Published in final edited form as:

J Biol Inorg Chem. 2018 January ; 23(1): 123–135. doi:10.1007/s00775-017-1515-7.

Incorporation of second coordination sphere D -amino acids alters Cd(II) geometries in designed thiolate-rich proteins

Leela Ruckthong^{1,2}, Aniruddha Deb¹, Lars Hemmingsen³, James E. Penner-Hahn¹, and Vincent L. Pecoraro^{1,2}

¹Department of Chemistry, University of Michigan, Ann Arbor, MI 48109-1055, USA

²Biophysics Program, University of Michigan, Ann Arbor, MI 48109, USA

³Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen Ø, Denmark

Abstract

We use a de Novo protein design strategy to demonstrate that the second coordination sphere of a metal site plays a key role in controlling coordination geometries of Cd(II)-tris-thiolate complexes. Specifically, we show that alteration of chirality within the core hydrophobic packing region of a three-stranded coiled coil (3SCC) can control the coordination number of Cd(II) by limiting steric encumbrance to the metal center. Within a specific class of 3SCCs [Ac-G-(LKALEEK)_{*n*}-G-NH₂], where *n* = 4 is **TRI** and *n* = 5 is **GRAND**, one L-Leu may be substituted by L-Cys to generate a planar tris-thiolate array capable of metal binding. In the native peptide containing only the L-configuration of leucine, the three-Cys ligand site leads to a mixture of 3- and 4-coordinate Cd(II). When the L-Leu above (toward the N-terminus) the tris-Cys site is substituted with D-Leu, solely a 3-coordinate structure [Cd(II)S₃] was obtained. When D-Leu is located below (toward the C-terminus), a mixture of two coordination geometries, presumably Cd(II)S₃O and Cd(II)S₃O₂, is observed, while substitution with D-Leu both above and below the tris-Cys plane yields a higher percentage of 4-coordinate Cd(II)S₃O species. Thus, the use of D-amino acids around a metal's coordination sphere provides a powerful tool for controlling the properties of future designed metalloproteins.

Keywords

Protein engineering; Heavy metal; Metallocenter assembly; Metal sensor

Introduction

As synthetic chemists, it is humbling to realize that nature has been able to construct all life forms, and the diversity of chemical reactions they complete, from essentially 20 canonical

Correspondence to: Vincent L. Pecoraro.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00775-017-1515-7>) contains supplementary material, which is available to authorized users.

Compliance with ethical standards

Conflict of interest The author(s) declare that they have no competing interests.

amino acids. And yet, one can only imagine the multitude of structures and functions available to proteins if a wider array of non-coded amino acids was exploited. While recent advances in molecular biology have allowed for the controlled incorporation of a myriad of different amino acids within proteins using amber codons or post-translational modification [1–3], the most straightforward method of incorporating alternate amino acids in protein engineering is through solid-phase synthesis [4–9]. The introduction of artificial amino acids greatly expands the versatility of the protein sequence pool. Modification of designed proteins with a broader range of amino acids might enhance protein stability, modify catalytic capabilities, or reveal novel physical properties not yet appreciated in native proteins [8, 10–12]. 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 [4, 8, 13]. Indeed, such changes may engender physical properties that could be useful for future biophysical and pharmaceutical applications.

The stereochemistry of D-amino acids is an interesting tool to extend the scope of de Novo protein design, because 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 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 [14, 15]. D-Proline (D-Pro) was used to force prime turns to nucleate antiparallel β -hairpin formation in LD dipeptide sequences [10, 15]. D-Aspartic 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 [16]. Moreover, the stereoisomer of D-amino acids has also been employed to create hetero-chiral fold designs of bracelet and boat-shaped molecules [17, 18], a canoe-like construct as an alkali receptor [19], a pi-cup scaffold as an acetylcholine receptor [20], and the Zn-finger hydrolase ($\alpha\beta\beta$ fold) structure [21]. 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* [15, 22–24]. Another prominent advantage of employing D-amino acids is as a helical terminator as evident in D-Alanine (D-Ala) and D-Leucine (D-Leu) facilitating the Schellman motif to terminate helix formation [25, 26]. 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 our work, no studies have been reported that use D-amino acids to engineer metal environments or to control metal geometries [4, 5, 27].

We have utilized the **TRI** family and **GRAND** peptide sequences (sequences as given in Table 1) to generate 3SCC peptide scaffolds for heavy metal studies [4, 6, 13, 28–43]. These peptides are amphipathic α -helix comprised of four and five heptad repeats [$L_aK_bA_cL_dE_eE_fK_g$], respectively, in which the N- and C-termini are capped with acetylated and amidated glycine residues [44]. Non-polar residues at the *a* and *d* positions make up the hydrophobic core and polar residues at the *e* and *g* positions provide electrostatic

interactions, leading to the assembly of a 3SCC (above pH 5.5) (Fig. 1, top). To generate a heavy metal binding site for Cd(II), Hg(II), As(III), or Pb(II), a Leu residue is substituted at either the *a* or *d* position with a cysteine (Cys) residue, resulting in a tris-thiolate site (Fig. 1, bottom).

The **TRIL16C** peptide, in which a Cys residue is replaced in lieu of Leu at the 16th position, was designed to study Cd(II)-binding interactions. This (**TRIL16C**)₃ peptide was spectroscopically shown to bind a series of metals in a three-coordinate environment [28, 29, 31, 32, 35, 36, 40, 45–47]; however, Cd(II) was found to bind to such peptide in two distinct geometries: Cd(II)S₃ and Cd(II)S₃O [35]. Much effort focused on how to engineer the binding site environment to control specific coordination geometries on Cd(II) in the 3SCCs. The knowledge gained helps clarify Cd(II)–protein interactions that might be found in nature, e.g., CadC and CmtR and to emphasize the strong ability of the “bottom-up approach” to build a small and stable metalloprotein model to understand basic principles of nature at an unprecedented level.

The coordination environment of the designed peptides was probed by the joint use of ¹¹³Cd Nuclear Magnetic Resonance (NMR) [48–51] and ^{111m}Cd Perturbed Angular Correlation (PAC) spectroscopies [52, 53], 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, ¹¹³Cd 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 previously published results of the *a* site-metal binding peptides, 3- and 4-coordinate thiolate sites have ¹¹³Cd NMR chemical shifts that are between 680–700 and 570–600 ppm, respectively. A mixture of 3- and 4-coordinate species falls in 600–680 ppm range [41]. These chemical shift values are now used as a reference for Cd(II) geometry assignment in this system. ^{111m}Cd PAC can provide information about the metal site electronic environment, symmetry, and dynamics. If the ligands bound to the metal exchange faster than the NMR timescale but slower than 0.1–100 ns, signals from the individual sites may be recorded by PAC spectroscopy, while in NMR spectroscopy, a single resonance will be observed due to fast exchange. The PAC spectrum provides information about first coordination sphere ligands (ω_0) and symmetry around the *z*-axis (η) [41, 53]. The ω_0 values gained from the previous research of the *a* site-metal binding peptides fall into the frequency regions; 0.450 rad ns^{−1}, indicating a trigonal planar Cd(II), and 0.350 rad ns^{−1}, indicating a pseudotetrahedral Cd(II) geometry [41].

