Published in final edited form as:

Isr J Chem. 2015 January ; 55(1): 85–95. doi:10.1002/ijch.201400146.

Sculpting Metal-binding Environments in *De Novo* Designed Three-helix Bundles

Jefferson S. Plegaria and Vincent L. Pecoraro

930 North University Ave, Department of Chemistry, University of Michigan, Ann Arbor, Michigan 48109 (USA)

Abstract

De novo protein design is a biologically relevant approach used to study the active centers of native metalloproteins. In this review, we will first discuss the design process in achieving α_3D , a *de novo* designed three-helix bundle peptide with a well-defined fold. We will then cover our recent work in functionalizing the α_3D framework by incorporating a tris(cysteine) and tris(histidine) motif. Our first design contains the thiol-rich sites found in metalloregulatory proteins that control the levels of toxic metal ions (Hg, Cd, and Pb). The latter design recapitulates the catalytic site and activity of a natural metalloenzyme carbonic anhydrase. The review will conclude with future design goals aimed at introducing an asymmetric metal-binding site in the α_3D framework.

Keywords

protein design; metal ions; metalloproteins; three-helix bundles; three-stranded coiled-coils

1. Introduction

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

Protein design is a biologically relevant approach used to study the concept of the structure-function relationship in native proteins.^[2–4] This emerging approach has two central design strategies: the first is protein redesign,^[2] and the second is *de novo* design.^[2–7] Protein redesign involves modifying or incorporating the desired metal-binding site in an existing

native protein scaffold. The second approach, which is unquestionably the most challenging strategy, employs first principles to design peptide or protein scaffolds from scratch, with an amino acid sequence not found in nature.^[6] It allows one to tailor-design a sequence that forms the proper hydrophobic, electrostatic, and hydrogen-bonding interactions that will manifest into a well-defined peptide scaffold, an important characteristic of native proteins. *De novo* protein design offers a novel approach in studying the mechanisms behind protein folding and exploring the active sites of native proteins in a simplified or unrelated fold. The knowledge gained from this approach could ultimately provide insight into the fundamental processes of biological systems, thus allowing the possibility of producing new metalloproteins, with higher stability and superior efficiency than native proteins, for many biotechnological applications.

We are actively involved in the *de novo* protein design of peptide scaffolds to understand the metal active sites of metalloenzymes (Figure 1a), as well as metalloproteins involved in electron transfer (ET, Figure 1b), metal-regulated gene expression,^[1,8] and metallochaperones.^[1] These last systems are a subset of metalloproteins used by microorganisms to control the levels of essential transition metal ions (Fe, Cu, Zn, and Mn), decrease levels of toxic metals (Hg, As, Pb, and Cd) within their cells, and to ensure proper trafficking and insertion of metals into enzymes or secretory vesicles. This review focuses on a *de novo* designed scaffold, α_3D ,^[9] a single polypeptide chain that folds into a three-helix bundle (THB). Section 2 will first briefly cover our work and progress with three-stranded coiled-coil (3SCC) constructs. It will then be followed by the achievement of DeGrado and coworkers in producing a THB scaffold. Next, section 3 will cover our

current work and first approach in functionalizing the α_3D scaffold through the incorporation of symmetric metal-binding sites to yield peptides α_3DIV ^[10] and α_3DH_3 .^[11] Our work comprises a short description of the solution structure of α_3DIV (unpublished work), an α_3D analogue that contains a tris(cysteine) motif, and its heavy metal binding properties. In addition, we will also describe our progress with a carbonic anhydrase model, α_3DH_3 , an iteration of α_3D that contains a tris(histidine) site. The last section will discuss future designs that aim to improve metal binding in α_3DIV and catalytic activity in α_3DH_3 , as well as to model an asymmetric His₂CysMet copper site.

2. Development of *De Novo* Designed Scaffolds

2.1. Three-stranded Coiled-coil Constructs

The most common *de novo* designed peptides use a heptad repeat sequence (abcdefg) that self-assemble into a parallel 3SCC tertiary structure (Figure 2a).^[7,14,15] Much of our effort and success in the *de novo* design field has been carried out using the 3SCC scaffold of TRI,^[16–31] Grand (Gr),^[24,26,32,33] and BABY^[18,33,34] peptides, as well as in CoilSer (CS),^[22,27,28,33,35–37] which serves as a crystallographic analogue for TRI (Table 1). A metal-binding site is generated by incorporating a cysteine or penicillamine residue at the “a” (Figure 2b) or “d” (Figure 2c) position in a 3SCC scaffold.^[37] We discovered that the subtle difference between the “a” and “d” positions can produce distinctive outcomes in heavy metal binding affinity and geometry, which can be attributed to the pre-organization of the sulfur ligands prior to metal binding.^[14,17] Overall, we have gained a deeper

understanding in the metallobiochemistry of heavy metals, such as As(III),^[19,28,35] Cd(II),^[14,19–22,24–26,32] Hg(II),^[16–19,22,23,25,27,34] and Pb(II),^[21,28,33,38] in a tristhiolate site. We have demonstrated how to control the coordination number and geometry of Cd(II)^[14,24,26,32] and Hg(II),^[18,19,23,25] determined the affinity for Cd(II)^[21] and Pb(II),^[21,33] based on site preferences for “a” or “d” sites of the Cys residues, and elucidated the effects of the core aliphatic groups in the second coordination sphere on the molecular recognition of Cd(II)^[14,24,26] and Pb(II).^[33] This work on heavy metal chemistry in 3SCC is a culmination of over 10 years of research and has given us a solid foundation for modeling catalytic sites of natural metalloenzymes.

Using the 3SCC scaffold, our work has progressed into modeling the symmetric tris(histidine) metal-binding site found in carbonic anhydrase (CA)^[27,30] and nitrite reductase (CuNiR).^[29,31] The [Hg(II)]₃[Zn(II)(H₂O/OH[−])]_N(TRIL9CL23H)₃ⁿ⁺ CA model (Figure 2a) of Zastrow *et al.* contains a structural Hg(II)S₃ site towards the N-terminal end of the TRI fold (Figure 3a) and a Zn(II)N₃O catalytic site at the C-terminal end (Figure 3b).^[27] This construct is an artificial metalloenzyme that catalyzes the hydration of CO₂ with an efficiency faster than any other small molecule model and is within ~500-fold of CAII, the most active isoform of carbonic anhydrase. Further, the CuNiR models of Tegoni *et al.*^[29] and Yu *et al.*^[31] are capable of multiple turnover catalysis for the one electron reduction of nitrite using ascorbate, and is the first mononuclear redox enzyme that was isolated via *de novo* protein design. Nevertheless, the metal-binding sites of native proteins, such as the ET site in cupredoxins, are often asymmetric and contain a mixed-ligand (O, N, or S) environment (Figure 3c).^[13] Further, the coordination environment of metal centers contains secondary residues that participate in hydrogen-bonding networks (Figure 3c), such as in CA,^[12,39] or in electrostatic interactions, which are essential in the catalytic or redox activity of many metalloproteins. Even though we have successfully modeled a variety of metal centers and performed catalytic reactions in our 3SCC scaffolds, its self-assembling nature makes it challenging to obtain asymmetric constructs. Therefore, to achieve asymmetric metal binding sites, we have expanded our work to acquire, and then later develop, a single polypeptide sequence that folds into a preformed three-helix bundle fold.

