal binding induces significant rotation of the interior Cys conformations by moving the thiols downward and to the side. This shift orients the cysteine sulfur atoms more toward the helical interface leading to a longer $S\gamma$-$S\gamma$ separation of 4.16 Å. This movement is required to provide space to accommodate a large Hg(II) atom,

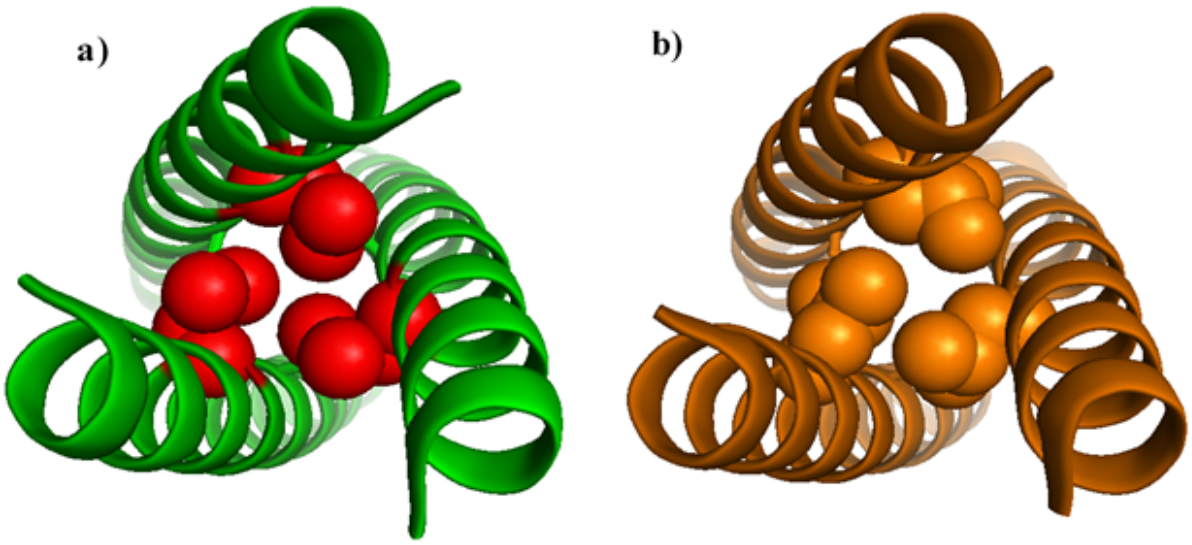

Figure 3-25. Packing of hydrophobic layers in the twelfth position above the 16Cys site of apo-(GRAND-CSL12$_D$LL16C)$_3$ and apo-(CSL16C)$_3$ shown as spheres. From top down view of the N-termini, representing a) the D-Leu packing in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and b) the L-Leu packing in apo-(**CS**L16C)$_3$.

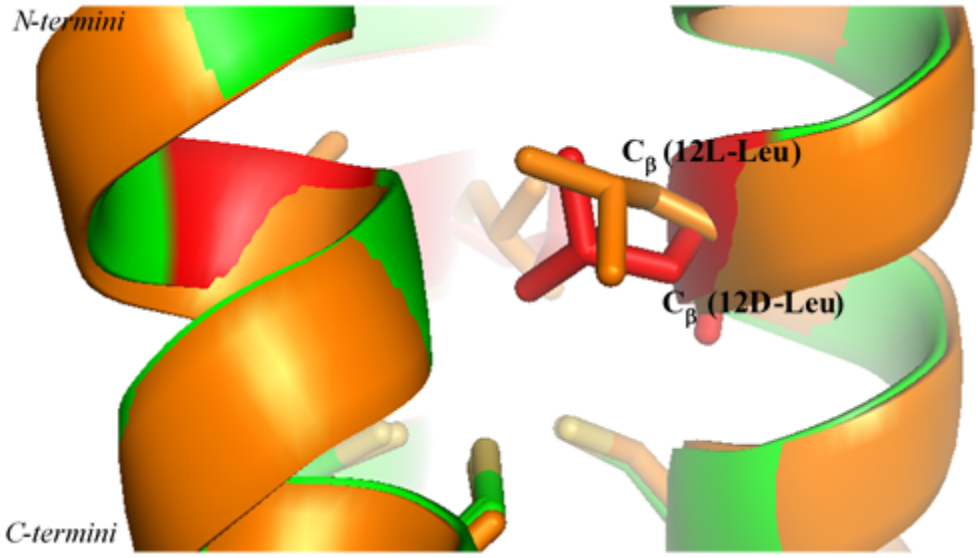

Figure 3-26. Side view cartoon diagrams of an overlay of apo-(GRAND-CSL12$_D$LL16C)$_3$ and apo-(CSL16C)$_3$ structures showing the difference in C_β carbon positions of the twelfth layer. In apo-(**GRAND-CS**L12$_D$LL16C)$_3$, D-Leu residues are shown as red sticks and main atoms labeled in green. In apo-(**CS**L16C)$_3$, both L-Leu and main atoms are orange-colored. Thiols are yellow.

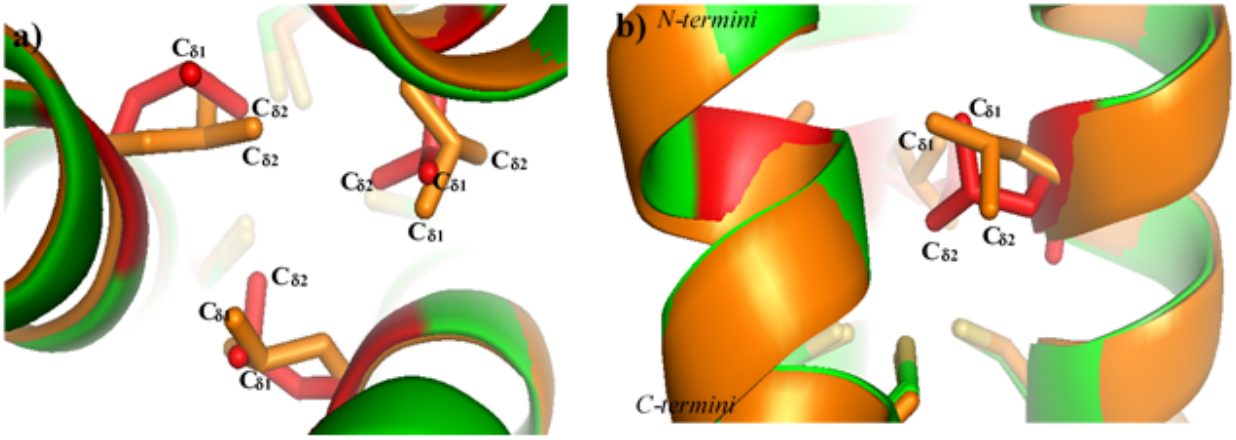

Figure 3-27. An overlay of apo-(GRAND-CSL12$_D$LL16C)$_3$ and apo-(CSL16C)$_3$ structures representing the Cδ$_1$, Cδ$_2$ atoms of leucine residue orientations in the twelfth layer. a) Top down view from the N-termini and b) Side-on view. In apo-(**GRAND-CS**L12$_D$LL16C)$_3$, D-Leu residues are shown as red sticks and main atoms labeled in green. In apo-(**CS**L16C)$_3$, both L-Leu and main atoms are orange-colored. Thiols are yellow.

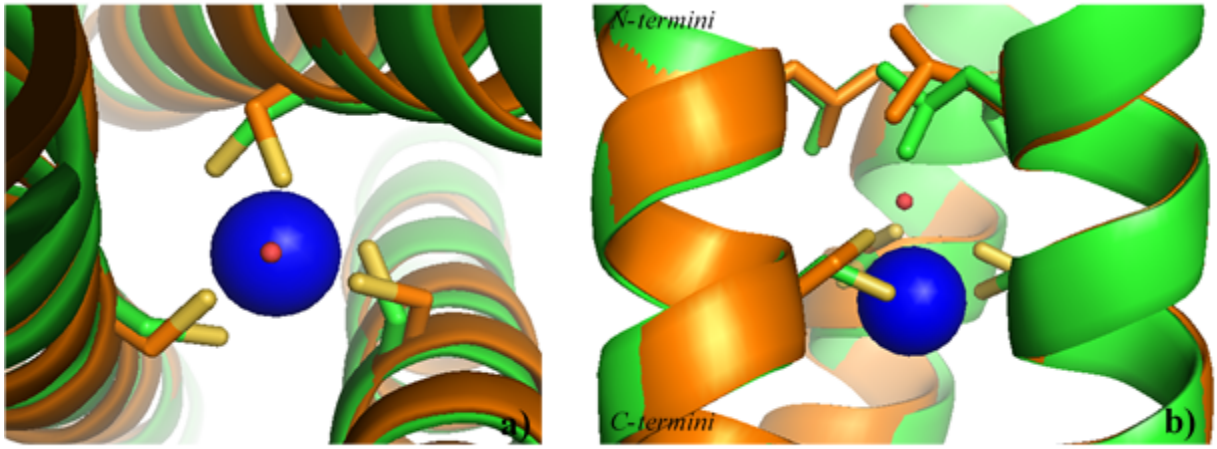

Figure 3-28. Ribbon diagrams demonstrating an overlay of the sixteenth layer between $Hg(II)_SZn(II)_N$(GRAND-CSL16CL30H)$_3^+$ and apo-(CSL16C)$_3$. a) Top down view from the N-termini and b) Side on view. Main atoms of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ are colored in green and apo-(**CS**L16C)$_3$ in orange. Cysteine side chains are shown as sticks. Hg(II) is present as a blue sphere where the observed water is shown in red.

with ionic radius of 1.16 Å. The χ_1 dihedral angles of these major thiols change from -69.20° in the apo-structure to -151.88° in the metal bound form. The S-Hg(II)-S angle is 118.36° and the Hg(II)-S distance (2.43 Å) is in excellent agreement with the 2.44 Å value determined from the X-ray absorption analysis given in **Table 3-5**. As mentioned in Chapter 2, there is a water molecule directly above the Hg(II) at 2.79 Å. The observed water is located within the hydrophobic cavity between the twelfth and sixteenth positions and does not affect the 12L-Leu layer. This conclusion is confirmed as no significant change in hydrophobic packing occurs between the apo- and metallated structures. The knowledge learned from the Hg(II)-binding behavior can be applied to explain the Cd(II) binding situation especially in a trigonal $Cd(II)S_3^-$ geometry because the Hg(II)-S and Cd(II)-S bond distances based on the X-ray absorption data are closely similar; 2.44 and 2.46 Å, respectively. Again, the coordinated Hg(II) in the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ crystal structure can be used to model Cd(II) complexes in this chapter: $Cd(II)S_3^-$ and $Cd(II)S_3O^-$ in L16C and $L12_DLL16C$ peptides. Before determining the impact of the 12D-Leu to yield the dehydrated Cd(II) structure, it is worth looking at the parent L16C peptides, containing 12L-Leu, in apo- and metallated forms to understand how the normal L-analog can allow for partial water access when a metal is bound.

Unlike Hg(II), Cd(II) intrinsically prefers to bind as a 4-coordinate complex in a thiolate rich environment. This preference is confirmed by the existence of a mixture of 60% $Cd(II)S_3O^-$ and 40% $Cd(II)S_3^-$ species in L16C and fully formed $Cd(II)S_3O^-$ in L12AL16C peptides when a hole above the metal site allows solvent access. Described in *Section III* Chapter 2, the evidence of four waters occupied in the space above the metal layer in the Hg(II)(**GRAND-CS**L12A16C)$_3^-$ compared to the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ where only one water is only present above the metal site reflects the potentials of Ala to remove the steric from the bulky Leu residues at the twelfth position. These observations allow for making potential models for $Cd(II)S_3O^-$ in L12AL16C and the mixtures [$Cd(II)S_3O^-$, $Cd(II)S_3^-$] in L16C peptides (**Figure 2-35** and **Figure 2-36**). Here in this chapter, the coordinates of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ in combination with the apo-(**GRAND-CS**$L12_DLL16C)_3$ are used to generate a possible model of $Cd(II)S_3^-$ obtained in the $L12_DLL16C$ peptides.