The main concept to enforce a desired geometry is to engineer the steric environment around the metal pocket. Previous work demonstrated that substitution of L-Leu by L-Ala above (toward the N-terminus), a tris-Cys binding site in **TRIL12AL16C** allowed for a 100% Cd(II)S₃O structure, because space was made available for solvent access to the metal [6, 37]. The isolation of a fully 3-coordinate species was achieved by two alternative designs. One increased the steric bulk of the first sphere coordination ligands around the metal binding site using a non-coded amino acid, L-Penicillamine (Pen), incorporated in the 16th position [5]. Cd(**TRIL16Pen**)₃[−] exhibited a fully 3-coordinate Cd(II) S₃ [5]. The other strategy was to use D-Leu to introduce steric encumbrance above the metal site in

TRIL12_DLL16C design. This approach exploited the alternative chirality which repositioned the side chain closer to the (Cys)₃ metal binding layer. As the chemical properties of L- and D-Leu remain invariant, it was the reorientation conferred by D-configuration in the 12th position that resulted in a 100% Cd(II)S₃ [4].

In the present study, a broader exploration of the application of stereochemical control of the second coordination sphere of metal ions to define coordination number has been completed. We show that when D-Leu is inserted below (toward the C-terminus) rather than above the Cys layer (**TRIL2WL16CL19_DL**), the reorientation of D-Leu produces 4- or 5-coordinate Cd(II) complexes, and substitutions of two D-Leu residues simultaneously above and below the 16-Cys layer in **GRAND-CSL12_DLL16CL19_DL** give similar results, although with a different relative population of the 4- and 5-coordinated species. An analysis of this work enhances our understanding of how the non-coordinating ligands around the metal center have an impact on metallocofactor structure, especially by defining the coordination chemistry of the metal site.

Experimental section

Materials

Peptide synthesis and purification—Peptides were synthesized on an Applied Biosystems 433A peptide synthesizer, purified and characterized as previously reported [54]. Concentration of peptide stock solutions was determined by quantitation of the cysteine thiols using Ellman's test, which uses dithionitrobenzoate (DTNB) as an indicator [55].

Physical characterizations

Circular dichroism (CD) spectroscopy—CD spectral measurements were performed at room temperature on an AVIV 62DS spectrometer using a 1 mm strain-free quartz cuvette. CD spectra for 20 μM (monomer concentration) apo-peptide in a solution containing 10 mM phosphate buffer pH 8.0 and 10:1 TCEP to monomer ratio (to prevent Cys oxidation) were collected from 280 to 190 nm every 1 nm with 1 s signal averaging. In the case of holo-peptide, Cd(II)Cl₂ 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 cm².

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)Cl₂ into a 3 mL solution containing 60 μM peptide monomer, 50 mM TrisHCl buffer pH 8.5 and 10 eq of TCEP solution per monomer ratio. Upon titrations, the solutions were stirred by flat rectangular magnet for 5 min and allowed to sit for 3 min before being scanned. The spectra between 350 and 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 μM 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 μM Cd(II)Cl_2 and 60 μM peptide monomer on an Ocean Optics SD 2000 fiber-optic spectrometer. The scans were collected from 350 to 190 nm and the absorbance at λ_{max} 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 $\text{Cd(ClO}_4)_2$ solution in D_2O . A spectral width was sampled using a 5 μs 90° pulse and 0.05 μs acquisition time, with no delay between scans. Samples were prepared by dissolving ~ 30 mg of lyophilized and degassed peptide in 500 μL of 10% $\text{D}_2\text{O}/\text{ddH}_2\text{O}$ 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 DTNB, yielding colormetric 2-nitro-5-thiobenzoate (NTB^-) [55]. The sample concentrations usually range from 4- to 6 mM 3SCC. An appropriate amount of a 250 mM $^{113}\text{Cd}(\text{NO}_3)_2$ solution (prepared from 95% isotopically enriched ^{113}CdO obtained from Oak Ridge National Laboratory) was then added to make a 1:1 $\text{Cd(II)}:3\text{SCC}$ ratio. The final pH was adjusted with concentrated KOH or HCl to reach pH 8.5. The data were processed using MestRe-C [56]. The free induction decays were zero-filled and processed with an exponential function with a line-broadening value of 100 Hz before Fourier transformation.

$^{111\text{m}}\text{Cd}$ perturbed angular correlation (PAC) spectroscopy—The experiments were performed with a setup using six detectors. The sample temperature was $1 \pm 2^\circ\text{C}$, which was controlled by a Peltier element. The radioactive $^{111\text{m}}\text{Cd}$ was produced on the day of the experiment at the University Hospital cyclotron in Copenhagen and extracted as described previously [57], except for the HPLC separation of zinc and cadmium, which is only required to lower Zn(II) . This procedure may lead to zinc contamination of the sample, but the level of contamination should not interfere with the experiment. The $^{111\text{m}}\text{Cd}$ solution (10–40 μL) 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. The pH of the solution was adjusted with H_2SO_4 or KOH. The pH reported was measured at room temperature after the PAC experiment 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 $\text{pH}(1^\circ\text{C}) = 0.964 [\text{pH}(25^\circ\text{C})] + 0.86$ [57]. 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 [52], 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 is denoted 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 ω_0/ω_1 ; dynamics similarly affects the PAC signal, and is reflected in the rotational correlation time τ_c . In summary, five parameters [A , η , ω_0 (or ν_Q), ω_0/ω_1 and τ_c] are determined for Cd(II)(**TRIL2W16CL19_DL**)₃[−] and Cd(II)(**GRAND – CSL12_DL16CL19_DL**)₃[−], see Table 2.

Extended X-ray absorption fine structure (EXAFS)—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 Lightsource (SSRL) on beam line 7–3. An Si(220) double crystal monochromator was detuned to 50% of the maximum intensity for harmonic rejection and the beam size was $1 \times 7 \text{ mm}^2$. The X-ray energies were calibrated by collecting the absorption spectrum of a Cd foil reference 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 Ge solid-state detector array. Spectra were measured with 10 eV increments in the pre-edge region (26,484–26,695 eV), 0.3 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 1 s in the pre-edge and edge and 1–20 s (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 of the scans was examined for glitches before averaging; and a total of six scans were used for averaging. The raw data were normalized to the tabulated X-ray absorption coefficients using MBACK [58]. Data were converted to k space, $k = [2m_e(E - E_0)/\hbar^2]^{1/2}$, using a threshold energy ($E_0 = 26\,716.3 \text{ eV}$) that was calibrated by fitting the EXAFS data for the crystallographically characterized reference compounds Cd(ClO₄)₂ and Cd(SC(ⁱPr)₃)₃[−] [59]. Background in EXAFS data was further removed using a least-squares spline with k^3 weighting. The EXAFS data, $\chi(k)$, were fit to Eq. (1), where N_s is the number of scatterers with atom types 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}), $\phi_{as}(k)$ is the phase-shift of the photoelectron wave traveling between the potentials of the absorbing and scattering atoms, and the sum is calculated over all shells of scatterers. The amplitude and phase functions, $A_{as}(k)$ and $\phi_{as}(k)$, were calculated using FEFF 7.02, with an amplitude reduction factor of 0.9. 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 [60]:

$$x(k) = \sum_s \frac{N_s R_{as} A_{as}(k)}{k R_{as}} e^{-2\sigma_{as}^2 k^2} \sin[2k R_{as} + \varphi_{as}(k)]. \quad (1)$$