2.2. From Three-stranded Coiled-coil to a Three-helix Bundle Fold

The THB fold is used as a molecular recognition domain found in many biological systems including immunoglobulin G, DNA binding proteins, and various enzymes. Inspired by its universal presence and diverse function in nature, DeGrado and coworkers aimed to create a *de novo* designed antiparallel THB scaffold. Bryson *et al.* used the sequence of CS as the foundation, because its X-ray crystal structure^[40] was observed to pack in an antiparallel manner, where the helices orient in an up-up-down manner, instead of the predicted parallel style.^[41] Using a hierarchical approach, α₃D^[9] was the final product and was isolated through a step-wise process (3 design rounds) of modifying helix-capping interactions which dictate the topology of the bundle, electrostatic interactions which orient the desired helix-helix pairing to avoid alternative states, and hydrophobic interactions to achieve a well-packed core.

In the first round, Bryson *et al.* shortened the sequence of CS by one heptad repeat to yield three 21-residue helices in α_3A (Table 2).^[41] To achieve an antiparallel strand, Glu and Lys residues in helix 2 at the “e” and “g” positions were reversed from the original positions in CS. Next, two simple loops of Gly-Asn and Pro-Gly-Asn were incorporated between helix 1 and 2 and helix 2 and 3, respectively, to serve as hairpin loops. This motif is essential for directing the topology of the bundle, which can adopt clockwise or a counterclockwise orientation. α_3A was observed to form monomer/dimer/trimer species in solution, indicating that the hairpin loops were not successful in stabilizing intermolecular interactions. Round two designs were a direct response to these issues. Sequence α_3B was designed to contain stronger helix stop signals to accurately direct the formation and conformation of the loops. This was achieved by lengthening residues in the loops and adding Asn residues as a helix stop signal in the form of helix capping boxes (Asn-Pro-Asp-Glu between helix 1 and 2 and Asn-Pro-Glu between helix 2 and 3). α_3B is monomeric in solution but still retains some characteristics of a molten globule, an undefined folded state with several energetically equal conformations.

Lastly, the final round of design focused on repacking the hydrophobic core, reordering the residues involved in interhelical electrostatic interactions, and further enhancing helix-capping interactions to prevent nonnative characteristics in α_3B . First, an Asn was replaced with a Ser as the helix-capping residue between helix 1 and 2. Next, the positions of the Lys and Glu residues were redesigned to force a counterclockwise topology in the bundle (Figure 4). The clockwise form was destabilized by careful placement of charged residues at the “e” and “g” positions. In helix 1, only positively (+) charged residues were placed both at the “e” and “g” positions; while in helix 3, negatively (–) charged residues were assigned in those corresponding sites. For helix 2, the “e” sites were given only – charged residues, whereas the “g” sites received + charged residues. Furthermore, some Lys residues were changed to Arg to reduce redundancy in the sequence and provide added stability gained from an Arg-Glu salt-bridge interaction. Lastly, using a genetic repacking algorithm, the hydrophobic residues at the “a” and “d” positions were altered to include various nonpolar residues such as Ala, Val, Ile, Leu, and Phe residues. The native-like property of α_3C was characterized in tandem with α_3B , and it was demonstrated to exhibit thermodynamic and spectroscopic properties of a well-defined and folded protein.

2.3 Solution Structure of α_3D , the Final Iteration in the THB Design

DeGrado’s and coworkers’ work had a significant impact on the field of *de novo* protein design through the design, preparation, and characterization of α_3D , a 73-residue peptide with a single conformation in solution and a unique native-like fold.^[9] Its well-packed core and single topology led to a solution structure; at the time, this was a very challenging feat to achieve (Figure 5). In contrast to its predecessors (α_3A , α_3B , and α_3C), which were chemically synthesized, Walsh *et al.* expressed α_3D in *E. coli*. Met1, Gly2, Gln9, Thr16, and Asp65 were changed in α_3C to generate the sequence of α_3D (Table 2). Thermodynamic studies (chemical and thermal denaturation) showed a scaffold that is fully folded at room temperature (pH 3–7), with a melting temperature (T_m) in the range 80–95°C, heat capacity (C_p) of 10–12 calmol^{–1}K^{–1} per residue, Gibbs free energy of unfolding (ΔG_U) of 5.1 kcalmol^{–1}, and an enthalpy (ΔH_{DSC}) value of –44 kcalmol^{–1}.

The solution structure of $\alpha_3\text{D}$ was obtained from several three-dimensional (triple resonance, TOCSY, and NOESY) NMR experiments at pH 5.5. About 1260 experimental restraints were used in solving the structure, which comprises of 1191 distances and 69 dihedral angles (ϕ and χ_1). The ensemble of 13 structures demonstrate a high quality structure, with RMSD values of 1.06, 0.75, and 1.61 Å for the backbone atoms (residues 1–73, N, C $^\alpha$, C), the backbone atoms in the structure regions (residues 4–21, 24–45, 51–70, N, C $^\alpha$, C), and the heavy atoms (residues 1–73), respectively. The helical bundle adopts a counterclockwise topology, which is confirmed by the interhelical tilt angles of the lowest energy structure ($\Omega_{1,2} = -165^\circ$, $\Omega_{1,3} = 17^\circ$ and $\Omega_{2,3} = -171^\circ$). Further, the χ_1 torsional angles of 14 core residues assumed a single conformation in the ensemble, demonstrating a well-packed apolar core. Ultimately, the success of $\alpha_3\text{D}$ demonstrates that the *de novo* design strategy could now serve as a practical method in constructing complex and multi-stranded scaffolds from a single sequence.

Furthermore, additional solution studies revealed that the backbone ^{15}N and ^{13}C atoms are well-ordered, with restrictive motion on the pico- to nanosecond scale,^[42] relative hydration of the backbone amides,^[43] and a folding time scale for $\alpha_3\text{D}$ in the 1–5 μs range.^[44] In addition, mutation studies showed that replacing Ala60 with a Leu or an Ile resulted in 1.5 kcalmol $^{-1}$ net gain in stability.^[45]

3. Functionalizing $\alpha_3\text{D}$ Framework

3.1. Construction and Structure of a Symmetric Heavy Metal Binding Peptide

$\alpha_3\text{D}$ offers a novel opportunity to add function to a well-defined *de novo* designed scaffold (Figure 6a). We redesigned the sequence of $\alpha_3\text{D}$ by introducing a tris(cysteine) motif to emulate the type of MS $_3$ environments that have been proposed for the metalloregulatory proteins MerR,^[46–48] ArsR/SmtB,^[49] and CadC/CmtR^[49–51] (an MS $_4$ or MS $_3\text{O}$ environment). At the C-terminal end of the bundle, three Leu residues at positions 18, 28, and 67 are inside a “hydrophobic box” which is formed by Ile14, Leu21, Phe31, Ile63, and Tyr70. Chakraborty *et al.* functionalized $\alpha_3\text{D}$ by mutating the “a” site Leu residues to Cys (Leu18Cys, Leu28Cys, and Leu67Cys) to produce $\alpha_3\text{DIV}$ (Figure 6b).^[10] This forms a metal-binding site with two “a” S γ ligands and one pseudo “a” site at the 28th position in the antiparallel strand.