To model $Cd(II)S_3^-$ and visualize the 12D-Leu perturbation in the second coordination sphere on the bound metal plane, the apo-(**GRAND-CS**$L12_DLL16C)_3$ structure was overlaid onto the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ structure. The helical backbones fit well together

along the length of the chains, except at the region of the 1Glu residues that point differently through space due to the frayed ends of both structures. Unsurprisingly, the packing of 12D-Leu in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ differs from the 12L-Leu in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ as found with apo-(**CS**L16C)$_3$. First, both δ-methyl groups of each D-Leu residue, point toward the core of the helices, whereas in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ only one Cδ atom (Cδ$_1$) of each 12L-Leu is in the core while the other is facing out toward the helical interface as shown in **Figure 3-29**. Moreover, the analysis of the aligned structures demonstrates that the Cβ atoms of D-Leu are drastically different in position from the L-chirality, causing the Cδ$_2$ atoms in the D-Leu layer to tuck toward the center and move closer to the observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. This causes the 12D-Leu layer(Cδ$_2$ plane) get really closer to the water with only a distance of 1.33 Å, while the interior Cδ plane (Cδ$_1$) of Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is at a distance of 3.80 Å from the water (**Figure 3-30**). Obviously, the D-Leu hydrophobic groups are located too close to the water to allow the solvent to occupy this position. Therefore, steric encumbrance appears to be the basis for water exclusion in L12$_D$LL16C.

To model the Cd(II)S$_3^-$ structure in the **GRAND-CS**L12$_D$LL16C peptide, the Hg(II) coordinate was used to locate the Cd(II) center when the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and apo-(**GRAND-CS**L12$_D$LL16C)$_3$ structures were well-aligned. The thiols in apo-protein need to rotate approximately 80 degrees from pointing into the core of the structure toward the helical interface in order to achieve the desired Cd(II)-S distance of 2.46 Å reported from EXAFS **(Table 3-5)**. This allows the Cys plane to increase the Sγ-Sγ distance to 4.25 Å and be able to accept the metal into the trigonal planar plane. Shown in **Figure 3-31**, the final model displays a trigonal Cd(II)S$_3^-$ with an S-metal-S angle of 120.07°. The model demonstrates the shift of sulfur plane toward the binding would cause the layers (12D-Leu versus 16Cys) above the metal site separated by 4.34 Å, while in the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ the related distance determined from 12L-Leu is 6.07 Å. This strongly emphasizes that the D-Leu layer above the metal site is tightly-packed suggesting that the water should no longer exist within this tiny space. The effect of 12D-Leu on water exclusion is also supported by a preliminary electron

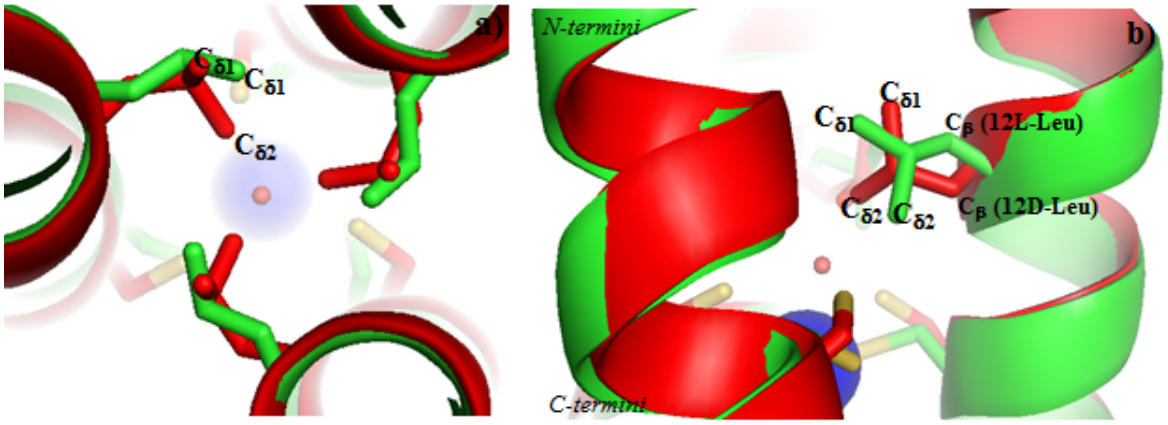

Figure 3-29. An overlay of apo-(GRAND-CSL12$_D$LL16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structures representing the Cδ_1, Cδ_2 atoms of leucine residue orientations in the twelfth position. a) Top down view from the N-termini and b) Side view. Main atoms of apo-(**GRAND-CS**L12$_D$LL16C)$_3$ are colored in red and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ in green. Hg(II) atom and the observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are present in blue and red spheres respectively.

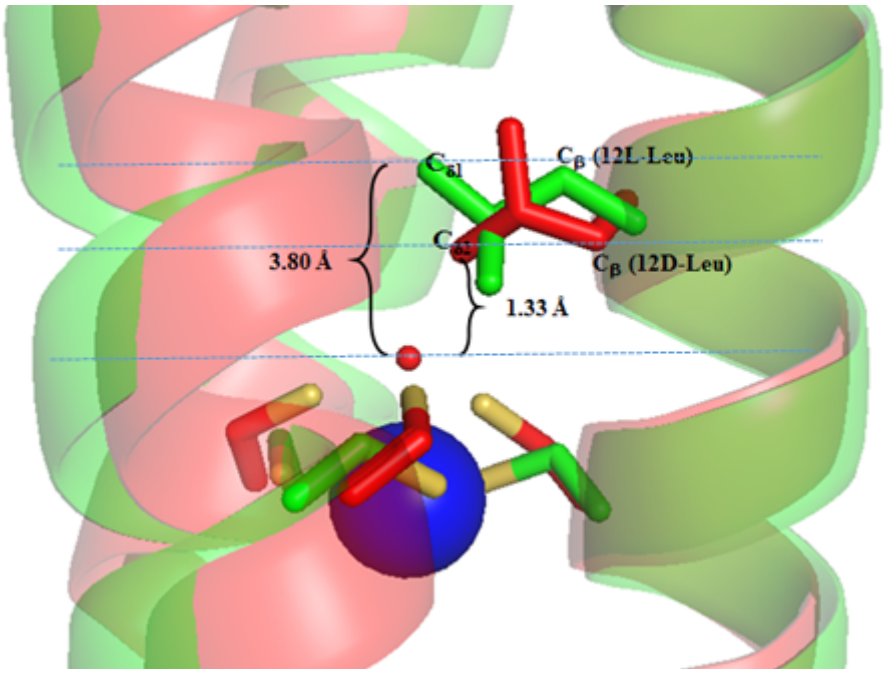

Figure 3-30. **PyMOL visualization representing the distance determined from the plane generated by the core δ-methyl atom of the 12Leu residue to the water in Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ compared to the related distance determined from the plane generated by the core δ-methyl atom of the 12D-Leu residue in apo-(GRAND-CSL12DLL16C)$_3$ to the water.** Main atoms of apo-(**GRAND-CS**L12$_D$LL16C)$_3$ are colored in red and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ in green. Hg(II) atom and the observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are present in blue and red spheres respectively.

distance determined from 12L-Leu is 6.07 Å. This strongly emphasizes that the D-Leu layer above the metal site is tightly-packed suggesting that the water should no longer exist within this tiny space. The effect of 12D-Leu on water exclusion is also supported by a preliminary electron density map (F_o-F_c) contoured at 3σ of a $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$structure in **Figure 3-32**. After several rounds of refinement, there is no evidence of any positive difference density with r.m.s.d. levels greater than 1 electron/A^3 that could represent waters between the stacking of 12D-Leu and 16Cys layers. Therefore, from the combination of all these crystallographic observations it can be concluded that when the metal is physically bound into the Cys plane one hydrophobic layer below D-Leu, the steric hindrance of D-Leu blocks the solvent access to the site.

Having successfully demonstrated the ability exploit hydrophobic interference induced with the D-chirality of leucine when positioned above the metal site, I next examined the impact of removing the bulky packing in the layer below the metal plane at the nineteenth position. The combination of ^{113}Cd NMR, ^{111m}Cd PAC and EXAFS spectroscopies confirmed that L16CL19$_D$L peptide binds Cd(II) with higher coordination numbers appearing equally as $Cd(II)S_3O^-$ and $Cd(II)S_3O_2^-$ species. The apo-peptide structural analysis was initiated to investigate the hydrophobic packing comparison around the metal site. The overlay of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ with apo-(**CS**L16C)$_3$ illustrates that the α-helical backbones are well-aligned with no kinks observed at the D-Leu residual region. In **Figure 3-33**, the Cys residues of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ display a single rotamer pointing toward the core of the structure resembling the major conformer of apo-(**CS**L16C)$_3$ and confirmed by their close values in side chain torsion angles; -61.17^o versus -69.20^o. The thiol arrangement appears to capable of forming a *preorganized* metal binding with a Sγ-Sγ separation of 3.31 Å, which is 0.02 Å longer than in apo-(**CS**L16C)$_3$. The similarity in Cys layers primarily reveals that D-Leu placed at the nineteenth layer does not affect the first coordination sphere ligands when the binding sites are not metal-bound.

In general, both the 19D-Leu and the 19L-Leu side chains appear to direct the δ-methyl groups out toward the helical interface; however, the reorientation of the Cβ atoms with the 19D-Leu in the apo-(**GRAND-CS**L16CL19$_D$L)$_3$ causes both of the δ-methylene groups ($C_{\delta 1}$ and $C_{\delta 2}$) to move to the outer face **(Figure 3-34)**, hence leading to an even greater separation of

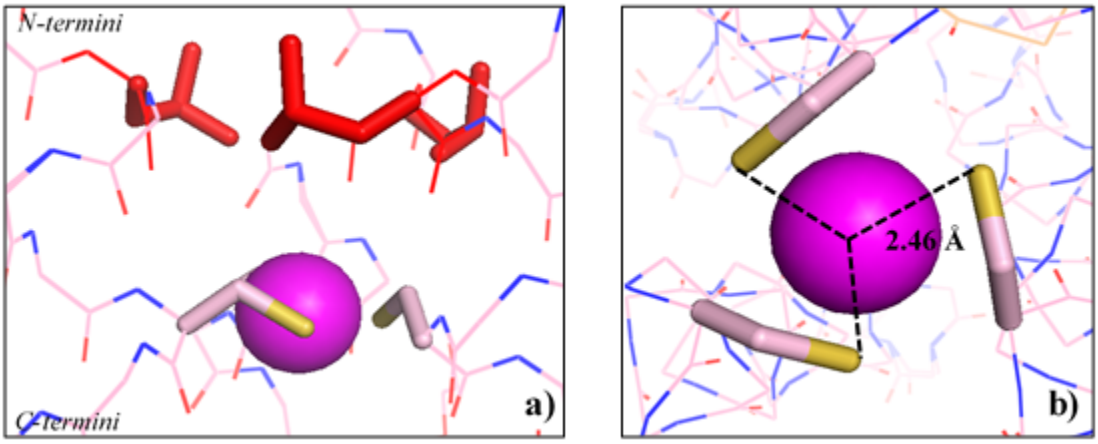

Figure 3-31. Model of $Cd(II)S_3^-$ using apo-(GRAND-CSL12$_D$LL16C)$_3$ crystal structure. a) Side view of the binding site. b) Top down view from the N-termini of the model representing Cd(II)-S distance of 2.46 Å. Cd(II) is modeled as a pink sphere. D-Leu residues are shown in red as sticks and Cys side chains in purple as sticks with sulfurs labeled in yellow. The backbones of helices are shown as lines (red color referred to oxygen atoms and blue referred to nitrogen atoms).