Results

Circular dichroism 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 designate folded coiled-coil structures, revealing an average of $-25,468 \text{ deg dmol}^{-1} \text{ cm}^2$ for apo-TRIL2WL16CL19_DL and of $-26,212 \text{ deg dmol}^{-1} \text{ cm}^2$ for metalated-TRIL2WL16CL19_DL. These values correspond to 71.7 ± 0.1 and $73.8 \pm 1.2\%$ initial folded coiled-coil structures, respectively (Fig. 2). GRAND-CSL12_DLL16CL19_DL has a negative peak at 222 nm with an average of $-25,052 \text{ deg dmol}^{-1} \text{ cm}^2$ for apo-peptide and of $-26,055 \text{ deg dmol}^{-1} \text{ cm}^2$ for metalated-peptide corresponding to 70.6 ± 0.9 and $73.4 \pm 0.8\%$ initial folding, respectively (Fig. S1). GnHCl denaturation studies were conducted to evaluate the stability of the peptides. It was found that G_{folding} of the TRIL2WL16CL19_DL is $1.8 \pm 0.1 \text{ kcal mole}^{-1}$ [in the absence of Cd(II)] and $1.9 \pm 0.1 \text{ kcal mole}^{-1}$ [in the presence of Cd(II)] (Fig. S2). For GRAND-CSL12_DLL16CL19_DL, the apo-peptide has G_{folding} of $0.80 \pm 0.05 \text{ kcal mole}^{-1}$ and the Cd(II)-peptide of $0.95 \pm 0.03 \text{ kcal mole}^{-1}$ (Fig. S3).

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 TRIL2WL16CL19_DL 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 equivalent of Cd(II) per 3SCC and leveling off afterwards (Fig. 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 $19,900 \text{ M}^{-1} \text{ cm}^{-1}$ ($\lambda_{\text{max}} = 232 \text{ nm}$). A similar observation was found with GRAND-CSL12_DLL16CL19_DL, where λ_{max} of LMCT band occurs at 232 nm with a final extinction coefficient of $19,700 \text{ M}^{-1} \text{ cm}^{-1}$ at pH 8.5 with 1:1 Cd(II):trimer (Fig. S4).

¹¹³Cd NMR

The ¹¹³Cd 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 and 610 ppm for TRIL2WL16CL19_DL and GRAND-CSL12_DLL16CL19_DL (Fig. 4), respectively.

¹¹¹mCd PAC

¹¹¹mCd PAC was used to assign the Cd(II) coordination geometry of the peptides. The PAC data of ¹¹¹mCd and **TRIL2WL16CL19_DL** monomer in a 1:12 ratio at pH 8.6, 1 °C cannot be analyzed with only one nuclear quadrupole interaction (NQI). A single NQI recorded for randomly oriented molecules must display the highest intensity of the first (lowest frequency) peak in the Fourier transformed data, and the current data clearly exhibit the highest intensity of the second peak, see Fig. 5. Thus, at least two NQIs are required for a satisfactory fit, reflecting the co-existence of two coordination geometries. The fact that only one ¹¹³Cd NMR signal is observed, indicates that the two structures are in fast exchange on the NMR time scale, while in slow exchange on the PAC time scale (~ nanoseconds). There are several PAC parameter sets that qualitatively account for the recorded spectra, but the best fit including two NQIs is achieved with one NQI at $\omega_0 = 0.316 \text{ rad ns}^{-1}$ ($\eta = 0.17$) and the other at $\omega_0 = 0.159 \text{ rad ns}^{-1}$ ($\eta = 0.33$) with equal amplitudes indicating 50:50 mixtures of two species (Fig. 5). **GRAND-CSL12_DLL16CL19_DL** also reveals two NQIs, both of which are similar to **TRIL2WL16CL19_DL**. The first NQIs at $\omega_0 = 0.329 \text{ rad ns}^{-1}$ ($\eta = 0.15$) and $\omega_0 = 0.183 \text{ rad ns}^{-1}$ ($\eta = 0.26$) with a 66:34 ratio, where the major species belongs to the higher angular frequency at pH 8.6, 1 °C (Fig. 5). The ¹¹¹mCd PAC parameters are given in Table 2.

EXAFS

X-ray absorption spectroscopy (XAS) was used to obtain precise metrical parameters for the Cd(II) coordination environment. The EXAFS data for Cd(II)(**TRIL2WL16CL19_DL**)₃⁻, shown in Fig. 6a, is representative of all of the EXAFS data (the spectra for the other three samples are shown in Fig. S5). The spectra are all quite similar, showing only minor variations in the shape of the amplitude envelope and in the frequency of the oscillations. In all cases, the Fourier transform of the data shows only a single peak, arising predominantly from Cd(II)-S scattering, and in all cases, the data can be fit using a single Cd(II)-S shell with a Cd(II)-S distance of ~ 2.5 Å. In every case, there is a modest improvement in the fit quality when a Cd(II)-O shell is added. Although the improvement in fit quality (15%) is slightly smaller for the authentic Cd(II)S₃ environment in **TRIL12_DLL16C** than it is for the samples shown by PAC to have Cd(II)S₃O ligation (21–26%), this difference is not sufficient on its own to distinguish between different possible ligations. This reflects the relative size of the Cd(II)-S and Cd(II)-O EXAFS signals. For this reason, we have instead used the PAC results to define the Cd ligation [Cd(II)S₃ vs. Cd(II)S₃O_{*n*}] and have relied on EXAFS to determine the metric parameters. The fitted Cd(II)-S parameters are independent ($\pm 0.01 \text{ Å}$) of whether or not an oxygen shell was included. The best fits to the *k*³-weighted EXAFS are shown in Figs. 6 and S5, and the parameters are summarized in Table 3.

Discussion

A true test for the complete understanding of metalloprotein design is the capability to force specific coordination structures on a metal site. One approach to achieve this control is to exploit changes in amino acid stereochemistry in de Novo metalloprotein design. While not used in nature, the advanced concept of using D-amino acids as an approach to control Cd(II) site properties has been shown to be a viable strategy [4, 5]. The placement of D-Leu at the

12th position in the **TRIL12_DLL16C** design [4] achieved an exclusive Cd(II)S₃ structure. We hypothesized that when D-Leu was placed in the 19th position, one layer below the metal site (the 16th position), the altered configuration of D-Leu would make the isobutyl side chain point down toward the C-terminus. Such a modification [**TRIL2WL16CL19_DL**, where the Tryptophan residue (W) at the second position is for spectroscopic tag] would provide room for an additional water ligand to access the Cd(II) center, which could allow for the formation of the first 5-coordinate Cd(II)S₃O₂ complex within this class of designed peptides.