We recently solved the apo structure of $\alpha_3\text{DIV}$ at pH 7.0 (unpublished work) to investigate the effects of a Leu to Cys mutation and the preorganization of the tris(cysteine) site prior to metal binding. Like the parent structure $\alpha_3\text{D}$, the three-helix bundle of $\alpha_3\text{DIV}$ adopts a counterclockwise topology with similar interhelical tilt angles. The structure shows that the tris(cysteine) metal-binding site can easily accommodate heavy metals like Cd(II), Hg(II), and Pb(II), with the 20 lowest energy conformations of the three Cys residues showing great uniformity, indicating a rigid structure. The structure of $\alpha_3\text{DIV}$ illustrates that we were successful in carving out a metal-binding site within $\alpha_3\text{D}$ without significantly compromising its overall structure or stability.

3.2. Characterizing Heavy Metal Binding Properties of α_3 DIV

Chakraborty *et al.* determined, via circular dichroism studies, that apo α_3 DIV is well-folded in solution in the range pH 6–9 and has a chemically-induced G_U of 2.5 kcalmol⁻¹.^[10] This

G_U is half of the reported value for α_3 D, exhibiting a loss in stability after the removal of the packing Leu residues. α_3 DIV stoichiometrically binds Hg(II), Pb(II), and Cd(II) in a pH-dependent manner. The tris(cysteine) site forms a linear [Hg(II)S₂(SH)] complex below pH 6.0 and a trigonal [Hg(II)S₃]⁻ complex above pH 8.5; a mixture of both species were observed under intermediate pH conditions (~pH 7.5). Above pH 5.0, Pb(II) and Cd(II) bound α_3 DIV to generate trigonal pyramidal [Pb(II)S₃]⁻ and pseudo-tetrahedral [Cd(II)S₃(N/O)]⁻ geometry, respectively. These coordination modes were determined using various spectroscopic methods and were compared with the physical properties of 3SCC analogues (Table 3). The absorption features of all three metallated species were characterized via UV/Vis spectroscopy. ¹¹³Cd and ¹⁹⁹Hg NMR, and ¹¹¹mCd and ¹⁹⁹mHg perturbed angular correlation spectroscopy (PAC) spectra were obtained for Cd(II)- and Hg(II)- α_3 DIV to confirm their binding modes in solution. These NMR^[52,53] techniques allow us to study the coordination environment at the millisecond timescale, while the PAC^[54] techniques can further confirm and elucidate speciation behavior at the micro- to nanosecond timescale. The chemical shift environments of ¹¹³Cd and ¹⁹⁹Hg NMR are especially sensitive to the coordination environment. The combination of metal NMR and PAC provides a powerful tool in identifying the primary ligand environment, as well additional coordinating ligands, such as solvent molecules or residues that are several layers removed. Furthermore, we have significantly advanced the development of ²⁰⁷Pb NMR,^[33,38] using our TRI and CS peptides, by proving how extremely sensitive this nucleus is to subtle changes in the apolar layer above the Pb(II)S₃ complex, with a chemical shift range of 5800–5500 ppm.

The ligand-to-metal charge transfer (LMCT) bands of metallated α_3 DIV^[10] in the UV/Vis studies exhibited metal-thiolate transitions comparable with the reported values for its TRI counterparts (Table 3). The Cd(II)- α_3 DIV species has a λ_{max} at 232 nm (18, 200 M⁻¹cm⁻¹) and Pb(II)- α_3 DIV exhibits four absorption bands, with a λ_{max} at 236 nm (18,000 M⁻¹cm⁻¹) and a characteristic band at 346 nm (3150 M⁻¹cm⁻¹). The trigonal complex that forms in Hg(II)- α_3 DIV demonstrates three bands, with a λ_{max} at 247 nm (12,500 M⁻¹cm⁻¹), while the linear species contains one λ_{max} at 240 nm (850 M⁻¹cm⁻¹). From further UV/Vis work, the [Pb(II)S₃]⁻ complex was determined to have a lower limit binding constant of 2.0×10⁷ M⁻¹, while the [Cd(II)S₃(N/O)]⁻ complex had a lower limit value of 3.1×10⁷ M⁻¹.

The linear complex of ¹⁹⁹Hg(II)- α_3 DIV has a chemical shift of -938 ppm, whereas the trigonal form experiences a downfield shift to -244 ppm. The ¹⁹⁹Hg NMR spectrum, at an intermediate pH (7.5), contains both the linear and trigonal planar species. The ¹⁹⁹mHg- α_3 DIV spectra from the PAC analysis confirm this pH-dependent speciation behavior (Figure 7). The PAC parameter, ν_Q , was determined to be 1.48(2) and 1.11(2) GHz for the linear and trigonal complex, respectively, which matches well with the reported ν_Q values for 2- and 3-coordinate Hg(II) in 3SCC constructs.^[31] These ν_Q values were also observed at pH 7.5, demonstrating the presence of both linear and trigonal planar species. In hindsight, these results show that we have created a peptide scaffold that can finely control

the coordination environment of Hg(II) ions. Additionally, Chakraborty *et al.* observed that the addition of 2 equivalents of Hg(II) induced the formation of a dimer species, where Hg(II) atoms bridge two peptides through a linear complex.^[55] This dimer species was confirmed with both $^{199}\text{Hg(II)}$ NMR and $^{199\text{m}}\text{Hg(II)}$ PAC.

The ^{113}Cd NMR spectrum of ^{113}Cd - $\alpha_3\text{DIV}$ shows overlapping resonance peaks at 583 and 595 ppm. Based on these chemical shift positions, this result indicates the presence of two 4-coordinate Cd(II) species. $^{111\text{m}}\text{Cd}$ PAC was again collected to supplement the ^{113}Cd NMR result, but in this case, it was used to define the two species observed in the NMR time scale. The $^{111\text{m}}\text{Cd}$ PAC showed three nuclear quadrupole interactions at 0.35, 0.27, and 0.17 rad ns⁻¹. The peaks at 0.35 and 0.27 rad ns⁻¹ agree well with a CdS₃O complex with two conformations, *endo* and *exo*. The frequency value at 0.17 rad ns⁻¹ fitted well with CdS₃N species, where the N ligand was proposed to originate from the imidazole ring of His72. Overall, not only did the combination of metal NMR and PAC provide a way to accurately characterize our *de novo* designed peptides, it also gave powerful insights into how toxic heavy metals may interact with native proteins.