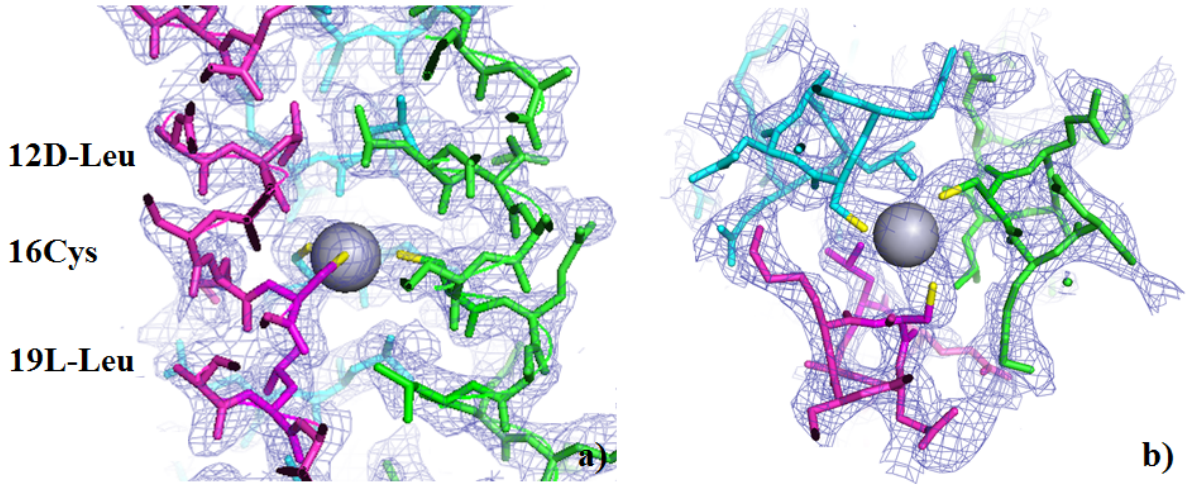

Figure 3-32. Electron density ($2F_O$-F_C) as contoured at 1.0 σ at the start of refinement of Hg(II)(GRAND-CSL12$_D$LL16C)$_3^-$ structure. The native $2F_O$-F_C map is shown in blue representing the electron density of the peptide. The peptide backbones and side chains are shown as sticks (chain A in blue, chain B in green and chain C in pink). Sulfurs are yellow. a) Side view confirming that there is no electron density for water observed around the metal site and b) top down view from the N-termini of the binding site.

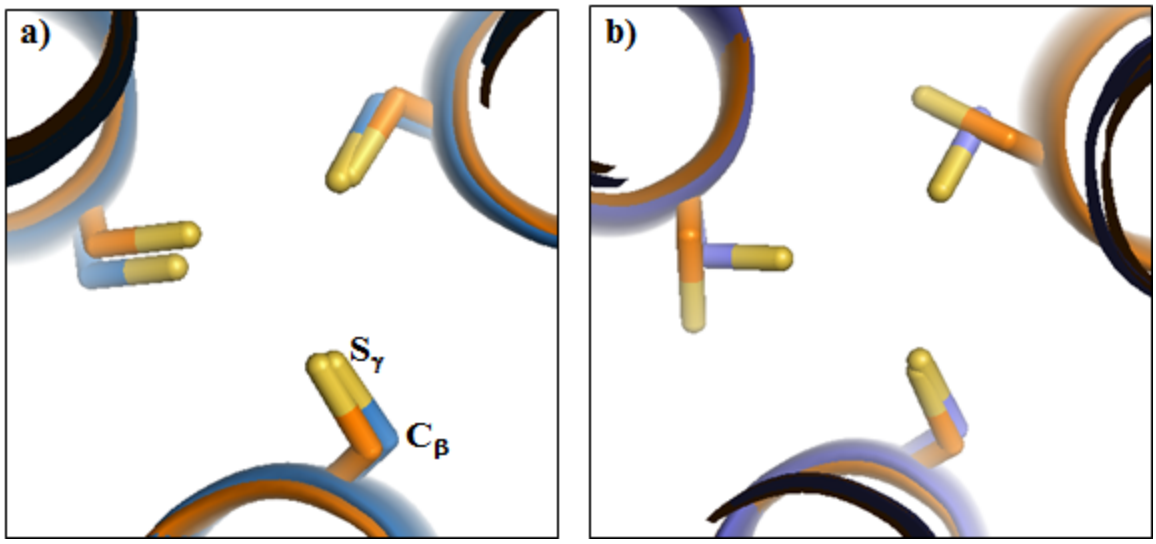

Figure 3-33. An overlay of apo-(GRAND-CSL16CL19$_D$L)$_3$ and apo-(CSL16C)$_3$ structures showing the orientations of 16Cys side chains. a) Top-down view from the N-termini representing the major Cys conformers. b) Top-down view of the minor Cys conformers observed in apo-(**CS**L16C)$_3$ (Only chain A and B that contain minor conformation) aligned with the major Cys conformers in apo-(**GRAND-CS**L16CL19$_D$L)$_3$. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ are shown in blue and apo-(**CS**L16C)$_3$ in orange. Cys side chains are present as sticks with the thiol groups labeled in yellow.

$C\delta_1$-$C\delta_1$ and $C\delta_2$-$C\delta_2$ compared to the 19L-Leu in apo-(**CS**L16C)$_3$ (see **Table 3-4**). Moreover, the pocket below the metal site made by 19D-Leu seems to be larger than 19L-Leu. This can be confirmed by the longer distance between the inner $S\gamma$ plane to the interior $C\delta$ plane ($C\delta_1$) of 19D-Leu layer compared to 19L-leu. In 19D-Leu, it is 5.18 Å, while the length is shorter by 0.30 Å with 19L-Leu. These structural affirmations clearly demonstrate that the altered chirality of D-Leu can remove the steric constraints when it is placed below the metal site by rearranging the bulky δ-methyl groups away from the center of the coiled coil, generating more open space with the less-packed hydrophobes **(Figure 3-35)**. This phenomenon is believed to allow for water access to the binding site, producing 4-coordinate Cd(II)S_3O^- species, corresponding to the first angular frequency observed at 0.316 rad/ns in the ^{111m}Cd PAC.

As previously shown with the L16C variant, the hydrophobic 12L-Leu packing can accommodate a solvent water to bind Cd(II) as a 4-coordinate Cd(II)S_3O^-. Hence, at this point, one cannot exclude that some portion of Cd(II)S_3O^- in L16CL19$_D$L could be generated by the water that enters above the metal site to bind Cd(II) from the top. Due to the limitation of spectroscopic techniques, these two conformations of Cd(II) products cannot be distinguished using the 0.316 rad/ns ^{111m}Cd PAC angular frequency value. In order to assess this assumption, the 12L-Leu packing in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ needs to be further investigated. It is proposed that if this hypothesis is true, the 12L-Leu in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ should show sufficient space for water, comparable to the 12L-Leu layer in apo-(**CS**L16C)$_3$. The overlay structures of both apo-peptides (**Figure 3-36**) reveals that all the Cβ carbons in 12L-Leu are directing toward the N-termini due to their L-configuration. The 12L-Leu in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ point one δ-methyl group ($C\delta_2$) to the core, while the other ($C\delta_1$) is oriented more toward the helical interface. The similar observation is also observed in the apo-(**CS**L16C)$_3$ parent peptide. In **Figure 3-37**, the packing in both structures look similar even though in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ the layer is slightly tighter-packed and the cavity is smaller. This can be confirmed by the shorter separation (4.72 Å) between the 12L-Leu and the inner $S\gamma$ conformer in apo-(**GRAND-CS**L16CL19$_D$L)$_3$, while it is 5.08 Å in the apo-(**CS**L16C)$_3$. However, both of which agree to generate a larger space above the metal site when compared to the smaller pocket made by the 12D-Leu in apo-(**GRAND-CS**L12$_D$LL16C)$_3$ (2.32 Å). The presentation of the corresponding

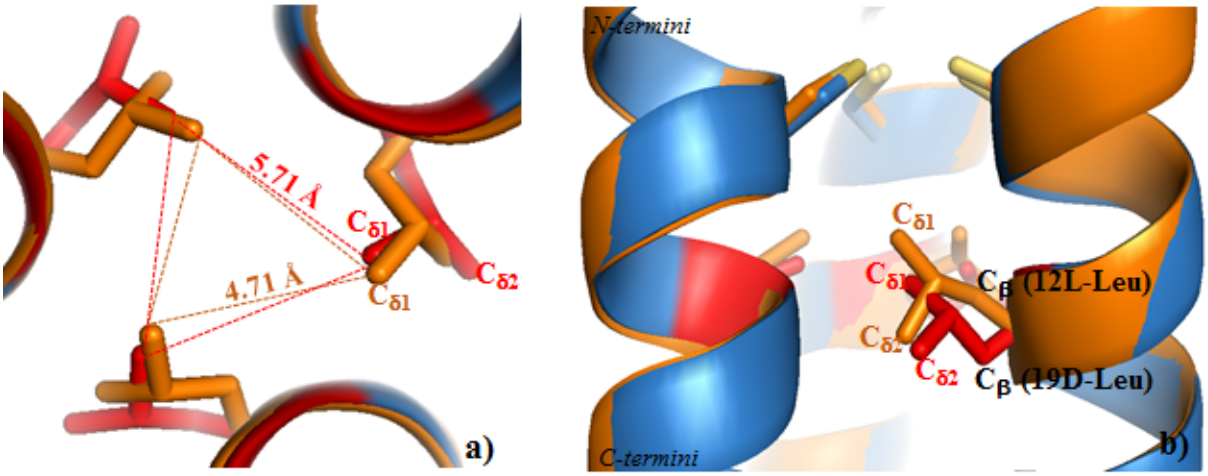

Figure 3-34. An overlay of apo-(**GRAND-CSL16CL19$_D$L)$_3$ and apo-(CSL16C)$_3$ structures representing the D-Leu layer compared to L-Leu in the nineteenth layer.** a) Top down view from the N-termini. b) Side view of the sixteenth (Cys, top) and nineteenth (bottom) layers. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ are shown in blue and apo-(**CS**L16C)$_3$ in orange. Cys side chains are present as sticks with the thiol groups labeled in yellow.