Analysis of Cd(II) interactions in **TRIL2WL16CL19_DL** peptide

The effect of introducing a D-amino acid on the secondary and tertiary structures of the coiled coils in aqueous solution at pH 8.5 was first determined by CD spectroscopy. In Fig. 2, the co-existence 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 [62]. Based on the molar ellipticity at 222 nm, Cd(II)(**TRIL2WL16CL19_DL**)₃⁻ is $73.8 \pm 1.2\%$ folded, which is less than the L-amino acid sequence Cd(II)(**TRIL16C**)₃⁻ (almost 100% folding) [63]. This indicated that the inclusion of a D-amino acid destabilizes the peptide. Moreover, GnHCl denaturation studies further emphasized that when Cd(II) was added into the peptide, Cd(II)(**TRIL2WL16CL19_DL**)₃⁻ exhibits G_{folding} of $1.9 \text{ kcal mol}^{-1}$ (G_{folding} of the apo-peptide is $1.8 \text{ kcal mol}^{-1}$) (Fig. S2), whereas the value of $\sim 6.4 \text{ kcal mol}^{-1}$ was reported for Cd(II)(**TRIL16C**)₃⁻. This result indicated that the coiled-coil system is destabilized when the opposite chirality of D-Leu is incorporated in the 19th position. 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(II)(**TRIL2WL16CL19_DL**)₃⁻ complex formed (Fig. 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) efficiently. 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 ϵ vs. equivalents of metal plotted for both **TRIL16C** and **TRIL2WL16CL19_DL**. However, the final extinction coefficients per metal of the peptides are different (Table 3). There is an increase in ϵ from $19,900 \text{ M}^{-1} \text{ cm}^{-1}$ to $22,600 \text{ M}^{-1} \text{ cm}^{-1}$ going from **TRIL2WL16CL19_DL** to **TRIL16C**. When these Cd(II) peptides are compared with the published **TRIL12_DLL16C** result, which gives a fully Cd(II)S₃ complex, a trend is observed from **TRIL2WL16CL19_DL** (having D-Leu below) $19,900 \text{ M}^{-1} \text{ cm}^{-1}$, **TRIL12AL16C** (exclusively CdS₃O) $21,200 \text{ M}^{-1} \text{ cm}^{-1}$ [6, 37], **TRIL16C** (mixture of 3- to 4- coordinate 40:60) $22,600 \text{ M}^{-1} \text{ cm}^{-1}$, and **TRIL12_DLL16C** [D-Leu above, 100% Cd(II)S₃ complex] $23,600 \text{ M}^{-1} \text{ cm}^{-1}$ [4, 35]. There is a trend in these observations suggesting that the higher extinction coefficients correlate with lower Cd(II) coordination numbers (Fig. S6). At one end, **TRIL2WL16CL19_DL** assumed to bind Cd(II) in higher coordination numbers (4 or 5) has the lowest ϵ , while **TRIL12_DLL16C** published with a fully trigonal site has the highest ϵ among the three. The pure 4-coordinate **TRIL12AL16C** is slightly lower than the **TRIL16C**, which contains that mixtures of the two species [Cd(II)S₃ and Cd(II)S₃O] have ϵ falling in between the two

diastereopeptides. While this correlation is likely not quantitative, it provides a qualitative method to differentiate geometries of the Cd(II) complexes.

The ^{113}Cd NMR spectrum of $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L})_3^-$ shows a single resonance at 600 ppm suggesting almost 100% of $\text{Cd(II)S}_3\text{O}$ (Fig. 4b); however, the PAC data exhibit two NQIs with ω_0 values of 0.316 and 0.159 rad ns^{-1} displaying approximately equal amplitudes (A) (Fig. 5). This suggests the co-existence of two well-defined binding sites in the interior of the coiled coil. The parameter at 0.316 rad ns^{-1} agrees with the presence of a $\text{Cd(II)S}_3\text{O}$ 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 angular frequency at 0.159 rad ns^{-1} is also present. Such a low angular frequency is different from the values that have usually been observed for Cd(II) within these protein environments, which strongly suggests that the Cd(II) site structure is distinct from $\text{Cd(II)S}_3\text{O}$ 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 charge distribution around the metal site is relatively symmetrical [where the perfect tetrahedral and octahedral geometries of (4 or 6) identical ligands ideally show ω_0 of 0 rad ns^{-1}] [57]. The trigonal planar Cd(II)S_3 metal site structure gives $\omega_0 \sim 0.450 \text{ rad ns}^{-1}$, and adding an axial water molecule while essentially maintaining the Cd(II) in the plane of the three Cys to form that $\text{Cd(II)S}_3\text{O}$ lowers the frequency to $\omega_0 \sim 0.340 \text{ rad ns}^{-1}$ [6]. Decreasing the frequency further toward the recorded 0.159 rad ns^{-1} may be achieved either by changing the structure from trigonal planar Cd(II)S_3 to pseudotetrahedral $\text{Cd(II)S}_3\text{O}$ with an axial water molecule or by adding another ligand at the opposite axial position forming, for example, $\text{Cd(II)S}_3\text{O}_2$. Structures of similar 3SCCs determined by X-ray diffraction show that the position of the metal ion is in, or close to, the plane of the three-Cys sulphurs even for $\text{Zn(II)S}_3\text{O}$, suggesting that formation of a tetrahedral $\text{Cd(II)S}_3\text{O}$ metal site is not favored [42]. The number of thiols bound to Cd(II), however, is still certainly considered as a tris-thiolate 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 **TRIL2WL16CL19_DL** designs have two potential spaces for solvent both at the 12th and 19th positions that are above and below the metal site. This peptide thus allows for the simultaneous access of two water molecules resulting in $\text{Cd(II)S}_3\text{O}_2$. The $\omega_0 = 0.159 \text{ rad ns}^{-1}$, therefore, suggests a trigonal bi-pyramidal Cd(II) structure and the heterogeneity observed from the PAC spectra indicates that the open space by 19_D-Leu in **TRIL2WL16CL19_DL** allows for more waters around the metal site. Moreover, the X-ray absorption data in Table 3 show that the best fit for Cd(II)-S and Cd(II)-O distances of $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L})_3^-$ (Fig. 6a) is 2.48 and 2.37 Å best explained by a mixed 4- and 5-coordinate system. While the pure $\text{Cd(II)S}_3\text{O}$ formed in $\text{Cd(II)(TRIL12AL16C)}_3^-$ showed 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 $\text{Cd(II)(TRIL12}_\text{D}\text{LL16C)}_3^-$ or $\text{Cd(II)(TRIL16Pen)}_3^-$ is shortened to 2.46 and 2.45 Å [61], respectively. As the coordination number increases, the average bond distances should increase by 0.07 Å [64]. Theoretically, a mixture of 4 and 5 coordinates in $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L})_3^-$ should have longer average bond lengths than the pure $\text{Cd(II)S}_3\text{O}$; however, based on these X-ray

absorption data, it suggests that the $\text{Cd(II)S}_3\text{O}_2$ 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 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 $\text{Cd(II)S}_3\text{:Cd(II)S}_3\text{O}$ species, the perturbation at the outer sphere one layer (12th 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 19th position) below the metal site of **TRIL16C**, **TRIL2WL16CL19_DL** allows access to a 5-coordinate Cd(II) complex.