3.3. Constructing a Symmetric Metalloenzyme Site in $\alpha_3\text{D}$

Our work with $\alpha_3\text{D}$ was further expanded to incorporate a tris(histidine) metal-binding site reminiscent of carbonic anhydrase (Figure 6c). This metalloenzyme plays a vital role in respiration, vision, cancer metastasis, regulation of acid-base equilibria, and other processes in animals, plants, and bacteria. Human carbonic anhydrase II is one of the most efficient enzymes (approaching the diffusion limit) which catalyzes the reversible interconversion between CO₂ and HCO₃⁻.^[56] Even though the mechanism, structure, and inhibition have been previously studied, *de novo* protein design still offers a novel approach to study and replicate an important function of a native metalloenzyme in a simplified peptide system. We have previously demonstrated a carbonic anhydrase model in a bimetallic 3SCC construct [Hg(II)]₅[Zn(II)(H₂O/OH⁻)]_N(TRIL9CL23H)₃⁺⁺ and Zastrow *et al.* reported this model to be within 500-fold of the fastest isozyme (CAII), which is the fastest CA model to date (Figure 2a).^[27,30] Nevertheless, CAII contains residues that participate in hydrogen-bonding networks, and the self-associating nature of our 3SCC construct limits its use in preparing asymmetric sites. Thus, we remodeled the zinc catalytic site of CA into $\alpha_3\text{D}$ to recapitulate CA activity. Now that this is established, we hope to add hydrogen-bonding residues in the second coordination sphere of the Zn(II) complex in future CA designs of $\alpha_3\text{D}$.

Cangelosi *et al.* incorporated a tris(histidine) site in $\alpha_3\text{D}$ to yield $\alpha_3\text{DH}_3$ (Figure 6c), a *de novo* designed metalloenzyme model that exhibited CA activity.^[11] The sequence of $\alpha_3\text{DH}_3$ (Table 2) was expanded by four residues (77 in $\alpha_3\text{DH}_3$), which led to a peptide with increased yields during expression, from ~100 to 230 mgL⁻¹. His residues were substituted at positions 18, 28, and 67, and a His72Val mutation was also incorporated to eliminate a competing ligand. At pH 9.0, the apo form folds well in solution (82% folded), according to the 208 and 222 nm bands that were observed in the CD spectrum, and has a chemically induced G_U of 3.1 kcalmol⁻¹. The Zn(II)- $\alpha_3\text{DH}_3$ complex was characterized using UV/Vis and X-ray absorption spectroscopies, while its CA activity was determined with Khalifah's stopped-flow indicator technique.^[57]

Using a UV/Vis Zincon colorimetric assay,^[58] the apparent Zn(II) binding constant to $\alpha_3\text{DH}_3$ was determined to be 150–40 nM at pH 7.5 and it strengthened to 59–9 nM at pH 9.0. When compared with CA, these affinity values are only two orders of magnitude weaker than the recent value determined for CAII (0.45 nM),^[56] and stronger than the $[\text{Hg}(\text{II})]_S[\text{Zn}(\text{II})(\text{H}_2\text{O}/\text{OH}^-)]_N(\text{TRIL9CL23H})_3^{++}$ (0.8–0.1 μM at pH 7.5 and 0.22–0.06 μM at pH 9.0)^[30] CA model. From extended X-ray absorption fine structure spectroscopy, the Zn(II)- $\alpha_3\text{DH}_3$ coordination environment (at pH 9.0) fitted well to a site that contains 1 oxygen (from an exogenous H_2O or OH molecule) and 3 nitrogen atoms from each His residue bound to a Zn(II) atom at 1.90 and 1.99 Å, which matches well with the Zn complex in CAII with a Zn–N/O bond length of 1.98 Å (pH 7.0).^[56] To demonstrate the success of $\alpha_3\text{DH}_3$ as a metalloenzyme, Cangelosi *et al.* performed a stopped-flow CO_2 hydration assay using Khalifah's indicator technique. The maximal catalytic efficiency (k_{cat}/K_M), which was derived from the k_{cat}/K_M values for pH 8–9.5, and the kinetic $\text{p}K_a$ for the deprotonation of Zn(II)-bound water to yield the active $[\text{Zn}(\text{II})\text{N}_3\text{O}]^{1+}$ hydroxide complex were calculated to be $6.9 \times 10^4 \text{ L mol}^{-1} \text{ s}^{-1}$ and 9.4, respectively. When compared with two small molecule models, Zn(II)(tris(4,5-di*n*-propyl-2-imidazolyl)-phosphine)^[59] and Zn(II)nitrilotris(2-benzimidazolylmethyl-6-sulfonate),^[60] which both have a Zn(II) N_3O complex and show CA activity, Zn(II)- $\alpha_3\text{DH}_3$ significantly outperformed these models, exhibiting a second-order rate constant (k_2) that was 14-fold higher. The CO_2 hydration efficiency of Zn(II)- $\alpha_3\text{DH}_3$ is 2.6-fold slower than its 3SCC counterpart,^[27] 1400-fold less efficient than CAII,^[61] but only 11-fold slower than CAIII.^[62] The decrease in the catalytic activity, as compared with our 3SCC CA model, could be as a result of a weaker dipole, less-symmetric environment for the imidazole rings, and difference in the electrostatics at the metal-binding site in the antiparallel bundle of $\alpha_3\text{DH}_3$. Furthermore, a product inhibition assay was performed on Zn(II)- $\alpha_3\text{DH}_3$, using acetate, since it serves as a more probable mimic of bicarbonate. At pH 8.5, the first-order rate constant (k_{cat}) experienced a modest decrease from 82–6 to 66–4 s^{-1} . This inhibition result indicates no significant loss in catalytic activity, and illustrates a CA model capable of preventing product inhibition, which is a major problem in small molecule models of enzymes. Overall, our work on Zn(II)- $\alpha_3\text{DH}_3$ exhibits that we are successful in recapitulating the primary active site and the function of carbonic anhydrase in a simplified antiparallel THB scaffold, a metalloenzyme that is found in a twisted β -sheet fold in nature.

4. Future Designs

4.1. Redesign of $\alpha_3\text{DIV}$ to Improve Metal Binding

The structure of $\alpha_3\text{DIV}$ demonstrates that we are still able to retain a well-folded peptide construct, even after mutating the core packing Leu to Cys residues to achieve a thiol-rich metal-binding site. With this knowledge, we can now attempt to prepare an antiparallel three-helix bundle construct (via modeling) that will contain a more preformed or more symmetric tris(cysteine) metal-binding site which resembles “a” site ligands. When overlaying the 20 lowest structures of $\alpha_3\text{DIV}$, we observed that the $\text{S}\gamma$ of Cys28 consistently points towards the C-terminal end, forming a skewed S3 plane. Thus, the first iteration involves rearranging the Cys28 by one layer towards the N-terminal end. This modification requires Cys28Phe and Phe31Cys mutations (Figure 8a). Along with Cys18 and Cys67, a Cys in the 31st position could produce an S3 plane which is more

perpendicular to the helical bundle than the current $\alpha_3\text{DIV}$ construct. The Phe in the 28th position could provide a hydrophobic capping interaction that forces Cys31 to take an “a” site orientation. We are currently working on these modifications and plan to characterize the heavy metal binding properties of this new $\alpha_3\text{D}$ construct fully.