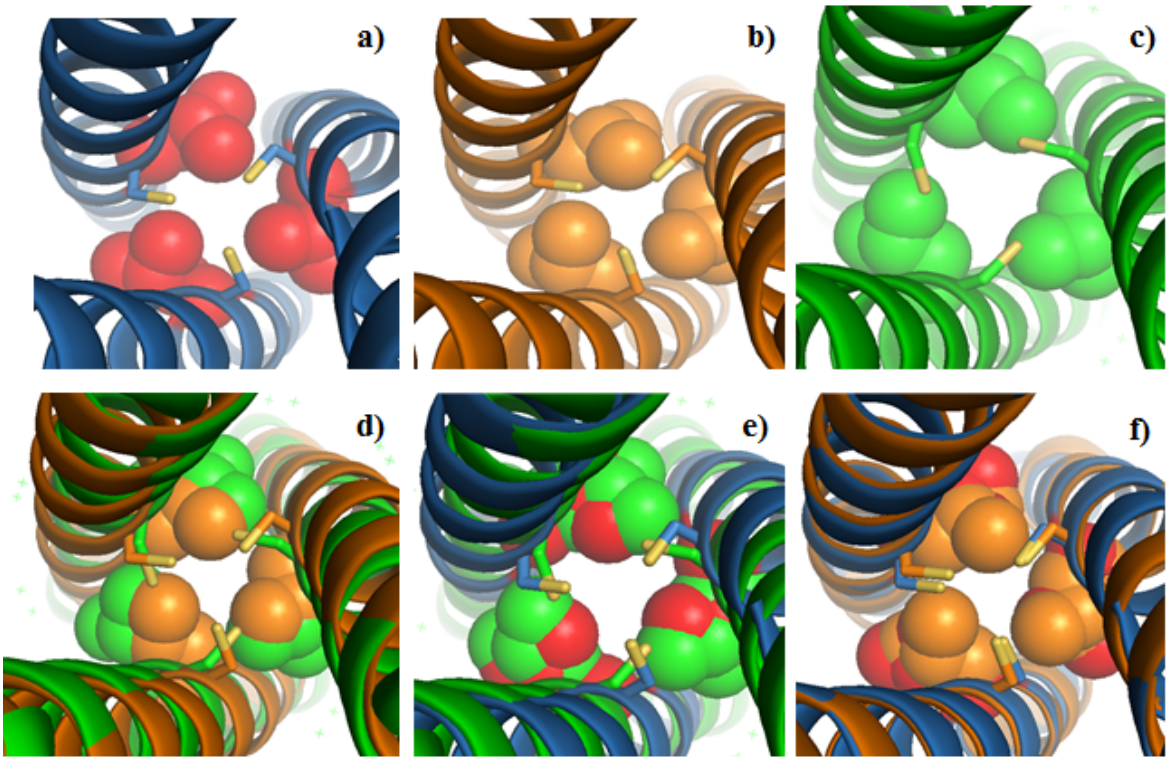

Figure 3-35. Packing of hydrophobic layers at the nineteenth position below the 16Cys site of the peptides shown as spheres. Top: From top down view of the N-termini, representing a) 19D-Leu packing in apo-(**GRAND-CS**L16CL19$_D$L)$_3$, b) 19L-Leu packing in apo-(**CS**L16C)$_3$ and c) 19L-Leu packing in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. **Bottom:** Representing the packing comparison in the nineteenth layers between d) apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, e) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, f) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and apo-(**CS**L16C)$_3$. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$, apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in blue, orange and green respectively. D-Leu residues are red and thiol atoms are yellow.

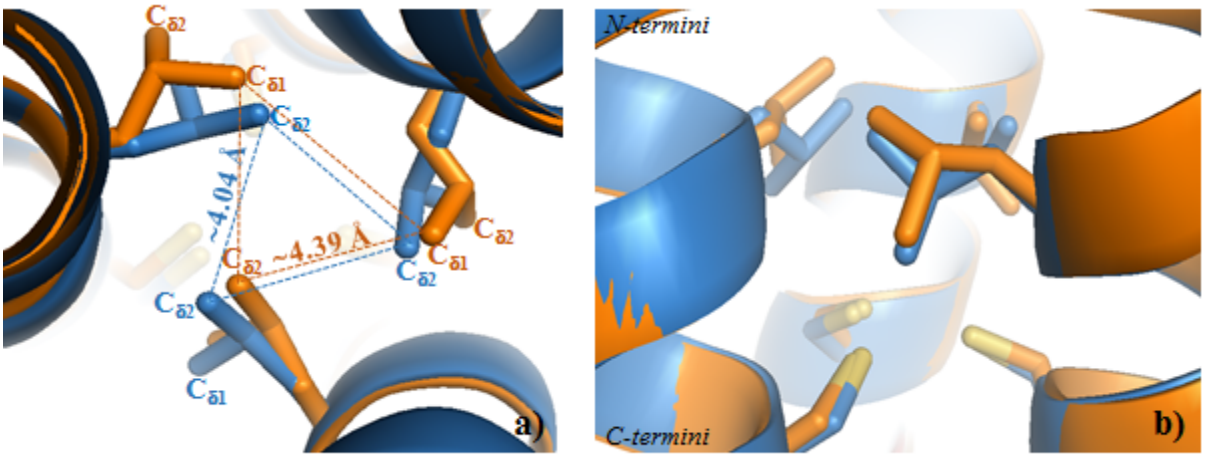

Figure 3-36. An overlay of apo-(GRAND-CSL16CL19DL)$_3$ and apo-(CSL16C)$_3$ structures showing the comparison of 12L-Leu packing between both structures. a) Top down view from the N-termini of the twelfth layer. b) Side view of the twelfth (L-Leu, top) and sixteenth(Cys, bottom) layers. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ are shown in blue and apo-(**CS**L16C)$_3$ in orange. Cys side chains are present as sticks with the thiol groups labeled in yellow.

interlayer spaces is demonstrated in **Figure 3-18**. As suggested from the crystal structures, however, the cavity generated by 12L-Leu in the apo-(**GRAND-CS**L16CL19$_D$L)$_3$ could be big enough to house a water ligand above the metal site that can allow for Cd(II)S$_3$O$^-$ formation. Crystallographically confirmed, there are two spaces available in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ for water access; one above and the other below the metal site. Therefore, the appearance of the 0.316 rad/ns angular frequency from ^{111m}Cd PAC could represent both 4-coordinate Cd(II) conformations where one has water bound on top with respect to the metal binding plane and the other has water bound below as suggested from the following described models.

To model the possible Cd(II) structures in a peptide with a D-Leu beneath the metal binding site, apo-(**GRAND-CS**L16CL19$_D$L)$_3$ was overlaid onto the Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ (**Figure 3-38**). The difference in Cys-torsion angles suggests that there must be a reorganization of the thiols in the apo-coiled coil to rotate their side chains down toward the outer face of the structure in order to accept the metal into a trigonal binding layer as previously seen in the case of **GRAND-CS**L12$_D$LL16C. As shown in **Figure 3-28**, when Hg(II) is bound to the metal site, the layer at the nineteenth position in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ opens and moves toward the helical interface causing the structure to become less well-packed compared to the apo-(**CS**L16C)$_3$, but within similar extent to the 19D-Leu packing of the 19D-Leu in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ (**Figure 35**,**e**).

The packing of leucine residues in the twelfth position between apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ structures are also similar as shown in **Figure 37**,**d**. The distance between the inner δ-methyl atom to the water with 12L-Leu is 4.44 Å in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ while it is 4.74 Å in apo-(**GRAND-CS**L16CL19$_D$L)$_3$. Subsequently the 12L-Leu cavity in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ should be able to accommodate a water molecule when the metal is bound as previously predicted. It is assumed that if the distance of 4.44 Å is sufficient to accommodate a water in the hydrophobic leucine environment in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, then the length of 4.74 Å in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ should also allow entrance of water as well. Therefore, the 12L-Leu layer comparison between these two structures indicates that there is an enough space for a coordinating water above the metal site. To model the Cd(II)S$_3$O$^-$ species with 0.316 rad/ns from ^{111m}Cd PAC, it is assumed that the water can enter both cavities either above (at the twelfth position) and below (at the nineteenth position) resulting in two conformations of 4-coordinate

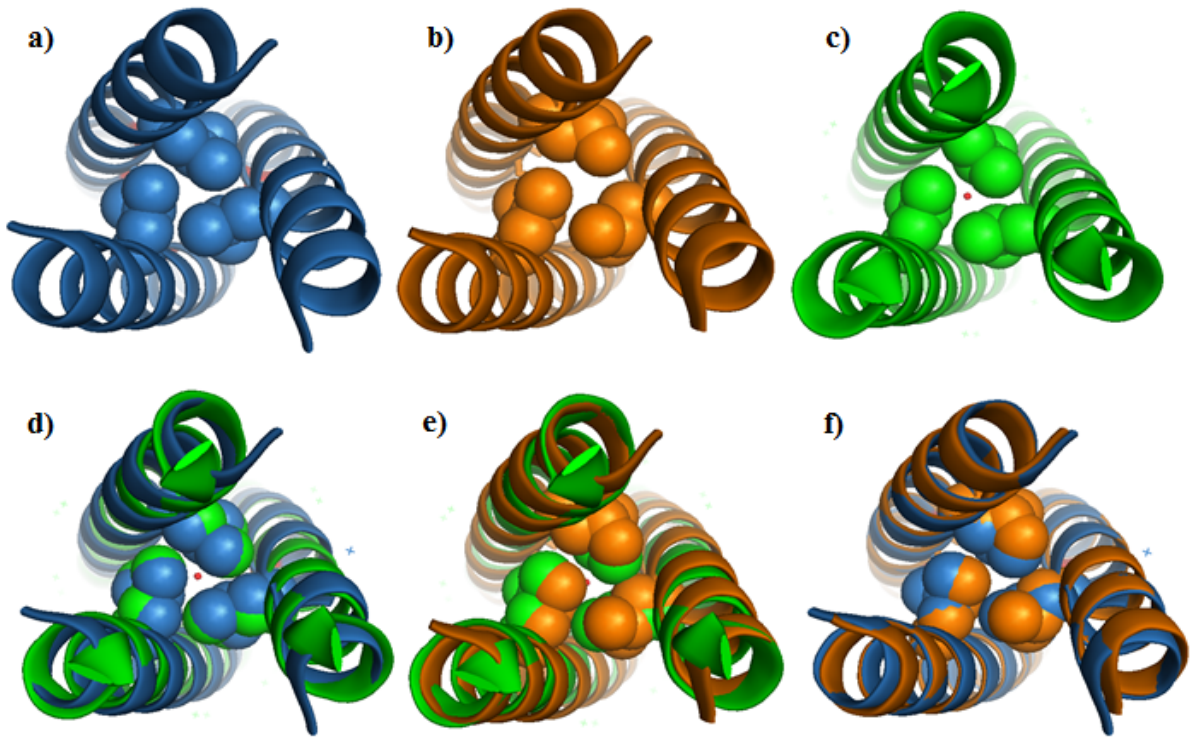

Figure 3-37. Packing of hydrophobic layers at the twelfth position above the 16Cys site of the peptides shown as spheres. Top: From top down view of the N-termini, representing the 12L-Leu layer of a) apo-(**GRAND-CS**L16CL19$_D$L)$_3$, b) apo-(**CS**L16C)$_3$ and c) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. **Bottom:** Representing the packing comparison in the twelfth layers between d) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, e) apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, f) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and apo-(**CS**L16C)$_3$. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$, apo-(**CS**L16C)$_3$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$are shown in blue, orange and green respectively. The observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is shown in a small red sphere.

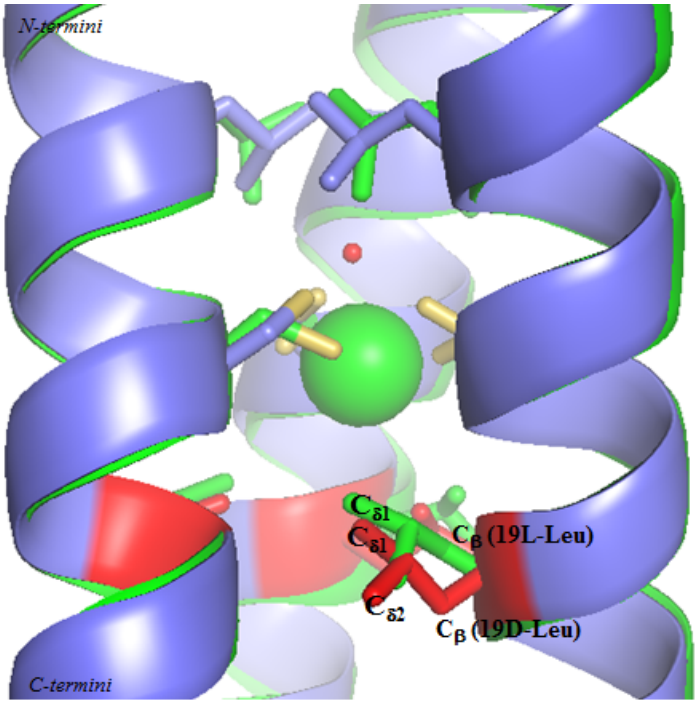

Figure 3-38. An overlay of apo-(GRAND-CSL16CL19$_D$L)$_3$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structures representing the Cδ_1, Cδ_2 atoms of leucine residue orientations in the nineteenth position. Main atoms of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ are colored in blue where the 19D-Leu residues are shown as red sticks. Main atoms and the Hg(II) ion in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are green. 19L-Leu are as green sticks. The observed water in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is red. Thiols of 16Cys residues are yellow.