Coordination geometry determination when both the secondary coordination environments of Cd(II) are occupied by D-Leu

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 asked whether D-Leu residues placed simultaneous at both the 12th and 19th positions leads to pure 4-coordination. The hypothesis is that the isobutyl 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 a 4-coordinate Cd(II) species $[\text{Cd(II)S}_3\text{O}]$, this time with a coordinated water pointing toward the C-terminus rather than the N-terminus as was seen for the **TRIL12AL16C** variant. If this approach were successful, it would provide a systematic method not only for controlling coordination number, but also for dictating the desired location of the exogenous fourth ligand. Such coordination control is a desirable attribute for metalloenzyme design, so that one may predictively generate a catalytic site in a defined region of the protein. The knowledge gained from this work will not only provide a better understanding of the benefits of using stereochemistry in metalloprotein design, but also extend insight into investigating how well a three-stranded coiled coil can handle multiple D-amino acid substitutions in close proximity within its sequence. The longer length scaffold is 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 both at the 12th and 19th positions of the GRAND-CoilSer (GRAND-CS) sequence, where the coordination site is at the 16th position generating **GRAND-CSL12_DLL16CL19_DL**. The **GRAND-CS** peptide is spectroscopically analogous to the GRAND peptide but also used for crystallographic studies [42]. The physical properties of the peptide were experimentally examined in the same manner as done with **TRIL2WL16CL19_DL**. 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-CSL12_DLL16CL19_DL** containing double D-Leu mutation as was observed for **TRIL2WL16CL19_DL** 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 ~ 5 kcal/mol per heptad per monomer) [65]. 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 apoprotein. 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 determine the effect of the double D-Leu mutation on the stability of **GRAND-CSL12_DLL16CL19_DL** further, GnaHCl denaturation titration studies were performed to compare the results with other single D-Leu containing **GRAND-CoilSer** variants (Fig. S3). The **GRAND-CS** is the crystallographic analogue of the **GRAND** peptide, in which the **GRAND-CS** contains His on the 35th position of the last heptad to facilitate the three-dimensional orientation of 3SCC packing. Previous studies have confirmed that the Cys-substituted variants of both peptides were spectroscopically identical [61]; however, the special feature of the **GRAND-CS** makes it also possible to perform structural studies. Obviously, the **GRAND-CSL12_DLL16CL19_DL** is not highly stable with the G_{folding} only 0.77 ± 0.15 kcal mole⁻¹ for the apo-peptide and 0.95 ± 0.11 kcal mole⁻¹ for the Cd(II) peptide. On the other hand, a single D-Leu substitution in **GRAND-CSL12_DLL16C** [61] has G_{folding} of 3.02 ± 0.12 kcal mole⁻¹ [in the absence of Cd(II)] and 3.07 ± 0.16 kcal mole⁻¹ [in the presence of Cd(II)]. The **GRAND-CSL16CL19_DL** (analogous to the **TRIL2WL16CL19_DL**) shows less stability compared to **GRAND-CSL12_DLL16C**. The apo-peptide has G_{folding} of 2.12 ± 0.05 kcal mole⁻¹ and the Cd(II) peptide of 2.60 ± 0.03 kcal mole⁻¹. From these observations, the replacement of D-Leu in the 12th position leads to the most stable peptide among the three in the series. This is most likely because the substitution is close to the N-termini, where the hydrophobes are more tightly-packed, and because of the side chain rotamer conformations of a *d* site amino acid, the isobutyl group is better accommodated within the hydrophobic core. Consequently, D-Leu residues in the 12th position could tuck the whole side chain into the center of the core. This conclusion is confirmed crystallographically [61]. In summary, it appears that D-Leu substitution in these 3SCCs toward the N-terminus leads to less destabilization than when incorporated near the C-terminus. The destabilization of these two peptides is not as great as seen with **GRAND-CSL12_DLL16CL19_DL**, 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 a double D-Leu substitution in helical assemblies.

Even though the coiled-coil construct of **GRAND-CSL12_DLL16CL19_DL** is less stably folded, the peptide could form a Cd(II)(**GRAND-CSL12_DLL16CL19_DL**)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 Fig. S4, 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-CSL12_DLL16CL19_DL** bind to one Cd(II) as a coiled coil with the best description for the complexes to be Cd(II)(**GRAND-CSL12_DLL16CL19_DL**)₃⁻. The final extinction coefficient is close to what was reported to **TRIL2WL16CL19_DL** (Table 3). From the UV-vis data, it is highly

likely that **GRAND-CSL12_DLL16CL19_DL** peptide shares a similar type of metal coordination geometry [a mixture of 4- and 5-coordinate Cd(II) as **TRIL2WL16CL19_DL**].

¹¹³Cd NMR experiments were performed to assign the Cd(II) binding environment. Cd(II) bound to **GRAND-CSL12_DLL16CL19_DL** results in a chemical shift of 610 ppm (Fig. 4a) which falls in an acceptable range of values that are reported for Cd(II)(SR)₃[−] complexes [66]. If constrained to a published correlation model comparing chemical shifts for Cd(II)S₃ and Cd(II)S₃O species for a site peptide system from Iranzo et al. [41], this observed shift would be consistent with predominantly Cd(II)S₃O present (~ 75%); however, the possible existence of Cd(II)S₃O₂ in **GRAND-CSL16CL19_DL** may render this assignment incorrect. The upfield shift does indicate the presence of water ligand(s) on the Cd(II) center in **GRAND-CSL12_DLL16CL19_DL**. Thus, ¹¹¹mCd PAC spectroscopy was applied to further characterize the Cd(II) coordination geometries. The ¹¹¹mCd PAC spectra of **GRAND-CSL12_DLL16CL19_DL** in Fig. 5 unambiguously demonstrate the presence of two NQIs; a minor species (34%) with $\omega_0 = 0.183 \text{ rad ns}^{-1}$ ($\eta = 0.26$) most likely reflects the three-Cys trigonal bi-pyramidal Cd(II) S₃O₂ structure, where O is from a solvent water and a major species (66%) at $\omega_0 = 0.329 \text{ rad ns}^{-1}$ ($\eta = 0.15$) suggesting Cd(II)S₃O with Cd(II) in or close to the plane of the three Cys and an axial water. Thus, the PAC results show that **GRAND-CSL12_DLL16CL19_DL** contains a heterogeneous metal center with two species, presumably Cd(II)S₃O₂ and Cd(II)S₃O. Because a single resonance was obtained from ¹¹³Cd NMR, this clearly illustrates that these two coordination geometries have faster dynamic exchange than the ¹¹³Cd NMR timescale. The ¹¹¹mCd PAC integrations demonstrate the co-existence of ca. 66% Cd(II)S₃O and 34% Cd(II)S₃O₂. It appears that when adding steric interference into the layer at the 12th position above the L16CL19_DL design, the coordination number of Cd(II) shifts from 50% 5-coordinate observed for **TRIL2WL16CL19_DL** to 34% for **GRAND-CSL12_DLL16CL19_DL**. The X-ray absorption data given in Table 3 support this model in that the best fit for Cd(II)-S and Cd(II)-O distances of this peptide (Fig. 6b) is 2.48 and 2.28 Å corresponding to a mixed 4- and 5-coordinate system. The Cd(II)-O distance in Cd(II)(**GRAND-CSL12_DLL16CL19_DL**)₃[−] is shorter than the corresponding result from Cd(II)(**TRIL2WL16CL19_DL**)₃[−] suggesting oxygens bind tighter to the Cd(II) center.