4.2. Redesign of $\alpha_3\text{DH}_3$ to Improve Catalytic Activity

The crystal structure of Human CAII revealed that Zn(II) is coordinated to three His residues and a water molecule/hydroxide ion.^[12,39] In addition, this structure exposed an elaborate H-bonding network, facilitated by Thr, Asn, and Gln residues in the second coordination sphere. Thr199 forms an H-bonding interaction with the hydroxide nucleophile of the Zn(II) center, which was determined to be significant for the acid-base catalysis activity of CA. Mutation studies of this residue resulted in a 1.5–2.5 unit increase in the pK_a (from 6.8) of the H_2O molecule to the OH^- active form, resulting in 100-fold loss in catalytic activity.^[12,39,63] Our current CA constructs Zn(II)- $\alpha_3\text{DH}_3$ and $[\text{Hg(II)}]_5[\text{Zn(II)}(\text{H}_2\text{O}/\text{OH}^-)]_N(\text{TRIL9CL23H})_3^{n+}$ have pK_a values for this deprotonation process 2–2.6 units higher, and demonstrate CO_2 hydration activities that are 500–1400-fold less efficient than CAII.^[11,27,30,61] Nevertheless, both constructs only contain the primary coordination sphere (ZnHis_3) and do not contain H-bonding interactions that greatly enhance CA activity. Consequently, to achieve native-like efficiency, we now have to consider incorporating secondary interactions in our designs. Since our 3SCC analogues are not ideal for adding H-bonding residues inside the core in a step-wise manner, we must turn to our $\alpha_3\text{DH}_3$ to design future CA models which investigate essential secondary interactions (as demonstrated in Figure 8b).

4.3. Design of an Electron-transfer Site in $\alpha_3\text{D}$

Cupredoxins are copper-containing proteins that shuttle electrons between two-membrane-bound proteins in photosynthesis, and for this reason, have been extensively studied to understand their fundamental function in nature.^[64] The copper ion is coordinated to a well-defined and pre-organized asymmetric metal-binding site encapsulated by a β -barrel fold. It contains two histidines, one cysteine, and one or two more weakly bound residues to form a 3- to 5-coordinate copper complex. The rack-induced bonding model has been used to describe the copper site in these cupredoxins, which states that the ligand environment in a β -barrel fold forces the copper ion, regardless of the oxidation state, into a strict geometry.^[65] This feature alone makes cupredoxins an ideal candidate for *de novo* design.

The most successful *de novo* designed model that demonstrates the spectroscopic properties of cupredoxins was achieved in a parallel four-stranded coiled-coil.^[66] Therefore, $\alpha_3\text{D}$ offers a viable framework in achieving an asymmetric cupredoxin site. Now knowing that we can incorporate cysteine ligands in $\alpha_3\text{DIV}$ and bulky His residues in $\alpha_3\text{DH}_3$, we have designed several constructs (Figure 8c) to test the importance of the β -barrel fold in copper ET proteins (unpublished work). Ultimately, these designed constructs will provide a novel approach for investigating biological long-range electron transfer.

5. Summary and Outlook

De novo protein design is a biologically relevant approach used to understand the structure-function relationship in the metal centers of native proteins, and is quickly becoming a viable method for developing heretofore unknown protein folds with nature-inspired metal-binding sites. As we have discussed herein, the field of *de novo* metalloproteins and enzymology has made great advances over the past decade and promises the future capability of preparing multicenter redox enzymes designed from first principles. At present, well-defined and well-folded scaffolds have been made available for helical bundle motifs,^[67–70] with the α_3D system being an excellent example. Here, we have demonstrated that it is possible to functionalize the α_3D scaffold to generate structural and catalytic sites, while studies are ongoing to generate electron-transfer centers within the same fold. We and others intend to build upon this work, constructing nature-inspired asymmetric metal-binding sites that form the foundation for future designs of artificial metalloenzymes and metalloproteins, which not only mimic natural reactions, but also confer new catalytic processes, ranging from CO_2 and H^+ reduction to water oxidation and fine chemical synthesis.

Acknowledgments

J. S. P. and V. L. P. would like to thank the National Institutes of Health (NIH) for financial support (ES012236).

References

1. Waldron KJ, Rutherford JC, Ford D, Robinson NJ. *Nature*. 2009; 460:823–830. [PubMed: 19675642]
2. Lu Y, Berry SM, Pfister TD. *Chem Rev*. 2001; 101:3047–3080. [PubMed: 11710062]
3. Lu Y, Yeung N, Sieracki N, Marshall NM. *Nature*. 2009; 460:855–862. [PubMed: 19675646]
4. Yu F, Cangelosi VM, Zastrow ML, Tegoni M, Plegaria JS, Tebo AG, Mocny CS, Ruckthong L, Qayyum H, Pecoraro VL. *Chem Rev*. 2014; 114:3495–3578. [PubMed: 24661096]
5. Zastrow ML, Pecoraro VL. *Coord Chem Rev*. 2013; 257:2565–2588. [PubMed: 23997273]
6. DeGrado WF, Summa CM, Pavone V, Nastri F, Lombardi A. *Annu Rev Biochem*. 1999; 68:779–819. [PubMed: 10872466]
7. Ghosh D, Pecoraro VL. *Curr Opin Chem Biol*. 2005; 9:97–103. [PubMed: 15811792]
8. Giedroc DP, Arunkumar AI. *Dalton Trans*. 2007:3107–3120. [PubMed: 17637984]
9. Walsh STR, Cheng H, Bryson JW, Roder H, De-Grado WF. *Proc Natl Acad Sci USA*. 1999; 96:5486–5491. [PubMed: 10318910]
10. Chakraborty S, Yudenfreund-Kravitz J, Thulstrup PW, Hemmingsen L, DeGrado WF, Pecoraro VL. *Angew Chem Int Ed*. 2011; 50:2049–2053.
11. Cangelosi VM, Deb A, Penner-Hahn JE, Pecoraro VL. *Angew Chem Int Ed*. 2014; 53:7900–7903.
12. Krebs JF, Ippolito JA, Christianson DW, Fierke CA. *J Biol Chem*. 1993; 268:27458–27466. [PubMed: 8262987]
13. Guss JM, Bartunik HD, Freeman HC. *Acta Crystallogr Sect B: Struct Sci*. 1992; 48:790–811.
14. Peacock AFA, Iranzo O, Pecoraro VL. *Dalton Trans*. 2009:2271–2280. [PubMed: 19290357]
15. Ghosh D, Pecoraro VL. *Inorg Chem*. 2004; 43:7902–7915. [PubMed: 15578824]
16. Dieckmann GR, McRorie DK, Tierney DL, Utschig LM, Singer CP, O'Halloran TV, Penner-Hahn JE, DeGrado WF, Pecoraro VL. *J Am Chem Soc*. 1997; 119:6195–6196.
17. Dieckmann GR, McRorie DK, Lear JD, Sharp KA, DeGrado WF, Pecoraro VL. *J Mol Biol*. 1998; 280:897–912. [PubMed: 9671558]