Cd(II) complexes. In the first case, the water binds on top as in the L16C variant generating an endo-Cd(II) conformation. Again the coordinate of the Hg(II) ion in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ is exploited to model the Cd(II) complexes. To make a pseudo-tetrahedral structure, the modeled Cd(II) was moved slightly out of the trigonal plane as well as requiring a small rotation of three Cys ligands to reach an agreement of 2.52 Å Cd(II)-S distance suggested from 4-coordinate $Cd(II)S_3O^-$ EXAFS distance in **Table 3-5**. The S-Cd(II)-S angle is 118.07°. The fourth ligand is the water, situated 2.35 Å away from the metal center. The final model of this situation strongly suggests that Cd(II) could bind in an endo-conformation, where both the β-carbon and the metal orient on the same side of the three atom sulfur plane as illustrated in **Figure 3-39**,**a**. The second possibility is when water access is at the open space made by 19D-Leu. In this case, the modeled metal was moved down by half an angstrom from the trigonal plane toward the C-terminus. This shift places the metal about 1 Å below the actual metal location. The Cd(II)-S distance of 2.52 Å was resulted in an averaged χ_1 Cys torsion angle of -141.81°. The water was placed axially at 2.35 Å below the modeled metal as the fourth coordinating ligand. The S-Cd(II)-S is modeled to be 113.60°. In this situation, the Cd(II) center which had been predicted to adopt an endo conformation, is better considered as an exo-conformation, where the metal plane is on the opposite side of the Cβ(Cys) plane with respect to the Sγ(Cys) layer as shown in **Figure 3-39**,**b**.

Apart from the 0.316 rad/ns angular frequency species observed in ^{111m}Cd PAC, another 50% of the products from L16CL19$_D$L was found to be 0.159 rad/ns which is uncommon in these peptide systems. However, based on the theoretical calculations by Professor Lars Hemmingsen, this lower frequency is consistent with a 5-coordinate $Cd(II)S_3O_2^-$ complex. Since the angular frequency is closer to zero, the NQI around the metal site is relatively symmetrical (where the prefect tetrahedral geometry ideally shows ω_o of 0 rad/ns).[43] Therefore, this 0.159 rad/ns value suggests an axially symmetric trigonal bipyramidal structure. As shown by the crystallography described above, the existence of two cavities above and below the metal site in the apo-(**GRAND-CS**L16CL19$_D$L)$_3$ supports this possibility as space is available for water to form $Cd(II)S_3O_2^-$ structure. In **Figure 3-39**,**c** Cd(II) is modeled to be situated in the trigonal Cys plane of apo-(**GRAND-CS**L16CL19$_D$L)$_3$ structure using Hg(II) coordinate of $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. The modeled Cd(II)-S distance is 2.48 Å based on the EXAFS derived distance (**Table 3-5).** Two waters are are modeled at 2.37 Å on top and below the metal, perpendicular to

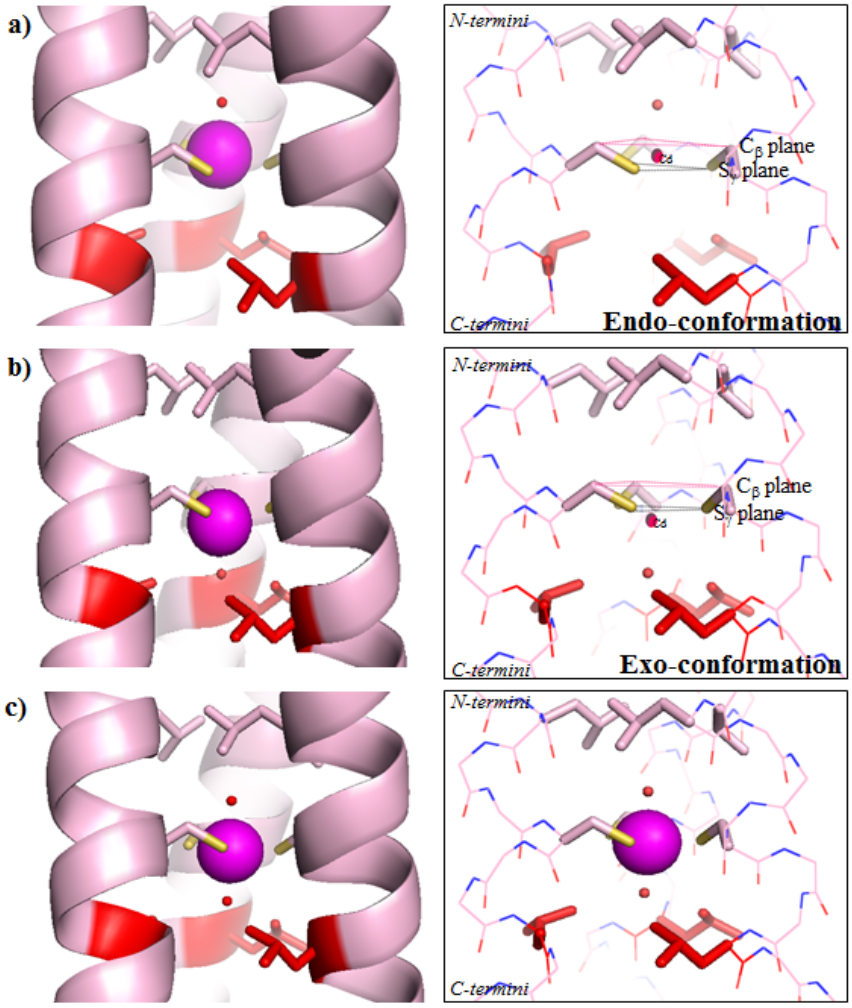

Figure 3-39. Models of $Cd(II)S_3O^-$ and $Cd(II)S_3O_2^-$ using apo-**$(GRAND\text{-}CSL16CL19_DL)_3$ crystal structure. Left** representing cartoon diagrams corresponding to the PyMOL line diagrams **(Right)** of a) an *endo*-$Cd(II)S_3O^-$; b) an *exo*-$Cd(II)S_3O^-$ and c) a $Cd(II)S_3O_2^-$ models. 12L-Leu(pink), 16Cys(pink, thiols in yellow) and 19D-Leu(red) layers are shown as sticks. Modeled Cd(II) is present a sphere in magenta. Modeled water(s) is labeled in red.

the Cys plane. Again, this Cd(II)-O distance is consistent with X-ray absorption data. The modeled S-Cd(II)-S of 119.49° . The final model yields a $Cd(II)S_3O_2^-$ structure displaying a trigonal bipyramidal polyhedron.

Excitingly, proof of the possibility for simultaneous water access in this peptide was obtained with the structure of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. Though the metal center is not Cd(II)-bound, the mercurated binding site reveals some interesting aspects related to the previous predictions. As shown in **Figure 3-40**, there are four spherical shapes of positive electron density (F_O -F_C map), each of which represents a water molecule at a contoured level of 3σ with an r.m.s.d. level greater than 1 electron/Å^3. The first density is above the metal site, at a 2.72 Å distance away from the Hg(II) center. This distance is similar to the previously observed water found in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$. It is stabilized by H-bonding interactions with Cys ligands and the second water which is located close to the helical interface between chains A and B. This second water is 3.13 Å from the first water and 4.39 Å from the Hg(II) center. The reason this second water is observed in this structure but not in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is probably because Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ was crystalized in $P2_12_12_1$ space group which does not impose three fold symmetry on the three helices. However, the R32 space group in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ is tightly packed and the crystallographic symmetry constraint may exclude any waters which are not aligned on the three fold axis, thus resulting in the appearance of there being only one axial water positioned above the Hg(II). The third and fourth waters are positioned within the nineteenth D-Leu cavity below the Cys plane; 3.75 Å and 5.84 Å away from Hg(II), respectively. The third water is positioned near chain B above the $C_{\delta 2}$ plane of the 19D-Leu, lying close to the helical interface and showing strong H-bonding interactions with the thiol and carbonyl oxygen of 16CysB. Moreover, it is at a distance of 3.09 Å from the fourth water situated toward the C-termini. This fourth water, unlike its neighbor, is located more to the center of the helix on the same plane as $C_{\delta 1}$ atom of the19D-Leu. This evidence strongly suggests that once the metal is bound in the Cys plane in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$, the pocket made by the 19D-Leu is relatively large enough to accommodate more than just one water molecule. To analyze the impact of D-Leu on hydrophobic packing in the metallated structures further, the Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is aligned to Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ in **Figure 3-41**. **Figure 3-42** represents an excellent

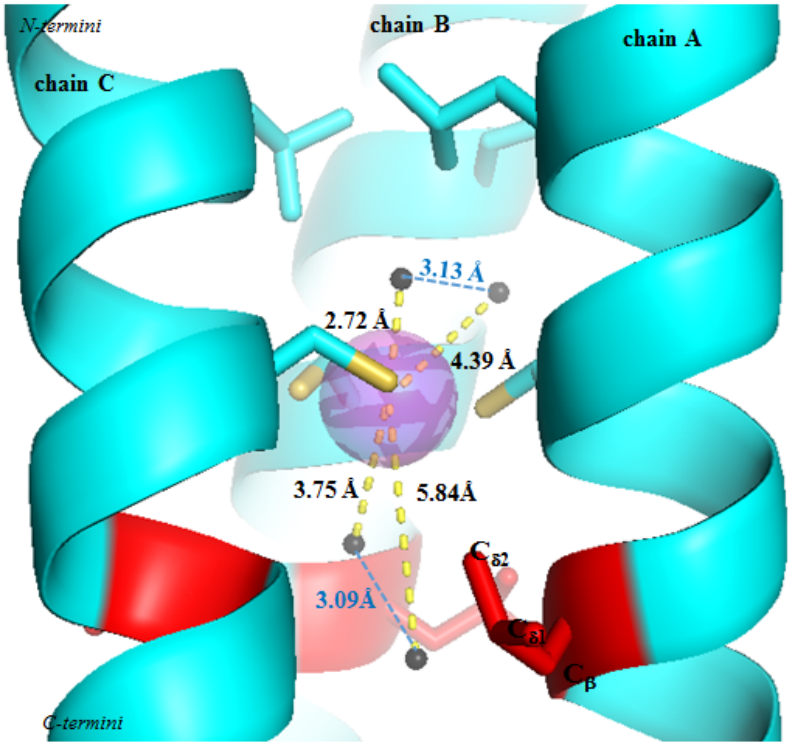

Figure 3-40. PyMOL visualization of Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ representing the existence of waters above and below the 16Cys metal site. Waters are shown as small spheres in gray and Hg(II) ion is shown as a big sphere in magenta. 19D-Leu, 16Cys and 12L-Leu are present as sticks in red, blue and blue which thiols of Cys are labeled in yellow.

overlay of the bound S_γ conformers between the two structures as confirmed from their similarity in χ_1 (**Table 3-4**). The Hg(II) ions of the two structures are in the same plane, but differ in position by 0.40 Å apart. The averaged Hg(II)-S distance is still 2.43 Å resembling the distances of trigonal planar Hg(II) structures in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and reported small molecule complexes.[29–31] As previously noted, all the waters observed in the second coordination sphere of Hg(II)-bound structures are non-bonded and believed to have H-bond interactions to the bound Cys ligands which help compensate the negative charge of the metallated site. The hydrophobic cavity of 12L-Leu in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is represented in **Figure 3-43**. This layer is slightly less-well packed compared to the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ due to the presence of the second water above the metal site that is located near the helical interface between chains A and B. The presence of this water causes the 12L-LeuB side chain to fray toward the helical interface and toward the N-termini in order to avoid steric clashing with the polar solvent molecule. The water observed directly on top of the Hg(II) atom in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is within 0.09 Å of the position observed in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. While the hydrophobicity of the layer at the twelfth positions are slightly different, both metallated-structures easily accommodate a water molecule axially above the Hg(II). Moreover, due to the low symmetry requirement of $P2_12_12_1$ space group, the cavity in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is also amenable for a second water to H-bond to the central water. These structural comparisons confirm that the space in the layer at position twelve is suitable for water access demonstrating that the $Cd(II)S_3O^-$ species with an endo-conformation in L16CL19$_D$L is reasonable.