These observations suggest that while the reorientation of 19_D-Leu definitely removes the steric hindrance below the 16-Cys site encouraging water access from the C-terminal side of the metal binding site, the 12_D-Leu placed one layer on top of the 16th 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 packing in the core causing the helices to separate further apart. This effect would lead to a difficulty of folding into a desired structure as indicated by the destabilization of the peptide. We conclude that the steric packing made by 12_D-Leu in the **GRAND-CSL12_DLL16CL19_DL** construct is not as efficient in excluding a water ligand as found in **TRIL12_DLL16C**. While not completely blocking the access of the second water molecule, these data clearly support the conclusion that the design succeeded to some degree. It may also be that placing two _D-Leu on either side of the cysteine layer allows the CH₂-S side chains to adopt different rotamers pushing the plane of the three-Cys layer

toward the C-terminus as compared with **TRIL12_DLL16C**. In this situation, additional space for a water molecule on the N-terminal side of the metal binding site, as compared to **TRIL12_DLL16C**, would be generated. These ideas will await corroboration through crystallographic determination of these proteins.

In summary, the replacement of D-Leu in the second coordination sphere around the metal binding site displays different effects on the Cd(II) structures. The orientation of D-Leu when positioned at the 12th position (close to the N-terminus) above the 16-Cys site follows the hypothesis that this residue can block the access of water giving pure Cd(II)S₃. 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)S₃O and a new species Cd(II)S₃O₂. Moreover, the combination of 12_D-Leu and 19_D-Leu placed in a peptide sequence also produces a mixture in which the major product is Cd(II)S₃O. Even though all of these physical properties are confirmed, there have 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 must be very important to perform crystallization studies to advance the complete understanding of these diastereopeptide designs 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.

Conclusion

De Novo protein design has as a 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 possible, 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 in varying degrees. 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This manuscript is dedicated to Prof. Dr. Helmut Sigel on the occasion of his 80th birthday. Prof. Sigel's contributions over 5 decades have been immense. His description of metal interactions with nucleotides has been groundbreaking and authoritative, while his inexhaustible energy to the community with his book series and service roles to the Society of Biological Inorganic Chemistry and the International Conference on Biological Inorganic Chemistry has helped place the discipline of bioinorganic chemistry as a major field of scientific endeavor. The authors acknowledge funding from the National Institutes of Health (L.R. and V.L.P., ES012236; A.D. and J.E.P.-H., GM 38047) for support of this research. Synchrotron measurements were made at the Stanford Synchrotron Radiation Laboratory, which is supported by the NIH Research Resource Program and the US Department of Energy.

References

1. Hu, C., Wang, J. Method for enzyme design with genetically encoded unnatural amino acids. In: Pecoraro, VL., editor. *Methods in enzymology*, vol 580 "peptide, protein and enzyme design". Vol. 580. Elsevier; United States: 2016. p. 109-134.
2. Guo JT, Wang JY, Lee JS, Schultz PG. *Angew Chem*. 2008; 47:6399–6401. [PubMed: 18624319]
3. Liu CC, Schultz PG. *Ann Rev Biochem*. 2010; 79:413–444. [PubMed: 20307192]
4. Peacock AFA, Hemmingsen L, Pecoraro VL. *Proc Natl Acad Sci USA*. 2008; 105(43):16566–16571. [PubMed: 18940928]
5. Peacock AFA, Stuckey JA, Pecoraro VL. *Angew Chem Int Ed Engl*. 2009; 48(40):7371–7374. [PubMed: 19579245]
6. Lee K-H, Cabello C, Hemmingsen L, Marsh ENG, Pecoraro VL. *Angew Chem Int Ed Engl*. 2006; 45(18):2864–2868. [PubMed: 16596690]
7. Petros AK, Shaner SE, Costello AL, Tierney DL, Gibney BR. *Inorg Chem*. 2004; 43(16):4793–4795. [PubMed: 15285646]
8. Lu Y. *Curr Opin Chem Biol*. 2005; 9(2):118–126. [PubMed: 15811795]
9. Kumar A, Ramakrishnan V. *ACS Synth Biol*. 2010; 4(4):221–247.
10. Baures PW, Ojala WH, Gleason WB, Johnson RL. *J Pept Res*. 1997; 50(1):1–13. [PubMed: 9273882]
11. Wang, A., Nairn, NW., Marelli, M., Grabstein, K. Protein engineering. Kaumaya, PP., editor. In *Tech*; Croatia: 2012. p. 253-290.
12. Privett HK, Reedy CJ, Kennedy ML, Gibney BR. *J Am Chem Soc*. 2002; 124(24):6828–6829. [PubMed: 12059195]
13. Zastrow ML, Peacock AFA, Stuckey JA, Pecoraro VL. *Nat Chem*. 2012; 4:118–123.
14. Balaram P. *J Pept Res*. 1999; 54(3):195–199. [PubMed: 10517156]
15. Mahalakshmi, R., Balaram, P. A new frontier in amino acid and protein research. Nova Science Publishers Inc; New York: 2006. p. 416-430.
16. Dhanasekaran M, Fabiola F, Pattabhi V, Durani S. *J Am Chem Soc*. 1999; 121:5575–5576.
17. Rana S, Kundu B, Durani S. *Chem Commun (Camb)*. 2004; 21:2462–2463.
18. Rana S, Kundu B, Durani S. *Chem Commun (Camb)*. 2005; 2:207–209.
19. Rana S, Kundu B, Durani S. *Bioorg Med Chem*. 2007; 15(11):3874–3882. [PubMed: 17407819]
20. Rana S, Kundu B, Durani S. *Biopolymer*. 2007; 87(4):231–243.
21. Petal K, Srivastava KR, Durani S. *Bioinorg Med Chem*. 2010; 18:8270–8276.
22. Ramachandran GN, Chandrasekaran R. *Indian J Biochem Biophys*. 1972; 9:1–11. [PubMed: 4642355]
23. Ghadiri MR, Granja JR, Buehler LK. *Nature*. 1994; 369(6478):301–304. [PubMed: 7514275]
24. Fernandez-Lopez S, Kim HS, Choi EC, Delgado M, Granja JR, Khasanov A, Kraehenbuehl K, Long G, Weinberger DA, Wilcoxon KM, Ghadiri MR. *Nature*. 2001; 412:452–456. [PubMed: 11473322]
25. Aravinda S, Shamala N, Pramanik A, Das C, Balaram P. *Biochem Biophys Res Commun*. 2000; 273(3):933–936. [PubMed: 10891350]
26. Aravinda S, Shamala N, Bandyopadhyay A, Balaram P. *J Am Chem Soc*. 2003; 125(49):15065–15075. [PubMed: 14653741]
27. Ruckthong L, Peacock AFA, Pascoe CE, Hemmingsen L, Stuckey JA, Pecoraro VL. *Chemistry*. 2017; 23(34):8232–8243. [PubMed: 28384393]
28. Dieckmann GR, McRorie DK, Lear JD, Sharp KA, DeGrado WF, Pecoraro VL. *J Mol Biol*. 1998; 280(5):897–912. [PubMed: 9671558]
29. 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(4):6195–6196.
30. Ghosh D, Pecoraro VL. *Inorg Chem*. 2004; 43(25):7902–7915. [PubMed: 15578824]
31. Farrer BT, McClure CP, Penner-Hahn JE, Pecoraro VL. *Inorg Chem*. 2000; 39(24):5422–5423. [PubMed: 11154553]