18. Farrer BT, Harris NP, Balchus KE, Pecoraro VL. *Biochemistry*. 2001; 40:14696–14705. [PubMed: 11724584]
19. Matzapetakis M, Farrer BT, Weng T-C, Hemmingsen L, Penner-Hahn JE, Pecoraro VL. *J Am Chem Soc*. 2002; 124:8042–8054. [PubMed: 12095348]
20. Lee KH, Matzapetakis M, Mitra S, Marsh EN, Pecoraro VL. *J Am Chem Soc*. 2004; 126:9178–9179. [PubMed: 15281796]
21. Matzapetakis M, Ghosh D, Weng TC, Penner-Hahn JE, Pecoraro VL. *J Biol Inorg Chem*. 2006; 11:876–890. [PubMed: 16855818]
22. Iranzo O, Ghosh D, Pecoraro VL. *Inorg Chem*. 2006; 45:9959–9973. [PubMed: 17140192]
23. Iranzo O, Thulstrup PW, Ryu S, Hemmingsen L, Pecoraro VL. *Chem Eur J*. 2007; 13:9178–9190. [PubMed: 17960740]
24. Iranzo O, Cabello C, Pecoraro VL. *Angew Chem Int Ed*. 2007; 46:6688–6691.
25. Łuczowski M, Stachura M, Schirf V, Demeler B, Hemmingsen L, Pecoraro VL. *Inorg Chem*. 2008; 47:10875–88. [PubMed: 18959366]
26. Peacock AFA, Hemmingsen L, Pecoraro VL. *Proc Natl Acad Sci USA*. 2008; 105:16566–16571. [PubMed: 18940928]
27. Zastrow ML, Peacock AFA, Stuckey JA, Pecoraro VL. *Nat Chem*. 2011; 4:118–123. [PubMed: 22270627]
28. Zampella G, Neupane KP, De Gioia L, Pecoraro VL. *Chem Eur J*. 2012; 18:2040–2050. [PubMed: 22231489]
29. Tegoni M, Yu F, Bersellini M, Penner-Hahn JE, Pecoraro VL. *Proc Natl Acad Sci USA*. 2012; 109:21234–21239. [PubMed: 23236170]
30. Zastrow ML, Pecoraro VL. *J Am Chem Soc*. 2013; 135:5895–5903. [PubMed: 23516959]
31. Yu F, Penner-Hahn JE, Pecoraro VL. *J Am Chem Soc*. 2013; 135:18096–18107. [PubMed: 24182361]
32. Iranzo O, Chakraborty S, Hemmingsen L, Pecoraro VL. *J Am Chem Soc*. 2011; 133:239–251. [PubMed: 21162521]
33. Neupane KP, Pecoraro VL. *Angew Chem Int Ed*. 2010; 49:8177–8180.
34. Farrer BT, Pecoraro VL. *Proc Natl Acad Sci USA*. 2003; 100:3760–3765. [PubMed: 12552128]
35. Touw DS, Nordman CE, Stuckey JA, Pecoraro VL. *Proc Natl Acad Sci USA*. 2007; 104:11969–11974. [PubMed: 17609383]
36. Peacock AFA, Stuckey JA, Pecoraro VL. *Angew Chem Int Ed*. 2009; 48:7371–7374.
37. Chakraborty S, Touw DS, Peacock AFA, Stuckey JA, Pecoraro VL. *J Am Chem Soc*. 2010; 132:13240–13250. [PubMed: 20825181]
38. Neupane KP, Pecoraro VL. *J Inorg Biochem*. 2011; 105:1030–1034. [PubMed: 21625408]
39. Avvaru BS, Kim CU, Sippel KH, Gruner SM, Agbandje-McKenna M, Silverman DN, McKenna R. *Biochemistry*. 2010; 49:249–251. [PubMed: 20000378]
40. Lovejoy B, Choe S, Cascio D, McRorie DK, De-Grado WF, Eisenberg D. *Science*. 1993; 259:1288–1293. [PubMed: 8446897]
41. Bryson JW, Desjarlais JR, Handel TM, DeGrado WF. *Protein Sci*. 1998; 7:1404–1414. [PubMed: 9655345]
42. Walsh STR, Lee AL, DeGrado WF, Wand AJ. *Biochemistry*. 2001; 40:9560–9569. [PubMed: 11583155]
43. Walsh STR, Cheng RP, Wright WW, Alonso DOV, Daggett V, Vanderkooi JM, DeGrado WF. *Protein Sci*. 2003; 12:520–531. [PubMed: 12592022]
44. Zhu Y, Alonso DOV, Maki K, Huang C-Y, Lahr SJ, Daggett V, Roder H, DeGrado WF, Gai F. *Proc Natl Acad Sci USA*. 2003; 100:15486–15491. [PubMed: 14671331]
45. Walsh STR, Sukharev VI, Betz SF, Vekshin NL, DeGrado WF. *J Mol Biol*. 2001; 305:361–373. [PubMed: 11124911]
46. O'Halloran TV, Walsh C. *Science*. 1987; 235:211–214. [PubMed: 3798107]
47. Wright JG, Natan MJ, MacDonnel FM, Ralston DM, O'Halloran TV. *Prog Inorg Chem*. 1990; 38:323–412.