The significant change in leucine orientations in the layer at the nineteenth position of both metallated structures shown in **Figure 3-44** strongly suggests that the19L-Leu layer in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ packs tighter than the 19D-Leu in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. In **Figure 3-45**, it is obvious that the Cβ deviation of D-Leu forces the whole side chain to be directed more toward the helical interface resulting in large separations of inner δ-methyl carbons (9.35 Å) while it is only 5.95 Å in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$. The alignment of both structures demonstrates that if waters were to be present in the metallated-L16C peptide, the tighter 19L-Leu packing would bring sufficient steric clashes between the 19L-LeuC residue in $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and the water below the metal site

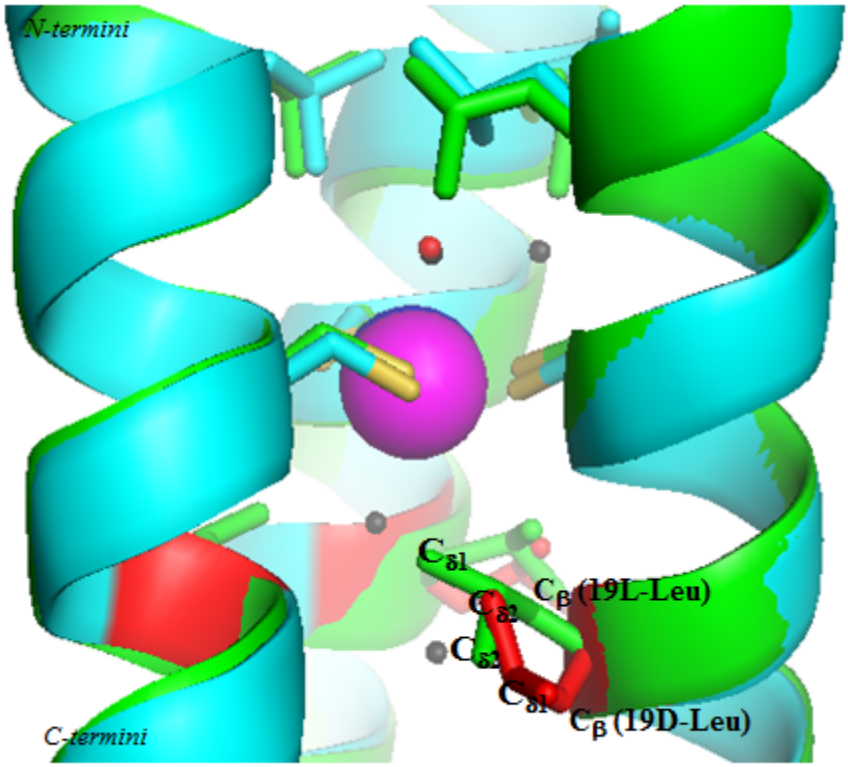

Figure 3-41. An overlay cartoon diagram of Hg(II)-(GRAND-CSL16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ representing the hydrophobic packing comparison in the twelfth and nineteenth layers as well as the 16Cys metal site shown as sticks. Main atoms of Hg(II)-(**GRAND-CS**L16CL19$_D$L)$_3^-$ are shown in cyan and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ in green. 19D-Leu residues are red-colored. Shown as spheres, the Hg(II) ion and observed waters in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ are labeled in magenta and gray, respectively whereas in blue and red for Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$.

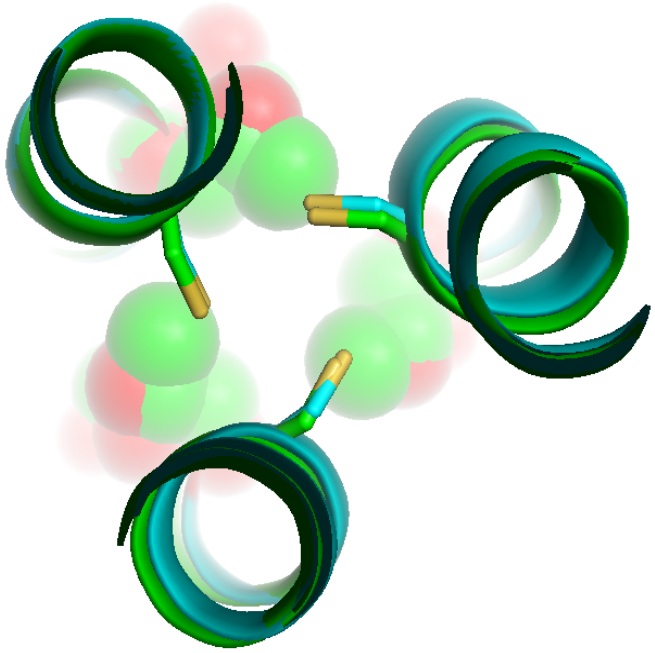

Figure 3-42. An overlay of Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ representing the inner Sγ conformers of the sixteenth layers shown as sticks. Main atoms of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ are shown in cyan and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ in green. Thiols of the 16Cys residues are colored in yellow. (Note: The experimentally observed Hg(II) and waters in both structures are not shown here in this figure.)

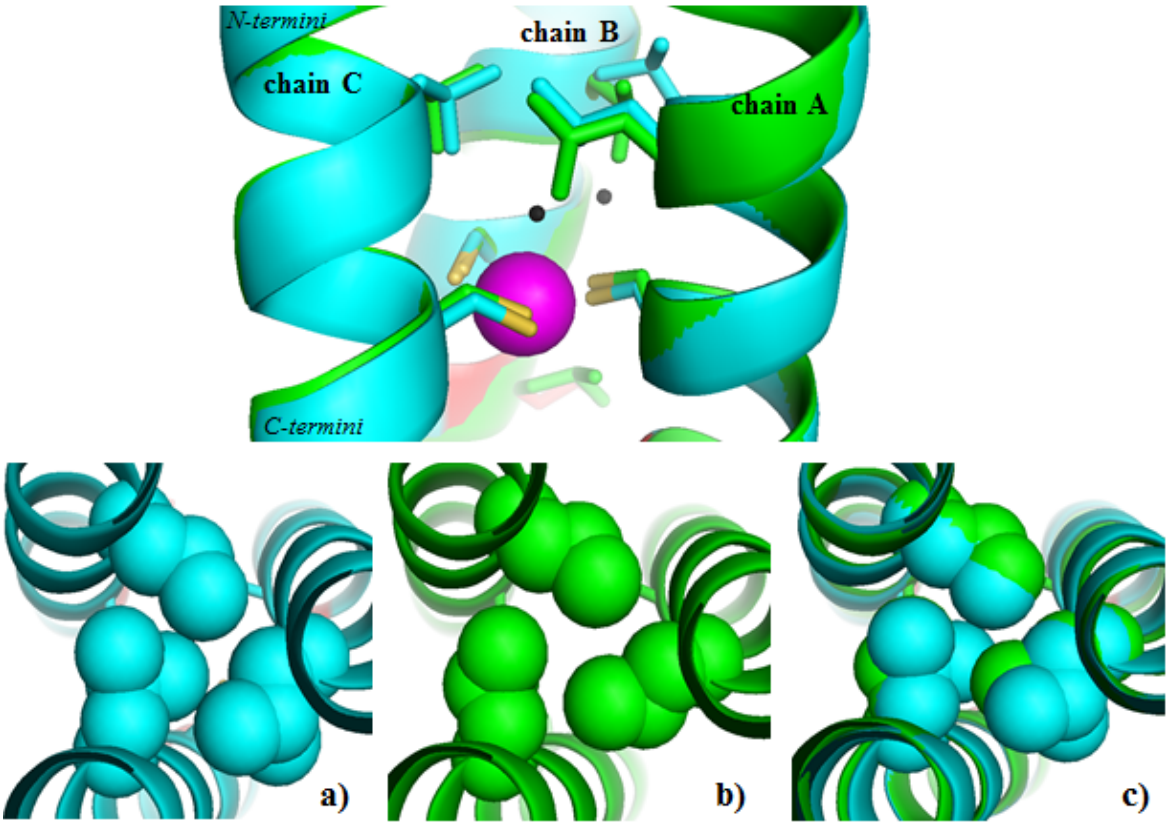

Figure 3-43. Top: An overlay of Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ showing the difference of 12L-LeuB side chain between two structures. 12L-Leu(top) and 16Cys(bottom) are shown as sticks. Main atoms of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in cyan and green respectively. The Hg(II) ion and two waters observed in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ are shown as sphere in magenta and gray respectively.

Bottom: Top down view from the N-termini illustrating the hydrophobic packing at the twelfth layer of a) Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$, b) Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and c) an overlay of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$.

(determined by using the water coordinates in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$) to block solvent access. The distance between $C_{\delta 1}$ of 19D-LeuC and water in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is 3.80 Å, while this length shortens to 2.50 Å when considered from $C_{\delta 1}$ of 19L-LeuC in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ to that water in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. This suggests that if this water was present in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$, it would lead to a drastic steric clash with the 19L-LeuC side chain. Moreover, it is believed that if this first water does not exist, the other one far below the metal site (the fourth water previously mentioned in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3$)$^-$ should no longer be present because it has no polar molecule nearby to form an H-bond. These structural comparisons not only explain the reasons why the altered D-Leu side chain removes the steric hindrance below the metal binding site, allowing for more water access in L16CL19$_D$L, but also hints at why there is no observation of water below the metal binding site that could bind to Cd(II) when the 19L-Leu configuration is placed in the L16C peptide.