32. Touw DS, Nordman CE, Stuckey JA, Pecoraro VL. *Proc Natl Acad Sci USA*. 2007; 104(29): 11969–11974. [PubMed: 17609383]
33. Chakraborty S, Touw DS, Peacock AFA, Stuckey J, Pecoraro VL. *J Am Chem Soc*. 2010; 132(38): 13240–13250. [PubMed: 20825181]
34. Matzapetakis M, Pecoraro VL. *J Am Chem Soc*. 2005; 127(51):18229–18233. [PubMed: 16366576]
35. Matzapetakis M, Farrer BT, Weng T-C, Hemmingsen L, Penner-Hahn JE, Pecoraro VL. *J Am Chem Soc*. 2002; 124(27):8042–8054. [PubMed: 12095348]
36. Matzapetakis M, Ghosh D, Weng T-C, Penner-Hahn JE, Pecoraro VL. *J Biol Inorg Chem*. 2006; 11(7):876–890. [PubMed: 16855818]
37. Lee K-H, Matzapetakis M, Mitra S, Marsh ENG, Pecoraro VL. *J Am Chem Soc*. 2004; 126(30): 9178–9179. [PubMed: 15281796]
38. Zampella G, Neupane KP, De Gioia L, Pecoraro VL. *Eur J Inorg Chem*. 2012; 18(7):2040–2050.
39. Neupane KP, Pecoraro VL. *Angew Chem Int Ed Engl*. 2010; 49(44):8177–8180. [PubMed: 20859984]
40. Iranzo O, Thulstrup PW, Ryu S-B, Hemmingsen L, Pecoraro VL. *Chemistry*. 2007; 13(33):9178–9190. [PubMed: 17960740]
41. Iranzo O, Jakusch T, Lee K-H, Hemmingsen L, Pecoraro VL. *Chemistry*. 2009; 15(15):3761–3772. [PubMed: 19229934]
42. Ruckthong L, Zastrow ML, Stuckey JA, Pecoraro VL. *J Am Chem Soc*. 2016; 138(36):11979–11988. [PubMed: 27532255]
43. Iranzo O, Cabello C, Pecoraro VL. *Angew Chem Int Ed Engl*. 2007; 46(35):6688–6691. [PubMed: 17582808]
44. Dieckmann, G., Heilman, S., DeGrado, W., Pecoraro, VL. De novo design of metalloproteins. In: Kessissoglou, DP, Coucouvanis, D., Kanatzidis, M., editors. *An inorganic perspective of life*. Elsevier; Amsterdam: 1995. p. 275
45. Pecoraro, VL., Peacock, AFA., Iranzo, O., Marek, Ł. Understanding the biological chemistry of mercury using a de novo protein design strategy. In: Long, E., et al., editors. *Bioinorganic Chemistry*, ACS Symposium Series. American Chemical Society; Washington, DC: 2009. p. 183–197.
46. Farrer BT, Pecoraro VL. *Proc Natl Acad Sci USA*. 2003; 100(7):3760–3765. [PubMed: 12552128]
47. Iranzo O, Chakraborty S, Hemmingsen L, Pecoraro VL. *J Am Chem Soc*. 2011; 133(2):239–251. [PubMed: 21162521]
48. Coleman JE. *Methods Enzymol*. 1993; 227:16–43. [PubMed: 8255225]
49. Summers F. *Coord Chem Rev*. 1988; 86:43–134.
50. Oz G, Pountney DL, Armitage IM. *Biochem Cell Biol*. 1998; 76(2–3):223–234. [PubMed: 9923691]
51. Hemmingsen L, Olsen L, Antony J, Sauer SPA. *J Biol Inorg Chem*. 2004; 9(5):591–599. [PubMed: 15221483]
52. Hemmingsen L, Stachura M, Thulstrup PW, Christensen NJ, Johnston K. *Hyperfine Interact*. 2010; 197(1–3):255–267.
53. Hemmingsen L, Sas KN, Danielsen E. *Chem Rev*. 2004; 104(9):4027–4062. [PubMed: 15352785]
54. Farrer BT, Harris NP, Balchus KE, Pecoraro VL. *Biochemistry*. 2001; 40(48):14696–14705. [PubMed: 11724584]
55. Ellman GL. *Arch Biochem Biophys*. 1959; 82:70–77. [PubMed: 13650640]
56. Cobas, C., Cruces, J., Sardina, FJ. 2.3 edn. Universidad de Santiago de Compostela, Santiago de Compostela. 2000.
57. Hemmingsen L, Bauer R, Bjerrum M, Adolph H, Zeppezauer M, Cedergren-Zeppezauer E. *Eur J Biochem*. 1996; 241:546–551. [PubMed: 8917454]
58. McMaster, WH., Del Grande, NK., Mallett, JH., Hubbell, JH. Lawrence Livermore Radiation Laboratory, Livermore. 1969.
59. Duhme AK, Strasdeit H. *Z Anorg Allg Chem*. 1999; 625:6–8.