48. Wright JG, Tsang HT, Penner-Hahn JE, O'Halloran TV. *J Am Chem Soc.* 1990; 112:2434–2435.
49. Busenlehner LS, Weng TC, Penner-Hahn JE, Giedroc DP. *J Mol Biol.* 2002; 319:685–701. [PubMed: 12054863]
50. Ye J, Kandegedara A, Martin P, Rosen BP. *J Bacteriol.* 2005; 187:4214–4221. [PubMed: 15937183]
51. Banci L, Bertini I, Cantini F, Ciofi-Baffoni S, Cavet JS, Dennison C, Graham AI, Harvie DR, Robinson NJ. *J Biol Chem.* 2007; 282:30181–30188. [PubMed: 17599915]
52. Öz G, Pountney DL, Armitage IM. *Biochem Cell Biol.* 1998; 76:223–234. [PubMed: 9923691]
53. Armitage I, Drakenberg T, Reilly B. *Met Ions Life Sci.* 2013; 11:117–144. [PubMed: 23430773]
54. Hemmingsen L, Stachura M, Thulstrup PW, Christensen NJ, Johnston K. *Hyperfine Interact.* 2010; 197:255–267.
55. Chakraborty, S. PhD thesis. University of Michigan (USA); 2011.
56. Song H, Wilson DL, Farquhar ER, Lewis EA, Emerson JP. *Inorg Chem.* 2012; 51:11098–11105. [PubMed: 23030313]
57. Khalifah RG. *J Biol Chem.* 1971; 246:2561–2573. [PubMed: 4994926]
58. Rush RM, Yoe JH. *Anal Chem.* 1954; 26:1345–1347.
59. Slebocka-Tilk H, Cocho JL, Frakman Z, Brown RS. *J Am Chem Soc.* 1984; 106:2421–2431.
60. Nakata K, Shimomura N, Shiina N, Izumi M, Ichikawa K, Shiro M. *J Inorg Biochem.* 2002; 89:255–266. [PubMed: 12062130]
61. Jackman JE, Merz KM Jr, Fierke CA. *Biochemistry.* 1996; 35:16421–16428. [PubMed: 8987973]
62. Jewell DA, Tu C, Paranawithana SR, Tanhauser SM, LoGrasso PV, Laipis PJ, Silverman DN. *Biochemistry.* 1991; 30:1484–1490. [PubMed: 1899618]
63. Liang Z, Xue Y, Behravan G, Jonsson B-H, Lindskog S. *Eur J Biochem.* 1993; 211:821–827. [PubMed: 8436138]
64. Solomon EI, Szilagyi RK, DeBeer George S, Basumallick L. *Chem Rev.* 2004; 104:419–458. [PubMed: 14871131]
65. Malmström BG. *FEBS.* 1994; 223:711–718.
66. Shiga D, Nakane D, Inomata T, Funahashi Y, Masuda H, Kikuchi A, Oda M, Noda M, Uchiyama S, Fukui K, Kanaori K, Tajima K, Takano Y, Nakamura H, Tanaka T. *J Am Chem Soc.* 2010; 132:18191–18198. [PubMed: 21126081]
67. Choma CT, Lear JD, Nelson MJ, Dutton PL, Robertson DE, DeGrado WF. *J Am Chem Soc.* 1994; 116:856–865.
68. Lombardi A, Summa CM, Geremia S, Randaccio L, Pavone V, DeGrado WF. *Proc Natl Acad Sci USA.* 2000; 97:6298–6305. [PubMed: 10841536]
69. Zhuang J, Amoroso JH, Kinloch R, Dawson JH, Baldwin MJ, Gibney BR. *Inorg Chem.* 2004; 43:8218–8220. [PubMed: 15606161]
70. Huang SS, Koder RL, Lewis M, Wand AJ, Dutton PL, Gray HB. *Proc Natl Acad Sci USA.* 2004; 101:5536–554. [PubMed: 15056758]

Biographies

Vincent L. Pecoraro is the John T. Groves Collegiate Professor of Chemistry at the University of Michigan. His research group is deeply involved in *de novo* metalloprotein design and metallamacrocyclic chemistry. Prof. Pecoraro has received numerous awards and honors, including the G. D. Searle Biomedical Research Scholarship, an Alfred P. Sloan Fellowship, the Vanadis Award, an Alexander von Humboldt Award for Senior Research Scientists, and The Blaise Pascal International Chair for Research. He is a Fellow of the American Association for the Advancement of Science and the American Chemical Society. He has served as an Associate Editor of Inorganic Chemistry since 1994.

Jefferson S. Plegaria received his B.A. degree in Chemistry from La Salle University, Philadelphia, PA, in 2007. He then worked in the pediatric-oncology lab of Dr. John Maris at The Children's Hospital of Philadelphia. In 2009, he began his Ph.D. degree with Professor Vincent L. Pecoraro. He is currently studying the spectroscopic and electron-transfer properties of *de novo* designed cupredoxin models using biophysical, electrochemical, and photo-physical methods.

Figure 1.

a) Crystal structure of a zinc metalloenzyme, human carbonic anhydrase II (PDB code 3KS3).^[12] b) Crystal structure of a copper electron-transfer metalloprotein, plastocyanin (PDB code 1PLC),^[13] which is part of the cupredoxin family.

Figure 2.

a) Crystal structure of $[\text{Hg}(\text{II})]_{\text{S}}[\text{Zn}(\text{II})(\text{H}_2\text{O}/\text{OH}^-)]_{\text{N}}\text{-(CSL9PenL23H)}_3^{n+}$ (PDB code 3PBJ), which is used as a crystallographic model for the $[\text{Hg}(\text{II})]_{\text{S}}[\text{Zn}(\text{II})(\text{H}_2\text{O}/\text{OH}^-)]_{\text{N}}(\text{TRIL9CL23H})_3^{n+}$ CA model.^[27, 30] This bimetallic 3SCC construct contains a Zn(II) and Hg(II) atom bound to a His- N_3 and Pen- S_3 site, respectively. b) Symmetric “a” site Cys residues in CSL9C (PDB code 3LJM)^[37] contain $\text{S}\gamma$ ligands that orient inside the core. c) Symmetric “d” site Cys residues in CSL19C (PDB code 2X6P)^[37] include $\text{S}\gamma$ ligands that orient towards the interhelical interface, forming a larger metal-binding site than Cys “a” sites.

Figure 3.

a) Top-down view of the trigonal Hg(II)S₃ site in [Hg(II)]_S[Zn(II)(H₂O/OH⁻)]_N(CSL9PenL23H)₃ⁿ⁺, which serves as a structural motif in the scaffold.^[27] b) Top-down view of the Zn(II)N₃O site in [Hg(II)]_S[Zn(II)(H₂O/OH⁻)]_N(CSL9PenL23H)₃ⁿ⁺, which is capable of CA activity.^[27] c) 2His, Cys and Met copper binding site in plastocyanin, illustrating an asymmetric metal center.^[13] d) Zinc metal-binding site in CAII, which contains a tris(histidine) site and hydrogen-bonding residues, such as Thr199, which are essential in catalysis.^[12, 39] Reprinted with permission from Ref. [4]. Copyright 2014 American Chemical Society.

Figure 4.

a) Carton scheme representing interhelical electrostatic interactions between α -helices. The circles symbolize α -helices and are illustrated from either the N- or C-terminal end. The numbers correspond to their sequential positions in the THB. The “-” signs denote negatively charged Glu residues in either the “e” or “g” positions, while “+” signs indicate positively charged Lys and Arg residues at the same positions. Solid and dashed lines represent loops that connect helices. a) CS was designed to form a parallel dimer; instead it was observed to pack into an antiparallel trimer with unfavorable interactions with like charges. b) In α_3A and α_3B , the arrangement of the interhelical electrostatic interactions allow for both topologies to be possible. c) The clockwise form was destabilized by careful placement of charged residues at the “e” and “g” positions. In helix 1, only positively (+) charged residues were placed both at the “e” and “g” positions; while in helix 3, negatively (-) charged residues were assigned in those corresponding sites. For helix 2, the “e” sites were given only - charged residues, whereas the “g” sites received + charged residues. Reprinted with permission from Ref [41]. Copyright 1998 The Protein Society.