The last part in the structural comparison is to address how the metal binding induces changes in the metal and second coordination spheres of the same designed peptides. Here, the Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ is analyzed and compared to apo-(**GRAND-CS**L16CL19$_D$L)$_3$. The diameter of the three stranded coiled coil aggregate in apo-(**GRAND-CS**L16CL19$_D$L)$_3$ is slightly smaller than that of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. In Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$, the peptide packing contains three-stranded coiled coil per asymmetric unit; consequently in the hydrophobic stacking layer each residue can display in different rotamers. The absence of a kink along the backbone helices demonstrates that Hg(II) binding does not disrupt the overall secondary structure of the coiled coil. However, the first metal sphere in **Figure 3-46**,**b** reveals the difference in Cys orientations from having S_γ positioned up toward the core of the structure and N-termini (χ_1 angle of -61.17^o) in the apo-peptide to having S_γ oppositely oriented to the outside and C-termini (χ_1 averaged angle of -152.49^o) when bound to Hg(II). The change in chi angles of Cys results in the difference in S_γ-S_γ separations; 3.31 Å in the apo-peptide compared to the averaged 4.21 Å in the Hg(II)-peptide. This longer S_γ-S_γ distance and the altered χ_1 in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ compared to the apo-structure in which all the side chains point toward the core reveals that Hg(II) induces changes in the Cys geometry, making them more

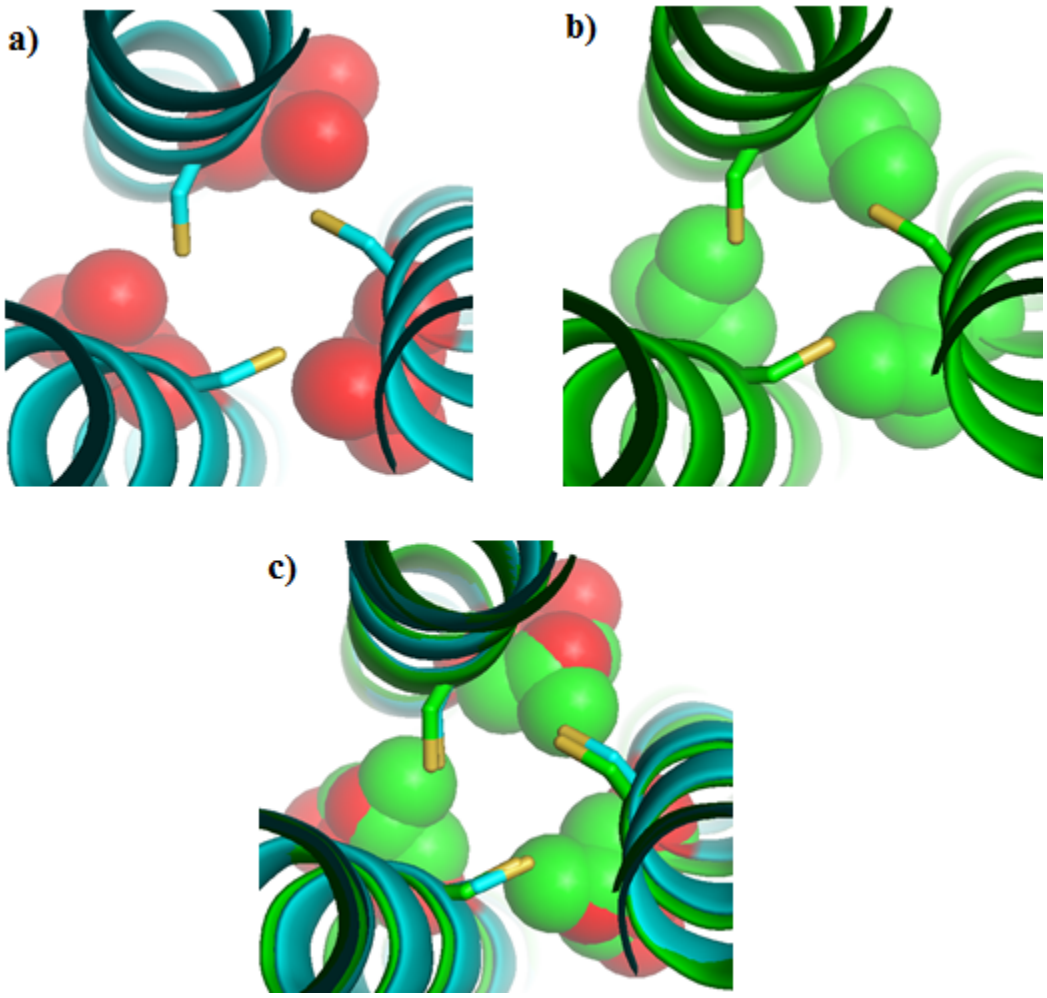

Figure 3-44. PyMOL representation of the hydrophobic packing comparison in the nineteenth position below the 16Cys site between the Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ structures. From top down view of the N-termini, representing a) 19D-Leu(red spheres) packing in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$, b) 19L-Leu (green spheres) packing in Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ and c) the packing comparison between a) and b). Main atoms of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in cyan and green respectively. 16Cys side chains are present as sticks which their thiols are labeled in yellow. (Note: The experimentally observed Hg(II) and waters in both structures are not shown here in this figure.)

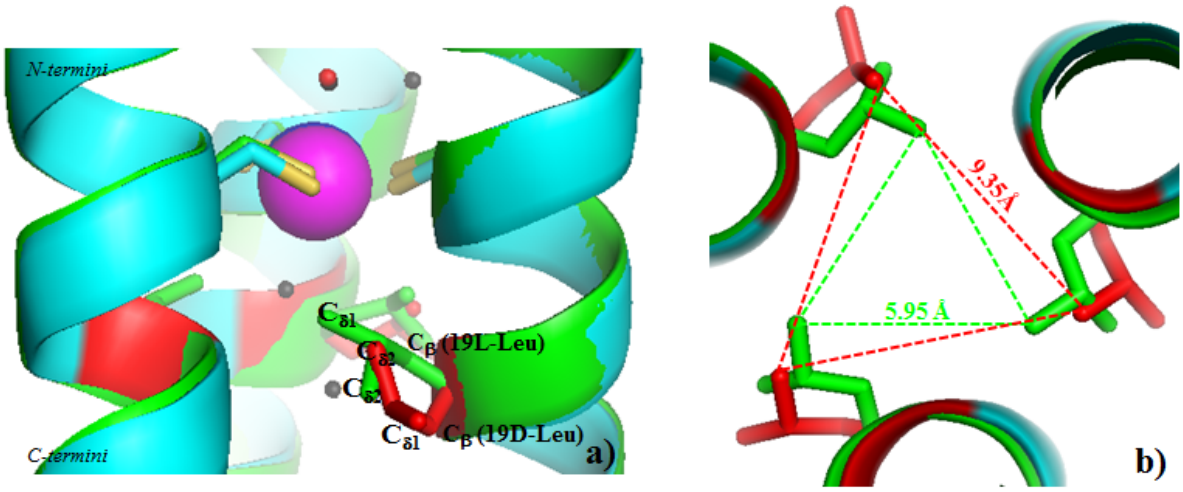

Figure 3-45. Ribbon diagrams showing an overlay of Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(GRAND-CSL16CL30H)$_3^+$ demonstrating the Cβ deviation of 19D-Leu (red sticks) from 19L-Leu (green). a) Side view of the sixteenth Cys site (top) and nineteenth (bottom) layers. b) Top down view from the N-termini of the nineteenth layer. Main atoms of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ and Hg(II)$_S$Zn(II)$_N$(**GRAND-CS**L16CL30H)$_3^+$ are shown in cyan and green respectively. 16Cys side chains are present as sticks which thiols are labeled in yellow.

constrained toward the bound metal. The rotation of thiols upon metal binding must require some reorganization energy to reorient their side chains to the helical interface, which is then overcome by the enthalpy gain from the favorable formation of Hg(II)-thiolate bonds. The relatively large size of Hg(II) with an ionic radius of 1.16 Å fits well within the trigonal plane generated by the major Cys conformer, resulting in an averaged S-Hg(II)-S of 119.75° and Hg(II)-S distance of 2.43 Å. These metrical parameters are also consistent with the results from $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ and other small molecule complexes.[29–31] Moreover, the thiols in chains A and B of Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ also show a minor contribution of a protonated form by moving from the major conformer toward the other side of the helical interface. These degrees of conformational flexibility in the minor conformations are not suitable for metal binding **(Figure 3-46,c)**. It is likely that the dynamics of these Cys side chains in the "metallated" structure comes from a small proportion of coiled coils that do not contain Hg(II). Surprisingly, the hydrophobic layers in the second coordination sphere both above (12L-Leu) and below(19D-Leu) the metal site show quite similar packing between the apo and metallated-(**GRAND-CS**L16CL19$_D$L)$_3^-$ structures. The 12L-Leu packing of the two peptides are similar as expected as shown in **Figure 47**, **top panels**. The leucine orientations direct their Cβ carbons up toward the N-termini and slightly toward the helical interface in the same degrees, as indicated by their χ_1 torsion angles that fall in the same range. The similar manner of 12L-Leu packing above the metal site suggests that the cavities of both structures are the same. On the other hand, the hydrophobic packing in the nineteenth layer becomes less well packed in the metallated structure (**Figure 3-47**, **bottom panels**). Even though all the 19D-Leu residues in both structures direct their Cβ atoms in the same manner toward the C-termini, the side chains in Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ are oriented more toward the helical interface, resulting in more open and wider cavities as compared to the 19D-Leu of the apo-peptide and also to the 12L-Leu layer. As indicated previously, the loosening hydrophobic packing upon Hg(II) binding is also observed between the $Hg(II)_SZn(II)_N$(**GRAND-CS**L16CL30H)$_3^+$ compared to the apo-(**CS**L16C)$_3$. Based on these observations, significant changes both to the thiolate conformations and adjacent layer packing are essential in order to accept the larger Hg(II) into the binding plane. Thus, one might expect that these changes would be similar for Cd(II), which has M-S bond lengths equal or larger than Hg(II).

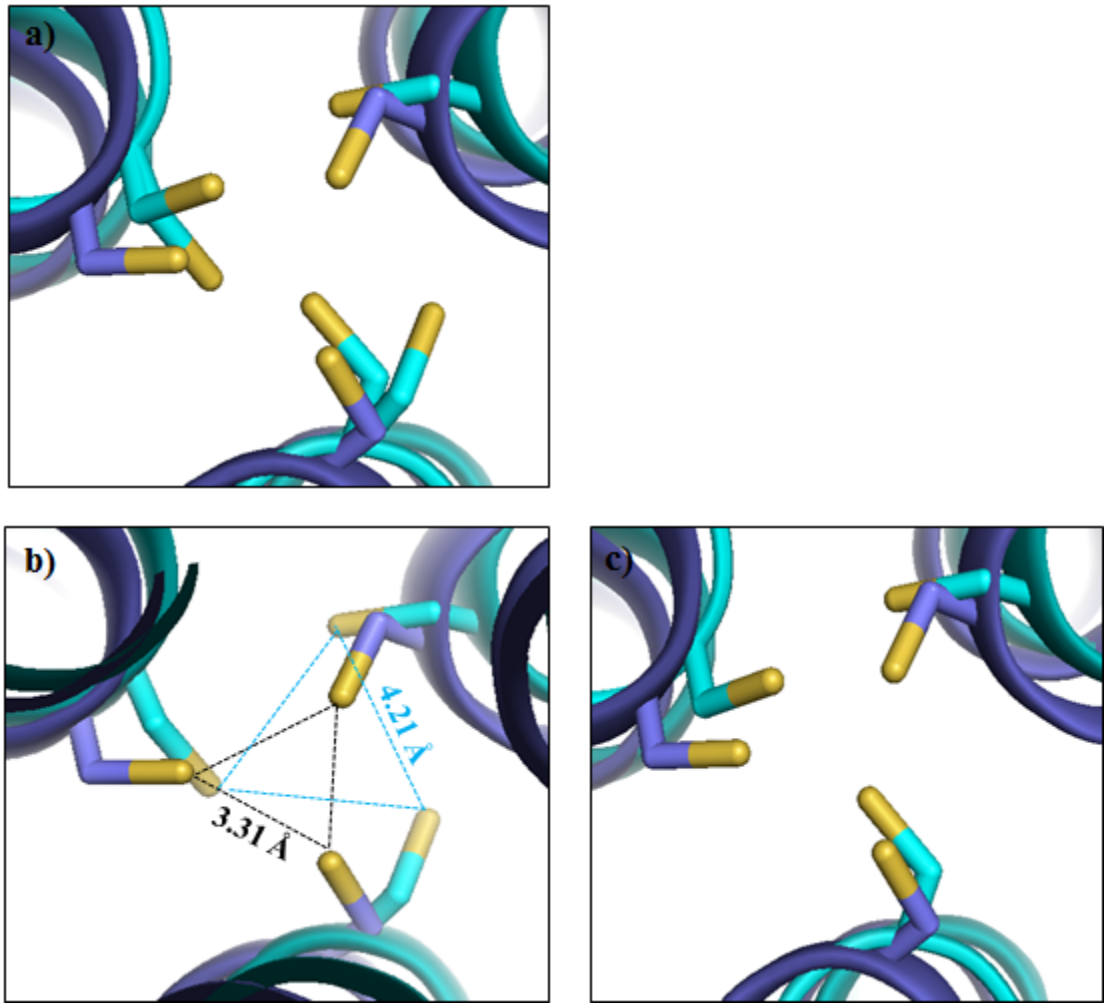

Figure 3-46. An overlay of apo-(GRAND-CSL16CL19$_D$L)$_3$ and Hg(II)(GRAND-CSL16CL19$_D$L)$_3^-$ structures showing the orientations of Cys side chains. Top-down views from the N-termini representing a) the major and minor conformers of Cys residues, b) the major conformers and c) the minor conformers(note: In Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$, only chain A and B contain minor conformation.). Main atoms are shown in blue for apo-(**GRAND-CS**L12$_D$LL16C)$_3$ and in cyan for Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. 16Cys side chains are shown as sticks with the thiol groups labeled in yellow.