60. Clark-Baldwin K, Tierney D, Govindaswamy N, Gruf ES, Kim C, Berg J, Koch SA, Penner-Hahn JE. *J Am Chem Soc.* 1998; 120:8401–8409.
61. Ruckthong, L. Dissertation. University of Michigan; Ann Arbor: 2016.
62. Chen YH, Yang JT, Chau KH. *Biochemistry.* 1974; 13:3350–3359. [PubMed: 4366945]
63. Dieckmann, GR. Dissertation. University of Michigan; Ann Arbor: 1995.
64. Shannon RD. *Acta Crystallogr A.* 1976; 32:751–767.
65. Ghosh D, Lee K-H, Demeler B, Pecoraro VL. *Biochemistry.* 2005; 44(31):10732–10740. [PubMed: 16060682]
66. Santos RA, Gruf ES, Koch SA, Harbison GS. *J Am Chem Soc.* 1991; 5:469–475.

Fig. 1.

Top: helical wheel diagram of three helices demonstrating hydrophobic residues occupies the interior face at *a* and *d* positions, while hydrophilic residues are on the exterior face at other positions. The grey circles indicate that the residues can form salt bridge interactions between *e* and *g* positions. Residues at *a* and *d* positions are shown as blue sticks, while the remainder are colored cyan (carbons = cyan, nitrogens = blue, and oxygens = red). The figure was made based on the apo-(CSL16C)₃ (PDB code: 5K92) [42]. Bottom: PyMOL visualization demonstrating how to engineer a tris-thiolate binding site into the interior of 3SCC. The substitution of Leu residues (shown as blue sticks) with Cys yields three-Cys environment that is capable of binding heavy metal atoms. The figure was made based on the Hg(II)_SZn(II)_N(GRAND – CSL16CL30H)₃⁺ structure (PDB code: 5KB0) [42]

Fig. 2.

Far UV circular dichroism spectra of apo-diastereopeptide and metalated-TRIL2WL16CL19_DL monitored by CD. The negatively doubled well at 208 and 222 nm reveals the α -helical coiled-coil structure. The grey and black lines represent the spectra of apo-(TRIL2WL16CL19_DL)₃ and Cd(II)(TRIL2WL16CL19_DL)₃⁻ [1 eq of Cd(II) per 3SCC], respectively, at micromolar concentration of monomer. The spectra were recorded at pH 8.5 and ambient temperature

Fig. 3. UV/Vis Cd(II) binding titrations to (**TRIL2WL16CL19DL**)₃ peptide. UV/Vis titration of Cd(II) into a solution of **TRIL2WL16CL19DL** at pH 8.5. The data were plotted as Absorbance ($\text{M}^{-1} \text{cm}^{-1}$) vs. wavelength (nm). The inset to each titration profile shows the titration curve plotted as ϵ at λ_{max} vs. equivalents of Cd(II) added

Fig. 4. ^{113}Cd NMR spectra of **a** $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L)}_3^-$ and **b** $\text{Cd(II)(GRAND - CSL12}_\text{D}\text{LL16CL19}_\text{D}\text{L)}_3^-$. The spectra were recorded at pH 8.5

Fig. 5.

^{111}mCd PAC data for top: $\text{Cd(II)(GRAND - CSL12}_\text{D}\text{LL16CL19}_\text{D}\text{L})_3^-$ and bottom $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L})_3^-$. Fourier transforms (the thin line is the Fourier transform of the experimental data and the bold line is the Fourier transform of the fit). The samples contained 20 mM Tris buffer pH 8.6 and 250–300 μM peptide and 1/12 eq of $\text{Cd(II)(NO}_3)_2$

Fig. 6.

Extended X-ray absorption fine structure data for **a** $\text{Cd(II)(TRIL2WL16CL19}_\text{D}\text{L)}_3^-$ and **b** $\text{Cd(II)(GRAND - CSL12}_\text{D}\text{LL16CL19}_\text{D}\text{L)}_3^-$. 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

Table 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 30
TRI	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH ₂			
TRIL16C	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH ₂			
TRIL12AL16C	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH ₂			
TRIL12_bLL16C	Ac-G LKALEEK LKALEEK LKALEEK LKALEEK G-NH ₂			
TRIL2WL16CL19_bL	Ac-G WKALEEK LKALEEK LKALEEK LKALEEK G-NH ₂			
GRAND-CoilSer	Ac-E WEALEKK LAALESK LQALEKK LQALEKK LEALEHG -NH ₂			
GRAND-CSL12_bLL16C	Ac-E WEALEKK LAAPLESK CQALEKK LQALEKK LEALEHG -NH ₂			
GRAND-CSL16CL19_bL	Ac-E WEALEKK LAALESK CQAPLESK LQALEKK LEALEHG -NH ₂			
GRAND-CSL12_bLL16CL19_bL	Ac-E WEALEKK LAAPLESK CQAPLESK LQALEKK LEALEHG -NH ₂			

Bold and underlined residues indicate substitutions.

C- and N-termini are capped by Ac and NH₂ groups, respectively.

C- and N-termini are capped by Ac and NH₂ groups, respectively

Bold and underlined residues indicate substitutions

Table 2¹¹³Cd NMR chemical shifts and parameters fitted to the ¹¹¹mCd PAC data

Peptide	¹¹³ Cd NMR (ppm)	¹¹¹ mCd PAC parameters ^d				η	ω_0/ω_0 ($\times 100$)	$1/\tau$ (μs^{-1})	A ($\times 100$)	χ_r^2
		pH	Temperature (°C)	ω_0 (rad ns ⁻¹)						
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)	
TRIL12pLL16C^c	697	9.1	1	0.456 (4)		0.10 (1)	0.7 (2)	14 (1)	8.2 (2)	1.18
TRIL2WL16CL19pL	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)	
GRAND-CSL12pLL16CL19pL	610	8.6	1	0.329 (5)		0.15 (6)	8 (2)	12 (2)	6.2 (8)	0.68
				0.183 (3)		0.26 (5)	4 (2)	12 (2)	3.2 (5)	

^aThe numbers in parentheses are the standard deviations of the fitted parameters^bData from Refs. [36]^cData from Ref. [4]

Table 3

Spectroscopic data for Cd(II) binding to the different peptides

Peptide	Cadmium species (%)		Apparent p <i>K</i> _{a2}	UV/Vis		EXAFS Cd(II)-S bond distance (Å)
	Cd(II)S ₃	Cd(II)S ₃ O	CdS ₃ O ₂	λ_{max} (nm)	ϵ (M ⁻¹ cm ⁻¹)	
TRIL16C	40	60	–	232 ^a	22,600 ^a	2.49 ^b
TRIL12pLL16C	100	–	–	232 ^c	23,600 ^c	2.46
TRIL2WL16CL19pL	–	50	50	232	19,900	2.48 [Cd(II)-O = 2.37 Å]
GRAND-CSL12pLL-16CL19pL	–	66	34	232	19,700	2.48 [Cd(II)-O = 2.28 Å]
TRIL12AL16C	–	100	–	231 ^e	21,200 ^e	2.50 [Cd(II)-O = 2.35 Å]

^aData from Refs. [26, 36]^bData from Refs. [3, 35]^cData from Ref. [4]^dData from Refs. [37, 40]^eData from Refs. [6, 38]