Figure 5.
Solution structure of α_3D (PDB code 2A3D) demonstrating a THB fold.^[9]

Figure 6.

a) Schematic representation of designing a metal center in the $\alpha_3\text{D}$ scaffold. b) Top-down view of the tris(cysteine) site in $\alpha_3\text{DIV}$ modeled from the $\alpha_3\text{D}$ structure. c) Top-down view of the tris(histidine) site in $\alpha_3\text{DH}_3$ modeled from the $\alpha_3\text{D}$ structure.

Figure 7.
Schematic representation of the pH- and stoichiometric-dependent behavior of Hg(II) species in α_3 DIV.^[10,55]

Figure 8.

Future α_3 D designs. a) α_3 DIV Cys28Phe/Phe31Cys mutant designed to achieve a more symmetric tris(cysteine) site. b) α_3 DH₃ Phe31Thr construct, where the Thr could serve as a hydrogen-bonding residue and lower the deprotonation pK_a of the water molecule bound to the ZnN₃ center. c) Asymmetric 2His, Cys and Met cupredoxin incorporated in the α_3 D scaffold.

Table 1

Amino acid sequence of *de novo* designed 3SCC analogues.^[a]

Peptide	Sequence	abcdefg	abcdefg	abcdefg	abcdefg	abcdefg	abcdefg
CoilSer	Ac-E	WEALEKK	LAALESK	LQALEKK	LEALEHG		-NH ₂
Baby	Ac-G	LKALEEK	LKALEEK	LKALEEK			G-NH ₂
TRI	Ac-G	LKALEEK	LKALEEK	LKALEEK	LKALEEK		G-NH ₂
Grand	Ac-G	LKALEEK	LKALEEK	LKALEEK	LKALEEK	LKALEEK	G-NH ₂

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

Table 2

Amino acid sequence of THB analogues.^[a]

Peptide	Sequence	Design Purpose/Function
	abcdefg	loop
α_3A	<u>E</u> WEALEKK L <u>N</u> ALESK LQALEK	Iteration of CS
	NWEAL <u>K</u> KE L <u>N</u> A <u>L</u> <u>K</u> SE LQAL <u>K</u> K	<u>PG</u>
	<u>E</u> WEALEKK L <u>N</u> ALESK LQALE <u>H</u> G	
α_3B	EWEALEKK L <u>A</u> ALESK LQALEK	Iteration of α_3A
	<u>N</u> PDEW <u>A</u> ALKKE L <u>A</u> ALESK LQALK	<u>GKG</u>
	<u>N</u> PEWEALEKK L <u>A</u> ALESK LQALEHG	
α_3C	<u>S</u> W <u>A</u> EFKER L <u>A</u> A <u>K</u> SR LQAL	Iteration of α_3B
	<u>S</u> E <u>A</u> ELAAFEKE L <u>A</u> A <u>F</u> ESE LQAYK	GKG
	NPEVEAL <u>R</u> KE <u>A</u> A <u>A</u> IRSE LQAYRH <u>N</u>	
α_3D	<u>M</u> GSWAEFK <u>Q</u> R LAAIK <u>T</u> R LQAL	GGs Molecular
	EAE <u>L</u> AAFEKE IAAFESE LQAY	KGKG recognition domain
	NPEVEALRKE AA <u>A</u> IR <u>D</u> E LQAYRH <u>N</u>	
α_3DIV	MGSWAEFKQR LAAIK <u>T</u> R <u>C</u> QAL	GGs Heavy metal
	EAE <u>C</u> AAFEKE IAAFESE LQAY	KGKG peptide
	NPEVEALRKE AA <u>A</u> IR <u>D</u> E <u>C</u> QAYRH <u>N</u>	Cd(II), Hg(II) and Pb(II)
α_3DH_3	MGSWAEFKQR LAAIK <u>T</u> R <u>H</u> QAL	GGs Carbonic anhydrase
	EAE <u>H</u> AAFEKE IAAFESE LQAY	KGKG
	NPEVEALRKE AA <u>A</u> IR <u>D</u> E <u>H</u> QAYRH <u>N</u> <u>SGA</u>	model

^[a]The sequences are prepared in heptads. Residues that are underlined and bolded were changed from the previous design. The α_3A was altered from the CS sequence.

Table 3

Physical parameters of metallated α_3 DIIV compared to 3SCC constructs.^[a]

Complex	λ [nm] (ϵ [M ⁻¹ cm ⁻¹])	δ (ppm) ¹¹³ Cd	¹⁹⁹ Hg	²⁰⁷ Pb	ω_0 (rad/ns), η ^{111m} Cd PAC	ν_Q (GHz), η ^{199m} Hg PAC	ref
Cd(α_3 DIIV)	232 (18 200)	583			0.350(6), 0.00(1)/ <i>b</i> /		[10]
		595			0.268(4), 0.18(7)/ <i>c</i> /		
					0.170(2), 0.50(2)/ <i>d</i> /		
Hg(α_3 DIIV)							[10]
3-coordinate	247 (12 500), 265 (8400), 295 (3900)		-244			1.11(2), 0.40(3)	
2-coordinate	240 (850)		-938			1.48(2), 0.15(5)	
Pb(α_3 DIIV)	236 (18 000), 260 (14 400), 278 (9100), 346 (3150)						[10]
Cd(TRIL12C) ₃	231 (20 600)	619			0.233(8), 0.25(12)		[19]
					0.468(9), 0.12(10)		
		625			0.337(2), 0.23(2)		[19]
					0.438(4), 0.20(3)		
Cd(TRIL16C) ₃	232 (22 600)				1.164(5), 0.25(2)		[23]
			-185		1.558(7), 0.23(1)		[23]
Hg(TRIL9C) ₃							[23, 25]
Hg(TRIL19C) ₃	230 (21 300), 247 (15 000), 297 (5500)		-316				
Hg(TRIL16C) ₃	247 (16 800), 265 (10 600), 295 (5000)		-179				[16]
Hg(TRIL16C) ₂	240 (2700)		-834				[16]
Pb(TRIL12C) ₃	238 (17 000), 278 (12 300), 343 (3700)			5814/ <i>e</i> /			[21, 33]
Pb(TRIL16C) ₃	236 (18 500), 260 (16 500), 278 (14 500), 346 (3400)			5612/ <i>f</i> /			[21, 33]

^[a]_{UV/Vis}, metal (¹¹³Cd, ¹⁹⁹Hg and ²⁰⁷Pb) NMR and (^{111m}Cd and ^{199m}Hg) PAC. L9C and L16C are "a" site constructs, while L12C and L19C are "d" sites.^[b]_{CdS₃O} in *exo* conformation.^[c]_{CdS₃O} in *endo* conformation.^[d]_{CdS₃N} species.^[e]₂₀₇Pb chemical shift for Pb(CSL12C)₃; ref 33.^[f]₂₀₇Pb chemical shift for Pb(CSL16C)₃; ref 33.