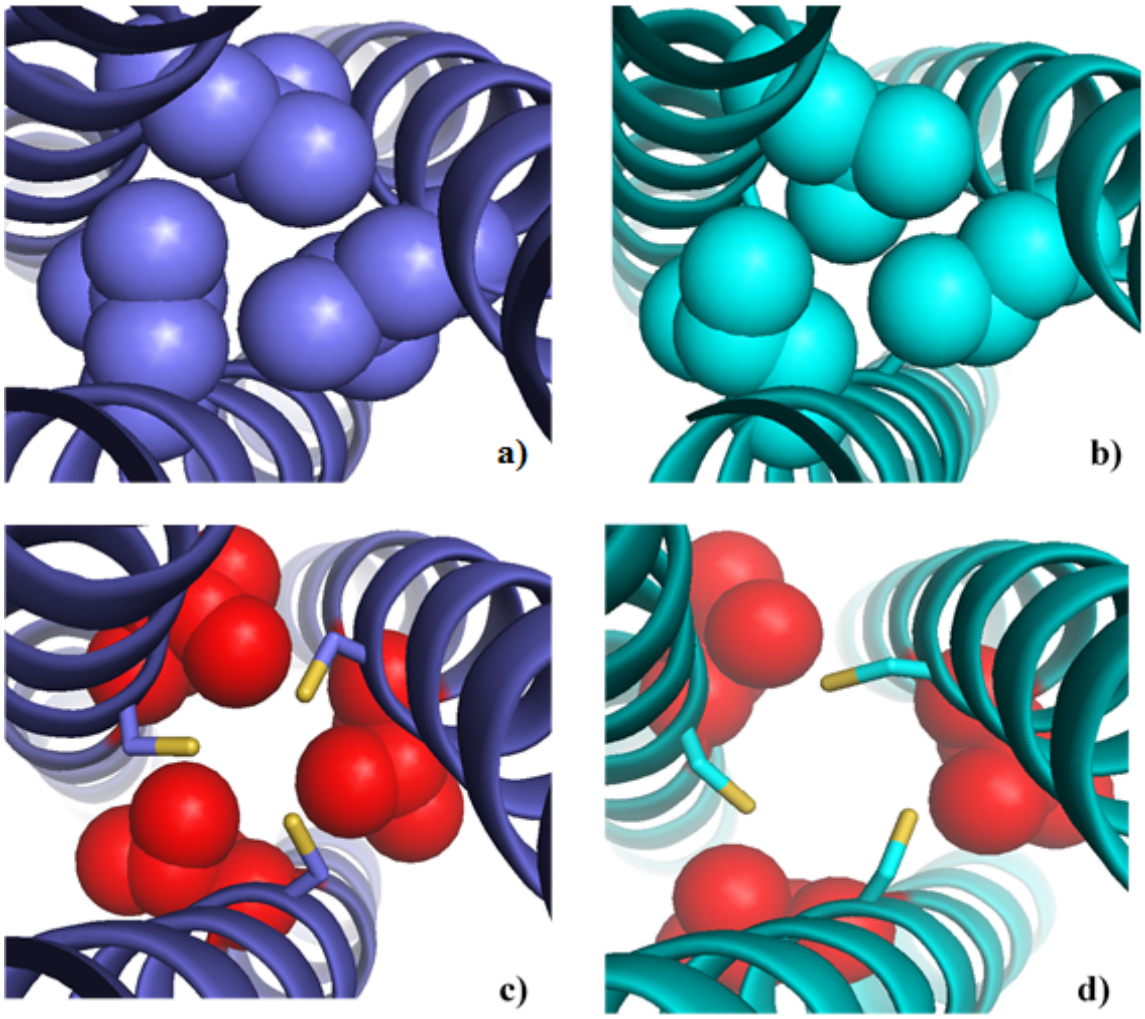

Figure 3-47. Top: **PyMOL representation of the twelfth position packing** in a) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ (blue) and b) Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$ (cyan). The backbones are shown as cartoon. Leucine residues are represent as spheres.

Bottom: **Packing of the nineteenth position** in c) apo-(**GRAND-CS**L16CL19$_D$L)$_3$ and d) Hg(II)(**GRAND-CS**L16CL19$_D$L)$_3^-$. D–Leu residues are shown as red spheres and Cys side chains are present as sticks. Thiols are yellow-coded.

Conclusion

De Novo protein design has as the promise to understand fully, and consequently build, complex biological architectures. Easily incorporated into peptides through solid phase synthesis, D-amino acids were used to explore the role of metal second coordination sphere effects on the first coordination structures of Cd(II). Here, we have shown that the introduction of a D-amino acid with a native L-amino acid sequence is plausible even though the overall stability of the scaffold decreases compared to the normal L-peptide construct. However; the desired control of metal sites was accomplished. D-amino acid incorporation, therefore, is potentially a powerful tool that significantly opens up entirely new vistas for metalloprotein design that will greatly expand a variety of biophysical, catalytic and therapeutic applications. Since D-Leu substitution in the second coordination sphere allows for significant opportunities to control Cd(II) structures, one wonders how metal structures could be altered when the first coordination sphere is perturbed by chirality. In the next chapter, D-Cys and D-Pen ligands are incorporated into the first coordination sphere of 3SCC peptides to examine this possibility.

References

(1) 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.

(2) 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**, *7863*, 6195–6196.

(3) 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.

(4) Peacock, A. F. A.; Hemmingsen, L.; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 16566–16571.

(5) Ellman, G. L. *Arch. Biochem. Biophys.* **1959**, *82*, 70–77.

(6) Cobas, C.; Cruces, J.; Sardina, F. J. In *2.3 ed.*; Universidad de Santiago de Compostela: Santiago de Compostela, Spain, **2000**.

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

(8) McMaster, W. H.; Del Grande, N. K.; Mallett, J. H.; Hubbell, J. H. *Lawrence Livermore Radiation Laboratory Report*, Livermore, CA, **1969**.

(9) Duhme, A. K.; Strasdeit, H.; Anorg, Z. *Allg. Chem.* **1999**, 625,6-8.

(10) Clark-Baldwin, K.; Tierney, D.; Govindaswamy, N.; Gruff, E. S.; Kim, C.; Berg, J.; Koch, S. A.; Penner-Hahn, J. E. *J. Am. Chem. Soc.* **1998**, *120*, 8401.

(11) Otwinowski, Z.; Minor, W. *Methods Enzymol.* **1997**, *276*, 307-326.

(12) Vagin, A.; Teplyakov, A. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2010**, 22–27.

(13) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2011**, *67*, 235–242.

(14) Potterton, E.; Briggs, P.; Turkenburg, M.; Dodson, E. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2003**, *59*, 1131–1137.

(15) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(16) Emsley, P.; Cowtan, K. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2004**, *60*, 2126-2132.

(17) Roversi, P.; Sharff, A.; Smart, O. S.; Vonrhein, C.; Womack, T.O. (**2011**). BUSTER version 2.11.2 Cambridge, United Kingdom: Global Phasing Ltd.

(18) Chakraborty, S.; Touw, D. S.; Peacock, A. F. A.; Stuckey, J.; Pecoraro, V. L. *J. Am. Chem. Soc.* **2010**, *132*, 13240–13250.

(19) Langer, G.; Cohen, S. X.; Lamzin, V. S.; Perrakis, A. *Nat Protoc.* **2008**, *3*, 1171-1179.

(20) Murshudov, G. N.; Skubak, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2011**, *67*, 355-367.

(21) Terwilliger, T. C.; Adams, P. D.;Read, R. J.; McCoy, A. J.; Moriarty, N. W.; Grosse-Kunstleve, R. W.; Afonine, P. V.; Zwart, P. H.; Hung, L. W. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2009**, *65*, 582-601.

(22) Grosse-Kunstleve, R.W; Adams, P.D. *Acta Cryst.* **2003**, *59*, 1966-1973.

(23) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658–674.

(24) Terwilliger, T. C. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2000**, *56*, 965-972.

(25) Chen, V. B.; Arendall III, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D.C. *Acta Crystallogr., Sec D: Biol. Crystallogr.* **2010***, 66,* 12-21.

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

(27) Ogihara, N. L.; Weiss, M. S.; DeGrado, W. F.; Eisenberg, D. *Protein Sci.* **1997**, *6*, 78–86.

(28) Ponder, J. W.; Richard, F. W. *J. Mol. Biol.* **1987**, *193*, 775–791.

(29) Christou, G.; Folting, K.; Huffman, J. C. *Polyhedron* **1984**, *3*, 1247-1253.

(30) Watton, S. P.; Wright, J. G.; MacDonnell, F. M.; Bryson, J. W.; Sabat, M.; O'Halloran, T. V. *J. Am. Chem. Soc.* **1990**, *112*, 2824.

(31) Gruff, E. S.; Koch, S. A. *J. Am. Chem. Soc.* **1990**, *112*, 1245-1247.

(32) Peacock, A. F. A.; Stuckey, J. A.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2009**, *48*, 7371–7374.

(33) Chen, Y. H.; Yang, J. T.; Chau, K. H. *Biochemistry* **1974**, *13*, 3350– 3359.

(34) Dieckmann, G. R. Ph.D. Dissertation, University of Michigan, 1995.

(35) Iranzo, O.; Jakusch, T.; Lee, K.-H.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry* **2009**, *15*,

3761–3772.

(36) Shannon, R. D. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1976**, *32*, 751–767.

(37) Iranzo, O.; Thulstrup, P. W.; Ryu, S.-B.; Hemmingsen, L.; Pecoraro, V. L. *Chemistry* **2007**, *13*, 9178–9190.

(38) Lee, K.-H.; Cabello, C.; Hemmingsen, L.; Marsh, E. N. G.; Pecoraro, V. L. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 2864–2868.

(39) Ghosh, D.; Lee, K.-H.; Demeler, B.; Pecoraro, V. L. *Biochemistry* **2005**, *44*, 10732–10740.

(40) Santos, R. A.; Gruff, E. S.; Koch, S. A.; Harbison, G. S. *J. Am. Chem. Soc.* **1991**, 469–475.

(41) Touw, D. S.; Nordman, C. E.; Stuckey, J. a; Pecoraro, V. L. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 11969–11974.

(42) Zastrow, M.; Pecoraro, V. L. *Coord. Chem. Rev.* **2013**, 2565-2588.

(43) Hemmingsen, L. *Encycl. Inorg. Bioinorg. Chem.* **2011**, 1–17.